

Copy free movement, Swiping and the ECP*

LING 895 paper

Chizuru Nakao

University of Maryland, College Park

(Last modified: 3/12/2007)

Section 1: Introduction

In the Minimalist Program framework (Chomsky 1993, 1995, and subsequent works), the operation “move α ” in the previous framework is reconsidered as a combination of two simpler operations: Copy and Merge. It is called the Copy Theory of Movement (CTM). This shift has welcome consequences both theoretically and empirically. Theoretically speaking, the CTM enabled us to dispense with the notion of “trace” as a grammatical formative. A grammatical formative such as the classical notion of trace violates the Inclusiveness Condition. The CTM also gives a principled account for why Merge is preferred to Move when both operations are available (Chomsky 1995: 348); Merge is a subpart of Move and is therefore more economical. On the empirical side, the CTM accounts for reconstruction phenomena such as (1).

- (1) John₁ wondered [which picture of himself_{1/2}] Bill₂ liked.
- a. John₁ wondered [[which picture of himself₁] [Bill₂ liked [~~which picture of himself~~]]].
 - b. John₁ wondered [[~~which picture of himself~~] [Bill₂ liked [~~which picture of himself~~₂]]].

The reconstructed interpretation is available when the part of the moved wh-phrase *picture of himself* is interpreted in the lower copy position as in (1)b. The CTM explains reconstruction as not an operation but as the interpretation of unpronounced copies.


Under this view, chains are reinterpreted as sets of multiple copies/occurrences of one lexical item. However, only one member of the set of copies gets to be pronounced (usually the head of the chain). That means that all the other copies have to undergo some sort of deletion operation (Chain reduction) before it reaches PF, presumably for linearization reasons (Nunes 1995, 1999, 2004). Consequently, some copies are deleted at PF, receiving interpretation at LF as in the reconstruction cases.

Note that, at LF, too, you cannot interpret more than one copy simultaneously. In (1), you can either delete *picture of himself* from the upper copy as in (1)b, or the lower copy as in (1)a. It cannot, however, leave the same part of the element in both copies. It would end up having both

* I would like to show my gratitude to my committee members, Norbert Hornstein, Howard Lasnik and Juan Uriagereka for their valuable and encouraging comments and suggestions. Some of the sections are based on my work with Hajime Ono and Masaya Yoshida. I am grateful to them for the fruitful discussion in our collaboration. I also thank Brian Agbayani, Takuya Goro, Jeff Lidz, Jason Merchant, Miki Obata, Akira Omaki, and Alexander Williams for their insightful comments on earlier versions of this paper. I also thank my graduate student colleagues, including Stacey Conroy, Tim Hunter, So-One Hwang, Rebecca McKeown, Phil Monahan, for providing me English judgments.

John and *Bill* as its antecedent simultaneously, which is not the case (Hornstein, Nunes and Grohmann 2005: 245). In sum, chain reduction is needed at both PF and LF. You need to pick one copy to interpret at both levels.

The purpose of this paper is to demonstrate that there is a type of movement that apparently does not leave a copy. There have been proposals in literature that some movement does not leave a trace/copy at all (Lasnik 1998, 1999; Fox 1999; Boeckx 2000 for A-movement¹; Baltin 1987; Tanaka 2005 for Rightward movement). I argue in this paper that PP shift is one such movement (Nakao, Ono, and Yoshida 2006). PP shift moves a PP rightwards as in (2)b.

- (2) a. John talked [_{PP} to Mary] yesterday.
 b. John talked _{PP} yesterday [**PP to Mary**].
- 

The shifted PP is pronounced at the right edge of the clause, and as I will show in a later section, it does not exhibit reconstruction effects or show other indications of the existence of the original copy. It seems that the PP movement did not leave a PF-copy or an LF-copy.

Existence of such movement apparently contradicts the core claim of the CTM. One might think we have to consider Move as a grammatical primitive again, instead of reducing it to Copy plus Merge. Instead I will argue that, a “copy-free” movement is not an instance of Moving without Copying, but an instance of chain reduction before reaching the split point of PF and LF (Spell-Out). The information of there being a copy downstairs does not reach PF or LF. As a result, only the upper copy of PP shift is pronounced at PF (in the same way as normal chain reduction at PF), and there is no choice of which copy to interpret at LF. LF acts as if there were no lower copy generated in the first place (it does not show reconstruction, etc.).

The organization of this paper is as follows. In section 2, I review the basic properties of PP shift and consider the possible landing site of PP shift. Some pieces of data to show that PP shift does not leave a copy will be presented. In section 3, I will present a puzzle on c-command relationships with shifted PPs and other elements: shifted PPs are attached to an adjoined position, but nevertheless can be c-commanded by elements on their left side. I will present a possible account of the puzzle, using a new set of assumptions about c-command. In section 4, I will examine so-called swiping construction. I will show that a complement PP licenses swiping only when it undergoes PP shift, which further supports the claim that PP shift is copy-less. Also, I will present a derivation of swiping which involves PP shift. If PP shift does not leave a copy, the fact that swiping is insensitive to the ECP effects will be accounted for. Section 5 discusses theoretical consequences of my analysis. Section 6 concludes the paper.

¹ Lasnik’s (1998) argues that A-movement does not leave a copy, based on the fact that Scope reconstruction is impossible with A-movement. He argues that the apparent “Quantifier Lowering” effect in (ib) (May 1977) is due to the special property of the indefinite subject.

- (i) a. Every coin is 3% likely to land heads. (every>likely, *likely>every)
 b. Some politician is likely to address John’s constituency. (some>likely, likely>some)

I do not commit myself to any particular analysis of A-movement. See Hornstein (2001, 2003), Boeckx (2001), Nevins and Anand (2003), among others, for alternative views.

Section 2: PP shift as copy-free movement

2.1. Argument against base-generation hypothesis

I will call movement of a PP to the right edge position of a clause, such as (2)b, “PP shift”. In (2)b, the PP *to Mary* is a complement of the verb *talk*, but is pronounced at the right edge of the clause.

One might argue that PP shift is not a movement operation; the PP in (2)b is not base-generated as a complement PP which later undergoes movement, but is originally base-generated at the right edge. If PP shift is not a movement operation in the first place, it would be nonsensical to argue if it leaves a trace or not. Such a hypothesis, however, is implausible given that the PP *to Mary* needs to be interpreted as a complement of the V, and hence, an argument. If the PP were base-generated separately from the V, how it gets theta-marked from the verb remains unclear. I will show below that the shifted PP in (2)b, repeated here as (3)a, is an instance of a complement PP, while other instances of PP shift involve an adjunct PP, such as (3)b.

- (3) a. John talked _{PP} yesterday [_{PP} to Mary].
b. John talked _{PP} yesterday [_{PP} near Mary].

Various syntactic tests show that the *to*-PP in (3)a is an argument while the *near*-PP in (3)b is an adjunct. First, *do-so* substitution (Lakoff and Ross 1976) shows that the *to*-PP must be included in *do-so* in (4)b, while the *near*-PP in (4)d does not need to be.

- (4) a. John talked [_{PP} to Mary]. Bill did so, too.
b. *?John talked [_{PP} to Mary]. Bill did so [_{PP} to Susan].
c. John talked [_{PP} near Mary]. Bill did so, too.
d. John talked [_{PP} near Mary]. Bill did so [_{PP} near Susan].

Second, the *to*-PP cannot be stranded by VP fronting as in (5)b, while the *near*-PP can as in (5)d.

- (5) a. **Talk** [_{PP} to Mary], though John did, ...
b. ***Talk**, though John did [_{PP} to Mary], ...
c. **Talk** [_{PP} near Mary], though John did, ...
d. **Talk**, though John did [_{PP} near Mary], ...

Third, there is an asymmetry in the interpretation among these PPs when they are fronted. The *to*-PP in (6)a forces topicalization interpretations, while the *near*-PP in (6)b does not.

- (6) a. ?**To Mary**, John talked. (The PP is a topic)
b. **Near Mary**, John talked. (The PP is not necessarily a topic)

These data show that the *to*-PP in (3)a has typical properties of a complement, while the *near*-PP in (3)b behaves as a typical adjunct PP.

Given these data, (3) shows that argument PPs as well as adjunct PPs can undergo PP shift. Under the standard theta-theory, an argument PP needs to be base-generated as a sister of a verb

to satisfy its theta-role. I will keep assuming that PP shift is a movement operation that moves a PP from its base position to the right-edge position, at least for argument PP shift cases.

2.2. Landing site of PP shift

The argument above shows that PP shift is a movement operation that moves a PP outside of the VP, to the right side of its original position. The next question is where the landing site of PP shift is. I assume that the landing site of PP shift is an adjoined position, given that the standard X-bar schema for English ((7)) does not allow a right-hand specifier.

(7) [_{XP} [_{XP} [_{XP} Spec [_{X'} X⁰ Complement]] Adjunct] Adjunct]

In the above schematic illustration, adjunction is the only available option to move a complement (or an adjunct) to a right side position.

In (2b), the existence of the adverb *yesterday* marks a right edge of the clause and indicates the existence of PP shift. The PP is presumably in the higher adjoined position than the adverb *yesterday*. Is this VP-adjunction or IP-adjunction?

(2) b. John talked _{PP} yesterday [_{PP} to Mary].

VP-fronting test in (8)a shows that even temporal adverbs such as *yesterday* can be VP-adverbs: they can be included in the fronted VP node. Furthermore, (8)b shows that PP shift can adjoin to VP, undergoing VP fronting together with the adverb.

(8) a. [_{VP} [_{VP} Talk to Mary] yesterday], though John did, ...
b. [_{VP} [_{VP} Talk _{PP}] yesterday] [_{PP} to Mary]], though John did, ...

The *do-so* test also shows that the temporal information can be included in the interpretation of *do so* in the second conjunct, as shown in (9).²

(9) John talked yesterday to Mary, and Bill did so, too. [= Bill talked *yesterday* to Mary, too.]

On the other hand, PP shift is also compatible with a sentential adverb such as *unfortunately* and *probably*.

² In (9), *yesterday* is only optionally included in the VP ellipsis site. In the case of (i), on the other hand, the temporal adverb is excluded from VP ellipsis, which indicates that it has a higher attachment site. Note that, in (i), the elided VP cannot involve PP shift; *do so* replaces *talked ... to Mary*, which is a disjoint element in the antecedent clause.

(i) John talked yesterday to Mary, and Bill did so today. [= Bill talked *to Mary* today]

As long as (i) is acceptable, it indicates that the existence of PP shift does not affect the licensing condition of VP-ellipsis. Under my analysis, PP shift does change LF configuration (See 2.3.). It would be incompatible with analyses where licensing of VP deletion solely relies on LF identity.

- (10) a. John gave a book, unfortunately, [PP to Mary].
 b. John talked, probably, [PP to Mary].

These adverbs modify the entire sentences and describe the modality or the speaker's attitude.³ They usually show up in the sentence initial position ((11)b) or the position immediately after the subject ((11)) but with an intonation break, they can appear in the sentence end (12).⁴

- (11) a. Unfortunately, John gave a book to Mary.
 b. John probably talked to Mary.

- (12) a. John gave a book to Mary, unfortunately.
 b. John talked to Mary, probably.

The *do-so* test in (13) shows that the sentential adverb *probably* cannot be included in the interpretation of *do so*. "Bill did so, too" here is not equivalent to "Bill also probably did so": the meaning of *probably* is not recovered. This suggests that the adverb cannot be included in the VP.

- (13) John gave a book, probably, to Mary, and Bill did so, too.
 a. [= Bill gave a book to Mary, too].
 b. *[= Bill gave a book, probably, to Mary, too].

If an adverb is IP-adjoined, the PP following it (i.e. *to Mary* in (10)) must also be IP-adjoined (or higher). Therefore, I will assume that the landing site of PP shift can also be an IP-adjoined position.

