Intertree Movement and Adjunct Control

Youngroung, Kim
(Woosuk University)

Kim, Youngroung. 2006. Intertree Movement and Adjunct Control. *The Linguistic Association of Korea Journal, 14*(4), 131-150. This paper purposes to see if sidewards movement is well motivated enough to take a place in grammar as one of grammatical operations. Hornstein (2001) takes up sidewards movement discussed extensively in Nunes (1995). Sidewards movement occurs when an element in one subtree is merged to a position in another unconnected subtree. Thus this movement is said to take place between two unconnected subtrees. A further assumption is required to lay a foundation for adjunct control as movement. Obligatory Control PRO is representative of Hornstein's analysis of sidewards movement. To look into the properties of PRO in adjunct in depth, we review some literature data, including Kayne (1984) and Lasnik/Saito (1992). Showing the need to replace Attract-based move with another alternative leads us to take into consideration several devices of grammar, including Superiority effect invoking Attract, linking, and WCO.

Key Words: OC PRO, sidewards movement, Attract, move, linking, WCO, Superiority effect, functionally-interpreted WH.

1. Introduction: A New Perspective of PRO

According to a device presented by Hornstein (2001), coreference is established between nominal elements in a sentence. His argument aims to eliminate the construal rules employed by Government-Binding theory (GB) in favour of a movement analysis compatible with the goals of the minimalist program. The analysis of obligatory control PRO (OC PRO) is representative of his approach. He argues that the empty category in the lower subject position of a sentence like (1) is a trace or a copy left by movement[1].

(1) a. John hopes PRO to leave.

b. [$_{IP}$ John [$_{VP}$ John [hopes [$_{IP}$ John to [$_{VP}$ John leave]]]]]

To treat PRO as part of a chain that also includes *John*, Hornstein (2001) substantially revises standard theta theory[2]. The derivation begins with *John* merging with *leave* thereby checking the verb's Θ-role. *John* then raises to (i.e. is copied and merges with) Spec of the embedded IP to check the D-feature of the IP. This is not a case marking position, so the case of *John* cannot be checked here. *John* raises again to Spec VP of *hope*. It checks the external Θ-feature of the verb. Each time *John* checks a Θ-feature of a predicate it assumes that Θ-role. Thus, *John* (or the chain it heads) has two Θ-roles, the "leaver-role" and the "hoper-role." *John* raises one last time to Spec IP of the matrix, where it checks the D-feature of the IP and nominative case. This is the only place where *John* checks case. On the assumption that it was inserted into the derivation with nominative case features, the derivation converges[3].

1) Hornstein (2001) has argued that the null hypothesis is that OC PRO is identical to NP-t, i.e. it is simply the residue of movement. NON PRO is 'pro,' the null pronomial analogous to the null pronoun found in various Romance and East Asian languages. Then, he adopts the following assumptions:

a. Θ-roles are features on verbs

b. Greed is enlightened self interest

c. A D/NP receives a Θ-role by checking a Θ feature of a verbal/predicative phrase that it merges with.

d. There is no upper bound on the number of Θ roles a chain can have.

2) He assumes that (a) theta-roles are features assigned by verbs to DPs; (b) theta-features can drive movement, which means that movement to theta positions is licit; and (c) a DP may receive more than one theta-role in the course of a derivation. Given that minimalist theories of syntax have independent reasons for positing the existence of features and feature-driven movement, this view of theta-role assignment fits rather neatly into MP. It also permits a theta-feature on *hopes* to drive the raising of an already theta-marked DP from the lower clause.

3) The relevant steps in the derivation of (1) are rehashed as follows:

(i) *John* merges with *leave*, checking the verb's theta-feature.

(ii) *John* moves to the subject position of the lower clause, checking D on INFL.

This movement analysis of (1) accounts for several well-known restrictions on PRO. Since the chain headed by *John* is produced by overt A-movement, PRO i.e., the copy of *John* in the lower subject position, must have a local c-commanding antecedent[4].

2. A Reading of PRO in the Adjunct

Hornstein (2001) posits sidewards movement by extending the movement analysis to account for control into adjuncts. In (2a), *John* is able to move from the adjunct to the main clause; and its paradigm is configured in (2b).

