The Processing of Wh-phrases and Interrogative Complementizers in Japanese

Edson T. Miyamoto,1,2 and Shoichi Takahashi1

1 Kanda University of International Studies
2 University of Tokyo

1. Introduction

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 relevant during parsing, 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) as each word is processed. We also have to consider whether 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 as well.
The present paper argues that, despite their syntactic differences, wh-phrases in English and Japanese present similar processing properties supporting the assumption that a single mechanism along the lines of Gibson (1998) may be at work in both cases. Thus, the parameterization is restricted to the grammars: In English, fronted wh-phrases require an associated gap, whereas in Japanese, in-situ wh-phrases require a question particle. Given the dependencies required by a grammar, cognitive mechanisms keep track of the constituents that are still needed as the sentence is read. Moreover, because of working memory constraints, shorter dependencies are favoured over longer ones, in other words, required constituents are preferred to occur early rather than late in the input string. Under these assumptions, we also discuss how certain types of movements without phonological effect (LF movement, Huang, 1982; May, 1985; and null-operator movement, Watanabe, 1992) may be computed by people.

2. The processing of wh-phrases

The fact that people read one word at a time and attempt to create a partial representation without delay may seem unsurprising and is supported by various experimental results (Marslen-Wilson & Tyler, 1980, *inter alia*). However, we immediately face problems in the processing of fronted wh-phrases in languages such as English. For example, in (1) below, the wh-phrase who is fronted to the beginning of the sentence and, therefore, it must be associated with a gap (i.e. an empty argument or adjunct position)\(^1\) in order for its relation with a predicate to be determined. But how is the representation built incrementally if the position of the gap only becomes clear at the end of the sentence?

(1) Who did the children sing for <gap>?

Moreover, if a uniform account of wh-phrases across languages is to be attained, there are clear differences that need to be taken into consideration. While languages such as English require wh-phrases to be fronted,\(^2\) East Asian languages allow wh-phrases to remain in situ. Thus, in a question such as (2a), Japanese allows the wh-phrase what kind of computer to occur in the same position as a lexical NP such as computer in (2b).

\(^1\)It will not be crucial in the present discussion how the gap is encoded (e.g. as a trace left by movement, Chomsky, 1981; as a slash feature, Gazdar, et al., 1985).

\(^2\)We will not address the processing of sentences with multiple wh-phrases in this paper.
To appear: Proc. 10th Japanese/Korean Linguistics Conference.  3

(2) a. senmu-ga donna-pasokon-o tukatteiru-no?
director-Nom what-kind-computer-Acc using-is-qp
‘What kind of computer is the director using?’

b. senmu-ga pasokon-o tukatteiru.
director-Nom computer-Acc using
‘The director is using the computer.’

It is possible that fronted wh-phrases in English and in-situ wh-phrases in Japanese are too different to be handled by a single processing mechanism. In fact, if we only consider the position where they occur, in-situ wh-phrases are more likely to be processed in the same manner as lexical NPs such as computer in (2b). However, the processing of in-situ wh-phrases cannot be reduced to the processing of lexical NPs because the former have a further requirement, namely, wh-phrases in Japanese have to be licensed by a question particle such as no in (2a) (Cheng, 1991; Nishigauchi, 1990). In the following, we will draw a parallel between this requirement and the gap-search process triggered by wh-phrases in English.

3. Fronted wh-phrases in English

One common assumption in the literature on the processing of fronted wh-phrases in English is that a wh-phrase triggers a search process for a gap (see Fodor, 1989, for a summary). This can be expressed as a dependency that is still pending; the constituent (namely a gap) that will complete this dependency has yet to come in the input string. The expectation generated by this process modulates the way how people process the following up-coming words. A successful line of work has suggested that people try to posit the required gap as soon as grammatically possible, in other words, people favour positing a gap rather than expecting an overt constituent (de Vincenzi, 1991; Frazier & Clifton, 1989; and references therein). Supporting this proposal, experimental evidence suggests that native speakers of English have difficulty when a potential position for a gap turns out to be unavailable because of an overt NP. For example, in the sentence below (adapted from Crain & Fodor, 1985), readers insert a gap immediately after force as indicated by the ∧ sign. Next, people slow down when the pronominal us is read because it is incompatible with the gap previously posited. The search is resumed and the eventual position of the gap is found after the preposition for.

(3) Filled gap effect (FGE)
Who did the children force ∧ us to sing the songs for <gap>?
This type of phenomenon, in which a gap posited too early turns out to be filled by a lexical NP, is known as the filled gap effect (FGE).