To summarize, I have seen that PP shift can adjoin a PP to a VP (as in (8)b) or an IP (as in (10)). In other cases (such as (2)b), the attachment site is still ambiguous.

2.3. PP shift does not leave a copy

As we have seen in the Introduction, the standard analysis under the CTM is that a moved element leaves a copy in its original position. However, there is several pieces of evidence to show that PP shift does not leave a copy behind. First, PP shift does not show reconstruction phenomena. Second, an (alleged) PP shift trace does not block contraction. Third, idiosyncratically selected PPs cannot undergo PP shift. These arguments suggest that PP shift is a copy-free kind of movement.

³ The adverbs in (10) are ambiguous between sentential modification and constituent (PP) modification. In the constituent modification interpretation, (10)a roughly means, "It was unfortunate that Mary was the one he gave a book to." In this interpretation, it is indicated that it would not have been unfortunate if John had given a book to someone else. On the other hand, there is no such connotation in the sentential modification reading. It just means "it was unfortunate that John gave a book to Mary," and it can be the case that the speaker feels bad about John giving away the book, no matter who the recipient is. We concentrate on this latter interpretation to claim the adverb here adjoins to IP, rather than to the specific constituent it modifies.

⁴ The necessity of the intonational break might indicate that these adverbs are inserted as some kind of interjection. If that is the case, we need to explore in what environments interjection is allowed. If it turns out that interjections can be inserted between a V and its complement, the argument that (10) involves IP-attached PP shift is undermined. I will leave this possibility open here.

- (15) a. I gave money [to the boys who had helped me clean the yard]₁ for themselves₁.
 b. *I gave money _ for themselves₁ [to the boys who had helped me clean the yard]₁.

The complement PP [*to the boys...*] in (15)a binds the adjunct PP [*for themselves*]. This binding relationship cannot be retrieved when the complement PP undergoes PP shift as in (15)b: [*to the boys...*] cannot bind [*for themselves*] anymore. This is naturally predicted under our hypothesis that PP shift does not leave a copy. Because there is no copy to be interpreted in the original position, the binding relation cannot be satisfied in the original position of the complement PP.

Similarly, an anaphor inside the complement PP in (16)a can not bound by the NP inside the adjunct PP, while that in the shifted complement PP in (16)b can. In this case, too, the shifted PP is interpreted in the shifted position rather than its original position. If it were interpreted in the base-generated position, there would be no contrast between the examples in (16).

- (16) a. *Sue spoke [to each other₁'s friends] [about these people₁]. (Pesetsky 1995: 161)
 b. Sue spoke [about these people₁] [to each other₁'s friends] in Bill's house.

Although these data nicely support the copy-free hypothesis of PP shift, they raise a question about the status of the landing site of PP shift and Binding Theory (Chomsky 1981). According to Chomsky's (1981) Binding Condition A, an anaphor must be bound by its antecedent. Given our earlier assumption that shifted PPs are adjoined to VP (or higher), the anaphors in (15)a and (16)a seem to violate Condition A in the configuration (17)a; the complement PP here does not c-command adjunct PP. Similarly, if the complement PP undergoes PP shift and is adjoined outside the adjunct PP, the structure will be (17)b. In this configuration, the complement PP does not seem to be bound by the adjunct PP; the acceptability of (15)b and (16)b is therefore unexpected.

- (17) a. [_{VP1} [_{VP2} Spec [_{X'} X⁰ Complement PP]] Adjunct PP]
 b. [_{VP1} [_{VP2} [_{VP3} Spec [_{X'} X⁰ PP₊]] Adjunct PP] Complement PP]

I will come back to the issue in Section 3.⁸

2.3.3. Contraction

Another piece of evidence to show the absence of a copy in PP shift is that PP shift does not block contraction (Lasnik 1984). Copulas such as *is* (e.g. (18)a) in English can be contracted as in (18)b. It is observed that the contraction is only possible when there is a morphologically realized category on its right (Bresnan 1971; Boeckx 2000; among others), although the orthographic form of *'s* makes it look like the copula is contracted onto the left-side element. The contraction is blocked when a wh-trace on the right is intervening between the copula and the next morphologically realized element. The copula in (18)c cannot be contracted as in (18)d.

⁸ The data in (15) and (16) also raise a question about the level where Binding Conditions apply. (15)b suggests that Binding Condition A is not satisfied at the base-structure, which is counterevidence to the claim that Condition A is an "anywhere" condition. Note that, under the theory where reflexives are residues of movement (Hornstein 2001: See Section 3), this question is irrelevant, at least for reflexives; licensing conditions of reflexives are derived from conditions on movement and there is no "Condition A" that applies to specific levels of representation.

- (18) a. John **is** in the room (now).
 b. John's in the room (now).
 c. I don't know where_i John **is** t_i (now)
 d. *I don't know where_i John's t_i (now).
 e. John **is** t_{PP} NOW [_{PP} in the room].
 f. John's t_{PP} NOW [_{PP} in the room].

The contraction, however, is not blocked by a trace of PP shift. If a PP was moved from the right side of the copula as in (18)e, it is freely contracted as in (18)f. This contrast between wh-traces and PP shift traces is naturally accounted for under the assumption that PP shift does not leave a copy, while wh-movement does.⁹

2.3.4. Idiosyncratic and idiomatic PPs

The final argument to show that PP shift does not leave a copy comes from the types of PP which can be shifted. We have seen in (3) that both argument and adjunct PPs can undergo PP shift. However, there is a small class of PPs which resists PP shift. For example, when the *in*-PP in the phrase *interested in (something)* is shifted, the sentence is degraded as shown in (19) (without an intonational break).

- (19) a. John was (most) interested in linguistics last year.
 b. ??John was (most) interested _{PP} last year [_{PP} in linguistics].

The *in*-PP in (19) is more strongly connected to the subcategorizing element than a normal complement PP such as (3)a in that it is obligatorily selected and cannot be omitted unless the previous context specifies the content of the *in*-phrase ((20)).

- (20) John was interested #(in linguistics) last year.
 (cf. A "John looked absent-minded in the class." B "He seemed interested last semester.")

Also, the choice of the preposition is restricted by the predicate in an idiosyncratic way. The adjective *interested* idiosyncratically selects an *in*-PP, and the use of the preposition here does not reflect the lexical meaning of *in* itself (i.e. 'be contained by', or 'be inside of'). Let us call this type of PP "idiosyncratic PPs".

Even idiosyncratic PPs can undergo wh-movement, as shown in (21) (although some speakers find pied-piping of PP in general marginal).

- (21) [_{PP} In which subject] was John most interested t_{PP} last year?

This asymmetry between PP shift and wh-movement can be accounted for in terms of the

⁹ This paradigm parallels the analyses of *wanna/gonna* contraction (Lightfoot 1976; Boeckx 2000; Hornstein 2001).

- (i) a. *Who do you wanna stay? (cf. Who_i do you want t_i to stay?)
 b. I'm gonna stay. (cf. I_i'm going t_i to stay.)

Boeckx (2000) claims that contraction is acceptable in (ib) because A-movement does not leave a copy (Lasnik 1998, 1999).

existence of a copy left by movement. Suppose that the idiosyncratic subcategorization of the *in*-PP by *interested* needs to be reflected at LF; the copy of the *in*-PP needs to be interpreted in a position subcategorized by the adjective, as opposed to a regular theta-marked PP. Under such a view, the acceptability of (21) indicates the availability of the lower copy, while the degradedness of (19)b suggests the lower copy is unavailable in PP shift.

There is a similar phenomenon in Dutch idiomatic PPs (Koopman 1984: 118). Both argument and adjunct PPs can be shifted to postverbal positions, as shown in (22)b and (23)b.

- (22) a. , dat ik **voor vijf uur** een baan gereserveerd heb
 that I **for five o'clock** a court reserved have
 ‘, that I reserved a court for five o'clock’
 b. , dat ik een baan gereserveerd heb **voor vijf uur**
 that I a court reserved have **for five o'clock**

- (23) a. , dat zij een boek **aan haar zusje** gegeven heft
 that she a book to her sister given has
 ‘, that she gave a book to her sister’
 b. , dat zij een boek gegeven heft **aan haar zusje**
 that she a book given has to her sister

However, if a PP is a part of an idiom, as in (24), it cannot undergo PP shift.

- (24) a. , dat zij eindelijk met dat geld **over de brug** gekomen zijn
 that they finally with that money **across the bridge** came be
 ‘, that they finally came across with the money’
 b. *, dat zij eindelijk met dat geld gekomen zijn **over de brug**
 that they finally with that money came be **across the bridge**

Chomsky (1993: 206-207) argues that idiomatic interpretations such as *take... picture* (‘to photograph’) is obtained only when the LF copy of *picture* is interpreted inside the VP headed by *take*. The phrase *picture of himself* in (25) needs to be interpreted in the lower copy to get the idiomatic interpretation, as illustrated by the unavailability of *John* as the antecedent of the reflexive.

- (25) a. John₁ wondered [which picture of himself_{*1/2}] Bill₂ took *t*.
 (in the idiomatic interpretation)
 b. John wondered [which ~~picture of himself~~] Bill took [which picture of himself].

Given this argument, (24) seems to suggest that Dutch PP shift does not leave a copy¹⁰. It remains to be seen if Dutch PP shift is a parallel operation to what we are considering in English.

In this section, we have seen evidence to show that PP shift is a movement operation that

¹⁰ However, the argument of English idioms under PP shift does not work so straightforwardly. Most PP idioms (e.g. *run into*) do not even allow wh-movement, which leaves a copy under Chomsky’s analysis ((ia)), as well as PP shift ((ib)).

- (i) a. *_[PP Into who(m)] did John ran _{t_{PP}} yesterday?
 b. *John ran _{PP_} yesterday _[PP into Mary].

does not leave a copy, unlike other movement such as wh-movement.

Section 3. A C-command puzzle and the position of adjuncts

3.1. A C-command puzzle

The argument in 2.3.2. on binding reconstruction raises a question as to how an anaphor inside an adjoined PP is bound by the antecedent within the complement PP in examples such as (16).

- (16) a. *Sue spoke [to each other₁'s friends] [about these people₁]. (Pesetsky 1995: 161)
 b. Sue spoke [about these people₁] [to each other₁'s friends] in Bill's house.

How do we account for the above data? Examples such as (26) show that a complement of a PP can bind outside the PP node.

- (26) *Mary seemed [_{PP} to him₁] to like John₁.

(26) induced Condition C violation because *him* binds the R-expression *John* from inside the PP. If the PP node blocked c-command, the sentence would be acceptable, contrary to the fact. With the assumption that a complement of the PP can bind outside, the acceptability of (16)b is problematic. Look at the structure in (17) again.

- (17) a. [_{VP1} [_{VP2} Spec [_{X'} X⁰ Complement PP]] Adjunct PP]
 b. [_{VP1} [_{VP2} [_{VP3} Spec [_{X'} X⁰ PP₁]]] Adjunct PP] Complement PP
-

In the standard 'first node' definition of c-command, the first node that dominates the complement PP in (17)a is X', which does not dominate the adjunct PP. On the other hand, the first node that dominates the adjunct PP (VP1) does dominate the complement PP. Therefore, would be incorrectly predicted that the Adjunct PP c-commands the Complement PP in (16)a and not vice versa; (16)a should be acceptable, contrary to the fact. Similarly, in (17)b, the first node that dominates the adjunct PP is the VP2 segment, which does not dominate the complement PP. On the other hand, the first node that dominates the complement PP is the VP1 segment and it also dominates the adjunct PP. In this definition, we would predict that the complement PP c-commands the adjunct PP in (17)b, and the acceptability of (16)b is unpredicted.

