言語研究 (*Gengo Kenkyu*) **129** (2006), 181〜226

【Forum】

Issues on the Scrambling Effects in the Processing of Japanese Sentences: Reply to MIYAMOTO and NAKAMURA (2005) Regarding the Experimental Study by KOIZUMI and TAMAOKA (2004)

Katsuo TAMAOKA (Hiroshima University)
Masatoshi KOIZUMI (Tohoku University)

Key words: scrambling effects, sentence processing, incremental processing, syntactic complexity, frequency

1. Introduction

The commentary offered by Miyamoto and Nakamura (2005, MN hereafter) in reference to our experimental study (Koizumi and Tamaoka, 2004, KT hereafter) provides us with an opportunity to re-examine our views regarding the effects of scrambling on the processing of Japanese sentences. The initiation of this kind of critical dialog furthers the scientific development of psycholinguistics, especially the area of syntactic processing. In their introductory remarks, MN claim that our 'conceptual justifications are problematic', citing a perceived failure to address 'critical issues in performance' with regards to our treatment of scrambling effects (MN, p. 113). More specifically, MN's comments opened up three important aspects for discussion; (1) incremental processing and syntactic complexity, (2) frequency effects, and (3) measurement methods. In the present rebuttal to these authors, we defend our 2004 findings by providing detailed explanations of analytical processes.

2. Incremental processing and syntactic complexity

Since Chujyo (1983) it has been repeatedly observed that Japanese speakers can judge whether or not a sentence is semantically plausible (through plausibility judgment task, or sentence correctness decision task) more quickly when its immediate constituents are ordered canonically (e.g., SOV) compared to when they are scrambled (e.g., OSV). Using a wide range of sentence-types, such as active, passive, potential and causative sentences in experiments with the plausibility judgment task, Tamaoka, Sakai, Kawahara, Miyaoka, Lim, and Koizumi (2005) argue that the primary source of the difference in reaction times between canonical and scrambled sentences is neither the linear sequence of thematic roles (such as agent and theme), nor the linear ordering of case particles (such as the nominative *ga* or the accusative *o*); but rather it is the difference in syntactic (i.e., representational and computational) complexity that is accountable. In other words, these authors suggest that scrambled sentences take longer to process than their canonical counterparts because the former have more complex syntactic representations due to the presence of the filler-gap dependency (associated with scrambling), which requires extra cognitive resources for processing. Assuming that this proposal is on the right track, in KT (2004), we investigated reaction times for complete, ditransitive sentences. We found that regardless of the verb type (*pass*-type or *show*-type), the *ga-ni-o* order was processed faster than the *ga-o-ni* order, from which we concluded that *ga-ni-o* is the canonical word order for ditransitive verbs1). Note that KT's assumption mentioned above regarding the relationship between the presence of scrambling and reaction times respects standard methodology in cognitive neuroscience, i.e. to assume 'all other things being equal, the more complex a representation ... [is], the longer it should take for a subject to perform any task involving the representation and the more activity should be observed in the subject's brain areas associated with creating or accessing the repre-

---

1) In support of this conclusion, Koizumi et al. (2005), a related fMRI study, report that regions in the inferior frontal gyrus, sensitive to syntactic complexity, showed greater cortical activation during the processing of *pass*-type ditransitive sentences with the *ga-o-ni* order compared to those with the *ga-ni-o* order.

sentation and with performing the task' (Marantz, 2005, see also Pritchett and Whitman, 1995)[2].

MN (2005) express concern about whether or not all other things are really equal in our experimental paradigm, given the incremental nature of sentence processing. MN state, in particular, that 'KT do not take into consideration potential ambiguities in mid-sentence' (p. 116) such as the one illustrated in (1).

(1) John-*ga* hanataba-*o* Mary-*ni*
John-NOM flowers-ACC Mary-DAT

This sequence was used in our study as part of a *pass*-type item. Within Matsuoka's (2003) theory, the order in (1) is canonical if followed by a *pass*-type verb, and is scrambled if followed by a *show*-type verb. Thus, if readers expect a *show*-type verb to follow, a gap (=trace) should be created after they see the third NP (i.e. *Mary-ni*), as shown in (2a). However, if a *pass*-type verb is seen next, the gap must be removed, as in (2b).

(2) a. NP-*ga* NP$_i$-*o* NP-*ni* *gap*$_i$
b. NP-*ga* NP$_i$-*o* NP-*ni* ~~*gap*~~$_i$ *pass*-type verb

MN maintain that the additional cognitive load required to perform this kind of reanalysis needs to be considered in order to provide a fair experimental assessment of Matsuoka's theory. This is a reasonable concern, considering that the Japanese parser is known to be *incremental.* For example, if presented with an NP-*o* NP-*ga* sequence, a Japanese reader would immediately posit an object gap before encountering the following verb (Miyamoto and Takahashi, 2002).

---

[2] We disagree with MN's assertion that 'syntactic complexity (as measured by number of merges, tree nodes, transformational rules) is not a factor' (p. 115) in sentence processing. Sentence comprehension necessarily involves the computation and maintenance of syntactic representations, which consume a large amount of cognitive resources such as working memory. Syntactic complexity therefore must be considered as a factor in sentence processing. As far as we know, this latter position has never been empirically denied.

(3) NP_i-*o* NP-*ga* gap_i

The creation of the gap after the nominative NP can be experimentally detected by observing the P600 component of event related brain potentials (ERP), which has been purported to be elicited at the point of expected gap sites, which would occur if the NP-*ga* is introduced in (3) (e.g., Hagiwara, Soshi, Ishihara and Imanaka, 2005; Phillips, Kazanina and Abada, 2005; Ueno and Kluender, 2003). Since a gap is created before the verb in a context like (3), it is conceivable that 'the decision to insert the gap is made before the verb is seen' (MN, p. 116) in a context like (1), as well. If a gap is indeed inserted here, as in (2a), then the reanalysis shown in (2b) may also occur, which in turn might cause a longer reaction time.

Although the scenario sketched in the previous paragraph is therefore *theoretically* possible, we do not think it probable for at least the following two reasons. First, consider the examples in (4).

(4) a. Accusative verb
Desi-*ga* atorie-*o* tukutta.
pupil-NOM atelier-ACC built
'The pupil built the atelier.'
b. Dative verb
Desi-*ga* atorie-*ni* komotta.
pupil-NOM atelier-ACC stayed
'The pupil shut himself/herself up in the atelier.'

The verb in (4a) takes an accusative object (i.e., accusative verb), and the verb in (4b) a dative object (i.e., dative verb). When an accusative verb is used in a syntactic causative construction, the causee (which corresponds to the subject argument in the simple transitive use) appears as an indirect object in the dative. On the other hand, in the causative construction with a dative verb, the causee appears as an indirect object in the accusative. The linear ordering of the indirect and direct objects can be freely altered by scrambling. These possible orders are shown in (5) and (6).

(5) Syntactic causatives of an accusative verb

a. Zyunko-*ga* desi-*ni* atorie-*o*
Junko-NOM pupil-DAT atelier-ACC
tukur-ase-ta. (canonical order)
build-CAUSE-PAST
'Junko made the pupil build the atelier.'

b. Zyunko-*ga* atorie-*o*$_i$ desi-ni t_i
Junko-NOM atelier-ACC pupil-DAT
tukur-ase-ta. (scrambled order)
build-CAUSE-PAST

(6) Syntactic causative of a dative verb

a. Zyunko-*ga* desi-*o* atorie-*ni*
Junko-NOM pupil-ACC atelier-DAT
komor-ase-ta. (canonical order)
stay-CAUSE-PAST
'Junko made the pupil shut himself up in the atelier.'

b. Zyunko-*ga* atorie-*ni*$_i$ desi-*o* t_i
Junko-NOM atelier-DAT pupil-ACC
komor-ase-ta. (scrambled order)
stay-CAUSE-PAST

By comparing reading times for the four types of sentences above, it should be possible to find out whether or not the linear ordering of dative and accusative case markers has a significant effect on syntactic processing, independent of the effect of syntactic complexity. Using the plausibility judgment task, we measured reaction times for these sentences (Experiment 5 of Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi, 2005; see also Koizumi, Tamaoka and Miyaoka, 2004). The results of this experiment revealed no main effect of either verb types or case particle orders. However, there was an interesting contrast between the two types of sentences. The causative sentences with accusative verbs were processed faster in the *ga-ni-o* order than in the *ga-o-ni* order, whereas the causative sentences with dative verbs yielded the opposite result that the *ga-o-ni* order was faster than the *ga-ni-o* order. In other words, the scrambled causative sentences took longer to process than the canonical causative sentences regardless of surface case orders. Linear ordering of accusative and dative case particles had no observable effect on speed (and accuracy) in sentence comprehension. This suggests that the kind of reanalysis (i.e., the creation of a gap after the third NP followed by its deletion when encountering the verb) illustrated in (2) either does not occur, or, if it does occur, its effect is so small that it cannot be detected in the plausibility judgment task. In either case, the putative effect of mid-sentence reanalysis, if any, does not affect our interpretation of the experimental results reported in KT.

Second, in an ERP study of ditransitive constructions in Japanese (Koso and Hagiwara, 2005), P600 (associated with a gap creation process as mentioned above) was observed in the *ga-o-ni* order compared to the *ga-ni-o* order *at the point of the verb* rather than at the third NP, which suggests that native speakers of Japanese would not posit a gap in a context such as (1) until they encounter the verb. This is a natural conclusion, given that either the *ga-ni-o* or *ga-o-ni* order can be syntactically canonical in Japanese as shown in (5) and (6). The mere parse of *ga-o-ni* (or *ga-ni-o*) is insufficient to reveal a filler-gap dependency. Therefore, until encountering a verb, the parser is incapable of recognizing the sequence as containing a filler-gap dependency. If no gap is created before the verb, there would be no reanalysis like the one in (2b).

To summarize, unlike a mono-transitive case like (3), in the processing of ditransitive sentences like (1), insertion of a gap is delayed until the verb is processed. A mid-sentence reanalysis such as the one depicted in (2) does not occur. Our 2004 interpretation of reaction times based on syntactic complexity therefore can be maintained even if incremental aspects of sentence processing are taken into consideration.

3. Three facets of frequency effects

Although MN repeatedly emphasize the importance of frequency as a 'potent confounding factor' (p. 126) affecting sentence processing, they provide neither a clear distinction of the types of frequencies nor theoretical frameworks of frequencies. Frequency is not merely a control condition for experiments, but has theoretical implications unto itself. There are three different avenues via which frequency influences the cognitive processing of Japanese sentences: word frequency (usually printed-frequency), syntactic structural frequency and collocation frequency. Frequency effects have been investigated by means of cognitive psychological experiments and computerized learning simulation (i.e., connectionism). Since these investigations focused primarily on *learning*, their theoretical explanations created a heated discussion between psychologists and linguists regarding the innateness of syntactic acquisition (see Makioka, 1999). This discussion has received increased attention as of late, particularly through the popular book *Rethinking Innateness* (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, and Plunkett, 1996). Without providing any theoretical background for frequency effects, MN's argument regarding our experimental findings can only be received as superficial. Thus, to comprehensively address this point, we now turn to a detailed discussion of the three aforementioned facets of frequency and their theoretical foundations.

