Understanding Sentences in Japanese Bit by Bit

Edson T. Miyamoto

How do people understand the meaning conveyed by sentences? What happens as each word in the sentence is read (or heard)? How much of the meaning intended can be understood before the end of the sentence is reached? These are some of the questions addressed in the field of sentence processing. The investigation of Japanese has been important in order to provide language-independent answers to such questions. This is a summary of some issues that have been investigated in the recent past (for a more detailed survey, see Miyamoto, in press b).

1. Working memory

Working memory (Baddeley, 1992, for a summary), which plays a role in much of the following discussion, is where information is held for short periods of time (up to 2 seconds) in order for various types of cognitive tasks to be performed. The severely constrained amount of information that working memory can hold is compensated by chunking (Miller, 1963), an operation to associate items together so that they occupy fewer slots in working memory. Consider how much harder it is to recall the sequence of ten digits 6-5-1-9-4-5-2-0-1-0 than the three numbers 65-1945-2010, or that it will be 65 years in 2010 since the end of World War II. A rough parallel can be drawn with sentence processing as words are associated (e.g., a subject with a verb) creating links with long term memory (e.g., that a war finished in 1945) in order to generate a more abstract representation that occupies fewer slots so that more processing capacity is available for upcoming words in the sentence.

Long term memory has been far less explored within the sentence processing literature, but its importance is underscored by results suggesting that knowledge of kanjis correlates with general reading abilities better than some working memory span measures (Jincho, Mazuka & Namiki, 2004).

2. Universal processing algorithm

One basic question is how similar the processing of sentences is across languages. Assume that it is possible to separate the knowledge base necessary to process language (e.g., grammar of the language one speaks, general knowledge about the world, social conventions), and the algorithm that uses this knowledge in order to process sentences. Each language (or each dialect) requires
a different knowledge base. Similarly, a different algorithm may be required for each language; another possibility is that a single algorithm sub-serves the processing of all languages (Fodor, 1998, for a discussion). There is no empirical way of showing that the processing algorithm is the same for all languages, thus this is taken to be the null hypothesis (i.e., true until proven wrong).

3. Incremental processing

One central topic within the universality of the processing algorithm debate is whether head-final languages like Japanese are processed incrementally, as is the case of head-initial languages, or whether comprehension lags behind the actual reading of the words in the sentence (A. Inoue & Fodor, 1995, for a discussion). The problem is clearest at the clausal level. In (1), each individual noun phrase (NP) can be created and interpreted immediately (i.e., there is ‘a young student’, ‘a dog’, and ‘an email’).

(1) Wakai gakusei-ga inu-ni imeeru-o
young student-nom dog-dat email-acc

The question is to what extent the NPs are interpreted together as part of a single representation before the verb is seen. For example, do readers assume that the dog belongs to the student? Do they realize that a dog is an unlikely recipient for an email? Is it possible to provide experimental evidence for such processes based on reactions collected before the verb is seen, and which is not explained by word-level correlations (e.g., that ‘email-acc’ cooccurs infrequently with ‘dog-dat’)?

Incremental models assume that information available (e.g. from case markers) allows readers to understand that the student did something to the email for (the benefit of) the dog, although the exact action is not known. But that is enough to bind the NPs together in one single chunk by predicting an underspecified predicate (i.e., partial information about the predicate is used to project the necessary structure, e.g., a verbal phrase node, VP). An increasing body of evidence supports this view, but our understanding of how rich a representation is created in these situations has evolved stepwise.

Intuitive judgements (A. Inoue, 1991), questionnaire data (Uehara & Bradley, 2002), reading times (Miyamoto, 2002) and event related potentials, ERPs (Garnsey et al., 2001), indicate that clause boundaries are determined based on case markers (also, Muraoka & Sakamoto, 2004, on the influence of animacy). According to reading times (Mazuka, Itoh & Kondo, 2002; Miyamoto & Takahashi, 2002a) and ERPs (Ueno & Kluender, 2003), non-canonical orders elicit deviant responses before the predicate is read. Reading time data indicate that preverbal pronoun-antecedent assignment obeys syntactic constraints (Aoshima, Phillips & Weinberg, 2003). These results indicate that a detailed syntactic structure is created before the verb is read. The next step is to determine how much semantic interpretation can take place without a predicate. For example, a hamburger is more likely to be expected after hearing (2a) than (2b) (Kamide, Altmann & Haywood, 2003).