The preference to posit a gap as soon as possible has been explained in terms of a requirement to minimize the length of chains (the minimal chain principle, MCP; de Vincenzi, 1991). One of the advantages of the MCP is that it provides a uniform account for the processing of wh gaps as well as other types of empty positions (e.g., pro taken as a singleton chain).

4. In-situ wh-phrases in Japanese

One problem with the MCP is that it does not easily generalize to constructions without gaps. For example, the MCP does not apply to the processing of in-situ wh-phrases in East-Asian languages because they do not involve a chain with an empty position (but see discussion below for a possible extension of the MCP).

In an alternative proposal, it has been suggested that the preference for early gaps in English is an instance of a more general preference for required constituents to come as soon as possible in the input string (Gibson, 1998). In this case, whether a chain is involved or not is irrelevant, and any type of predicted constituent should preferentially occur as soon as possible. This model can be extended to East-Asian languages as it predicts that people would prefer a constituent required by in-situ wh-phrases to occur early rather than late in the input string.

The question then is whether in-situ wh-phrases predict an upcoming constituent. It has been observed that East-Asian languages and various other languages with in-situ wh-phrases require a question particle in order to indicate that a sentence should be typed as interrogative (Cheng, 1991). Moreover, in-situ wh-phrases have been argued to require unselective binding by a relevant element in complementizer position (Pesetsky, 1987, and references therein). In Japanese, in particular, wh-phrases require a question particle such as ka or no in order to be licensed. Thus, following Gibson (1998), Japanese speakers should prefer such a particle to occur as soon as possible after a wh-phrase. If the parallel with the processing of fronted wh-phrases in English holds, we should be able to produce in Japanese an effect similar to the FGE. The rest of this section describes a phenomenon which we will argue to be the in-situ counterpart of the FGE.

In the following sentence, the wh-phrase *donna*-*pasokon* requires a

\[ ^3 \text{The actual requirement is for a particle that can determine the force of the wh-phrase (as interrogative, existential or universal; Nishigauchi, 1990). For example, a non-interrogative particle such as mo could also license the wh-phrase, in this case, as a universal quantifier.} \]
question particle, whose earliest possible position would be immediately after the verb using-is. Therefore, Japanese speakers should slow down when they read the complementizer to because its affirmative typing clashes with the expectation for a question particle and gives rise to a typing mismatch effect (TME).

(4) Typing mismatch effect (TME)

\[
\text{senmu-ga} \quad \text{donna-pasokon-o} \quad \text{tukatteiru} \land \text{to} \\
\text{director-Nom} \quad \text{what-kind-computer-Acc} \quad \text{using-is} \quad \text{Comp[aff]}
\]

\[
\text{kararichoo-ga} \quad \text{ittano?} \\
\text{supervisor-Nom} \quad \text{said-qp}
\]

‘What kind of computer did the supervisor say the director is using?’

In order to test whether Japanese readers slow down at the affirmative complementizer, we conducted a non-cumulative moving-window self-paced reading experiment (Just, Carpenter & Woolley, 1982; the segmentation used is indicated with spaces in the example sentences). We compared the sentence above (repeated as (5a) below) with the sentence in (5b). In the latter sentence, the embedded verb is followed by the question particle ka, thus providing a base-line reading time for the embedded verb with the affirmative complementizer in (5a).

(5) a. Wh/Aff

\[
\text{senmu-ga} \quad \text{donna-pasokon-o} \quad \text{tukatteiru-to} \\
\text{director-Nom} \quad \text{what-kind-computer-Acc} \quad \text{using-is-Comp[aff]}
\]

\[
\text{kararichoo-ga} \quad \text{ittano?} \\
\text{supervisor-Nom} \quad \text{said-qp}
\]

‘What kind of computer did the supervisor say the director is using?’

b. Wh/Int

\[
\text{senmu-ga} \quad \text{donna-pasokon-o} \quad \text{tukatteiru-ka} \\
\text{director-Nom} \quad \text{what-kind-computer-Acc} \quad \text{using-is-qp}
\]

\[
\text{kararichoo-ga} \quad \text{kiitano?} \\
\text{supervisor-Nom} \quad \text{asked-qp}
\]

‘Did the supervisor ask what kind of computer the director is using?’

However, there is the possibility that the affirmative complementizer to is always read more slowly than the question particle ka, independent of the environment in which these two types of particles occur. In order to counter this possibility, we created two more conditions in which the wh-phrase was replaced with a lexical NP, new-computer, as shown in
(6ab).