3.1 Effects of word frequency

The effect of word frequency is one of the most heavily investigated characteristics in psychology (e.g., Besner and McCann, 1987; Carroll and White, 1973; Caza and Moscovitch, 2005; Forster and Chambers, 1973; Gardner, Rothkopf, Lapan and Lafferty, 1987; Hino and Lupker, 1998; Jescheniak and Levelt, 1994; Morrison and Ellis, 1995; Paap, McDonald, Schvaneveldt and Noel, 1987; Taft, 1979, 1991; Tamaoka and Takahashi, 1999). This effect is essentially based on the assumption that the more frequently a word is seen, the more quickly and accurately the word will be recognized. Printed frequencies of morphemes (including single kanji units) and words are easy to find using various lexical databases. One of the most-frequently used lexical databases to find Japanese word frequency is that of Amano and Kondo (2000). The CD-ROM version of the lexical database (Amano and Kondo, 2003a, 2003b) is available through the *Sanseido* web-site (www.sanseido-publ.co.jp/publ/ntt_database.html). However, it should be noted that the frequency index of Amano and Kondo (2000, 2003b) is a lexical database (not a *corpus*), which counted all inflected forms as single lexical items. Thus, this database provides neither frequency counts for inflected forms nor collocation frequency; it only suffices to determine word frequency in general.

According to Amano and Kondo (2003b), a word 危険 /kikeN/ (/N/ refers to a moraic nasal sound) meaning 'danger' in English appears 14,254 times in the *Asahi Newspaper* over fourteen years (1985 to 1998) while a word 危害 /kigai/ meaning 'harm' appears 595 times. Both words contain the same kanji 危 /ki/ combined with other kanji of an almost-equal visual complexity 険 /keN/ and 害 /gai/, constructing two different lexical items. Among experimental psychologists focusing upon a study of *lexical access* (e.g., Becker and Killion, 1977; McClelland, 1987; McClelland and Rumelhart, 1981; Morton, 1969; Taft, 1991), it is commonly known that a high frequency word like 危険 'danger' is understood more quickly and accurately than a low frequency word like 危害 'harm'.

3.1.1 Cognitive processing models for effects of word frequency

The classical *logogen* model (Morton, 1969), which explained word frequency effects as an activation level of *threshold*, is predetermined by word frequency. High frequency words had their thresholds set lower while low frequency words had their thresholds set higher. As depicted in Figure 1, since the word 'danger' is higher in printed-frequency than the word 'harm', the word 'danger' rests upon a lower threshold than does the word 'harm'. Until the activation level exceeds the threshold, the word can not be recognized. The reaction time required for a lexical decision task asking whether or not a visually presented word exists in Japanese is actually an indirect indicator of an activation level. Basically, as the reaction time after the visual presentation of the target word becomes longer, the activation level becomes higher. Therefore, as shown in Figure 1, the activation can reach the lower threshold of the high frequency word 'danger' (T_1 in the Figure 1) faster than the higher threshold of the low frequency word 'harm' (T_2 in the Figure 1), resulting in the faster recognition of 'danger' than 'harm' (the difference of reaction times is expressed by T_2 minus T_1). Naturally, activation will reach the saturation level, and after a certain period of time, will attenuate.

The act of 'understanding a word' actually involves three different types of representations in the metal lexicon. For example, the word 'danger' has the orthographic representation of two kanji symbols 危険, the phonological representation of three morae /kikeN/, and the semantic representation of the meaning 'danger' (English is used for representing

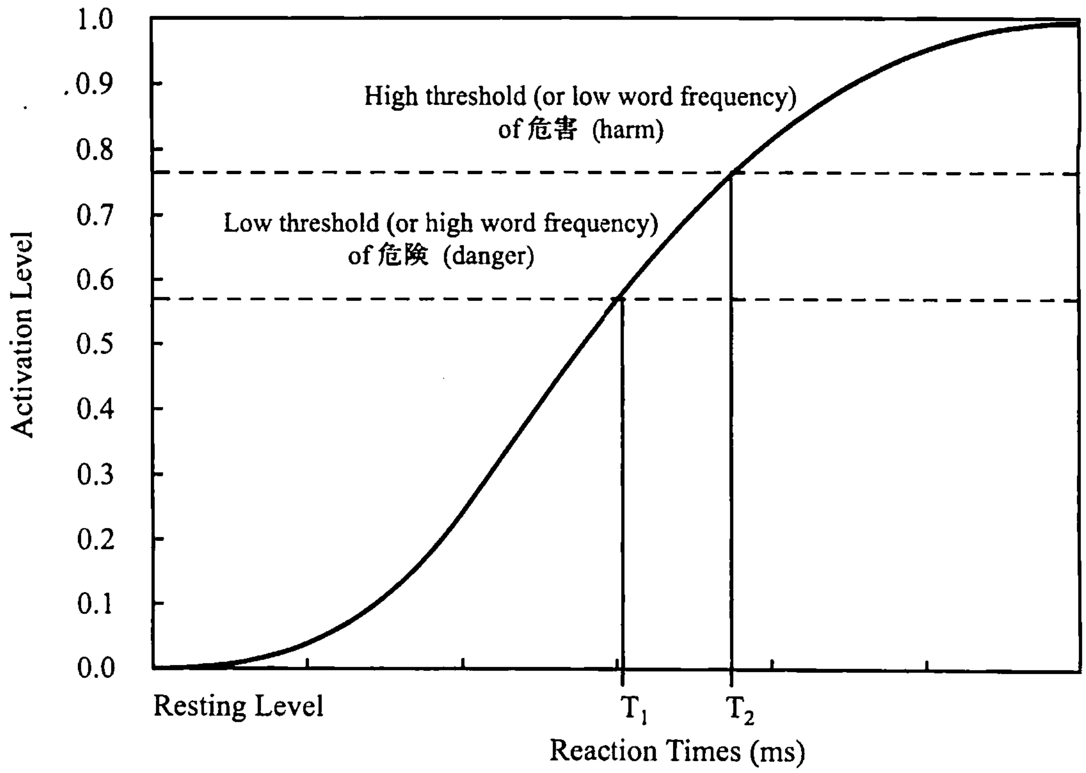

Figure 1. Height of thresholds depending on high and low word frequencies

the semantic meaning in this paper). The three representations are interactively activated in parallel distributed processing, the model for which being commonly referred to as the *triangle model* (e.g., McClelland and Elman, 1986; McClelland and Rumelhart, 1981; Seidenberg and McClelland, 1989; Taft, 1991; Tamaoka, 2005a). Each of three different representations in the mental lexicon can be activated fundamentally on the basis of a task requirement, but, to some degree, with automatic unintentional inter-activations.

Connectionists, who construct computer simulations of *learning* (or *acquisition*) according to neurological features of the human brain, provide a slightly different explanation from a broader perspective, by considering activations of multiple lexical items at the same time. They use the concept of *weights*, which is expressed by the following formula:

$net_i = \Sigma w_{ij} a_j$.

In this formula, each node (or a simulation unit for a neuron) receives multiple activations from multiple notes. Each connection between two nodes has a weight. A high frequency word has a heavier weight depending on the number of nodal connections it contains. The net of activations is the sum of a series of multiplications of weights and nodes as described by the formula above.

Furthermore, in the process of activating the target item, other lexical items are automatically activated as described in Figure 2. The thresh-

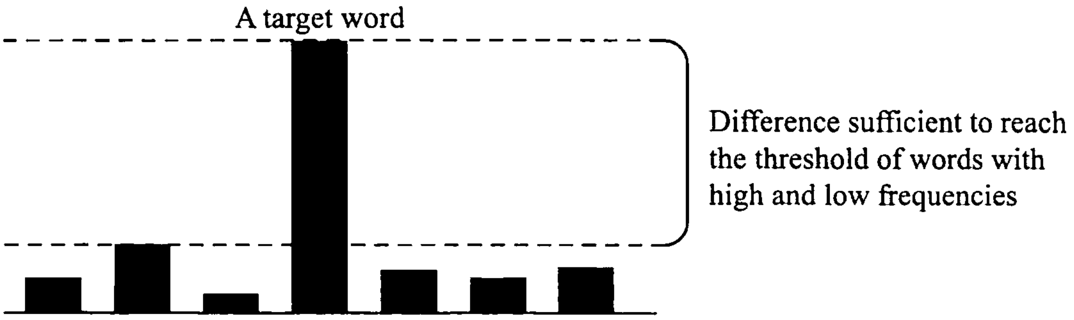

Figure 2. Difference in activation levels sufficient to reach the *threshold* of words with high and low frequencies

old is reached when the difference between the activation level of the target word and the secondly highly-activated lexical item becomes sufficiently large. In this sense, the threshold is expressed by a certain degree of constant difference. Since a high frequency word has heavier weights as determined by its connections, it will reach a larger difference within a shorter reaction time. In turn, a low frequency word has fewer weights in connections than a high frequency word; consequently, a low frequency word requires a longer reaction time to reach the threshold than a high frequency word. Again, this explanation rests upon the behavioral assumption that the more frequently words are seen, the faster they can be recognized. Since *frequency*—in either phonological or orthographic form—is determined by environment (e.g., parent-child communication in early childhood), any pre-assumed (i.e., innate) basic lexical or syntactic structure would not fall into this model. The connectionist explanation and approach does not automatically imply disregarding syntactic innateness, but so far, they do not consider innateness as fundamental to neurological *learning* simulation and its applied processing models.

3.1.2 Indexes of word familiarity and word frequency

Word frequency is an objective index taken from a lexical database as printed-counts, word familiarity is a subjective index asking (usually) native speakers to check 1-to-5 or 1-to-7-point scales which are designed to measure how familiar they are with certain words. Participants for the familiarity scaling questions are required only to *think* of each word. Longer thinking times often indicate a degree of unfamiliarity, which likewise indirectly result in longer reaction times. If an experimenter obtains a common sample group, such as university students, familiarity scaling and reaction times can normally be expected to have a similar index. This is one of the major reasons that word printed-frequencies receive priority for a control index for lexical stimuli in experiments compared to word familiarities.

In their commentary, MN discuss word frequency and word familiarity in critiquing the findings of our 2004 experiment. MN state 'KT's *pass*-type verbs are more frequent than their *show*-type verbs' (p. 122) and report that, according to Amano and Kondo (1999, 2003a for the CD-ROM version of lexical familiarity), the mean familiarity for nine *pass*-type verbs was 5.993 while the mean of seven *show*-type verbs was 5.728. As MN did not provide any statistical result for the difference in familiarity scales between the two types of verbs, we can only assume that they conducted an analysis for the difference. Without such information, we have to guess there was no significant difference in familiarity in light of the small difference of 0.265. Additionally, without any explanation provided in MN's comments, we fail to understand why they do not report having checked all ten verbs in each condition. We again must assume that MN did not find familiarity of certain verbs in Amano and Kondo (1999, 2003a). This null difference in familiarities between the *pass*-type and the *show*-type verbs cannot merely be disregarded, as it strengthens the overall findings of our 2004 experiment. However, being a subjective index, word familiarities are likely to show similar scales for relatively frequently-used verbs, especially comparing a very large number of lexical items in the 7-point scale as did Amano and Kondo (1999, 2003a).

