Weakest Islands

MIKI OBATA

Abstract

This paper focuses on some properties of wh-islands in terms of finiteness. Wh-extraction out of infinitival wh-clauses is much better than the one out of finite wh-clauses, as reported so far. The amelioration of wh-island effects in infinitival clause, which I term "weakest island" effects in this paper, has been captured as "Tensed-island". However, a close look at some data seems to show that weakest island effects do not purely depend on finiteness. In this paper, I present supporting evidence for this view, and I argue that weakest island effects are selectively observed in A-to-A' movement across an island. Meanwhile, A'-to-A' movement across an island in either finite or infinitival clause equally makes a sentence bad. Also, I will show that at least some existing approaches to locality of movement fail to capture these generalizations and that Pesetsky and Torrego (2001, 2002, 2004) give us a clue to understand them in terms of "T-relatedness".

Introduction

Locality of movement is one of the central topics in generative grammar. Since Ross (1967) reported several types of island constraints, "how to derive those island effects" has been discussed from various points of view: Subjacency (Chomsky, 1973), Connectedness (Kayne, 1983), Barriers (Chomsky, 1986), Relativized Minimality (Rizzi, 1990), Locality (Manzini, 1992) among others.

In this paper, I focus on so-called "Tensed-island constraint". Generally speaking, infinitival T tends to ameliorate extractions out of islands as in the following examples:

(1) a. ??Which book did John ask Mary [when he should buy t_{which} t_{when}]?
b. Which book did John ask Mary [when to buy t_{which} t_{when}]?

(2) a. ??What did John ask Mary [when he should buy t_{what} t_{when}]?
b. What did John ask Mary [when to buy t_{what} t_{when}]?

In (1) and (2), an extraction out of infinitival wh-clauses is much better than the one out of finite wh-clauses, despite the fact that wh-islands are involved in both cases. Let us call cases like (1b) and (2b) "weakest" islands, in that the island effects seem to be much weaker than any other ones. The contrast observed in (1) and (2) implies that properties of T have some effect on deciding whether or not movement is allowed.

 University of Maryland Working Papers in Linguistics 14, ed. N. Kazanina, U. Minai, P. Monahan and H. Taylor, pp. 143-162. College Park, MD: UMWPiL.

However, there are some data which seem to show that finiteness is not the only factor giving rise to the contrast in (1) and (2). In this paper, I present supporting evidence for this view, and I argue that weakest island effects are selectively observed in A-to-A' movement across an island. Meanwhile, A'-to-A' movement across an island in either finite or infinitival clause equally makes a sentence bad. Moreover, I suggest that this observation can be induced from the Case-system proposed in Pesetsky and Torrego (2001, 2002, 2004).

This paper is organized as follows: Section 1 briefly introduces two existing approaches to weakest islands. Section 2 points out that weakest island effects seem to be observed only in a certain step of movement. Section 3 captures the generalization obtained in section 2 in terms of "T-relatedness". Section 4 concludes our study.

1 Locality of Movement: How to Derive Island Effects

Since Ross (1967), locality of movement has been discussed from various points of view. Especially, the central concern is how to derive the same effects as Ross's islands without actually referring to "islands". In this section, some existing approaches to this topic are briefly reviewed, focusing on the relation between islands and finiteness.

1.1 Chomsky's (1986) Barriers *Barriers* proposed in Chomsky (1986) is one of the intriguing works on locality of movement. Unlike the bounding node type approach suggested in Chomsky (1973), every maximal projection potentially becomes a barrier for movement in this system. That is, *Barriers* can derive island effects without specifying categories of bounding node. On the other hand, the system seems to have troubles to dealing with Tensed-island effects. A set of basic assumptions of *Barriers* is not equipped with a device to distinguish finite clauses and infinitival clauses. Therefore, the contrast observed in (1) and (2) is problematic as follows[1]:

(3) a. ??Which book did $[_{IP}$ John $[_{VP}$ t_{which} $[_{VP}$ ask Mary $[_{CP}$ when $[_{IP}$ he should $[_{VP}$ t_{which} $[_{VP}$ buy t_{which} $t_{when}]$?

b. Which book did $[_{IP}$ John $[_{VP}$ t_{which} $[_{VP}$ ask Mary $[_{CP}$ when $[_{IP}$ to $[_{VP}$ t_{which} $[_{VP}$ buy t_{which} $t_{when}]$?

[1] In (3), I abstract away from movement of "when" for making the discussion simpler.

In (3a), "which book" is first adjoined to the embedded VP. Since the embedded [Spec,CP] is already filled with another wh-phrase, "which book" is adjoined to the matrix VP at the second step of movement.[2] At this point of derivation, the moved wh-phrase acrosses 1 barrier, because IP is not L-marked with C and, CP, which immediately dominates IP, becomes a barrier. In terms of Subjacency, the system predicts that (3a) has a weak island violation. On the other hand, (3b) is better than (3a) despite the fact that they are structurally the same. The only different point is finiteness in the embedded clauses. This point becomes an obstacle in *Barriers* framework.

