Scrambling and operator movement

Jun Abe *

Department of English, Tohoku Gakuin University, 1-3-1 Tsuchitoi, Aoba-ku, Sendai-shi, 980-8511, Japan

1. Introduction

This paper aims to demonstrate that what appears to be an instance of overt operator movement in Japanese in fact involves the operation of scrambling. It is shown that such an instance of overt operator movement shares with scrambling the property that it is insensitive to the wh-island. On the basis of this fact, I argue that such movement is analyzed as resulting from scrambling of the target phrase and then assigning the feature [Focus] to it, a feature that establishes an operator-variable chain. Furthermore, I argue that the present hypothesis is motivated by the phenomena of what Takano (2002) calls “surprising constituents,” whose presence in Japanese can be attributed to that of multiple scrambling. I suggest that the reason why scrambling is involved in an apparent instance of overt operator movement in Japanese comes from the defectiveness of functional categories, following Fukui (1986).

© 2011 Elsevier B.V. All rights reserved.

ARTICLE INFO

Article history:
Received 7 June 2011
Received in revised form 23 November 2011
Accepted 23 November 2011

Keywords:
Operator movement
Scrambling
Surprising constituent
Functional category

ABSTRACT

This paper aims to demonstrate that what appears to be an instance of overt operator movement in Japanese in fact involves the operation of scrambling. It is shown that such an instance of overt operator movement shares with scrambling the property that it is insensitive to the wh-island. On the basis of this fact, I argue that such movement is analyzed as resulting from scrambling of the target phrase and then assigning the feature [Focus] to it, a feature that establishes an operator-variable chain. Furthermore, I argue that the present hypothesis is motivated by the phenomena of what Takano (2002) calls “surprising constituents,” whose presence in Japanese can be attributed to that of multiple scrambling. I suggest that the reason why scrambling is involved in an apparent instance of overt operator movement in Japanese comes from the defectiveness of functional categories, following Fukui (1986).

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

This paper aims to demonstrate that scrambling is involved in what has been standardly recognized as “overt operator movement” in such a language as Japanese which allows free word order and hence has access to such an operation. Among those named overt operator movement is a type of movement in which its trigger will not be appropriately attributed to the inherent property of the head of a phrase to be moved. Such movement is illustrated below:

(1) a. [John], Mary likes very much.
   b. [The man you’re talking about], Mary likes very much.
   c. John met yesterday [the woman you had been talking about].

(1a) and (1b) are instances of Topicalization and (1c) an instance of Heavy NP Shift. In these cases, it is not obvious what is the locus of the trigger of the movement involved. Based upon this observation, Abe (2002) proposes that a relevant triggering feature is not carried by a particular head but rather assigned to a whole category that undergoes movement. Identifying the relevant feature as [Focus] and assuming that this feature is licensed in a “peripheral” position, ¹ Abe then puts forward the following hypothesis:

(2) The feature [Focus] can be assigned to a syntactic object during the derivation.

¹ See Abe (2002) for a precise definition of peripheral position. Even though the feature [Focus] is a trigger for overt operator movement, it cannot be an uninterpretable feature in the sense of Chomsky (1995) for an obvious reason: it contributes to interpretation at LF. Abe (2002) characterizes this feature as a semantic feature rather than a formal feature, just like the feature [Scope], which triggers so-called Quantifier Raising, and claims that the way of licensing it is different from that of checking formal features, especially in that the former does not involve matching between a triggering head and its target.
Furthermore, Abe (2002, 2005) claims that even those cases of overt operator movement whose triggers can be attributed to an inherent property of a head involve assignment of [Focus] according to (2), so that this enables us to capture the process of pied-piping. A typical case of such overt operator movement is \( \text{wh} \)-movement; compare the following examples:

(3)  
\[ \begin{align*}  
\text{a. Which book did you like that John read?} \\
\text{b. Which book that John read did you like?} 
\end{align*} \]

Both examples share the property that the triggering feature of the movement in question is the \( \text{wh} \)-feature borne by \textit{which}, but they differ in what size of phrases that include this feature is pied-piped. Abe (2002, 2005) captures this fact by claiming in effect that all instances of overt operator movement involve assignment of [Focus] in accordance with (2). Given this hypothesis, the difference in question between (3a) and (3b) is attributed to which phrase the feature [Focus] is assigned to: \textit{which book} in (3a) and \textit{which book that John read} in (3b).

Since such a language as English does not allow the option of scrambling, the Last Resort Principle dictates that overt operator movement involve the assignment of [Focus] to a phrase before it undergoes movement. On the other hand, if we assume, following Fukui (1993) among others, that scrambling is immune from the Last Resort Principle, such a language as Japanese that does allow scrambling gives rise to the possibility that [Focus] is assigned to a phrase after it undergoes trigger-free movement, i.e., scrambling. The following discussions aim to demonstrate that this is in fact the case. In Section 2, it will be shown that what appears to be an instance of overt operator movement in Japanese shares with scrambling the property that it is insensitive to the \( \text{wh} \)-island. In Section 3, I demonstrate, based upon the property of scrambling pointed out by Hoji (1985), according to which string-vacuous application of scrambling is prohibited, that scrambling cannot be involved in real instances of null operator movement. This is motivated by the fact that those instances exhibit \( \text{wh} \)-island sensitivity. In Section 4, the hypothesis that an apparent instance of overt operator movement involves scrambling is further motivated by the phenomena of what Takano (2002) calls “surprising constituents.” I will argue that the presence of these phenomena in Japanese is attributed to that of multiple scrambling. In Section 5, I will claim that there is no genuine instance of overt operator movement in Japanese, apparent instances of such movement all involving scrambling. I will attribute this fact to the defectiveness of functional categories of Japanese, following the Fukui’s (1986) line of parametrization of functional categories and agreement.

2. A similarity between scrambling and operator movement

A curious property of some Japanese constructions that are standardly claimed to involve overt operator movement is lack of \( \text{wh} \)-island effects, despite the fact that they show other typical island effects. Among these types of Japanese constructions is the Cleft Construction (CC). Hoji (1987) demonstrates that one type of the Japanese CC shows a general sensitivity to islands, comparable to the English CC, as illustrated below:

\[
\begin{align*}  
\text{a. [John-ga [Bill-ga Mary-ni t_i ageta to] omotteiru no]-wa [sono hon-o] i da.} \\
\text{Nom} \quad \text{Nom} \quad \text{Dat gave} \quad \text{Comp think} \quad \text{NL-Top that} \quad \text{book-Acc be} \\
\text{‘It is that book, that John thinks Bill gave} \ t_i \text{ to Mary.’} \\
\text{b. ?*[John-ga [Mary-ni t_i ageta] hito-o kiratteiru no]-wa [sono hon-o] i da.} \\
\text{Nom} \quad \text{Dat gave} \quad \text{person-Acc hate} \quad \text{NL-Top that} \quad \text{book-Acc be} \\
\text{‘It is that book, that John hates the person [who gave} \ t_i \text{ to Mary].’} \\
\text{c. ?*[John-ga [Mary-ga t_i syokuzi-ni sasotta node] rakutsanteiru no]-wa [sono hito-o] i da.} \\
\text{Nom} \quad \text{Nom dinner-to invited because be-depressed NL-Top that} \quad \text{person-Acc be} \\
\text{‘It is that person, that John is depressed [because Mary invited} \ t_i \text{ to dinner].’} 
\end{align*} \]

(4a) shows that the Japanese CC can involve long distance movement of a cleft phrase. (4b) and (4c) show that this construction is sensitive to the complex NP island and the adjunct island, respectively. Interestingly, this construction is different from the English counterpart in that it does not show \( \text{wh} \)-island effects or if any, the effects are very week, as illustrated below:

\[
\begin{align*}  
\text{?*[John-ga [Bill-ga dare-ni t_i ageta ka] siritagatteiru no]-wa [sono hon-o] i da.} \\
\text{Nom} \quad \text{Nom who-Dat gave Q want-to-know NL-Top that} \quad \text{book-Acc be} \\
\text{‘It is that book, that John wants to know who Bill gave} \ t_i \text{ to} t_j \text{.’} 
\end{align*} \]

The other type, which involves a Case-less cleft phrase, is island-insensitive, as Hoji (1987) shows. See Section 5 for relevant discussions. NL put for the morpheme -no in the glosses of (4) stands for a nominalizer, though I am not committed to the exact status of this morpheme.
This situation is reminiscent of the property of scrambling regarding island effects: as observed by Saito (1985), scrambling shows very weak effects of the wh-island, as illustrated in (7), although it does show other typical island effects, as illustrated in (6):

(6)  
a. ?*[Sono hon-o] [John-ga [Mary-ni t, ageta] hito-o kirateiru (koto)]  
that book-acc -nom -dat gave person-acc hate fact  
That book, John hates the person [who gave t, to Mary].

b. ?*[Sono hito-o] [John-ga [Mary-ga t, syokuzi-ni sasotta node] rakutansiteiru (koto)]  
that person-acc -nom dinner-to invited because be-depressed fact  
That person, John is depressed [because Mary invited t, to dinner].

(7) ?*[Sono hon-o] [John-ga [Bill-ga dare-ni t, ageta ka] siritagatteiru (koto)]  
that book-acc -nom who-dat gave Q want-to-know fact  
That book, John wants to know who Bill gave t, to t,

Such a similarity between scrambling and the Japanese CC with respect to lack of wh-island effects indicates that scrambling may be involved in this construction.

There are also many other constructions with which much the same situation emerges. First, let us consider the Japanese Sluicing Construction (SC), which is discussed by Takahashi (1994) in a way relevant here. He observes that this construction exhibits general sensitivity to islands, as illustrated below:

(8)  
John hates the person who gave something to Mary, but I don’t know what.

John is depressed because Mary invited someone to dinner, but I don’t know who.

Under the assumption made by Takahashi (1994) that the second conjuncts in (8) are derived by moving the wh-phrases to the embedded Specs of CP and then deleting the following TPs, which are identical to the corresponding parts of the first conjuncts, (8a) shows that the Japanese SC exhibits complex NP island effects and (8b) shows that it exhibits adjunct island effects. Although Takahashi claims that the Japanese SC shows general island sensitivity, including the wh-island, he himself seems to notice a difference in acceptability between complex NP and adjunct island effects on the one hand and wh-island effects on the other in the Japanese SC, as witnessed by the fact that he puts ?? to the examples which show the former type of island effects and ?? to those which show the latter type. I find that such a difference in acceptability is as significant as in cases of scrambling and the Japanese CC; a relevant example is provided below:

3 Tanaka (1999) makes much the same observation, providing the following examples, which show the contrast with respect to island-sensitivity of scrambling between the complex NP and wh-islands (the judgments are Tanaka’s):

(i)  
a. ?*LGB-o, Mary-ga [t, kaita] hito-o sagasiteiru.  
LGB-acc, Mary is looking for the person [who wrote t,].

Bill-dat, John met whether wanted-to-know

4 See the discussion around (22) for the evidence that the final chain produced in the Japanese CC must be an operator-variable chain, irrespective of whether scrambling is involved or not.

5 See Nishiyama et al. (1996) for the claim that what is taken as a case of the Japanese SC by Takahashi (1994) is in fact a species of the CC, a reduced one in which the presupposition part is elided. In such a case, this construction is subsumed under the Japanese CC. See footnote 12 for a relevant discussion.