In contrast to the result of familiarities, MN reported that word printed-frequencies of appearance in the *Asahi Newspaper* (Amano and Kondo, 2000, 2003b, CD-ROM for the index of word-printed frequency) showed a great difference between the *pass*-type verbs (mean frequency=26,443 times) and the *show*-type verbs (mean frequency=7,169 times). Again, we assume that since the difference in the average word frequencies between the two types of verbs is large at 19,274 times, the *pass*-type and *show*-type verbs will differ significantly without conducting a statistical test. As MN acknowledge, it is well-known (e.g., Besner and McCann, 1987; Carroll and White, 1973; Forster and Chambers, 1973; Gardner, Rothkopf, Lapan and Lafferty, 1987; Jescheniak and Levelt, 1994; Morrison and Ellis, 1995; Paap, McDonald, Schvaneveldt and Noel, 1987; Taft, 1979; Tamaoka and Takahashi, 1999) that word frequency affects the speed of performing tasks such as lexical decision, word naming, and word writing-initiation, and likely also plausibility judgment. As we explained earlier in this section, high frequency words have lower thresholds or heavier weights, which make these words more easily (quickly and accurately) recognizable than low frequency words. As such, MN stated that 'if the plausibility task is affected by verb frequency, as seems to be the case, then it may also be affected by the frequency of the word orders' (p. 122).

However, the scrambling condition in our study (as well as other studies such as Muraoka, Tamaoka, Miyaoka, 2004, 2005; Tamaoka, 2005b; Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi, 2005; Tamaoka, Miyaoka, Ito and Sakai, 2004; Tamaoka, Lim and Miyaoka, 2005) was created by simply re-ordering a noun phrase. For example, the assumed canonical sentence, *Kenzi-ga Kazuko-ni nooto-o kaesita* [Kenzi-NOM [Kazuko-DAT [notebook-ACC return-PAST]]] is scrambled to *Kenzi-ga nooto-o Kazuko-ni kaesita* [Kenzi-NOM [nootbook-ACC$_1$ [Kazuko-DAT [*gap*$_1$ return-PAST]]]]. The plausibility judgment task requires decisions as to whether or not a sentence is correct, after reading through the whole sentence. As a result, all lexical items at the end of sentences including verbs are identical between the canonical and scrambled sentences; consequently, all lexical items used in the both canonical and scrambled sentences have equal threshold levels or equal weights in connections, leading to equal effects of word frequencies. In other words, if only word frequencies affect sentence processing, null effects are expected to be obtained from the performance of the plausibility judgment task when comparing reaction times and error rates between the two conditions of canonical and scrambled sentences (whichever is canonical). Furthermore, we did not conduct a direct comparison of sentences between *pass*-type and *show*-type verbs, but rather compared these lexically identical canonical and scrambled sentences in each condition, which led to the same results for both verb conditions; namely the word order of *ga-ni-ō* (NOM-DAT-ACC) was faster than *ga-o-ni* (NOM-ACC-DAT). Again, because word frequency effects are equalized in the canonical and scrambled sentences in our study, MN's comment (p. 112) applies to neither our findings, nor those of other studies.

3.2 Syntactic structural frequency

MN's comment on frequency effects is fundamentally focused on the second type of frequency, *syntactic structural frequency*. As discussed by Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi (2005), case particles are not the lone source of information in identifying word order: thematic roles and grammatical functions are also equally important candidates. In fact, the main purpose of Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi (2005) was to clarify priority information for identifying canonical word order as determined by scrambling effects. A series of five experiments in the above-mentioned Tamaoka et al. (2005) study suggested that only grammatical functions could provide sufficient information for canonical word order for various types of sentences. Putting the argument of priority information aside, case particles are easier to search for in the corpus by simply using a collocation for nouns with particle kana symbols (i.e., が, は, に, を) appearing with a verb.

3.2.1 Fundamental chicken-or-egg question of canonical order

Taking an example from the Appendix of our study, a sentence of a *pass*-type verb *Taro-ga Zyunko-ni dengon-o tutaeta* [Taro-NOM [Zyunko-DAT [message-ACC convey-PAST]]] is defined as a NOM-DAT-ACC word order. As Hoji (1985) proposed, once this NOM-DAT-ACC word order is defined as canonical, the word order NOM-ACC-DAT of *Taro-ga dengon-o Zyunko-ni tutaeta* [Taro-NOM [message-ACC$_1$ [Zyunko-DAT [*gap*$_1$ convey-PAST]]] can be considered as scrambled. Miyagawa (1997) claimed that both word orders can be base-generated (i.e., canonical) without syntactic movement (or *gap*). Furthermore, Matsuoka (2003) distinguished two types of canonical word order based on verbs; *show*-type for NOM-DAT-ACC and *pass*-type for NOM-ACC-DAT. Our study supported the proposal by Hoji (1985) that this NOM-DAT-ACC word order is canonical regardless of any verb type. MN's commentary paper argued that since syntactic structural frequency of NOM-DAT-ACC was much higher than NOM-ACC-DAT in both *pass*-type and *show*-type verbs, the results of our study were caused by this difference in syntactic structural frequency (hereafter, simply syntactic frequency).

The issue of syntactic frequencies in corpora fundamentally contains a chicken-or-egg question. If the assumed canonical order of the NOM-DAT-ACC is base-generated, native Japanese speakers are likely to produce it naturally. As a result, syntactic frequency in the corpora would also indicate the same direction; the NOM-DAT-ACC order would appear more frequently than the NOM-ACC-DAT order. On the contrary, it is possible to argue that because the NOM-DAT-ACC order is used more frequently than the NOM-ACC-DAT order, Japanese children acquire this word order as canonical. The former explanation accepts the innateness of syntactic structure providing canonical order, whereas the later explanation fits nicely into the connectionist idea of enhancement by learning from one's environment; namely, children *frequently* hear or see a certain word order of sentences.

3.2.2 Tentative solution for the chicken-or-egg question on canonical order

The third (passive sentences) and fourth (potential sentences) experiments of Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi (2005, comprising two of five experiments of a five-experiment project) showed an interesting contrast, providing a tentative solution for the chicken-or-egg question of precedence. In the third experiment, the *ga-ni* word order in passive sentences (e.g., *Kazuko-ga Taro-ni odos-are-ta*, Kazuko-NOM Taro-DAT threaten-PASS-PAST, 'Kazuko was threatened by Taro') was performed more quickly (*ga-ni* for 1,521 ms and *ni-ga* for 1,722 ms in reaction times) and accurately (*ga-ni* for 1.85% and *ni-ga* for 6.25% in error rates) than the word order of *ni-ga* (e.g., *Taro-ni Kazuko-ga* *odos-are-ta*, Taro-DAT Kazuko-NOM threaten-PASS-PAST). In contrast, in the fourth experiment, the word order of *ni-ga* or DAT-NOM in potential sentences (e.g., *Takashi-ni girishago-ga kak-eru-daroo-ka*, Takashi-DAT Greek-NOM write-POT-wonder-Q, 'I wonder if Takashi can write Greek?') was performed for the plausibility judgment task more quickly (*ni-ga* for 1,326 ms and *ga-ni* for 1,542 ms in reaction times) and accurately (*ni-ga* for 4.17% and *ga-ni* for 29.86% in error rates) than the word order of *ga-ni* or NOM-DAT (e.g., *Girishago-ga Takashi-ni kak-eru-daroo-ka*, Greek-NOM Takashi-DAT write-POT-wonder-Q). It should be also added that Muraoka, Tamaoka and Miyaoka (2004) and Fukumitsu, Kim and Koizumi (2004) also found that *ga-ni* was processed faster than ni-ga in active sentences with dative verbs (e.g., *Suzuki-ga Yamada-ni at-ta*, Suzuki-NOM Yamada-DAT meet-PAST 'Suzuki met Yamada' versus *Yamada-ni Suzuki-ga at-ta*, Yamada-DAT Suzuki-NOM meet-PAST). As such, *ni-ga* is the canonical order for potential sentences while *ga-ni* is for passive sentences and active sentences with dative verbs. Since an overall assumed-estimation from the wider perspective of syntactic frequency would indicate *ga-ni* to be a canonical order, these contrasting experimental results imply that syntactic structure can be independent from syntactic frequency in the broad sense. The same point can be made with the study of causative sentences mentioned above in relation to the examples in (5) and (6) (Experiment 5 of Tamaoka, et al., 2005). The word order of *ni-o* or DAT-ACC is more frequently used than the *o-ni* or ACC-DAT order. If the more frequently used word order in this sense is processed faster regardless of the type of verb involved, the reaction times for (6b) should be shorter than (6a). However, this assumption is not held up by the results of reaction times in sentence processing indicated by Experiment 5 of Tamaoka et al. (2005).

Yet, MN state their position that our conclusion regarding syntactic frequency was 'premature' (p. 119). In so doing, MN argue the importance of syntactic frequency from a narrower perspective, stating that 'we may need to count instances of *ni-ga* followed by potentials separate from those instances followed by passive' (p. 119). In this sense, syntactic frequency still remains unsolved, while syntactic frequency can be separated from syntactic structure from a wider perspective. From the narrower perspective of syntactic frequency, once again, we theoretically encountered the same chicken-or-egg question in terms of the acquisition of syntactic structure.

3.2.3 Effects of animacy

Since MN briefly mentioned 'animacy may also be an issue' (p. 119) with no detailed explanation, we also report an interesting experiment on animacy using the same experimental paradigm of the plausibility judgment task as conducted by Muraoka, Tamaoka and Miyaoka (2005). Muraoka et al. (2005) used four types of sentences to simultaneously investigate factors of animacy and word order; (1) *Yoshikawa-san-ga seito-o tataku*, Yoshikawa-san-NOM student-ACC hit-PRET 'Yoshikawa hits a student' for the condition of canonical word order with an animate object, (2) the same sentence in scrambled order as *Seito-o Yoshikawa-san-ga tataku*, student-ACC Yoshikawa-san-NOM hit-PRET, (3) *Yoshikawa-san-ga taiko-o tataku*, Yoshikawa-san-NOM drum-ACC hit-PRET 'Yoshikawa-san hits a drum' for the condition of canonical word order with an inanimate object, (4) the same sentence in scrambled order as *Taiko-o Yoshikawa-san-ga tataku*, drum-ACC Yoshikawa-san-NOM hit-PRET. Muraoka et al. (2005) created 28 sets (112 sentences in total) of these correct 'Yes' responses (i.e., correct sentences) with the same number of correct 'No' responses (i.e., incorrect sentences) in the counter-balanced (or the Latin-square method) condition, and dummy sentences (both correct and incorrect sentences).