In order to capture the contrast observed in (3), Chomsky stipulates that tensed IP is an inherent barrier (possibly weak) to wh-movement, this effect being restricted to the most deeply embedded tensed IP (Chomsky 1986:37). With this stipulation, the second step of movement in (3a) crosses 2 barriers at once: the lowest tensed IP and the inherited barrier CP. On the other hand, the lowest embedded clause is not tensed in (3b). Therefore, the wh-phrase crosses only 1 barrier (CP) at the second step. The system predicts the contrast in (3). Chomsky has to limit this stipulation to the "most deeply embedded" clause because the following example has to be correctly ruled in:

(4) Who did you think that John said that Bill saw?

(Chomsky 1986: 38)

If every tensed IP is an inherent barrier, each step of movement in (4) always crosses 1 barrier. As a result, the system wrongly predicts that an example like (4) is degraded despite the fact that it is perfectly acceptable. Therefore, the "most deeply embedded" clause is an important key for Chomsky's system to derive the contrast in (3).

As mentioned earlier, the underlying spirit of *Barriers* is to derive locality effects without specifying certain categories. That is, *Barriers* gives a more principled account to locality effects than the bounding node type approach. With respect to Tensed-islands, however, a specific category, the most deeply embedded IP, has to have a special status against this spirit. To get rid of that stipulation for Tensed-islands, Manzini (1992) suggests another type of approach.

[2] Chomsky (1986:32) stipulates that wh-phrases may not adjoin to IP, then moving to [Spec,CP].

1.2 Manzini's (1992) Locality Manzini (1992) also attempts to elucidate locality effects in a different way, although she partly adopts *Barriers* of Chomsky (1986). In her system, the availability of "sequence" between antecedents and traces holds the key to the solution of the locality problem. There are two kinds of sequence in her system: Categorial index dependency and Address-based dependency.[3] If either of them can be established, the system predicts that the sentence is grammatical.

Let us review how the system deals with the contrast in (1)-(2). Manzini assumes that T(ense) has the following two parametric values: [+Tns] or [-Tns]. When T is [+Tns], it has an address of its own. Meanwhile, [-Tns]T does not have an address of its own.[4]

(1) a. ??Which book did John ask Mary [when he should buy t_{which} t_{when}]?
b. Which book did John ask Mary [when to buy t_{which} t_{when}]?
(2) a. ??What did John ask Mary [when he should buy t_{what} t_{when}]?
b. What did John ask Mary [when to buy t_{what} t_{when}]?

Under this system, (1) and (2) can be properly explained. Since "which book" is an argument wh-phrase in (1), it is Case-marked and has an address. The expected address-based sequence is (which, C, T, V, C, T, V, t_{which}). Since [-Tns] does not have its own address, a uniform sequence can be obtained in (1b). Therefore, the address-based sequence can be established and the system correctly predicts that (1b) is grammatical. On the other hand, (1a) cannot establish the sequence, because [+Tns] has its own address and a uniform sequence cannot be obtained. In addition, categorial index sequences cannot be formed in (1), because one barrier (the embedded CP) is crossed. As a result, the system can successfully predict the contrast in (1). The same thing applies to (2).

In sum, both Chomsky (1986) and Manzini (1992) can capture the contrast in (1) and (2). Keeping these approaches in mind, let us view some other types of data in the next section. I will show that weakest island effects involve another factor as well as finiteness: weakest island effects are only observed in A-to-A' movement across an island, not A'-to-A' movement.

---

[3] The categorial index and the address are defined as follows:

(1) α has a categorical index if $\alpha$is lexical.
(2) α has an address if there is a head β that Case-marks α. (Manzini 1992: 38)

[4] Manzini (1992) seems to assume this parametric variation only in the non-root clause, although she does not make it explicit. Otherwise, the matrix T is supposed to wrongly block address-based sequence every time.

2 Further Observations on Weakest Islands

2.1 Wh-Islands In this section, we will further consider weakest island effects in a more complicated context. First of all, let us make sure of the basic data set:

(5) a. ??Which book did John ask Mary [when he should buy t_{which} t_{when}]?
b. Which book did John ask Mary [when to buy t_{which} t_{when}]?

The contrast between (5a) and (5b) sets up a generalization that infinitival T ameliorates wh-island violation.

However, the story is not that simple: according to the generalization sofar, all of the examples in (6) are supposed to be good.

(6) a. Which book did John decide [to try [to buy t]]?
b. ?Which book did John decide [to ask Mary [when to buy t_{which} t_{when}]]?
c. *Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]]?[5]
d. John asked Mary [when to decide t [to buy the book]].[6]

---

[5] Richards (2002:240) also shows the same kind of data. However, he develops a different analysis from this paper.

[6] Coopmans and Stevenson (1991) suggests an interesting asymmetry of wh-extraction from finite and infinitival complements: extraction of "how" across an infinitival is worse than extraction across a finite clause, as shown below.

(1) a. *I wonder how to believe [John fixed the car t].
b. *I wonder how to believe [John to have fixed the car t].
(2) a. I wonder how Mary believed [John fixed the car t].
b. I wonder how Mary believed [John to have fixed the car t].

On the other hand, extraction of object wh-phrases from the same circumstance is fine as follows:

(3) a. They wondered what to believe [John to have fixed t].
b. They wondered what to believe [John had fixed t].
(4) a. They wondered what Mary believed [John to have fixed t].
b. They wondered what Mary believed [(that) John had fixed t].