We tackle this problem in this section. I will show that some pieces of evidence seem to show that an adjoined position is 'lower' than the phrase it adjoins to, while other pieces of evidence shows that it is 'higher'. I will review some previous studies that try to account for the contradictory evidence, and point out their problems. Instead I will argue that, employing the notion of adjunction by Hornstein and Nunes (2006), an adjoined position and the phrase it adjoins to have no c-commanding relations. The anaphors in adjoined positions in (16) are not licensed by c-command configuration, contrary to the standard Binding Conditions. The analysis of reflexives as residues of movement (Hornstein 2001) can account for the distribution of anaphors inside adjuncts in terms of sideward movement.

3.1.1. Evidence to show adjoined positions are 'lower'

Pesetsky (1995) points out that the data of anaphors, bound pronouns and NPIs show that an adjunct PP on the right is c-commanded by the complement of a PP on the left. The below examples show that the left-side (complement) PP always c-commands the right-side (adjunct) PP, but not vice versa.¹¹

(27) *Anaphor binding* (cf. (16)a)

- a. Sue talked [to these people₁] [about themselves₁].
- b. *Sue talked [to themselves₁] [about these people₁].

(28) *Bound variable pronoun*

- a. Sue talked [to each employee₁] [about his₁ paycheck].
- b. ??Sue talked [to his₁ friends] [about each employee₁].

(29) *Negative polarity*

- a. Sue talked [to no linguist] [about any conference].
- b. *Sue talked [to any linguist] [about no conference].

When the complement PP undergoes PP shift, however, the c-command relationship seems to be reversed: the shifted PP seems to be asymmetrically c-commanded by the adjunct PP on its left.

(30) *Anaphor binding* (cf. (16)b)

- a. (?)John talked _{PP_} [about the men₁] [_{PP} to themselves₁].
- b. *I talked _{PP_} [about themselves₁] [_{PP} to [John and Bill]₁].

(31) *Bound variable pronoun*

- a. Rosa talked _{PP_} [about each defendant₁] [_{PP} to his₁ lawyer].
- b. ??I talked _{PP_} [about her₁ mother] [_{PP} to every girl₁].

(32) *Negative polarity*

- a. Bill talked _{PP_} [about no conference] [_{PP} to any linguist].
- b. *I talked _{PP_} [about any of the boys] [_{PP} to none of the girls].

The above data seems to show that an adjoined element (adjunct PPs and shifted complement PPs) is lower than the element it adjoins to. The outer on the right it adjoins to, the lower the element is. As a result, an adjunct/shifted PP is always c-commanded by a PP on its left side.

3.1.2. Evidence to show adjoined positions are 'higher'

On the other hand, other data suggest the opposite. The farther right the adjunction is, the higher the adjunct is. Especially, I observe that when PP shift is long-distance, the shifted PP cannot be bound by an element on its left. (33)-(35) show that the anaphor etc. in the complement PP past the temporal adverb *yesterday* cannot be bound by an adjunct PP anymore, unlike the

¹¹ The examples (27)-(32) are slight modification from Pesetsky (1995: 161-162, 266-267). Some speakers find the anaphor binding examples with PP shift ((30)a) less acceptable than the other examples.

cases above. (Recall that PP shift past a temporal adverb is possible, as seen in (2)b.)

(33) *Anaphor binding*

??John talked_{PP_} [about the men₁] yesterday [PP to themselves₁].

(34) *Bound variable pronoun*

??Rosa talked_{PP_} [about each defendant₁] yesterday [PP to his₁ lawyer].

(35) *Negative polarity*

??Bill talked_{PP_} [about no conference] yesterday [PP to any linguist].

If adjunction is always to a lower position, the contrast between the (a) examples of (30)-(32) on one hand, and (33)-(35) on the other is hard to capture.

The example in (14), which we used to show the lack of reconstruction points to the same direction.¹²

- (14) a. [IP Someone [VP gave every book₁ [PP to its₁ prize winning author]] yesterday].
(some>every, every>some)
b. [IP Someone [VP gave every book₁ **t_{PP}**] yesterday] [**PP to its₁ prize winning author**].
(^{??}some>every, every>some)

The binder of the bound variable (i.e. the quantifier *every*) needs to be in a higher position than the subject (*someone*) only in (14)b. This indicates that the PP in (14)b is located higher than that in (14)a. This shows that a PP that undergoes PP shift past a temporary adverb moves higher. This does not follow if a landing site of adjunction is necessarily lower than the element it adjoins to.

¹² The configuration in (14)b apparently induces Weak Crossover violation (WCO), because the trace of the QP does not c-command the variable *its* in the shifted PP. (i) exemplifies classic cases of WCO violation.

- (i) a. *Who₁ does [his₁ mother] love t₁?
b. *His₁ mother loves everyone₁.

The fact that derivation in b is not excluded as WCO violation seems to suggest that WCO should be defined in terms of ‘leftness’ rather than ‘c-command.’

- (ii) *Crossover* (Chomsky 1976: 201)
A variable cannot be co-indexed with a pronoun to its left.

If this descriptive generalization is on the right track, (14)b is not problematic. The bound variable is not c-commanded by the original position of *every*, but still sits on its right side. (See Epstein 1987, 2000 for a related discussion.)

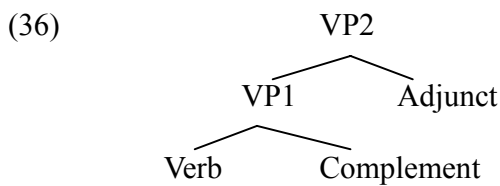
Of course, syntactic conditions solely relying on linear order are not a desired move. Uriagereka (1999) suggests that WCO effect should be derived from conditions on novelty/familiarity (See Ferro 1994 for a similar analysis). In (ib), for example, the familiar pronoun *his* cannot be the antecedent of the novel *everyone*. The same should hold for the pronoun *his* and the trace of *who* in (ia), which means that the *wh*-phrase is not considered ‘familiar’ in the discourse until its trace comes along. Under his definition of WCO, too, (14)b is not problematic; the bound pronoun is not ‘familiar’ at the point where *every* is interpreted.

3.1.3. Evidence to show adjoined positions are ‘outer’

The above subsections illustrate the puzzle about the status of an adjoined position. Moreover, you need to take into consideration which elements make a constituent when you consider hierarchical structure of adjunction. Recall that, in 2.1., we used VP-fronting and the *do-so* test as constituency tests to distinguish complement PPs and adjunct PPs, as illustrated in (4) and (5).

- (4) b. *?John talked [PP to Mary]. Bill did so [PP to Susan].
d. John talked [PP near Mary]. Bill did so [PP near Susan].
- (5) b. *Talk, though John did _ [PP to Mary], ...
d. Talk, though John did _ [PP near Mary], ...

The contrast shows that complement PPs are obligatorily inside the target of these operations, while adjunct PPs can be excluded from them. This asymmetry is standardly considered to be evidence that a complement is a sister of the V head and included in the innermost VP segment, while adjunct is excluded from the innermost VP segment, as in (36). Here, if the operation targets VP1, the adjunct can remain excluded from the operation. On the other hand, the complement is obligatorily inside the VP, whether the operation targets VP1 or VP2.



If you abandon the configuration in (36) and claim that the adjunct is somewhere below the complement, you need to explain why adjuncts can freely be outside of the operations that target a VP.

3.2. Previous analyses

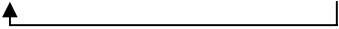

I will review three analyses on the landing site of PP shift and configuration of adjunction: Kayne (1994), Pesetsky (1995) and Phillips (1996). They all succeed in explaining how adjoined positions can be lower, and hence are c-commanded by the element on the left (3.1.1.), and the latter two also account for the constituency facts in 3.1.3. However, they do not consider the examples where a shifted position is higher (3.1.2.). Given this empirical problem and some theoretical considerations, I will claim that none of the analyses is fully satisfactory.

3.2.1. Kayne (1994)

Kayne (1994) claims that no rightward movement is allowed, using his Linear Correspondence Axiom (LCA) system. The LCA assumes a complete mapping between asymmetrical c-command and left-to-right precedence relations. As a consequence, no element can go upward and rightward at the same time. In his system, extraposition (e.g. (37)b) involves a leftward movement of the head NP as in (38)a, instead of the rightward movement from inside an

NP, which is illustrated in (38)b.

- (37) a. A picture of Mary was sent to her. (Kayne 1994: 122)
 b. A picture was sent to her of Mary.

- (38) a. [[A picture]₁ was sent to her [[e]₁ of Mary]].

 b. [[A picture [e]₁] was sent to her] [of Mary]₁.


In such a system, a shifted PP (which is the rightmost element in the clause) would have to be asymmetrically c-commanded by the rest of the clause, in all situations.

Details aside, if ‘leftness’ and ‘c-command’ completely go together, the data to show that the landing site of PP shift is higher than its original position (3.1.2.) would not be explainable. I am not taking the strict LCA analysis here. For the extraposition case, too, there is evidence that an extraposed relative clause is higher than the original position. Fox and Nissenbaum (1999) provide the following data.

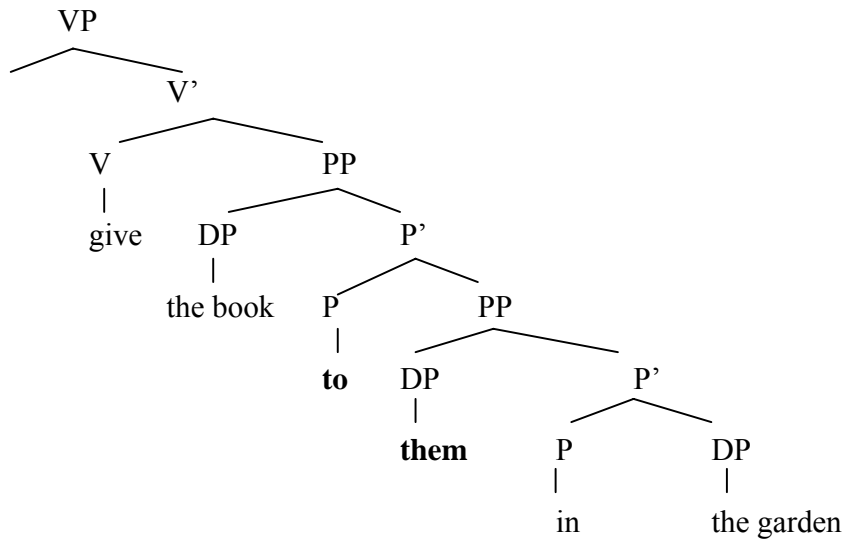
- (39) a. I looked very intensely for anything that would help me with my thesis.
 b. *I looked for anything very intensely that would help me with my thesis.

Free choice *any* must be inside a modal operator such as *look for*. In (39)a, *anything* is interpreted inside the scope of *look for* and the sentence is acceptable. When the relative clause undergoes extraposition as in (39)b, on the other hand, the free choice *any* is not licensed under the scope of the verb. Fox and Nissenbaum explain the contrast by arguing that the relative clause that attaches to a nominal decides its scope. The scope of *anything* needs to be as high as the attachment site of the extraposed clause in (39)b. As a result, *anything* cannot have a narrower scope than the verb and is not licensed in the extraposition example. This argument shows that the attachment position of extraposed material is higher than its original position. I point out that this is a counterexample to Kayne’s analysis of extraposition in (38), although I will not go into the analysis of extraposition in this paper in any more detail.

3.2.2. Pesetsky (1995) on cascade structures

To explain the c-command facts in 3.1.1., Pesetsky (1995) proposes what he calls ‘cascade structure’ for adjunction. Under cascade structure, adjoined materials are embedded deeper than the original structure. (40) illustrates the configuration of a complement PP and an adjunct PP before PP shift.