Since word printed-frequencies of animate and inanimate objects were controlled to be constant (other lexical items were the same) using the database of Amano and Kondo (2000, 2003b), reaction times required for correct 'Yes' responses of the plausibility judgment task were directly analyzed by a 2 (word order: canonical or scrambled)×2 (animacy: animate or inanimate) analysis of variance (ANOVA). The means and standard deviations are depicted in Figure 3. As one can guess from Figure 3, the results indicated significant main effects of both word order and animacy, but no interaction of both variables. The results suggested two independent trends: (1) sentences with canonical word order were more quickly processed than those with scrambled word order regardless of object animacy, and (2) sentences with inanimate objects were more quickly processed than those with animate objects regardless of word order. Responding to MN's comment on animacy, the findings of Muraoka et al. (2005) clearly indicated that the factor of animacy was independent from the factor of word order.

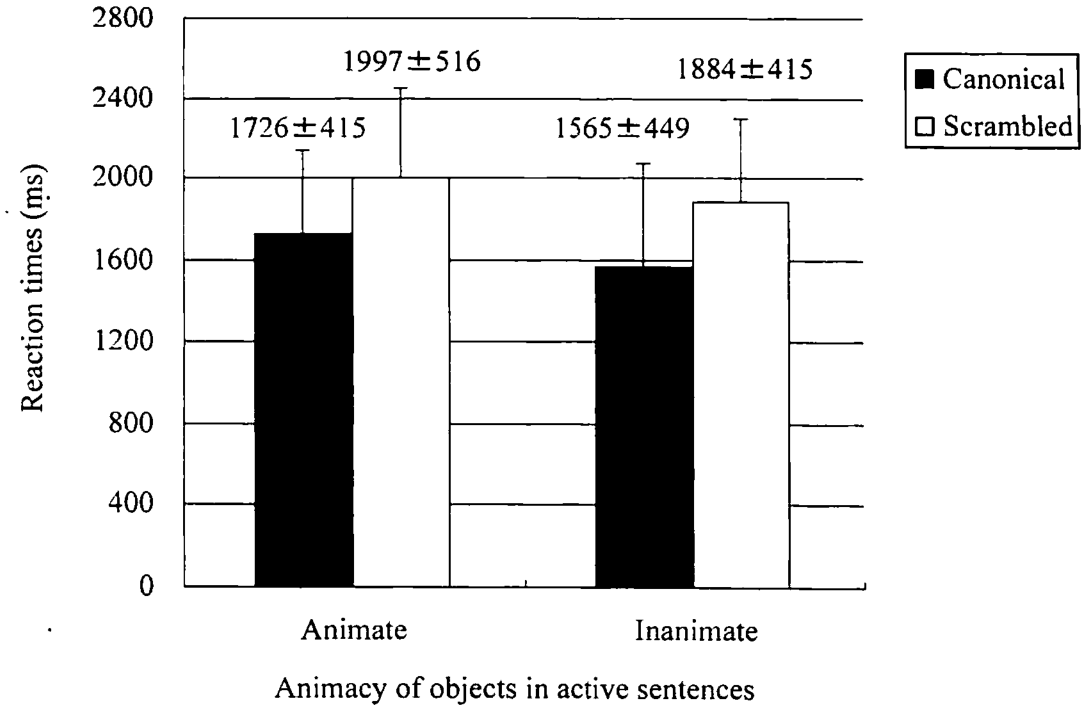

Figure 3. Scrambling effects of active sentences with animate and inanimate objects

Note 1: Values are taken from Muraoka, Tamaoka and Miyaoka (2005).

Note 2: Values before ± refer to the means of reaction times in *milliseconds* while values after it refer to standard deviations.

3.2.4 MN's corpus study on sentences with ditransitive verbs

Using 38,383 sentences in the Kyoto University Corpus (Kurohashi and Nagano, 1997), MN reported two corpus studies counting syntactic frequency regarding sentences with ditransitive and transitive verbs. Of all these sentences, MN found 305 instances (0.79% of the total) with a NP-*ni* and NP-*o* in the same clause. MN refers to these sentences as the *coarse-grained sentences*. According to verb types used in our study, these 305 sentences were classified into three types of verbs, *pass*-type, *show*-type for plausible items and *pass*-type verbs used for implausible items (i.e., items used for correct 'No' responses) which are reported in MN's Table 1 (p. 121). With an insufficient explanation for analysis of syntactic frequency in Table 1, it is difficult to know how MN reached the statement that 'overall, the *ni-o* order is more frequent' (p. 121). MN (p. 121) reported χ^2=94, p<.0001 without specifying the type of chi-square test they used and without reporting the degree of freedom (even a chi-square value of the decimal point was 0). Since we do not know how MN analyzed the frequency data, we conducted an analysis by ourselves.

Our experiment used ten *pass*-type and ten *show*-type verbs for analyzing reaction times of correct 'Yes' responses. Since correct 'No' responses differ from correct 'Yes' responses in cognitive nature of sentence processing, implausible items used for correct 'No' responses was separately analyzed. Thus, if MN intended to compare their findings in terms of tendency of syntactic frequency with our results in reaction times, MN should conduct a statistical test for syntactic frequencies among the coarse-grained sentences using a table of *pass-/show*-type and *ni-o/o-ni* order as shown in Table 1 of this paper. Therefore, we conducted the chi-square test of goodness-of-fit, expecting a chance level of frequency in *ni-o* and *o-ni* of *pass*-type (expected frequency of 90.50) and *show*-type (expected frequency of 26.00) verbs as shown in Table 1 (see expected frequencies of Expected A), yielded a chi-square value of 70.547 with a degree of freedom of 1 for the *pass*-type verbs and a chi-squire value of 24.923 with a degree of freedom of 1 for the *show*-type verbs. Both the

Table 1. Syntactic frequency for coarse-grained ditransitive sentences based on MN approach

	Verb types						
	Pass-type frequency			*Show*-type frequency			
Word order	Observed	Expected A	Expected B	Observed	Expected A	Expected B	Total
ni-o	147	90.50	148.37	44	26.00	42.63	191
o-ni	34	90.50	32.63	8	26.00	9.37	42
Total	181			52			233

Note 1: Expected frequency A is used for the chi-square test of goodness-of-fit.
Note 2: Expected frequency B is used for the chi-square test of independence.

chi-square values were significant at the 0.1 percent level, p<.001 (we use p<.001 for the maximum probability in reporting). Thus, our conclusion for the chi-square test of goodness-of-fit suggests that the *ni-o* and *o-ni* order do not appear at the chance level, further indicating that the *ni-o* order was likely to be observed more frequently than the *o-ni* order in both cases of the *pass*- and *show*-type verbs.

Our chi-square test of goodness-of-fit, however, does not provide information concerning a frequency difference of the *ni-o* and *o-ni* word order between *pass*-type and *show*-type verbs. To compare the difference in frequency between the two types of verbs, a chi-square test of independence should be conducted using the expected frequencies in Expected B of Table 1. For example, using observed frequencies in Table 1, an expected frequency of the *ni-o* order of the *pass*-type verbs was calculated by the row total (i.e., 191) multiplied by the column total (i.e., 181) with this result divided by the grand total (i.e., 233), which resulted in the expected frequency of 148.37. Other cells follow the same manner of calculation. The result of the chi-square test of independence (the calculation formula is the same as the chi-square test of goodness-of-fit. Please refer to the calculation procedure in an appropriate statistic textbook), examining whether or not the two factors of word order and verb type are independent (in other words, not related), did not reach significance [$\chi^2(1)=0.316$, p=.574, *n.s.*]. This result suggested an equal proportion of frequency of the *ni-o* and *o-ni* word order under the conditions of the both *pass*-type and *show*-type verbs. Therefore, putting both chi-square test results of goodness-of-fit and independence together, we conclude that syntactic frequency of the coarse-grained ditransitive sentences shows that, regardless of the *pass*- or *show*-type verbs, the *ni-o* order appears more frequently than the *o-ni* order. This result is congruent with our earlier experimental findings, indicating that sentences of *ni-o* order are processed more quickly than the *o-ni* order in both cases of *pass*- and *show*-type verbs. This likelihood of syntactic frequency of the coarse-grained ditransitive sentences supported the proposal of Hoji (1985), but not Matsuoka (2003). However, once again, since MN's frequency counts and our experimental results indicated the same pattern, we again have to face the chicken-or-egg question: what determines word order—syntactic frequency or syntactic structure?

MN (p. 121) provided an interesting figure of syntactic frequency using the 63 fine-grained sentences specified in our experiment, which actually gives favorable evidence to support our experimental results. For the same reasons regarding the analysis shown in Table 1, the two different chi-square tests were conducted with expected frequencies indicated in Table 2 of this paper. First, the chi-square test of goodness-of-fit using expected frequencies (see Expect A in Table 2) showed no significance in either the *pass*-type verbs [$\chi^2(1)=0.022$, $p=.881$, *n.s.*] or the *show*-type verbs [$\chi^2(1)=0.222$, $p=.637$, *n.s.*]. This result indicated no difference in syntactic frequencies between the *ni-o* and *o-ni* orders in either *pass*-type or *show*-type verbs. In addition, the chi-square test of independence also showed no significance [$\chi^2(1)=0.229$, $p=.633$, *n.s.*]. As such, both chi-square tests suggest that syntactic frequencies of the *ni-o* and *o-ni* word order do not differ among the fine-grained sentences in either the *pass*-type or *show*-type verbs. Therefore, unlike the results from the coarse-grained sentences in Table 1, word order was neutral on the basis of the fine-grained sentences in Table 2, which supports neither Hoji (1985) nor Matsuoka (2003) from the perspective of syntactic frequency. Instead, these results can be interpreted as indicating that our experimental results were caused only by syntactic structure or incremental syntactic processing with no influence of syntactic frequency, supporting our experimental findings, which in themselves supported those proposed earlier by Hoji (1985).

However, despite the analysis and conclusion described above, MN state that 'frequency in this narrow sense cannot explain the advantage for *ni-o* in KT's study' (MN, p. 121). Then, MN arrive at the statement 'the *ni-o* order is more frequent' (p. 121) based on the subtotals of MN's Table 2 (p. 121). MN's calculation of the chi-square test indicated $\chi^2=8.58, p<.01$, again with no explanation for the type of chi-square test they used and no report of the degree of freedom. Thus, we conducted the chi-square test of goodness-of-fit for the *ni-o* and *o-ni* word order separately in the *pass*- and *show*-type verbs using expected frequency indicated in Expected A of Table 3. The result showed that there is significance in frequency for the *ni-o* and *o-ni* orders in the *pass*-type verbs [$\chi^2(1)=9.574, p<.01$]. In contrast, no significance is found in the *show*-type verbs [$\chi^2(1)=0.222, p=.637$, *n.s.*] which is identical to the above-mentioned result since frequencies have not changed. Thus, the *ni-o* word order appears more frequently than the *o-ni* order in the *pass*-type verbs, but not the *show*-type verbs.

Furthermore, to examine the relations of factors between word order

Table 2. Syntactic frequency of the fine-grained ditransitive sentences as specified as in KT (2004)

	Verb types						
	Pass-type frequency			*Show*-type frequency			
Word order	Observed	Expected A	Expected B	Observed	Expected A	Expected B	Total
ni-o	22	22.50	22.86	10	9.00	9.14	32
o-ni	23	22.50	22.14	8	9.00	8.86	31
Total	45			18			63

Note 1: Expected frequency A is used for the chi-square test of goodness-of-fit.
Note 2: Expected frequency B is used for the chi-square test of independence.

