ISSN 2959-6122

Semantic Effects of Long-Distance Wh-Scrambling in Japanese

Zhen Wang

School of Arts, Languages and Cultures, The University of Manchester, Manchester, The United Kingdom
zhenwang@ldy.edu.rs

Abstract:

This paper investigates the semantic effects of long-distance *wh*-scrambling in Japanese. Through qualitative analysis, it demonstrates that long-distance scrambling is a semantically vacuous movement in declarative sentences where a *wh*-phrase undergoes radical reconstruction for interpretation. It is demonstrated that it need not flow into semantic interpretations, in contrast to *wh*-movement. Long-distance *wh*-scrambling, on the other hand, has an impact on sentential readings in interrogative sentences since the *wh*-phrase must draw scope from its nearest question marker, which must be understood through the c-command relationship. The difference is discovered to originate from a processing technique, with each alternative interpretation being more prominent in both the jumbled sentence and the canonical form. This study provides a syntactic-semantic perspective and offers new insights into the semantic implications of scrambling, thus contributing to the existing literature.

Key words: Long-distance *wh*-scrambling; semantically vacuous; scope; processing strategy

1. Introduction

In Japanese, word disposition is deemed free. With the verb being positioned sentence-finally, phrases like determiner phrase (DP) and complementizer phrase (CP) can be moved to a clause-initial position. This free word permutation develops from the notion of "scrambling" [1]. There are two types of scrambling: long-distance (LD) scrambling, in which a phrase is transferred from its original location to a higher clause; and short-distance (SD) scrambling, which occurs clause-internally. However, regardless of the kind of scrambling, it is thought that, in the absence of other variables like prosody, it has no effect on the meaning of a sentence when compared to its canonical form [2-3].

Japanese is also a *wh*-in-situ (i.e., *wh*-in-place) language, meaning when eliciting *wh*-questions, *wh*-phrases like *nani* 'what', *doko* 'where', *itsu* 'when', etc. remain in their original position within the sentence rather than obligatorily moving them to the front. It is not the same as English in this regard, where all *wh*-phrases must be inserted sentence-initially in a *wh*-question. *Wh*-phrases can also be jumbled, but this is thought to be optional [4].

This paper explores the semantic effects of LD *wh*-scrambling in Japanese. This research investigates the conditions under which *wh*-scrambling is regarded as semantically vacuous or not at logical form (LF) in various Japanese sentence forms, based on qualitative analysis. This work attempts to add new insights to the existing research by investigating this phenomenon and providing a syntactic-semantic level viewpoint on the interaction.

2. Literature Review

2.1 Scope

A critical aspect when examining the performance of *wh*-phrases is their scope. The location of the *wh*-phrase and the question (Q) marking *ka*, which c(onstituent)-commands it, determine the *wh*-scope in Japanese [5].

(1-a) Haruto-wa [$_{CP}$ Saki-ga itsu satta ka] sitte-iru
Haruto-TOP Saki-NOM when leave.PST Q know.GER-be.PRS
'Haruto knows (the time) when Saki left.'

Example (1-a) is a declarative sentence, and *itsu* takes embedded scope since the Q-marker appears in the embedded CP. However, if *itsu* were to be LD scrambled into the matrix clause, the sentence would be totally grammatical without changing its original meaning:

(1-b) [$_{IP}$ Itsu$_i$ [$_{IP}$ Haruto-wa [$_{CP}$ Saki-ga t_i satta ka]] sitte-iru]
when Haruto-TOP Saki-NOM leave.PST Q know.GER-be.PRS

Given that (1-b) is not ungrammatical and *itsu* takes only embedded scope, Saito argued that scrambling at LF can be undone unlimitedly [6]. It means that the *wh*-phrase in the scrambled position does not disrupt the necessary c-command relationship with the Q-morpheme *ka*. Further supporting this hypothesis, Tsutsumi stated in his recent work that an LD scrambled quantifier gets scope in an embedded CP instead of a quantified subject in the matrix CP [7]. This goes beyond *wh*-scrambling.

2.2 Previous Studies

In 1995, Chomsky proposed that the syntax of a language is shaped by its semantics, and that syntax must not eliminate any semantically substantial material [8]. However, Brody proposed in a critique of Chomsky's work that there is a propensity for empirically illogical transformations to happen even when syntactic perfection is sought after [9]. This suggests that even while syntax and semantics are tightly related, their alignment may not always coincide.