(2) a. $John_i$ saw $Mary_j$ before PRO_{i+*j} entering the room.
  b. NP_i V NP_j $[_{adjunct}$ $PRO_{i/*j}$ …]

Adjuncts also manifest properties of obligatory control. PRO headed adjuncts requires a local, c-commanding antecedents. Consider the sentences below.

(3) a. *$John_i$ said [that Mary left after PRO_i dressing himself
  b. *$John's_i$ picture appeared after PRO_i shaving himself
  c. *It seemed that Bill left before PRO_{arb} noticing

(3a) is unacceptable under the indicated indexation as the antecedent *John* is too remote. (3b) involves a non-c-commanding antecedent. The arbitrary reading of PRO in (3c) is unavailable, as it would be if PRO

(iii)*John* moves to the specifier position of the higher vP, checking the external theta-feature of 'hopes.'
(iv) *John* moves to the matrix subject position, checking D and case.

4) The antecedent must be a single DP constituent:
(i) *$John_i$ told $Mary_j$ PRO_{i+j} to leave together.
Because the moved element receives a theta-role in the higher clause, expletives and idiom chunks are not eligible controllers
(ii) *There hope PRO to be three cats in the room
(iii) *The shit expects PRO to hit the fan.

here did not require an antecedent[5]. The PROs in these adjuncts do not tolerate split antecedents:

(4) a. *John$_i$ said that Mary$_j$ left after PRO$_{i+j}$ washing themselves
    b. *John$_i$ told Mary$_j$ a story after PRO$_{i+j}$ washing themselves

PRO headed adjuncts only have sloppy readings under ellipsis.

(5) John left before PRO singing and Bill did too

Thus, (5) only has the reading paraphrased in (6a). It cannot be understood as (6b).

(6) a. . . and Bill left before Bill sang
    b. . . and Bill left before John sang

In sentences (7), (7a) cannot take 'Churchill' as antecedent as shown in (7b). In other words, (7b) is not an adequate paraphrase of (7a).

(7) a. Only Churchill left after PRO giving the speech
    b. Only Churchill left after Churchill gave the speech

In sum, adjunct control structures pattern like OC structures. They have a further distinctive property noted in (2b): they do not permit object control. This property is typically accounted for by assuming that objects do not c-command adjuncts and so cannot bind the PRO[6].

---

5) That is, the absence of a reading for (3c) analogous to "It seemed that Bill left before anyone noticed" follows from the requirement that this PRO needs a grammatical antecedent.

6) In MP terms, this requires assuming that objects fail to c-command adjuncts at LF, the locus of binding requirements in a minimalist theory. This assumption, however, is doubtful. Objects can license bound pronouns within adjuncts as in (i).

Sidewards movement occurs when an element in one subtree is merged to a position in another unconnected subtree. In short, such movement is interarboreal. The characteristic property of such movement is that movement is not to a c-commanding position. Let's see how it can be put to use to account for the properties of adjunct control.

Turn to the derivation of (2a). The numeration for (2a) consists of the set of items {John, heard, Mary, before, entering, the, room, assorted functional categories}. We build the adjunct phrase by merging *the* to *room* then merging *the room* with *entering* and merging *John* with *entering the room*. The two theta-roles of *enter* are checked by the merger of the two D/NPs. The 'ing' heads its own Infl projection. This merges with the previously formed VP small clause. The strong feature of this Infl is checked by raising *John*[7]. At this point we have a structure like (8) after the adjunct has merged with the IP.

(8) a. [$_{\text{adjunct}}$ before [$_{\text{IP}}$ John [$_{\text{I}}^{0}$ ing [$_{\text{VP}}$ John [entering the room]]]]]

(i) John read every book$_i$ without reviewing it$_i$

If *every book* can bind *it* then *every book* c-commands *it* at LF. If so, *it* c-commands PRO as well.

There is a larger problem with this proposal, however. It does not account for the OC properties of PRO headed adjuncts. The OC properties of these constructions suggest that the PRO found here is actually the residue of movement.

Analyzing adjunct control as movement requires a further assumption. We need to assume that movement out of an adjunct is possible. This in turn requires reanalyzing standard CED effects. For the present, let's simply assume that movement out of an adjunct is indeed possible in some cases. Moreover, we must assume that this movement is a species of sidewards movement discussed extensively in Nunes (1995).

7) The derivation of (2a) involves the following steps:

(i) The subtrees *before John entering the room* and *saw Mary* are constructed by normal applications of Merge and Move.