(6) a. NP/Aff
senmu-ga atarasii-pasokon-o tukatteiru-to
director-Nom new-computer-Acc using-is-Comp(aff)
kokarichoo-ga itta.
supervisor-Nom said
‘The supervisor said that the director is using the new computer.’

b. NP/Int
senmu-ga atarasii-pasokon-o tukatteiru-ka
director-Nom new-computer-Acc using-is-QP
kokarichoo-ga kiita.
supervisor-Nom asked
‘The supervisor asked whether the director is using the new computer.’

If the wh-phrase is indeed predicting a question particle, we should expect a slow-down in the third region (the embedded verb and the complementizer) of (5a) in comparison to the corresponding region in (5b). In contrast, there should be no such a difference between the NP conditions in (6) because new computer does not require a question particle. The results of the experiment supported our claims (see Figure 1). The Wh/Aff condition was slower than the Wh/Int condition; whereas, the NP/Aff condition was faster than the NP/Int condition (interaction Wh/NP × Aff/Int: $F_1(1,24) = 5.39, P < 0.05; F_2(1,23) = 5.36, P < 0.05$; see Miyamoto & Takahashi, 2000, for further details). The difference between the two NP conditions can be explained if we assume that the lexical NP does not require a question particle, and furthermore, that the affirmative typing is the default. Under those two assumptions, the slow-down in the NP/QP condition occurs because some extra processing is required when the question particle is detected; in other words, the question particle is the first constituent that indicates that the sentence must be typed as interrogative, contradicting the default affirmative typing that was probably assumed up to that point.

The present result supports the assumption that there is a TME taking place at the affirmative complementizer in the Wh/Aff condition, (5a). The parallel with the FGE in (3) suggests that a similar process is taking place in both cases: a wh-phrase is read and requires a certain type of constituent, and a slow-down is observed when the required constituent does not occur at the earliest possible position.
5. Grammars and cognitive resources

The previous section suggests that parameterizations can be restricted to the grammars, allowing the processing algorithm to be characterized uniquely for both types of languages at least with respect to wh-phrases.

As observed earlier, the syntactic properties of wh-phrases in English and Japanese are so different that, at first sight, it would seem more natural to propose two separate mechanisms to deal with their processing. The FGE and the TME certainly differ in many respects. In the FGE, the fronted wh-phrase is not associated with a predicate, consequently, its thematic role remains uncertain until the gap is found. In the TME, by contrast, the in-situ wh-phrase is associated with a local predicate in order to get a thematic role, but its scope is undetermined until the question particle is found.

The model we are adopting here implies that the processing mechanism may be blind to the fact that, at the grammar level, thematic relations are distinct from scope relations. Even though linguists have various reasons to keep those two types of relations separate, the present model suggests that the only relevant aspects for the processing algorithm are the positions that have to be associated and how far apart they are from each other.

More precisely, a syntactic model has to account for two aspects of wh-phrases — their requirements for a thematic role and for a scope position. In most cases, it is not possible to satisfy both requirements locally; one of them has to be encoded as a long distance dependency. As observed earlier, fronted wh-phrases indicate their scope position...
based on where they are pronounced in a sentence, but their thematic role has to be determined through a long distance dependency with a gap. In contrast, the thematic role of in-situ wh-phrases is indicated by the position where they are pronounced, whereas their scope is expressed as a long distance dependency with a question particle. In our processing model, the long distance dependency is the factor that exposes the cognitive bottleneck which is being measured in experimental settings. The presumable bottleneck is the restricted capacity that working memory has to hold various items (in the present case, various unresolved dependencies) for relatively long periods of time. Clearly, it is entirely possible that future experimental techniques will uncover more intricate patterns that will require differentiated treatments of these dependencies even at the processing level.

Note that, in the present model, even if the exact nature of the dependencies is not crucial, nevertheless, the grammar is strictly obeyed during parsing as people do not make random predictions in order to circumvent working memory constraints.

6. Extending the minimal chain principle

The minimal chain principle (MCP) was originally proposed for chains with empty positions, in other words, for constituent omission (e.g. instances of pro) or overt movement (de Vincenzi, 1991). We contended that the MCP cannot apply to the processing of in-situ wh-phrases because these constructions do not involve a gap. However, an extended version may be possible, in which chains generated by movement without phonological effect may also be subject to the MCP. The present section discusses how such movements may be processed by people as they read a sentence, and whether such a proposal could explain the TME.