6 Fukaya (2002) reports that many speakers accept such examples as (8a) that involve complex NP islands. I do not want to take issue over such a disagreement of judgments here (If such sluicing examples as those in (8) do not show island sensitivity, then I would say that this construction is simply irrelevant for the points I have been making in the text). But, Fukaya makes an important claim on such sluicing examples: the island insensitivity that seems to be exhibited by such examples is only apparent due to the indistinguishability of the local vs. non-local readings, and if such a factor is successfully controlled, then such examples do show island sensitivity. I am indebted to one of the reviewers for bringing Fukaya’s (2002) work to my attention.
Second, Simon (1989) and Rosen (1996) observe that the Japanese Right Dislocation Construction (RDC) shows general island sensitivity, as illustrated below:

(10) a. ?*John-wa [Mary-ni t_i ageta] hito-o kiratteiru, sono hon-o,.
   -Top -Dat gave person-Acc hate that book-Acc
   'John hates the person who gave t_i to Mary, that book.'

b. ?*John-wa [Mary-ga t_i syokuzi-ni sasotta node] rakutansiteiru, sono hito-o,.
   -Top -Nom dinner-to invited because be-depressed that person-Acc
   'John is depressed because Mary invited t_i to dinner, that person.'

Again, wh-island effects are very weak with this construction, as noted by Haraguchi (1973) and Kuno (1978); a relevant example is given below:

   -Top -Nom who-Dat gave Q want-to-know that book-Acc
   'John wants to know to whom i Bill gave t_i, that book.'

Third, Saito (1985) demonstrates that PP-Topicalization, unlike NP-Topicalization, exhibits island effects, as shown below⁷:

   -to-Top -Nom went fact-Nom have person-Acc found seem
   'To Pekingi, it seems that John found a person who has been t_i.'

   there-to-Top -Nom -Nom want-to-go though ignoring be seem
   'There, it seems that Mary is ignoring although John wants to go t_i.'

(Saito, 1985, pp. 332–333)

These sentences are fairly degraded, compared with those counterparts which involve NP-Topicalization, namely those which omit the postposition -ni ‘to’ from the preposed phrases. In contrast, PP-Topicalization exhibits very weak effects with respect to the wh-island, as illustrated below:

   -to-Top -Nom went fact-Nom have person-Acc found seem
   'To Pekingi, it seems that John found a person who has been t_i.'

b. *Hirosima-kara-wa i minna-ga [hito-ga oozee t_i kuru ka] gimon-ni
   -from-Top everyone-Nom people-Nom many come Q wondered
   omotteita.
   'From Hiroshimai, everyone wondered whether many people would come t_i.'

The above facts indicate that the constructions under consideration all may involve scrambling.⁸

Given the standard assumption that scrambling is an overt movement operation,⁹ it will be predicted under the present hypothesis that such weak effects of the wh-island as observed above will not be extended to those constructions that

---

⁷ The crucial morphological distinction between PP-Topicalization and NP-Topicalization that is relevant to the present discussion is whether or not a category that is marked with the topic marker -wa is accompanied with a postposition or a Case-marker. Thus, what PP-Topicalization means here is Topicalization of not only a PP but also a DP accompanied by a Case-marker.

⁸ See Saito (1985, 1987) for the claim that PP-Topicalization is a subcase of scrambling, and see Tanaka (2001) for independent pieces of evidence that the RDC involves scrambling. But see the discussions around (23) and footnote 12 that indicate that these constructions involve operator-variable chains as their final outputs.

⁹ As will be discussed later, given the single-cycle model of the computational component under the Minimalist Program (see, among others, Chomsky (2000, 2001, 2004)), overt movement simply means movement whose head is pronounced whereas covert movement means movement whose tail is pronounced.
involve covert operator movement. This is in fact borne out by sentences involving wh-adjuncts in situ, which show very clear island effects, as illustrated below:

(14)  a. ‘John-wa [Mary-ni naze sono hon-o ageta] hito-o kiratteiru no?  
     -Top -Dat why that book-Acc gave person-Acc hate Q  
     ‘Lit. John knows the person [who gave that book to Mary why]?’  
     b. ‘John-wa [Mary-ga naze sono hito-o syokuzi-ni sasotta node]  
        rakutansiteiru no?  
        -Top -Nom why that person-Acc dinner-to invited because be-depressed Q  
     ‘Lit. John is depressed [because Mary invited that person to dinner why]?’

These facts can be accounted for by the assumption, made by Abe (1993) and Tsai (1994), that adjunct wh-phrases in situ must move covertly to an appropriate Spec of CP to be licensed, unlike argument wh-phrases, which can be licensed by way of binding, hence not exhibiting island effects. As predicted, such a construction exhibits as strong wh-island effects as those observed in (14); a relevant example is given below:

(15)  John-wa [Bill-ga naze Mary-ni sono hon-o ageta ka] siritagatteiru no?  
      -Top -Nom why -Dat that book-Acc gave Q want-to-know Q  
     ‘Lit. John wants to know [Q Bill gave Mary that book why]?’

This sentence is interpreted only as a yes or no question, naze ‘why’ taking scope over the embedded clause. This fact is straightforwardly captured by the assumption that covert movement of naze is sensitive to the wh-island, hence naze being unable to reach the matrix Spec-CP.

Another construction that is claimed to involve covert movement is that involving a phrase with a focus particle (FP). Aoyagi (1994) observes that Japanese FPs such as -sae ‘even’ and -mo ‘also’, when attached to NPs, show ambiguity with respect to its scope, as illustrated below (I modified Aoyagi’s original example slightly):

(16)  John-ga [Mary-ga oisii ringo-mo/-sae tabeta to] omotte-inai (koto)  
      -Nom -Nom tasty apple-also/-even ate Comp think-not fact  
     ‘John does not think that Mary ate also/even a tasty apple.’  
     i.  [embedded scope] John does not think that Mary ate A TASTY APPLE in addition to some other things.  
     ii.  [matrix scope] Even for A TASTY APPLE, John does not have an idea that Mary ate it (in addition to some other idea about some other things).

Aoyagi claims that such a different scope can be naturally captured by assuming that a phrase with an FP undergoes covert movement. Thus, (16) can have the following two LF representations:

(17)  a. John-ga [CP [TP [oisii ringo-mo/-sae], [TP Mary-ga t, tabeta]] to] omotte-inai  
     b. [TP [oisii ringo-mo/-sae], [TP John-ga [CP [TP Mary-ga t, tabeta ] ] to] omotte-inai]  

(17a) represents the embedded scope of the FP phrase and (17b) its matrix scope. It is predicted under this analysis that FP phrases may not extend their scope across islands. This is in fact borne out, as observed by Aoyagi (1994):

(18)  Mary-ga [e, gakubusee-zidai-ni Barriers-mo/-sae yonda hito-ni atta.  
      -Nom undergraduate-time-at -also/-even read person-Dat met  
     ‘Mary met a person who read also/even Barriers when he/she was in undergraduate school.’

This sentence, unlike (16), is unambiguous: it has only the reading where Barriers-mo/sae ‘even Barriers’ takes scope within the relative clause, and crucially it does not have the reading where this phrase takes scope over the matrix clause. This is straightforwardly accounted for by assuming that Japanese FP phrases undergo covert movement, thereby being sensitive to islands. Interestingly, such a phrase also exhibits wh-island effects, as shown below:

(19)  John-ga [Mary-ga itu wariatarera ringo-mo/-sae tabeta ka] wakara-nai (koto)  
      -Nom -Nom when allotted apple-also/-even ate Q know-not fact  
     ‘John does not know when Mary ate also/even an allotted apple.’
This sentence has only the reading where *wariaterareta ringo-mo/sae* ‘also/even an allotted apple’ takes scope over the embedded clause.\(^{10}\) This fact will follow as a result of the *wh*-island effect under the covert movement approach, since in this case, the option of relying on scrambling is unavailable.

We have argued that unlike covert movement, overt operator movement may involve scrambling, by showing that a variety of constructions that are claimed to involve overt operator movement are insensitive to the *wh*-island. Abe (1993) claims that the fact that scrambling does not show *wh*-island effects is attributed to its nature as a “costless” movement in that it does not require any triggering feature. Abe follows Chomsky and Lasnik (1993) in assuming that *wh*-island effects are derived by an economy condition such as Minimize Chain Links (MCL), which dictates that an application of movement cannot skip its possible landing site. Given that a “possible landing site” is defined in terms of triggering features, as follows:

\[(10)\text{ The possible landing sites for movement of } \alpha \text{ are possible checking positions of the feature(s) borne by } \alpha.\]

it follows that scrambling is insensitive to the MCL because it has no triggering feature. This explains why scrambling does not exhibit *wh*-island effects.\(^{11}\)

Further, if we assume (2), repeated here:

\[(21)\text{ The feature } [\text{Focus}] \text{ can be assigned to a syntactic object during the derivation.}\]

then we can account for the insensitivity to the *wh*-island of those Japanese constructions considered above by claiming that the “overt operator movement” involved is derived by first applying scrambling to a given syntactic object and then assigning a [Focus] feature to this syntactic object.

Notice that assignment of a [Focus] feature in those constructions under consideration is necessary since the chains created by scrambling must be finally interpreted at LF as operator-variable chains, unlike those “semantically vacuous” chains in the sense of Saito (1989). This is most clearly shown by the contrast between the Japanese CC and a genuine instance of scrambling regarding the possibility of so-called radical reconstruction, as one reviewer points out. Consider the following examples, kindly provided by the reviewer:

\[(22)\]

\[a. \text{ Dono hon-o, Taro-ga } [\text{oootoo-ga tosyokan-kara } t_i \text{ karita ka}] \text{ siritagatteiru (koto)}\]  
which book-Acc -Nom brother-Nom library-from checked out Q want-to-know fact

‘Lit. Which book, Taro wants to know [Q his brother checked out } t_i \text{ from the library.’}.

\[b. \text{ *}[\text{Taro-ga } [\text{oootoo-ga tosyokan-kara } t_i \text{ karita ka}] \text{ siritagatteiru no]-wa dono hon-o, } \text{ da.}\]  
-Nom brother-Nom library-from checked out Q want-to-know NL-Top which book-Acc be

‘Lit. It is which book, that Taro wants to know [Q his brother checked out } t_i \text{ from the library.’}.

According to Saito (1989), (22a) is grammatical since the scrambled phrase *dono hon-o* ‘which book-Acc’ can be moved back to its original position, hence properly located within the *wh*-scope. Then, the ungrammaticality of (22b) indicates that although the movement involved in this construction may be scrambling, the created chain cannot be undone because the cleft *wh*-phrase is assigned the feature [Focus], hence the *wh*-phrase being improperly located outside the *wh*-scope. The same point can be made with PP-topicalization. Compare the following examples:

\[(23)\]

\[a. \text{ Dono hito-ni-wa, Taro-ga } \text{ tosyokan-de } t_i \text{ atta no?}\]  
which person-Dat-Top -Nom library-in saw Q

‘Which person, did Taro see } t_i \text{ in the library?’

\[b. \text{ ‘Dono hito-ni-wa, Taro-ga } [\text{oootoo-ga tosyokan-de } t_i \text{ atta ka}] \text{ siritagatteiru.}\]  
which person-Dat-Top -Nom brother-Nom library-in saw Q want-to-know

‘Lit. Which person, Taro wants to know [Q his brother saw } t_i \text{ in the library.’}.

The grammaticality of (23a) shows that PP-topicalization can get along with a *wh*-phrase, though in this case, the topicalized PP is forced to have a contrastive reading. Then, the ungrammaticality of (23b) suggests that the topicalized *wh*-phrase

---

\(^{10}\) Compare (19) with the following example, in which *wariaterareta ringo-mo/sae* ‘also/even an allotted apple’ is scrambled to the sentence initial position:

\[(i) \text{ [Wariaterareta ringo-mo/sae], John-ga [Mary-ga itu } t_i \text{ tabeta ka] wakara-nai (koto)}\]  
allotted apple-also-even -Nom -Nom when ate Q know-not fact

‘[Also/Even an allotted apple], John does not know when Mary ate } t_i.’

This sentence clearly has the reading where *wariaterareta ringo-mo/sae* takes matrix scope.