Table 3. Syntactic frequency of fine-grained ditransitive sentences with all usages (subtotals) based on MN

	Verb types						
	Pass-type frequency			*Show*-type frequency			
Word order	Observed	Expected A	Expected B	Observed	Expected A	Expected B	Total
ni-o	62	47.00	60.43	10	9.00	11.57	72
o-ni	32	47.00	33.57	8	9.00	6.43	40
Total	94			18			112

Note 1: Expected frequency A is used for the chi-square test of goodness-of-fit.
Note 2: Expected frequency B is used for the chi-square test of independence.

and verb types, we conducted a chi-square test of independence for the observed and expected frequencies (see Expected B of Table 3). The chi-square test of independence, examining whether or not the two factors of *ni-o/o-ni* word order and *pass/show*-type verbs are independent, did not indicate significance [$\chi^2(1)=0.712$, $p=.399$, *n.s.*]. Although the chi-square test of goodness-of-fit showed that the *ni-o* order appeared more frequently than the *o-ni* order only in the *pass*-type verbs, the chi-square test of independence suggested that the syntactic frequency of *ni-o/o-ni* word order has nothing to do with the *pass/show*-type verbs.

We disagree with the use of subtotals in calculating syntactic frequency of *ni-o/o-ni* word order (see in MN's Table 2, p. 121). First, MN classified the fine-grained sentences into the four subcategories which clearly show syntactic frequencies of the *pass*- and *show*-type verbs in relation to our experiment. Their results turned out to strongly support those of our experiment. Although the other three subcategories appear only with the *pass*-type verbs, MN insist on using subtotals in their Table 2. However, the small frequency of the *show*-type verbs has no increase of syntactic frequency in either KT subcategory or subtotal; there is no difference in frequency between subcategory and subtotals. Second, since MN nicely identified syntactic frequency for verbs used in our experiment, MN should not add other impure frequencies. We prefer to support the results using the fine-grained sentences subcategorized only in our experiment (i.e., Table 2 of this paper), suggesting the *ni-o* and *o-ni* word order appears at the chance level regardless of the *show-* or *pass*-type of verbs.

3.2.5 MN's corpus study on sentences with transitive verbs

MN further analyzed syntactic frequency of transitive sentences, extracting sentence-initial accusative NPs out of the total of 38,383 sentences of the Kyoto University Corpus (Kurohashi and Nagano, 1997). The results of classification and frequencies are reported in MN's Table 3 (p. 123), suggesting that 4,537 transitive sentences or 98.1 percent of the total of 4,625 had no subject after NP-*o* while only 88 sentences or 1.9 percent of the total had subjects after NP-*o*. As MN provided a very high chi-square value of 4365.4, again without any explanation, we assume that they conducted the chi-square test of goodness-of-fit expecting a chance level of occurrences with and without subjects. Our calculation, however, indicates a result of $\chi^2(1)=4279.70$, $p<.001$, which is different from MN's value of 4365.4. Again, we do not know how MN calculated the chi-square value.

Besides the issue of chi-square calculation, based on frequency of transitive sentences in MN's Table 3, MN argued that 'in the majority of the *no subject* instances, the accusative NP was next to the predicate' (p. 123), and concluded 'a fronted accusative NP is highly unlikely to be scrambled prior to an overt subject' (p. 123). This argument is reasonable; since a subject phrase only appears in 88 cases or 1.9 percent after NP-*o*, MN can assume that the scrambled cases are very rare. Again, experimental results of scrambled effects could be caused by the rarity of scrambled conditions of sentences in general. However, MN used the argument that 'we may need to count instances of *ni-ga* followed by potentials separate from those instances followed by passive' (p. 119) to disagree with the comparison of the third and fourth experiments by Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi (2005). This argument now returns to MN: we do not know how this argument related to our 2004 experiment conducted for ditransitive verbs. The frequency figures given in MN's Table 3 (p. 123) seems too broad to effectively contest our 2004 experimental results.

MN report on their sentence completion study for object NPs which required participants to complete a sentence after accusative and dative NPs such as *hukuoka syussin-no syain* 'the employee from Fukuoka' with accusative *-o* or dative *-ni*. Using instances with no subjects, MN report the results in Table 5 of their paper (p. 125). MN conclude 'NP-*o* tends to be adjacent to the predicate (64.7%) and is rarely followed by another object (2.4%), whereas a NP-*ni* is equally likely to be followed by the verb (36.7%) or another object (33.3%)' (p. 124). Once again, MN report a chi-square value without mentioning the degree of freedom, χ^2=82.47, p<.0001. This single chi-square test cannot serve as a basis for this conclusion, although, looking only at frequencies, we can guess this trend. To examine MN's claim using proper statistic procedures, we produced a table for the chi-square tests. First, as shown in Table 4 of this paper, the chi-square test of goodness-of-fit should be conducted separately for NP-*o* and NP-*ni*. As MN indicated, observed frequencies of NP-*o* (using a chance level of expected frequency 83.50 in Expected A of Table 4) show that NP-*o* is likely to be adjacent to the predicate [$\chi^2(1)$=143.862, p<.001]. Likewise, the same chi-square test of goodness-of-fit (expected frequency of 83.00) indicate that NP-*ni* is equally followed by a verb

Table 4. Frequency with no subject in the completion task for object NPs (MN Table 5, p. 125)

Intervention between object and verb	Initiation of sentence						Total
	NP-*o*			NP-*ni*			
	Observed	Expected A	Expected B	Observed	Expected A	Expected B	
adjacent	161	83.50	124.37	87	83.00	123.63	248
NP-*ni* or NP-*o*	6	83.50	42.63	79	83.00	42.37	85
Total	167			166			333

Note 1: Expected frequency A is used for the chi-square test of goodness-of-fit.
Note 2: Expected frequency B is used for the chi-square test of independence.

and another object [$\chi^2(1)=0.386$, $p=.535$, *n.s.*]. The relation between NP-*o*/NP-*ni* and following/adjacent object is examined by the chi-square test of independence. As shown in Table 4 of this paper, the test was conducted using expected frequencies in Expected B, and showed significance [$\chi^2(1)=84.773$, $p<.001$]. Although our conclusion is the same as MNs, their claim is not entirely convincing without explanations regarding the type of chi-square test and correct calculations.

Overlooking the incorrect chi-square value calculated by MN, their claim as to *anticipatory processes* is very insightful and nicely supported by the syntactic frequency data in their completion task. Based on MN's frequency data shown in Table 4 of this paper, it will be hypothesized that NP-*ni* will be followed by both NP-*o* and a verb while NP-*o* anticipates only a verb. This assumption provides a general likelihood that NP-*ni* will be followed by NP-*o* or a verb when NP-*ni* is initiated in a sentence, and furthermore it is generalized that the *ni-o* word order is expected to occur more frequently in ditransitive sentences. Although MN's task is not exactly equivalent to a printed or speech corpus, we do think that the index of stratified human production as used in MN's task can be considered as a type of syntactic frequency. MN again proposed a possible factor of syntactic frequency influence upon scrambling effects as observed by psychological experiments including our 2004 study, as they also did in the aforementioned corpus study. Needless to say, this argument faces the chicken-or egg question of innateness, and fails to account for the results of Tamaoka et al.'s (2005) Experiment 5. One should also keep in mind, however, that frequency data of a corpus, including MN's sentence competition task, does not provide information of real-time sentence processing and is far from the self-paced reading which MN seem to strongly support.

3.2.6 Summary—syntactic complexity and frequency

To summarize, syntactic complexity is inversely correlated with syntactic structural frequency, and syntactic complexity predicts relative reaction

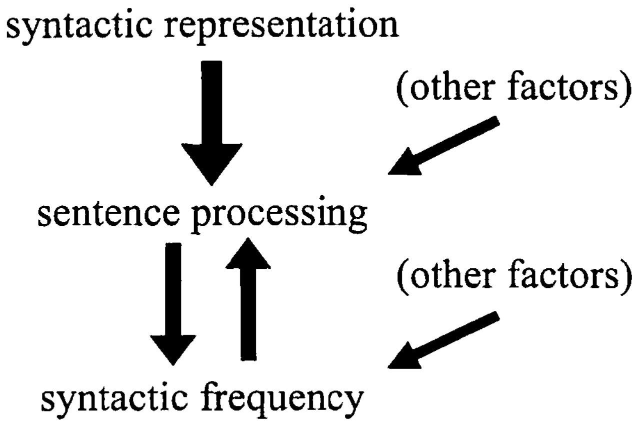

Figure 4. Relationship among syntactic representation, processing, and frequency

times for the sentence plausibility judgment task more reliably than the syntactic frequency in the broad sense and at least as reliably as the syntactic frequency in the narrow sense. Given that syntactic structure has ontological priority over syntactic frequency and reaction times, we suggest the following relationship among these factors: Syntactic complexity affects syntactic processing (both comprehension and production) in the sense that the former partially determines cognitive load of the latter. Syntactic processing in turn affects syntactic frequency because, all other things being equal, a structure which is easier to process is more likely to be used than one which is more difficult to process. As more frequently used structures may be easier to process, syntactic frequency also affects syntactic processing. This model is schematically shown in Figure 4.

3.3 Collocation frequency

The third type of frequency is *collocation frequency*, referring to how frequently a certain lexical item is combined with others within the same syntactic structure. For example, the intransitive verb *okureru* 'to be delayed' (hereafter, simply 'delay') can take the subject *kaihuku* 'recovery' to construct an active sentence *Kaihuku-ga okureru*, recovery-NOM delay-PRET 'Recovery is delayed'. In the same manner, the verb 'delay' can be combined with the subject *happyoo* 'announcement' as *Happyoo-ga* *okureru*, announcement-NOM delay-PRET 'Announcement is delayed'. Both sentences have the same syntactic structure of NP-*ga* plus V and share the same intransitive verb *okureru* 'delay'. According to the web-based corpus search program *Chasen* designed by Jun Fukada at Purdue University, USA using a corpus of the nine years of *Mainichi Newspaper* published from 1991 to 1999, we can find that *kaihuku* 'recovery' appears 30,748 times while *happyo* 'announcement' similarly appears 30,254 times. However, collocation frequency of these two nouns (i.e., subject) and the verb differs: the verb 'delay' co-appears 39 times with 'recovery', but it co-appears only 5 times with 'announcement'. In this case, the difference of collocation frequency is 34 times (39 minus 5). Taking that the syntactic structure is the same, if any difference in sentence processing is found, it must be caused by collocation frequency of a subject noun and an intransitive verb.