(ii) *John* undergoes interarboreal movement to check the external theta-feature of *saw*.

(iii) The adjunct merges with the matrix vP.

(iv) The matrix INFL merges and *John* moves to its specifier position.

b. before $[_{IP}$ John $[_{I}$ $[_{vP}$ John v $[_{VP}$ leaving the room]]]]
+Θ/+Nom +Θ/+Nom

Next we build the main clause. *Mary* merges with *read.* The internal theta role is thereby discharged.

(9) $[_{VP}$ saw Mary], $[_{adjunct}$ before $[_{IP}$ John $[_{I}^{0}$ ing $[_{VP}$ John [entering the room]]]]]

Note that in (9) we have two unconnected subtrees in the derivation. This is where sidewards movement becomes relevant. Let's continue with the derivation. The external Θ-feature of *saw* must be checked. If we move *John* then the derivation proceeds as follows: *John* raises and discharges the external Θ-role by merging with the VP of *saw*. This is an instance of sideways movement as the target of movement is not a c-commanding position. Indeed it is not even in the same subtree.

(10) $[_{VP}$ John [saw Mary]], $[_{adjunct}$ without $[_{IP}$ John $[_{I}^{0}$ ing $[_{VP}$ John [entering the room]]]]]

The next step is to merge the adjunct and the VP.

(11) a. $[_{VP/VP}$ $[_{VP}$ John [saw Mary]] $[_{adjunct}$ without $[_{IP}$ John $[_{I}^{0}$ ing $[_{VP}$ John [entering the room]]]]]]

b. $[_{IP}$ John $[_{I}$ $[_{vP/vP}[_{vP}$ John $[_{VP}$ saw Mary]] [after
+Θ/+Θ/−Nom +Θ/+Θ/+Nom
$[_{IP}$ John $[_{I}$ $[_{vP}$ John v $[_{VP}$ leaving the room]]]]]]]]
+Θ/+Nom +Θ/+Nom

The derivation then terminates with *John* raising to Spec IP to check its own case, those of Infl, and the latter's D-features. At LF, *Mary* raises to check accusative case in either the outer Spec of VP or in AgroP. The derivation converges with the overt structure in (12).

(12) [$_{IP}$ John [$_{I}^{0}$ past [$_{VP/VP}$ [$_{VP}$ John [saw Mary]] [$_{adjunct}$ without [$_{IP}$ John [$_{I}^{0}$ ing [$_{VP}$ John [entering the room]]]]]]]]

This derivation illustrates a number of economy principles. First, all movement is greedy: it must check some feature of either the moved element or the target. Movement cannot be implemented as Attract, because the element to be moved (*John*) is not within the search space c-commanded by the target *saw*. Locality is thus evaluated from the perspective of the moved element: movement of an XP to a target T can be blocked by an intervening landing site T′, but not by an intervening XP′. Second, Move is deemed less economical than Merge, on the familiar grounds that Move consists of two steps (Copy plus Merge), while Merge involves only one. This is what allows subjects but not objects to control into adjuncts like the one in (2). Otherwise, the internal theta-feature of *saw* could be checked by moving *John* instead of by merging *Mary*, ultimately yielding *Mary saw John$_i$ before PRO$_i$ entering the room. Finally, economy principles evaluate only convergent derivations. In the derivation of *before John entering the room*, moving *John* from its base position to Spec IP is less economical than merging *Mary* in Spec IP. However, if *Mary* merged as the lower subject, *John* would be unable to check case. The economy of a contemplated step in a derivation thus cannot be evaluated in isolation, but must depend on whether the derivation can ultimately converge.

We see that adjuncts at different heights will permit control from different positions in the main clause. Adjuncts like the one in (2a) attach at the level of light vP, and can be controlled only by subjects.

Sidewards movement from adjuncts also makes it possible to account for some phenomena traditionally analyzed as involving null operators. For example, tough-movement, treated in GB as shown in (13a), is analyzed by sidewards movement as in (13b):

(13) a. [Moby-Dick is [$_{AP}$ easy [$_{CP}$ Ø$_i$ PRO$_{arb}$ to read t$_i$]]].

b. [[Moby-Dick$_i$ is t$_i$ easy] [$_{CP}$ t$_i$ pro to read t$_i$]].