(40) ... give the book [to them] [in the garden]

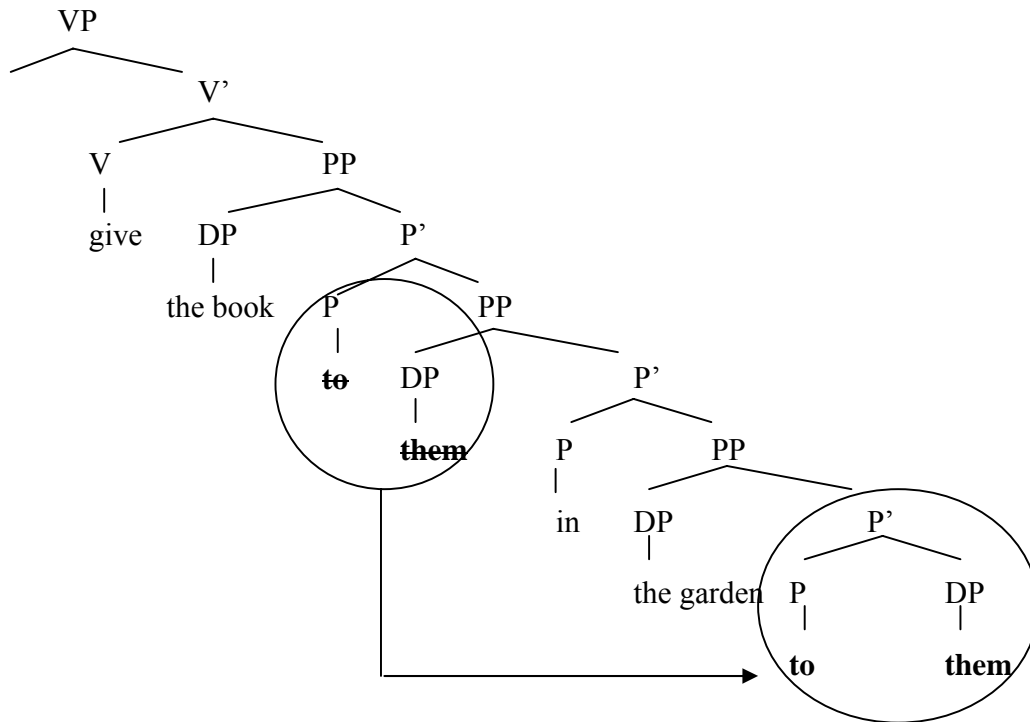


In this structure, the adjunct PP (*in the garden*) is embedded inside the complement PP (*to them*). The fact that the adjunct PP is asymmetrically c-commanded by the complement (See (27), repeated from 3.1.1.) naturally follows from the cascade structure.

- (27) a. (?)Sue talked [to these people₁] [about themselves₁].
 b. *Sue talked [to themselves₁] [about these people₁].

When the argument PP undergoes PP shift, it is further embedded deeper than the adjunct PP, as illustrated in (41).

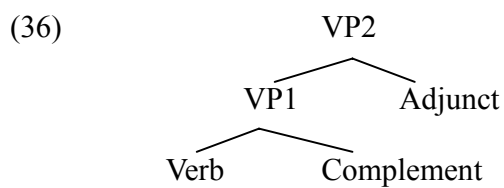
(41) ... give the book PP_ [in the garden] [PP to them]



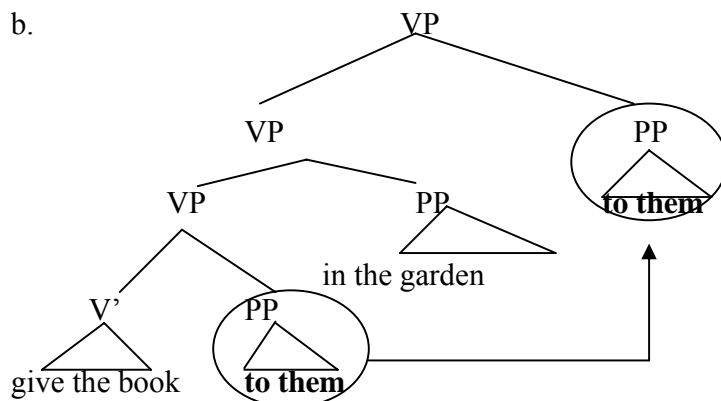
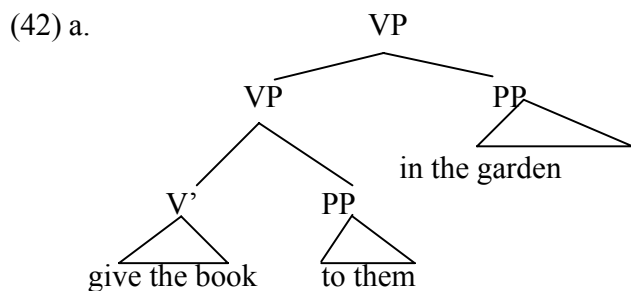
Hence, the fact that a complement PP that is shifted to the right side of an adjunct PP is asymmetrically c-commanded by the adjunct PP (See (30)) follows.

- (30) a. (?)John talked PP_ [about the men_i] [PP to themselves_i].
 b. *I talked PP_ [about themselves_i] [PP to [John and Bill]_i].

He also assumes that the standard 'layered structure' such as (36) is also constructed in the grammar. This explains the fact that adjuncts can be outside of the target VP in VP constituency tests (3.1.3).



Under the layered structure, (40) and (41) are represented as (42)a and (42)b, respectively.



In this structure, the adjoined PPs are layered higher and higher to the right edge. The same operation which is a lowering operation in the cascade structure ((41)) is a raising operation in the layered structure ((42)b).

In Pesetsky's system, the layered structure and the cascade structure are both constructed for every sentence in a parallel way. The layered structure is used for determining the constituency (such as *do-so* test and VP-fronting we discussed in 3.1.3.), while the cascade structure is used to determine the c-command relationships. In sum, the c-command facts such as anaphor licensing, bound variable pronoun licensing and NPI licensing in (27)-(32) (3.1.1.) are derived from the cascade structure, where the adjoined PP is lower than the existing elements.

Although Pesetsky's analysis correctly captures the conflict between the constituency tests and c-command facts, there are several problems in this analysis. First, similarly to Kayne's (1994) analysis, it cannot explain the fact that the 'long' PP shift that crosses a temporal adverb is to a higher adjoined position ((33), 3.1.1.).

(33) ??John talked _{PP_} [about the men₁] yesterday [_{PP} to themselves₁].

Although the position for the adverb *yesterday* is unclear in his analysis, the shifted PP always buries inside the existing elements. If this is the case, the PP *to themselves* should still be lower than *about the men*, and we would predict that the anaphor should be licensed, contrary to the fact.

Secondly, PP shift in his view must involve a non-constituent movement as illustrated in (41). Pesetsky calls elements that are constituents in layered structure but not in cascade structure 'L-constituents,' while elements that are constituents in both structures are called 'LC-constituents.' For example, the PP *to them* in (41) and (42)b is an L-constituent; it is a constituent in the layered structure (42)b, but not in the cascade structure (41). He assumes that L-constituents as well as LC-constituents undergo movement, but it is unclear why movement of

L-constituents is not problematic in the cascade structure. On the cascade structure side, such movement should count as operation on non-constituents.

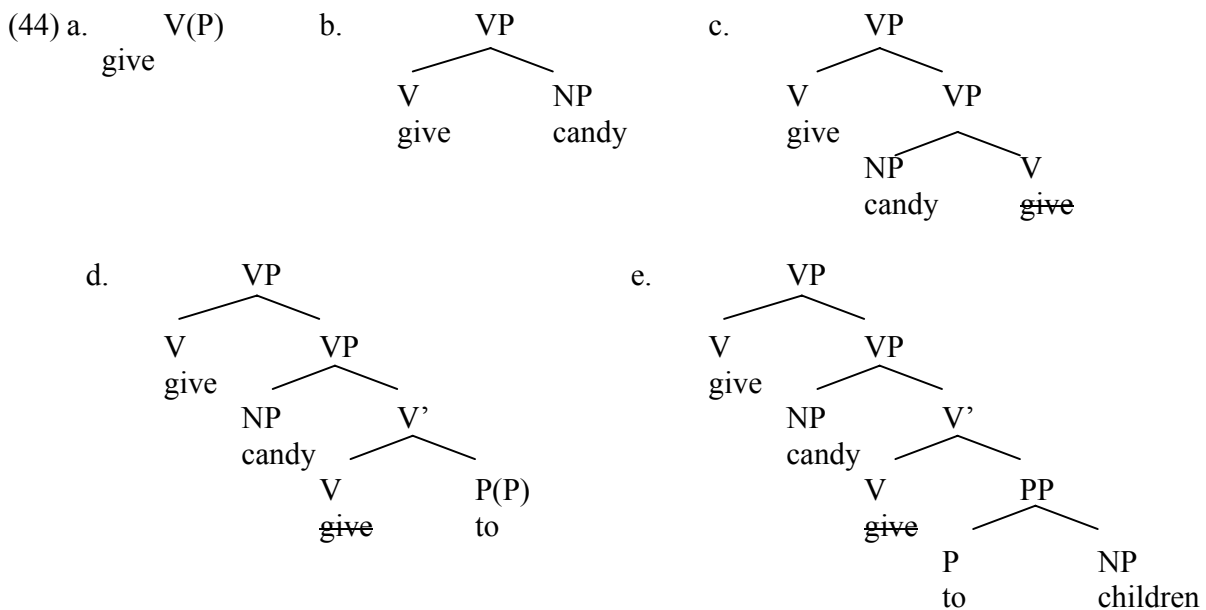
Moreover, positing two totally different hierarchies for one sentence can be a burden for computation. If we can get by with a single structure per sentence, it would be theoretically more desirable.

3.2.3. Phillips (1996) on Merge right

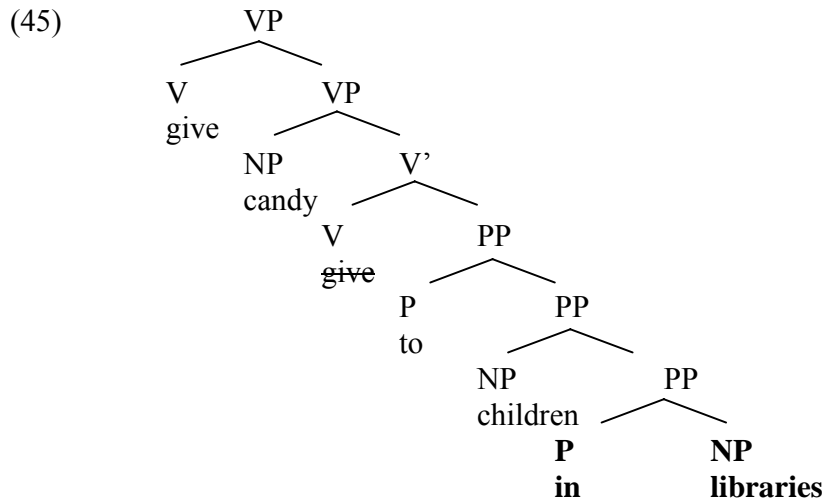
Phillips (1996) accounts for the same paradigm explained in Pesetsky (1995) without positing two parallel structures. He proposes that derivation proceeds incrementally from left to right, as opposed to the bottom-up tree building advocated by Chomsky (1995) and many others. He proposes two conditions on tree building, Merge Right and Branch Right.

- (43) a. *Merge Right*: New items must be introduced at the right edge of a structure.
 b. *Branch Right*: Where a terminal can be attached to more than one position in the existing structure with no effect on interpretation, the attachment that results in the more right-branching structure must be chosen.

Under this system, the derivation of the VP *give candy to the children* proceeds in the following way.