The effects of collocation frequencies were investigated by Tamaoka and Tanaka (2005) using 24 sets of the aforementioned experimental

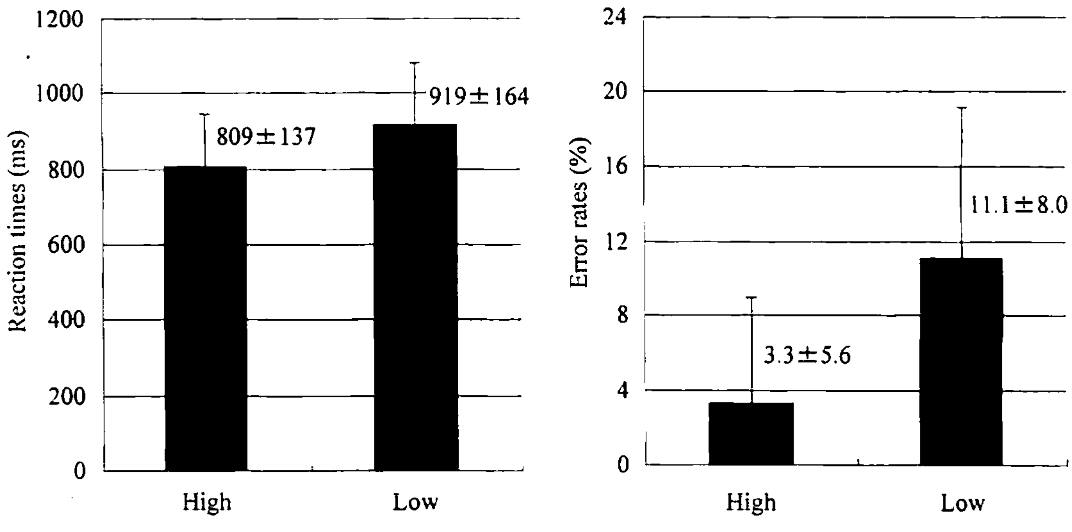

Figure 5. Reaction times for high and low collocation frequencies of active sentences with intransitive verbs

Note 1: Values are taken from Tamaoka and Tanaka (2005)

Note 2: Values before ± refer to the means of reaction times in milliseconds (left) or error rates in percentage (right) while values after it refer to standard deviations.

stimulus pair as correct 'Yes' responses. The means of reaction times and error rates required for the plausibility judgment task are depicted in Figure 5. A paired-sample *t*-test indicated a significant difference between the high and low collocation frequency conditions in both reaction times and error rates. Intransitive sentences with high collocation frequencies were processed more quickly and accurately than those with low collocation frequencies. Consequently, this experiment clearly illustrates co-activations of two lexical items affecting sentences constructed by a subject noun and an intransitive verb. As graphically seen in Figure 5, independent of syntactic structure, collocation frequency, which implies how frequently two lexical items are combined, is also the important factor. Although MN only refer to frequencies at the word and syntactic levels in their commentary, the third type of collocation frequency must play an important role for sentence processing. However, it should be noted that simple collocation frequency of two lexical items (i.e., a subject noun and an intransitive verb) can be found using a corpus available today, but collocation frequencies of three lexical items in transitive sentences and of four lexical items in ditransitive sentences requires an enormous number of corpora, so that investigation in the psycholinguistic arena is still in its infancy.

4. Measurement methods—plausibility judgment and self-paced reading

MN note that 'the plausibility task only measures the total time to read the whole sentence and therefore it is not very informative' (p. 118). By this statement, MN place great value on the method of *self-paced reading*, stating that 'recent work in sentence processing has measured behavioral or neurological responses time-locked with the processing of individual words in sentences' (p. 118). The self-paced reading requires participants to read one part (often a single phrase) of a sentence at a time and press a button (usually a space key) to see the next part. The duration time between key pressings is interpreted as a reading time for each sentential part. The idea that self-paced reading can measure reaction times required for reading each part (often phrase), seems to be appropriate for investigating sentence processing, especially from the syntactic perspective. However, in reality, this method has seldom detected scrambling effects (e.g., Nakayama, 1995; Tamaoka, Sakai, Kawahara and Miyaoka, 2003; Yamashita, 1997), and as MN themselves mention, scrambling effects measured by self-paced reading 'have only reached statistical significance when complex structures were used' (p. 118). What initially appears to be an ideal method, therefore, is seen to contain inherent problems upon further consideration.

4.1 Scrambling effects as shown by plausibility judgment

Prior to examining the method of self-paced reading, we will briefly review recent studies that utilized the plausibility judgment task. As shown in Table 5, four psycholinguistic experimental studies found scrambling effects on the processing of Japanese sentences in both participant and item analyses. The size of scrambling effects is measured by reaction times needed for assumed canonical sentences minus reaction times for scrambled sentences. First, our study in question, found 98 milliseconds of scrambling effects in the case of active sentences with *pass*-type verbs while 109 milliseconds for those with *show*-type verbs. The size of scrambling effects observed in Koizumi and Tamaoka (2004) was somewhat smaller. Muraoka, Tamaoka and Miyaoka (2004) observed greater than 200 milliseconds of scrambling effects; 229 milliseconds for sentences with accusative verbs and 214 milliseconds for those with dative verbs. The paper by Muraoka, Tamaoka and Miyaoka (2005) regarding animacy observed even stronger scrambling effects of 271 milliseconds for animate objects and 319 milliseconds for inanimate objects.

The Tamaoka et al. (2005) experiments listed in Table 5 were conducted using various types of sentences. Although the primary objective of this study was to identify priority information for sentence processing, we now are considering only the size of scrambling effects. Scrambling

Table 5. Scrambling effects on correct 'Yes' responses for sentence correctness decision task

Papers and experimental conditions	ΔRT (ms)	F_1sig.	F_2sig.
(1) Japanese sentences processed by native Japanese speakers			
1. Koizumi and Tamaoka (2004)			
Exp. 1(1)—Active sentences with *pass*-type verbs	98	*	*
Exp. 1(2)—Active sentences with *show*-type verbs	109	*	*
2. Muraoka, Tamaoka and Miyaoka (2004)			
Exp. 1(1)—Active sentences with accusative verbs	229	***	*
Exp. 1(2)—Active sentences with dative verbs	214	***	*
3. Muraoka, Tamaoka and Miyaoka (2005)			
Exp. 1—Active sentences with animate objects	271	*Scrambling*	
Active sentences with inanimate objects	319	**	***
4. Tamaoka, Sakai, Kawahara, Miyaoka, Lim and Koizumi (2005)			
Exp. 1—Active sentences with transitive verbs	223	***	***
Exp. 2—Active sentences with ditransitive verbs	604	***	***
Exp. 3—Passive sentences	201	***	***
Exp. 4—Potential sentences	216	***	***
Exp. 5(1)—Causative sentences with accusative verbs	187	*	**
Exp. 5(2)—Causative sentences with dative verbs	185	**	*
(2) Turkish sentences processed by native Turkish speakers			
Tamaoka, Kuribayashi and Sakai (2005)			
Exp. 1(1)—Active sentences with animate objects	155	*	*
Exp. 1(2)—Active sentences with inanimate objects	148	**	*
Exp. 2—Active sentences with subject incorporation	-209	***	*n.s.*
(3) Korean sentences processed by native Korean speakers			
Tamaoka, Lim and Miyaoka (2005)			
Exp. 1—Active sentences with transitive verbs	328	***	***
Exp. 2—Active sentences with ditransitive verbs	615	***	***
(4) Japanese sentences processed by native Chinese speakers			
Tamaoka (2005b)			
Exp. 1—Active sentences with transitive verbs	367	*	*
Exp. 2—Potential sentences	369	*n.s.*	*n.s.*

Note 1: ΔRT (reaction times in milliseconds) refers to the degree of scrambling effects calculated by RTs of scrambled condition minus RTs of canonical condition.

Note 2: *$p<.05$. **$p<.01$. ***$p<.001$. *n.s.* refers to not significant.

Note 3: F_1 refers to the results of participant analyses while F_2 refers to those of items analyses.

Note 4: Muraoka, Tamaoka and Miyaoka (2005) was analyzed by 2 (animacy: animate and inanimate)×2 (word order: canonical or scrambled) ANOVA. Both main effects were significant in participant and item analyses.

effects for active sentences with transitive verbs is 223 milliseconds. This size shows a dramatic increase of up to 604 milliseconds for active sentences with ditransitive sentences. Although the two experiments of transitive and ditransitive sentences cannot be directly compared (involvement of different lexical items), this great increase of scrambling effects is likely due to the phrase being positioned just prior to a verb which is moved to the sentence initial position. Judging from the large, 10.00 percent, scrambling effects on error rates (only 1.79 percent for canonical sentences with 11.79 for scrambled sentences), this stimulus manipulation must create a syntactically unfamiliar condition. The third experiment of passive sentences again shows the size of 201 milliseconds. Even the fourth experiment of potential sentences in which case particle and grammatical information conflict (details already explained in Section 3.2) produces scrambling effects of 216 milliseconds. Furthermore, causative sentences also show 187 milliseconds with accusative verbs and similarly 185 milliseconds with dative verbs.

Similarly, scrambling effects were also found in languages with free word order such as Turkish and Korean. Turkish native speakers showed scrambling effects in the processing of active sentences of 155 milliseconds with animate objects and 148 milliseconds with inanimate objects (Tamaoka, Kuribayashi and Sakai, 2005). Unlike Muraoka, Tamaoka and Miyaoka (2005), this Turkish experiment cannot examine the interaction between word order and animacy, since word frequencies are not controlled over the factor of animacy. Yet, scrambling effects are apparent in either case of active sentences with animate and intimate objects. Likewise, translated sentences from Japanese to Korean (Experiments 1 and 2 of Tamaoka, et al., 2005 listed in Table 5) also showed scrambling effects of 328 milliseconds for active sentences with transitive verbs and 615 millisecond for those with ditransitive verbs (Tamaoka, Lim and Miyaoka, 2005). As such, the plausibility judgment tasks can serve as a kind of litmus test to detect scrambling effects not only with Japanese but also other languages such as Turkish and Korean.

In addition, the plausibility judgment task was conducted to the second (or foreign) language condition of native Chinese speakers learning Japanese to investigate how native Chinese speakers learning Japanese comprehend active and potential sentences with canonical and scrambled word order (Tamaoka, 2005b). Based on over 91.7% scores on a Japanese grammar test, 24 of 87 students were selected as participants for the experiments. Experiment 1 showed that active sentences with canonical order were processed 367 milliseconds more quickly and 9.5 percent more accurately than the same sentences with scrambled order. As with native Japanese speakers, Japanese learners must also establish the base structure for active sentences and process scrambled sentences using the gap-filling parsing. In Experiment 2, potential sentences whose case particles conflicted with the grammatical information of subject and object were conducted with the same Chinese participants. Unlike the active sentences, potential sentences with canonical order did not differ in reaction times from the same sentences with scrambled order (i.e., there were no scrambling effects). This result implies that Japanese learners have not constructed the base structure of potential sentences and therefore cannot apply the gap-filling parsing to potential sentences with scrambled order. Once again, Japanese learners with a good knowledge of Japanese grammar also indicated scrambling effects.

All these studies, including those conducted with speakers and learners of two other languages exhibiting similar characteristics of free word order, clearly show scrambling effects strong enough to reach significance. Additionally, Japanese learners also display scrambling effects of active sentences with transitive verbs. Even though MN claim that these studies only provide the end result of whole sentence processing, the plausibility task at least has successfully measured scrambling effects. These studies showing consistent scrambling effects implies the existence of the base structure, which acts as a basis for the gap-filling parsing. Scrambling effects shown by a series of studies as presented in Table 5 can be hardly ignored by the virtue of the task nature itself.