If Coopmans and Stevenson's (1991) observation is correct, the base form of (6c) could be degraded because "when" is also an ajunct wh-phrase. As showed in (6d), however, the base form is acceptable. Although I do not investigate why there is a difference "how" and "when" in this paper, it could be one of the factors that "when" is closer to an argument wh-phrase like "what" in terms of the locality of movement as mentioned in Huang (1982).

The example in (6c) is worse than (6b), despite the fact that the embedded clauses are all infinitival in both cases. The only difference is where wh-islands are located: (6b) in the lowest clause, (6c) in the second clause. Suppose that this observation is true and that a wh-island in the "second" clause makes a sentence degraded. If so, is finiteness irrelevant to the acceptability in (6)?

(7) a. Which book did John decide [he should promise that [he would buy t]]?
b. *Which book did John ask Mary [when he should decide t_{when} [he should buy t_{which}]]?

The sentence in (7b) has the same configuration as (6c), but (7b) is a tensed clause. Comparing (6c) with (7b), both of them are unacceptable regardless of finiteness. Therefore, these data seem to imply that a wh-island in the "second" clause is a crucial factor in giving rise to the difference between (6b) and (6c); however, the contrast in (1) and (2) is still problematic, because they do not have the second clause.

In sum, we have so far considered the following generalizations of weakest island effects:

(8) Generalization 1 (from (5))
Infinitival T ameliorates wh-island violation.
(9) Problematic case of (8): (6c)
(10) Generalization 2 (from (6c))
A wh-island in the "second" clause makes a sentence degraded.
(11) Problematic cases of (10): (1), (2)

In conclusion, in order to describe all of the data presented here, both of the generalizations in (8) and (10) are necessary. Either of them is not sufficient.

Again, let us consider the same kinds of data from a different point of view. ((13) = (6), (15a) = (7b))

(12) Finite-infinitival
a. *Which book did John ask Mary [when he should decide t_{when} [to buy t_{which}]]?
b. ?Which book did John decide [he should ask Mary [when to buy t_{which} t_{when}]]?
(13) Infinitival-infinitival
a. *Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]]?

b. ?Which book did John decide [to ask Mary [when to buy t_{which} t_{when}]]?

(14) Infinitival-finite

a. *Which book did John ask Mary [when to decide t_{when} [he should buy t_{which}]]?

b. *Which book did John decide [to ask Mary [when he should buy t_{which} t_{when}]]?

(15) Finite-finite

a. *Which book did John ask Mary [when he should decide t_{when} [he should buy t_{which}]]?

b. *Which book did John decide [he should ask Mary [when he should buy t_{which} t_{when}]]?

Let us focus on the positions of clauses with wh-islands. In (a) examples, a wh-island is in the second clause, while in (b) examples, it is in the lowest clause. Notice that (b) examples include both the original position (Case-marked position) of "which book" and a wh-island within the same clause. That is, "which book" crosses a clause with a wh-island when it moves from an A-position to an A'-position as in the following configuration:

(16) [which book . . .[$_{CP}$ . . . [$_{CP}$ when . . . t]]]

Also, comparing (12b)/(13b) with (14b)/(15b), it is clear that this movement is sensitive to finiteness: (12b)/(13b) is acceptable where the lowest clause is infinitival, while (14b)/(15b) is unacceptable where the lowest clause is finite.

Next, as for (a) examples, they do not include a wh-island and the original position of "which book" within the same clause. From the same point of view as the discussion in (b) examples, "which book" crosses a wh-island when it moves from an A'-position ([Spec,CP]) to an A'-position ([Spec,CP]) as in the following configuration:

(17) [which book . . .[$_{CP}$ when . . . [$_{CP}$ . . . . t]]]

Let us compare (12a)/(15a) with (13a)/(14a). Both of the examples are unacceptable despite the fact that finiteness is different in each case. In other words, movement from an A'-position to an A'-position is not sensitive to finiteness, in contrast to A-to-A' movement in cases like (b).

Further, if an adjunct wh-phrase is extracted out of a wh-island, the examples are unacceptable regardless of finiteness as in (18).

(18) a. *How did John wonder where he should buy the book t_{where} t_{how}?
b. *How did John wonder where to buy the book t_{where} t_{how}?

Since adjunct wh-phrases are in A'-positions from the beginning, every step of movement is an A'-position to an A'-position. As mentioned in the discussion of (12a)-(15a), this type of movement is not sensitive to finiteness.[7] Therefore, both of the examples become unacceptable in (18).

In sum, we obtain the following generalization:

(19) Generalization 3 (from (12)-(18))
a. A-to-A' movement crossing an island is sensitive to finiteness.
(infinitival clause: acceptable, finite clause: unacceptable)
b. A'-to-A' movement crossing an island is NOT sensitive to finiteness.
(infinitival clause: unacceptable, finite clause: unacceptable)

Under this generalization, the data we have considered so far can be successfully described.

2.2 An Implication for Successive Cyclic A'-movement In the above section, we have discussed weakest island effects in terms of both finiteness and A/A' distinctions. As a result, the generalization in (19) is gained and successfully describes the data presented here. In addition, (19) has an interesting implication for the potential landing site of successive cyclic A'-movement. Let us think about this problem before going into further discussion on weakest island effects.

