Phase Out*

Toshifusa Oka

1. Introduction

In this paper I will propose an approach for the island phenomena on the basis of Chomsky's (1998, 1999) idea of the phase-based derivation.

The following paradigm illustrates the *wh*-islandhood:[1]

(1) a. ?? [who$_i$ do you wonder [when/whether John kissed t_i]]
b. * [why$_i$ do you wonder [when/whether John kissed Mary t_i]]
c. [who$_i$ do you think [that John kissed t_i]]
d. [why$_i$ do you think [that John kissed Mary t_i]]

As has been observed in the literature, *wh*-clauses, unlike *that*-clauses, behave as a weak island for argument extraction and as a strong island for adjunct extraction.

Below I will present discussion of these and related phenomena, including subject-object asymmetries.

2. Phase and successive cyclicity

A traditional assumption is that *wh*-extraction from a *that*-clause is carried out successive cyclically from clause to clause, as in (2):

(2) $[_{CP^2}$ wh$_i$ C^2 [$[_{CP^1}$ t_i' C^1 [t_i]]]]

Here, the extracted *wh*-phrase is moved through Spec of COMP, given the current CP analysis of the clause.

This is also required by Chomsky's (1999) Derivation-by-Phase (DBP) theory, which imposes the Phase Impenetrability Condition (PIC) on syntactic operations:

(3) The domain of H is not accessible to operations at ZP, but only H and its edge, for $[_{ZP}$ Z ... $[_{HP}$ α [H YP]]] , where HP is a strong phase and ZP is the next highest phase.

When the derivation proceeds to the next strong phase ZP, no element in the previous strong phase HP is any longer accessible unless it is in the edge of HP, namely the residue outside of H-bar.

Given the assumption that CP constitutes a strong phase, an extracted *wh*-phrase

---

* I would like to thank Hiroyuki Ura, who never fails to give me suggestions and encouragements while I am working on a paper. I'm also obliged to TES reviewers for their valuable comments on an earlier version of this paper.

[1] Throughout the paper, indices are used only for the expository purpose, and therefore they should not be considered as linguistic entities.

should be moved through the edge of COMP. For the purpose of this paper, I will keep aside Chomsky's assumption that *v*P counts as a strong phase, just for simplicity of discussion.

For the phase-by-phase derivation, Chomsky (1999) also proposes the following:

(4) Interpretation/evaluation for PH_1 is at the next relevant phase PH_2, where PH_1 is strong and PH_2 is the next highest strong phase.

When the construction of a syntactic structure reaches a strong phase, interpretation is applied to the structure constructed at the previous strong phase.

Thus, Spell-Out applies to CP^1 in (2) when CP^2 is constructed. At this stage, the *wh*-phrase in the Spec of CP^1 is a trace, so that it is subject to deletion by virtue of its trace status, satisfying the recoverability condition. If Spell-Out had applied just as soon as CP^1 was constructed, it would have been necessary to spell out the *wh*-phrase in the edge of CP^1, since it was a non-trace copy at that stage. The principle (4) provides a simple solution to the problem that no phonetic realization is found in the position of an intermediate trace.

A natural assumption is that semantic interpretation is applied in parallel with Spell-Out. The semantic interpretation of a syntactic structure must be able to wait until the derivation proceeds to the stage where the structure is contained in a larger one. Under the principle (4), the interpretation of CP^1 in (2), for example, is done when the derivation proceeds up to the CP^2 phase. The situation is the same as in the case of Spell-Out. If the interpretation had been applied as soon as CP^1 is constructed, the *wh*-phrase would have been wrongly interpreted in Spec of CP^1, rendering CP^1 interrogative, under the condition of full interpretation. When the derivation reaches the CP^2 phase, the *wh*-phrase in Spec of CP^1 has become a trace copy. In the following sections I will discuss how such an intermediate trace is treated.

The principle (4) implements the requirement that interpretation of a syntactic structure should be able to be postponed. This principle, just as the PIC does, also realizes the idea that derivations proceed in a local way, stating that interpretation can be postponed, but only until the next strong phase. For interpretation, it is only possible to go back to the immediately previous phase. The principle should be understood to be vacuously satisfied in the case of the final structure yielded by a derivation. Thus, when the derivation proceeds to the stage where syntactic operations do not apply any more, the maximal structure constructed so far undergoes interpretation, terminating the derivation.