Hornstein (2001) claims that *Moby-Dick* originates as the object of *read* and moves to Spec CP of the adjunct in order to check a 'relative' WH feature. It then undergoes sidewards movement to check the theta-feature of *easy*, and finally moves to Spec IP of the main clause, whereupon the adjunct CP merges with the matrix IP.

A parasitic-gap construction, illustrated in (14), is also analyzed by Hornstein's analysis of sidewards movements:

(14) Which book$_i$ did you read t_i before t_i Fred reviewed t_i?

Which book merges in the position of the parasitic gap, WH-moves to Spec CP of the adjunct, moves sidewards to check the internal theta-feature of *read*, and finally WH-moves to the matrix Spec CP. The fact that the parasitic gap in the adjunct must be licensed by the true gap in the matrix clause follows from the fact that the adjunct attaches to the matrix clause too low to permit *which book* to move directly to the higher Spec CP; it must pass through the matrix VP first. In contrast to (2), the element that moves out of the adjunct must target the matrix verb's internal theta-position rather than its external one. As in (2), economy principles would prefer that the internal theta-feature of *read* be discharged by merging *you* rather than by moving *which book*[8].

---

8) In (14), however, this alternative, like the economical derivation that yields the ill-formed passive version as in (i), is ruled out by the requirements of convergence.

(i) *Which book$_i$ t_i was read t_i by you before t_i Fred reviewed t_i?

In (i), the trace of *which book* in the matrix Spec IP c-commands the parasitic gap in the lower object position. In GB terms, this configuration violates Principle C: the lower case-marked trace (variable) is illicitly A-bound by the higher one. Hornstein recasts Principle C in minimalist terms by proposing a Scope Correspondence Axiom (SCA) similar to the Linear Correspondence Axiom (LCA) of Kayne (1994). The SCA holds at LF and states that if α c-commands β, then α takes scope over β. Assuming that both case-marked copies of *which book* are subject to this axiom, neither may c-command the other, because the SCA would then require *which book* to scope over itself. Since this is impossible,

3. Sidewards Movement vs Attract

3.1 Move vs Attract

Every approach to movement imposes locality restrictions on the operation. Standard minimalist approaches incorporate some version of relativized minimality which prohibits moving an expression over another expression of the same type. The intuition behind minimality can be implemented in at least two distinct ways. The first takes a target centered view of the Move operation. The second is launch based and focuses on the expression being moved rather than the position to which it moves.

The target centered version of minimality measures the relative length of a movement from the perspective of the target of the operation. For example, in a structure like (15), where XP1 c-commands XP2 the distance between XP1 and the target T is shorter than the distance between XP2 and T so this blocks the movement of XP2 to T.

(15) T [. . XP1 . . . XP2 . .]
(16) a. XP2 [T [. . . XP1 . . XP2 . .]]
b. XP1 [T [. . . XP1 . . . XP2 . . .]]

In effect, (16a) is blocked because the span of the indicated movement, i.e. the length of the XP2-chain, is longer than the equally available, but "shorter"XP1-chain in (16b).

A second implementation of the minimality idea centers on the expression moving rather than the position it targets. For a movement to be licit it must be greedy. Thus, some feature must be checked as a result of the move operation. Say that (at least one of) the features that

---

the economical derivation fails to converge at LF, and so (14), although uneconomical, is permitted. In addition to sidewards movement, Hornstein's analyses of tough-movement and parasitic gaps depend on the possibility of movement from an A'-position (Spec CP in the adjunct) to an A-position.

requires checking is on the expression being moved. minimality requires that an XP's movement be the shortest possible to check this feature. For example, in configurations in which P1 c-commands P2, XP cannot pass over a position P1 on its way to P2 in which it checks feature F if F could have just as well been checked in P1. Given this conception, the movement indicated in (16a) is perfectly licit so long as the feature on XP2 could not have been checked in any position between XP2 and T. The fact that XP1 is closer to T (the position in which a feature is checked) is irrelevant so long as XP2's move is the shortest one possible for XP2.

These different implementations of minimality have important consequences for sidewards movement. Consider why. Both conceptions define the metric of locality in terms of c-command. The target centered notion only compares the relative distances of expressions c-commanded by a target T[9]. The second conception also invokes c-command but in a different manner. The only intervening positions relevant for computing minimality are the ones that c-command the expression that moves[10]. Of these two ways of thinking about minimality, the second is incompatible with sidewards movement for the target of movement can never c-command the expression that moves in cases where the movement is between subtrees. This is illustrated in (17).