[7] To be honest, there are some differences in acceptability between adjunct cases like (18) and argument cases like (12a)-(15a). Although both of the cases are unacceptable, the adjunct cases are worse than the argument cases. However, the generalization presented in (19) cannot make any distinction between them. I do not have a good answer to this problem. Therefore, I put it aside in this paper.

According to Chomsky (2000, 2001a,b), wh-phrases have to go through every [Spec,CP] and [Spec,vP] on the way to the edge of [+Q]C under the Phase Impenetrability Condition.[8] Every phase head (v, C) has EPP features to attract elements to the edge positions. Therefore, [Spec,CP] and [Spec,vP] are potential landing sites of A'-movement. This kind of idea has been maintained since Chomsky (1986).

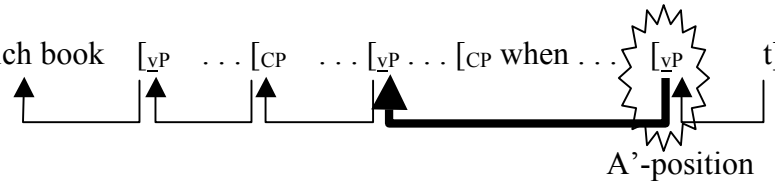
On the other hand, our generalization of weakest island effects tells us that the potential landing site of wh-movement is only [Spec,CP], not [Spec,vP]. Let us recall the generalization in (19): A-to-A' movement crossing an island is sensitive to finiteness (→(20)), while A'-to-A' movement crossing an island is not sensitive to finiteness (→(21)).

(20) [which book . . .[$_{CP}$ . . . [$_{CP}$ when . . . t]]] (→(12b)-(15b))

(21) [which book . . .[$_{CP}$ when . . . [$_{CP}$ . . . . t]]] (→(12a)-(15a))

Consider movement in (20). If "which book" stops at [Spec,vP] as suggested in Chomsky's work, the step of movement crossing a wh-island is no longer A-to-A' movement, but A'-to-A' movement as follows[9]:

(22) [which book [$_{vP}$ . . . [$_{CP}$ . . . [$_{vP}$ . . . [$_{CP}$ when . . . [$_{vP}$ t]]]



If a derivation like (22) is assumed, every movement across a wh-island is A'-to-A' movement, and we cannot maintain the distinction between (20) and (21). From a viewpoint of weakest island effects, it can be concluded that only [Spec,CP] is available as a potential landing site of wh-movement. Putting this conclusion to more recent terms, only CP, not vP, is a phase.[10] Also, Du Plessis (1977), McCloskey (2002) and others show that footprints of wh-movement are visible in some languages, and that they are observed only at [Spec,CP]. Therefore, they seem to support the present view that wh-movement stops at only [Spec,CP], not [Spec,vP].

---

[8] Phase Impenetrability Condition (Chomsky 2000:108): In phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

[9] I am assuming that CP has only one spec-position, not multiple-spec, in English.

[10] Simpson and Wu (2002) and Obata (to appear) also reaches the same conclusion.

2.3 Section Summary Let us summarize what we have discussed so far. In this section, we have considered the relation between weakest island effects and A/A' distinctions. Based on wh-island examples, we obtained the generalization in (19): A-to-A' movement crossing an island is sensitive to finiteness, while A'-to-A' movement crossing an island is not sensitive to finiteness. This generalization shows that weakest island effects depend on not only finiteness but also A/A' distinction.

Looking back to the approaches in Chomsky (1986) and Manzini (1992), their systems would fail to deal with the data presented here, because they do not take account of A/A' distinction in the sense of (19).

(23) Chomsky (1986:37):
Tensed IP is an inherent barrier (possibly weak) to wh-movement, this effect being restricted to the most deeply embedded tensed IP.

(24) Manzini (1992:118):
[+Tns] blocks the address-based sequence between wh-phrase and trace. (because [+Tns] T has an address of its own, while [-Tns] does not.)

Problematic Cases:

(25) Chomsky (1986):
(1a) ??Which book did John ask Mary [when he should buy t_{which} t_{when}]?
→crossing 2 barriers
(6c) *Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]?
→crossing 1 barrier

(26) Manzini (1992):
(6c) *Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]]?

Under Chomsky's system, (6c) involves movement that crosses only 1 barrier. However, the example is much worse than what the system predicts. On the other hand, (1a) involves movement that crosses 2 barriers despite the fact that it is better than (6c). Therefore, this system fails to capture the weakest island effects considered here. In Manzini's (1992) approach, [+Tns] in non-root clauses blocks an address-based sequence. Under this system, (6c) is problematic, because the embedded clauses are all infinitival ([-Tns]) but the example is unacceptable for some reason. Therefore, this system also suffers from the data discussed in this paper. In order to capture these data, we need to refer to A/A' positions as well as finiteness. In section 3, I attempt an account of weakest island effects from a different point of view.

3 An Alternative Approach to Weakest Islands

As discussed so far, weakest island effects seem to have something to do with A/A' distinction. Before going through the main subject in this section, let us make sure of several basic points in the following sections.