3. *Wh*-islandhood

Along the traditional lines of reasoning, the islandhood of *wh*-clauses will be explained by assuming that the edge of a *wh*-clause is not available for an intermediate

landing site for an extracted *wh*-phrase. This will block extraction from a *wh*-clause unselectively, whether the extracted *wh*-phrase is an argument or an adjunct. An important observation that has been attracting much attention, however, is that although the acceptability of argument extraction is far from uncontroversial, it is sharply contrasted with that of adjunct extraction. This suggests that argument extraction should be permitted from a *wh*-clause just as from a *that*-clause, attributing its degrading character to some other factor. If so, *wh*-clauses, just like *that*-clauses, should be able to provide an extra landing site for an extracted *wh*-phrase, allowing a successive-cyclic derivation, analogous to the one illustrated in (2).

The derivation of (1a) yields the following:

(5) $[_{CP^2}$ who ... $[_{CP^1}$ t'_{who} when/whether ... t_{who} ...]]

Given the principle (4), the semantic interpretation of the CP^1 is done when the derivation proceeds up to the CP^2 phase.

At this stage, a trace copy of *who* is left in the edge of CP^1. If it is interpreted there as an interrogative operator, it will yield a multiple question, taking the CP1 scope. In this case, however, the non-trace copy of *who* in CP2 will be left without being properly interpreted to constitute an independent operator-variable construction.

Here a question arises whether an interrogative COMP allows multiple operators for interpretation in the first place, considering the fact that there is no syntactic multiple *wh*-fronting, at least in English, as is shown in the following:

(6) a. I wonder [who_i John persuaded t_i to buy what]
b. *I wonder [$what_j$ who_i John persuaded t_i to buy t_j]
c. *I wonder [who_i $what_j$ John persuaded t_i to buy t_j]
d. *I wonder [whether John persuaded Mary to buy what]
e. *I wonder [$what_j$ whether John persuaded Mary to buy t_j]

It may be the case that multiple *wh*-fronting is impossible only for some syntactic reasons. Or there may be an interpretive reason as well. Thus it may be that a multiple question interpretation with respect to an interrogative COMP is only obtained by absorption in the following configuration:

(7) $[_{CP}$ wh_0 C $[_{TP}$ ... wh_1 ... ... wh_n ...]]

The absorption relation is one-to-many, one and only one *wh*-operator being in the edge of COMP and others inside its complement, so that no extra *wh*-operator is allowed in the CP. If so, the *wh*-trace in Spec of CP^1 in (5) cannot be interpreted in that position.

There is no way to interpret the intermediate trace, but there should be a way to solve or avoid this problem. Note that this trace is deleted in the PF side. This suggests that there may be a semantic version of deletion applied to traces under the recoverability condition.

When CP^1 in (5), which is (8a), is subject to interpretation, the intermediate trace undergoes deletion, yielding (8b), where Ø indicates the deletion site:

(8) a. $[_{CP1}$ t'_{who} when/whether ... t_{who} ...]

b. $[_{CP1}$ Ø when/whether ... t_{who} ...]

In (8b) the original copy is interpreted as a variable, unbound by any operator.

At the final stage, interpretation applies to CP^2, which contains an already interpreted structure, namely CP^1, so that the interpretation obtained from (8b) must be incorporated to yield the entire interpretation. This amounts to taking consideration of the structure (9):

(9) $[_{CP2}$ who ... $[_{CP1}$ Ø when/whether ... t_{who} ...]]

Here the free variable is linked to the matrix *wh*-phrase so as to get properly bound, yielding a well-formed interpretation with the matrix scope.

Note that this linking is part of the interpretive process, so that it is not subject to the PIC (3), which is imposed on syntactic operations. The linking will be operative only after the interpretation obtained from CP1 is combined with the interpretation of the residue of CP2. It is a process analogous to such as the binding of (long-distance) anaphors and pronouns and the absorption of in-situ *wh*-phrases. These elements can be associated with elements outside the minimal clause containing them. They can be left unassociated in its minimal clause. If they remain unassociated with a proper element in a proper configuration throughout the derivation, it will simply yield an ill-formed interpretation of the entire structure.

Finally, let us consider the degrading character of the (argument) extraction from a *wh*-island, as is exemplified in (1a). We may be able to attribute it to the total deletion of the trace copy in the CP edge. This deletion differs from the partial deletion of a *wh*-phrase for reconstruction, which is discussed in Chomsky (1995), in that the latter does not totally destroy the effect of movement since it leaves an operator-variable construction. The total deletion, on the other hand, has the effect of "undoing" of movement. This gives the derivation an uneconomical flavor, resulting in a degrading status of the derived expression.

Chomsky (1986b) discusses the "double *Wh*-island violation," observed by Rizzi (1982), as is exemplified in the following:

(10) $[_{CP3}$ what did you wonder $[_{CP2}$ who knew $[_{CP1}$ who saw t]]]

This is more degraded than the case of extracting from a single *wh*-island. Chomsky argues that violations are cumulative. In our terms, the example (10) undergoes deletion of an intermediate trace twice, at the interpretation for CP^1 and for CP^2, lessening the acceptability.