(2) a. Ueitoresu-ga kyaku-ni...
waitress-nom customer-dat
b. Ueitoresu-ga kyaku-o...
waitress-nom customer-acc

The result is compatible with the view that the NPs are being interpreted together and used to predict what is coming next. But a more shallow process may also explain the result (Kamide, Altmann & Haywood, 2003, for a discussion) given that a dative NP is often followed by another NP, while an accusative NP is seldom followed by another NP (Miyamoto & Nakamura, 2005).

A related issue involves the kinds of properties of the verb that can be anticipated based on the case-marked NPs alone. Some work indicates that argument structures are predicted, but it remains to be seen whether other types of information are equally predicted (e.g., tense, aspect).

The evidence for the prediction of argument
The processing of filler-gap dependencies has been shown to underlie a variety of phenomena across a number of different languages. The typical example is fronted wh-phrases in English as in the following sentence (adapted from Crain & Fodor, 1985).

(3) Filled gap effect

Who, did the children force us to sing the songs for gap, last week?

The filler who in (3) triggers the search for a gap (the thematic position of the wh-phrase) and a gap is posited at the earliest possible position, that is, at ‘∧’ as the direct object of force. A slowdown occurs when the word us is read as it indicates that the position of the gap is incorrect (Crain & Fodor, 1985). This type of filled-gap effect is evidence for strict incremental processing as the required gap is inserted without lookahead to check the next word in the sentence.

This effect has been discussed extensively and has been characterized in a number of ways (the active filler hypothesis, Frazier & Clifton, 1989; the minimal chain principle, de Vincenzi, 1991; locality, Gibson, 1998). For example, it has been proposed that long dependencies consume cognitive resources such as working memory necessary to keep the filler and its gap requirement active until the gap position is found (e.g., King & Just, 1991). In some models, this is a special case of a general phenomenon as readers are assumed to keep track of the syntactic heads necessary to complete the sentence. The number of heads predicted and the amount of material processed since each head was first predicted determine the memory cost at each point in the sentence. Thus, the more predictions are made and the longer they are kept active in working memory, the larger the processing cost (Gibson, 1998).

4.1 Scrambled wh-phrases

The filled-gap effect has also been observed in Japanese for dative-marked wh-phrases scrambled across two clauses to the front of the sentence (Aoshima, Phillips & Weinberg, 2004). Although the grammatical nature of the fronting of the wh-phrase is arguably different in the
two languages (fronting to the beginning of the clause is obligatory in English, whereas scrambling in Japanese is optional and can place the wh-phrase virtually anywhere), the result indicates that processing is similar as readers attempt to find the base position of the wh-phrase as soon as possible. This is expected if the processing mechanism is the same for all languages. Although grammatical knowledge is important in order to determine where the gap can occur, the exact grammatical nature of the fronting operation is not important for processing, what is crucial is how to interpret the wh-phrase and, in both languages, this is done by determining the gap position.

4.2 Wh-phrases and question particles

Wh-phrases in Japanese have to be associated with a question particle (see Nakagome et al., 2001, for a P600 response to the declarative particle yo replacing a question particle such as ka after a wh-phrase). In (4), ‘what kind of computer-acc’ is in-situ and is licensed by the question particle no.

(4) Senmu-ga donna-pasokon-o
director-nom wh-kind-computer-acc
tukatteiru-to kkarichoo-ga itta-no?
using-is-that supervisor-nom said-qp
‘What kind of computer did the supervisor say the director is using?’

Self-paced reading studies (in which participants press a button to reveal sentence segments one at a time; Just, Carpenter & Woolley, 1982, for a comparison of different types of self-paced reading with eye-tracking) indicate that wh-phrases in Japanese trigger the search for a question particle in a manner similar to gap searching for wh-phrases in English. Readers slow down at the embedded verb with the declarative complementizer to ‘that’ in (4) compared to a similar sentence with the question particle ka replacing to. In this typing mismatch effect, the expectation for a question particle at the earliest point allowed by the grammar is contradicted by a constituent with declarative typing (Miyamoto & Takahashi, 2002b, c). This effect has also been reported for scrambled dative wh-phrases (Aoshima, Phillips & Weinberg, 2004), negative polarity items (Yoshida, 2004) and exclamatives (Ono et al., 2004). Here as well, wh-phrases reveal how the same processing strategy can surface in different guises. To be interpreted, wh-phrases require their scope and thematic role to be determined. English indicates scope by pronouncing the wh-phrase at the relevant clausal level, and the thematic-role requirement is satisfied by searching for a gap. In Japanese, in contrast, wh-phrases are pronounced at the thematic position, and scope requires searching for a question particle. In short, the constituent sought (a gap or a question particle) is determined by the grammar of each language, but the search process is similar as expected if the same processing algorithm is used for all human languages.

5. Relative clauses

Relative clauses are interesting for a number of reasons. Some of the reasons are specific to the construction in Japanese, but several issues are related to the different forms that relative clauses assume across languages.

5.1 Clause boundaries

One crucial property of embedded clauses in Japanese, and of relative clauses in particular, is that it is often not clear that a new clause has started until material after the embedded verb is seen (A. Inoue, 1991, for a detailed discussion; also Venditti, in press, for prosodic cues that can indicate the beginning of the embedded clause). In general, there is a preference to assume that the fragment being read is part of a simple clause, even when one of the arguments is missing. For example, when shown an accusative animate NP (e.g., tosiyorino obaasan-o ‘elderly woman-acc’) and asked to complete it into a full sentence, native speakers completed it
as a simple matrix clause with the subject missing 74.7% of the time, 15.7% of the completions included an overt subject following the direct object, 7.6% were imperatives, clefts, and adverbial clauses, and the remaining 2% were relative clauses (Miyamoto & Nakamura, 2005, who also report similar numbers for dative NPs). In short, the missing subject prior to the accusative NP is rarely interpreted as the gap of a relative clause.

For (5), the single-clause interpretation ‘Yoko saw the child at the intersection’ is initially favored (Mazuka & Itoh, 1995).

(5) Yoko-ga kodomo-o koosaten-de
Yoko-nom child-acc intersection-loc
mikaketa ...

saw

When the verb is followed by a noun, it becomes clear that the fragment includes another clause, and only then is the structure for the relative clause built by inserting a gap and associating it with the modified noun (M. Inoue, 1990, for experimental evidence that a slowdown occurs at the noun; also Imui et al., 1998, for related fMRI data).

This model has been generalized to explain the processing of double-gap relative clauses, in which two gaps are associated with two modified nouns, by applying the same procedure recursively in order to create each gap-noun dependency in turn (Nakamura, 2003).

Moreover, the difficulty at the postverbal noun is modulated by the type of relative clause created as illustrated in the following two continuations for (5).

(6) a. Early opening
   ... onnanoko-ni koe-o-kaketa.
   girl-dat called

   ‘Yoko called the girl who saw the child at the intersection.’

b. Late opening
   ... takusii-ni noseta.
   taxi-dat put

   ‘Yoko made the child get on the taxi she saw at the intersection.’

In the early opening type in (6a) the relative clause starts at the direct object ‘child’, whereas in the late opening type in (6b) it starts after the direct object. Mazuka & Itoh (1995), based on intuitive judgments, observed that late opening is harder than early opening because of the number of arguments that have to be displaced in each case from the single clause initially built. In (6a), the subject ‘Yoko’ is displaced to the newly built matrix clause, whereas in (6b) both ‘Yoko’ and the object ‘child’ have to be displaced (Hirose & Inoue, 1998; Miyamoto, 2002, for experimental results; also, Hirose, 2003, on the labels late and early opening).

5.2 Implicit prosody

There is some debate as to when speakers produce prosodic contours to disambiguate an utterance, and to what extent listeners take advantage of such cues (e.g., Hirose, 2006; also Venditti, in press, for a summary). Moreover, based on the proposal that during silent reading implicit prosodic contours are imposed on the input and influence the way how ambiguous sentences are read (the implicit prosody hypothesis, Fodor, 2002), various phenomena have been re-evaluated in formal syntax as well as in psycholinguistics (e.g., Kitagawa & Fodor, 2003, on the assignment of scope for wh-phrases; Fodor, 2002, on length and modifier attachment).

A series of experiments suggest that during silent reading, putative prosodic contours affect the resolution of clause boundary ambiguities. Until ‘trusted’ in (7a, b), the single clause interpretation ‘Morishita (and Hosokawa) truly trusted the medicine’ is favored. When ‘friends’ is read, the beginning of the relative clause can be set at ‘medicine-acc’ (in the early opening interpretation) or after it (in the late-opening interpretation).

The nominative subjects in fragments (7a, b) differ in their lengths (as measured in minor phrases or accentual phrases), which affect prosodic contours when the sentences are read

5.2 Implicit prosody

There is some debate as to when speakers produce prosodic contours to disambiguate an utterance, and to what extent listeners take advantage of such cues (e.g., Hirose, 2006; also Venditti, in press, for a summary). Moreover, based on the proposal that during silent reading implicit prosodic contours are imposed on the input and influence the way how ambiguous sentences are read (the implicit prosody hypothesis, Fodor, 2002), various phenomena have been re-evaluated in formal syntax as well as in psycholinguistics (e.g., Kitagawa & Fodor, 2003, on the assignment of scope for wh-phrases; Fodor, 2002, on length and modifier attachment).

A series of experiments suggest that during silent reading, putative prosodic contours affect the resolution of clause boundary ambiguities. Until ‘trusted’ in (7a, b), the single clause interpretation ‘Morishita (and Hosokawa) truly trusted the medicine’ is favored. When ‘friends’ is read, the beginning of the relative clause can be set at ‘medicine-acc’ (in the early opening interpretation) or after it (in the late-opening interpretation).

The nominative subjects in fragments (7a, b) differ in their lengths (as measured in minor phrases or accentual phrases), which affect prosodic contours when the sentences are read
(7) a. Conjoined-names subject

Hosokawa-to Morisita-ga | sinyaku-o kokinokara sinyoosita yuujintati-ni . . .
Hosokawa-and Morisita-nom medicine-acc truly trusted friends-dat

b. Single-name subject

Morisita-ga sinyaku-o | kokinokara sinyoosita yuujintati-ni . . .
Morisita-nom medicine-acc truly trusted friends-dat

aloud. A major phrase boundary (also known as intermediate phrase) tends to be produced immediately after the conjoined names in (7a), as indicated by the vertical bar, whereas with the single-name subject in (7b) it is produced after ‘medicine-acc’ (Hirose, 2003, Experiment 2). Note that these two positions for the major phrase boundary correspond to the beginning of the embedded clause in the early-opening and the late-opening interpretations.

When asked to write sentences starting with the fragments in (7a, b), Japanese speakers produce completions with early-opening interpretations more often with conjoined subjects (Hirose, 2003, Experiment 1). This is compatible with the claim that speakers assign implicit prosodic contours while reading in silence and that the position of the major phrase boundary influences the insertion of the embedded clause boundary.

However, the prosodic manipulation (as is often the case) leads to an unintended side-effect, namely sentences that are prosodically longer also contain more characters. Thus, it is conceivable that non-phonological factors are responsible for the effect reported (but see Kobayashi, Miyamoto & Sato, 2006, for data from deaf readers of Japanese that support the phonological explanation).

5.3 Crosslinguistic differences

Apart from the issues specific to Japanese in the previous sections, relative clauses are interesting from a number of other reasons. They allow the investigation of how people deal with multiple events being described simultaneously and the burden this causes on cognitive resources such as working memory and attention. Moreover, the structure of relative clauses can vary within and across languages. In some languages, relative clauses are postnominal as they come after the noun they modify (e.g., English, French, German), whereas in Japanese (as well as Chinese, Korean), relative clauses are prenominal as they are placed before the modified noun.

Within the same language, relative clauses can also vary in terms of the position that is relativized (i.e., the position of the gap) and the grammatical role of the modified noun.

5.4 Gap position

The memory load imposed by long dependencies has been used to explain why relative clauses with the gap in subject position are easier to understand than those with the gap in object position in postnominal relative clauses (e.g., King & Just, 1991, on English).

(8) a. Subject-gap relative clause: the senator, that gap, accused the reporter

b. Object-gap relative clause: the senator, that the reporter accused gap,

In (8), the distance between the gap and the filler (i.e., the modified noun senator) is larger with the gap in object position (as indicated by the underlines), thus the filler has to be kept in working memory longer until the gap is found. In Japanese, in contrast, relative clauses are prenominal and gaps in object position are closer to the modified noun.
(9) a. Subject-gap relative clause:
\[
\text{gap}, \text{obasan-o mita} \quad \text{onnanoko,}
\]
\['\text{the girl who saw the woman}'\]

b. Object-gap relative clause:
\[
\text{obasan-ga gapi mita} \quad \text{onnanoko,}
\]
\['\text{the girl who the woman saw}'\]

While in postnominal relative clauses, the cost is usually associated with storage of the filler until the gap is found, in Japanese the cost may be related to retrieval given that it is necessary to recover the gap position when the postverbal noun is detected. Thus, distance may be important in terms of how far back it is necessary to return in order to find the gap. In this case, object gaps should be easier to understand than subject gaps, contrary to results using questionnaires (Sheldon, 1976), self-paced reading (Ishizuka, Nakatani & Gibson, 2003, Miyamoto & Nakamura, 2003; also Kwon, Polinsky & Klunder, 2004, for a similar preference for Korean) and ERPs (Ueno & Garnsey, 2005), according to which subject gaps are easier to understand.

Although the grammatical role of the modified noun has an effect, it is unlikely to be the decisive factor (the object gap relative clause is hardest when modifying a nominative-marked noun, compared to when it modifies topics and accusative-marked nouns; Miyamoto & Nakamura, 2003).

Interestingly, self-paced reading results for prenominal relative clauses in Chinese have been claimed to favor object gaps (Hsiao & Gibson, 2003; but see Hsu, 2003; Lin, Fong & Bever, 2005). Thus, Japanese may be exceptional and its preference for subject gaps may be caused by the syntactic properties of its relative clauses (Miyamoto & Nakamura, 2003), which may also explain the availability of constructions usually not seen in other languages (e.g., Yamashita, 1995, on gapless relative clauses; Yoshida & Sano, 2001, on head-internal relative clauses).

5.5 Relative-clause attachment

Another line of research started with the finding that when reading the servant of the actress who was on the balcony in their native language, English speakers prefer the interpretation in which the relative clause attaches to the local noun (actress; i.e., the actress was on the balcony), whereas in Spanish the far noun (servant) is preferred (Cuetos & Mitchell, 1988; see http://www.lingua.tsukuba.ac.jp/etm/rc/ for a list of results for various languages). This has been one of the strongest cases against the single processing mechanism assumption as it suggests that relative clause attachment is different across languages. Numerous proposals have been made in order to explain these preferences without parameterizing the processing mechanism for each language (see Fodor, 1998, for a summary).

In Japanese, questionnaire data have found a preference for the far noun ‘servant’ in the example below (Kamide & Mitchell, 1997, Experiment 1; also, Hirose, 2001; Sturt, Branigan & Matsumoto-Sturt, 1999).

(10) Barukonii-ni iru joyuu-no
\[
\text{balcony-on is actress-gen}
\]
meshitsukai
\[
\text{servant}
\]
‘the servant of the actress who is on the balcony’

Reading times have also been measured for unambiguous sentences (disambiguated through plausibility). The far attachment condition was read more slowly initially, but it was faster at sentence end (Kamide & Mitchell, 1997, Experiment 2). One explanation for this reversal in reading times is based on the segmentation used in the self-paced reading presentation, which may have artificially favored the local noun (Kamide et al., 1998). But even when the segmentation is modified, the overall results remain (Miyamoto, Nakamura & Takahashi, 2004).

A second possibility is that readers are nondeterministically considering different interpretations at different points in the sentence.
(Kamide & Mitchell, 1997), but recent evidence suggests otherwise as the same interpretation is maintained throughout (Aoyama & Inoue, 2005). A third explanation is that the relative clauses are initially interpreted as matrix clauses and the reanalysis process to turn them into relative clauses favors the local noun as the attachment site (Miyamoto, Nakamura & Takahashi, 2004).

Relative clauses with three candidate nouns, instead of two, have also been investigated (Miyamoto et al., 1999).

In sum, many aspects of relative clauses (gap position, grammatical role and position of the modified noun) have been investigated. However, unifying accounts for each phenomenon, let alone for their possible interactions, still remain elusive.

6. Word order

Japanese allows the same propositional content (i.e., *who did what to whom*) to be expressed with the exact same words in different orders. There is an extensive literature investigating the processing of scrambled as opposed to *canonical* orders using a variety of methodologies (e.g., reading times, word recognition latencies, plausibility judgment latencies, brain responses) and the issues are too involved to be included here (see Miyamoto, in press a, for a summary), but the overall conclusion is that scrambled orders are more difficult to understand. The following preliminary results relate lexical access within sentence processing, to the earlier discussion on the kinds of information used to predict an upcoming verb.

6.1 Method

In a *stop-making sense* experiment, using a non-cumulative self-paced moving-window presentation, participants pressed a button to reveal each segment of the sentence, and pressed another button as soon as they thought the sentence did not make sense. Plausible items as in (11) were used (the segmentation shown is the same as the one used in the experiment; Miyamoto et al., 2005).

6.1 a. Canonical order

\[
\begin{align*}
\text{Yonakano} & \quad \text{denwade} & \quad \text{byoukino} \\
\text{middle-night} & \quad \text{phone} & \quad \text{sick} \\
\text{shufu-ga} & \quad \text{kinjono} & \quad \text{isha-o} \\
\text{housewife-nom} & \quad \text{neighbor} & \quad \text{doctor-acc} \\
\text{yonda rasii.} & \\
\text{called seems}
\end{align*}
\]

‘In a phone call in the middle of the night, the sick housewife called the local doctor.’

The only difference is whether the subject comes before the direct object in the more common order in (11a), or after the direct object in the scrambled order in (11b). Apart from word order, the $2 \times 2$ design manipulated the frequency of the matrix verb. There were 12 sentences that used high-frequency verbs (with more than 1.2 million hits on the internet search engine Google, $M=3.185$ million hits; e.g., *yobu* ‘call’) and 12 sentences with low frequency verbs (fewer than 400,000 hits, $M=129,998$; e.g., *karakau* ‘make fun’).

For the 80 items (24 test items, 16 plausible fillers, and 40 fillers that became implausible at various points), 32 native Japanese speakers judged the implausible items to be implausible at some point (97.4%) more often than for the plausible items (7.2%; $P < 0.001$). For the test items, the low-frequency verb items were judged to be implausible (10.7%) more often that the ones with high-frequency verbs.
(7.6%) marginally in the participant analysis ($F_1(1, 31)=3.72, P<0.065; F_2<1$). There was no word order or interaction effects ($Ps>0.14$).

6.2 Reading time results

If at some point a sentence was judged to be implausible, presentation was halted and the remaining regions not shown, hence reading time analyses only included regions judged to be plausible.

Residual reading times compensating for differences in word length (Ferreira & Clifton, 1986) were analyzed with the following results. The NP ‘doctor-acc’ in the scrambled (11b) was read more slowly than ‘housewife-nom’ in the canonical (11a) in the participant analysis ($P_1<0.001; P_2<0.08$) perhaps because readers expected the subject to come at that point or because more time is needed in order to create a null subject (and determine its antecedent) in the scrambled order.

The following adjective in the scrambled order was read more slowly than the one in the canonical order ($Ps<0.001$) as readers usually expect the verb to come immediately after the direct object (Miyamoto & Takahashi, 2002a; also Ueno & Kluender, 2003, for a P600 at a demonstrative after a scrambled direct object). There were no differences between ‘doctor-acc’ in (11a) and ‘housewife-nom’ in (11b) ($Fs<1$; Mazuka, Itoh & Kondo, 2002, for longer reading times to the scrambled order at that point according to eye-tracking but not to self-paced reading data).