The derivation starts as a single node *give* ((44)a). The next element, that is, the complement NP *candy* merges to its right ((44)b). A null copy of the verb *give* is added to the further right to make a VP-shell structure of the oblique dative construction (Larson 1988) ((44)c). When a PP is added, it is not merged as the whole PP. First, the P node is merged right ((44)d), followed by its complement ((44)e). When another adjunct PP is added in the derivation, it is merged further to the right, as illustrated in (45).

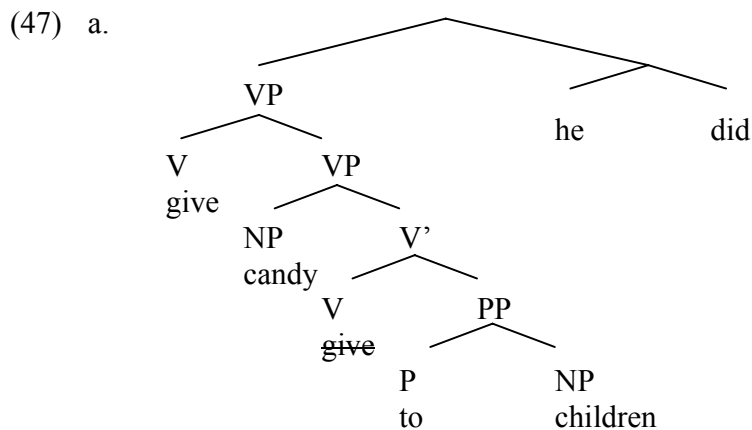


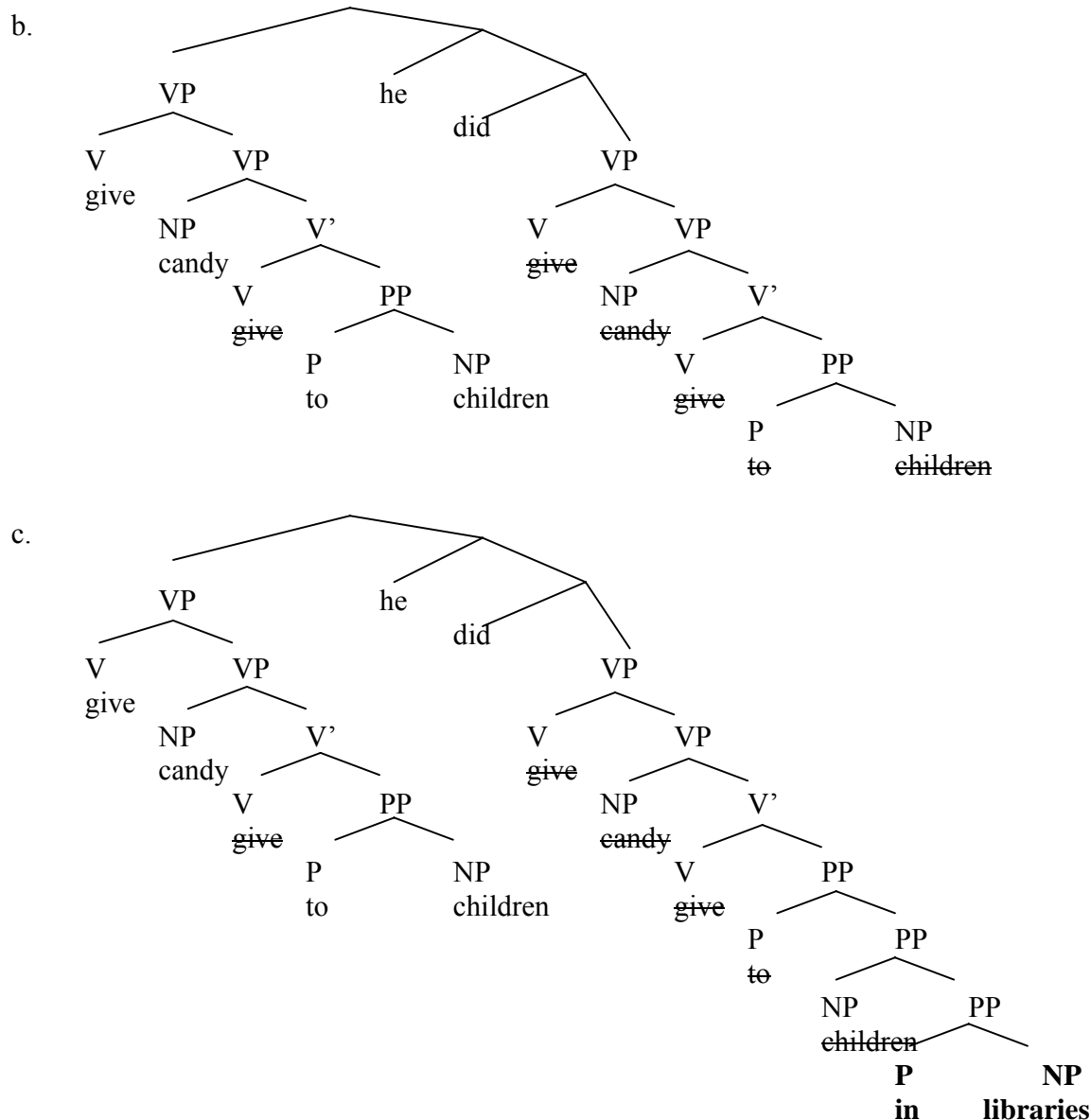
This structure, just like Pesetsky's (1994) cascade structure, can explain why an anaphor in the right side PP can be licensed by the antecedent in the left side PP. Note that, the former PP *to children* is not a constituent in the outcome representation (45), although it used to be a constituent in (44)e. In this system, constituency changes as the derivation proceeds in this way.

In the structure in (45), the string *give candy to children*, which is a part of the VP that excludes the adjunct PP *in libraries*, does not make a constituent. How is (46) explained, where VP-fronting targets this string, excluding the adjunct PP (See also 3.1.3.)?

(46) Give candy to children he did in libraries.

Phillips (1996) explains this example using the incremental system in the following way. The fronted VP *give candy to children* is constructed in the same way as (44)e. This structure internally conforms to Merge Right. Then, the VP is merged to the main clause *he did*.





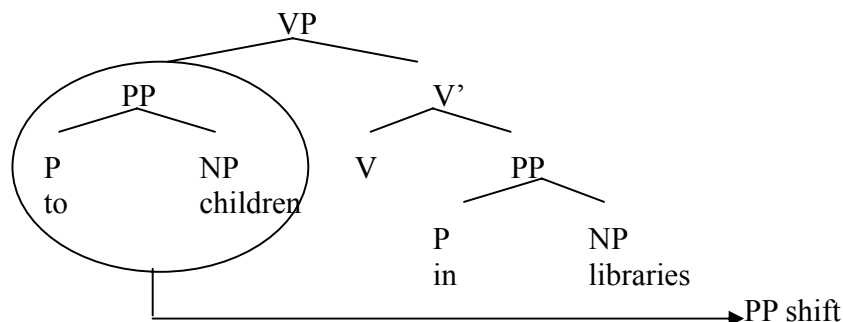
Then, the elements in the VP *give candy to children* are inserted as null copies after *he* auxiliary *did* ((47)b). Finally, the adjunct PP is merged right to the null copy structure ((47)c). Although *give candy to children* is not a constituent in the second VP (which is constructed after *he did*), it is a constituent in the fronted position. In this way, the part of the VP which is not a constituent in the full VP can be a target of a VP fronting operation.

In Phillips' system, Branch Right (43)b is a preference principle, and not an absolute principle. Left-branching structure is allowed when Merge Right is not appropriate. For example, PP shift to the right is impossible from a strictly right-branching structure. (45), for example, is incompatible with the rightward movement of PP shift of *to children*, because this complement PP does not make a constituent. The PP shift example (48)a, hence, needs to be derived from another source (48)b, which involves left-branching.

(48) a. I give candy PP_{-} in libraries [PP to children].



b.



In (48)b, the PP *to children* is a constituent, so it can undergo rightward PP shift.

Note that, whether it is upward or downward does not affect his discussion. If you add an assumption that the ‘short’ PP shift in (30)a is a lowering operation while the ‘long’ PP shift in (33) is a raising operation, the contrast between the two examples are accounted for. The shifted PP is lower than the adjunct PP in (30)a., but not in (33).

(30) a. John talked PP_{-} [about the men₁] [PP to themselves₁].

(33) ??John talked PP_{-} [about the men₁] yesterday [PP to themselves₁].

Although the analysis has wide empirical coverage, there is some potential theoretical concern about optionally allowing left-branching structure. First, the assumption is needed that the theta-role of a PP complement can be discharged in either the right-branching configuration (44)e or the left-branching configuration in (48)b. This means that the configurational relation between the theta-role assigner (the verb) and the PP complement is not uniform. To posit two different subcategorization frames for the exact identical verb *give*, which takes a *to-PP*, leads to massive redundancy. Second, the mechanism is unclear that chooses the left-branching configuration (48)b over the right-branching configuration only when the PP is to undergo rightward movement in a later stage. You could argue that rightward PP shift is induced by a feature, and only a PP with that feature allows left-branching structures, but such a requirement is just a statement of the observation. As I do not find good answers to these questions, I would like to propose an alternative analysis in the next section.

3.3. An alternative analysis

In this section, I review the theory of adjunction proposed by Hornstein and Nunes (2006). They argue that adjuncts can be concatenated with a phrase without labeling. Using their notion of adjunction, I will propose a new account of the c-command puzzle. I will claim that there is no c-commanding relation between an adjunct and the phrase it adjoins to. Assuming that c-command relationship comes free with Merge (Epstein 1999), and that Merge is concatenation plus labeling (Hornstein and Nunes 2006), it follows that c-command does not necessarily follow from adjunction, which is ‘concatenation without labeling.’ If there is no c-commanding relationship of adjuncts, an alternative way to account for the binding facts without recourse to c-command is necessary, at least for examples of binding inside adjuncts. I introduce Hornstein’s

- (52) John could eat the cake [in the yard] [with a fork] and...
- [eat the cake] he did in the yard with a fork.
 - [eat the cake with a fork] he did in the yard.
 - [eat the cake in the yard] he did with a fork.
 - [eat the cake in the yard with a fork] he did.

In a standard ‘layered’ structure such as (53), the outer adjunct (*with a fork*) is necessarily excluded from the operation whenever the inner adjunct (*in the yard*) is excluded (under the assumption that the constituency structures of the preceding sentence (52) and the following sentences (52)a-d are parallel).

- (53) [_{VP1} [_{VP2} [_{VP3} eat the cake] in the yard] with a fork]

In such a representation, when the VP fronting targets the innermost VP3, (52)a arises. When it targets VP2, (52)c is generated. When it targets VP1, both adjuncts are included in the fronted VP as in (52)d. However, the VP fronting in (52)b, which only targets the outer adjunct, should be impossible.

Note that, the data such as (52)b is not observed in previous literature we reviewed in 3.2. It is problematic for Pesetsky (1995) because he appeals to the layered structure in (53), when it comes to the VP-fronting test. It is also problematic for Phillips’s (1996) analysis, as *eat the cake with a fork*, does not make a constituent in any stage of the derivation in the preceding sentence, but it needs to make a constituent in the fronted material.

- (54) John could [_{VP} eat [_{NP} the cake [_{PP} in [_{NP} the yard [_{PP} with [_{NP} a fork]]]]], and
 [_{VP} eat [_{NP} the cake [_{PP} with [_{NP} a fork]]]] he did
 [_{VP} eat [_{NP} ~~the cake~~ [_{PP} ~~with~~ [_{NP} ~~a fork~~ [_{PP} in [_{NP} the yard]]]]]]].

In this way, the hanging adjunct analysis explains the behavior of adjoined elements.

3.3.2. *Adjunct attachment without c-command*

Adopting the hanging adjunct analysis, I will attempt to explain the facts about the PP shift and c-commanding relations. The core of my argument is that an adjoined position is not in a c-command relationship with a phrase it adjoins to. According to Hornstein and Nunes (2006), adjuncts are in an adjoined position, which is a hanging position that does not have a labeled node. If there is no labeled node that dominates the adjoined PP, the adjoined position is not hierarchically higher than the inside element of the VP. As a result, the adjoined PP cannot c-command inside the adjoined phrase. In other words, using the “first-node” definition of c-command in (55), there is no c-command domain for the adjoined PP because there is no “node” that immediately dominates the adjunct.

- (55) A c-commands B if the node that immediately dominates A also dominates B.

Therefore, the fact that an adjunct cannot c-command the complement PP ((27)b) is naturally derived.

3.3.3. Epstein (1999) on derivational c-command

My claim that adjuncts are not involved in hierarchical relationships with the phrase they adjoin to is based on the fact that no labeled node immediately dominates an adjunct; the ‘first node’ definition of c-command does not work for adjuncts. Before moving on, however, let me note that this claim is not crucially dependent on this particular interpretation of c-command. It is also compatible with the ‘derivational’ type of definition of c-command, advocated by Epstein (1999).

Epstein (1999) points out that treating the ‘representational’ definition of c-command such as (59) as a primitive grammatical definition leaves a number of questions open.

(59) *The representational definition of c-command* (Reinhart 1979)

A c-commands B iff

- a. The first branching node dominating A dominated B, *and*
- b. A does not dominate B, *and*
- c. $A \neq B$.

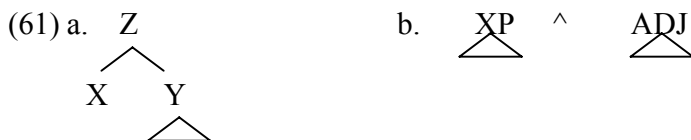
Under this definition, it remains unclear why this particular relationship that involves “the first branching node,” as opposed to “the second branching node,” or “the first non-branching node.”¹⁴ It lacks the explanation of why such relationship needs to exist, which is undesirable in the Minimalist point of view.

Instead, he proposes that c-command inherently comes along with the operation Merge, which is considered to be an indispensable primitive operation to create syntactic structures. As long as Merge is a fundamental syntactic operation, c-command is a fundamental relationship. He proposes the ‘derivational’ definition of c-command in (60).

(60) *Derivational c-command* (Epstein 1999: 329)

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.

If the CTM is on the right track, Move is an instance of Merge which includes a copying operation, so you can attribute all the instance of c-command to Merge. Under this definition, X *derivationally* c-commands Y and all the elements inside Y in (61)a, by virtue of the Merging operation of X and Y.



What about the adjunction in (61)b? We defined Merge as concatenation plus labeling, and adjunction as concatenation without labeling. Then, XP and ADJ in this example are not paired by Merge or Move; instead, they are merely concatenated. If you take this difference crucially, the lack of the ‘dominating node’ in (61)b can be restated as lack of Merge under derivational

¹⁴ See Kupin (1978) for the arbitrariness of the definition of c-command configurations under the traditional phrase markers, as opposed to his reduced phrase markers (RPM).

The derivation of (67) starts from the numeration in (68).

(68) N = {which₁, paper₁, Q₁, you₁, did₁, v₂, file₁, without₁, C₁, PRO₁, T₁, reading₁}

Both the adjunct subtree and the main clause subtree are constructed from the items in N. First, the adjunct subtree K is constructed as in (69)a. Here, the NP *which paper* is base-generated as an object of the verb *reading*. When you construct the main clause L in (69)b, *which paper* is moved from K to L. It is merged with another verb *file* and discharges the object theta-role of *file*. This analysis makes crucial use of Hornstein's assumption that DPs can have more than one theta role and movement into a theta position is possible. Note also that this movement does not violate CED, because K is not adjoined to L and is not an adjunct island yet.

(69) a. K = [CP C [TP PRO₁ [T' T [VP t₁ [v' v [VP reading [which paper]₂]]]]]]
 b. L = [VP file [which paper]₂]

As the derivation proceeds, the adjunct clause becomes (70)a, while the main clause becomes (70)b. It is at this point that the adjunct subtree can be adjoined to the main clause subtree.

(70) a. P = [PP without [CP C [TP PRO₁ [T' T [VP t₁ [v' v [VP reading [which paper]₂]]]]]]]
 b. Q = [VP you [v' v [VP file [which paper]₂]]]

(71)a is the structure after the adjunction of (70)a to (70)b. (It is a traditional layered adjunction structure, because Nunes (1995, 1999, 2004) has not adopted the hanging adjunct analysis of Hornstein and Nunes (2006) reviewed in 3.3.1.) After the complementizer Q is introduced to the structure as in (71)b, *which paper* moves to [Spec, CP] to satisfy the strong feature of Q, forming the final representation in (71)c.

(71) a. [VP [VP you [v' v [VP file [which paper]₂]]] [PP without [CP C [TP PRO₁ [T' T [VP t₁ [v' [VP reading [which paper]₂]]]]]]]]]
 b. [C' did+Q [TP you₃ [VP [VP t₃ [v' v [VP file [which paper]₂]]]]] [PP without [CP C [TP PRO₁ [T' T [VP t₁ [v' [VP reading [which paper]₂]]]]]]]]]
 c. [CP **[which paper]₂** [C' did+Q [TP you₃ [VP [VP t₃ [v' v [VP file ~~[which paper]₂]]]]]]]
 [PP without [CP C [TP PRO₁ [T' T [VP t₁ [v' [VP reading ~~[which paper]₂]]]]]]]]]~~~~

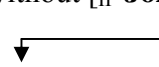
Similarly, the derivation of adjunct control in Hornstein's (2001) system involves sideward movement of the control subject. Let us take a look at the derivation of the adjunct control sentence (72).

(72) John saw Mary without PRO entering the room.

The adjunct clause (73)a and the main clause (73)b are constructed as subtrees.

(73) a. [PP without [IP John₁ [I⁰ [VP John₁ [entering the room]]]]]
 b. [v' saw Mary]

After the Merge of the verb *saw* in (73)b, *John* in the adjunct clause undergoes sideward movement into the subject position of the main clause to satisfy the subject theta role of *saw*, as illustrated in (74).

- (74) a. [PP without [IP **John**₁ [I⁰ [VP John₁ [entering the room]]]]]


 b. [VP **John**₁ [saw Mary]]

After the VP in the main clause is complete, the adjunct tree is adjoined to it, forming (75)a. The derivation proceeds to make the complete sentence (75)b.

- (75) a. [VP [VP John₁ [saw Mary]] [PP without [IP John₁ [I⁰ [VP John₁ [entering the room]]]]]]]
 b. [IP John [I^{past} [VP [VP John₁ [saw Mary]] [PP without [IP John₁ [I⁰ [VP John₁ [entering the room]]]]]]]]]

In this subsection, we have reviewed that sideward movement from within an adjunct is possible in Nunes (1995, 1999, 2004) and Hornstein's (2001) system.

3.3.6. Analysis: binding without c-command

I pointed out in 3.1. that there are three kinds of data we need to account for about adjunction and PP shift. First, a reflexive in an adjoined position is licensed by the complement position ((27)a), while the opposite is the case when the complement PP is on the right side of the adjunct PP as in (30)a.

- (27) a. Sue talked [to these people₁] [about themselves₁].

- (30) a. John talked PP_ [about the men₁] [PP to themselves₁].

Second, a reflexive is degraded in the complement PP when the PP shift further passes a temporal adverb, as in (33). I call it 'long' PP shift.

- (33) ??John talked PP_ [about the men₁] yesterday [PP to themselves₁].

Third, an adjunct can optionally be outside of VP-fronting or *do-so* substitution, as the following examples show.

- (4) d. John talked [PP near Mary]. Bill did so [PP near Susan].

- (5) d. Talk, though John did _ [PP near Mary], ...

The ambition of this subsection is to account for these apparently conflicting data, without recourse to c-command. Employing the assumption that an adjoined position does not c-command or is not c-commanded by the phrase it adjoins to, I try to account for the distribution of reflexives by sideward movement.

3.3.6.1. Data where adjoined positions seem to be ‘lower’

Recall that Hornstein (2001) analyzes reflexives as movement residues. In his system, the distribution of reflexives is not defined in terms of traditional Binding Conditions anymore (i.e. the definition of Condition A in (58).) Instead, distribution of anaphors conforms to the restriction of movement. It predicts that binding by a non-c-commanding antecedent is possible as long as the movement to a non-c-commanding position conforms to general constraints on movement (e.g. Extension Condition). Hornstein (2001: 181) gives examples of bound pronouns with a non-c-commanding antecedent.

- (76) a. I will drink no wine₁ before it₁ is ready.
 b. Nobody₁'s mother loves him₁.
 c. Someone from every city₁ hates it₁.

Bound pronouns have been believed to require a c-commanding antecedent, just like reflexives. The above examples, however, are counterexamples to this assumption. Hornstein claims that these examples are adequately explained in terms of movement: the bound pronouns are residues of movement of their antecedents.

Under the assumption that non-c-commanding antecedents are allowed, the fact that reflexives are licensed within an adjunct PP (e.g. (56)) is not incompatible with my claim that adjunct PP is not c-commanded by its antecedent.

- (56) [VP/V' talked [PP to these people₁]] ^ [PP about themselves₁]

Even if there is no c-commanding relationship between the VP and the PP, the anaphor inside the adjunct PP is licensed via sideward movement, as illustrated below. First, the adjunct PP is constructed as in (77)a. The reflexive is base-generated as *the people* plus *self*, followed by the movement of *the people* into the complement PP. The reflexive *the people+self* has a theta role of the complement of *about*, while the people further discharges a theta role of the complement of *to*.


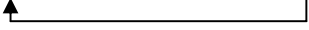

- (77) a. [PP about **the people₁+self**]
 ↓
 b. [PP to **the people₁**]

The adjunct PP can be adjoined to the main tree after the VP is complete as in (78). Thus, the adjunct PP is interpreted as a VP-adjunct. The adjoined structure is (79), under the hanging adjunct notation.

- (78) a. [PP about **the people₁+self**]
 b. [VP talked [PP to **the people₁**]]

- (79) [VP talked [PP to the people₁]] ^ [PP about ~~the people₁~~ **themselves**]

The original position of the movement of *the people* is unpronounced because it is the tail of the chain. As a result, *self* is pronounced as *themselves* because of its morphophonological requirements. This is how a reflexive is generated inside an adjunct PP. It does not necessarily

- (81) a. John₁ heard Mary₂ [without/while/before/after [PRO_{1/*2} entering the room]].

 b. Which book did you read t before Fred reviewed pg?

 c. John talked PP_ [about the men₁] [PP to themselves₁].


This distribution of sideward arrows can be explained if there is a constraint that the starting subtree of the sideward movement is pronounced later. I do not know how to motivate such a constraint at this point, but if there is a general direction of sideward movement from the right-side adjunct clause to the left-side main clause, as illustrated in (81)a/b, it can be extended to the explanation of (81)c, too.

3.3.6.2. Data where adjoined positions seem to be ‘higher’

The sideward movement account can distinguish the data of ‘long’ PP shift. We have seen in 3.1.2. that when PP shift crosses a temporal adverb as well as an adjunct PP, the shifted PP seems to be higher than the ‘short’ PP cases, as the contrast between (30)a and (33) shows.

(30) a. John talked PP_ [about the men₁] [PP to themselves₁].

(33) ??John talked PP_ [about the men₁] yesterday [PP to themselves₁].