4. Argument-adjunct asymmetry

Now let us consider the question: How is adjunct extraction totally barred from a *wh*-clause?

The derivation of (1b) proceeds in the same way as that of (1a), to yield (11):

(11) $[_{CP^2}$ why ... $[_{CP^1}$ t'_{why} when/whether ... t_{why} ...]]

The interpretation of CP1 deletes the intermediate trace, changing (12a) into (12b):

(12) a. $[_{CP^1}$ t'_{why} when/whether ... t_{why} ...]

b. $[_{CP^1}$ Ø when/whether ... t_{why} ...]

Here some difference between arguments and adjuncts should become relevant, barring free adjunct variables while permitting free argument variables.

One possibility is to appeal to the referentiality difference, incorporating the insight of Chinque (1990). Suppose that that free variables in the relevant phase cannot be simply permitted as such. Rather it can be only permitted by being reinterpreted as if it were a resumptive pronoun. This interpretation is only temporary, and the free variable is properly bound in the interpretation of the next relevant phase. Then an adjunct variable will be unable to be interpreted as a resumptive pronoun, since it does not range over individuals. Therefore, adjuncts cannot be extracted from a *wh*-island. The impossibility to extract non-referential arguments will be explained in the same way.

An alternative approach is to consider a more fundamental difference between arguments and adjuncts: an argument enters into a θ-relation to render the structure well-formed, but an adjunct is, in essence, just added to an already well-formed structure. Suppose that an economy condition requires that interpretation should be minimal. Under this condition the original trace copy of *why* in (12b) is deleted, yielding (13):

(13) $[_{CP^1}$ Ø when/whether ... Ø ...]

This is possible, because the resulting structure is the well-formed structure that would have been yielded if *why* had not been adjoined in the first place. If the intermediate trace had not been deleted, it would be impossible to delete the original trace, since it would destroy the operator-variable construction, leaving an operator binding no variable.

Deletion of a non-*wh*-adjuncts that is not subject to *wh*-movement and therefore stay in situ is blocked by the recoverability condition. Note also that this approach implies that if such an adjunct is moved to some non-operator position where it could be interpreted, then its trace should be deleted for minimal interpretation, revealing no (A-)chain property.

For interpretation of CP^2, we will have (14):

(14) $[_{CP^2}$ why ... $[_{CP^1}$ Ø when/whether ... Ø ...]]

Here *why* is an operator binding no variable, an illegitimate object. If it could have moved through some adequate position in the matrix, it could bind a variable left there, but this would never yield an interpretation where *why* binds into the embedded clause.

In the case of argument extraction, deletion of the original trace is blocked, because it would destroy a θ-relation so as to render the structure ill-formed. Thus the argument trace stays as a free variable waiting to be properly bound at the next relevant phase. This time we may be able to refrain from assuming that a free variable has to act as a resumptive pronoun. An advantage of this approach is that it exploits the most fundamental difference in character between arguments and adjuncts, just as the approaches of Lasnik and Saito (1992) and Chomsky (1986b) do. Moreover, it will more readily explain the fact that not only true adjuncts such as *why* and *how* but also allegedly referential adjuncts such as *when* and *where* are not possible to extract from a *wh*-island.

I presented in Oka (1995) an explanation for the argument-adjunct asymmetry with respect to island sensitivity by implementing the essentially same idea in a different way within a more economy-based framework. I argued there that non-referential *wh*-arguments can be treated in the same way as *wh*-adjuncts because reconstruction applied to such an argument leaves an adjunct variable behind. Consider the following:

(15) a. [how many books]$_i$ did you read t_i yesterday
b. [how many]$_j$ did you read [t_j books] yesterday
c. [how]$_k$ did you read [t_k many books] yesterday

To obtain a non-referential reading for *how many books*, reconstruction must apply to (15a) in such a way as to yield (15b) or (15c). In either case the trace left behind counts as an adjunct rather than as an argument. Thus argument extraction from a *wh*-clause will be barred under a non-referential reading in the same way as adjunct extraction in the present framework.

5. *That*-clause

Now the question is: How is it possible to extract a *wh*-phrase from a *that*-clause, whether it is an argument or adjunct, and with no degrading effect?

The derivation proceeds in the way illustrated in (16):

(16) $[_{CP^2}$ wh_i ... $[_{CP^1}$ t_i' that ... t_i ...]]

For interpretation of CP^1, there needs to be some other way to save the structure than deleting the trace copy in the CP edge, which is the only option for the *wh*-clause.

A fundamental character of *that* as a COMP, compared with the interrogative COMP, is that it simply works to indicate the presence of a clause and does not add any semantic content to the clause. It is transparent and unselective, so that it allows

an empty operator in its edge, unlike the interrogative COMP, which is exemplified in relative clauses and the cleft sentences as in (17), where O stands for an empty operator.

(17) a. the girl [O_i that John kissed t_i]
b. the reason [O_i that John kissed Mary t_i]
c. it was Mary [O_i that John kissed t_i]
d. it was for that reason [O_i that John kissed Mary t_i]

The operator is linked to its antecedent to yield a legitimate interpretation. What is important here is that there is no argument-adjunct asymmetry with respect to the operator-variable construction.

The empty operator is empty in the sense that it is phonetically unrealized, whether it is assumed that it is listed in the lexicon as an item lacking a phonetic content in the first place, or that it is just a *wh*-form, *which* or *who*, that undergoes deletion in the PF side. Moreover, it is not only phonetically empty, but also semantically empty. It formally enters into construction of an operator-variable relation, binding a variable from an operator position. It has no semantic content to determine the range or value of its variable by itself. It needs to have its antecedent to complete the strong binding of the variable in the sense of Chomsky (1986a: 85). Notice that in these respects an intermediate trace in the CP edge has a similar character, being deleted phonetically and bound by its antecedent.

In these considerations, suppose that for interpretation a trace copy of a *wh*-phrase in the CP edge is identified with an empty operator. Then, the structure (18a) becomes (18b) for interpretation of CP^1:

(18) a. $[_{CP^1}$ t_i' that ... t_i ...]
b. $[_{CP^1}$ O_i that ... t_i ...]

In (18b) the original trace is not free, so that it does not have to be deleted even if it is an adjunct. At the level of CP^2, we will have the following:

(19) a. $[_{CP^2}$ wh ... $[_{CP^1}$ O_i that ... t_i ...]]
b. $[_{CP^2}$ wh_i ... $[_{CP^1}$ Ø that ... t_i ...]]

Here the empty operator in (19a) is linked to the matrix *wh*-phrase and then deleted, so that the semantically incomplete operator-variable construction (O_i, t_i) turns into the complete one (wh_i, t_i) in (19b). The deletion of the intermediate trace is part of the linking procedure applied to an empty operator, so that it requires no extra cost. Thus *wh*-extraction from a *that*-clause is successful, regardless of the argument-adjunct status of the extracted element.

To complete the explanation for the *wh*-island, it is necessary to assume that the interrogative COMP, unlike *that,* does not allow an intermediate trace in its edge to be interpreted as an empty operator. This follows if we assume that the interrogative

COMP can have only one interrogative operator so that no kind of operator is permitted as an extra in the CP edge. However, this property is subject to parametric variation, since there are languages that do allow the extraction from a *wh*-clause. To explain for the lack of the *wh*-islandhood, Reinhart (1981), Rudin (1988), Koizumi (1995) have proposed that a *wh*-phrase can be moved through the intermediate CP edge. In the present framework, this means that the interrogative COMP must be able to have an extra operator in its edge in such languages. This is probably a marked option for the interrogative COMP, but in the languages such as Bulgarian and Romainian that Rudin refers to as [+MFS] languages, the presence of overt multiple *wh*-fronting to the CP edge gives direct evidence to determine the COMP property.

Finally, note that once we permit an empty operator in the edge of *that* for interpretation, we must somehow rule out the improper movement that is illustrated in (20):

(20) *John$_i$ seems [t$_i$' that [it appears [t$_i$ to be intelligent]]]

Here *John* is successively moved through the edge of *that*-clause. If the intermediate trace is interpreted as an empty operator and is relinked to *John*, it will yield the following:

(21) John$_i$ seems [O$_i$ that [it appears [t$_i$ to be intelligent]]]

Compare this with the well-formed construction of *tough*-movement in (22):

(22) John$_i$ is difficult [O$_i$ to talk to t$_i$]

The relevant difference here is that *John* and O is not associated by movement but "base-generated" as two separate elements.

Chomsky (1986a: 98) proposes that the case of improper movement should be ruled out as a Condition (C) violation, stating as follows:

(23) An r-expression is A-free (in the domain of the head of its maximal chain)

The situation is more complicated, however. Suppose that the empty operator is deleted by linking to *John* for interpretation of the matrix clause so that the chain (*John*, t) is reinterpreted as a single argument chain, just as in the normal case of raising. Then, no violation of the Condition (C) will arise here.

It may be able to take a more economy-based approach. Suppose that every operation in the course of derivation should produce a legitimate object so as to make the derivation economical. Movement of an argument to an argument position just extends the argument in terms of chain. Movement of an element to an operator position creates an operator-variable construction. Movement of an operator in an operator position to another operator position just extends the operator-variable construction. However, movement of an operator to a non-operator position does not yield any legitimate object. Thus the movement of *John* from the CP edge to the

matrix position in (20) counts as an uneconomical operation, blocking the derivation. Afterward the interpretive procedure could delete the intermediate operator to recreate a legitimate object. But it would be too late. The economy calculation is performed more locally at every step in the course of derivation.