\(^{11}\) The fact that scrambling does show other island effects, then, indicates that such effects have nothing to do with economy conditions that regulate movement triggered by feature-checking, but rather that such effects are derived by purely configurational terms such as barriers exploited by Chomsky (1986).
dono hito-ni-wa ‘which person-Dat-Top’ cannot be moved back to its original position. This follows on the assumption that a topicalized PP is assigned the feature [Focus] after it is scrambled, so that it produces an operator-variable chain. 12

3. Implications for null operator movement

In this section, I discuss some empirical consequences of the present analysis with respect to how to analyze null operator movement. In particular, I argue that the present analysis provides a means to identify whether a given instance of null operator movement is “overt” or “covert.”

The CC has been standardly analyzed as involving null operator movement. The fact that the Japanese CC does not show wh-island effects, as illustrated in (5) above, suggests that it involves “overt” movement, namely scrambling. Interestingly, there is another construction in Japanese which is standardly assumed to involve null operator movement but which does show wh-island effects; namely, the Comparative Deletion Construction (CDC). Kikuchi (1989) demonstrates that this construction exhibits island sensitivity, including the wh-island, as illustrated below:

   -Top -Nom read person-Dat met than more book-Acc read
   ‘John read more books than Paul met a person who read e.’
   b. *[Minna-ga Paul-ga e yonda atode sanpo-ni dekaketa] yorimo John-wa takusan hon-o yonde-ita.
   everyone-Nom -Nom read after walking-to went-out than -Top more book-Acc had-read
   ‘John had read more books than everyone went for a walk after Paul read e.’

   everyone-Nom why -Nom read Q wondered than -Top more book-Acc had-read
   ‘John had read more books than everyone wondered why Paul read e.’

(Kikuchi, 1989, 12–13)

Given the present analysis, these facts indicate that the Japanese CDC does not involve scrambling; we may, then, characterize this construction as involving “covert” null operator movement.

As far as I can determine, there are at least two more constructions that show sensitivity to islands, including the wh-island. One is concerned with the type of the Japanese Tough-Construction (TC) that involves PP subjects, as illustrated below:

(26) a. John-kara, ga [hon-o e i nusumi] yasui (koto)
   -from-Nom book-Acc steal easy fact
   ‘Lit. From Johni is easy to steal a book e.’

12 I am indebted to one of the reviewers for pointing out the importance of providing such examples as in (22) and (23) to guarantee that the constructions in question involve operator-variable chains though these chains may be created via scrambling. I have shown such examples for only the CC and the PP-Topicalization construction in the text, and there are two more, namely, the SC and the RDC, that are taken to have the same derivation. With these constructions, it is not easy to demonstrate that they really involve operator-variable chains, and yet they show some properties that may suggest that they do. As noted in footnote 5, Nishiyama et al. (1996) demonstrate rather convincingly that what is taken as a case of the Japanese SC by Takahashi (1994) is in fact a species of the CC. If this is the case, then the Japanese SC should also involve an operator-variable chain as a special case of the CC. As for the RDC, Saito (1985) argues that this construction should not be derived simply from right scrambling, on the basis of the following properties: First, it is strictly a root phenomenon, and hence the chain created by a right-dislocated phrase cannot be identified as one created by scrambling, which can take place in an embedded clause. Second, unlike scrambling, subject can be right-dislocated in this construction, as shown below:

(i) Bill-ga [t Mary-ni sono hon-o watasita to] ittan da, John-ga.
   -Nom -Dat that book-Acc handed Comp said -Nom
   ‘Bill said that t, had handed that book to Mary, John.’
   -Nom -Nom -Dat that book-Acc handed Comp said
   ‘Johni, Bill said that t, had handed that book to Mary.’

Third, a resumptive pronoun can be inserted into the trace position of a right-dislocated phrase, unlike scrambling, as shown below:

(ii) John-wa kanozyo-o nagutta, Mary-o.
   -Top her -Acc hit -Acc
   ‘John hit her, Mary.’
   -Acc -Top her -Acc hit
   ‘Maryi, John hit her.’

It will not be unreasonable to attribute these differences to the fact that the RDC involves an operator-variable chain, but I must leave further investigation of this matter for future research. As far as I can see, such a pursuit will not be incompatible with Tanaka’s (1999) scrambling approach to the RDC, given the present mechanism of producing an operator-variable chain.
b. John-kara i-ga [Mary-ga hon-o e₁ nusum to kitaisi] yasui (koto)
   -Nom -Nom book-Acc steal Comp expect easy fact
   ‘Lit. From John, it is easy to expect that Mary will steal a book e₁.’

Unlike English, the Japanese TC allows its subject to be a PP that is associated with an argument within the complement clause of such a main predicate as yasui ‘easy’, as illustrated in (26a,b). (26b) shows that such an association can be “long distance” in the sense that the argument position that is associated with the PP subject is within the embedded clause of kitaisi ‘expect’, which heads the complement clause of yasui. Given the standard assumption that empty pronominals, which are available to Japanese, are NPs (see Saito, 1985, among others), it is reasonable to assume that in this construction, the device that establishes the association in question involves a null operator whose categorial status is PP. Given this, (26a) will have the following representation:

(27) John-kara_i-ga [CP OP_i [TP PROarb hon-o t_i nusumi]] yasui

It will then be predicted that such null operator movement shows sensitivity to islands. Takezawa (1987) observes that this is in fact the case, as illustrated below ((28a), a case of the relative clause island, is taken from his original example):

   such financial agency-from-Nom always money-Acc a lot borrow person-Acc trust hard
   ‘Lit. From such a financial agency, it is difficult to trust a person who always loans a lot of money e_i.’
   (Takezawa, 1987, p. 216)

b. *John-kara_i-ga [[Mary-ga hon-o e₁ nusunda node] iiwakesi] yasui (koto)
   -Nom -Nom book-Acc stole because excuse easy fact
   ‘Lit. From John, it is easy to make an excuse because Mary stole a book e₁.’

This construction also shows sensitivity to the wh-island, as illustrated below:

(29) *John-kara_i-ga [[Mary-ga dooyatte e₁ hon-o nusunda ka] wakari] yasui (koto)
   -Nom -Nom how book-Acc stole Q understand easy fact
   ‘Lit. From John, it is easy to understand how Mary stole a book e₁.’

This fact indicates that the null operator movement that is operative in the Japanese TC is “covert” in the sense that it does not involve scrambling.

Another construction that seems to involve covert null operator movement is relativization of time and place expressions, as illustrated below:

(30) a. [Mary-ga Bill-to Tokyo-de atta] zikan
   -Nom -with -in met time
   ‘the time when Mary met Bill in Tokyo’

b. [Mary-ga Bill-to kinoo atta] basyo
   -Nom -with yesterday met place
   ‘the place where Mary met Bill yesterday’

This construction allows long distance dependency between the relative head expressing time or place and its modifying event, as shown below:

(31) a. [John-ga [Mary-ga Bill-to Tokyo-de atta to] omotta] zikan
   -Nom -Nom -with -in met Comp thought time
   ‘the time when John thinks that Mary met Bill in Tokyo’

b. [John-ga [Mary-ga Bill-to kinoo atta to] omotta] basyo
   -Nom -Nom -with yesterday met Comp thought place
   ‘the place where John thinks that Mary met Bill yesterday’

In these NPs, the relative heads expressing time or place can be associated with the events expressed by the embedded clauses as well as the matrix clauses. Given that time and place expressions typically do not have the categorial status of NP, it is again reasonable to assume that this construction involves null operator movement within the relative clause. This is in fact supported by the fact that this construction is sensitive to islands, including the wh-island, as illustrated below13:

13 Murasugi and Saito (1992) argue that empty pronominals can be of the category PP as long as they are arguments, contrary to the assumption made in the text. They base their argument upon their factual claim, on which I must disagree, that such sentences as given in (32)–(34) are all acceptable.
(32)  a. *[John-ga [Mary-ga e, Tokyo-de atta] hito-o sitteiru] zikan_i
   -Nom -Nom -in met person-Acc know time
   ‘the time, when John knows the person [who Mary met in Tokyo e,]’

    b. *[John-ga [Mary-ga e, kinoo atta] hito-o sitteiru] basyo_i
   -Nom -Nom yesterday met person-Acc know place
   ‘the place, where John knows the person [who Mary met yesterday e,]’

(33)  a. *[John-ga [Mary-ga e, Bill-to Tokyo-de atta node] hara-o tateteiru] zikan_i
   -Nom -Nom -with -in met because be-angry time
   ‘the time, when John is angry [because Mary met Bill in Tokyo e,]’

    b. *[John-ga [Mary-ga e, Bill-to kinoo atta node] hara-o tateteiru] basyo_i
   -Nom -Nom -with yesterday met because be-angry place
   ‘the place, where John is angry [because Mary met Bill yesterday e,]’

(34)  a. *[John-ga [Mary-ga e, Bill-to doko-de atta ka] sitteiru] zikan_i
   -Nom -Nom -with where met Q know time
   ‘the time, when John knows [where Mary met Bill e,]’

    b. *[John-ga [Mary-ga e, Bill-to itu atta ka] sitteiru] basyo_i
   -Nom -Nom -with when met Q know place
   ‘the place, where John knows [when Mary met Bill e,]’

Given the above facts, a question immediately arises as to why null operator movement is “overt” in some cases and “covert” in
others, which amounts, in the present context, to the question why scrambling is involved in some instances of null operator
movement and not in others. First of all, given the single-cycle model of the computational component under the Minimalist
Program (see, among others, Chomsky, 2000, 2001, 2004), it will not be a meaningful question whether null operator movement
is overt or covert, since under this model, covert movement will be reinterpreted as a movement that does not carry a
phonological feature of a target phrase, unlike overt movement. It might be said that null operator movement is simply a special
case of “covert” movement in that its target phrase lacks a phonological feature. Given this conception, I claim that scrambling
cannot be involved in any instance of null operator movement. We can attribute the reason for this to the following condition on
scrambling posited by Hoji (1985) ((35) is slightly modified from Hoji’s original condition):

(35)  A scrambling operation cannot apply if it does not change the order of the overt lexical string.

The intuition behind this condition is clear; since the function of scrambling is mainly to change word order, its application is
illegitimate if the resulting structure gives rise to no change of word order. Given this condition, it is illegitimate to scramble
a null operator, since it would not change the order of the overt lexical string. From this condition, we can derive the fact that
the CDC, the TC with PP subjects, and relativization of time and place adverbials all show sensitivity to the
wh-island as well
as other islands.

Further, it also follows that the Japanese CC does not involve null operator movement; otherwise it would show
sensitivity to the wh-island. Thus, the present theory of movement lends support to an analysis of the CC in terms that do not
use null operators. There are at least two such analyses that come to mind. One is a PF deletion approach according to which
an overt phrase that corresponds to that in the focus position of the CC undergoes overt operator movement within the
presupposed clause and then gets deleted under strict identity to the phrase in the focus position; this approach is recently
adopted, for example, by Takano (2002). The other approach is to adopt Vergnaud’s (1974) analysis of relative clauses,
according to which this construction will involve overt movement of a phrase directly into its focus position.14 Either

14 If we adopt this option, we will probably have to assume a more sophisticated derivation, since to move a phrase within the presupposition clause
directly to the focus position appears to involve sideward movement. One possibility to resolve this potential problem is to assume that the Japanese CC
involves double application of operator movement, i.e., not only movement into the focus position but also remnant movement into the topic position, as is
proposed by Hasegawa (1997) and Hiraishi and Ishihara (2002). Thus, under this analysis, (36a), for instance, will have the following derivation:

(i)  a.  [e]-wa [Bill-ga Mary-ni sono hon-o ageta no] da
   \[scrambling of sono hon-o and assigning of [Focus] to it

    b.  [e]-wa [sono hon-o, [Bill-ga Mary-ni t, ageta no]] da
   \[Focus]
   \[remnant movement of the clause following sono-hon-o
   \[into [e]

    c.  [Bill-ga Mary-ni t, ageta no]-wa [sono hon-o, t,] da
   \[Focus]
approach seems to have independent plausibility when we consider the Japanese CC, since in this construction, a phrase that appears in the focus position accompanies a Case particle or a postposition, as illustrated below, so that strict identity holds between a phrase in the focus position and the corresponding missing phrase in the presupposition clause:

(36) a. [Bill-ga Mary-ni titi ageta no]-wa [sono hon-o]i da.
   -Nom -Dat gave NL-Top that book-Acc be
   'It is that book that Bill gave to Mary.'