Indubitably, the studies on the effects of *wh*-scrambling in Japanese on syntax and semantics have been widely debated. In support of *wh*-scrambling in Japanese counts as *wh*-movement, Takahashi claimed that *wh*-scrambling can be seen as a complete *wh*-movement, representing an overt relocation of the *wh*-phrase to the position where it takes scope [10]. In a similar vein, Abe suggested that LD *wh*-scrambling functions like *wh*-movement and generates a focus chain [11]. On the other hand, Sauerland proposed—contrary to Chomsky's suggestion [8]—that scrambling in Japanese might eliminate semantic information. He claimed that one of the scrambling features, [-interpretable], deletes itself once the syntactic action is finished, meaning that it has no bearing on interpretation [12]. Sabel argued that *wh*-scrambling in Japanese cannot be seen as a complete *wh*-movement and cannot be undone at LF, viewing LD *wh*-scrambling as an adjunction movement [13]. More recently, Shimamura and Tanaka contended against the need to postulate *wh*-scrambling as an independent syntactic operation in Japanese [14].

However, these diverse perspectives reveal a research gap between discussing the relationship of *wh*-scrambling and *wh*-movement and understanding the semantic implications of LD *wh*-scrambling. Further research is required to examine the relationship between the Q-marker *ka* and the LD syntactic performance, as well as how different processing strategies in Japanese can affect the way sentences are understood. Thus, the purpose of this work is to close the gap.

3. Discussions

3.1 LD Wh-Scrambling as Semantically Vacuous Movement in Declaratives

The idea that the understanding of a scrambled sentence stays semantically unaltered even after a *wh*-phrase has been moved is known as semantically vacuous *wh*-scrambling. To put simply, it appears as though there was no scrambling at all. (1-a) and (1-b) exemplify a perfect pair, proving that LD scrambling of a *wh*-phrase is considered a semantically vacuous movement.

The following pair of sentences shows a slightly more complex situation:

(2-a) [$_{TP}$ Kaito-wa [$_{CP}$ [$_{TP}$ Sana-ga [$_{CP}$ Akari-ga dono basu-o notta to]
Kaito-TOP Sana-NOM Akari-NOM which bus-ACC take.PST that
omotte-iru] ka] siritagatte-iru]
think.GER-be.PRS Q want.to.know.3.GER-be.PRS
lit. '[Kaito wants to know [$_{Q}$ [Sana thinks [that Akari took which bus]]]]'
= 'Kaito wants to know which bus Sana thinks that Akari took.'

(2-b) [$_{TP}$ [$_{CP}$ Akari-ga dono basu-o notta to]$_{i}$ [$_{TP}$ Kaito-wa [$_{CP}$ [$_{TP}$ Sana-ga
Akari-NOM which bus-ACC take.PST that Kaito-TOP Sana-NOM
t_{i} omotte-iru] ka] siritagatte-iru]]
think.GER-be.PRS Q want.to.know.3.GER-be.PRS
lit. '[[That Akari took which bus]$_{i}$, Kaito wants to know [$_{Q}$ [Sana thinks t_{i}]]]'

Sentence (2-b) conveys the same meaning as (2-a). To these declarative statements with an implicit query, Kaito would like to know (the fact) that Sana thinks Akari took bus number 86. In (2-a), *ka*, the head of the outside CP, assigns scope to *dono basu*, which is positioned deeply within the embedded CP. Because of the LD scrambling of the embedded CP from the outer CP in (2-b), *dono* is no longer able to remain at its initial c-commanded place where it takes scope. However, given (2-b)'s grammaticality, its identical meaning compared with the canonical form (2-a), and their shared answer, *dono* must take scope within the embedded *wh*-question, in spite of its scrambled position outside of it. *Dono* goes through "radical reconstruction," what Saito advocated to re-enter the embedded *wh*-question [6], as though this scrambling had never occurred. As was previously established, the only way this LD scrambling at LF can be "freely undone" is because it is semantically meaningless and has no bearing on anything.

There is further evidence showing LD *wh*-scrambling is semantically vacuous, different from *wh*-movement in that *wh*-scrambling need not feed into semantic interpretation while *wh*-movement must:

(3-a) [$_{CP}$ Who$_{i}$ [$_{TP}$ t_{i} knows [$_{CP}$ [which book by whom]$_{j}$ [$_{TP}$ she read t_{j}]]]]

(3-a-i) Peter knows which book by William Shakespeare she read, and Robin knows which book by Jane Austen she read.

(3-a-ii) Peter knows which book by whom she read.

(3-b) *[$_{CP}$ [Which book by whom]$_{j}$ does [$_{TP}$ she know [$_{CP}$ who$_{i}$ [$_{TP}$ t_{i} read t_{j}]]]]

(3-b-i) She knows who read *Hamlet* by William Shake-

speare.

Sentence (3-a) is a *wh*-question that has experienced *wh*-movement, containing an echo question. In the embedded CP, *which* ends at the initial position, pied-piping *book of whom* along with it. This determines that the scope of *which* is larger than *whom*. *Who* accepts either scope because they are in both matrix and embedded clauses. As a result, for (3-a), (3-a-i) and (3-a-ii) are both appropriate responses. On the other hand, trying to induce *wh*-scrambling in English leads to (3-b)'s ungrammaticality. *Wh*-phrases are removed from the embedded question, degrading it because *which* and *whom* must assume the matrix scope. Thus, the unambiguous answer (3-b-i) is a proper response to (3-b).

Similarly, (2-b) parallels (3-b) in configuration, and it is also degraded from (2-a). While (3-a) and (3-b) demonstrate that, in order to ascertain the *wh*-scope in a sentence following *wh*-phrases having undergone *wh*-movement, both their final landing site and their original position need to be taken into consideration, (2-b) clearly permits the interpretation of the *wh*-phrase within the embedded question, even if it is LD scrambled.

In all, this section demonstrates how LD *wh*-scrambling in Japanese is regarded as a semantically vacuous movement in declarative sentences with embedded questions. Additionally, it provides evidence that *wh*-scrambling differs from *wh*-movement is due to the fact that it need not feed into semantic interpretations.

3.2 Semantic Effects of LD Wh-Scrambling on Interrogatives

Although LD *wh*-scrambling in declarative sentences in Japanese is considered a semantically vacuous movement, it affects sentence interpretation in interrogatives:

(4-a) [$_{CP1}$ Suzuki-ga [$_{CP2}$ Inoue-ga nani-o kaita ka] siritagatte-imasu ka]

Suzuki-NOM Inoue-NOM what-ACC write.PST Q_2 want.to.know.3.GER-be.POL.PRS Q_1

(4-p) 'Does Suzuki want to know what Inoue wrote?'

(4-b) [$_{CP1}$ Nani-o$_i$ Suzuki-ga [$_{CP2}$ t_i' Inoue-ga t_i kaita ka]

what-ACC Suzuki-NOM Inoue-NOM write.PST Q_2 siritagatte-imasu ka]

want.to.know.3.GER-be.POL.PRS Q_1

(4-q) 'What does Suzuki want to know whether Inoue wrote?'

The above pair differs from those previous pairs in that it contains each Q-marker *ka* in the matrix and the embedded CP. It is noteworthy that either (4-a) or (4-b) has both grammatical readings of (4-p) and (4-q). Though (4-a) is more commonly understood as a yes/no question with an embedded *wh*-question, the embedded scope reading (4-p) is preferred because *ka* can serve as a complementizer as well as a scope marker for a *wh*-phrase. In contrast, (4-b) causes the *wh*-phrase *nani* to undergo LD scrambling and land at a sentence-initial position, degrading the embedded scope reading. For this reason, the reading of (4-q), an embedded *whether*-clause within a *wh*-question, is more salient. Although other readings are grammatically valid, they are less intuitive and often harder to get.

Beyond this idiolect superficiality, the deeper reason for each preferred reading for (4-a) and (4-b) is that the *wh*-phrase must take scope of its closest Q-marker at LF, which must be understood in the c-command relationship. This can be better explained through the following simplified syntactic trees (Figure 1):

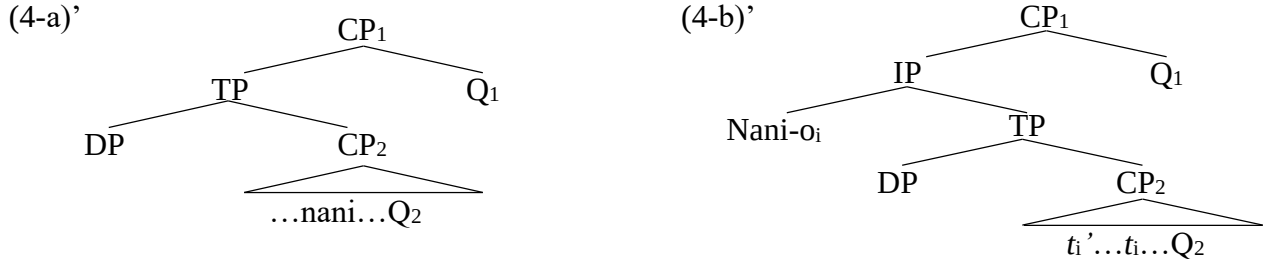

Fig. 1 Simplified syntactic trees for (4-a) and (4-b)

In (4-a)', *nani* takes scope from Q_2, making Q_1 a complementizer marking the sentence as a yes/no question. However, in (4-b)', the Q-marker that assigns scope to *nani* is Q_1, marking the sentence as a *wh*-question. This LD scrambling locates the *wh*-phrase outside its original c-command domain of the lower Q_2. Therefore, Q_2 *ka* is construed as *kadooka* ('whether') when it does not assign scope. This explains why (4-p) and (4-q) are two more salient readings for (4-a) and (4-b).