6. Subject-object asymmetries

6.1. Surprising asymmetry

There is a well-known subject-object asymmetry, which is noted by Pesetsky (1984) as a "surprising asymmetry." The following examples and their judgments are due to Rizzi (1990):[2]

(24) a. ?? who do you wonder [whether we can help t]
b. ?? who do you wonder [whether we believe [we can help t]
c. ?* who do you wonder [whether we believe [t can help us]

All the examples are expected to be degraded because of is a *wh*-island violation. However, the contrast between (24b) and (24c) is surprising. Compare these examples with the following pair of fully acceptable examples:

(25) a. who do you think [we believe [we can help t]
b. who do you think [we believe [t can help us]

There is no asymmetry found here.

The derivations of the examples in (25) are schematically illustrated in the following, where C stands for an empty variant of *that*:

(26) $[_{CP^3}$ who$_i$ $[_{CP^2}$ t_i" C^2 $[_{CP^1}$ t_i' C^1 t_i]]]

Interpretation proceeds as follows:

(27) a. $[_{CP^1}$ O_i^1 C^1 t_i]
b. $[_{CP^2}$ O_i^2 C^2 $[_{CP^1}$ Ø C^1 t_i]]
c. $[_{CP^3}$ who$_i$ $[_{CP^2}$ Ø C^2 $[_{CP^1}$ Ø C^1 t_i]]]

When the derivation reaches the CP^2 phase, CP^1 undergoes interpretation. The trace t_i' is interpreted as an empty operator O_i^1, as in (27a). When the derivation reaches the CP^3 phase, CP^2 undergoes interpretation. The trace t_i" is interpreted as an empty operator O^2. O_i^1 is linked to O^2 and deleted so that a new operator-variable

---

[2] Following Rizzi (1982), Chomsky (1986b) claims that examples like (24c) are more degraded than examples like (24a), and discusses the relevant tense effect to compare English with Italian. Basically on the same factual assumption, Frampton (1990) gives more data concerning the tense effect. I will keep away from this matter, and continue to assume, with Rizzi (1990), Lasnik and Saito (1992) and others, that there is no significant difference in status between (24a) and (24b) for the purpose of this paper.

construction (O_i^2, t_i) is created, as in (27b). Finally, CP^3 undergoes interpretation. O_i^2 is linked to *who* and deleted so that a new operator-variable construction (*who*$_i$, t_i) is created, as in (27c). Here there is no reason to think that there is a difference between the subject and object traces.

The derivations of (24b) and (24c) yield the following:

(28) $[_{CP^3}$ who$_i$ $[_{CP^2}$ t_i" whether $[_{CP^1}$ t_i' C^1 t_i]]]

Interpretation proceeds as follows:

(29) a. $[_{CP^1}$ O_i^1 C^1 t_i]

b. $[_{CP^2}$ Ø C^2 $[_{CP^1}$ O_i^1 C^1 t_i]]

c. $[_{CP^3}$ who$_i$ $[_{CP^2}$ Ø C^2 $[_{CP^1}$ Ø C^1 t_i]]]

When the derivation reaches the CP^2 phase, CP^1 undergoes interpretation. The trace t_i' is interpreted as an empty operator O_i^1, as in (29a). When the derivation reaches the CP^3 phase, CP^2 undergoes interpretation. The trace t_i" cannot be interpreted in any way so that it is deleted, as in (29b), leading to a degraded status. Finally, CP^3 undergoes interpretation. O_i^1 is linked to *who* and deleted so that a new operator-variable construction (*who*$_i$, t_i) is created, as in (29c). Just as in the case of (27), the original trace is bound by the empty operator in CP^1, and that operator is linked to an element outside CP^1. In this respect there is no way to differentiate the subject trace from the object trace in this case, either.

Now the situation makes it reasonable to suspect that there is something wrong with the interpretive process illustrated in (29). Note that that the linking of the empty operator in (29) is not to something inside CP^2 but to something in the CP^3 edge. The crucial difference between (27) and (29) seems to be whether the linking is done between adjacent phases or not.

The linking of the empty operator in CP^1 involves the deletion of that operator. This is a reinterpretation of CP^1, because it changes the already obtained interpretation of CP^1. However, this kind of reinterpretation could be argued to be impossible, given the principle (4), repeated below:

(4) Interpretation/evaluation for PH_1 is at the next relevant phase PH_2, where PH_1 is strong and PH_2 is the next highest strong phase.

This principle requires that the interpretation of CP^1 should be done at the CP^2 phase, but at the same time it requires that CP^2 or any larger structure should not be subject to interpretation yet. Therefore, if everything is left as it is, the linking from CP^1 will be impossible, requiring the interpretation of a larger structure containing CP^1. More generally, any linking will be impossible as far as it involves deletion.

Note, however, that in order to complete the interpretation of the entire structure of CP^2, it is necessary to incorporate the already obtained interpretation of CP^1, so as to combine it with the interpretation of the residue of CP^2. Suppose further that

reinterpretation of CP^1 counts as part of interpretation of CP^2 as far as it is done when CP^1 is incorporated. Then the reinterpretation of CP^1 will be possible only when CP^2 is subject to interpretation. Thus the linking from CP^1 is possible to CP^2 but not to CP^3, because CP^1 is incorporated to CP^2, not to CP^3.

Under this locality of derivation/interpretation, the examples (24b) and (24c) must have a different way of derivation. It is possible for the original trace in CP^1 to be directly linked to who in CP^3, as in the case of simple *wh*-island examples like (24a), without being mediated by an intermediate empty operator.

A conceivable process of interpretation is as follows:

(30) a. $[_{CP^1}\ \emptyset\ C^1\ t_i\ ]$

b. $[_{CP^2}\ \emptyset\ C^2\ [_{CP^1}\ \emptyset\ C^1\ t_i\ ]]$

c. $[_{CP^3}\ who_i\ [_{CP^2}\ \emptyset\ C^2\ [_{CP^1}\ \emptyset\ C^1\ t_i\ ]]]$

For interpretation of CP^1, the trace t_i' in the CP^1 edge is deleted in the same way as in a *wh*-clause, as in (30a). For interpretation of CP^2, the trace t_i" in the CP^2 edge is deleted, as in (30b). Finally, for interpretation of CP^3, the original trace t_i in CP^1 is linked to *who* to yield the legitimate operator-variable construction (who_i, t_i), as in (30c). The linking of the original trace to the matrix *who* does not change the interpretation of CP^1, because it does not delete that trace. Although the trace is bound by *who* to yield an operator-variable construction in CP^3, it is still free in CP^1. Therefore, there is no reinterpretation of CP^1 involved here, so that there is no violation of the principle (4). However, a problem is that this interpretive process is the same as that of a double *wh*-island example like (10) in the respect that it involves double trace-deletion. Therefore, the examples (24b) and (24c) should be degraded as much as the example (10). But at least (24b) is as good as the single *wh*-island example (24a).[3]

There is still another possible derivation:

(31) a. $[_{CP^1}\ O_i'\ C^1\ t_i\ ]$

b. $[_{CP^2}\ \emptyset\ C^2\ [_{CP^1}\ \emptyset\ C^1\ t_i\ ]]$

c. $[_{CP^3}\ who_i\ [_{CP^2}\ \emptyset\ C^2\ [_{CP^1}\ \emptyset\ C^1\ t_i\ ]]]$

For interpretation of CP^1, the intermediate trace is interpreted as an empty operator in the same way as in (29), as in (31a). For interpretation of CP^2, not only the intermediate trace in the CP^2 edge but also the intermediate trace in CP^1, which was identified as an empty operator for interpretation of CP^1, is deleted under the

---

[3] Chomsky (1986b) treats such an example as (24b) as a double violation basically in the same way as a double *wh*-island example such as (10), observing no significant contrast between the two. See also the previous footnote. I am on a different factual assumption throughout this paper.

recoverability condition, as in (31b). The deletion of the latter causes a reinterpretation of CP^1, but this reinterpretation counts as part of the interpretation of CP^2 because it is done when the already obtained interpretation of CP^1 is incorporated into CP^2, satisfying the principle (4). For interpretation of CP^3, the original trace is linked to the matrix *who* in the same way in (30), as in (31c).

Note that the two intermediate traces are deleted for interpretation of CP^2 at the CP^3 phase in the case of (31), whereas they are deleted separately for interpretation of CP^1 at the CP^2 phase and for interpretation of CP^2 at the CP^3 phase in the case of (30). In the former case just one phase is relevant, whereas two phases are relevant in the latter. Suppose that the cumulative effect of the degrading trace-deletion is determined not in terms of the number of the deleted traces, but in terms of the number of the phases where the deletion is carried out. Then, with respect to the degrading by trace-deletion the examples (24b) and (24c) will be on a par with (24a) rather than (10).

In all *wh*-island cases of (24) the original trace is linked directly to the matrix *wh*-phrase. Then the question is, why is there a further degrading when the trace is a subject?