(17) [DP_1 [v [V DP_2]]] [$_{adjunct}$ DP_1 . . .]

In (17), DP_1 moves from the adjunct to the Spec vP of a separate subtree. Two features of this movement are noteworthy in the present context. First, as it is a case of sidewards movement, the target, in this case v, does not c-command the adjunct and so fails to c-command the

9) That is, elements not c-commanded by T cannot be compared with respect to distance and so can be considered equi-distant from a given target.

10) In other words, if YP fails to c-command XP then its presence is irrelevant for computing shortest moves. This notion is also invoked in many target centered theories that incorporate Attract into the definition of move.

expression which has moved, i.e. DP1, prior to movement. On the assumption that movement requires attraction and that attraction requires c-command, this sort of movement should be illicit. Second, DP_2 does not prevent the indicated movement[11]. DP2 does not c-command DP1 so cannot block the movement of DP1 to Spec vP. Moreover, this is the shortest move DP1 can licitly make, e.g. that respects Greed. Thus, this case of sidewards movement respects the launch centered version of minimality[12].

3.2 Attract and Superiority Effect

The principal argument in favor of a definition of movement incorporating Attract is empirical. It is critical to most current analyses of Superiority Effects.

(18) a. I wonder who saw what
b. *I wonder what who saw

The contrast between (18a) and (18b) can be explained if move is understood as involving Attract. Consider the details. The structure of the sentences in (18) prior to movement to Spec CP is (19).

(19) . . . [C^0 [$_{IP}$ [$_{IP}$ who I^0 [$_{vP}$ who [$_{VP}$ saw what]]]]]

The C^0 has a strong WH feature that requires checking by a WH element. Both *who* and *what* can satisfy this need. The movement of *who* is preferred for the distance it must travel is shorter than the

11) This would be surprising if Attract were operative. The reason is that v does c-command DP2 so one might expect the movement of DP2 to v to be preferred to (and so block) the movement of DP1 from the adjunct. Note that the postulated sidewards movement operation is not problematic if move is subject to the launch based conception of minimality outlined above.

12) Thus, the approaches to movement that incorporate Attract are target centered in the sense applicable here and so are incompatible with sidewards movement.

distance *what* must traverse to check the feature. More specifically, the set of nodes between the head of the A′-chain formed by moving *who* is a subset of those formed by moving *what*.

(20) a. . . . [$_{CP}$ who [$_{IP}$ who I^0 [$_{vP}$ who [$_{VP}$ saw what]]]]
b. . . . [$_{CP}$ what [$_{IP}$ who I^0 [$_{vP}$ who [$_{VP}$ saw what]]]]

The path traversed by moving *who* is the set of nodes {CP, IP}. Moving *what* yields the path {CP, IP, vP, VP}. As the former is a subset of the latter, moving *who* is shorter and so is required. This accounts for the unacceptability of (18b).

This basic account of Superiority Effects can also account for a variety of apparent exceptions to the condition. For example, Superiority Effects are obviated when 'which' phrases are moved. Thus the relative acceptability of the sentences in (21) is roughly on a par.

(21) a. I wonder which book which man saw
b. I wonder which man saw which book

This follows if we are careful in computing the paths formed by the two movements. What drives the movement in these examples is the requirement to check a strong WH feature of C^0. What is attracted, therefore, are WH features. These features reside in 'which' in the DPs in (21). The relevant structures of (21) are provided in (22).

(22) a. . . . [$_{CP}$ [$_{DP1}$ which book] C^0 [$_{IP}$ [$_{DP2}$ which man] [$_{vP}$ [$_{DP2}$ which man] [$_{VP}$ saw [$_{DP1}$ which book]]]]]
b. . . . [$_{CP}$ [$_{DP2}$ which man] C^0 [$_{IP}$ [$_{DP2}$ which man] [$_{vP}$ [$_{DP2}$ which man] [$_{VP}$ saw [$_{DP1}$ which book]]]]]

The 'C^0'attracts the WH features in *which*. The relevant paths therefore are {CP, DP1, IP, vP, VP} for (22a) and {DP2, IP} for (22b). Observe that neither path is a subset of the other. As such, either *which* phrase may move to Spec CP as neither move is longer than the other[13]. This general approach to Superiority has been successful in accounting for the basic Superiority data[14].