Under the sideward movement account, the degradedness of (33) can be explained in terms of the strict cycle condition. The crucial assumption is that temporal adverbs such as *yesterday* need to attach to (at least) the vP level, where the subject is base-generated.¹⁶ This is not an unreasonable claim, given that the temporal information modifies the whole event of ‘someone doing something’, rather than the action of ‘doing something,’ irrelevant of the agent. Given this assumption, the derivation of (33) is shown in (82). The reflexive is born as *the men+self* within the complement PP of *talked*, in the same way as (80) ((82)a). The temporal adverb *yesterday* is a vP-adverb, and thus is generated after vP is built as in (82)b. Note that the shifted PP in (33) is pronounced after *yesterday*. This indicates that the shifted PP also needs to be a vP-adjunct rather than a lower adjunct. Therefore, the PP shift of *to the men+self* occurs after the vP is built ((82)c). Note that, trying to create the VP-adjunct *about the men* by sideward movement here is against strict cycle; at this point, the vP and vP-adjuncts are already created. This is the cause of the degradedness of (33).

¹⁶ In 2.2., I showed that a temporal adverb can be inside the target of operations such as VP-fronting.


(i) [_{VP} [_{VP} Talk to Mary] yesterday] though John did, ... (= (8))

I assume that VP fronting can target vP as well as VP, and that the VP in above example should be actually vP.

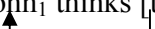
pronouns and NPIs. There are movement-based analyses on these other phenomena proposed in the literature. For example, Heim, Lasnik and May (1991) (HLM) propose that interpretation of the reciprocal *each other* is obtained by covert movement of *each* to the position of the antecedent of the reciprocal. Under this analysis, the LF of (85)a is roughly the same as (85)b.

- (85) a. The men like each other.
 b. The men each like the other.

In (85)a, the *each* part of *each other* moves and adjoins to its antecedent, leaving an anaphoric empty category *e*, as shown in (86).

(86) [[The men]₁ each₂] [like [e₂ other]]


Kayne (2002) gives an analysis of bound pronouns which involves movement. He proposes that a pronoun and its antecedent are in a same relation as a clitic and its double. The pronoun *he* and the antecedent *John* in (87) are base-generated as [*John he*] as in (87)a. *John* inside it then moves out to its surface position as shown in (87)b.

- (87) a. thinks [John he] is smart
 b. John₁ thinks [t₁ he] is smart.


If these approaches are on the right track, the analysis on adjunction here can cover phenomena other than reflexives. I will leave the issues on these phenomena for future research.

3.4. Summary

In this section, I tried to investigate the exact status of the landing site of PP shift, and further, the status of adjunct positions in general. The binding facts (and other hierarchy phenomena) display mixed status of adjoined positions: adjoined positions on the right side of a clause sometimes behave like a low position, sometimes like a high position, and it behaves like it is an outer position. The latter two facts are hard to capture in a theory where a right side element is always lower (Kayne 1994). I proposed one way to account for the data, claiming that an adjoined position does not need to enter into a c-commanding relationship with a phrase it attaches to. I went on to show that the reflexive licensing, for example, does not need to be explained in terms of c-command, using ideas in Hornstein (2001). In such an account, a reflexive inside a PP is licensed by sideward movement.

Section 4: On swiping

In this section, I will discuss the so-called swiping construction in English (Ross 1969; Rosen 1976; Kim 1997; Merchant 2002; van Craenenbroeck 2004, among others), which is considered to be a variation of sluicing (Ross 1969; Chung et. al. 1995; Merchant 2001, among others).

Sluicing is a construction where a *wh*-phrase shows up as a fragment, as in (88)a. In the standard analysis, sluicing is considered to be *wh*-movement followed by IP-deletion, as

illustrated in (88)b (Ross 1969; Merchant 2001; contra Chung et. al. 1995).

(88) *Sluicing*

- a. John danced with someone. I don't know **who**.
 b. I don't know [_{CP} **who**₁ [_{TP} ~~John danced with t₁~~]].
-

In swiping, a sluiced wh-phrase is followed by a preposition as in (89)a. This word order in swiping is peculiar, given the basic word order in English PPs is 'P DP,' as in (89)b.

- (89) a. *Swiping*
 (?)John danced, but I don't know **who with**.
 b. *Pied-piped sluicing*
 John danced, but I don't know **with who**.

I will call the sluicing of whole PP in (89)b 'pied-piped sluicing', as it involves the sluicing of the whole PP; it parallels with the pied-piped wh-question in (90)a.

- (90) a. [**With who**]₁ did John dance t₁?
 b. **Who**₁ did John dance **with** t₁?
 c. *[**Who with**]₁ did John dance t₁?

Note that the swiped word order is never allowed in non-sluiced wh-questions ((90)c). Non-sluiced wh-questions only allow pied-piping ((90)a) or P-stranding ((90)b).

I will consider (i) the licensing condition of swiping and (ii) the internal syntax of swiping in this section. Both arguments serve as additional empirical evidence of the analysis of PP shift as Copy-free movement. First, in section 4.1., I will review a generalization on the licensing condition on swiping (Rosen 1969; Merchant 2002), and show that a complement PP licenses swiping only when it undergoes PP shift. I will show that this fact follows from the assumption that PP shift does not leave a copy. Second, I will point out in section 4.2. that swiping is not sensitive to the ECP, unlike pied-piped sluicing. This is naturally accounted for if we assume that the swiped P undergoes PP shift, and it does not leave a trace behind. I will also consider the implications of these arguments on vacuous movement, the ECP and the nature of islands.

4.1. Licensing of swiping and PP shift

4.1.1. The Givenness Condition in swiping: Merchant (2002)

There is a longstanding generalization regarding the licensing condition of swiping. Swiping is impossible if the antecedent PP is a complement, while it is possible when there is no antecedent PP or when the antecedent is an adjunct PP (Rosen 1976; Merchant 2002). In (91)a, the swiped phrase is *what with* and there is no corresponding *with*-PP adjunct in the antecedent clause: *John fixed it*. Similarly, in (91)b, the swiped phrase *who to* does not have an antecedent PP inside the antecedent clause; it corresponds to the implicit argument of *talk*. These cases of swiping without an antecedent (called 'sprouting' by Chung et. al 1995) are the best cases of swiping. When there is an adjunct PP antecedent as in (91)c, the swiping is slightly degraded (indicated by (?)). Swiping with a complement PP antecedent such as (91)d, on the other hand, is unacceptable for most speakers.

- (91) a. John fixed it, but I don't remember what with.
 b. John was talking, but I don't remember who to.
 c. (?)John fixed it *with something*, but I don't remember what with.
 d. *John talked *to someone*, but I don't remember who to.

In order to account for the above paradigm, Merchant (2002) argues that swiping is licensed only if the relevant PP is not “given” in the antecedent of the elided IP. He observes that a swiped preposition is always focused, as indicated by focal stress pattern in (92).

(92) ... I don't remember who TO/ *WHO to.

He argues that the focused P in swiping cannot be “given” due to the condition in (93) (Merchant 2002: 306).¹⁸

(93) **The Givenness Condition:** The content of the focused P should not be given.

The Givenness Condition correctly accounts for the data in (91). (91)d violates the Givenness Condition because the complement PP is present in the antecedent clause and hence it is “given”. On the other hand, (91)a and (91)b satisfy it because the swiped PP does not show up in the antecedent clause and therefore is not “given.”

The adjunct PP example (91)c is apparently problematic; the presence of the overt antecedent PP seems to violate (93). Merchant, however, circumvents this problem by (i) defining “givenness” in a structurally constrained manner, and (ii) claiming that the VP can be the antecedent of the elided IP under the semantic identity approach to ellipsis (Merchant 2001). First, Merchant assumes that an adjunct PP is adjoined to VP (as in (94)) and therefore is not “given” *within the lowest VP segment*.

(94) John [_{VP} [_{VP} **t_{John} fixed it**] with something],
 but I don't remember [_{CP} [_{PP} what with] [~~PP~~ **John fixed it** [~~PP~~ with what]]].

Note that, under this view, existing information (the PP) can remain “not given” as long as it is structurally outside of the domain under consideration (the lowest VP segment).

Second, Merchant's (2001) theory of ellipsis allows a VP segment to be a potential antecedent for IP-ellipsis (See also, Yoshida 2006). He argues that an elided constituent and its antecedent must satisfy a mutual entailment condition (Merchant 2001). Under this view, the lowest VP segment in (94) is a potential antecedent for the elided IP, because a VP segment containing the trace/copy of the subject (VP-internal Subject Hypothesis: Fukui and Speas 1986, Kitagawa 1986, Kuroda 1988, Koopman and Sportiche 1991) is a full proposition and thus can hold a mutual entailment relation with an IP.^{19, 20} In the above case, the VP [*t_{John} fixed it*] and the

¹⁸ Merchant calls this condition “AvoidF,” following the insight of Schwarzschild (1999). Merchant's definition of “AvoidF”, however, slightly diverges from Schwarzschild's original definition in that “givenness” is structurally determined (see the explanation for (94)). Throughout this paper, we will use the term “the Givenness Condition” to exclusively refer to Merchant's definition of AvoidF.

¹⁹ See Merchant (2001) and Yoshida (2006) for more discussion. See also, Hornstein (1994) for a similar argument on Antecedent Contained Deletion.

²⁰ The analysis of (94) we employ makes a crucial use of the VP-internal subject trace. This contradicts the analysis

elided IP[*John fixed it (with what)*] mutually entails each other and therefore the VP can be the antecedent of the IP-deletion. (For his full analysis of swiping, see 4.2.4.1.) As the adjunct PP *with something* is excluded from the antecedent of IP-deletion (the lowest VP-segment), (94) satisfies the Givenness Condition.

4.1.2. PP shift and the Givenness Condition: Nakao, Ono and Yoshida (2006)

It is reported in Nakao, Ono and Yoshida (2006) that swiping is licensed with a complement antecedent PP, when the antecedent PP undergoes PP shift (as in (95)b).²¹

- (95) a. *John talked [_{PP} to someone] yesterday, but I don't know who to.
 b. ?John talked t_{PP} yesterday [_{PP} to someone], but I don't know who to.

This seems to be direct counterevidence to the Givenness Condition. The antecedent clause in (95)b contains a complement PP antecedent *to someone*. The Givenness Condition crucially relies on the assumption that a complement needs to be contained inside the lowest VP segment, unlike an adjunct. If the complement is generated within the VP segment, how does it license swiping in (95)b?

Nakao et. al. show that this example is not problematic under the assumption that PP shift does not leave a copy. Without the copy of PP shift, the structure of (95)b is as in (96).

- (96) [_{IP} John [_{VP} t_{John} talked [_{PP} _] yesterday] [**PP to someone**],
 but I don't [_{VP} know [_{CP} who+to [_{IP} John [_{VP} talked [_{PP} to who]]]].

The lowest VP segment of the antecedent clause does not contain the copy of *to someone*. Hence, the PP is not “given” inside the lowest VP segment, observing the Givenness Condition.²² On the other hand, if the PP shift leaves a copy, the PP would still be “given” inside the lowest VP segment, as illustrated in (97) and the Givenness Condition would be violated, contrary to the observation in (95).

- (97) [_{IP} John [_{VP} t_{John} talked [**PP to someone**]] yesterday] [**PP to someone**],
 but I don't [_{VP} know [_{CP} who+to [_{IP} John [_{VP} talked [_{PP} to who]]]].

In this way, the data of complement PPs licensing swiping further supports the hypothesis that PP shift does not leave a copy: when a complement PP undergoes PP shift, it does not count

where A-movement does not leave a trace/copy, which I mentioned in section 1 (Lasnik 1998, 1999; Fox 1999; Boeckx 2000). However, these authors' claims are based on other instances of A-movement (e.g. raising, object shift) and they do not consider the status of subject trace in [Spec, VP]. There is still a possibility that only some instances of A-movement leave a trace. I will leave the validity of their claim open in this study.

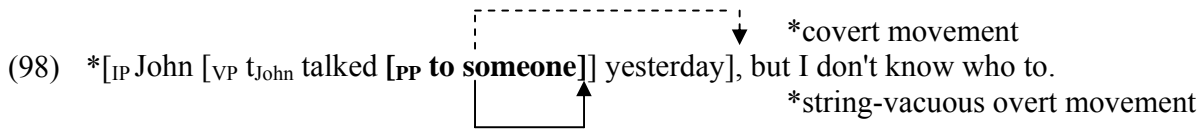
²¹ The “*” and “?” indicate relative judgment rather than absolute judgment. Some of the speakers we consulted did not find the contrast. However, no one gave us the reversed judgment between these examples, which suggests that this subtle contrast is not merely accidental. The argument in this section focuses on judgments by people who consistently observe the contrast in (95).

²² We suspect that the speakers who do not accept (95)b any more than (95)a are less tolerant in taking the lowest VP as the antecedent, because many of such speakers also did not find (91)c better than (91)d. If they do not take the lowest VP as the antecedent, both the adjunct PP in (91)c and the shifted complement PP in (95)d would violate the Givenness Condition. If this speculation is on the right track, there is idiosyncratic variation in the licensing conditions of ellipsis. Merchant's semantic identity account (5.1.1.) only works for people who share our judgment.

in the calculation of “givenness” in the lowest VP.

4.1.3. Implications for LF movement and string-vacuous movement

The account of the contrast in (95) above has an implication for the status of LF movement and vacuous movement. Recall that swiping of the complement PP in (95)a is excluded by the Givenness Condition, while the shifted PP in (95)b escapes this condition. The unacceptability of (95)b, therefore, shows that swiping is not licensed when the complement PP undergoes no “explicit” PP shift. That is, there must not be a derivation like (98) where covert PP shift or string-vacuous PP shift licenses Swiping.



This indicates either that covert/string-vacuous PP shift is impossible, or that such PP shift does not affect swiping even if it is ever possible.²³ On the other hand, Brian Agbayani (p.c.) observes that (95)a improves when there is a focal stress on the PP.

(99) (?)John talked TO SOMEONE (yesterday), but I don't know who to.

The generalization seems to be that covert movement or string-vacuous movement is allowed only when the PP is focused (See 4.2.5. for a related discussion).

Furthermore, the fact that PP shift in (95)b makes swiping possible indicates that PP shift cannot be PF-movement, on the assumption that “givenness” is presumably calculated at LF, rather than PF. If PP shift were PF-movement, the PP in (95)b would remain in-situ in the LF-representation of the sentence and the Givenness Condition would still be violated.

4.2. Swiping and the ECP

In 4.1., we have seen that the PP shift in the antecedent clause affects the licensing condition of swiping. This subsection concentrates on the internal syntax of swiping itself. Following Nakao and Yoshida (2006), I will propose a PP shift analysis of swiping.²⁴

²³ This might naturally follow under Chomsky's (1995) account of covert movement as feature movement. If covert or string-vacuous movement of PP in (98) only involves feature movement (presumably movement of some feature related to focus), the PP category is still in the original position in this example. Under the assumption that givenness is calculated based on existence of the category inside a certain structure (i.e. VP), the PP in (98) should be given even with the covert/string-vacuous PP shift.

²⁴ Our analysis is basically the same as the focus movement analysis by Kim (1997), except for the landing site of the PP shift. Kim assumes the rightward movement of PP is to [Spec, FocusP], which is on the right-side of the Focus head. I assume that the landing site of PP is on the adjoined position, based on the argument in section 3 that shifted PPs have the same status as hanging adjuncts.

(100) *The PP shift analysis of swiping*

- a. [IP [IP John danced PP] [PP **with who**]] PP shift
- b. [CP **who**₁ [IP [IP John danced PP] [PP with **t**₁]]] wh-movement of *who*
- c. [CP **who**₁ [IP [IP John danced PP] [PP with **t**₁]]] IP-deletion at PF

I will show that this analysis nicely captures the fact that swiping is free from ECP violations, unlike pied-piped sluicing. Under our assumption that PP shift does not leave a copy, the absence of the relevant “empty category” in the above derivation explains the fact that swiping is not constrained by the ECP.²⁵

4.2.1. Sluicing out of islands

Under Merchant’s (2001) theory of islands, islands such as complex NPs and adjunct islands are classified as “propositional islands”, and distinguished from PF-islands such as Left Branch Condition and Comp-trace effect on the face of it. Both island effects are ameliorated under sluicing (as is well-known since Ross 1967), but Merchant gives different explanations for the two types of islands under sluicing.

He claims that sluicing involves PF-deletion of IP, and it ameliorates violations of PF-islands, as in (101) and (102).

(101) *Left Branch Condition* (PF-island)

- a. ***[How big]**₁ did she buy [a **t**₁ car]?
- b. She bought a big car, but I don’t know [CP **[how big]**₁ [~~IP she bought [a **t**₁ car]]]].~~

(102) *Comp-trace effect* (PF-island)²⁶

- a. ***Who**₁ is she wondering [CP if **t**₁ left]?
- b. She is wondering if someone left, but I don’t know [CP **who**₁ [~~IP she was wondering [CP it **t**₁ left]]]].~~

Propositional island violations are also not observed under sluicing, as illustrated in (103) and (104).

²⁵ The traditional ECP (Huang 1982; Lasnik and Saito 1984, 1992) is formulated under the assumption that empty categories are traces, rather than copies. An important question remains as to how to formulate the ECP under the CTM; under the CTM, the ECP cannot be principles on empty categories as grammatical formatives but it has to be principles on unpronounced copies of any lexical item. However, as far as the argument of this section is relevant, the distinction between traces and copies is not crucial. I will use the term ‘trace’ interchangeably with ‘copy’ in this subsection just to conform to the standard definition of the ECP.

²⁶ Merchant (2001) also gives examples of *that*-trace effect such as (i), but it is hard to tell whether the underlying structure of the sluice in (i) really includes *that*, given that *that* does not have an obvious semantic content. Therefore, it is not clear if this is an example of amelioration by sluicing.

(i) She said that someone left, but I don’t know [CP **who**₁ [~~IP she said [CP **(that)** **t**₁ left]]]].~~

(103) *Complex NP island* (propositional island)

- a. ***What**₁ does John want to hire [someone who fixes cars with **t**₁]?
- b. John wants to hire [someone who fixes cars with something], but I don't know **what**.

(104) *Adjunct island* (propositional island)

- a. ***Who**₁ will John be mad [if Mary dances with **t**₁]?
- b. John will be mad [if Mary dances with a certain guy], but I don't know **who**.

Merchant claims that, unlike the case of PF-islands, (103)b and (104)b are not real 'island-repair' examples. They are derived from a source that does not involve an island, as illustrated in (105).

- (105) a. John wants to hire [someone who fixes cars with something], but I don't know **what**₁ [*she* fixes cars with **t**₁]. (*She* = e-type pronoun)
- b. John will be mad [if Mary dances with a certain guy], but I don't know **who**₁ [Mary dances with **t**₁] (*in a situation where John will be mad*).

In (105)a, the source of the sluicing employs an e-type pronoun subject, and therefore does not involve the complex NP. In (105)b, the sluiced clause is bound by a null subjunctive operator, which roughly means 'in a situation where John will be mad', and it is not in an adjunct clause anymore. In sum, according to Merchant, the repair effect in these examples is only apparent.

4.2.2. *Argument-adjunct asymmetry and the ECP*

Merchant's analysis of islands above, however, cannot predict the following data: there is an argument-adjunct asymmetry in sluicing out of a propositional island, as pointed out in Lasnik (2005).

- (106) a. John wants to hire [someone who fixes cars with something], but I don't know **what**.
 - b. *John wants to hire [someone who fixes cars for a certain reason], but I don't know (exactly) **why**.
 - c. *John wants to hire [someone who fixes cars in a certain way], but I don't (exactly) know **how**.
- (107) a. John will be mad [if Mary dances with a certain guy], but I don't know **who**.
 - b. *John will be mad [if Mary dances for a certain reason], but I don't know **why**.
 - c. *John will be mad [if Mary dances in a certain way], but I don't know **how**.

When the sluiced wh-phrase is an argument such as *who* and *what*, sluicing out of a complex NP or an adjunct clause is possible ((106)a, (107)a). On the other hand, sluicing of adjunct wh-phrases such as *why* and *how* cannot cross an island ((106)b/c, (107)b/c).

The ungrammaticality of adjunct sluicing out of a propositional island is problematic for Merchant's analysis, as there is no underlying island in these examples under his account.

- (108) a. John wants to hire [someone who fixes cars in a certain way], but
I don't know **how**₁ [*she* fixes cars **t**₁]. (*She* = e-type pronoun)
b. John will be mad [if Mary dances for a certain reason], but
I don't know **why**₁ [Mary dances **t**₁] (*in a situation where John will be mad*).

Therefore, the analysis that employs underlying e-type pronouns and null subjunctives in these examples is hard to maintain.

Following Lasnik's insight, Nakao and Yoshida (2006) suggests an ECP-based account of argument-adjunct asymmetry. According to the traditional definition of the ECP (Huang 1982; Lasnik and Saito 1984, 1992), an empty category must be either antecedent-governed or head-governed by a lexical head. Under the CTM, the notion of 'an empty category' must be rephrased as 'an unpronounced copy'. I will simply assume here that a copy corresponds to a trace in the ECP theory, and just use the term *trace* in this section.

Nakao and Yoshida's assumptions are roughly as follows. First, locality violations are averted by PF-deletion (i.e. sluicing) as Fox and Lasnik (2003) assumes. Second, a propositional island is a barrier for antecedent-government, as in Lasnik and Saito's (1984, 1992) theory. As a result, only adjunct traces violate the ECP when included inside a propositional island (as in (109)b), because an argument trace satisfies the ECP by the lexical government by the preposition (e.g. *with* in (109)a).

- (109) a. I don't know **what**₁ [~~John wants to hire [someone who fixes cars with **t**₁]].
b. *I don't know **why**₁ [~~John wants to hire [someone who fixes cars **t**₁]].~~~~

The ungrammaticality of (109)b suggests that the ECP violation is not remedied by sluicing. This is compatible with the view that the ECP is an LF constraint (Huang 1982)²⁷, while sluicing is PF-deletion (Merchant 2001).

4.2.3. Asymmetry between swiping and pied-piped sluicing with respect to the ECP

Nakao and Yoshida (2006) point out that pied-piped sluicing and swiping show different behavior with respect to propositional islands. Pied-piped sluicing past a complex NP or an adjunct clause is bad ((110)a, (111)a), while corresponding examples of swiping are as good as local swiping of adjunct PPs ((110)b, (111)b; cf. (91)c).²⁸

²⁷ Lasnik and Saito (1992) argue that the ECP is applied at S-structure and LF for argument traces and only at LF for adjunct and intermediate traces. This asymmetry is proposed to exclude (i) as a *that*-trace violation at S-structure, while including (ii) and (iii) using *that*-deletion at LF.

- (i) *Who₁ do you think [that [t₁ left]]?
(ii) Why₁ do you think [that [John left t₁]]
(iii) Who₁ do you think [that [Mary said [t'₁ [t₁ won the race]]]]?

Merchant (2001) claims that the *that*-trace effect is a type of PF-island (see (102)) and is remedied by sluicing. In sum, (i) is an S-structure violation for Lasnik and Saito and a PF violation for Merchant. I will concentrate on ECP violation of adjunct traces *at LF*, claiming that LF violations cannot be ameliorated by sluicing.

²⁸ I owe the judgment of the examples (110)-(112) to my graduate student colleagues in the syntax seminar, Fall 2006 in University of Maryland. