Sentence-level Processing in Japanese : A Preliminary Report

著者名(英)	Miyamoto,Edson T./Takahashi,Shoichi
journal or publication title	言語科学研究 : 神田外語大学大学院紀要
volume	7
page range	89-102
year	2001-03
URL	http://id.nii.ac.jp/1092/00000362/

Sentence-level Processing in Japanese – A Preliminary Report

Edson T. Miyamoto
Shoichi Takahashi

The present paper discusses some issues in sentence-level processing in Japanese, their relation to the processing of English, and how they are being investigated at the Sentence Processing Laboratory at Kanda University of International Studies (SPL/KUIS). The phenomena discussed include the processing of case markers, wh-phrases, question particles and negative polarity items.

*** language processing * case markers * wh-phrases**
*** question particles * negative polarity items**

Researchers in sentence processing are concerned with how people read (or hear) words in a sentence and produce an appropriate mental representation. Emphasis is on the process of incorporating each word into a partial representation until the end of the sentence is reached and a complete representation is obtained. Assuming that the constraints proposed by syntacticians are correct, the question is how people compute mental representations that conform with such constraints. In particular, one should ask how knowledge of the grammar interacts with other sources of information (e.g. world knowledge) and with cognitive resources (e.g. working memory capacity). Inevitably, we also have to consider where cross-linguistic parameterization is necessary. Clearly, grammars have to be parameterized, but the more interesting question is whether the parsing algorithm, that presumably uses grammars as one of its knowledge sources, has to be different for each human language.

1. Some universal principles

Given a sentence, there are various ways in which a representation can be built for it.

For example, it is possible to write a computer algorithm which looks for verbs in the input sentence; then, if the algorithm finds a NP immediately after a verb, it assigns the NP as the direct object of the verb. Next, a NP preceding a verb is assumed to be its subject. Finally, extra-linguistic sources of information such as world knowledge rectify any implausible associations. For an English sentence such as "The fingerprints examined incriminated the defendant.", the algorithm would first find two verbs – **examined** and **incriminated**. It would then create a representation in which **the defendant** is the direct object of the second verb. Next, **the fingerprints** would be incorporated in the representation as the subject of **examined**. In the final stage, because **fingerprints** are an improbable agent, **examined** should be reanalysed from a main verb to a relative clause interpretation, with the resulting complex NP ("the fingerprints (that were) examined") being assigned as the subject of **incriminated**. Apart from its limited scope, the algorithm behaves in ways that people in general do not.

First, the algorithm violates *incremental processing*, according to which people incorporate words into the representation immediately, in the order that they read them (Marslen-Wilson, and Tyler, 1980; 1981). For example, the algorithm decided that **the defendant** is the direct object of **incriminated** before it associated **the fingerprints** with **examined**. In contrast, English speakers read sentences left to right and, thus, they first determine the relation between **the fingerprints** and **examined** before proceeding to the rest of the sentence.

Second, there is no evidence that English speakers associate **the fingerprints** as the subject of **examined** (Trueswell, Tanenhaus, and Garnsey, 1994); even though, being a preverbal NP, it belongs to the appropriate category and occurs in a compatible position. If some types of information (such as linear position and part of speech) were prioritized over other types of information (e.g. animacy of an entity), it would be conceivable that people initially associate **the fingerprints** as the subject of **examined**; but no experimental evidence is available supporting this hypothesis. Experimental work so far suggests the *immediate use of multiple sources of information*, in other words, all information available is used at once as each word is incorporated in the mental representation. There is no delay in the use of sources of information such as plausibility (Trueswell, Tanenhaus, and Garnsey, 1994), discourse complexity (Altmann, and Steedman, 1988), argument structure frequency (Garnsey et al., 1997).

Third, English speakers not only incorporate the word they just read into the mental representation, but they also make predictions as to what words will follow in the input string (*constituent prediction*). This type of expectation has been argued to play a crucial role in acquisition (see for example Elman, 1991, for its role in a neural network model in order to avoid poverty of stimulus) as well as processing (e.g. in the processing of fronted wh-phrases in English, Frazier, and Clifton, 1989; de Vincenzi, 1991, *inter alia*). Moreover, because of working memory constraints, it has been argued that people expect those predicted constituents to occur as soon as grammatically permissible in the input string (Gibson, 1998, and references therein; see also below).

Although much of the experimental and theoretical work in language processing has been based on English, the three tenets above (incrementality, immediate use of multiple sources of information, constituent prediction) are likely to apply to the processing of any human language. Despite this common assumption, it is also true that almost as often those tenets have been ignored tacitly or explicitly when the processing of Japanese was considered.

In what follows, we describe experimental work conducted at the SPL/KUIS that provides support for an approach to Japanese sentence processing in which the three tenets above are strictly obeyed. Section 2 describes work on case markers and incremental parsing. Section 3 discusses the processing of in-situ wh-phrases in regard to two aspects: their requirement for a question particle and their interaction with negative polarity items.

2. Case markers and incremental parsing

At first, incremental processing may seem at odds with head-final languages in the most flagrant manner possible even in a simple Japanese sentence like the following.

(1) Mary-ga John-ni hana-o ageta.
Mary-Nom John-Dat flower-Acc gave
"Mary gave flowers to John."

Incrementality requires each incoming word[1] to be associated with the previous words within a partial representation of the sentence. But how can a Japanese speaker associate **Mary**-Nom with **John**-Dat if neither word can take the other as an argument? Or for that matter, how can the first three NPs in the sentence be associated in a coherent representation if there is no predicate to take them as arguments?

As usual, there are two ways of dealing with this kind of dilemma. One possibility is to assume that sentence processing takes place in different manners in head-final languages when compared to head-initial languages. In the present case, this entails proposing that incrementality is obeyed in English, but not in Japanese. A second possibility is to propose that there is an extra factor that has to be taken into consideration in order for a uniform account to emerge for the two languages.

Non-incremental proposals in the literature assume that the NPs in (1) are initially stored as unrelated items in a list and can only be associated within a representation when the verb is detected. The basic justification is that a non-terminal node, such as a VP, cannot be projected until its head has been encountered (a performance-level version of the *projection principle*, Pritchett, 1991). However, this type of proposal runs into a number of problems. The most obvious is that, intuitively speaking, readers are able to extract a partial interpretation of the sentence before the verb is read ("Mary did something to the flowers for John. . . "), suggesting that some type of representation is built.

There is mounting evidence that rather complex representations are built in head-final clauses even before a verb is read, obeying incrementality just like in English (Miyamoto, 2000, and references therein). Incremental processing in Japanese is relatively unsurprising once we take into account a source of information that is not usually available in English, namely, case markers. In English, linear position and the early availability of the verb (and its argument structure) guide the parsing process. In Japanese, in contrast, case markers and the relative position of the NPs are the main sources of information. Moreover, the assumption that case markers have an immediate effect on the way how NPs are interpreted within a partial representation of the sentence is in accordance with the immediate use of multiple sources of infor-

1 In the case of Japanese, the term *word* will be used to refer to a content word plus inflection material and particles (such as case markers) that follow it.

mation during parsing.

It is assumed here that case markers help determine the kind of verb that is coming up, allowing an underspecified V node (and all its projections) to be created even before the verb itself is read (see Miyamoto, 2000, for a more detailed discussion and experimental evidence). Such a node is underspecified in that only some of its features are known in advance (e.g. it is a ditransitive verb), but this is enough to permit the NPs to be associated together in a coherent representation. When the verb is read, only the more specific aspects of its lexical information have to be added to the node previously created, and the features that were underspecified can be instantiated. Note that a similar process seems to occur in head-initial languages like English, so that a verb will predict the arguments that are needed to complete the clause. Thus, the parsing process of predicting what needs to come next is the same in both languages, but because of differences in word order, arguments in Japanese predict/subcategorize for the verb, whereas, in English, the verb predicts/subcategorizes for the arguments. Nowadays, to *subcategorize* is an asymmetrical relation in which the verb is the category doing the subcategorization, but in earlier work the term was used without this endocentric bias in the sense that both arguments and predicates were assumed to be subcategorized by the environment in which they occur (see, for example, Chomsky, 1965; a similar observation can be made for categorial grammars, Ades, and Steedman, 1982).

Experimental work being conducted at the SPL/KUIS explores the exact manner in which case markers influence parsing decisions on-line. For example, in a sentence fragment as the following, readers associate the first two NPs as part of a single clause (thus, predicting a transitive verb such as **saw**: "Ishii saw the baby").

(2)	Ishiisan-wa	akanboo-o	Tanakasan-ga. . .
	Ishii-Top	baby-Acc	Tanaka-Nom

When the third nominative NP is read, readers immediately create a clause boundary between the second and the third NPs because no verb in Japanese can take three NPs with such case markers as its arguments. Thus, the first two NPs predict a transitive verb and the third NP predicts an embedded verb (e.g. **sleep** as in "Ishii saw the baby in the room where Tanaka was sleeping.").

If a transitive verb such as **saw** follows after **Tanaka**, there are two possible interpretations. In one interpretation, the direct object in the embedded clause is taken to have been dropped (i.e., an empty category, *pro*, is inserted). Alternatively, the accusative NP **baby** is reanalysed so that it becomes part of the embedded clause (a possible sentence in this case would be "Ishii was sleeping in the room where Tanaka saw the baby."). Preliminary results suggest that readers prefer the former interpretation. This is consistent with proposals such as *reanalysis as last resort* (Frazier, 1990) according to which people avoid modifying the mental representation whenever possible. In the present case, the insertion of **pro** in the embedded clause allows for the initial interpretation (with **Ishii** and **baby** in the same clause) to be maintained.

The prediction of constituents is not restricted to arguments and verbs, but rather, it seems to be pervasive in language processing. The next section discusses the predictions generated by the processing of wh-phrases.

3. In-situ wh-phrases

Fronted wh-phrases in English occur in a non-argument position; thus, their processing requires that they be associated with a gap (i.e., an empty position for an argument or adjunct).[2] In East-Asian languages, in contrast, in-situ wh-phrases do not have an associated gap. In the following, we discuss the processing of in-situ wh-phrases in Japanese with respect to their requirement for a question particle and their interaction with negative polarity items.

3.1 Question particles

The processing of fronted wh-phrases in English has generated an extensive literature due to its special requirements (Frazier, and Clifton, 1989; de Vincenzi, 1991; *inter alia*). For example, in the following sentence, how is the representation built incrementally if the position of the gap only becomes clear at the preposition **for**? If people are building a partial representation as each word is read, what are they doing before the gap position is found in the sentence?

---

[2] It will not be relevant to the present discussion whether the gap is encoded as an empty category (the trace left by movement, Chomsky, 1981) or as a slash feature (Gazdar et al., 1985).

(3) Who$_i$ could the child have sung for <gap$_i$ >?

It has been proposed that wh-phrases create the expectation for an incoming gap and people try to insert such a gap as soon as grammatically permissible (the *active filler hypothesis*, Frazier, and Clifton, 1989). In general, if people have to keep track of all the expected constituents that are necessary to create a complete grammatical sentence, it is unsurprising that they should try to satisfy those requirements as soon as possible so that working memory is not overloaded with pending dependencies. However, the early creation of a gap can lead to mistakes as later words in the sentence may disconfirm the position of the gap. In the sentence below (adapted from Crain, and Fodor, 1985), readers attempt to create a gap for the wh-phrase **who** immediately after the verb **forced** (as indicated by the ∧ symbol). But this gap is incompatible with the next word, **us**, which turns out to be the true argument of the verb. The correct position for the gap occurs after the preposition **for**.

(4) Who$_i$ could the child have forced ∧ us to sing the song for <gap$_i$ >?

Experimental evidence using similar sentences indicates that readers have difficulty at **us** because they realize that the position of the gap created at ∧ must be filled by this pronoun. This phenomenon is known as the *filled-gap effect.*

Work in Italian suggests that the processing of wh-phrases and their gaps is similar to the processing of other types of empty positions, in that both involve the processing of syntactic chains. The ***minimal chain principle***, thus, proposes that people try to minimize the length of chains (e.g. **pro**, as a singleton chain, would be favoured in subject position over a postposed inverted subject; de Vincenzi, 1991).

One problem with the minimal chain principle is that it is not easily extendable to the processing of in-situ wh-phrases because this type of construction does not involve chains with gaps (but see Miyamoto, and Takahashi, 2000, for a potential solution in terms of chains created by LF movement or null operator movement). Two approaches are possible in order to resolve this problem. One can assume that the processing of wh-phrases involves two different mechanisms, thus maintaining the validity of the minimal chain principle for fronted wh-phrases in English and Romance, meanwhile proposing a different principle for in-situ wh-phrases in East-Asian languages. In an alternative approach, one may try to find a unifying principle that can explain the processing of wh-phrases in both types of languages.

We pursue the latter type of approach as we argue that the processing of fronted wh-phrases should not be characterized in terms of chains. As observed in the previous section, constituents very often create the expectation for a constituent to come (e.g. case marked NPs predict a specific type of verb). Similarly, wh-phrases can be argued to trigger the expectation for a constituent, independent of whether a chain is involved or not. But, the type of constituent predicted will vary depending on the type of wh-phrase being processed. Whereas fronted wh-phrases predict a gap, their counterparts in East-Asian languages require a question particle (Cheng, 1991). Moreover, following Gibson (1998), we assume that people have a preference for required constituents to occur as soon as possible in the input string because of working memory constraints. Thus, in Japanese, people should prefer a question particle to occur as soon as possible in the input string, in the same way that early gaps are preferred in English and Romance.

A self-paced reading experiment (Just, Carpenter, and Woolley, 1982) conducted at the SPL/KUIS supports this proposal in that Japanese speakers were slow to read the embedded verb with the affirmative complementizer in (5a) when compared to the question particle in (5b).

(5) a. Affirmative complementizer:

Senmu-ga	donna-pasokon-o	tukatteiru-**to**
director-Nom	what-type-computer-Acc	using-is-Comp
kakarichoo-ga	ittano?	
supervisor-Nom	said-QM	

"What type of computer did the supervisor say the director is using?"

b. Question particle:
Senmu-ga donna-pasokon-o tukatteiru-**ka**
director-Nom what-type-computer-Acc using-is-QM
kakarichoo-ga kiitano?
supervisor-Nom asked-QM
"Did the supervisor ask what type of computer the director is using?"

Readers know that the wh-phrase **what type of computer** requires a question particle and expect it to come in as soon as grammatically possible (i.e., after the embedded verb). However, when an affirmative complementizer is detected in (5a), readers slow down having to re-assess their expectations. Simply put, the present phenomenon is the in-situ counterpart of the filled-gap effect observed for English.

The result suggests that the processing of fronted wh-phrases should not be characterized in terms of a gap necessary to complete a chain. Instead, it should be generalized to the processing of expected constituents. One advantage in this assumption is that the algorithm for the processing of wh-phrases in English and in Japanese can be taken to be the same: it associates elements thus far in the input string and predicts the constituents that are necessary in order to complete the sentence. Parameterizations necessary to explain cross-linguistic variation are confined to the grammars of the two languages. In English, the grammar determines that a gap should be expected; whereas, in Japanese, a question particle is necessary. In both cases, the restricted amount of working memory available leads to the preference for expected constituents to be posited as soon as possible.

3.2 Intervention effects with negative polarity items

It has been known for quite a while that fronted wh-phrases in English are subject to island constraints, in the sense that a gap within certain types of constituents is prevented from being associated with a fronted wh-phrase (Ross, 1967). Recently, it has been observed that in-situ wh-phrases inside quantificational environments are precluded from being associated with their scope positions (Beck, 1996; Tanaka, 1997). In both types of phenomena, the association of a wh-phrase with another position is prevented by the interference of a special type of environment that works as a barrier.

Such a simplifying characterization does not do justice to the vast literature investigating the syntax of wh-phrases (some recent developments can be found in Pesetsky, 2000; Richards, 1997; among many others), but it makes explicit the generative assumption that both types of phenomena are to be treated in similar fashion in terms of environments that induce barriers preventing certain types of associations to take place across them. Naturally, apart from grammaticality judgements, independent evidence supporting this unified treatment of the two phenomena would be desirable. Below, we describe an event related potential (ERP) experiment being conducted in collaboration with professor Kenji Itoh at the University of Tokyo, in order to provide such an evidence.

Recent developments in neurolinguistics have made possible the investigation of neurological responses to semantic anomalous sentences (the N400, a negative wave form peaking around 400 msec after the onset of the anomalous word; Kutas, and Hillyard, 1984) in contrast to syntactically ill-formed sentences (the left anterior negativity or LAN, Kluender, and Kutas, 1993; the P600, Osterhout, and Holcomb, 1992). For example, subjacency violations generate a LAN peaking at around 200 msec after the relevant gap (Kluender, and Kutas, 1993).

If intervention effects with quantificational features are syntactic in nature, then they should produce a LAN or a P600. In contrast, it is conceivable that the interpretation of quantificational features belongs to the realm of semantics, in which case, intervention effects should yield a semantic anomaly response, the N400. In order to test these two alternative predictions, an ERP experiment is being conducted using sentences like the following.

(6) a. NPI intervention effect

Takagisan-shika	nani-o	gakushoku-de	tabenaika
Takagi-NPI	what-Acc	cafeteria-at	eat-not-QM
Taro-ga	tazunetano?		
Taro-Nom	asked		

b. Scrambled control

nani-o	Takagisan-shika	gakushoku-de	tabenaika
what-Acc	Takagi-NPI	cafeteria-at	eat-not-QM
Taro-ga	tazunetano?		
Taro-Nom	asked		

"Did Taro ask what nobody except Takagi eats at the cafeteria?"

In sentence (6a), the negative polarity item (NPI) **shika** has to be associated with a negative element (the negation following the embedded verb), and the sentence is uninterpretable because of the interaction between the in-situ wh-phrase **what** and the NPI/negation complex. In (6b), in contrast, the wh-phrase was dislocated to the beginning of the sentence; therefore, it does not intervene between the NPI and the negated verb, and the sentence is correctly interpreted. Note that the interpretability of the latter sentence suggests that the problem with the first sentence is unlikely to be due to semantics.

The outer clause ("did Taro ask') is not relevant and was only included so that the crucial region for comparison (**eat-not**-QM) did not coincide with the end of the sentences. This is because both behavioural and neurological evidence suggests that the last word in a sentence involves extraneous wrap-up effects of no interest for our present claims. The relevant contrast between the two sentences holds even when the outer clause is removed.

A behavioural pre-test confirmed that native speakers of Japanese find the first sentence significantly less acceptable than its control. In the actual ERP experiment, participants were not asked to judge the acceptability of the sentences and only a probe recognition task followed after each sentence.

If the waveforms for (6a) minus its control (6b) yield an early LAN after the onset of the embedded verb, it would indicate that subjacency violations and NPI intervention effects generate similar brain responses.

Moreover, the present experiment can test the proposal that the amplitude of the LAN is modulated by working memory load and not by syntactic violations (Kluender, and Kutas, 1993). In this type of model, subjacency violations are argued to be generated by wh-phrases being held in working memory across nodes that constitute syntactic barriers. In the present experiment, in contrast, the larger working memory load is expected for (6b), the sentence in which the wh-phrase is further from its base position. Consequently, the working memory interpretation of the LAN should predict a larger LAN for (6b).

4. Conclusion

The present paper described work in progress at the SPL/KUIS which supports a model of sentence processing in which parameterization is restricted to the grammar component, whereas the parsing algorithm itself is maintained constant across languages. We believe that work in Japanese processing can support findings in head-initial languages by adducing evidence from new constructions (such as the processing of case-marked NPs and question particles discussed above). Moreover, this type of work can also refine previous proposals by exploring their predictions in different environments (such as the working memory load interpretation of the LAN in NPI intervention effects). These are natural avenues that should be pursued in the investigation of general cognitive mechanisms underlying language processing.

Acknowledgements and contact information

The experimental results reported in the present paper were supported by a Grant-in-Aid for COE Research (08CE1001) from the Japanese Ministry of Education, Science and Culture granted to Kanda University of International Studies.

Correspondence should be sent to the first author at University of Tokyo, Faculty of Science, Dept of Information Science, Tsujii Lab., 7-3-1 Hongo Bunkyo-ku Tokyo 113-0033 JAPAN; or etm@is.s.u-tokyo.ac.jp.

References

Ades, Anthony E., and Mark Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4:517-558.

Altmann, Gerry, and Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30:191-238.

Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4:1-56.

Cheng, Lisa, L.-S. 1991. On the typology of wh-questions. Doctoral dissertation, MIT, Cambridge, Mass.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Crain, Stephen, and Janet D. Fodor. 1985. How can grammars help parsers? In *Natural language parsing*, ed., David R. Dowty, Lauri Karttunen and Arnold M. Zwicky, 94-128. Cambridge: Cambridge University Press.

de Vincenzi, Marica. 1991. *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer Academic Publishers.

Elman, Jeffrey L. 1991. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning* 7:195-225.

Frazier, Lyn. 1990. Identifying structure under X0. In *Yearbook of Morphology* 3, ed. Geert Booij and Jaap van Marle, 87-105. Amsterdam: Foris.

Frazier, Lyn, and Charles Clifton. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4:93-126.

Garnsey, Susan, Neal Pearlmutter, Elizabeth Myers and Melanie A. Lotocky. 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 37:58-93.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, Mass.: Harvard University Press.

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

Just, Marcel A., Patricia A. Carpenter and Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology*: General 3:228-238.

Kluender, Robert, and Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8:573-633.

Kutas, Marta, and Steven A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307:161-163.

Marslen-Wilson, William, and Lorraine K. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8:1-71.

Marslen-Wilson, William, and Lorraine K. Tyler. 1981. Central processes in speech understanding. *Philosophical Transactions of the Royal Society*, Series B, 295:317-332.

Miyamoto, Edson T. 2000. Case markers as clause boundary inducers in Japanese. Under review.

Miyamoto, Edson T., and Shoichi Takahashi. 2000. The processing of wh-phrases and interrogative complementizers in Japanese. To appear in the *Proceedings of the 10th Japanese and Korean Linguistics Conference*. UCLA, Los Angeles, CA.

Osterhout, Lee, and Phillip J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31:785-806.

Pesestky, David. 2000. *Phrasal movement and its kin*. Cambridge, Mass.: MIT Press.

Pritchett, Bradley L. 1991. Head position and parsing ambiguity. *Journal of Psycholinguistic Research* 20:251-270.

Richards, Norvin. 1997. What moves where in which language? Doctoral dissertation, MIT, Cambridge, Mass.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Tanaka, Hidekazu. 1997. Invisible movement in sika-nai and the linear crossing constraint. *Journal of East Asian Linguistics* 6:143-188.

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