Outside of traditional bounds, the grammaticality of both readings in (4-a) and (4-b) suggests that a processing technique is probably the cause of their differences. It is expected that the Japanese parser would create a preliminary syntactic representation in real time, even before all the lexical information related to the structure is ready. This means that when a *wh*-phrase is LD preposed, it is

considered co-indexing itself and the complementizer of the sentence in which it appears. Such co-indexation occurs the instant the *wh*-phrase is contacted and must be fed into the existing sentence structure. In (4-a), the parser posits that the *wh*-phrase is positioned in CP_2 and is linked with Q_2 in the same embedded clause, thus marking it as a yes/no question immediately; so is the situation in (4-b) where the *wh*-phrase and Q_1 are in CP_1, marking it as a *wh*-question. Despite being grammatical, the original structure must be reconstructed because the matrix scope reading for (4-a) and the embedded scope reading for (4-b) contradict with the parser's initial structure assignment. It becomes more difficult to obtain these alternative readings since achieving them increases processing difficulty.

Overall, this section examines how LD *wh*-scrambling in interrogatives leads to different sentential interpretations, shedding light on the behavior of the Q-marker *ka* and the processing strategy.

4. Conclusion

This paper demonstrates that LD *wh*-scrambling in Japanese is considered a semantically vacuous movement in declarative sentences in that the *wh*-phrases undergo radical reconstruction back to their original positions for interpretation, as if the scrambling never occurred; it also differs from *wh*-movement because it need not feed into semantic interpretations. However, LD *wh*-scrambling in interrogative phrases can result in equivocal readings, with just one reading being highly preferred over its canonical version. This choice results from two factors: (1) the *wh*-phrase at LF must draw scope from its nearest Q-marker *ka*, understood through the c-command relationship; and (2) the processing strategy's initial structure assignment highlights only one reading. This paper focuses on two types of sentences—declaratives and interrogatives. Other sentence types, such as imperatives, still need to be explored. Additionally, the semantic effects of *wh*-scrambling with elements like quantifiers (e.g. *dare-ka*, 'someone'; *dare-mo*, 'everyone') and negative polarity items (e.g. *sika*, 'only', used with negation) in different sentence types require further investigation.

References

[1] Ross, J. R. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.

[2] Yoshimura, N. 2017. Scrambling. In M. Shibatani, S. Miyagawa & H. Noda (Eds.), *Handbook of Japanese Syntax*, 807-844. Berlin, Boston: De Gruyter Mouton.

[3] Ishihara, S., & van de Weijer, J. 2024. Effects of word order and embedded clause boundary on intonation in Tokyo Japanese. *Speech Prosody, 246*, 1220-1224.

[4] Saito, M. 2009. Optional A-scrambling. *Japanese/Korean Linguistics, 16*, 44-63.

[5] Nishigauchi, T. 2017. Quantification and *wh*-Constructions. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics*, 269-296. Oxford: Blackwell Publishers Ltd.

[6] Saito, M. 1989. Scrambling as Semantically Vacuous A'-movement. In M. R. Baltin and A. S. Kroch (Eds.), *Alternative Concepts of Phrase Structure*, 182-200. Chicago: The University of Chicago Press.

[7] Tsutsumi, H. 2022. *Scope Economy and Scope Reconstruction*. Sendai: Tohoku University dissertation.

[8] Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

[9] Brody, M. 1998. Review Article: The Minimalist Program and a Perfect Syntax: A Critical Notice of Noam Chomsky's The Minimalist Program. *Mind & Language, 13*(2), 205-214.

[10] Takahashi, D. 1993. Movement of *wh*-phrases in Japanese. *Natural Language & Linguistic Theory, 11*, 655-678.

[11] Abe, J. 2024. How to apply multiple scrambling in Japanese. *Syntax (Oxford, England), 2024*, n. p.

[12] Sauerland, U. 1996. The interpretability of scrambling. *Proceedings of Formal Approaches to Japanese Linguistics, 2*, 213-234.

[13] Sabel, J. 2001. *Wh*-questions in Japanese: Scrambling, reconstruction, and *wh*-movement. *Linguistic analysis, 31*(1-2), 1-31.

[14] Shimamura, K., & Tanaka, H. 2023. Eliminating Scrambling from Japanese via Neo-Davidsonian Event Semantics. *Proceedings of WAFL, 16*, n. p.