   -Nom money-Acc got NL-Top that person-from be
   'It is from that person that Bill got money titi.'

This property also fits nicely with the present claim that scrambling is involved in the derivation of Japanese CC, since it also requires strict identity between the scrambled phrase and the corresponding missing phrase. This construction contrasts with real cases of null operator movement such as the CDC and relativization of time and place adverbials, which seem to require only referential or "content-wise" identity between the null operator and its antecedent. Hence, we are led to the conclusion that scrambling is involved only in those constructions in which strict identity holds between the displaced phrase and the corresponding missing phrase.15

However, this must be only a necessary condition, since it appears that strict identity is at stake in the case of the TC with PP subjects. This immediately raises the question why the TC with PP subjects involves null operator movement instead of scrambling. The most natural answer will be that direct movement of a phrase into the subject position in the TC via scrambling is barred due to a locality condition on A-movement or whatever condition that precludes so-called improper movement. Hence, the only way that makes this construction possible is to rely on null operator movement, as shown in (27), reproduced below:

(37) John-kara-ga [CP OPi [TP PROarb hon-o titi nusumi]] yasui

4. Multiple occurrences of focus phrases

In this section, I argue that the present analysis gives rise to a significant consequence with respect to what Takano (2002) calls "surprising constituents" in Japanese, those constituents that appear to consist of more than one phrase and nonetheless can appear in a position that requires a single constituent. It has been observed in the literature that in the CC, the SC and the RDC in Japanese, more than one phrase can appear in the focus positions of these constructions, as illustrated below:

15 One reviewer points out the possibility that the dichotomy in question may have nothing to do with the availability of scrambling, noting that a similar effect with respect to the sensitivity to such a weak island as the wh-island is observed even in non-scrambling languages. The reviewer provides the following examples ((ia) is taken from Corver (2006)):

(i) a. *John bought more oranges than Sue wondered whether to buy?
   b. Which orange did Sue wonder whether to buy?

I have implicitly assumed that in such a non-scrambling language as English, an operator movement is basically sensitive to the wh-island. Although it is usually observed that non-tensed wh-clauses exhibit very weak island effects, as shown in (ib), tensed ones do show wh-island effects, so that the following example is much worse than (ib):

(ii) ?*Which orange did Sue wonder whether Mary had bought?

Notice that the Japanese examples that show insensitivity to the wh-island all involve tensed wh-clauses, so it is not unreasonable to assume that the dichotomy shown in (i) is not the same kind as that under consideration in the text. There is another dichotomy, extensively discussed by Cinque (1990), according to which extraction of PP out of a weak island is more difficult than extraction of NP, as shown below (the examples, in which (iiib) is taken from Szabolcsi (2006), are kindly provided by the reviewer):

(iii) a. *About which politician did John ask whether to worry?
   b. Which politician did John ask whether to worry about?

As the reviewer correctly notes, this NP vs. PP dichotomy is not relevant for the dichotomy under consideration in the text, either, since PP-Topicalization is not sensitive to the wh-island, unlike the TC involving subject PPs, for instance. Further, the time/place relativization is sensitive to the wh-island, whereas extraction of the place adverbial where in English does not seem to be, as shown below (the example is taken from den Dikken and Szabolcsi (1999)):

(iv) ?Where did John ask whether to put/read this book?

Thus, I think it safe to conclude that the dichotomy discussed in the text does have to do with the availability of scrambling.
'Lit. It was [an apple to Mary] that John gave.'

'Lit. John gave something to someone, but I don’t know [what to whom].'

John-ga kinoo ageta yo, [Mary-ni ringo-o].
'Lit. John gave yesterday, [an apple to Mary].'

Note that the English counterparts of these constructions are all ungrammatical, as shown below:\(^{16}\):

(41) a. "It was [an apple to Mary] that John gave.
       b. "John said someone bought something, and Mary wonders [who what].
       c. "John gave (it to her) yesterday, [the apple to Mary].

I argue in what follows that the possibility of surprising constituents in Japanese is attributed to the possibility of multiple scrambling in this language, which is illustrated below:\(^{17}\):

(42) Mary-ni ringo-o John-ga ageta (koto)
'Lit. Mary an apple, John gave.'

Then the option of surprising constituents is unavailable in such a language as English, since this language does not have the option of scrambling.

An initial indication that supports the present hypothesis comes from the fact that those constructions which allow surprising constituents correspond, to a significant extent, to those which are claimed above to involve scrambling.\(^{18}\) The CC, the SC and the RDC, illustrated in (38)–(40), exhibit such a correlation. The case of PP Topicalization, which is also claimed to involve scrambling, allows surprising constituents, as illustrated below:

(43) a. [Hon-o John-kara]-wa daremo [Mary-ga nusunda to] kitaishite-inai.
       book-Acc -from-Top anybody -Nom stole Comp expect-not
       'Lit. A book from John, nobody expects that Mary stole.'

       b. [Sono mondai-o Mary-to]-wa daremo [Bill-ga hanasita to] kitaishite-inai.
       that matter -Acc -with-Top anybody -Nom talked Comp expect-not
       'Lit. That matter with Mary, nobody expects that Bill talked about.'

Note that the same array of word strings can be derived by scrambling hon-o ‘book-Acc’ in (43a) and sono mondai-o ‘that matter-Acc’ in (43b) to the top of the sentences, with the topic marker -wa only modifying John-kara ‘John-from’ and Mary-to

\(^{16}\) Takahashi (1994) claims that “[i]n English, one and only one wh-phrase appears as a remnant of Sluicing,” providing such an example as (41b). However, Nishigauchi (1998) notes that there are some instances that seem to instantiate multiple Sluicing in English, providing the following example, taken from Bolinger (1978):

(i) I know that in each instance one of the girls got something from one of the boys.
       But which from which? (Bolinger, 1978, p. 109)

I will not commit myself to the status of multiple Sluicing in English here, simply noting that this issue does not directly affect the argument given in the text.

\(^{17}\) In this section, I will defend Koizumi’s (2000) remnant VP movement approach to multiple scrambling. Hence, when I mention “multiple scrambling”, it just refers to a phenomenon rather than an analysis.

\(^{18}\) The only case that I found to be an apparent counterexample to this correlation is the Japanese CDC. Kennedy (2002) observes that this construction allows simultaneous comparison of more than one phrase in a sentence, as illustrated below:

(i) John-wa kimi-ga kawaseta yori(mo) motto ooku-no dansee-ni motto ooku-no kuruma-o kawaseta.
       Top you-Nom made-buy than more many-Gen man-Dat more many-Gen car-Acc made-buy
       ‘Lit. John made more men buy more cars than you made e bought e.’

It is highly likely that the two separate null operators involved in this sentence constitute surprising constituents, given the fact that in English, which does not allow such constituents in general, the counterpart of such a sentence as (i) is ungrammatical. I must leave a more detailed analysis of this property for future research.
‘Mary-with’. It seems, however, that the two cases are distinguished by their intonation patterns: in the intended cases, the two phrases constitute an intonational chunk so that no intonational break intervenes between them, whereas in the other cases, a clear intonational break is attested between the two. Taking this intonational property into consideration, the sentences in (43) can still be judged as acceptable. In contrast, let us consider relevant cases of the Japanese TC with PP subjects, which are illustrated below:

(44) a. ?* [Hon-o John-kara]-ga [[Mary-ga nusumun to] kitaishita] yasuis (koto)
    book-Acc -from-Nom -Nom steal Comp expect easy fact
    ‘Lit. A book, from John is easy to expect that Mary will steal.’

b. ?* [Sono mondai-o Mary-to]-ga [[Bill-ga hanasu to] kitaishita yasuis (koto)
    that matter -Acc -with-Nom -Nom talk Comp expect easy fact
    ‘Lit. That matter, with Mary is easy to expect that Bill will talk about.’

These sentences seem to be unacceptable when uttered with appropriate intonations; they are, in contrast, acceptable when uttered with an intonational break right after the first NPs, which suggests that these NPs are moved to the front of the sentences by scrambling. The unacceptability of (44a,b) will follow immediately under the present hypothesis, since this construction does not involve scrambling, which in turn excludes the possibility of making surprising constituents.

It is rather evident that what Takano calls “surprising constituent” makes a real constituent in the constructions observed above, since they are subject to the so-called clause-mate condition. Koizumi (2000) observes that where more than one constituent appears in the focus position of the Japanese CC, each phrase must come from the same clause, as illustrated in the ungrammaticality of the following example:

    -Nom -Nom bought Comp said NL-Top -Dat apple-Acc be
    ‘Lit. It was [Nancy, apples] that Mary told t[i] that John had bought t[j].’

Similarly, Takahashi (1994) observes that what he calls Japanese multiple Sluicing follows the requirement that the remnant wh-phrases in the sluice be clause-mates, as illustrated below:

(46) *John-ga dareka-ni [Bill-ga nanika-o tabeta to] itteita ga,
    -Nom someone-Dat -Nom something-Acc ate Comp said but
    Mary-wa [dare-ni nani-o ka] oboetei-nai.
    -Top who-Dat what-Acc Q remember-not
    ‘Lit. John told someone that Bill had eaten something, but Mary doesn’t remember [who(m) what].’

Furthermore, Cecchetto (1999) notes that the clause-mate condition is observed in Japanese multiple Right Dislocation, as illustrated below (the example is adapted from Cecchetto (1999, p. 67)):

    -Nom -Nom love Comp said -Dat that village-Acc
    ‘Lit. John told t[i] that Mary loved t[j], Bill that village.’

The same restriction applies to multiple PP topicalization, as shown below:

(48) *[Nancy-ni, John-kara]-wa Mary-ga t[i] [daremo hon-o t[j] nusumanakatta to] itta.
    -Dat -from -Top -Nom anyone book-Acc stole-not Comp said
    ‘Lit. Nancy, from John, Mary told t[i] that nobody stole a book t[j].’

There are mainly two approaches to explain this clause-mate condition as well as other properties of surprising constituents: one is the VP preposing approach advocated by Koizumi (1995, 2000) and the other is the oblique movement approach advocated by Takano (2002) for the CC, by Abe (1999) for the RDC, and by Takahashi (1994) for the SC. According to the former approach, a surprising constituent is derived by overt head movement of the V whose maximal projection contains the elements of this surprising constituent and then preposing of this remnant VP. Thus, (38), for instance, has the following structure:

(49) [TP John-ga [VP t] age+ta no]-wa [VP Mary-ni ringo-o t[j]] da

According to this analysis, the clause-mate condition observed above is attributed to the locality condition on head movement of the V that feeds remnant VP movement. Thus, in order to derive the sentence given in (45), for instance, the
embedded V *katta* ‘bought’ must be raised at least to the matrix T, so that the remnant VP to be preposed can contain both the embedded object *ringo-o* ‘apple-Acc’ and the matrix object *Nancy-ni* ‘Nancy-Dat’. Koizumi (1995, 2000) claims that such a head movement is prohibited since it crosses a tensed clause boundary.

