

PRO gate and movement

Hirohisa Kiguchi

University of Maryland at College Park

0. Introduction

There has been a controversy concerning the approach to grammars. The issue is whether to take a representational approach or derivational approach. To capture syntactic phenomena, the former exploits the notion of chains and the latter uses histories of movement. Brody (1995) who advocates a representational version of the Minimalist Program argues that if both the notion of chain and movement are constrained by *c*-command, one of them must be discarded because there would exist too much redundancy. Sideward movement proposed by Nunes (1995), which de-couples movement from *c*-command supports a derivational approach, for the notion of chain would not cover this operation. But Nunes (1995) also argues that the notion of chain is necessary at the end, to license the result of movement. Thus, his system remains a hybrid one.

In this paper I will examine PRO gate phenomena first discussed in Higginbotham (1980). One of the main interests of this phenomenon is that I believe its correct analysis argues against having a chain at all. (1) illustrates the phenomenon:

- 1) a. ?? [His₁ getting his₁ car fixed] upset everyone₁.
- b. [PRO₁ getting his₁ car fixed] upset everyone₁.

In (1b), PRO replaces the overt pronoun and shows no weak cross over effects. Higginbotham (1980) dubs the phenomena “PRO gate”. As shown above, the object controls PRO inside the subject. Apparently, the relation between *everyone* and PRO is not a chain relation.

Here, adopting Hornstein’s (1999) suggestion that PRO is a residue of NP-movement, we show that a strict derivational approach which exploits sideward movement can explain PRO gate phenomena. This implies that the notion of chain cannot capture all the same facts that a movement can. If so, the concept of chain is empirically flawed, and a derivational approach is favored. In addition, this paper argues that the subject is not intrinsically an island and that movement out of a subject should be possible. We also show that the derivational approach taken here can solve the apparent contradiction between the movement out of the subject and the CED.

This paper organized as follows. In section 1 we make explicit the basic issue in this paper. In section 2 we argue that the movement out of the subject, in principle, is possible. We show that this option follows from the Extension Condition and the assumption that movement is the complex operation of Copy + Merge as in Nunes (1995). In section 3 we discuss the existence of sentential subjects. This has been under debate since Koster (1978). We conclude that they indeed exist, at least in the case of gerunds. In section 4 we rebut Belletti & Rizzi’s (1988) proposal that Phyc verbs are unaccusatives. Instead, we regard the surface subjects of Phyc verbs as the genuine subject i.e., the external argument. In section 5 given the assumptions discussed in the previous sections, we show how the derivational approach with sideward movement derives PRO gate phenomena, and how it handles the tension between the movement out of the subject and the CED. In section 6 we investigate various independent reasons to regard PRO in PRO gate phenomena as a residue of movement. Hornstein (1999) shows that Obligatory Control PRO (=OCPRO) is a result of movement. Basically following his argument, we provide evidence that PRO in PRO gate phenomena belongs to OCPRO category. In section 7 we discuss the implementation of the deletion

of copies. We reject the chain requirement for the deletion by Nunes (1995) and adopt an alternative i.e. Hornstein's (Forthcoming) deterministic approach that the deletion of copies must be forced. In section 8 we discuss representational approach and derivational approach to grammar. We conclude that the representational approach such as Brody's (1995) is problematic if the present analysis of PRO gate phenomena is (roughly) correct. We further reject Epstein's (1999) First Law that no relations can hold between members of two trees that are unconnected at any point of the derivation. We also argue that the present analysis requires rejecting Attract-based approaches to movement. Section 9 is a brief summary.

1. Issue

1.1 Hornstein (1999)

Raising and control structures have long been considered to be different. The former is analyzed as (2), whereas the latter is (3) in Government and Binding approach.

- 2) a. John seems to leave.
 b. John₁ seems [t₁ to leave].
- 3) a. John wants to leave.
 b. John₁ wants [PRO₁ to leave]

Hornstein (1999) argues that raising and control sentences are both generated through movement. Let us see the derivation of the control construction (3).

- 4) a. [to John leave] (building the embedded IP)
 b. [John₁ [to John₁ leave] (moving 'John' to the Spec IP)
 c. [wants[John₁ [to John₁ leave]] (building the matrix VP)
 d. [John₁ [wants[John₁ [to John₁ leave]]] (moving John to the Spec VP)
 e. [John₁ [John₁ wants[John₁ [to John₁ leave]]]] (building TP and moving 'John' to its Spec)
 f. [John₁ [~~John₁~~ wants[~~John₁~~ [to ~~John₁~~ leave]]]] (deleting the copies all but one)

Now that D-structure and the theta-Criterion are history, assuming that theta-roles are features in the sense of inducing greedy movement, he permits movement from one theta-position to another. In the example above, 'John' gets the first theta-role when it merges with 'leave' in (4a). Then, it gets the second one when it moves to the specifier of the matrix verb, 'wants' in (4d). The derivation keeps going to converge until (4f). What this shows is that allowing multiple theta-marking enables us to regard a control construction as a version of raising. It follows that we can dispense with the entity, PRO, which is now identical to NP-trace i.e. a residue of movement in Hornstein (1999)'s approach.

1.2 PRO in the subject

Hornstein (1999) assumes that PRO is small pro, (a null pronominal) in cases where the movement is prohibited. The empty category below serves as an example. Since subject sentences are islands, we cannot extract anything out of the subject position of a sentential subject. That is to say, we

cannot have so-called NP trace in the place. As a result, we need the help of small *pro*, which is the pronominal found in many Romance or East Asian languages.

5) John₁ thinks that it is believed that [PRO₁ (=pro) shaving himself] is important.

However, if all gaps in subjects are null pronouns (=pro) why are Weak Cross Over (WCO) effects not observed in the so-called PRO gate sentences such as (6) first discussed in Higginbotham (1980)?

6) Who₁ did [PRO₁ cooking his₁ lunch] annoy t₁?

WCO-effects appear where a pronoun is interpreted as a bound variable but neither the pronoun nor the variable it is semantically tied to (cf. (7)) c-commands the other.

7) ??Q₁.... pronoun₁....variable₁ (where the pronoun and the variable are not in a c-command relation)

But as Higginbotham (1980) points out, WCO-effects disappear when PRO controlled by the Q-element intervenes between the Q-element and the pronoun as shown (6). That is, descriptively, the configuration (8) cancels WCO-effects:

8) Q₁...PRO₁...pronoun₁....variable₁ (when the pronoun and the variable are not in a c-command relation)

Then, if PRO in (6) is indeed small *pro*, which is pronominal, this sentence should be prey to WCO-effects, just like the sentence (9a).

9) a. ??Who₁ did his₁ cooking his₁ lunch annoy t₁?

Note that it is unlikely that the phonologically null status of *pro* is relevant factor in exempting PRO gate cases from WCO-effects. Hoji (1985) observes that the phonologically null pronouns in Japanese are subjects to WCO-effects.

b. ?? John-ga [Mary-ga pro₁ atta]-atode dono hito₁-ni mo atta
 John-NOM Mary-NOM pro met after every man-DAT-too met
 'John met every person₁ after Mary had met him (pro)₁'

However, if we can say that PRO in PRO gate sentences is a residue of A-movement, then we can use the fact that A-movement cancels WCO-effects (as in (10)) in order to account for the absence of WCO-effects in PRO gate sentences.

10) a. ??Who₁ did it seem to his₁ mother t₁ is handsome.
 b. Who₁ t₁ seems to his₁ mother t₁ to be handsome.

Note that the fronting of the Q-element with local clause internal scrambling also cancels WCO-effects in (9b). This is analogous to the cases in (10) if, as Saito (1992) argues, local clause internal scrambling is a species of A-movement.

- c. John-ga [dono hito-ni mo₁]₂ [Mary-ga pro₁ atta]-atode t₂ atta.

1.3 Why can't we simply say that PRO cancels WCO-effects?

Given this description we have seen so far, the question to follow is whether we can simply stipulate that PRO has the power to cancel WCO-effects when it binds the pronoun. Suppose that PRO is an empty category with Null Case, as suggested in Chomsky & Lasnik (1993), one might stipulate that this entity and the pronoun bound by it are not subject to WCO-effects. However, we cannot simply say that PROs act as gates. (11) displays WCO-effects even though PRO intervenes between the Wh-element and the pronoun in the same way as in (6).

- 11) ??Who₁ does the fact that PRO₁ cooking his₁ lunch is mandatory annoy t₁?

The contrast between this sentence (11) and (6) suggests that the PROs in these sentences are different. If they were identical, we would expect them to act the same way, with respect to the PRO gate, contrary to the fact. If we show that PRO in (6) is a residue of A-movement but the one in (11) is not, we can account for the contrast between (6) and (11) as follows. Since A-movement cancels WCO-effects, PRO which is a product of A-movement should do so as well. Whereas when A-movement is prevented, PRO is small pro, which being pronominal has no power to obviate WCO-effects. Observe that this line of thought requires movement from inside the subject to a position outside the subject. However subjects are generally taken to be islands. And it is the case that we cannot freely extract elements out of the subject. For example, Wh-movement from subjects is unacceptable. It violates the CED, which bans the extraction from subjects and adjuncts:

- 12) a. *Who do you think that [pictures of t] resemble photos of Bill?
b. *Who do you get jealous [because I talked to t]?

In other words, the question we must address now is how movement out of the subject is possible without violating the CED. We will see that movement from inside the subject is possible only before it merges with another tree. This allows PRO gates sentences to show no WCO-effects.

2. Movement from the inside of the subject

2.1 Sideward movement: Nunes (1995)

Nunes (1995) argues that if movement is the interaction of the distinct operations, Copy and Merge, sideward movement is possible. Sideward movement is the derivation in which an element is copied out of one syntactic tree and then, the copy merges with another independent syntactic tree.