4.2 Merits and demerits of self-paced reading

There are at least three potential reasons the self-paced reading approach is not sufficiently sensitive to detect scrambling effects in the processing of simple sentences.

First, we have actually conducted several experiments using self-paced reading to find scrambling effects of long and short sentences, but results of these experiments did not meet our expectations in terms of providing conclusive results. Tamaoka, Sakai, Kawahara and Miyaoka (2003) examined both factors of long/short phrase-length order and canonical/scrambled order on relatively long sentences. No significant main effects were found in either phrase-length or canonical/scrambled factor in reaction times, but scrambling effects were found in error rates. Based on these results, Tamaoka et al. (2003) suggested the Japanese preference of 'long-before-short' phrase-length (Yamashita and Chang, 2001) is an indication of preference rather than cognitive processing. In their concluding remarks, Tamaoka et al. (2003) add, in evaluating the self-paced reading method, the comment that participants seem to establish a rhythm when reading a sentence. This tendency becomes extreme in a simple active sentence with transitive verb. Since this type of sentence is read by pressing the space key three times, participants are likely to repeat a three-beat rhythm for NP-*ga*, NP-*o* and V. Likewise, a four-beat rhythm is used for active sentences with ditransitive verbs as NP-*ga*, NP-*ni*, NP-*o* and V. Since target stimulus sentences used for the experiments listed in Table 5 were constructed by a maximum of four phrases, participants could complete reading a sentence by pressing the space key four times, at most. With this repetitious behavior, reaction times were flat in each phrase.

Second, self-paced reading usually requires performing an additional task parallel to reading a sentence. This task often puts heavy memory loading on participants. In experimental psychology, this can be regarded as a dual task condition. One example is to ask participants (i.e., Japanese students in grades 4–6) to verbalize one to five while reading a sentence for the plausibility judgment task (Leong and Tamaoka, 1995; Tamaoka, Leong and Hatta, 1992). This method is the dual task condition by 'phonological suppression' or 'vocal interference'. Since it was hypothesized that students at an elementary school processed sentences with a heavy dependency on phonology, phonological suppression caused interference when reading a sentence. In fact, phonological suppression was very strong for students prolonging reading times of sentences in comparison with no vocal interference. For an extreme example, Muraoka and Sakamoto (2005) reported reaction times of sentences constructed of six phrases; where pressing the key six times was required to complete reading these sentences. Surprisingly, average reading times were 1,100–1,200 milliseconds for the first phrase, 1,200–1,400 for the second phrase, and 1,100 to 1,600 for the third phrase. A question about the content was asked after reading a sentence. In contrast, in our study, four phrases were read 1,414 milliseconds for canonical sentences with *pass*-type verbs, 1,512 milliseconds for their scrambled condition, and 1,570 for canonical sentences with *show*-type verbs and 1,679 milliseconds for their scrambled condition. There is an extreme difference in reading times between the two experiments. In addition, event-related brain potential (ERP) studies (e.g., Friederici, 2002; Friederici and Kotz, 2003; Kuperberg, Sitnikova, Caplan, and Holcomb, 2003; Schlesewsky and Bornkessel, 2004) cite the 400 milliseconds of negative electrical activity (N400) as an indication of semantic and pragmatic violations and the 600 milliseconds of positive electrical activity (P600) as one of syntactic reanalysis and repair for the phrasal processing. Actual reading times of phrases should be much smaller than those in previous self-paced reading studies. Even including a key-pressing time, the average of reading times for each phrase in Muraoka and Sakamoto (2005) seems very long. We assume that the task of Muraoka and Sakamoto (2005) must have required a heavy memory load to remember the content, which then acted like an experiment with a dual task condition. Since working memory plays an importance role for sentence processing (e.g., Just and Carpenter, 1992; King and Just, 1991;

MacDonald, Just and Carpenter, 1992), the dual task condition requiring heavy working memory will result in an extremely long reaction time to read each phrase, mostly spent for remembering the context.

A third potential weakness of self-paced reading methods is that it is not totally confirmed whether participants actually read the target phrase between key pressings. When a difficult memory task is required with reading a sentence, participants may have to make sure that they remember all items in a sentence. As a result, participants probably stop during sentence reading, and possibly stay longer in unrelated phrases. In fact, longer reading times are likely to be observed just after the target phrase expecting an extra cognitive (or working memory) load for parsing. However, this tendency could be minimized by averaging reaction times over an enough number of participants and target stimuli, and experimenters anticipate this trend in reading times when they interpret the results.

In voicing these three concerns, it is by no means our intention to suggest that self-paced reading is entirely unsuitable for studies of sentence processing. Indeed, this methodology could be selectively used depending on the nature of target stimulus sentences. Examples of areas which have enjoyed some success using self-paced reading include studies on ambiguity resolution in Japanese garden-path sentences (Inoue, 1998; Inoue and Den, 1999) and on locally ambiguous homonyms between nouns and verbs (Tokimoto, 2005). In these studies, readers required longer reading times for resolving ambiguity; self-paced reading would be an excellent means to measure how readers cope with ambiguity. Likewise, when target stimulus sentences are longer and more complex (e.g., Muraoka and Sakamoto, 2005), self-paced reading will be useful, especially to obtain reading times of each phrase. This selective use could be applied to our 2004 study of the plausibility judgment task, reflecting an approach that will be useful for measuring the processing of simple sentences. Again, however, information regarding reaction times is limited to the whole sentence reading. Respective implementations of self-paced reading and plausibility judgment seem to contain case-specific advantages and disadvantages.

5. Ending Remarks

In their commentary, MN draw attention to three important aspects for discussion.

First, assuming that the decision to insert the gap is made before the verb appears in processing ditransitive sentences, MN stress the putative effect of reanalysis when encountering the *pass*-type verb. We have argued that the insertion of a gap is likely to be delayed until the verb is processed in the case of ditransitive sentences used in our experiment. A mid-sentence reanalysis either does not occur, or, if it does occur, its effect is so small that our overall conclusions are not affected.

Second, MN argue that the scrambling effects observed in our experiment and other related studies using the plausibility judgment task could be caused, if only partially, by word and syntactic frequencies. While acknowledging the well-intended aims of MN's statistical tests and calculations, we must disregard them as being both inaccurate as well as inappropriate to furthering the results of the present experiment. According to our calculations, MN's corpus study of the fine-grained ditransitive sentences subcategorized in our experiment suggests that the *ni-o* and *o-ni* word order appear at the chance level regardless of the *show*-type or *pass*-type verbs (see Table 2). Although this procedure is commonly practiced to control stimulus items for experiments, MN simply deny this corpus data. However, when MN identify an interesting general contrast regarding canonical orders for passive sentences as *ga-ni*, and potential sentences as *ni-ga* shown by Tamaoka et al. (2005), they simply consider this appealing contrast as 'premature' (p. 119), claiming that this overall trend is too broad to argue. MN imply the need to distinguish between potential and passive sentences for counts of syntactic frequency. Their corpus study of transitive sentences shows an overall general trend indicating that 'a fronted accusative NP is highly unlikely to be scrambled prior to an overt subject' (p. 123), proposing that scrambled word order is rare. As MN critically comment a contrasting result of the third and fourth experiments of Tamaoka et al. (2005), we can now in turn state that this result seems to be too general to pertain to the stimulus sentences of our experiment. MN's comments are too narrow in scope when the corpus data tends not to support their argument, while their argument regarding the contrast over canonical order is not fully realized. Their approach, having no clear direction, save for the claim that frequency should be considered, seems to be opportunistic.

Thirdly, from a methodological perspective, MN seem not to place much value on the scrambling effects as shown by the plausibility judgment task. As listed in Table 5, experiments conducted on various types of sentences have indicated consistent results of scrambling effects. Nevertheless, MN repeatedly state that plausibility judgment provides only reading times for a whole sentence, which diverges from MN's self-stated position that 'the self-paced reading have only reached statistical significance when complex structures were used' (p. 118). Ultimately, self-paced reading would not be sufficiently sensitive to obtain differences between canonical and scrambled order in *simple sentences*. We put forward three possible problems regarding the self-paced reading method: (1) participants unthinkingly establish a rhythm pressing a key while reading a sentence phrase-by-phrase; (2) participants necessarily perform a heavily loaded memory task while reading a sentence such as with the dual task condition, and (3) there is no guarantee that participants will actually read target phrases between the pressing of keys. We propose that a method can be selectively employed depending on the nature of target stimulus sentences. In cases where conditions require longer reading times, such as those involving the resolution of ambiguity, self-paced reading would be an excellent means to determine how readers cope with ambiguity and particularly to obtain reading times of each phrase. On the other hand, the plausibility judgment can be used to measure the processing of simple sentences with limited information regarding reaction times for whole sentences. We therefore feel that MN would have done well to have been more flexible in their selection of experimental methodology. Their commentary on our experiment nevertheless constitutes a worthwhile contribution to the professional dialog on sentence processing.

References

Amano. N. and Kondo, K. (1999). *Nihongo-no goi tokusei—Dai-1-6 kan [Lexical properties of Japanese—Volume 1–6].* Tokyo: Sanseido.

——————— (2000). *Nihongo-no goi tokusei—Dai-7-kan [Lexical properties of Japanese—Volume 7].* Tokyo: Sanseido.

——————— (2003a). *NTT Database series: Nihongo-no goitokusei—Dai-1-ki CD-ROM-ban [Lexical properties of Japanese—The first period of CD-ROM verison].* Tokyo: Sanseido.

——————— (2003b). *NTT Database series: Nihongo-no goitokusei—Dai-2-ki CD-ROM-ban [Lexical properties of Japanese—The second period of CD-ROM verison].* Tokyo: Sanseido.

Becker, C. A. and Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 389–401.

Besner, D. and McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance, XII* (pp. 201–219). London: Lawrence Erlbaum Associates.

Carroll. J. B. and White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, 25, 85–95.

Caza, N. and Moscovitch, M. (2005). Effects of cumulative frequency, but not of frequency trajectory, in lexical decision times of older adults and patients with Alzheimer's disease. *Journal of Memory and Language*, 53, 456–471.

Chujyo. K. (1983). Nihongo tanbun-no rikai katei—Bunrikai sutoratejii no sougo kankei [The interrelationships among strategies for sentence comprehension].

Japanese Journal of Psychology, 54, 250–256.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. (1996) *Rethinking innateness*. Cambridge, MA: MIT Press.

Forster, K. I. and Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12, 627–635.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.

———— and Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage*, 20, S8–S17.

Fukumitsu, Y., Kim, J. and Koizumi, M. (2004) Me-de miru 'yokaku mokutekigo kakimazebun' no nounai shori [An fMRI study of the scrambling of dative object]. *Proceedings of the 128th Conference by the Linguistic Society of Japan*, 287–292.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., and Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency based component. *Memory and Cognition*, 15, 24–28.

Hagiwara, H., Soshi, T., Ishihara, M. and Imanaka, K. (2005). A topographical study on the ERP correlates of scrambled word order in Japanese complex sentences. Manuscript of Tokyo Metropolitan University.

Hino, Y. and Lupker, S. L. (1998). Effects of word frequency for Japanese kana and kanji words naming and lexical decision: Can the dual-route model save the lexical-selection account? *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1431–1453.

Hoji, H. (1985). *Logical form constraints and configurational structures in Japanese*. Unpublished doctoral dissertation, University of Washington.

Inoue, M. (1998). Gaaden-pasu bun no yomi to bun no rikai [Reading of garden-path sentences and sentence understanding]. In N. Osaka (Ed.), *Yomi: Noo to kokoro no joohoo shori [Reading: Information processing of brain and mind]* (pp. 721–789), Tokyo: Asakura Shoten.

———— and Den, Y. (1999). Influence of verb-predictability on ambiguity resolution in Japanese. *Proceedings of the 2nd International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society*

Joint Conference, 499–502.

Jescheniak, J. D. and Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 824–843.

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

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

Koizumi, M. and Tamaoka, K. (2004). Cognitive processing of Japanese sentences with ditransitive verbs. *Gengo Kenkyu [Journal of the Linguistic Society of Japan]*, 125, 173–190.

———, Tamaoka, K. and Miyaoka, Y. (2004). Bun-dokkai jikan ni ataeru toogo koozoo to kakujyoshi hairetsu jyunjyo no eikyoo —'*o*-kaku mokutekigo dooshi shiekibun' to '*ni*-kaku mokutekigo dooshi shiekibun' tono hikaku [Effects of syntactic complexity and case-particle order on sentence reading times—A comparison of 'passive sentences with *o*-case-marked objects' and 'passive sentences with *ni*-case-marked objects']. *Proceedings of the 129th Conference by the Linguistic Society of Japan*, 277–282.

———, Kim, J., Kimura, N., Watanabe, J., Yokoyama, S., Ikuta, N., Uchida, S., Akitsuki, Y., Iwata, K., Jeong, H., Miura, N., Sassa, Y., Yusa, N., Sato, S., Horie, K. and Kawashima, R. (2005). Cognitive brain imaging and the argument structure of ditransitive constructions in Japanese. Poster presented at the Second International Workshop on Evolutionary Cognitive Sciences "In Pursuit of Language-Brain Interactions: Language Acquisition, Sentence Processing, and Neurolinguistics", June 2005, University of Tokyo.

Koso, A. and Hagiwara, H. (2005). Exploring the type of filler integration: Behavioral and ERP evidences for a non-canonical word order of Japanese ditransitive sentences. Paper presented at the Sixth Tokyo Conference on Psycholinguistics, Keio University, March 2005.

Kuperberg, G. R., Sitnikova, T., Caplan, D. and Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17, 117–129.

Kurohashi, S. and Nagano, M. (1997). Kyoto University text corpus project. *Proceedings of the Third Annual Meeting of the Association for Natural Language Processing*, 115–118.

Leong, C. K. and Tamaoka, K. (1995). Use of phonological information in processing kanji and katakana by skilled and less skilled Japanese readers. *Reading and Writing*, 7, 377–393.

MacDonald, M. C., Just, M. A. and Carpenter, P. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, 24, 56–98.

Makioka, S. (1999). A view from a connectionist approach: Serious questions about the innateness of language. *English Linguistics*, 16, 541–562.

Marantz, A. (2005). Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review*, 22, 429–445.

Matsuoka, M. (2003) Two types of ditransitive constructions in Japanese. *Journal of East Asian Linguistics*, 12, 171–203.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance, XII* (pp. 3–36). London: Lawrence Erlbaum Associates.

——————— and Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–16.

——————— and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375–407.

Miyagawa, S. (1997) Against optional scrambling. *Linguistic Inquiry*, 28, 1–25.

Miyamoto, E. T. and Nakamura, M. (2005). Unscrambling some misconceptions: A comment on Koizumi and Tamaoka (2004). *Gengo Kenkyu [Journal of the Linguistic Society of Japan]*, 128, 113–129.

——————— and Takahashi, S. (2002). Antecedent reactivation in the processing of scrambling in Japanese. In T. Ionin, H. Ko, and A. Nevins (Eds.), *MIT Working Papers in Linguistics*, Vol. 43 (pp. 123–138). Cambridge, MA: MITWPL.

Morrison, C. M. and Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 116–133.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165–178.

Muraoka, S., Tamaoka, K. and Miyaoka, Y. (2005). Bun-rikai ni okeru gojyun to yuusei-sei no kankei [The relation between word order and animacy in sentence understanding]. *Proceedings of the 69th Conference of Japanese Psychological Association*, 639.

————, Tamaoka, K. and Miyaoka, Y. (2004). Kakimaze-bun no rikai ni okeru kaku-jyoshi no eikyoo [The effects of case markers on processing scrambled sentences]. *Technical Report of IEICE (Institute of Electronics, Information and Communication Engineers) 'Thought and Language'*, 37–42.

———— and Sakamoto, T. (2005). Mokuteki-go meishi-ku no kaiseki ni okeru kakujyoshi-no eikyo [Effects of case-markers on the parsing of object noun phrases]. *Proceedings of the 131st Conference by the Linguistic Society of Japan*, 204–209.

Nakayama, M. (1995). Scrambling and probe recognition. In R. Mazuka and N. Nagai (Eds.), *Japanese sentence processing* (pp. 257–273), Hillsdale, NJ: Lawrence Erlbaum Associates.

Paap, K. R., McDonald, J. E. and Schvaneveldt, R.W., and Noel, R.W. (1987). Frequency and pronounceability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and performance, XII* (pp. 221-243). London: Lawrence Erlbaum Associates.

Pritchett, B. L. and Whitman, J. B. (1995). Syntactic representation and interpretive preference. In Reiko Mazuka and Noriko Nagai (Eds.) *Japanese sentence processing* (pp. 65–76). Hillsdale, New Jersey: Laurence Erlbaum.

Phillps, C., Kazanina, N., and Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22, 407–428.

Schlesewsky, M. and Bornkessel, I. (2004). On incremental interpretation: Degrees of meaning accessed during sentence comprehension. *Lingua*, 114, 1213–1234.

Seidenberg, M. S. and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7, 263–272.

——— (1991). *Reading and the mental lexicon.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tamaoka, K. (2005a). Meimei kadai ni oite kanji 1-ji no shoji to on'in no tan'i wa itti suruka [Is an orthographic unit of a single Japanese kanji equivalent to a kanji phonological unit in the naming task?]. *Cognitive Studies* (Bulletin of the Japanese Cognitive Science Society), 12(2), 47–73.

——— (2005b). Chuugoku-go o bogo tosuru nihon-go gakushuu-sha ni yoru seijyun to kakimaze gojyun no noodoobun to kanoobun no rikai [Comprehension of Japanese active and potential sentences with canonical and scrambled word orders by native Chinese speakers learning the Japanese language]. *Journal of Japanese Grammar*, 5(2), 92–109.

——— and Takahashi, N. (1999). Kanji niji jukugo no shoji kodo ni okeru goi shiyo hindo oyobi shoji-teki fukuzatu-sei no eikyo [The effects of word frequency and orthographic complexity on the writing process of Japanese two-morpheme compound words]. *Japanese Journal of Psychology*, 70, 45–50.

——— and Tanaka, J. (2005). Meishi-ku to noodoo-bun ni okeru korokeeshon hindo no eikyoo [Effects of collocation frequencies on the processing of noun phrases and active sentences]. *Technical Report of IEICE (Institute of Electronics, Information and Communication Engineers) 'Thought and Language'*, 23–27.

———, Kuribayashi, Y. and Sakai, H. (2005). Psycholinguistic investigation of subject incorporation in the processing of Turkish active sentences with transitive verbs. *Proceedings of the 130th Conference by the Linguistic Society of Japan*, 116–121.

———, Leong, C. K. and Hatta, T. (1992). Effects of vocal interference on identifying kanji, hiragana and katakana words by skilled and less skilled Japanese readers in Grades 4–6. *Psychologia*, 35, 33–41.

———, Lim, H. and Miyaoka, Y. (2005). Nihon-go to Kankoku-go no bunshori ni okieru sukuranburu kooka [Scrambling effects on the processing of

Japanese and Korean sentences]. *The proceedings of the sprint international academic conference by the Japanese Cultural Association of Korea*, 90–96.

———, Miyaoka, Y., Ito, T. and Sakai, H. (2004). The canonical position of instructional and locative adverbs in the cognitive processing of Japanese sentences and noun phrases. *Proceedings of the 128th Conference by the Linguistic Society of Japan*, 275–280.

———, Sakai, H., Kawahara, J. and Miyaoka, Y. (2003). The effects of phrase-length order and scrambling in the processing of visually-presented Japanese sentences. *Journal of Psycholinguistic Research*, 32, 431–455.

———, Sakai, H., Kawahara, J., Miyaoka, Y., Lim, H. and Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions. *Journal of Psycholinguistic Research*, 34, 273–324.

Tokimoto, S. (2005). Disambiguation of homonyms in real-time Japanese sentence processing: Case-markings and thematic constraint. *Language and Speech*, 48, 65–90.

Ueno, M. and Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, 86, 243–271.

Yamashita, H. (1997). The effects of word-order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research*, 26, 163–188.

——— and Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition*, 81, B45–B55.

日本文の処理におけるスクランブル効果の諸問題—Koizumi and Tamaoka(2004)に対するMiyamoto and Nakamura(2005)のコメントへの回答

玉　岡　賀津雄（広島大学）

小　泉　政　利（東北大学）

Koizumi and Tamaoka（2004，以下KT）の実験結果に対して，Miyamoto and Nakamura（2005，以下MN）が寄せたコメントについて，3つの視点から回答した．第1に，KTの実験条件ではMNが指摘しているような再解析が起こらないことを示唆する経験的証拠を提示し，少なくともKTの実験においては，統語構造の複雑さが文正誤判断課題の反応時間と関係するという仮定が成り立つことを再確認した。第2に，頻度について，語彙，統語，共起の3種類があることを説明した．語彙頻度についてはKT（2004）の実験では比較条件で一様であるため，影響がない．統語頻度については，KTが使用した実験刺激の頻度をMNが示している（MN, Table 2, p. 121）が，これが実験統制に用いられる通常の方法であるにもかかわらず，結果を不十分としているのは不適切な議論であることを指摘した．MNの文完成課題については，興味深い結果を得ているものの，オフラインの結果であるため，それを支持するためのオンラインの実験が必要であろう．共起頻度については，今後の研究を待つことになろう．第3に，文正誤判断課題と自己制御読みに関して，MN自身も指摘しているように，自己制御読みについては複雑な文でない限り有意なスクランブル効果が観察されていない．本稿では，近年行われた文正誤判断課題の実験がスクランブル効果を一貫して観察していることを示した．その上で，自己制御読み実験では，キー押しのための運動が文処理に影響すること，読みのために与えられる課題が「二重課題法」のように機能して，文処理以外に過度な記憶負荷をかけていることなどを指摘した．

（受理日　2006年1月31日　　最終原稿受理日　2006年2月10日）