However, there are various examples in the literature that raise problems for an Attract based approach to Superiority Effects. First consider the following examples from Kayne (1984). Kayne observes that Superiority Effects are mitigated when additional WH elements are added. He observes the contrast in (23).

(23) a. *John wonders what$_i$ who saw t_i then
b. John wonders what$_i$ who saw t_i when

(23a) is a standard superiority violation. What is curious, as Kayne observes, is that the effect seems to go away when we replace the pronominal adverb 'then' with the WH word 'when', as in (23b). Observe that this is hard to account for if Superiority violations are analyzed in terms of shortest move based on Attract. The distance that *what* moves in (23a) and (23b) is the same. In both cases *who* is closer to 'C^0' than *what* is. Thus, both sentences should have the same status. This is not what we find. (23a) is unacceptable while (23b) is fine. Lasnik and Saito (1992) discuss a second problematic case.

(24) Who wonders what who bought

(24) is ambiguous. *Who* in the subject position of the embedded clause can be either paired with *what* in the embedded Spec CP or with *who* in the matrix comp. The former pairing leads to a structure in which we expect a standard Superiority violation and under this reading the sentence is indeed judged unacceptable. This contrasts with the

13) Thus, on the assumption that 'which' phrases are complex DPs while simple WH phrases like 'who', 'what', etc are simplex, we can account for the lack of Superiority Effects with complex WH DPs.

14) Extensions have also been elaborated to account for a wide array of cross linguistic differences manifested by multiple interrogative constructions in a variety of languages

acceptability of the sentence on the interpretation in which the embedded *who* is paired with the matrix *who*. An appropriate answer to this question would be: "John wonders what Bill bought and Mary wonders what Sheila bought." On this reading the sentence is fine.

Once again, this makes little sense given the standard analysis in terms of Attract. Whether or not the embedded subject is paired with the matrix or embedded WH does not alter the fact that *who* is closer to the embedded 'C^0' than *what* is. This should prevent moving *what* over *who* and should result in a violation of Superiority irrespective of the reading being considered. Put this another way. The analysis of Superiority Effects in terms of Attract is purely formal. It identifies the shortest path to 'C^0' and this alone is decisive in determining which WH element can move. Lasnik and Saito's case indicates that this is not sufficient.

3.3. An Alternative to Superiority Effect

Referring to an approach to Superiority first outlined in Hornstein (1995), we find the virtue of yielding an account of Superiority Effects without exploiting target centered conceptions of minimality based on Attract. The basic idea is as follows. A WH in situ is interpreted functionally. The idea is to treat functionally interpreted WHs as similar to bound pronouns within complex DPs. An example should make the core of the proposal clear.

Consider a sentence like (25) in which *his* is understood as bound by *everyone*. This sentence is true just in case there is a pairing of every relevant person in the domain with another individual, namely that person's mother, and the relevant individual loves the person he is paired with[15].

(25) Everyone loves his mother

---

15) In other words, one can understand a bound pronoun in cases such as this as establishing a function whose range is a set of individuals and whose domain consists of mothers of those individuals.

Assume that the same sort of relation is established in multiple interrogative constructions. In other words, assume that in (26) the 'what' in object position functions like 'his mother' in (25).

(26) a. Who saw what
b. [$_{CP}$ Who$_i$ [$_{IP}$ t$_i$ saw [pro$_i$ thing](=what)]]

That is, we assume that a functionally interpreted WH actually resolves into a bound pronominal part and a nominal restrictor, i.e. a part corresponding to *his* in (25) and a part corresponding to *mother*. This provides (26a) with a structure like (26b).

Given this much, it is possible to assimilate Superiority Effects to Weak Cross Over (WCO) Effects. For example, the unacceptability of (37a) can be analyzed as a Weak Cross Over violation. The indicated binding conforms to a Weak Cross Over configuration and the unacceptability of (27) is roughly on a par with that of (28), a standard WCO violation. Observe that the binding indicated in (26) is not of the Weak Cross Over variety and so it is fully acceptable.

(27) a. *What did who see
b. [$_{CP}$ What$_i$ [$_{IP}$ [pro$_i$ person](=who) see t$_i$]]
(28) *Who$_i$ did his$_i$ mother see t$_i$?