According to the oblique movement approach, on the other hand, a surprising constituent is derived by one element being adjoined to the other, a process that Takano (2002) names oblique movement. Thus, (38), for instance, has the following structure:

(50) \[ \text{TP John-ga} \ t_i \ t_j \ \text{ageta no]-wa} \ \text{DP Mary-ni} \ \text{DP ringo-o]} \ da \]

According to this analysis, the clause-mate condition in question is attributed to whatever locality condition applies to oblique movement.\(^\text{19}\)

Either approach may be consistent with the present hypothesis that surprising constituents are created by scrambling, as long as the processes proposed by each approach involve scrambling. Takano (2002) in fact suggests that oblique movement may be a special instance of scrambling. This is a natural claim in that oblique movement does not seem to be triggered by any morphological feature. On the other hand, Koizumi (1995, 2000) does not clearly identify the nature of the remnant VP movement that is involved in deriving surprising constituents, but simply regards such an operation as an available option of UG. As Takano (2002) points out, however, this leaves open the question why the surprising constituents observed in Japanese do not equally constitute grammatical forms in such a language as English, as exemplified in (41). This problem is immediately resolved if it is established that the remnant VP movement in question is in fact an instance of scrambling. Assuming that this is in fact the case for the moment, I provide a couple of pieces of evidence to show that Koizumi’s (1995, 2000) remnant VP movement analysis is superior to Takano’s (2002) oblique movement analysis.

In so doing, let us turn to the question how instances of multiple scrambling such as the one illustrated in (42), repeated below, are derived.

(51) Mary-ni ringo-o John-ga ageta (koto)
    \[-Dat apple-Acc -Nom gave fact\]
    \[\text{Lit. Mary an apple, John gave.}\]

Koizumi (2000) provides several pieces of evidence for the claim that instances of multiple scrambling can involve remnant VP movement. While his claim is based exclusively upon those instances which involve long distance multiple scrambling, I provide a couple of pieces of evidence for this claim that involve instances of clause-internal multiple scrambling such as (51). Thus, I demonstrate that this sentence has the following structure:

(52) \[ \text{TP} \ [\text{VP Mary-ni ringo-o} \ t_i] \ [\text{TP John-ga} \ [\text{VP} \ t] \ \text{age\+ta}]] \]

The evidence has to do with scope interaction phenomena in Japanese. There is a well-known fact about scope interaction among QPs, first noted by Kuroda (1971), according to which the structural height among QPs directly reflects the scope relation among them; this is named the rigidity condition. Japanese is one of the well-attested languages that observe this condition. Let us consider the following examples\(^\text{20}\):

(53) a. Sannin-no otoko-ga hutari-no onna-o hihansita.
    \[\text{three-Gen man -Nom two-Gen woman-Acc criticized}\]
    \[\text{‘Three men criticized two women.’}\]

b. Hutari-no onna-o sannin-no otoko-ga t_i hihansita.
    \[\text{two-Gen woman-Acc three-Gen man -Nom criticized}\]
    \[\text{‘Two women, three men criticized.’}\]

What these sentences show is that when the surface order reflects the basic word order, as in (53a), it observes the rigidity condition; that is, the structurally higher QP must take scope over the lower one. On the other hand, when the surface order is

\(^{19}\) See Saito (1994) and Sohn (1994) for the exact formulation of the locality condition on oblique movement. For relevant discussions, see Takahashi (1994), Abe (1999) and Takano (2002).

\(^{20}\) Hoji (2003) stresses the importance of controlling relevant examples in order to obtain the generalization of scope rigidity, namely, the importance of what he calls repeatability of judgments. Hoji notes that this generalization appears to be falsified by counterexamples, but that by choosing the right kind of QPs involved in scope interaction, the generalization still holds true. One factor is concerned with whether a QP refers to a specific group of entities. Thus, Hoji provides the following example as an apparent counterexample to the generalization in question:

(i) Dareka-ga uti-no subete-no sensyu-o bikoositeiru.
    \[\text{Someone-Nom our-Gen every-Gen athlete-Acc is-shadowing}\]
    \[\text{‘Someone is shadowing every athlete of ours.’}\]

As Hoji notes, the inverse scope order is relatively accessible with this example. That is why I basically used indefinite numeral NPs in the text to make relevant points. I am indebted to one of the reviewers for bringing my attention to the importance of constructing the right kind of QP examples.
a derived word order, as in (53b), either QP can take scope over the other. This shows that a QP that undergoes scrambling, crossing another QP, makes the scope order ambiguous. In order to capture these facts, let us first assume, following Abe (1993, 2002), that a QP carries a [Scope] feature and that when a QP undergoes movement, this feature can be carried along or left in situ. Let us then assume the following interpretive rule:

(54) If [Scope] \(i\) c-commands [Scope] \(j\), then [Scope] \(i\) \(\geq\) [Scope] \(j\).

Given these assumptions, (53a,b) can have the following representations:

(55) Sannin-no otoko-ga hutari-no onna-o hihansita.

(56) a. Hutari-no onna-o sannin-no otoko-ga \(t_i\) hihansita.

b. Hutari-no onna-o sannin-no otoko-ga \(t_i\) hihansita.

(55) represents the reading of (53a) where the subject sannin-no otoko ‘three men’ takes scope over the object hutari-no onna ‘two women’, the only reading available to this sentence. (56a) is derived when the [Scope] feature of hutari-no onna ‘two women’ is carried along when it undergoes scrambling, representing the reading where it takes scope over sannin-no otoko ‘three men’. On the other hand, (56b) is derived when the [Scope] feature of hutari-no onna ‘two women’ is left in situ when it undergoes scrambling, representing the reading where it takes scope under sannin-no otoko ‘three men’.

Keeping this mechanism of scope interaction in mind, let us consider the following examples:

(57) a. John-ga hutari-no onna-ni sannin-no otoko-o syookaisita.

b. Hutari-no onna-ni sannin-no otoko-o John-ga \(t_i\) \(t_j\) syookaisita.

Hoji (1985) observes that (57a) has only the reading where the indirect object hutari-no onna ‘two women’ takes scope over the direct object sannin-no otoko ‘three men’, and argues, based on this scope fact, that the basic word order is S–IO–DO–V in Japanese. Interestingly, as observed by Yatsushiro (1996) and Nishida (1999), even if these indirect and direct objects are scrambled together, as in (57b), the resulting sentence still has only the reading where the IO takes scope over the DO. Suppose that literal multiple scrambling took place in deriving (57b). Then, at least the following two representations could be derived:

(58) a. \([\text{TP} \text{ hutari-no onna-ni} \ [\text{TP} \text{ sannin-no otoko-o} \ [\text{TP} \text{ John-ga} \ t_i \ t_j \text{ syookaisita}]]]\)

b. \([\text{TP} \text{ hutari-no onna-ni} \ [\text{TP} \text{ sannin-no otoko-o} \ [\text{TP} \text{ John-ga} \ t_i \ t_j \text{ syookaisita}]]]\)

(58a) is derived when the [Scope] features of both the IO and DO are carried along when these phrases undergo scrambling, representing the reading of 2 \(\geq\) 3, the one available to (57b). (58b) differs from (58a) in that when hutari-no onna-ni ‘two women’-Dat undergoes scrambling, it leaves behind its [Scope] feature; hence it represents the reading of 3 \(\geq\) 2, the scope opposite to that represented in (58a), which is lacking in sentence (57b). Thus we cannot obtain the right result if we apply literal multiple scrambling to derive (57b).

On the other hand, the remnant VP movement approach can capture the scope fact of (57b) correctly. According to this approach, this sentence has the following representation:

(59) \([\text{TP} \text{ [VP hutari-no onna-ni sannin-no otoko-o} \ t_v] \ [\text{TP} \text{ John-ga} \ t_i \ \text{ syookaisita}]]\)

This correctly represents the only reading available to (57b), namely that of 2 \(\geq\) 3.

Note that it is not immediately clear how the oblique movement approach advocated by Takano (2002) can correctly capture this scope fact. Given that the relative scope of QPs are determined by the notion of c-command, as formulated in (54), the IO and DO in (57b), which become a constituent thanks to oblique movement, will c-command each other, so that not only the reading of 2 \(\geq\) 3 but also the opposite reading should be obtained. Furthermore, if trace positions of QPs can count to determine the relative structural height among QPs, as assumed here in the form of leaving [Scope] in situ, again the two
readings should be obtained according to the oblique movement approach. Thus, such a scope fact as in (57b) lends strong support to the need of remnant VP movement.

Further evidence for the remnant VP movement approach is provided in the following array of examples:

(60) a. Sannin-no otoko-ga hutari-no onna-ni dareka-o syookaisita.
   three-Gen man-Nom two-Gen woman-Dat someone-Acc introduced
   ‘Three men introduced someone to two women.’

b. Hutari-no onna-ni sannin-no otoko-ga ti dareka-o syookaisita.
   two-Gen woman-Dat three-Gen man-Nom someone-Acc introduced

c. Hutari-no onna-ni sannin-no otoko-o ti dareka-ga ti syookaisita.
   two-Gen woman-Dat three-Gen man-Acc someone-Nom introduced
   ‘Lit. To two women, three men, someone introduced.’

Sentence (60a) reflects the basic word order of Japanese according to Hoji (1985) and, as expected, it has the reading where sannin-no otoko ‘three men’ takes scope wider than both hutari-no onna ‘two women’ and dareka ‘someone’ and between the latter two, hutari-no onna takes scope over dareka. Thus, this sentence can be uttered in the situation in which three men each did the act of introduction to different pairs of women and further each man introduced a different person to each woman. This reading will be represented as follows under the present assumptions:

(61) [TP sannin-no otoko-ga hutari-no onna-ni dareka-o syookaisita]
    [Scope] [Scope] [Scope]

According to the interpretive rule given in (54), (61) represents the reading of \(3 > 2 > \exists\). Now let us consider (60b), which is derived from (60a) by moving the indirect object hutari-no onna-ni ‘two women’ to the top of the sentence. This sentence has the reading derived from the scope order that reflects the hierarchical order of the indirect object, the subject and the direct object, namely the reading of \(2 > 3 > \exists\). Thus, this sentence can be uttered in the situation in which two women were each involved as recipients in the act of introduction by different groups of three men and further each man introduced a different person to each woman. This reading can be represented as follows under the present assumptions:

(62) [TP hutari-no onna-ni [TP sannin-no otoko-ga ti dareka-o syookaisita]]
    [Scope] [Scope] [Scope]

According to the interpretive rule given in (54), this representation correctly represents the reading in question. An interesting case is (60c), which involves multiple scrambling of indirect and direct objects. This sentence does not have the reading of scope order that is expected from the hierarchical order of the indirect object, the direct object and the subject, namely the reading of \(2 > 3 > \exists\). Thus, this sentence is infelicitous when uttered in the situation in which two women were each involved as recipients in the act of introduction of different groups of three men and further each man was introduced to each woman by a different person. This reading could be represented as follows if we derived (60c) by applying literal multiple scrambling:

(63) [TP hutari-no onna-ni [TP sannin-no otoko-o ti dareka-ga ti syookaisita]]
    [Scope] [Scope] [Scope]

Thus, the fact that (60c) does not have the reading represented in (63) shows clearly that we cannot obtain the right result if we apply literal multiple scrambling to derive this sentence.

Under the remnant VP movement approach, this sentence will have the following representation:

(64) [TP [VP hutari-no onna-ni sannin-no otoko-o ti dareka-ga ti syookaisita+ta]]
    [Scope] [Scope] [Scope]

Given that in this representation, hutari-no onna ‘two women’ asymmetrically c-commands sannin-no otoko ‘three men’ and no c-command relation exists between these QPs and dareka ‘someone’, this represents the reading where hutari-no onna takes scope over sannin-no otoko and no scope interaction exists between these two QPs and dareka. This reading amounts to that in which two women each acted as the recipients of introduction of different groups of three men by a particular person. In fact, sentence (60c) has this reading. Thus, the remnant VP movement approach brings us the right result regarding such a case of scope interaction. Note again that it is not immediately clear that the oblique movement approach can capture such a scope fact correctly, given that the two objects that become a constituent by way of oblique movement will c-command each other.

We have so far established the need of remnant VP movement to derive cases of multiple scrambling. Recall the present hypothesis that surprising constituents allowed in the Japanese CC, SC, RDC, and PP Topicalization are possible due to the availability of multiple scrambling. Given that “surprising constituents” are real constituents, as indicated by the
clause-mate requirement observed in (45)–(48), it is natural to reason that such surprising constituents are derived by remnant VP scrambling. It will then be predicted that the scope facts observed above should be carried over in cases of "surprising constituents" observed in the four constructions. I demonstrate that this is in fact the case. First, the following examples show that the prediction is borne out for the Japanese CC:

(65) [John-ga syookaisita no]-wa [hutari-no onna-ni sannin-no otoko-o] da.
    -Nom introduced NL-Top two-Gen woman-Dat three-Gen man-Acc be
    ‘Lit. It was [to two women, three men] that John introduced.’

(66) a. [Sannin-no otoko-ga dareka-o syookaisita no]-wa [hutari-no onna-ni] da.
    three-Gen man-Nom someone-Acc introduced NL-Top two-Gen woman-Dat be
    ‘It was to two women that three men introduced someone.’

b. [dareka-ga syookaisita no]-wa [hutari-no onna-ni sannin-no otoko-o] da.
    someone-Nom introduced NL-Top two-Gen woman-Dat three-Gen man-Acc be
    ‘Lit. It was [to two women, three men] that someone introduced.’

(65) has only the reading of $2 \rightarrow 3$, just like (57b), and while (66a) allows the reading of $2 \rightarrow 3 \rightarrow \exists$, just like (60b), (66b) does not have this reading, exactly like (60c). These scope facts immediately follow if we assume that the constituents in the focus positions of (65) and (66b) are derived by remnant VP scrambling, forming the following structure:

(67) [VP hutari-no onna-ni sannin-no otoko-o $\{\text{Scope}\} \{\text{Scope}\}$]

Likewise, the Japanese RDC shows exactly the same pattern of facts, as illustrated below:

(68) John-ga syookaisita yo, [hutari-no onna-ni sannin-no otoko-o].
    -Nom introduced two-Gen woman-Dat three-Gen man-Acc
    ‘Lit. John introduced, [to two women, three men].’

(69) a. Sannin-no otoko-ga dareka-o syookaisita yo, [hutari-no onna-ni].
    three-Gen man-Nom someone-Acc introduced two-Gen woman-Dat
    ‘Lit. Three men introduced someone, to two women.’

b. Dareka-ga syookaisita yo, [hutari-no onna-ni sannin-no otoko-o].
    someone-Nom introduced two-Gen woman-Dat three-Gen man-Acc
    ‘Lit. Someone introduced, [to two women, three men].’

Again, (68) has only the reading of $2 \rightarrow 3$, and while (69a) allows the reading of $2 \rightarrow 3 \rightarrow \exists$, (69b) does not have this reading. These scope facts immediately follow if the dislocated phrases of (68) and (69b) are derived by remnant VP scrambling, forming the structure given in (67). The following examples show that the prediction in question is borne out for PP topicalization:

(70) [Hutari-no onna-ni sannin-no otoko-o]-ba John-ga syookaisita.
    two-Gen woman-Dat three-Gen man-Acc-Top -Nom introduced
    ‘Lit. [To two women three men], John introduced.’

(71) a. [Hutari-no onna-ni]-wa sannin-no otoko-ga dareka-o syookaisita.
    two-Gen woman-Dat-Top three-Gen man-Nom someone-Acc introduced
    ‘To two women, three men introduced someone.’

b. [Hutari-no onna-ni sannin-no otoko-o]-ba dareka-ga syookaisita.
    two-Gen woman-Dat three-Gen man-Acc-Top someone-Nom introduced
    ‘Lit. [To two women three men], someone introduced.’

In (70) and (71b), the topic is marked by -ba, which is an archaic variant of the topic marker -wa. For an unknown reason, this archaic form seems to be the only choice when the topic marker is preceded by a NP ended with an accusative marker -o.
Again, (70) has only the reading of 2 > 3, and while (71a) allows the reading of 2 > 3 > Ξ, (71b) does not have this reading. These scope facts are immediately explained if the topicalized phrases of (70) and (71b) are derived by remnant VP scrambling.22

5. Mechanism of assignment of [Focus]

So far I have argued that surprising constituents that appear in focus positions of a variety of Japanese constructions are derived in the same way that multiple scrambling is derived, and that the derivations involve remnant VP movement, as advocated by Koizumi (1995, 2000). Thus, according to the present hypothesis, a case of the CC such as (38), for instance, will be derived in the following way:

\[
(72) \quad \begin{align*}
   & a. \quad \text{[TP John-ga [VP Mary-ni ringo-o age] ta] no]-wa [e] da} \\
      & \quad \downarrow \text{V-movement} \\
   & b. \quad \text{[TP John-ga [VP Mary-ni ringo-o \text{ t}_v] age+ta] no]-wa [e] da} \\
      & \quad \downarrow \text{remnant VP scrambling} \\
   & c. \quad \text{[TP John-ga [VP t] age+ta] no]-wa [VP Mary-ni ringo-o \text{ t}_v] da} \\
      & \quad \downarrow \text{assignment of [Focus]} \\
   & d. \quad \text{[TP John-ga [VP t] age+ta] no]-wa [VP Mary-ni ringo-o \text{ t}_v] da} \\
      & \quad \quad \text{[Focus]}
\end{align*}
\]

It is crucial here to note that assignment of the feature [Focus] takes place after remnant VP scrambling. This process is possible under the hypothesis given in (21), repeated below:

\[
(73) \quad \text{The feature [Focus] can be assigned to a syntactic object during the derivation.}
\]

I argued in the previous section that the possibility of surprising constituents in Japanese is attributed to the availability of remnant VP scrambling and that the impossibility of surprising constituents in such a language as English is attributed to the unavailability of scrambling in this language. What has been implicitly assumed in this discussion is that such a derivation as shown in (72) is possible only with scrambling, and not with operator movement; if remnant VP movement could be an instance of operator movement, then such a derivation should be possible even in English. We thus need to address the question why remnant VP movement cannot be operator movement.

Arguing against the remnant VP movement approach advocated by Koizumi (1995, 2000), Takano (2002) points out that in such languages as German and Dutch, which allow VP remnant movement, the following restriction is at work ((74) is taken from Takano (2000)):

\[
(74) \quad \text{Remnant movement of } \alpha \text{ is impossible if the head of } \alpha \text{ has moved out of } \alpha.
\]

(75) shows various kinds of VP remnant movement in Dutch, and (76) shows the impossibility of remnant VP movement without its head (the examples are taken from Takano (2002, pp. 253–254)):

\[
(75) \quad \begin{align*}
   & a. \quad \text{[VP Het boek aan Marie gegeven] heeft Jan waarschijnlijk.} \\
      & \quad \text{the book to given has probably} \\
      & \quad \text{`Jan has probably given the book to Marie.'} \\
   & b. \quad \text{[VP Het boek t\_i gegeven] heeft Jan waarschijnlijk aan Marie\_i.} \\
      & \quad \text{the book given has probably to} \\
      & \quad \text{`Jan has probably given the book to Marie.'} \\
   & c. \quad \text{[VP t\_i Gegeven] heeft Jan het boek, waarschijnlijk aan Marie\_i.} \\
      & \quad \text{given has the book probably to} \\
      & \quad \text{`Jan probably gave the book to Marie.'}
\end{align*}
\]

\[
(76) \quad \text{Remnant movement of } \alpha \text{ is impossible if the head of } \alpha \text{ has moved out of } \alpha.
\]

22 In order to complete the argument given in the text, the same paradigm of scope facts should be provided for the SC. However, it is not immediately clear whether the relevant examples demonstrate what is intended here. Consider, for instance, the following example:

\[
(\text{i}) \quad \text{John-ga hutari-no onna-ni sannin-no otoko-o syookaisita ga.} \\
   \quad \text{Nom two-Gen woman-Dat three-Gen man-Acc introduced but} \\
   \quad \text{[dono hutari-no onna-ni dono sannin-no otoko-o] ka oboetei-nai.} \\
   \quad \text{which two-Gen woman-Dat which three-Gen man-Acc Q remember-not} \\
   \quad \text{`Lit. John introduced three men to two women, but I don’t remember which three men to which two women.'}
\]

This example has only the reading of 2 > 3, as predicted, but it is not clear whether this scope fact is attributed to the latter sluicing part of this example or to its former part. One might claim that the former part, which has the basic word order, only allows the reading of 2 > 3, according to (54) and that a parallelism constraint of some sort imposes this reading upon the second part of this example. For this reason, I have avoided discussing relevant examples of SC cases in the text.
(76) "het boek aan Marie [vaarschijnlijk] gaf Jan.
the book to gave probably
'Jan (probably) gave the book to Marie.'

Takano (2000) derives the restriction given in (74) from the Attract/Move F mechanism of movement; that is, movement operations involve attraction of a feature that belongs to the head of a category to move. Thus, to move VP to the top of the sentences in (75) under this mechanism requires that the relevant feature of the head V be attracted by the topmost functional category. The unacceptability of (76) is then attributed to the fact that Attract F cannot apply to the head of VP since it has moved out of this VP.

Notice, however, that this account does not preclude all instances of remnant VP movement that are not accompanied by their heads; nothing will prevent such remnant VP movement if it does not involve any feature checking. Hence, this account is compatible with our claim that remnant VP movement involves feature checking and yet is not accompanied by its head. A qualification is necessary at this moment. Recall that we have been crucially assuming (73) and that the conceptual motivation for this hypothesis comes from the fact that there are instances of operator movement such as those illustrated in (1), repeated below, whose triggering feature will not be appropriately attributed to the inherent property of the head of a phrase to move.

(77) a. John, Mary likes very much.
   b. The man you're talking about, Mary likes very much.
   c. John met yesterday [the woman you had been talking about].

Thus, the present approach appears to be in conflict with the Attract/Move F approach, which crucially involves attraction of a feature that belongs to the head of a category to move. In order to solve this tension, I propose a slight modification of this movement mechanism. Let us take (77a) for illustration. Under the present assumptions, its derivation will be the following:

(78) a. Mary likes John very much.
   ∎ [Focus] assignment
   b. Mary likes John very much
      [Focus]
      ∎ [Focus]
   c. John, Mary likes t₁ very much
      [Focus]

Suppose that when the feature [Focus] is assigned to a syntactic object, it is registered to its label, as is natural given that no other place is available in the notation of the bare phrase structure theory proposed by Chomsky (1995).²⁴ Then, I suggest that the reason why the restriction stated in (74) operates is that when a target category is devoid of the content of its head, the probe cannot see the label of the whole category, which then leads to the impossibility of attracting the category in question. In order to implement this idea, I first assume the following:

(79) When a head is moved out of any of its projections, that projection is devoid of its label.

Then, I suggest that something like the following condition is at work:

(80) A category that lacks its label is invisible to the operation Attract.

With these assumptions, we can account for the ungrammaticality of (76), since the VP to move loses its label as a result of its head moving out of this VP and hence is invisible to the operation of attracting its [Focus]. On the other hand, these assumptions do not prevent remnant VP scrambling from taking place in deriving surprising constituents, as illustrated in (72). This is because scrambling does not involve any feature checking and hence is not subject to the operation Attract but simply is an instance of Move.²⁵

²³ As one reviewer correctly points out, we are led to the conclusion that German and Dutch do not have access to remnant VP scrambling, given the ungrammaticality of (76), although these languages are supposed to have scrambling. I have nothing interesting to say about why this must be the case.

²⁴ I also assume that the feature [Focus] is in the numeration of a given lexical array in accordance with the inclusiveness condition. The peculiarity of this feature resides only in the fact that it can be a member of a numeration without being incorporated into a lexical item.

²⁵ See Abe (2002) for the claim that movement operations that involve feature checking consist of Attract (or Select in his terminology, which amounts to Probe in Chomsky's (2000, 2001, 2004)) and Move, and further that scrambling involves only Move.
To recapitulate, what emerges out of the discussion so far about the Japanese constructions that are considered to involve overt operator movement is that Attract is not involved as far as generation of surprising constituents is concerned and that what appears to be an instance of overt operator movement can be derived by scrambling and assignment of [Focus] in the scrambled position. This may lead to the question whether Attract is inoperative in principle in all cases of the Japanese constructions under consideration. Notice that what we have shown in Section 2, which is concerned with cases that involve focalization of single constituents, is simply that scrambling may be involved in what appears to be an instance of operator movement. In what follows, I suggest that scrambling must be involved in such cases, leading to the following conclusion:

(81) [Focus] is not attracted in Japanese.

If this is the case, then it follows under the present system that Japanese does not have any genuine overt operator movement.

Before going into this issue directly, let us note that those cases which involve checking of [Focus] should be distinguished from those which involve “covert” movement of an inherent focus feature. One of the cases that belong to the latter category is a case involving a phrase with a focus particle (FP), as we have seen in Section 2. We have shown there that such phrases show ambiguity with respect to its scope, as illustrated in (16), repeated below:

(82) John-ga [Mary-ga oisii ringo-mo/-sae tabeta to] omotte-inai (koto)
   -Nom -Nom tasty apple-also/-even ate Comp think-not fact
   'John does not think that Mary ate also/even a tasty apple.'
   i. [embedded scope] John does not think that Mary ate A TASTY APPLE in addition to some other things.
   ii. [matrix scope] Even for A TASTY APPLE, John does not have an idea that Mary ate it (in addition to some other idea about some other things).

We can capture this scope fact by making the following assumption:

(83) FPs have an inherent feature that has to be checked in peripheral positions (call this feature [I(nherent)-Focus]).

Given this assumption, the embedded scope of the FP phrase of (82) is derived from the representation in which the [I-Focus] of the FP is covertly adjoined to the embedded TP, as indicated in (17a) above, whereas the matrix scope of the FP phrase is derived from the representation in which the [I-Focus] is covertly adjoined to the matrix TP, as indicated in (17b). There is evidence that such covert movement of [I-Focus] involves the operation Attract. This is concerned with the fact that such covert movement shows minimality effects, as observed by Sano (2001, p. 20):

(84) a. Taro-wa [Aiko-ga Jiro-no okane-sae nusunda] to syutyoosita.
   -Top -Nom -Gen money-even stole Comp claimed
   'Taro claimed that Aiko stole even Jiro’s money.'
   b. Taro-wa [Aiko-wa Jiro-no okane-sae nusunda] to syutyoosita.
   -Top -Top -Gen money-even stole Comp claimed

(84a) has not only the embedded scope reading of *Jiro-no okane-sae ‘even Jiro’s money’* but also its matrix scope reading, as is the case with (82) above. In (84b), the embedded subject is changed into *Aiko-wa ‘Aiko-Top’*, which functions as contrastive. Sano claims, as I think correctly, that this sentence lacks the matrix reading of *Jiro-no okane-sae*. Given that a phrase with contrastive -wa carries [I-Focus], this fact follows straightforwardly from the assumption that Attract is involved in the movement operation in question; that is, *Jiro-no okane-sae* cannot be attracted by the matrix TP since *Aiko-wa* is closer to the probe than *Jiro-no okane-sae*. Note that English FP phrases also show scope ambiguity, as observed by Taglicht (1984):

(85) I knew he had taught only SPANISH (to students).
   i. [embedded scope] I knew he hadn’t taught any other language.
   ii. [matrix scope] I didn’t know he had taught any other language.

From these facts, it will be safe to conclude that as far as “covert” movement is concerned, the operation of Attract is at work universally.

26 It will be reasonable to claim that this feature is basically the same as [Focus], which is assigned during the course of a derivation and that the difference lies only in the way the feature is introduced in the derivation. I will not pursue this matter here, simply using the name [I-Focus] to make it explicit that this feature is incorporated into a lexical item, unlike [Focus].
Turning back to those cases that involve licensing of [Focus], let us first note that the operation licensing it involves “overt” movement. Thus, in contrast with the way of licensing [I-Focus], as stated in (83), [Focus] is licensed in the following way:

(86) A phrase that has received [Focus] must be located overtly in a peripheral position.

An illustration of licensing [Focus] was given in (78), repeated below:

(87) a. Mary likes John very much.
   ↓ [Focus] assignment

b. Mary likes John very much
   [Focus]
   ↓ Attract [Focus]

c. John, Mary likes $t_i$ very much
   [Focus]

If the same process could happen in Japanese, simply fronting a DP assigned [Focus] to the top of a sentence would give rise to a Topicalization construction, just like English. This is clearly not the case, however; consider the following examples:

(88) a. Mary-ni $i$ John-ga kinoo $t_i$ atta.

   -Dat -Nom yesterday saw

   'Mary, John saw yesterday.'

b. Pekin-ni $i$ John-ga kinoo $t_i$ itta.

   Peking-Dat -Nom yesterday went

   'To Peking, John went yesterday.'

In both sentences, the moved phrases Mary-ni ‘Mary-Dat’ and Pekin-ni ‘to Peking’ do not function as topics but simply contribute to changing the basic word orders. This will indicate that the feature [Focus] is not attracted in the way shown in (87) in Japanese. Instead, this language uses the topic marker -wa to indicate a topic phrase, as shown below:

(89) a. Mary-ni-wa $i$ John-ga kinoo $t_i$ atta.

   -Dat-Top -Nom yesterday saw

   'Mary, John saw yesterday.'

b. Pekin-ni-wa $i$ John-ga kinoo $t_i$ itta.

   Peking-Dat-Top -Nom yesterday went

   'To Peking, John went yesterday.'

(90) a. Mary-wa $i$ John-ga kinoo $e_i$ atta.

   -Top -Nom yesterday saw

b. Pekin-wa $i$ John-ga kinoo $e_i$ itta.

   Peking-Top -Nom yesterday went

(89) illustrates instances of PP-Topicalization, in which an NP is accompanied with a postposition or Case marker, whereas (90) illustrates instances of NP-Topicalization in which such a postposition or Case marker is dropped. Note further that, as observed by Saito (1985), NP-Topicalization, unlike PP-Topicalization, does not show island sensitivity. Thus, all the examples in (12), which show island sensitivity of PP-Topicalization, become acceptable if the postpositions or Case markers of the topic phrases are dropped, as shown below:


   -Top -Nom went fact-Nom have person-Acc found seem

   'To Peking, it seems that John found a person who has been $e_i$.'


   there-Top -Nom want-to-go though ignoring be seem

   'There, it seems that Mary is ignoring although John wants to go $e_i$.'
This indicates that in NP-Topicalization, the topic phrase is base-generated in the top of a sentence. Suppose for the sake of argument that there is a functional category above TP (call it Fo(cus)P) that is responsible for licensing such a phrase as a topic, as schematically shown below:

(92) \[ \text{FoP} \]
    \[ \text{Fo'} \]
    \[ \text{TP} \]
    \[ \text{Fo} \]

Suppose further that the head Fo selects in its Spec position a phrase carrying [I-Focus], i.e., a FP phrase. Then, in a case of NP-Topicalization such as (90a), the topic phrase Mary-wa 'Mary-Top' is base-generated in the Spec of FoP and is associated with the pro base-generated in the complement position of atta 'saw'. Given this base-generation option, it is natural to claim, following Saito (1985, 1987), that a case of PP-Topicalization such as (89a) is simply derived by scrambling from the following sentence with basic word order:

(93) John-ga kinoo Mary-ni-wa atta.
    -Nom yesterday -Dat-Top saw

It is interesting to note that the Japanese CC and SC have variants which have characteristics of NP-Topicalization in relevant respects. Hoji (1987) observes that there are two types of the Japanese CC: one involves a phrase in the focus position that is accompanied with a postposition or a Case marker, as shown in (4a), repeated below as (94a) and the other involves a focused phrase which is a bare NP, as shown in (94b):

(94) a. [John-ga [Bill-ga Mary-ni t\_i ageta to] omotteiru no]-wa [sono hon-o], da.
    -Nom -Nom -Dat gave Comp think NL-Top that book-Acc be
    'It is that book\_i that John thinks Bill gave t\_i to Mary.'

b. [John-ga [Bill-ga Mary-ni t\_i ageta to] omotteiru no]-wa [sono hon], da.
    -Nom -Nom -Dat gave Comp think NL-Top that book be

As expected, the latter case of the CC does not show island sensitivity, unlike the former case; compare (4b, c) with the following examples:

(95) a. [John-ga [Mary-ni e\_i, ageta to] kiratteiru no]-wa [sono hon], da.
    -Nom -Dat gave person-Acc hate NL-Top that book be
    'It is that book\_i that John hates the person [who gave e\_i to Mary].'

b. [John-ga [Mary-ga e\_i, syokuzi-ni sasotta node] rakutansiteiru no]-wa [sono hito], da.
    -Nom -Nom dinner-to invited because be-depressed NL-Top that person be
    'It is that person, that John is depressed [because Mary invited e\_i to dinner].'

Suppose that we adopt the PF deletion approach to the CC and that its presupposed clause contains a FoP whose head selects in its Spec position a phrase that is strictly identical with the one that appears in the focus position. Note that in this case, a phrase that can appear in the Spec of FoP does not have to be one carrying [I-Focus]. Let us then assume that in such a case, a phrase in the Spec of FoP must bear [Focus]. We can then summarize the selectional condition of FoP as follows:

(96) FoP selects in its Spec position a phrase bearing [Focus] or accompanied with [I-Focus].

In the case of the Japanese CC in which a bare NP appears in the focus position, it is reasonable to claim that the same bare NP is also base-generated in the Spec of FoP in the presupposed clause and is assigned the feature [Focus]. This NP then gets deleted under strict identity with the one in the focus position. Given this base-generation option, it is reasonable to claim that in the other type of the Japanese CC, a phrase that is strictly identical with the one in the focus position is moved by scrambling to the Spec of FoP in the presupposed clause and then is assigned the feature [Focus].

Much the same account can be given to the Japanese SC, which also has two variants that correspond to NP- and PP-Topicalization. Thus, besides the one in which the sluice retains a postposition or a Case marker, as observed above, there is another type in which such a marker is dropped; compare the following examples:
John gave something to Mary, but I don’t know what.

John hates the person who gave something to Mary, but I don’t know what.

John is depressed because Mary invited someone to dinner, but I don’t know who.

Thus, again, in the latter type of the SC, the sluice is base-generated in the Spec of FoP, (in fact, the Spec of CP; see the following discussion), is assigned [Focus] during the derivation, and the TP below it gets deleted under identity with the previous clause. Given this base-generation option, it is natural to claim that the former type of the SC is derived by scrambling the sluice from its base-generated position to the Spec of FoP and is assigned [Focus].

Though far from decisive, all the above considerations lead to the conclusion stated in (81), repeated below:

[Focus] is not attracted in Japanese.

I would like to speculate here on why this is the case, especially from a parametric point of view, comparing this language with such a language as English. A natural idea will be to exploit the “defectiveness” of functional categories in Japanese, characterized by Fukui (1986), which is closely associated to the lack of agreement in this language. I suggest that the following parametrization is at work with respect to the way [Focus] is licensed:

Any functional category is able to attract [Focus] overtly in its peripheral position in a language in which functional categories are “active.”

In English, a typical language in which functional categories are active, any functional category is able to attract [Focus] in principle, so that the Topicalization construction can be derived in the way indicated in (87). In Japanese, on the other hand, functional categories cannot attract [Focus], since they are defective. Instead, a particular functional category such as FoP is responsible for licensing [Focus] in such a way that it requires a phrase in its Spec position to bear this feature, but since it is incapable of attracting a phrase, it needs assistance of scrambling.

If the above reasoning is on the right track, we are led to the conclusion that there is no genuine overt operator movement in Japanese, attributing apparent cases of such movement to cases of scrambling, and that this fact is derived from the

There remains one construction that should be considered: the Japanese RDC. Unlike the other constructions considered in the text, this construction does not seem to have a counterpart that corresponds to NP-Topicalization. Thus, it is mandatory to retain a postposition or a Case marker of a dislocated phrase in this construction, as illustrated below:

(i) John-ga Mary-ni e, ageta yo, sono hon-o i.
    -Nom -Dat gave that book -Acc
    ‘John gave e, to Mary, that book.’

I have no account to offer for this fact here, but note that if the undeletability of such a Case marker or a postposition is given an independent account, this may not be incompatible with what is claimed in the text. Note further that this construction appears to involve rightward movement, hence not movement into a Spec but rather an adjunction operation. There is an approach, however, advocated by Abe (1999) and Tanaka (2001), among others, according to which the dislocated phrase initiates a separate sentence that involves deletion. Thus, (i) will have roughly the following derivation:

(ii) a. John-ga Mary-ni e, ageta yo, sono hon-o, [John-ga Mary-ni i, ageta]
    (leftward movement of sono hon-o)

b. John-ga Mary-ni e, ageta yo, sono hon-o, [e]
    (deletion under identity)

Given this approach, we can claim that the leftward movement of sono hon-o ‘that book-Acc’ indicated in (iiia) is in fact scrambling into the Spec of FoP.
defectiveness of functional categories. Given this conclusion, there is a possibility of providing a new perspective of analyzing phenomena that appear to run counter to the consequences derived from the defectiveness of functional categories. While Fukui (1986) claims that the lack of wh-movement in Japanese will be derived from the defectiveness of functional categories or the lack of agreement, Takahashi (1993) provides the data that seem to show that there are instances of wh-movement in Japanese. The relevant examples are reproduced below:

(101) a. John-wa [Mary-ga nani -o tabeta ka] siritagatteiru no?
    -Top -Nom what-Acc ate Q want-to-know Q
    ‘Does John want to know what Mary ate?’ or
    ‘What does John want to know whether Mary ate?’

b. Nani,-o John-wa [Mary-ga ti tabeta ka] siritagatteiru no?
    what-Acc -Top -Nom ate Q want-to-know Q
    ‘What does John want to know whether Mary ate?’

In each of these examples, there are two Q-markers; one is ka in the embedded clause and the other is no in the matrix clause. Takahashi observes that (101a) is ambiguous according to whether the wh-phrase in situ nani ‘what’ takes scope over the embedded clause or the matrix clause. When nani takes embedded scope, the whole sentence is interpreted as a yes or no question, and when it takes matrix scope, the embedded ka is interpreted as ‘whether’. Takahashi makes an interesting observation with (101b), where nani-o is moved out of the embedded question clause. He notes that in this sentence, nani can only take matrix scope. Given this fact, he claims that if the movement operation involved in (101b) is simply an instance of scrambling, then it will not account for the scope fact correctly, since, according to Saito (1989), scrambling may be undone, as exemplified by the following example:

    which book-Acc -Nom -Nom library-from checked-out Q want-to-know fact
    ‘Lit. which booki, Mary wants to know [Q John checked out ti from the library].’

In this sentence, the wh-phrase dono hon-o ‘which book-Acc’ is scrambled out of the embedded question clause. From the fact that the matrix clause is declarative, it is obvious that the scrambled wh-phrase should take scope over the embedded clause initiated by the Q-morpheme ka. To account for this fact, Saito proposes that scrambling can be freely undone in the LF component due to its nature as semantically vacuous movement. Under this proposal, the wh-phrase dono hon-o, which has undergone long distance scrambling, can be moved back to its original position, in which it is properly bound by the Q-morpheme ka, representing the embedded scope. Going back to example (101b), the fact that the fronted wh-phrase takes only matrix scope suggests that the movement operation is not scrambling but rather wh-movement, an instance of operator movement that resists LF undoing, Takahashi claims.

This apparent counterexample to Fukui’s (1986) line of theory of agreement and functional categories can be solved in an interesting way under the present mechanism of movement and licensing [Focus]. Let us first assume, following Abe (2002, 2005), that a wh-phrase has [WH] as its inherent feature and that this feature can be licensed in the following way:

(103) [WH] is licensed:
    (i) in the Spec of [+WH] Comp; or
    (ii) by means of being bound by [+WH] Comp.

Besides this, let us make the following assumption, proposed by Abe (2005):

(104) Wh-phrases can have [Focus] features in Japanese.

Recall that generally [Focus] is licensed in a peripheral position and that in Japanese it is licensed in the Spec of FoP, as stated in (96). Suppose that it is in fact the Spec of a [+WH] Comp that licenses a wh-phrase bearing [Focus]. Then, when a wh-phrase takes the option of bearing [Focus], it must be in the Spec of CP headed by a [+WH] Comp, whereas, if it does not bear such a feature, it can be in situ or undergo a genuine instance of scrambling, an adjunction to any phrase. Let us then adopt the following economy condition, proposed by Abe (2005):

(105) “Operator movement” is preferred over scrambling unless the derivation crashes.

We need to interpret “operator movement” here as movement that involves licensing [Focus], since we have been claiming that a genuine instance of overt operator movement does not exist in Japanese. Given the two options available to a
wh-phrase XP when it undergoes movement, depending upon whether or not it carries [Focus], as schematically indicated below:

(106)  
  a. \(^*\text{XP } \ldots t\)  
        \[\text{WH}\]
  b. \text{XP } \ldots t  
        \[\text{Focus}\][WH]

(105) dictates that the option of (106b) be chosen over (106a) as long as the derivation converges. Given these assumptions, (101b) will be derived schematically in the following way:

(107)  
  a. \[\text{CP } \ldots \text{wh-phrase } \ldots +\text{WH}] +\text{WH}\]  
        \[\text{WH}\]
        ↓ scrambling of the wh-phrase to the Spec of the matrix CP  
  b. \[\text{CP wh-phrase}, \ldots \text{CP } \ldots t_i \ldots +\text{WH}] +\text{WH}\]  
        \[\text{WH}\]
        ↓ assignment of [Focus]
  c. \[\text{CP wh-phrase}, \ldots \text{CP } \ldots t_i \ldots +\text{WH}] +\text{WH}\]  
        \[\text{Focus}\][WH]

Note that this derivation does not involve wh-movement in its strict sense, but rather involves scrambling. In this respect, it is compatible with Fukui’s (1986) line of characterization of Japanese. On the other hand, the output of this derivation is as if it were derived by wh-movement, thanks to the assignment of [Focus] to the wh-phrase, and the chain created by scrambling is not regarded as semantically vacuous, hence resisting LF undoing. This correctly derives the fact that nani ‘what’ in (101b) takes only matrix scope. In contrast, if (102) took the same derivation as given in (107), it would lead to a violation of (103), as schematically shown below:

(108) *\[\text{CP wh-phrase}, \ldots \text{CP } \ldots t_i \ldots +\text{WH}] -\text{WH}\]  
        \[\text{Focus}\][WH]

Suppose that (103) is a condition on convergence of a derivation. Then, economy condition (105) dictates that the other option be taken, i.e., the one that involves a genuine instance of scrambling with no assignment of [Focus]. In this case, “LF undoing” in the sense of Saito (1989) is allowed. Let us reinterpret this operation as allowing the feature [WH] to be left behind. Then, the legitimate output of (102) will be roughly as follows:

(109) \[\text{CP } \ldots \text{wh-phrase}, \ldots \text{CP } \ldots t_i \ldots +\text{WH}] \ldots -\text{WH}\]  
        \[\text{WH}\]

where the wh-phrase is adjoined to any phrase XP, as long as the adjunction derives the right word order. Here, the wh-phrase is properly licensed according to (103ii) since its [WH] feature, which is left behind in its original position, is bound by the embedded [+WH] Comp.

6. Conclusions

I have demonstrated that what appears to be an instance of overt operator movement in such a language as Japanese in fact involves the operation of scrambling. It was shown that such an instance of overt operator movement shares with

---

28 What is implicitly assumed here is that [WH] cannot be separated from [Focus] in the derivation given in (107). See Abe (2002, 2005) for the claim that these two features must be merged into one feature called [Wh-Focus].

29 Takahashi (1993) provides another case that demonstrates an instance of wh-movement in Japanese, one exhibiting a superiority effect, as shown below:

(i)  
  a. John-ga dare-ni [Mary-ga nani-o tabeta to] itta no?  
      -Nom who-Dat -Nom what-Acc -Comp said Q  
      ‘Who did John tell that Mary ate what?’
  b. ??Nani-o John-ga dare-ni [Mary-ga t_i tabeta to] itta no?  
      -what-Acc -Nom who-Dat -Nom ate -Comp said Q

Given that superiority effects are accounted for by a minimality condition on Attract such as Chomsky’s (1995) Minimal Link Condition, such a fact as given in (i) might constitute another counterexample to our conclusion given in (99). I will leave this matter for future research, however.
scrambling the property that it is insensitive to the wh-island. On the basis of this fact, I argued that an apparent instance of overt operator movement in this language is analyzed as resulting from scrambling of the target phrase and then assigning the feature [Focus] to it, a process available under the assumption that [Focus] is assigned during the derivation. Given the property of scrambling pointed out by Hoji (1985), according to which string-vacuous application of scrambling is prohibited, I argued that scrambling cannot be involved in real instances of null operator movement. This is motivated by the fact that those instances exhibit wh-island sensitivity. Furthermore, the hypothesis that an apparent instance of overt operator movement involves scrambling is motivated by the phenomena of what Takano (2002) calls “surprising constituents.” I argued that the presence of these phenomena in Japanese is attributed to that of multiple scrambling. After showing that multiple scrambling should be analyzed as an instance of remnant VP movement, an approach advocated by Koizumi (1995, 2000), I argued that those phenomena of surprising constituents are analyzed in the same way. Finally, I showed that multiple scrambling should be analyzed as an instance of remnant VP movement, an approach advocated by Fukui’s (1986) line of parametrization of functional categories and agreement: since functional categories are defective in this language, the feature [Focus] cannot be attracted, hence requiring the assistance of scrambling in order to move the target phrase.30

Finally, if the conclusion that we have reached in this paper is on the right track, it lends strong support to the conception revealed by Chomsky (2004) according to which Move (or internal Merge according to his terminology) is “an operation that is freely available,” and “accordingly, displacement is not an ‘imperfection’ of language; its absence would be an imperfection.” (p. 8) If we follow this conception, scrambling, a movement operation that lacks its own driving feature, should also be an operation freely available to UG in principle.

Acknowledgments

I am indebted to the three anonymous reviewers of this journal for their clarification and helpful suggestions. The usual disclaimer applies.

References


30 One may be naturally led to the question whether it is possible to claim that overt attraction of a phrase is inoperative at all in Japanese, a claim stronger and surely more desirable than that given so far. We have not touched upon so-called A-movement in Japanese. If we succeed in claiming that such movement also involves scrambling, then the stronger claim will be verified. But I must leave this pursuit for future research.