3.1 Properties of Movement: A/A' distinctions First, we need to consider A/A' distinctions, because those distinctions perform an important factor to describe weakest island effects. What distinguishes A-movement and A'-movement? The landing sites decide which types of movement are taking place: if the landing site is an A-position, that movement is A-movement, or if an A'-position, it is A'-movement. On the other hand, the starting points have varieties. A'-movement can start from either an A-position or an A'-position, while A-movement can start only from an A-position.

In the case of successive cyclic movement, the "first" step/landing site of movement makes a distinction between A-movement and A'-movement. At the same time, only the chain formed by the first step of movement holds information on the original position of a moved wh-phrase in that all other steps start from intermediate landing sites ([Spec,CP]) in the case of A'-movement. Moreover, only in the first step of movement, movement from an A-position to an A'-position is allowed. In this sense, it seems that the first step of movement has a special status.

Let us recall the generalization of weakest island effects we obtained in the previous section. The generalization also focuses on the first step of movement, especially the position from which movement originates. If movement starts from an A'-position, it does not care about finiteness. On the other hand, if movement starts from an A-position, it is sensitive to finiteness. That is, the finiteness sensitivity is observed in A-to-A' movement. This type of movement is allowed only in the first step of successive cyclic A'-movement as mentioned in the above paragraph. The first step is an important factor also in weakest island effects.

3.2 T-Feature and T-Related Elements In this section, let us consider how to relate finiteness with movement. Generally speaking, A-movement is more sensitive to finiteness than A'-movement, because A-movement out of finite clause is impossible, while A'-movement out of a finite clause is possible:

(27) a. John seems [t to be genius].
b. It seems [that John is genius].
c. *John seems [that t is genius].

(28) a. John is easy [for Bill to please t].
b. It is easy [that Bill pleases John].

c. *John is easy [that Bill pleases t].

(29) What do you think [that John bought t]?

As shown in (27c) and (28c), A-movement from a finite clause makes examples unacceptable. On the other hand, A'-movement from a finite clause is perfectly acceptable as in (29). As already discussed, however, in some limited environment, A'-movement is also sensitive to finiteness. If our generalization is on the right track, only the A-to-A' movement crossing an island cares about finiteness of the clause that it crosses. What is common between A-movement and A-to-A' movement is that both types of movement start from an A-position:

(30) If a step of movement starts from an A-position, that step is sensitive to finiteness.

In the case of successive cyclic A-movement, every step of movement starts from an A-position. Therefore, it is always sensitive to finiteness. In the case of successive cyclic A'-movement, on the other hand, only the first step of movement could originate from an A-position, then it might be sensitive to finiteness. Further, movement of a wh-adjunct and movement from an intermediate landing site do not involve A-positions. Therefore, those kinds of movement do not show sensitivity to finiteness. (30) can capture some properties of finiteness sensitivity observed in A-movement and A'-movement.

The next question is how to derive the property in (30). Let us think about the way of Case-assignment. As widely assumed, nominative Case is assigned by T, while accusative Case is assigned by v. For example, in Chomsky (2000, 2001a,b), Case-assignment/feature valuation takes place by means of Agree between v and an object (for accusative Case) and between T and a subject (for nominative Case). Under this approach, Case-assignment is done by two different heads. On the other hand, Pesetsky and Torrego (2001, 2002, 2004) argue that structural case on DP is an uninterpretable instance of T-feature and that the Case Filter of the Government-Binding Theory (Chomsky, 1981) can be understood as the following Argument-Tense Condition:

(31) Argument must bear T.

(Pesetsky and Torrego, 2002:9)

That is, both a subject and an object equally have T-features, and Case-assignment is accomplished through Agree between T and a subject DP or an object DP. The single head T takes care of both accusative Case and nominative Case. This point is different from Chomsky's system mentioned above. Under this system, DPs are associated with T for the Case. In this sense, it can be thought that elements at A-positions are "T-related" elements. Note that adjuncts are not T-related elements, because they have no T-features (Case) and are not associated with T. In terms of T-relatedness, let us posit the following working hypothesis with respect to finiteness:

(32) Working Hypothesis
T-related elements (elements at A-positions) are sensitive to finiteness.

If the working hypothesis in (32) is empirically supported, T-relatedness might be able to capture argument/adjunct (A/A') asymmetry observed in finiteness.

Moreover, Pesetsky and Torrego (2001) claims that C also has an uninterpretable T-feature. In their system, this assumption is an important factor in explaining the availability of T-to-C movement. Also in this paper, I adopt this assumption.

In the next section, we will capture the generalization of weakest island effects based on the discussion in 3.1 and 3.2.

3.3 Weakest Islands: Defective Intervention Constraint and T-feature First of all, let us make sure of the paradigms we considered in section 2 again:

(33) Finite-infinitival
a. *Which book did John ask Mary [when he should decide t_{when} [to buy t_{which}]]?
b. ?Which book did John decide [he should ask Mary [when to buy t_{which} t_{when}]]?

(34) Infinitival-infinitival
a. *Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]]?
b. ?Which book did John decide [to ask Mary [when to buy t_{which} t_{when}]]?

(35) Infinitival-finite
a. *Which book did John ask Mary [when to decide t_{when} [he should buy t_{which}]]?
b. *Which book did John decide [to ask Mary [when he should buy t_{which} t_{when}]]?