Hornstein (1995) argues that the correct interpretation of pronoun binding is in terms of linking. Assume that this is correct. The Weak Cross Over Condition can then be stated as (29).

(29) A pronoun cannot be linked to a WH-t/variable to its right[16)]

We have seen how these assumptions suffice to account for standard cases of Superiority like (27) above. Consider the details. The Lasnik

16) The gist of the proposal then is that WHs in situ are analyzed as functionally interpreted expressions on a par with bound pronouns. Moreover, binding is understood as linking.

and Saito cases have the structure in (30), repeated.

(30) Who wonders [$_{CP}$ what$_i$ [$_{IP}$ who bought t$_i$]]

The current proposal says that the embedded *who* must be functionally interpreted in a multiple question. It can be so interpreted with respect to the WH in the matrix or the embedded CP. The two relevant readings are diagrammed in (31).

(31) a. Who$_i$ [$_{IP}$ t$_i$ wonders [$_{CP}$ what$_j$ [$_{IP}$ [pronoun$_j$ person] bought t$_j$]]]
b. Who$_i$ [$_{IP}$ t$_i$ wonders [$_{CP}$ what$_j$ [$_{IP}$ [pronoun$_j$ person] bought t$_j$]]]

In (31a) the "functional" pronoun is linked to the matrix trace 't$_i$'. This linking is in conformity with the Weak Cross Over Condition (29) as 't$_i$' is to the left of the bound pronoun. Hence the structure is well formed and the sentence is acceptable. In (31b), in contrast, the pronoun is linked to the trace/variable of the embedded *what*. This trace/variable, 't$_j$', is to the right of the bound pronoun. This linking violates the Weak Cross Over Condition and so the structure is illicit. This is why the sentence under this reading is unacceptable.

Consider now the case discussed by Kayne (1984). He observes that the addition of WHs obviates the effects of Superiority in multiple interrogatives. The sentences, repeated here, have the structures in (23), repeated, with the indicated functional dependency of the WH in situ.

(23) a. *John wonders what who saw then
b. John wonders what who saw when

(32) a. John wonders [$_{CP}$ what$_i$ [$_{IP}$ [pronoun$_i$ person] saw t$_i$ then]]
b. John wonders [$_{CP}$ what$_i$ [$_{IP}$ [pronoun1$_i$ person] saw t$_i$ [pronoun2$_i$ person]]]

In (32a), the pronoun is linked to a WH-t/variable to its right, in violation of (29). The addition of an extra WH, which must also be functionally interpreted, allows the structure to conform to (29). Consider how. Assume that the following linking relations obtain: $pronoun_2$ is linked to ti and pronoun1 is linked to $pronoun_2$. This linking is perfectly legitimate. Moreover, the linkings conform to (29) as no pronoun is linked to a WH-t/variable on its right. If Superiority is simply an instance of the WCO effect then we expect the presence of additional WHs which require functional interpretations to ameliorate otherwise unacceptable sentences.

The same account extends to explain the improvement in (33b) when an additional bound pronoun is introduced. In (33a) the pronoun is linked to a WH-t/variable to its right. In (33b) it can be linked to the genitive pronoun in the 'mother'-DP, which is in turn linked to the WH-t/variable to its left. This multiple linking, which extra pronouns make available leads to the attenuation of the WCO effect in a manner parallel to that witnessed in (32). This parallel behavior of bound pronouns and functionally interpreted WHs in WCO and Superiority contexts supports the type of analysis proposed here.

(33) a. *Who_i did his_i father talk to t_i about Mary
b. Who_i did his_i father talk to t_i about his_i mother

There is an additional argument in favor of the approach that assimilates Superiority to Weak Cross Over. It comes from considering the fact that most cases of multiple A'-movement fail to manifest Superiority Effects. Superiority Effects are absent in focus constructions and echo questions. Given the above analysis, we only expect Superiority to appear in cases where a functional dependency among WH expressions is established. Where there is no such relation, linking does not obtain and the Weak Cross Over Condition should be inoperative.

Romanian and Bulgarian are languages that require all WHs to front in multiple interrogative constructions. For example, the translation of (34a) in Romanian has the 'ce'(=what) overtly move to pre-sentential position.