At the matrix verb (‘called’ in (11a, b)), both main effects were reliable (see Figure 1). The scrambled conditions were slower than the canonical conditions ($Ps<0.005$). And the low-frequency verbs were read more slowly than the high-frequency verbs ($Ps<0.05$), as less frequent words take longer to read (Ashby, Rayner & Clifton, 2005, for recent results and references).

The result of interest was an interaction between verb frequency and word order at the matrix verb ($F_1(1, 31)=4.7, P<0.05; F_2(1, 22)=4.53,$ $P<0.05$) as the advantage of the canonical over the scrambled order was reliable for the verbs with low frequency ($F_1(1, 31)=10.75, P<0.005; F_2(1, 11)=13.48, P<0.005$), but not for the ones with high frequency ($F_1(1, 31)=2.25, P=0.143; F_2(1, 11)=1.36, P=0.27$). The next section investigates this effect in more detail.

6.2.1 Correlation analyses

Linear correlation analyses were conducted between reading times and frequencies. In all analyses, the mean raw reading times to the main verb in each item (the penultimate region, ‘called’ in (11)) were used (RTs for short). Number of hits on Google was used to estimate the frequency of the verb alone in (I) and of the verb with each word order in (II) and (III). Corpus counts results are reported in (IV).

(I) When the RTs were paired with the number of hits for each verb on Google, reliable correlations were found using raw frequencies (i.e., number of hits: $R=-0.35, P<0.05$) and log frequencies (logarithm of the number of hits: $-0.43, P<0.005$). Hence, the more frequent the verb, the faster it was read regardless of the word order that preceded it. This is the general effect that high-frequency words are read faster. However, there was no correlation between the frequency of each verb and the reading time difference (scrambled minus canonical) for that verb (raw: $R=-0.19, P=0.37$; log: $-0.24, P=0.26$).
Thus, the relative difficulty with the scrambled order does not seem to be related to the frequency of each verb. In other words, these correlation results are compatible with the main effect of frequency described in the previous section, but they do not explain the interaction between frequency and word order.

(II) Searches were entered on Google as ᵃ^t * ᵖ V for the canonical order, and ᵖ * ᵃ^t V for the scrambled order, where ᵃ^t was intended as the nominative marker (but several clausal conjunctions were returned), ᵖ for the accusative marker, V was the exact form of the verbs in the penultimate region of the test sentences, and * is the regular-expression wild card that can stand for any sequence of letters. In total, there were 3,758,027 hits (94.53%) for the canonical order and 217,590 (5.47%) for the scrambled order. There was a negative correlation between RTs and the number of hits for each verb in each word order (raw: $R = -0.29, P < 0.05$; log: $R = -0.42, P < 0.005$). Hence, even when the frequency takes into account the word order preceding the verb, the correlations are not better than the correlations conducted with the frequency of the verb alone. Because these two types of frequencies correlate with each other (raw: $R = 0.67, P < 0.0001$; logs: $R = 0.8, P < 0.0001$), their individual effects on RTs are unclear.

(III) The first 100 sentences returned for each Google search were manually checked and irrelevant cases eliminated. Of the 2,372 sentences checked for the canonical order, 1228 (51.77%) were found to be relevant. Out of 2,300 sentences for the scrambled order, 1050 (45.65%) were valid. The proportion of correct sentences in each search was multiplied by the corresponding total number of hits in order to get a second frequency estimate (the filtered number of hits), which correlated with the RTs at levels similar to the earlier analyses (raw: $R = -0.41, P < 0.005$; log: $R = -0.39, P < 0.01$).

(IV) A total of 9,604,304 sentences from the Mainichi Shinbun (1995) and the Nippon Keizai Shinbun (1990 to 1995) were automatically parsed (Kudo & Matsumoto, 2000) and 508,574 occurrences of the verbs used in the experiment (not necessarily with the same inflection) retrieved, of which 2,067 had accusative and nominative NPs in the same clause. These 2,067 sentences were manually checked and 286 were removed because they had been incorrectly parsed and were not relevant. For the remaining, there were 1662 (93.3%) in canonical order and 119 (6.7%) in scrambled order. Reliable correlations were found between RTs and the corpus frequency for each verb in each word order (raw: $R = -0.34, P < 0.02$; log: $R = -0.37, P < 0.01$).