(36) Finite-finite

a. *Which book did John ask Mary [when he should decide t_{when} [he should buy t_{which}]]?

b. *Which book did John decide [he should ask Mary [when he should buy t_{which} t_{when}]]?

In every (a) example, a wh-island is crossed in the movement from an A'-position to an A'-poisition. In every (b) example, on the other hand, it is crossed in the movement from an A-position to an A'-position. Behind these data, it seems that the following generalization exists:

(37) Generalization on Weakest Island Effects

a. A-to-A' movement crossing an island is sensitive to finiteness.
(infinitival clause: acceptable, finite clause: unacceptable)

b. A'-to-A' movement crossing an island is NOT sensitive to finiteness.
(infinitival clause: unacceptable, finite clause: unacceptable)

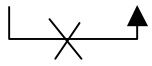
In addition, we established the following working hypothesis:

(38) T-related elements are sensitive to finiteness.

Next, we need to consider how to derive wh-island constraints. According to Chomsky (2000), wh-island effects can be explained by means of "Defective Intervention Constraint (DIC)".

(39) $\alpha > \beta > \gamma$

(40) $\alpha \quad \beta \quad \gamma$



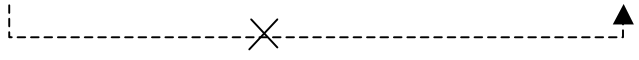
In (39), α c-commands β and β c-commands γ. In this configuration, if β and γ match the probe α but β is inactive, the relation in (40) is prohibited. The wh-island configuration is an instance of the configuration in (40):

(41) *What do you wonder where John bought?

[C you wonder [where John bought what ~~where~~]

[+Q] [+Q] (inactive) [+Q]

α β γ



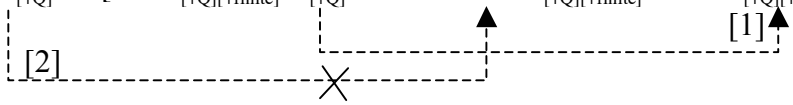
In terms of DIC, Agree between the matrix [+Q] and "what" is prohibited, because [+Q] (β) is intervening between them. As a result, the derivation crashes, and wh-island effects can be successfully captured.

Now, let us consider the following basic examples:

(42) a. ??Which book did John ask Mary [when he should buy t_{which} t_{when}]?

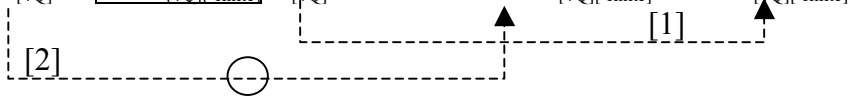
b. Which book did John ask Mary [when to buy t_{which} t_{when}]?

(43) $[C_{[+Q]}$ $[when_{[+Q][+finite]}$ $C_{[+Q]}$ which $book_{[+Q][+finite]}$ ~~$when_{[+Q][+finite]}$~~$]$



(43) is the configuration for (42a). As mentioned in the above section, it is assumed that D and C have T-features along the line of Pesetsky and Torrego (2001, 2002, 2004). Suppose that T-features on D and C include information on finiteness. D and C know whether the clause in which they are is finite or not as illustrated in (43)[11]: D[±finite], C[±finite]. First, the embedded [+Q]C agrees with "when". Then, it attracts the wh-phrase to [Spec,CP] by an EPP feature as in [1]. At this point, [+Q] on "when" is already valued. Hence, it is inactive. Next, the matrix [+Q]C enters the derivation. The matrix [+Q]C has to agree with "which book". However, DIC blocks this Agree relation because the intervening [+Q] on "when" is already valued. As a result, an example like (42a) is ruled out. Now, what is going on in a case like (42b)?

(44) $[C_{[+Q]}$ $[when_{[+Q][-finite]}$ $C_{[+Q]}$ which $book_{[+Q][-finite]}$ ~~$when_{[+Q][-finite]}$~~$]$



The same derivation as (43) takes place in (44). The only difference between (43) and (44) is finiteness in the embedded clause: [+finite] in (43), [-finite] in (44). Suppose that [-finite] has the following effect on DIC:

(45) [-finite] hides a valued features from DIC.

(46) The effect in (45) is only visible/relevant to Agree which involves T-related elements.

[11] For space limitations, [±finite] on C is omitted.

T-related elements are sensitive to finiteness as mentioned in (38). Under that working hypothesis, Agree with a T-related element can see the effect in (45), because that effect is induced by a property of finiteness. On the other hand, Agree which does not involve a T-related element cannot see the effect in (45), even if that effect comes out. In (44), "when" is marked with [-finite][+Q], and [+Q] is already valued. Hence, "when" can induce the effect in (45). The matrix C has to agree with "which book". Since "which book" is at the original position (an A-position), it is a T-related element. Agree between the matrix [+Q] and "which book" can see the effect in (45) because of (46). Under (46), (44) does not induce the violation of DIC. As a result, the system predicts that (42b) is grammatical. The same discussion applies to (33b)-(36b).

(33b) Finite-infinitival
?Which book did John decide [he should ask Mary [when to buy t_{which} t_{when}]]?