(34) a. who saw what
b. Cine ce a vazut
"who what saw"

This requirement carries over to echo-questions. Comorovski (1996: 59ff) observes that Superiority Effects are cancelled in echo questions. This is illustrated by (35).

(35) Despre ce cine ti a vorbit
about what who to you has told
"Who told you about what"

(35) is ill formed under a standard question intonation. This is expected given that Romanian displays Superiority Effects and in this case we are raising a PP containing a WH over a subject. In terms of the analysis we have made, the problem with (35) as a standard multiple question is that it induces a Weak Cross Over Effect. The structure is that of (36) at LF.

(36) what$_i$ [[pronoun$_i$ person] told you about t$_i$]

The pronoun in (36) is linked to a WH-t/variable to its right in violation of (29). Thus, the multiple interrogative is ill-formed as it violates Weak Cross Over/Superiority[17].

---

17) However, Comorovski (1996: 59) notes that (35) is perfectly acceptable with an echo interpretation. She similarly observes that the English counterparts of (35) seem immune to Superiority in echo questions. Comorovski (1996: 60) makes a methodological point concerning these data.
(i) a. It was a unicorn that Esmeralda saw
b. *What* did *who* see (Italics indicate stress for echo intonation)

The absence of Superiority Effects is expected if the latter are simply the result of linkings in violation of Weak Cross Over. Echo questions do not

4. Conclusion

This paper scrutinizes Hornstein (2001) taking up sidewards movement, which was extensively discussed in Nunes (1995), with a view to examining whether or not his argument to make the grammar safe for sidewards movement is justified. As shown above, in order to adopt sidewards as a viable grammatical option, it is imperative to eliminate the analysis of the attracted-approached move, because sidewards movement and Attract-based move does not fit properly. In addition, Superiority Effects invoking Attract based approaches to move should be eliminated. Hornstein claims that the data, dealt with Attract-based Superiority Effects, can be accounted for through other grammatical devices such as linking/WCO, instead. We see to it that whether linking/WCO is so adequate as to replace Superiority Effects.

Granting that Hornstein's argument to reject Superiority Effects to make room for sidewards movement is made to some justifiable extent, we cannot but think that his such argument is artificial. In addition, his idea--that views a WH in situ as being interpreted functionally and treats functionally interpreted WHs as similar to bound ronouns within complex DPs--is considered awkward with no justifiable rationale. Thus, prior to bring sidewards movement into a grammatical operation, a reasonable clue to resolving the issue concerning these awkward claims should be presented first. Worse still, the cost of replacing Superiority Effect by linking/WCO is expensive in that we have to make separate assumptions to account for the data, dealt with Superiroty Effects, through linking/WCO. Such attempts cannot be considered economical at all.

References

establish functional relations between the WHs. There is no appropriate pair-list answer available to echo questions. Consequently, there is no linking and so no threat of a violation of (34). As such, Superiority should be, and is, inoperative.

Adger, D. (2003). *Core Syntax, A Minimalist Approach.* Oxford: Oxford University Press

Carnie, A. (2002). *Syntax, A Generative Introduction.* Oxford: Blackwell Publishers

Chomsky, N. (1995). *The Minimalist Program,* 351-353. Cambridge, Mass.: MIT Press

Chomsky, N. (2000). Minimalist Inquiries: The Framework. Martin., R, D. Michaels, and J. Uriagereka, (eds). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik,* 89-155. Cambridge, Mass.: MIT Press.

Comoroski, I. (1996). *Interrogative Phrases and the Syntax-Semantics Interface.* Kulwer. Dordrecht

Hornstein, N. 1995. *Logical Form, From GB to Minimalism.* Oxford: Blackwell Publishers

Hornstein, N. (2001). *Move! A Minimalist Theory of Construal.* Oxford: Blackwell Publishers

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

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

Nunes, J. (1995). *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program,* Ph.D. dissertation, University of Maryland, College Park.

Radford, A. (2004). *Minimalist Syntax, Exploring the structure of English.* Cambridge, Cambridge University Press.

Youngroung, Kim
Dept. of English, Woosuk University
Wanju-Gun, Chonbuk, 565-701
Phone: 82-63-290-1336
Email: youngkim@woosuk.ac.kr

Received: 30 Oct, 2006
Revised: 29 Nov, 2006
Accepted: 2 Dec, 2006