6.1 Movements without phonological effect

Two types of syntactic movements have been proposed in order to explain in-situ wh-phrases. In one proposal, in-situ wh-phrases move covertly at the level of logical form (LF; Huang, 1982). In another proposal, a phonologically-null operator associated with wh-phrases moves to the relevant CP Spec before LF (Watanabe, 1992). In either case, movement does not have any phonological effect and a crucial question is when people compute such types of movements. Take LF movement, for example. In syntax, LF movement is assumed to take place after spell-out, or after S-structure has been sent to the phonological component. Thus, one possibility in processing is that people only compute LF movements at sentence end, after all overt movements have taken
place. However, coupled with the MCP, this proposal does not explain the TME. If LF chains are only computed at the end of the sentence, they cannot be related to the mid-sentence slow-down observed at the affirmative complementizer in (5a).

Alternatively, it can be argued that the steps in the processing of a sentence (i.e., in the performance module) do not have to mirror the order in the grammatical formalization (i.e., the order observed in the competence module). In this case, the computation of LF movement may occur at any point during the processing of a sentence; in particular, given the general processing preference for relations to be asserted as soon as possible (some sort of don’t procrastinate policy in psycholinguistics), LF movement may be processed as soon as it is clear that such a movement is required. (A similar rationale is true for null-operator movement.) The earliest point in which it is clear that movement is necessary is at the wh-phrase. In the sentences below, assume that covert movement occurs as soon as the wh-phrase is detected. As observed earlier, such a movement does not have phonological consequences, and is represented in (7a) with the null operator \( \wedge \) in Spec of CP1.\(^4\)

\[
\begin{align*}
(7) \quad & a. \quad \underline{\text{[} \, r_1 \, \text{]} \wedge \text{senmu-ga  donna-pasokon-o}} \\
& \quad \text{director-Nom wh-computer-Acc} \\
& b. \quad \underline{\text{[} \, r_2 \, \text{]} \wedge \underline{\text{[} \, r_1 \, \text{]} \, \text{senmu-ga  donna-pasokon-o  tukatteiru-to \]}} \\
& \quad \text{director-Nom wh-computer-Acc using-is-Comp[aff]} \\
\end{align*}
\]

\(^4\)We crucially assume in the present analysis that the specifier in Japanese is on the left of the complementizer, thus preceding the IP node.

\[\text{\textbullet Figure 2 The TME as reanalysis of a null operator.}\]

When the affirmative complementizer to is read, an outer clause
is created (as indicated by CP2 in (7b); see also the tree structure in Figure 2). Because the null operator in Spec of CP1 is not compatible with the affirmative to, reanalysis must take place and the null operator is shifted to Spec of CP2 indicating that the wh-phrase takes scope in the higher clause, CP2, and not in the embedded CP1. Consequently, the slow-down observed at the affirmative complementizer is predicted to have been caused by the mismatch between the null operator and the complementizer, or by the reanalysis process necessary to change the scope of the wh-phrase.

6.2 The processing of LF movement: QR in English

As discussed above, the extended MCP may account for the TME, but it requires a number of unattested assumptions about the computation of movements without phonological effect. It would be desirable to obtain independent evidence that LF chains or null-operator chains are computed on-line and influence processing time. The present section discusses potential evidence in support of the processing of LF movement during on-line parsing; the next section considers the same discussion for null-operator chains.

If LF chains are relevant during processing, in other words if they can add difficulty to the processing of a sentence, then the extended MCP should predict that the direct object in (8a) would take less time to process than the direct objects in (8bc) because the latter NPs require quantifier raising (QR; May, 1989) to a position adjoined to IP, as indicated by ∧.

(8) a. Object NP without quantifier
   A man bought the painting.

b. Object NP with quantifier, near its scope position
   ∧ A man bought every painting.

c. Object NP with quantifier, far from its scope position
   ∧ A man who the reporter talked about bought every painting.

The NP the painting does not contain a quantificational feature and, consequently, does not involve LF movement. In contrast, every painting and other NPs with a quantifier require QR. If QR is computed as soon as the quantificational feature is detected, we should expect the direct objects in (8bc) to take longer to process than the direct object in (8a). Moreover, if longer LF chains are harder to process than shorter LF chains, the direct object in (8b) should be easier to process than the direct object in (8c). However, intuitive judgements suggest that there are no noticeable differences between the sentences
in (8). One possible confound in this case is that discourse complexity factors may be masking the LF-chain effect given that each determiner requires a different set of presuppositions in those sentences (Crain & Steedman, 1985, inter alia).

In related work, it has been proposed that the more elements in a chain, the more complex is the syntactic representation perceived to be (the representational theory of complexity, RTC; Pritchett & Whitman, 1995). The RTC was proposed as an updated version of the derivation theory of complexity (see an extensive critique of the DTC in Fodor, Bever & Garrett, 1974). It applies to chains in S-structure as well as in LF, and it is supposed to explain preferences in the interpretation of ambiguous sentences in English and Japanese. However, the RTC runs into a number of conceptual and empirical problems. Like its predecessor, it fails to give an explicit parsing algorithm and consequently only provides an off-line measure of complexity; in other words, the RTC only makes predictions of people’s preferences at sentence end and does not explain the word-by-word process. Moreover, the examples provided by Pritchett and Whitman seem to be explained by independent factors, making the contribution of the RTC unclear. Finally, and perhaps most problematic is the assumption that only the number of chain elements influences processing difficulty, whereas the distance between two consecutive elements in a chain does not (Pritchett & Whitman, 1995, pp73-6), given that, at least in S-structure chains, it is well known that both measures are relevant (de Vincenzi, 1991; Gibson, 1998).

In short, we would argue that the influence of LF chains during on-line parsing is dubious at this point. But this is a little explored area and further work is clearly necessary.

6.3 The processing of null-operator movement

If it is eventually shown that LF chains for QR do not involve a processing cost, it will be harder to argue for the processing cost of LF chains in Japanese wh-phrases. However, independent of the conclusion about QRs, it should still be possible to argue that null-operator movement (Watanabe, 1992) has some processing impact. This type of movement was originally proposed for the level of S-structure and is not directly related to LF chains.

As a corollary of the MCP, assume that longer chains take longer to be processed than shorter chains. Thus, a wh-phrase should take longer to process, the farther it is from the CP Spec containing its null operator. For example, the null operator $\Lambda$ in CP Spec is closer to the the subject wh-phrase in (9a) than to the object wh-phrase in (9b).
Consequently, less processing time should be required for the subject wh-phrase in (9a) than the object wh-phrase in (9b). Sentence (9c) is given as a possible baseline for each wh-phrase in the first two conditions.\textsuperscript{5}

(9) a. Subject wh
\[ \wedge \text{dono-senmu-ga pasokon-o tukatteiru-ka...} \]
which-director-\textit{Nom} computer-\textit{Acc} using-is-QP

b. Object wh
\[ \wedge \text{senmu-ga dono-pasokon-o tukatteiru-ka...} \]
director-\textit{Nom} which-computer-\textit{Acc} using-is-QP

c. Lexical NPs
\[ \text{senmu-ga pasokon-o tukatteiru-to...} \]
director-\textit{Nom} computer-\textit{Acc} using-is-\textit{Comp(aff)}

The distance between the null operator and the question particle ka is the same in (9ab), thus the extended MCP should not predict any differences at the particle. Moreover, in this model, an in-situ wh-phrase has a dependency with a null operator, which in turn is associated with a question particle; but there is no direct dependency between the wh-phrase and the question particle.

In the original explanation we provided for the TME, the crucial dependency is between the wh-phrase and the question particle; there is no null operator in Spec of CP serving as an intermediary. In this case, there should not be any differences in the processing times of the wh-phrases in (9ab), because no null operator is posited. However, the question particle should be read more slowly in (9a) than in (9b) because it is farther from the wh-phrase in the former sentence.\textsuperscript{6}

We are currently investigating the extended MCP and the TME in a number of different environments. As a long term project, we would like to explore other possible constructions in which LF movement or null-operator movement may affect processing times.

\textbf{Conclusion}

The present paper argued that the processing of in-situ wh-phrases in Japanese can be explained by the same procedure that underlies the processing of fronted wh-phrases in English. In this model, cognitive

\textsuperscript{5}This example could also be explained with LF chains.

\textsuperscript{6}We are adopting a linear distance metric here (according to Gibson, 1998, the crucial measure should take into account the number of intervening discourse referents). A structural metric would yield different predictions because, in a tree structure, the question particle in complementizer position is closer to the subject than to the object NP.
mechanisms attempt to satisfy as early as possible the dependencies dictated by the grammar of a given language (Gibson, 1998). The proposal was supported by the results of a reading time experiment which revealed a typing mismatch effect similar to the filled gap effect previously reported for English (Frazier & Clifton, 1989, inter alia). An alternative explanation was also discussed, which takes into consideration chains created by LF movement (Huang, 1982; May, 1985) or null-operator movement (Watanabe, 1992).

Acknowledgements and contact information
The experimental result reported in the present paper was 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 etm@is.s.u-tokyo.ac.jp or University of Tokyo, Faculty of Science, Dept of Information Science, Tsuji Lab., 7-3-1 Hongo Bunkyo-ku Tokyo 113-0033 JAPAN.

References