(34b) Infinitival-infinitival
?Which book did John decide [to ask Mary [when to buy t_{which} t_{when}]]?

(35b) Infinitival-finite
*Which book did John decide [to ask Mary [when he should buy t_{which} t_{when}]]?

(36b) Finite-finite
*Which book did John decide [he should ask Mary [when he should buy t_{which} t_{when}]]?

In these cases, wh-islands are crossed in the movement from an A-position to an A'-position. A wh-phrase in an A-position is a T-related element. In (33b) and (34b), "when" marked with [+Q][-finite] can induce the effect in (45). Since "which book" at the original position is a T-related element, Agree between the matrix [+Q] and "which book" can see that effect. Hence, it can escape the violation of DIC. On the other hand, "when" in (35b) and (36b) is marked with [+Q][+finite]. The effect in (45) is not induced here because of [+finite], and DIC blocks Agree between the matrix [+Q] and "which book". Hence, these examples are ruled out. This is the derivation for the first half of the generalization: (37a) A-to-A' movement crossing an island is sensitive to finiteness.

Let us go on to the second half of the generalization: (37b) A'-to-A' movement crossing an island is not sensitive to finiteness. We can derive this generalization by means of the technology suggested above.

(33a) Finite-infinitival
*Which book did John ask Mary [when he should decide t_{when} [to buy t_{which}]]?

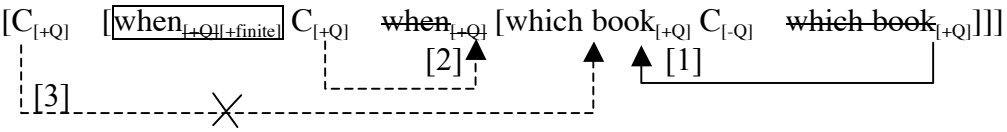
(34a) infinitival-infinitival
*Which book did John ask Mary [when to decide t_{when} [to buy t_{which}]]?

(35a) infinitival-finite
*Which book did John ask Mary [when to decide t_{when} [he should buy t_{which}]]?

(36a) finite-finite
*Which book did John ask Mary [when he should decide t_{when} [he should buy t_{which}]]?

"Which book" in (33a) through (36a) crosses a wh-island in A'-to-A' movement. Let us take a look at the derivations of 0) and (36a).

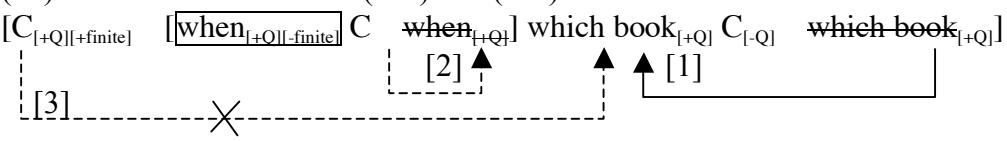
(47) The derivation for (33a) and (36a):



First, "which book" in the lowest clause moves to the edge position of CP by EPP as in [1]. And then, the second clause enters the derivation. As in [2], [+Q]C agrees with "when" and EPP attracts it to the edge position. Furthermore, the matrix clause enters the derivation. The matrix [+Q]C has to agree with "which book" at [Spec,CP] in the lowest clause. However, [+Q] on "when" in the second clause is already valued and inactive. Therefore, Agree across the inactive [+Q] violates the DIC as in [3]. "When" is marked with [+finite], not [-finite]. Hence, the effect in (45) does not come out. That is how examples like (33a) and (36a) are ruled out.

Next, let us turn to the derivations for (34a) and (35a). These examples are also unacceptable. In (34a) and (35a), the second clauses are infinitival. Only this point is different from (33a) and (36a). The following derivation takes place in (34a) and (35a):

(48) The derivation for (34a) and (35a):



The same derivation as (47) is executed. Note that "when" in the second clause is marked with [-finite], and it is already valued by Agree with C. As mentioned earlier, the effect in (45) should be observed also in this case. [-finite] hides a valued [+Q] from DIC in the same manner as the previous cases. However, there is another point to be focused on in this case. When the Agree between the matrix [+Q] and "which book" shown in [3] takes place, "which book" is already at [Spec,CP]. That is, "which book" is a non-T-related element. According to (46), only Agree which involves a T-related element can see the effect in (45). Under these assumptions, although the effect in (45) comes out, Agree of [3] cannot see that effect because a T-related element is not included in the Agree relation. Therefore, Agree of [3] violates DIC, and the system predicts that examples like (34a) and (35a) are ungrammatical.

Again, the same mechanism can appropriately rule out wh-adjunct cases as shown in (18), repeated as (49).

(49) a. *How did John wonder [where$_{\text{[+Q][+finite]}}$ he should buy the book t_{where} t_{how}]?

b. *How did John wonder [where$_{\text{[+Q][-finite]}}$ to buy the book t_{where} t_{how}]?

Adjunct wh-phrases are at A'-positions from the beginning. Therefore, every step of movement is from an A'-position to an A'-position. Even if the effect in (45) appears in (49), no one can see that effect because no T-related elements are involved in Agree relations. In fact, the embedded clause in (49b) can provide the effect in (45) because "where" is marked with [-finite]. For the reason mentioned above, (49b) cannot escape the violation of DIC. Eventually, both of the examples in (49) are ruled out by DIC.