6.3 Discussion

The results indicate that the reading times to the verb correlate with the frequency of each verb in isolation as well as with the frequency of each verb preceded by each word order. Because these two types of frequencies are highly correlated, it is not clear which one contributes the most to the reading time effect. The lack of correlation between verb frequency and reading time difference (scrambled minus canonical), suggests that lexical frequency is not the critical factor. But further evidence is needed to determine whether word order frequency with each verb is actually responsible for the reading time differences. If this is the case, there are two ways we could explain the effect. Records on word-order frequency kept for each verb may help predict the verb that is upcoming; thus, not only the types of the NPs, but also their order may be a cue (but see Yamashita, 1997, Experiment 2, for a different conclusion). A second possibility is that there is no predictive process going on, but this type of information may be used when the verb is read helping associate the NPs with the slots in the argument structure of the verb.

One may wonder whether this type of contextual information can be used so rapidly during processing. But, factors such as plausibility of prior contexts have been shown to have an early
effect in English (e.g., on the first-fixations of the target word, although without interacting with lexical frequency; Ashby, Rayner & Clifton, 2005). It has also been argued that in English argument-structure frequency is used in order to interpret an ambiguous NP following a verb (e.g., whether the NP in the fragment *found the answer* is interpreted as a direct object or as the subject of an embedded clause, depends on how often this verb is used with each type of complement; Garnsey et al., 1997). The effect in Japanese may be the counterpart of this effect. While in English, argument-structure frequency determines the interpretation of a postverbal NP; in Japanese, the order of case-marked NPs may help predict the argument structure of the upcoming verb.

Two points are worth stressing out. First, the SOV order was overwhelmingly more frequent than the OSV order both on the internet (94.53% to 5.47% in (II)) and in corpora (93.3% to 6.7% in (IV)) confirming previous reports (Miyamoto & Nakamura, 2005). Second, correlations between corpus frequency and hits on the search engine indicate that the latter provides reasonable approximations for the former (raw hits: $R=0.52$, $P<0.0005$; filtered hits: $R=0.63$, $P<0.0001$). Although the search engine numbers are noisy (roughly half of the returned sentences were irrelevant, see (III)) their correlations with the reading times are at levels similar to the ones between the corpus counts and the reading times. Sparseness is a major problem with the corpora as only 12 verbs had relevant sentences in both word orders, four verbs only had canonical sentences, and eight verbs had no relevant sentences, in spite of the fairly large number of sentences (over 9 million) checked. For the internet searches, in contrast, all verbs returned sentences in both orders with the exception of *torokashita* ‘melted’ which was missing sentences in the scrambled order. Thus, the sheer size of the internet seems to compensate for the noisy results returned by search engines (Keller & Lapata, 2003).

### 7. Conclusion

Results on the processing of sentences in Japanese were surveyed under the assumption that the same mechanism is used in the processing of all human languages. In particular, processing was assumed to be incremental as readers try to make sense of the input based on whatever information available at each point. Different types of constructions (e.g., with filler-gap dependencies) that are processed in a similar manner as well as possible counter-examples to the universality of the processing mechanism (e.g., cases involving relative clauses) were discussed. Preliminary results on the effects of word-order frequency on the processing of the upcoming verb were reported. Overall, the aim of the discussion was to highlight the importance of investigating typologically diverse languages in order to understand the cognitive mechanisms involved in language processing.

### References


(Received 30 March 2006)

(Accepted 30 April 2006)

Edson T. Miyamoto

Edson T. Miyamoto received his Ph.D. in 1999 (Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology). He was a postdoctoral fellow at the Graduate School of Information Sciences at the University of Tokyo supported by the MIT-Japan Program (1999–2001), assistant professor at the Graduate School of Information Sciences at the Nara Institute of Science and Technology (2001–2003), and is currently a lecturer at the University of Tsukuba (2003– at present). His research interests are primarily concerned with how people understand sentences.