This recalls what Chomsky (1981) refers to as RES(NIC), based on Kayne's (1981) insight. The following examples are from Lasnik and Saito (1992):

(32) a. who_i t_i said (that) John bought what

b. ?who_i t_i said (that) who bought the book

The in-situ *wh*-phrases in (32) are interpreted as a free variable in the embedded clause, so that they must be linked to the matrix *wh*-phrase in terms of absorption. We find another type of subject-object asymmetry here. In the relevant respect, we can regard RES(NIC) as stating that the linking of a variable from a subject position leads to a degrading of acceptability. Seeking for an explanation for RES(NIC) is far beyond the scope of this paper. For the present purpose, it is sufficient to suggest that the surprising subject-object asymmetry under consideration may be reduced to RES(NIC).

6.2. Empty operator constructions

It has been noted in the literature that a subject-object asymmetry of a similar kind is found in empty operator constructions. Consider the following examples of the parasitic gap construction:[4]

(33) a. who did you kiss t [before you talked to e]

b. ?? who did you kiss t [before you said [you loved e]]

c. ?* who did he kiss t [before you said [e loved him]]

---

[4] See Taraldsen (1981), Chomsky (1982, 1986b), Cinque (1990) and Frampton (1990), among

There is a contrast between (33b) and (33c), revealing a subject-object asymmetry. We also find another contrast between (33a) and (33b), which is interesting in light of the absence of such a contrast in the case of the *wh*-island examples in (24).

Chomsky (1986b) argues that the parasitic gap construction involves the movement of an empty operator within the adjunct clause. Thus, the examples in (33) are derived as in the following, assuming for concreteness that *before* is a preposition taking TP as its complement:

(34) a. who$_i$ did you kiss t_i [$_{PP}$ O_j before [$_{TP}$ you talked to t_j]]

b. who$_i$ did you kiss t_i [$_{PP}$ O_j before [$_{TP}$ you said [$_{CP}$ t_j' C you loved t_j]]]

c. who$_i$ did he kiss t_i [$_{PP}$ O_j before [$_{TP}$ you said [$_{CP}$ t_j' C t_j loved him]]]

Here the empty operator is not a reinterpretation of an intermediate trace, but an independent element "base-generated" in the position of the parasitic gap.

Compare the example (33a) with the following example of extraction from an adjunct:

(35) ??who did you kiss Mary [before you talked to t]

The fact is not crystal-clear here, but (35) is clearly worse than (33a), which shows that the adjunct PP is an extraction island. Within the present framework, we may assume that the adjunct PP, just like the *wh*-clause, constitutes a strong phase.[5] The extraction should be through the edge of the PP, as in (36):

(36) who$_i$ did you kiss Mary [$_{PP}$ t_i' before [$_{TP}$ you talked to t_i]]

The degraded acceptability of (35) follows if the trace in the PP edge should undergo deletion because *before* does not allow it to be interpreted as an empty operator.

Thus the empty operator in (34a) is not allowed when the PP is interpreted, and therefore must be deleted. However, if it is simply deleted under the interpretive process for the PP, it should yield the degraded acceptability of (35).

Suppose that the empty operator, lacking a substantial content, can undergo a syntactic deletion when it is moved to the PP edge, keeping aside the question why it must moved there in the first place. Then, the examples in (33) will be not as in (34) but as in the following:

---

others, for more data and discussion.

[5] More generally, if there is an island, it suggests that there is a strong phase, too. The inner (or negative) island, for example, suggests that *v*P or something around there works as a strong phase at least in that case.

(37) a. who$_i$ did you kiss t_i [$_{PP}$ Ø before [$_{TP}$ you talked to t_j]]

b. who$_i$ did you kiss t_i [$_{PP}$ Ø before [$_{TP}$ you said [$_{CP}$ t_j' C you loved t_j]]]

c. who$_i$ did he kiss t_i [$_{PP}$ Ø before [$_{TP}$ you said [$_{CP}$ t_j' C t_j loved him]]]

When PP in (37a) is interpreted, the empty operator has been already deleted. Just as in the case of (36), the original trace is interpreted as a free variable in PP, and is linked to the matrix *who* when the matrix clause is interpreted later. Assuming that syntactic deletion as well as movement is costless, no degrading effect is produced on the example (33a). Note that the example (35) cannot escape from the degrading effect because the intermediate trace is not subject to syntactic deletion, assuming that syntactic operations generally do not apply to traces.

As for (37b) and (37c), under the principle (4) it is impossible to interpret the intermediate trace in the embedded CP as an empty operator and link it to the matrix *who*, since there will be an intervening phase, namely PP. Therefore, this intermediate trace must be deleted without being linked to the matrix. It may be deleted when the embedded CP is interpreted. Or it may be interpreted as an empty operator at that time and deleted when the adjunct PP is interpreted. In either way, it will be just one phase that involves interpretive deletion with a degrading effect, the empty operator in the PP edge being syntactically deleted. This explains why the example (33b) has the degraded status found in the case of extraction from an island. The more degraded status of the example (33c) is again reduced to RES(NIC).

Finally, consider another empty operator construction, namely the *tough*-construction. This construction reveals the same subject-object asymmetry as the parasitic gap construction, which is discussed by Schachter (1981), Stowell (1986), Cinque (1990), and others. Cinque attributes the following examples to Schachter:

(38) a. ?? Mary is hard for me to believe John kissed

b. ?* Mary is hard for me to believe kissed John

It is generally assumed that an empty operator is moved to the edge of the infinitival CP, yielding the following:

(39) a. Mary is hard [O_i for me to believe [t_i' C John kissed t_i]]

b. Mary is hard [O_i for me to believe [t_i' C t_i kissed John]]

If the empty operator is syntactically deleted as in the case of the parasitic gap construction, we will have (40) rather than (39):

(40) a. Mary is hard [Ø for me to believe [t_i' C John kissed t_i]]

b. Mary is hard [Ø for me to believe [t_i' C t_i kissed John]]

Just as in the case of (37b) and (37c), the intermediate trace in the most deeply embedded CP must be deleted without being linked in the interpretive process, degrading the acceptability. The original trace is linked to the matrix subject, not to an operator, resulting in such a situation as is found in the raising construction. Again the

subject-object asymmetry here is reduced to RES(NIC).

At the moment it is unclear why the empty operator must be deleted in the *tough*-construction. I will leave it to the future research, in hope that the present work will lead to a deeper understanding of this and related constructions.

References

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, MA.

Chomsky, N. (1986a) *Knowledge of Language*, Praeger, New York.

Chomsky, N. (1986b) *Barriers*, MIT Press, Cambridge, MA.

Chomsky, N. (1995) *The Minimalist Program*, MIT Press, Cambridge, MA.

Chomsky, N. (1998) *Minimalist Inquiries: the Framework*, *MIT Occasional Papers in Linguistics* 15, MITWPL.

Chomsky, N. (1999) *Derivation by Phase*, *MIT Occasional Papers in Linguistics* 18, MITWPL.

Cinque, G. (1990) *Types of A'-Dependencies*, MIT Press, Cambridge, MA.

Frampton, J. (1990) "Parasitic Gaps and the Theory of *Wh*-Chains," *Linguistic Inquiry* 21, 49-77.

Frampton, J. (1999) "The Fine Structure of Wh-movement and the Proper Formulation of the ECP," *The Linguistic Review* 16, 43-61.

Kayne, R. (1981) "Two Notes on the NIC," *Theory of Markedness in Generative Grammar*, eds. A. Belleti, L. Brandi and L. Rizzi, Scuola Normale Superiore di Pisa, Pisa.
[Reprinted in Kayne (1984).]

Kayne, R. (1984) *Connectedness and Binary Branching*, Foris, Dordrecht.

Koizumi, M. (1995) *Phrase Structure in Minimalist Syntax*, Doctoral dissertation, MIT.

Lasnik H. and M. Saito (1992) *Move α: Conditions on Its Application and Output*, MIT Press, Cambridge, MA.

Oka, T. (1995) "Fewest Steps and Island Sensitivity," *MIT Working Papers in Linguistics 27: Papers on Minimalist Syntax*, 189-208, MITWPL.

Pesetsky, D. (1984) "Extraction Domain and a Surprising Subject/Object Asymmetry," *GLOW Newsletter* 12 58-60.

Reinhart, T. (1981) "A Second COMP position," *Theory of Markedness in Generative Grammar*, eds. A. Belleti, L. Brandi and L. Rizzi, Scuola Normale Superiore di Pisa, Pisa.

Rizzi, L. (1982) *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, L. (1990) *Relativized Minimality*, MIT Press, Cambridge, MA.

Rudin, C. (1988) "On Multiple Questions and Multiple WH Fronting," *Natural Language and Linguistic Theory* 6, 1-24.

Taraldsen, T. (1981) "The Theoretical Interpretation of a Class of Marked Extractions," *Theory of Markedness in Generative Grammar*, eds. A. Belleti, L. Brandi and L. Rizzi, Scuola Normale

Superiore di Pisa, Pisa.

Schachter, R. (1981) "Lovely to Look At," *Linguistic Analysis* 8, 431-448.

Stowell, T. (1986) "Null Antecedents and Proper Government," *Proceedings of the Sixteenth Annual Meeting, NELS*, GLSA, University of Massachusetts, Amherst.

Fukuoka University of Education

e-mail: okatoshi@fukuoka-edu.ac.jp