3.4 Section Summary In this section, we have captured the generalization we obtained in section 2. One of the central problems to be tackled was how to deal with finiteness sensitivity observed in only A-to-A' movement. As for this point, a series of works by Pesetsky and Torrego gave us an important key to solving the problem. Based on their work, the following working hypothesis was established:

(50) T-related elements are sensitive to finiteness.

Under (50), moreover, the following two points were suggested:

(45) [-finite] hides a valued feature from DIC.

(46) The effect in (45) is only visible to Agree which involves T-related elements.

Under these assumptions, our system successfully predicted the paradigms shown in (33) through (36) and (49).

As discussed so far, it can be concluded that the discussion in this paper supports the view of Pesetsky and Torrego (2001, 2002, 2004): structural case on DP is an uninterpretable instance of T. The existence of T-feature on D enables us to capture A/A' asymmetry in weakest island effects.

4. Conclusion and Some Remaining Problems

In this paper, we have discussed some properties of weakest island effects. We obtained the generalization as follows: A-to-A' movement crossing an island is sensitive to finiteness, while A'-to-A' movement crossing an island is not sensitive to finiteness. Moreover, I showed some existing approaches to weakest islands, and that they did not straightforwardly capture these properties. In section 3, we tried to give an alternative account to capture the generalization on weakest island effects. By means of a device suggested in Pesetsky and Torrego (2001, 2002, 2004), I established a working hypothesis: T-related elements are sensitive to finiteness. Under the hypothesis, we derived the data presented in section 2. The system proposed here could be a possible approach to weakest island effects.

However, there are some problems to be considered. One of them concerns a problem with Cyclicity. In the discussion here, long-distance Agree relations take place especially in the case of Agree between the matrix C and the most deeply embedded wh-phrase. This type of Agree crosses several numbers of phases (vP or CP). However, Chomsky (2000, 2001a,b) argues that operations must obey Phase Impenetrability Condition (PIC). If operations do not obey PIC, Cyclicity disappears from the grammar. Moreover, the system has to stipulate the effect in (45) in order to deal with weakest island effects. These are all pending problems.

Acknowledgments

I have received various comments and suggestion from Tomohiro Fujii, Norbert Hornstein, Tomohiko Ishimori, Howard Lasnik, Chizuru Nakao, Andrew Nevins, Jairo Nunes, Hajime Ono, David Pesetsky, Norvin Richards, Yuji Takano, Koichi Takezawa, and Masaya Yoshida on several points in this paper. I also thank Stacey Conroy, Scott Fults, Rebecca McKeown and Philip Monahan for judgments on the English examples. But responsibility for the text rests entirely upon the author.

References

Chomsky, N. 1973. Conditions on Transformations. In S.R. Anderson and P. Kiparsky, (eds.) *A Festchrift for Morris Halle*. New York, NY: Holt, Rinehart, & Winston, pp.232-286.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris Publications.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka, (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.

Chomsky, N. 2001a. Derivation by Phase. In M. Kenstowicz, (ed.) *Ken Hale: A Life in Language.* Cambridge, MA: MIT Press, pp.1-52.

Chomsky, N. 2001b. Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20, pp.1-28.

Coopmans, P, and S. Stevenson. 1991. *How* Extraction from Finite and Infinite Complements: A Surprising Asymmetry. *Linguistic Inquiry* 22: 359-367.

Huang, J. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Doctoral dissertation. MIT.

Kayne, R. 1983. Connectedness. *Linguistic Inquiry* 14: 223-249

Manzini, R. 1992. *Locality*. Cambridge, MA: MIT Press.

McCloskey, J. 2000. Quantifier Float and Wh-Movement in an Irish English. *Linguistic Inquiry* 31: 57-84.

Obata, M. To appear. Phase and Convergence. *Proceedings of the 29th Annual Penn Linguistics Colloquium* (*Penn Working Papers in Linguistics*).

Pesetsky, D, and E. Torrego. 2001. T-to-C Movement: Causes and Consequences. In M. Kenstowicz, (ed.) Ken Hale: A Life in Language. Cambridge, MA: MIT Press, pp. 355-326.

Pesetsky, D, and E. Torrego. 2004. Tense, Case, and the Nature of Syntactic Categories. In J. Guéron and J. Lecarme (eds.) *The Syntax of Time*. Cambridge, MA: MIT Press, pp.495-538.

Pesetsky, D, and E. Torrego. 2004. The Syntax of Valuation and the Interpretability of Features. Ms., MIT and UMass Boston.

du Plessis, H. 1977. Wh-Movement in Africaans. *Linguistic Inquiry* 8: 723-726.

Richards, N. 2002. Very Local A'-Movement in a Root-First Derivation. In S.D. Epstein and T. D. Seely. (eds.) *Derivation and Explanation in the Minimalist Program.* Oxford: Blackwell, pp.227-248.

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

Ross, J.R. 1967. *Constraints on Variables in Syntax*. Doctoral Dissertation, MIT.

Simpson, A, and Z. Wu. 2002. Understanding Cyclic Spell-Out. *NELS* 32: 499-518.