- 13) a. [k...T...] [l....] (two independent trees)
b. [k...T...] T [l....] (copying T in k)
c. [k...T...] [T[l....]] (merging the copy with the other tree, l)

Nunes (1995) makes use of this sideward fashion to account for the movement from the adjunct to the main clause. (i.e., Parasitic Gap constructions and Across The Board constructions)

2.2 The subject as an individual tree

If we assume that sideward movement is possible, following Nunes (1995), we can also expect an element to move “sideward” from the subject to the main clause. Assuming the Extension Condition, the subject itself must be an independent whole tree at some point of the derivation.

14) The man loves Mary.

In this simple sentence, we need two independent trees-one [love[Mary]] and the other [the[man]]. Then, these two merge to form the whole sentence [[the[man]][love [Mary]]. But the following derivation is excluded by the Extension Condition:

- 15) a. [love[Mary]]
 b. [[man][love[Mary]]]
 c. [[the[man]][love[Mary]]]

The operation in (15c) violates the Extension Condition because “the” is inserted inside the tree instead of merging with its root. So, if subject is an independent tree before it merges with the VP, sideward movement should in principle be possible from subjects in a manner similar to the movement which Nunes (1995) proposes from the adjunct to the main clause.

3. Do sentential subjects exist?

3.1 Why subject sentences don't exist? : Koster (1978)

We have shown in the last section that movement out of the subject should be possible. Then, since we are focusing on PRO gate phenomena which are involved with sentential subjects, the next question to follow is whether sentential subjects really exist. In this section and the next one, we will discuss this question in detail.

Since Koster (1978), whether sentential subjects exist has been under debate. He argues that sentential subjects do not exist but they are topic clauses which are base-generated in the specifier of CP in the current sense and bind the phonologically null pronoun in the subject position. He introduces several cases originally from Emonds (1970), where Extraposition is obligatory.

i. Extraposition is usually obligatory in subordinate clauses:

- 16) a. *That for Bill to smoke batters the teacher is quite possible.
 b. *Although that the house is empty may depress you, it please me.

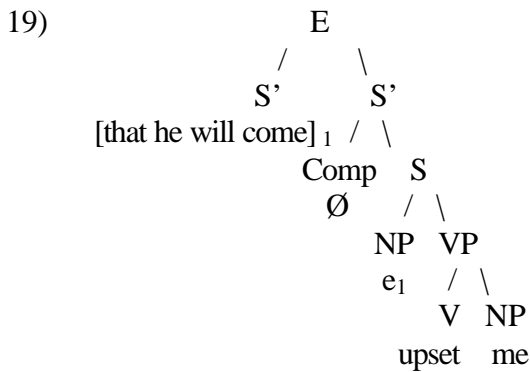
ii. Extraposition is obligatory after Subject Aux Inversion

17) *What does that he will come prove?

iii. Extraposition is obligatory after preposing like Topicalization

- 18) a. That he reads so much does not prove such things.
 b. *Such things that he reads so much does not prove.

Given these anomalies, Koster (1978) argues that the sentential subjects actually sit in leftmost daughters of E, the specifier of CP in the current sense and it binds the phonologically null pronoun in the subject position as (19):



With this structure, the first anomaly is solved, because sentences involved with topicalization are never embedded in Dutch and only very restricted conditions in English. The second one is solved, because the sentential subject is higher than inverted AUX sits. The third one is solved, because the sentential subject fills as high a place as the place in which the topicalized NP is supposed to be. Therefore, the topicalized NP cannot precede the sentential subject.

3.2 But sentential subjects do exist

Here, we show that what Koster (1978) argues is not quite right. According to Delahunty (1983), when the clause is relatively shorter than the predicate around which it is extraposed, the first and second anomalies disappear:

i. Sentential subjects can be embedded:

- 20) a. I think that for Bill to remain would so upset so many people that he and everybody else would be very much more comfortable if he left quietly but immediately.
 b. That for Bill to smoke bothers the teacher is not only possible, but quite certain, given what we know of his brand of tobacco, and of the teacher's asthma.
 c. It seems that that Fred left early so bothered all of the people who have been waiting for him that they now refuse to do business with him.

Furthermore, clausal gerunds can be embedded freely.

- 21) a. I think that Bill('s) smoking bothers the teacher.
 b. Although John ('s) kissing Mary may upset her parents, it pleases his.

ii. Extraposition is not obligatory after Subject Aux Inversion.

- 22) a. To what extent did that Fred failed to show up anger those of his devoted fans who had waited by the stage door since dawn of the previous day?
 b. Why does that Fred wants to marry her so upset Mary's mother, father, brothers, sisters and four grandparents that they haven't ceased to harangue her about it since they discovered the proposal?

In addition, PRO gate sentences, which include clausal gerunds, serve as examples which allow Subject Aux Inversion.

- 23) a. Who did cooking his lunch annoy?
 b. Who did getting his car fixed upset?
 c. Who did forgetting what he said to him annoy?

iii. Extraposition can not salvage the sentence after preposing like topicalization, anyway.

24) *Such things it doesn't prove that he reads so much.

Though Koster (1978) shows that the sentential subject can not tolerate topicalization in (18), the sentence (24) which undergoes extraposition is still ungrammatical. This nullifies his argument that Extraposition is obligatory after topicalization.

Delahunty (1983) also provides support for the existence of sentential subject. Sentential subjects can appear in the subject position of the embedded infinitive clause:

- 25) a. Bill wants that Fred lied to be obvious to everyone.
 b. I don't want [[to win games] to be our only goal].¹

Delahunty points out that we can not do topicalization there:

- 26) a. Bill wants to give a raise to Fred.
 b. *Bill want [to Fred]₁ to give a raise t₁.

If this is correct, Koster's argument that sentential subjects are topics is weakened.

Finally, Koster (1978) points out that sentential subjects can not be passivized. Then he argues that this is one of the most important points in favor of his analysis.

- 27) a. That the children are always late shows the necessity of discipline.
 b. * The necessity of discipline is shown by that the children are always late.

However, at least clausal gerunds can be passivized:

- 28) a. The children ('s) always being late shows the necessity of discipline.
 b. The necessity of discipline is shown by the children ('s) always being late.

¹ From D. Lightfoot in his lecture.

Furthermore, Pires (1999) points out that clausal gerunds can not tolerate extraposing like nominal subjects.

- 29) a. *It surprised me Mary receiving a medal
 b. *It surprised me the news.

We, therefore conclude that some subject sentences do exist (at least clausal gerunds), contrary to what Koster (1978) proposes.

4. On psych verbs

4.1 The problem that Psych verbs present

There are some attractive hypotheses that the thematic information and the structural information are uniformly connected. Baker's (1988) UTAH is one of them:

30) UTAH (=Uniformity of Theta Assignment Hypothesis):

Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-Structure.

For example, in the sentence, "John hit Bill.", we know that "John" is the Agent connected to subject position and "Bill" is the Patient connected to object position. Baker (1988) and others hypothesize that this kind of relationship should be extended to other lexical items.

However, Psych-verbs, which describe psychological states are problematic for this kind of hypothesis. Generally, they discharge the two theta-roles, the Experiencer and the Theme. But the relationship between these theta roles and the syntactic configuration is arbitrary.

- 31) a. John's comment worried Mary.
 b. Mary worried about John's comment.

In this pair, (31a) has the Theme on the subject position and Experiencer on the object position, but (31b) has the Experiencer on the subject and the Theme on the object. It seems that UTAH problematically holds for Psych verbs.

4.2 Belletti & Rizzi's (1988) analysis

In defense of the uniformity hypothesis, Belletti & Rizzi (1988) argue that both of the sentences which include a Psych verb should be derived from the structure in (32) where the Experiencer is in the subject position and the Theme is in the object position.

- 32)
- $$\begin{array}{l}
 \text{VP} \\
 / \quad \backslash \\
 \text{V}' \quad \text{Experiencer} \\
 / \quad \backslash \\
 \text{V} \quad \text{Theme}
 \end{array}$$

Then, overt position of the Theme results from movement in a very similar to unaccusatives.

- 33) a. John's comment worries Mary.
 b. [[John's comment]₁ [VP [V worry [t₁]] Mary]]
- 34) a. The ice melt.
 b. [[The ice]₁ [VP[V melt [t₁]]]]

So-called "backward-binding" is the key to their argument that the Experiencer always occupies a position higher than the Theme at D-structure.

- 35) a. These rumors about himself₁ worry Gianni₁ more than anything else.
 b. Each other₁'s comments annoyed [John and Mary]₁.

The antecedents do not appear to bind their anaphors in these examples, violating Binding condition A. Belletti & Rizzi (1988) argues that if (36) is the D-structures of (35a) and (35b), we can account for this phenomenon, with the assumption that Binding condition A is applied at D-structure, or via reconstruction at LF, given current assumptions.

- 36) a.[VP [worry [these rumors about himself₁]] Gianni₁]
 b.[VP [annoy [each other₁'s comment]] [John and Mary]₁]

This seems to strongly support their proposal. If what they argue is true, the PRO gate phenomena which have Psych verbs are irrelevant to the issue of this paper, namely movement out of the subject. Since the Theme of Psych-verbs always occupies the object position in their analysis, the surface subject of (37) could move from the object position.

- 37) a. Who does [PRO cooking his lunch] annoy ?
 b.
- $$\begin{array}{c}
 \text{VP} \\
 / \quad \backslash \\
 \text{V}' \quad \text{who} \\
 / \quad \backslash \\
 \text{V} \quad [\text{PRO cooking his lunch}]
 \end{array}$$

4.3 The argument in Pesetsky (1995)

Pesetsky (1995) extensively criticizes this argument made by Belletti & Rizzi (1988). Here we mainly discuss his critique about backward binding phenomena because they seem to serve as the strongest evidence for the argument that the surface subject begins in the object position in Psych verbs.

Pesetsky (1995) shows that not only Psych verbs but also causatives allow this phenomena.

- 38) a. Pictures of ourselves₁ caused us₁ to start laughing.
 b. Each other₁'s comments forced [John and Mary]₁ to confront their problems.

Therefore, backward binding phenomena cease to be the evidence for the structure of Psych verbs which Belletti & Rizzi (1988) proposed, because these phenomena are not peculiar to Psych verbs.

In addition to (38), backward-bindings appear in the causative structures where the causative verb does not assign any theta-role to the “causee”.

39) [Each other’_{2s} remarks]₁ made t₁ [John and Mary]₂ seem t₂ to be angry.

“John and Mary” starts from the lowest clause in (39), receiving its theta-role from “angry”. Therefore, even if we suppose that “each other’s remarks” starts from the object position in the similar way to unaccusatives as in (39), there is no chance for it to c-command the reciprocal at D-structure. But backward binding still holds. This argues against Belletti & Rizzi’s (1988) proposal.

Moreover, there are other reasons to reject Belletti & Rizzi’s proposal. Recall that they consider Psych verbs as species of unaccusatives. The questions to follow are (i) why the Theme can be raised to the subject position, skipping over the Experiencer, which sits in the external argument in apparent violation of shortest Move and (ii) how to get away with the violation of Burzio’s generalization, which says that any verb which does not discharge external theta-role does not assign accusative Case. Their solution is that unaccusative verbs may assign inherent accusative Case which is linked to a theta-role, together with the assumption that Burzio’s generalization applies only to structural Cases. Thus, in their system Psych verbs may assign inherent accusative Case to the Experiencer argument. However, going back to (39), even if we allow the causative verb “made” to assign inherent accusative Case to the nominal “John and Mary”, it receive the theta role from “angry” in the lowest clause as we have already seen. Therefore, “John and Mary” can not receive inherent accusative Case from the verb because they are not thematically related to each other at D-structure.

These arguments made by Pesetsky (1995) I believe scotch the proposal by Belletti & Rizzi (1988). Added to these, in the light of Minimalism, Belletti & Rizzi (1988) would need the special assumption that inherent Case does not block NP-movement to get its structural Case checked. Otherwise, the Theme can not move to the subject position in place of the Experiencer, which sits the higher position within the vP. But now the benefits of their proposal are almost gone, given the required stipulation needed to make it work. We might as well assume that the subjects in Psych verbs always start from the external argument position, i.e. Spec of vP in the current sense.

4.4 Further argument against Belletti & Rizzi (1988)

The fact that Psych verbs show WCO-effects is unexpected by Belletti & Rizzi (1988).

40) ??Who₁ did his₁ shaving his₁ face annoy t₁?

According to their analysis, the Experiencer, “who” is supposed to bind the pronoun in the Theme at D-structure. So, we would predict that WCO-effects should not appear, contrary to the fact. Furthermore, It is reported that raising constructions do obliterate WCO-effects², unlike Psych verbs.

- 41) a. [A picture of his₁ dog]₂ seemed to everyone₁ to t₂ be out of focus.
 b. [His₁ mother]₂ is believed by every boy₁ t₂ to be a saint.

In these examples, the raised subject starts from the lower clause with bound by the quantifier at D-structure as Belletti & Rizzi’s (1988) analysis on Psych verbs. Then, these raising sentences dilute

² See Hornstein (1995 p159).

WCO-effects, contrary to the case of Psych verbs. This fact makes Belletti & Rizzi's (1988) proposal more implausible. Furthermore, their proposal can not account for the irregularity of the occurrence of PRO gate phenomena.

- 42) Who₁ did PRO₁ shaving his₁ face annoy t₁?
 43) a. Who₁ t₁ thinks that PRO₁ shaving his₁ face is mandatory?
 b. ??Who₁ did the fact that PRO₁ shaving his₁ face is mandatory annoy t₁?

It is problematic for Belletti & Rizzi that (43b) shows WCO-effects unlike (43a). Though in their analysis, "who" in the external argument in (43b) binds PRO and the pronoun in the internal argument at D-structure like (43a), the sentence displays WCO-effects.

4.5 Anaphors in backward binding are nonlocal anaphors

Having denied Belletti & Rizzi's (1988) proposal, we need an alternative account for the backward binding. According N. Hornstein (p.c.), it is doubtful whether the anaphors in the backward binding are local ones. Lebeaux (1984-85) observes that there are two types of anaphors, local and nonlocal. The properties of local anaphors and nonlocal anaphors can be distinguished uniformly in several ways:

(i) Split antecedents: Nonlocal anaphors may have split antecedents but local anaphors may not.

- 44) a. John₁ told Mary₂ that there are some pictures of themselves₁₊₂ in the building. (nonlocal)
 b. *The men₁ introduced the women₂ to themselves₁₊₂. (local)

(ii) Free variation with pronouns: Many nonlocal anaphors appear in free variation with pronouns but local anaphor can not.

- 45) a. John₁ knew that there were some pictures of himself/him₁ in the building. (nonlocal)
 b. The man₁ likes himself/*him₁. (local)

(iii) C-command: Nonlocal anaphors do not have to be c-commanded by the antecedent but local anaphors have to.

- 46) a. Bush₁'s campaign said that pictures of himself₁ were placed all over town. (nonlocal)
 b. *Bush₁'s campaign admires himself₁. (local)

(iv) Interpretation in VP ellipsis: Non local anaphors allow both strict and sloppy readings inside the elided VP but local anaphors only allow a sloppy reading.

- 47) a. Bush₁ thought that there were some pictures of himself₁ in the building, and Gore₂ did too. (nonlocal) (=Gore did <think that there were some pictures of him_{1/2} in the building.>)
 b. Bush₁ seated himself₁ and Gore₂ did too. (local) (=Gore did <seat himself_{2/*1}.>)

Let me apply to the anaphors in backward bindings the first three of these four properties which differentiate nonlocal anaphors from local ones. Since subjects cannot undergo VP ellipsis, the

fourth property is irrelevant. As a beginning, we will examine split antecedents. The anaphors in backward bindings do allow the split antecedents.³

48) Nude pictures of themselves₁₊₂ in the Time made Hillary₂ tell Bill₁ that he should resign.

This is the first piece of evidence that these are nonlocal anaphors. And we already know that they are not c-commanded by the antecedent. Therefore, they meet the second condition. Finally, they in fact alternate with pronouns. See the next pair of sentences.

49) a. Those stories about herself₁ made Bill hate Mary₁.⁴
 b. Those stories about her₁ made Bill hate Mary₁.

(49b) is no worse than (49a). This also implies that reflexives in backward bindings are nonlocal anaphors. To sum up, the anaphors in backward bindings suffice all properties of nonlocal anaphors should possess. Thus, if they are nonlocal anaphors, they need not to be c-commanded by the antecedent.

4.6. Relaxing the tension

Finally, one may notice that the tension between Psych verbs and a hypothesis like UTAH still remains after we rejected Belletti & Rizzi's proposal. Addressing this lingering tension, Pesetsky (1995) shows that the pair of sentences in (50) need not be treated as twins because their meanings are not identical. He suggests a division of the Theme role into Subject Matter and Causer.

50) a. John worried about the TV set.
 b. The TV set worried John.

According to Pesetsky (1995 p57), in (50a), "whatever John was experiencing the worry described in the example, he was thinking in some way about the television set... Whatever the nature of John's specific concern, *the television set* is the Subject Matter of Emotion." But in (50b), "on the other hand, the DP *the television set* bears the...role of Causer, It is sufficient that the television set causes John to experience worry, but the Subject Matter of his thoughts while experiencing worry could have nothing to do with the television set. There is simply a casual relationship between the set and some state of worry."

Then, he proposes that Psych verbs may assign either the Subject Matter role to the object position or the Causer role to the subject position, instead of assigning the Theme to object position uniformly.

51) a. John worried about the TV set. (Subject matter) assigned to object (i.e. internal argument)
 b. The TV set worried John. (Causer) assigned to subject (i.e. external argument)

This proposal can relax the tension between Psych verbs and the hypotheses like UTAH, and make more sense of the assumption that the subjects in Psych verbs always start as the external argument.⁵

³ From N. Hornstein (p.c.)

⁴ From Pesetsky (1995) note #185.

In this section, we have shown that Phych verbs are not a variety of unaccusatives, having denied the proposal by Belletti & Rizzi (1988). Furthermore, we have proved that anaphors in backward bindings are nonlocal, which do not require any c-commanding antecedent. Given what has been said, subjects in Psych verbs can be simply regarded as the genuine subjects which start from the external argument position, namely Spec vP.

5. Justifying derivations

5.1 How to Cancel WCO-effects

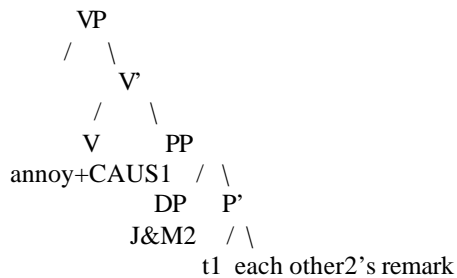
The last two sections let us suppose that sentential subjects including the ones in Psych verbs, are in fact external arguments, which start from Spec vP. So, now we are ready to consider how the movement out of the subject accounts for PRO gate phenomena we have introduced in the section 1.

Recall that A-movement cancels WCO-effects. To put it more precisely, when A-movement makes the copy of the Q-element precede the pronoun, WCO-effects disappear:

- 52) a. ??Who₁ does it seem to his₁ mother t₁ is handsome?
 b. Who₁ t₁ seems to his₁ mother t₁ to be handsome.
- 53) a. ??John-ga [Mary-ga pro₁ atta]-atode [dono hito₁-ni mo] atta.
 b. John-ga [dono hito₁-ni mo]₂ [Mary-ga pro₁ atta]-atode t₂ atta.

Assuming this is descriptively accurate, WCO-effects are canceled in the configuration as in (54).

⁵ Pesetsky (1995) concludes that these Phych verbs and causative structures, which show backward bindings have the structure below, dubbing it as cascade structures:



Then, “each other’s remark”, the complement of CAUSER (zero-morpheme), preposition moves to the external argument position, while the zero-morpheme incorporate into the matrix verb. However, this analysis fails to solve the problems that Belletti & Rizzi (1988) has. It is hard to see why the lowest nominal can move all the way to the external argument position of the matrix verb, especially, in these sentences where the subject on the surface must start from the bottom of the tree in order to be bound by the antecedent at D-structure:

- [Each other’s remarks]₁ made [John and Mary]₂ seem t₂ to be angry t₁.
 [Those books about himself]₁ taught Bill the meaning of caution t₁.

Furthermore, his proposal also fails to explain why WCO-effects arise in Phych verbs and the irregularity of PRO gate phenomena. Thus, we do not adopt his ultimate proposal.

- 54) $Q_1 \dots Q_1$ in $A_{\text{position}} \dots \text{pronoun}_1 \dots \text{variable}_1$ (when the pronoun and the variable are not in C-command relation)

5.2 Justifying the Derivations

Let us see how all I have said so far, works. Compared with (56), the sentence (55) has the antecedent (the controller) of PRO in the higher clause than the clause in which PRO occurs. On the other hand, in (56) the antecedent (the controller) of PRO is within the same clause, namely in the object position of the verb.

55) John₁ thinks that it is believed that PRO₁ (=pro) shaving himself is important.

56) Who₁ did [PRO₁ cooking his₁ lunch] annoy t₁?

Keeping this in mind, let us take sentence (56) at first and start the derivation from the subject:

- 57) a. [IP who₁ [vP who₁ cooking [his lunch]]] (building up the subject)
 b. [IP who₁ [vP who₁ cooking [his lunch]]] annoy (taking “annoy” from the numeration)
 c. [IP who₁ [vP who₁ cooking [his lunch]]] [vP annoy [who₁]] (sideward-movement)
 d. [vP [IP who₁ [vP who₁ cooking [his lunch]]] [annoy [who₁]]] (subject merging with vP)
 e. [CP who₁ [TP [IP who₁ [vP who₁ cooking [his lunch]]] [annoy [who₁]]]] (Wh-movement)
 f. [CP who₁ [TP did [IP who₁ [vP who₁ cooking [his lunch]]] [annoy [who₁]]]] (do-support)
 g. [who₁ did [[~~who~~_± cooking [his lunch]]] [annoy ~~who~~_±]] (deleting copies all but one)

In (57a), the subject tree undergoes the merge operations in which “who” starts from the specifier of vP and then moves to the subject position of the clausal gerund. So, if the derivation converges, this copy obviates WCO-effects. (57b) picks up “annoy” from the lexical array. In (57c), here “who” moves sideward to merge with “annoy”. Why is the movement out of the subject allowed here? Recall that this kind of operation is supposed to violate the CED. However, at the point where sideward movement applies the constituent we have moved from has not yet been merged and so need not be considered a subject. In effect, at this point in the derivation the CED will not apply. This can be seen if we adapt Barriers’ interpretation of the CED to the derivational approach to grammar, which I assume here.

Suppose CED is the reflex of moving over more than one barrier (cf. Chomsky (1986b)). After it merges with the main clause, the subject becomes a ‘barrier’ since it is not L-marked, in other words it is not a sister of any lexical head but the specifier of vP. The IP, which dominates the subject also becomes a barrier. And we may not move anything over two barriers. Therefore, we cannot extract anything from inside the subject beyond the IP. On the other hand, the sideward movement shown in (56c) does not go across two barriers. Even if the sentential subject itself is a barrier since it is not L-marked, the movement to another tree does not step over any barrier other than it. So, this movement should be licit because it just goes across one barrier. Notice that, crucially, we must execute the sideward movement before the subject is attached to the predicate because movement from the subject to the object position results in a violation of the Extension Condition, even if it just steps over less than two barriers. Namely, movement out of the subject to the object position does not target the root of the tree, if the sentential subject has already merged here. Thus, there is no way to move an element out of the subject to the object position after the subject tree merges with the predicate.

There is another interpretation of CED in effect. Uriagereka (1999) argues that a unit of locally total command (command unit =CU) is sufficient for linearization or Spell-out. In short, in order for the tree of the subject or adjunct to merge with another tree, it must be spelled-out. Since it is already shipped to the interface at this point, we cannot execute any further syntactic operation inside of the tree. This explanation of CED effects is also useful here, because it prevents the operation inside of the subject only after the subject is attached to the main tree. To put this the other way around, in Uriagereka's (1999) interpretation of the CED, we may move elements from one tree to another as long as they are not connected.⁶ Moreover, an advantage to adopting this approach is that we can exclude the representational interpretation of CED effects as Barriers. To sum, whichever approach we take, we are now able to see that we can move elements sideward from inside the subject only before it merges with another tree.

Going back to the derivation in (57), both of trees merge in (57d) where the subject is in the specifier of the matrix vP. Notice the next step. In (57e), Wh-movement occurs and here the moved Wh must be the lower one (=the one in the complement of "annoy") not in the subject. The subject already merged with the main clause, so we cannot extract anything out of the subject anymore. So, the lower one should move to the specifier of CP. This movement is valid because the Wh in the subject does not c-command the other one in the VP. It means that both Whs are equi-distant in the eye of the C⁰. Therefore, this movement is not a violation of Minimal Link Condition. Then, In (57f) and (57g), Do-support and the deletion of copies come into play to complete the derivation of the sentence. Note: given this derivation, PRO in the PRO gate sentence is a residue of A-movement. Thus, we account for why it is that in such sentences WCO-effects fail to arise.

Finally, one more question that arises is why we do not merge the whole subject with "annoy" at (57c), instead of moving "who" sideward. The problem is that this derivation seems to violate Merge Over Move. Here, I want to say that some kind of selectional (categorical or semantic) restriction matters. In particular "annoy" cannot select either an infinitival clause or a clausal gerund as its internal argument. We can predict that the verb should merge with the gerund as its complement in the sentence below in which "finish" selects a clausal gerund as its object.⁷

58) Who finished PRO getting his car fixed?

- 59) a. [who getting [[his car]fixed]]
 b. [who getting [[his car] fixed]] finished (taking "finish" from the numeration)
 c. [finished [who₁ getting [[his car] fixed]]] (merging the two trees)
 d. [who₁ [finished[who₁ getting [[his car] fixed]]]] ("who" moving to Spec vP)
 e. [who₁ [who₁ [who₁ [finished[who₁ getting [[his car] fixed]]]]]](A-movement and Wh-movement)
 f. [who₁ [~~who₁~~ [~~who₁~~ [finished[~~who₁~~ getting [[his car] fixed]]]]]]] (deleting copies all but one)

The next task is to explain why PRO in the sentence (55), which is re-introduced as (60), can not be the product of movement. As we have seen, the antecedent (the controller) of PRO is in the higher clause than the one that PRO occurs in (55). Therefore, long before the position that the antecedent (=John) sits is open, the subject that contains PRO merges with the clausal tree:

⁶ In fact, this is totally the opposite standing against the one that Epstein (1999) takes. We will come back this issue later.

⁷ See Pires (1999) for more detailed implementation of the derivation of this kind of sentence.

- 60) John₁ thinks that it is believed that PRO₁ (=pro) shaving himself is important.
 61) a. [[John shaving himself]is important] (the complete derivation of the embedded clause)
 b. [thinks [that it is believed that [[John shaving himself]is important]]]
 c. *[John₁ thinks [that it is believed that [[John₁ shaving himself] is important]]] (illicit movement)

In the stage of (61a), the subject is already an island since it is merged with the main tree. As we have seen, once a subject merges with another tree we can not extract anything from it. So, in (61b) when the matrix verb, “thinks” needs external argument, we cannot move “John” from inside the sentential subject in the lowest clause as shown in (61c) because this movement crosses several barriers. (or the sentential subject has already been Spelled-Out.) Therefore, PRO in (55) cannot be a residue of the movement unlike the one in (56).

Now, we can also see why WCO-effects appear in (62) where the clause that contains PRO is embedded:

- 62) ??Who₁ does the fact that PRO₁ cooking his₁ lunch is mandatory annoy t₁?
 63) a. [who cooking his lunch] (building the subject)
 b. [who cooking his lunch] is mandatory (taking ‘is’ and ‘mandatory’ from the numeration)
 c. [who cooking his lunch] [is mandatory] (building the predicate)
 d. [[who cooking his lunch] [is mandatory]] (merging the two)
 e. [the fact [that[[who cooking his lunch] [is mandatory]]]] (building up the subject)
 f. [the fact[that[[who cooking his lunch] [is mandatory]]]] annoy (taking ‘annoy’ from the numeration)
 g. *[the fact[that [[who₁ cooking his lunch] is mandatory]]] [annoy [who₁]] (illicit movement)

In (63a), we build the subject in the same way as we did in (57a). If the derivation converged, WCO-effects should not appear because the copy of ‘who’ in the subject position of the clausal gerund would intervene between the Q-element and the pronoun in the same way as in (57). At the point in (63d), when the subject merges with the vP, it becomes an island. So, when we need to move ‘who’ later in (63g), we may not extract it from inside the subject. Since this movement is prevented, PRO in this sentence cannot be the copy of ‘who’ in A-position. Instead, it is small pro, which is pronominal. Therefore, this sentence is prey to WCO-effects just like the sentence (64) or (65).

- 64) ??Who₁ did his₁ cooking his₁ lunch annoy t₁?
 65) ??Who₁ does the fact that his₁ cooking his₁ lunch is mandatory annoy t₁?

To put it more precisely, the sentence (62) involves two islands, the complex NP island and subject island. The next example shows that as we predict, a subject island is enough to prevent the movement:

- 66) ??[That [PRO₁ leaving early] upset his₁ friend] made nobody₁ feel lousy. (subject island).

In addition, we correctly predict that PROs confined within the other kind of islands also fail to be a gate.

- 67) a. ??The girl [whom PRO₁ shaving his₁ face fascinated] kissed everyone₁.⁸ (relative clause)
 b. ??[Any attempt PRO₁ to argue with his₁ mother] would leave everyone₁ exhausted. (complex NP island)
 c. ??Who₁ did [the question [whether PRO₁ to hit his₁ mother]] upset t₁? (complex NP island +Wh island)

Since in addition to subject island, the embedded sentential subject in each sentence is confined within other kinds of island (Relative clause island, Complex NP island or Wh-island), These data support the claim that PRO gate phenomena are a result of movement. Namely, since PRO is embedded in islands, it can not be a residue of A-movement but must be small pro, a pronoun. Thus, WCO-effects should arise in these examples in which PROs are confined within islands, just as in the sentences (28) and (29) in which an overt pronoun replaces PRO. On the other hand, in the cases where movement from inside subjects is possible via sideward movement as shown in (21), PRO must be a result of A-movement. This analysis can reduce PRO gate phenomena to the well-known fact that A-movement cancels WCO-effects.

5.3 Further discussion

Here, let us extend into other examples the argument made so far. What is wrong with these sentences below?

- 68) a. *Who₁ did [t₁ washing his₁ car] upset him₁.
 b. *Who₁ did [t₁ shaving himself] annoy John?⁹

In these sentences, “who” moves directly from inside the subject sentence. Thus, by the time the matrix C⁰ causes the Wh-movement, the subject already has become an island:

- 69) a. [Who washing his car] (building the subject.)
 b. [upset him] (building the main clause.)
 c. [[Who washing his car] [upset him]] (merging both a and b, a becomes an island at this point)
 d.*[Who₁ [[Who₁ washing his car] [upset him]]] (illicit movement.)

Therefore, the derivation cannot converge. One might argue that “Who” cannot get its Case checked in the course of the derivation and the sentence can be excluded in either way. However, it is unclear that subjects of gerunds need to get its Case checked especially when they are realized as accusatives.¹⁰

- 70) Bill washing his car upset John.

⁸ In this sentence, the subject, *The girl whom shaving his face fascinated* can merge with the verb first. But notice the derivation can not converge anyway:

- a. [the girl whom everyone shaving his face fascinate]
 b. kissed [the girl whom everyone shaving his face fascinated]
 c.*everyone₁ kissed [the girl whom everyone₁ shaving his face fascinated]

⁹ This is from Safir (1996)

¹⁰ From N. Hornstein (p.c.)

Given this, the sentences in (68) indicate that the island bars “who” from moving directly from the subject to the specifier of CP. We can observe the exact same phenomenon when we try to execute the explicit Wh-movement out of the subject sentence:

- 71) *Who₁ does it seem that acquitting t₁ causes a controversy?
 72) a. [pro acquitting who] (building the subject.)
 b. [causes a controversy] (building the VP.)
 c. [[pro acquitting who] [causes a controversy]] (merging both (a) and (b), (a) becomes an island at this point)
 d. *[Who₁ it seem that [[pro acquitting who₁] [causes a controversy]]] (illicit movement.)

In the derivation above, the subject of the embedded clause becomes an island when it merges with the vP in (72c). Thus, we cannot extract ‘who’ out of it at the later stage of the derivation in (72d).

Let us consider more complicated examples originally from Safir (1984).

- 73) a. *Who₁ did PRO₁ shaving his₁ face convince Mary to trust t₁?
 b. *Everyone who PRO₁ shaving his₁ face convinced Mary to trust t₁?

What is the problem with these examples? Considering the derivation of the sentence step by step gives us an answer. At first, the subject is built up to form [who [shaving [his face]]].¹¹ Again, if the derivation converged, “who” in the subject position on the clausal gerund would cancel WCO-effects. Then, the derivation picks up “trust” from the numeration. Notice that we cannot move “who” sideward out of the subject to merge it to “trust” because we still have “Mary” in the array. That is, this operation violates Merge Over Move. Hence, we must simply merge “Mary” with “trust” instead of moving “who” from the subject.

- 74) a. [who [shaving his face]] (building the subject)
 b. [who [shaving his face]] trust (picking up “trust” from the numeration)
 c. *[who₁ [shaving his face]] [trust [who₁]] (illicit movement, violating Merge over Move)

We correctly predict that the examples below are good.

- 75) a. Who₁ did PRO₁ inviting his₁ girl friend encourage t₁ to face Pete?
 b. Who₁ did PRO₁ shaving his₁ face convince t₁ to grow a beard?¹²
 c. Everyone who PRO₁ shaving his face convinced t₁ to date Mary

As for the sentence (75a), after we build the subject, [who [inviting his girl friend]] in (76a), we pick up “face” in (76b). At this point, unlike (74), we merge “Pete” with it instead of moving “who” out

¹¹ More specifically, to reach the step of (74a), we already undergoes a movement. That is, “who” which starts in the SPEC of vP presumably moves to the Spec IP within the subject. But this movement is licit. Even if we have “Mary” in the array, we cannot directly merge it to Spec IP, since this position is not a theta-position, following Chomsky (1998). Therefore we have no choice but to move “who” at this point of time.

¹² This is from Safir (1984).

of the subject in (74c). There is no violation of Merge Over Move. Then, “who” moves to the specifier of vP in (76d), The movement should be valid, for the numeration is exhausted. A few steps later, the subject merges with the main clause (76f). Finally ordinary operations take place to complete the derivation (76g) with the copy of “who” in the subject position of the clausal gerund inbetween the Q-element and the pronoun, canceling WCO-effects:

- 76) a [who [inviting his girl friend]] (building the subject)
 b [who [inviting his girl friend]] face (picking up “face” from the numeration)
 c [who₁ [inviting his girl friend]] [face [Pete]] (Merging “Pete” with the verb)
 d [who₁ [inviting his girl friend]] [who₁ [face [Pete]]] (sideward-movement)
 e [[who₁ [inviting his girl friend]][encourage [who₁ [to [who₁ [face [Pete]]]]]]]
 f [who₁ [[who₁ [inviting his girl friend]][encourage [who₁ [to [who₁ [face [Pete]]]]]]]]]
 g [who₁ did [[~~who₁~~ [inviting his girl friend]][encourage [~~who₁~~ [to [~~who₁~~ [face [Pete]]]]]]]]]
 (the deletion of copies and Do-support)

Given what we have seen in this section, we can regard PRO gate phenomena as an instance of A-movement. PRO which is a residue of movement cancels WCO-effects. On the other hand, PRO which can not be a residue of movement retains WCO-effects. This conclusion supports the treatment of PRO (as a residue of A-movement) and Small pro in Hornstein (1999). Importantly, the movement out of the subject, which is an application of sideward movement, drives the derivation outlined in this section. That is, to the extent that what we have shown is on the right track it argues that sideward movement out of a subject is possible and so supports the proposal in Nunes (1995) and Hornstein (forthcoming).

6. Are PRO in the PRO gate sentences OCPRO?

In the last section we have shown that PROs in PRO gate sentences are a residue of A-movement. Here, we provide independent evidence for this claim.

6.1 Obligatory Control

Hornstein (1999) introduces the various properties of Obligatory Control PRO, which he argues, is a residue of NP-movement.^{13 14}

- 77) a. *It was expected PRO to shave himself.
 b. *John thinks that it was expected PRO to shave himself.
 c. *John’s campaign expects PRO to shave himself.
 d. *John expects PRO to win and Bill does too. (=Bill win)
 e. *John₁ told Mary₂ PRO₁₊₂ to wash themselves/each other.
 f. The unfortunate expects PRO to get a medal.
 g. Only Churchill remembers PRO giving the BST speech.

(77a) shows that OCPRO must have an antecedent. (77b) shows that the antecedent must be local and (77c) indicates that it must c-command PRO. (77d) shows that OC PRO only licenses a sloppy

¹³ They are from Hornstein (1999) P73

¹⁴ (77a-e) originates from Lebeaux(1984-5). (77f) is discussed by Higginbotham (1992). (77g) originates from Foder (1975).

reading under ellipsis. (77e) shows that OC PRO cannot have split antecedents. (77f) shows that OCPRO only allows a *de se* reading. That is, in the sentence the unfortunate must believe that s/he, herself/ himself will get a medal. In other words, the sentence does not allow the interpretation that the unfortunate who turns out to be awarded a medal knows that someone other than him will get the medal. In (77g), Only+NP must be the controller of PRO. That is, the sentence has the only meaning that only Churchill could have this memory because he was the sole person who gave the BST speech.

6.2 Non Obligatory Control

Hornstein (1999) mentions that Non Obligatory PRO lacks the properties which OCPRO possess. The next sentences have instances of NOCPRO.

- 78) a. It was believed that PRO shaving was important
 b. Who₁ thinks that it is believed that PRO₁ shaving his face is important?
 c. Bush₂'s campaign₁ believes that PRO_{1/2} keeping his sex life under control is necessary.
 d. John thinks that PRO getting his resume in order is crucial and Bill does too.
 e. John₁ told Mary₂ that PRO₁₊₂ washing themselves would be fun.
 f. The unfortunate believes that PRO getting a medal would be boring.
 g. Only Churchill remembers that PRO giving the BST speech was momentous.

NOCPRO does not need an antecedent as shown in (78a). Its antecedent does not have to be local as in (78b). The antecedent does not have to c-command NOCPRO as shown in (78c). NOCPRO allows a strict reading. That is, Bill also thinks about John's resume in (78d). It permits split antecedents as shown in (78e). PRO in (78f) allows non-*de se* reading. In (78g), other people could remember that the BST speech was momentous.

Recall that Hornstein (1999) argue NOCPRO is not a residue of movement but a small pro. We show that movement can not generate NOCPRO, illustrating the derivation of (78b):

- 79) a. [who shaving his face] (building the subject.)
 b. [is important] (building the lowest predicate.)
 c. [[who shaving his face] [is important]]
 (merging both a and b, the subject becomes an island at this point!)¹⁵
 d. *[who₁ thought that [it is believed that [[who₁ shaving his face] [is important]]]]
 (illicit movement.)
 e. [who₁ thought that it is believed that [[pro shaving his face] [is important]]]
 (putting 'pro' in the SPEC IP)
 f. [who₁ [who₁ thought that it is believed that [[pro shaving his face] [is important]]]]
 (Wh-movement)
 g. [who₁ [~~who₁~~ thought that it is believed that [[pro shaving his face] [is important]]]]
 (the deletion of the copies)

In this derivation, 'who' in the subject of the embedded clause becomes an island in (79c) when it merges with the predicate. Then, at the point of (79d) where we need to put 'who' in the specifier of

¹⁵ The subject might as well merge to Spec AP of "important" if one prefer the small clause analysis. But it does not make any difference for the purpose of the discussion here.

the matrix vP, we are forced to move ‘who’ directly from the inside of the embedded subject, which is already an island. Since the movement is prevented, we should regard PRO in the sentence as something other than NP-trace, and following Hornstein (1999), it is small pro.

6.3 PRO in the PRO gate is OCPRO

Given the distinction between OCPRO and NOCPRO that we have seen, let us examine PROs in the PRO gate configurations, namely the configurations where the object, internal argument of the verb controls PRO in the specifier of the subject, external argument of the same verb. There are a few reasons to believe that PRO in such a configuration is indeed OCPRO, which is a result of movement. Consider the next example at first.

80) PRO_{1/2*/1+2*} going home early made Bill₁ kiss Mary₂.

(80) indicates that PRO in this sentence does not allow split antecedents. This is one property that OCPRO possesses. In addition, the controller must be ‘Bill’ the higher nominal not ‘Mary’, because of Merge Over Move: movement may occur only after the numeration is exhausted. The movement to the lower position (=the object position of ‘kiss’) out of the subject is illicit with one more nominal left in the numeration. The next pair of sentences shows a similar contrast.¹⁶

81) *a. PRO₁ washing herself caused John to hit Mary₁.
 b. PRO₁ washing himself caused John₁ to hit Mary.

Since the control relation depends on the constraint on movement in these examples above, they imply that PRO in the PRO gate configuration is a residue of movement, which explains the fact that PRO must be backward-controlled by the highest nominal.

Secondly, although the control relation in PRO gate configuration does not involve *e* command, the nominal embedded in DP may not be the controller. Compare the three examples below:

82) a. PRO_{1/2*} To have to give an important speech tomorrow makes [Bush₂’s campaign]₁ nervous.
 b. [Bush₂’s campaign]₁ expects PRO_{1/2*} to give an important speech tomorrow.
 c. [Bush₂’s campaign]₁ believes that PRO₂ keeping his sex life under control is necessary.

The nominal in the specifier of DP (=Bush) can not be an antecedent of PRO in the PRO gate configuration as shown in (82a). Instead, the whole DP (=Bush’s campaign) must be the controller.

¹⁶ These data are problematic to Pesetsky (1995)’s proposal (see ft.2). Since the surface subject starts from the very bottom in his proposed construction and assuming Binding condition A is applied at D-structure, it predicts an incorrect result.

(81) *a. [PRO washing herself]₁ caused John to hit Mary₁ t₂.
 a’. caused John to hit Mary₁ [PRO shaving herself]₁
 b. [PRO washing himself]₁ caused John₁ to hit Mary t₂.
 b’. caused to John₁ to hit Mary [PRO shaving himself]₁

As ‘Mary’ more locally binds the reflexive at his D-structure of (81a) in (81a’) rather than that of (81b) in (81b’), (81a) is analyzed to be the grammatical sentence, contrary to the fact.

This indicates that PRO in PRO gate configuration acts as OCPRO in (82b) rather than NOCPRO in (82c) where the nominal embedded in the DP can be the antecedent of PRO.

From a semantic point of view, PRO in PRO gate configuration only allows the *de se* reading. The sentence below requires the circumstance where the unfortunate should know that s/he herself/ himself will be awarded a medal.¹⁷

83) PRO receiving a medal annoyed the unfortunate.

As to “only NP constructions”, (84b) solely means that only Churchill was annoyed for he was the sole person to give the BST speech. Thus, it requires *only Churchill* to be the antecedent.

84) a. (??) His₁ giving the BST speech upset only Churchill₁.¹⁸
 b. PRO₁ giving the BST speech upset only Churchill₁.

In the pair of sentences, the truth-value of (84a) would be false if Churchill’s wife were upset by his speech. On the other hand, the truth-value of (84b) would be still true in the given circumstance because she did not give the BST speech. In short, viewed in the semantic light, PRO in the PRO gate configuration can be regarded as OCPRO.

Finally, according to N. Hornstein (p.c.) we may count the complementary distribution with pronouns as the diagnostic factor. He points out that obviation arises when an overt pronoun replaces PRO in PRO gate configuration.¹⁹

85) a. PRO₁ kissing Mary annoys John₁.
 b. *His₁ kissing Mary annoys John₁.

Lebeaux (1984-5) observes that NOC PRO does appear in free variation with pronoun and OC PRO on the other hand, is in complementary distribution.

86) a. John₁ knew that PRO/for him₁ to leave early would be difficult.
 b. John₁ expects PRO/him*₁ to leave.
 c. *John persuaded Mary for her to leave early.

In our account, PRO in (86a) is NOCPRO. We have already noticed that the movement from inside the subject to the higher clause is not allowed since the subject is merged with another tree by the point in the derivation that the higher clause induces the movement. This PRO appears in free

¹⁷ Contrary to (83) the sentence below allow the non *de-se* reading:

His receiving a medal annoyed the unfortunate.

¹⁸ Certainly, this sentence is not perfectly fine because of WCO-effects. But it is fine as long as “Churchill” is the antecedent, not “only Churchill”.

¹⁹ One might wonder if the issue here is really WCO-effects, since the sentence below might be irrelevant to WCO-effects, given the obviation in (85b).

Who₁ did his₁ cooking his₁ lunch annoy t₁?

However, Higginbotham (1980) introduces the data that show WCO-effects are at issue.

?* Mary’s seeing his₁ father pleased every boy₁.

?* Who_{1/2} did her forgetting what he₁ said to him₂ annoy t₁?

These data show that WCO-effects arise regardless of the status of the subjects of gerunds.

variation with the overt pronoun. However, it does not happen in OCPRO as shown in (86bc). On the other hand, curiously, the overt pronoun may appear where the control relation is prohibited.

- 87) a. *PRO₁ going home early made Bill kiss Mary₁.
 b. Her₁ going home early made Bill kiss Mary₁.
 88) a. *PRO₁ shaving himself convinced Mary trust John₁.
 b. Him₁ shaving himself convinced Mary to trust John₁.

This suggests that when movement is possible, pronominalization should be barred because it is more costly. That is, the pronoun may allow coreference with the DP only when movement is impossible.

To sum up, from the evidence surveyed, we conclude that PRO in PRO gate configurations belongs to OCPRO. What we have seen so far enables us to solve a long-lived puzzle i.e. Control of the PROs in the subjects in which anti-c-command and backward binding are concerned:

- 89) a. PRO₁ cooking his lunch annoys John₁.
 b. PRO₁ shaving himself is important for John₁.

We have shown that this phenomenon is one of Obligatory Control, one of the possible syntactic relations, even though the two elements are not in c-command relation. And we have argued that it should be regarded as a mere instance of A-movement.

7. On how to delete copies

Let me take a close look at the sentence discussed in the last section.

- 90) PRO₁ cooking his lunch annoys John₁.

We have concluded that PRO in (90) is a residue of a movement. The derivation illustrated in (91) raises a question.

- 91) John₁ shaving himself annoys John₁.

Notice that in this construction, the pair of the copies of “John” are not in a c-command relation. Nunes (1995) argues that the deletion of copies, more generally syntactic operations, require chains that represent c-command relations. Therefore, if we followed Nunes (1995), we could not delete either copy as they are not parts of a common chain (at least if chains are connected via c-command (cf. Nunes (1995 p 84)). This would cause an LCA violation at PF and prevent the derivation from converging.

Let us look into Nunes (1995)’s implementation of copy deletion, in detail.

- 92) *You read every book₁ before Fred reviewed t₁.

The derivation of (92) proceeds as follows:

- 93) a. [before Fred reviewed [every book]] (building the adjunct)
 b. read [before Fred reviewed [every book]] (picking up the “read” from the numeration)
 c. [read [every book]] [before Fred reviewed [every book]] (moving “every book” sideward to the main clause)
 d. [[read [every book]] [before Fred reviewed [every book]]] (merging these two trees)
 e. [you [[read [every book]] [before Fred reviewed [every book]]]] (keeping building the tree to the end)

After completing the derivation, we need to delete one of the copies of “every book”, otherwise this would violate the LCA at PF. Nunes (1995) suggests that we can delete copies only when they form chains and chains require copies to be in c-command configuration. In the sentence above, the two copies of “every book” do not form chain, and we cannot delete either of them. Unless a derivation deletes all copies but one, the derivation will not converge.

Turn now to (94), a Parasitic Gap construction:

- 94) Which book₁ did you read t₁ before Fred reviewed t₁?

The derivation of this sentence goes as follows:

- 95) a. [before Fred reviewed [which book]] (building the adjunct)
 b. read [before Fred reviewed [which book]] (picking up the “read” from the numeration)
 c. [read [which book]] [before Fred reviewed [which book]]
 (moving “which book” sideward to the main clause)
 d. [[read [which book]] [before Fred reviewed [which book]]] (merging these two trees)
 e. [you [[read [which book]] [before Fred reviewed [which book]]]] (keeping building the tree)
 f. [which book did [you [[read [which book]] [before Fred reviewed [which book]]]]] (Wh-movement)

After completing the derivation, the copy of “which book” in the specifier of CP c-commands both lower copies. It enables us to delete these copies. The next question is why the highest copy survives at PF. Nunes (1995) argues that if we assume that Wh-feature of the nominal expressions is uninterpretable at PF, there is no choice but to keep the copy in the specifier of CP. Since the rest of the copies retain at least one more uninterpretable feature at PF, the highest copy is the only one to get its Wh-feature checked (with +wh C^0). Thus, if we retained either one of the lower copies instead, we would have to get rid of more unchecked uninterpretable features, namely Wh-feature here than if the higher copy were retained.

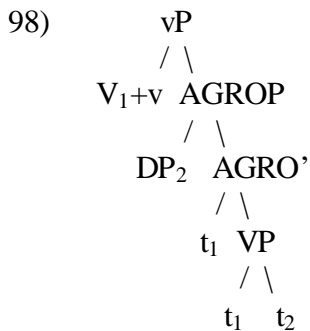
Hornstein (forthcoming) tries to exclude (92) without chains. He suggests that the deletion of the copies be executed “deterministically”. He argues that if, as we have seen, there is no choice in terms of which copy survives, it is redundant to resort to a c-command relation when executing the deletion of copies. In his reasoning, we cannot derive sentence (92) because the two copies have the same status in terms of PF restrictions. Assuming that theta-features are invisible at PF, these two copies in (93e) have the same feature composition. The only difference between them is that the one in the matrix clause receives an extra theta role. That is, we cannot decide which one is to survive at PF. Therefore, if the deletion of copies has to be deterministic as Hornstein (forthcoming) argues, we also cannot delete any copy in (92), which results in the violation of the LCA.

Given what have been said, let me turn back to the sentence (91) repeated in (96) whose complete structure of the derivation is (97).

96) PRO₁ shaving himself annoys John₁.

97) [[John₁ shaving himself] annoys John₁]

By adopting Hornstein's (forthcoming) approach to copy-deletion in place of Nunes' chain based requirement, we can regard PRO in (96) as a result of movement. One remaining problem is what makes difference between these two copies in (97)? In other word, why are we forced to retain the lower copy? Both of the copies seem to have the identical status with the exception that the lower one has a theta-feature. If this were the relevant difference, we would incorrectly allow the sentence (92). The surviving copy in (92) has one more theta-feature than the deleted one. Recall that we must have assumed that theta-feature is not visible at PF, to avoid the unwanted result. To overcome this problem, I will adopt the approach proposed by Johnson (1991), Koizumi (1993, 1995) and Lasnik (1998) that the Object Case is checked overtly in English. In other words, there exists overt object shift in English. For example, Koizumi (1993, 1995) developing the proposal by Johnson (1991), argues that objects move to the specifier of AGROP in the middle of VP shell where Case features get checked and verbs move to small v overtly.



Combined with this kind of over object shift approach, adopting the deterministic deletion allows us to retain the only copy in the object position in (97). With the copy in the object checked its Case feature overtly, it possesses less uninterpretable features than the copy in the subject does. Thus, we can say the lower copy is the one to survive in the two copies in (97). This also enables us account for these PRO gate sentences below, which do not involve overt Wh-movement.²⁰

99) a. PRO₁ seeing his₁ father pleased [every boy]₁.

b. PRO₁ to visit his₁ old neighborhood would embarrass [someone I know]₁.

c. PRO₁ getting letters from their₁ sweethearts is important for [many of the soldiers]₁.

100) a. [every boy seeing his father] (building the subject)

b. [every boy seeing his father] pleased (taking "please" from the numeration.)

c. [every boy₁ seeing his father] pleased everyboy₁ (sideward movement)

d. [[every boy₁ seeing his father][pleased every boy₁]] (merging these two trees.)

²⁰ They are from Higginbotham (1980) P688

After completing the derivation, we need to say that the lower copy of “every boy” is the one to survive. Then, if the deletion of copies requires chains as Nunes (1995) argues, this derivation should cause an LCA violation at PF. Since the copies are not in c-command relation we can not delete the copy in the subject. On the other hand, the deterministic approach allows us to retain only the copy in the object position, with the same reasoning as we treat sentence (90). Since the derivation converges with a copy of the Q-element in the subject position of the clausal gerund intervening between Q-element in the operator position and the variable, we have PRO gate phenomena in these sentences, cancellation of WCO-effects.

8. On representation approach and derivational approach

8.1 On chains and representational approaches

We have argued that the control from the object to inside a sentential subject is an instance of Obligatory Control and should be regarded as the result of NP-movement, as in (101).

- 101) a. PRO₁ cooking his lunch annoys John₁.
 b. PRO₁ cooking his lunch annoys everyone₁.

If this is correct, “chain” in the traditional sense is not a primitive notion in syntax, contrary to Brody (1995), Nunes (1995), in favor of Hornstein (1999, forthcoming), and Chomsky (1998).

Following Brody (1995 p6), chains in the traditional sense are defined as “an ordered set where every member binds the next and every member except the first is a nonpronominal empty category”. With this notion of chain, Brody (1995) explores a representational approach to grammar. He points out that with the assumption that movement is constraint by c-command, the system which exploits both chains and movement is conceptually suspicious because they correspond to each other so strictly that such a system would incorporate too much redundancy. He, then, argues that if the notion of chain is necessary at LF, or if chains are the sole licensed elements at LF as Chomsky (1991) assumes, then movement must be the one that we should discard. However, if we adopt sideward movement, which is not restricted to c-commanding expressions, then, in PRO gate configurations, where the controller and PRO are never in c-command relation (unlike Parasitic Gap construction), we get into trouble if we assume that a chain is needed as a licensing condition for movement. That is, chains are not only redundant or unnecessary but also they are empirically undesirable. It is clear that the notion of chain cannot cover the phenomenon which Movement can account for. Thus the notion of chains is the one that we should discard, given Brody’s methodological observations. Since the notion of chain is necessary for a representational approach to grammar as Brody (1995) notes, this conclusion is very problematic for this kind of approach to syntax. If our analysis above is correct, it provides the following challenge to a representational approach to grammar: to redefine the well-formedness conditions of chain without c-command. Failing this, chains will be empirically inadequate and, so dispensable. Without chains as primitive, notions like the Chain Condition (102) lapse into mere empirical generalization without any explanatory value.

102) The Chain Condition (Chomsky (1995) P46):

Every argument chain must be headed by a Case position and must terminate in a θ -position.

To sum, if, as argued above, PRO gate phenomena are the result of sideward movement from sentential subject it suggests that derivational approaches to grammar are superior to representational alternatives.

8.2 On c-command and derivational approach

Epstein (1999) argues that c-command is a primitive in syntactic operations, and he tries to deduce it in terms of a strict derivational approach to grammar. His derivational definition of c-command is as follows:

103) Epstein (1999 p.329) Derivational c-command:

X c-commands all and only the terms of the category Y with which X was paired *by Merge or by Move* in the course of the derivation.

However, as Pesetsky (1995) and others have already pointed out, the control relation we have seen here is an apparent exception to the c-command property of syntactic relations. And we have accounted for it with sideward movement, which is a non-c-command operation. If so, one might ask why c-command is around in ordinary cases. The reason is that we usually deal with a single rooted tree. If we assume that the Extension Condition holds, it forces the moved element to land in a c-commanding position. There is no other choice in such derivation.²¹ However, in addition to PRO gate phenomena, we can find a few more exceptions of the anti-c-command relation.

- 104) a. Which book₁ did you file t₁ before reading t₁ (Parasitic Gap)
 b. Which book₁ did John file t₁ and Mary read t₁? (Across The Board Extraction)
 c. John read every book₁ before it₁ was published.
 d. Someone from every city₁ hates it₁. (Some cases of bound pronoun)

In the first two examples, each variable is not in c-command relation with other. The quantifiers in (104cd) successfully bind the pronouns without c-command relating them.

Epstein (1999) claims that “no relations hold between members of two trees that were unconnected at any point in the derivation”. Then he introduces the First Law as a derivational law.

105) The First Law (Epstein (1999) P334):

- T₁ can enter into c-command (perhaps, more generally, syntactic) relations with T₂ only if there exists no derivational point at which
- T₁ is a term of K₁ (K₁ ≠ T₁), and
 - T₂ is a term of K₂ (K₂ ≠ T₂), and
 - There is no K₃ such that K₁ and K₂ are both terms of K₃.

This clearly contrasts with what we have been assuming. Recall that movement out of the subject is possible only before it merges with another tree, via sideward movement, and we can not extract any element out of the subject after the merge. In other words, we need a point where two trees are unconnected in the derivation in order to execute sideward movement.²² Worse, PRO gate phenomena never get into c-command relations at any point in a derivation. Epstein (1999) attempts

²¹ From N. Hornstein (p.c)

²² See Nunes (1995) P85 foot note 50 for the similar argument.

to attribute to the first law the consequence that “there are no relations between members of the specifier and members of X”. This is empirically inappropriate. PRO gate phenomena serve as an obvious counter example.

So far, we have provided a few examples in which syntactic relations hold without c-command. Thus, c-command may not be a primitive notion of grammar. Moreover, in the face of PRO gate phenomena, Epstein’s (1999) First Law ceases to be a foundation for syntax. Even if we relax the First Law to one in which syntax operation must be done within a single rooted tree, it does not help. Recall that the movement from the subject to the object position causes a serious violation of the Extension Condition if it occurs after these two trees merge to become single rooted.

Consider a consequence of this. It implies that we cannot replace Move with Attract. Chomsky (1995) construes movement as K, the landing site attracting α , the element(s) (features, categories) that relate to K. For Chomsky (1995) K must dominate α to attract it. And Attract affects only the most proximate element it dominates. In effect K attracts β iff K dominates β and there is no α which c-commands β that K also dominates (cf. Chomsky (1995) p. 297). This always holds for movement in single rooted trees but it never holds in cases of sideward movement. Thus, if sideward movement exists, as argued here, we must reject an Attract based theory of “movement”.

In conclusion, to the extent that what we have seen is true, this paper supports a strictly derivational approach to grammar. If we just look at the result of the operation that we have applied to PRO gate phenomena, it apparently violates the Extension Condition, Chain condition and Sentential Subject Condition. We have seen how adopting sideward movement allows us to preserve the first two while showing the latter is superfluous.

Furthermore, we must dispense both with chains in the traditional sense and c-command constraint on movement, contrary to Brody (1995) and Epstein (1999). The notion of chain, if anything, should be regarded as simply the history of a legitimate series of movements of an element. This view corresponds with Chomsky’s (1998) (i.e., “Basic properties of chains should then follow from elementary derivational principle.”), which is return of the interpretation that Chomsky (1986a) gives to them: chains are a “reflection of ‘history of movement.’”

Finally, rejecting the c-command and single rooted trees as the geographical requirements for movement, makes the notion of Attract untenable and argues for a return to Move as the basis for handling displacement phenomena.

9. Conclusion

To summarize, we have argued that the correct analysis of PRO gate phenomena involves sideward movement and this is incompatible with chains in traditional sense. This means that not all legitimate movement creates a chain, contrary to Brody (1995). Since the concept of chain is necessary for a representational approach to grammar, our analysis supports a derivational approach to Universal Grammar, which also allows sideward movement. At the same time, we have denied that c-command is a primitive notion in syntax, contrary to Epstein (1999), while arguing that his First law is empirically unsatisfactory. Added to these, the present account indicates that Attract-based approaches to movement are inadequate.

Furthermore, we have shown that the subject is not intrinsically an island. Rather, we conclude that the timing in a derivation decides whether a subject is an island for a given operation. The Sentential Subject Constraint proposed by Ross (1967) has been assumed to be descriptively valid, whatever efforts to derive it have been made. But this paper suggests that the movement out of the inside of the subject is indeed possible.

APPENDIX: A note on OCPRO

We have seen in section 6.3 that the nominal embedded in DP may not be the controller of PRO in the subject:

106) PRO_{1/2*} to have to give an important speech tomorrow makes [Bush₂'s campaign]₁ nervous.

However, there arises a problem. What is wrong with the next derivation?

- 107) a. [give an important speech] (building the VP)
 b. Bush [give an important speech] (picking up 'Bush' from the numeration)
 c. [Bush to [give an important speech]] (building up the subject)
 d. [Bush to [give an important speech]] campaign (picking up 'campaign')
 e. [Bush to [give an important speech]] [Bush's campaign] (building up DP with the sideward movement)
 f. [[Bush to give an important speech] [annoy [Bush's campaign]]] (building up the whole clause.)
 g. [[~~Bush~~₁ to give an important speech] [annoy [Bush's campaign]]] (the deletion of copies all but one)

Give the derivation above, we incorrectly predict that the specifier of the object DP can control PRO in the subject. Here, we suggest that the movement to the specifier of DP is not possible, following Williams (1982). He discusses the several pieces of evidence for the impossibility of the operation.

i. Raising is not allowed (following Chomsky (1970))

- 108) a. John₁ appeared t₁ to leave.
 b. *John₁'s appearance t₁ to leave

Raising construction of nominalization is impossible as shown above.²³

ii. The relation between the specifier of DP and the rest of DP can be any relation at all.

- 109) a. Your cat = "the cat you stepped on"
 b. Your destruction of Rome = "your account of the destruction of Rome"
 c. Yesterday's destruction of the city = "the destruction of the city that happened yesterday"²⁴

These examples show that the specifier of DP need not be thematically linked to the head noun. From this, Williams (1982) postulates Det Rule that relates genitive nominals to the rest of DP, assuming that they are base-generated in the specifier of DP.

²³ On the other hand, "John's desire to score a goal" is good. Here I simply assume that PROs in Noun Complements are NOC PRO.

²⁴ I thank M. Kamiya for reminding me of this.

iii. Raising Passive nominal and Pseudo-passive

- 110) a. The city₁'s destruction t₁ by the army
 b. The city₁ was destroyed t₁ by the army.

The passive nominals above seem to show the movement within DP. However, Raising passive and Pseudo-passive nominals do not exist.

iv. Raising passive

- 111) a. *John₁'s belief t₁ to have been there
 b. John is believed to have been there.

v. Pseudo-passive

- 112) a. *The news₁' talk about t₁ by John
 b. The news₁ was talked about t₁ by John.

These types of passive must be involved with the movement because the specifier of DP starts from the subject position of the lower clause in Raising passive and the complement of the preposition in Pseudo-passive. Thus, these facts support his argument that the movement to the specifier of DP is not possible.

Given this, we assume that the specifier of DP can not be a landing site of movement, and the nominal that sits there must be base-generated. Hence the sideward movement to the specifier of DP as shown in (107e) is illicit.

Reference

- Belletti, A. and L. Rizzi (1988) 'Phych verbs-Verbs and Theta Theory.' *NLLT* 6, 291-352
 Brody, M. (1995) *Lexico-Logical Form*. Cambridge, MA: MIT Press
 Brody, M. (1997) 'Perfect Chain', In L. Haegeman (ed), *Elements of Grammar*, 139-167
 Chomsky, N. (1981) *Lectures on government and binding*. Foris: Dordrecht
 Chomsky, N. (1986a) *Knowledge of Language: Its Nature, Origin and Use*. NY: Praeger.
 Chomsky, N. (1986b) *Barriers*. Cambridge, MA: MIT Press.
 Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press
 Chomsky, N. (1998) Minimalist Inquiries: The Framework. ms. MIT.
 Delahunty (1983) 'But sentential subjects do exist.' *Linguistic Analysis* 12.4:379-398.
 Epstein, S. D. (1999) 'Un-principled syntax: The derivation of syntactic relations', In S. D Epstein and N. Hornstein (eds), *Working Minimalism*, 317-345
 Higginbotham, J. (1980) 'Pronoun and bound variable.' *LI* 11: 679-708.
 Hoji, H. (1985) 'Weak crossover and Japanese phrase structure', In T. Imai and M. Saito (eds), *Issue in Japanese Linguistics* , 163-201
 Hornstein, N. (1995) *Logical Form*. Oxford: Blackwell.
 Hornstein, N. (1999) 'Movement and control.' *LI* 30.1: 69-96.
 Hornstein, N. (Forthcoming) *Move! A minimalist theory of construal*. Oxford: Blackwell.

- Johnson, K. (1991) 'Object position.' *NLLT* 9, 577-636.
- Koizumi, M. (1993) 'Object agreement phrases and the Split VP Hypothesis.' In: *MIT Working Papers in Linguistics* 19.
- Koster, J. (1978) 'Why subject sentences don't exist', In S. J. Keyser (ed), *Recent transformational studies in European languages*.
- Lasnik, H. (1998) *Minimalist Analysis*. Oxford: Blackwell.
- Lebeaux, D. (1984-5) 'Locality and anaphoric binding' *The linguistic review* 4, 343-363.
- Nunes, J. (1995) The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland.
- Pesetsky, D. (1995) *Zero syntax*. Cambridge, MA: MIT Press.
- Pires, A. (1999) A minimalist approach to Clausal gerunds. ms. University of Maryland.
- Safir, K. (1984) 'Multiple variable binding.' *LI* 15: 603-638.
- Safir, K. (1996) 'Derivation, representation, and resumption: The domain of Weak crossover.' *LI* 27. 2: 313-339
- Uriagereka, J. (1999) 'Multiple Spell-Out', In S. D Epstein and N. Hornstein (eds), *Working Minimalism*, 251-282
- Williams, E. (1982) 'NP cycle.' *LI* 13. 2: 277-295

Acknowledgements

I thank Norbert Hornstein for his advice and encouragement. And I want to thank the other members of my committee, David Lightfoot and Juan Uriagereka. I also thank people in our department, especially who play soccer on every Sunday morning.

Hirohisa Kiguchi
Dept. of Linguistics
University of Maryland
1401 Marie Mount Hall
College Park, MD 20742
kiguchih@wam.umd.edu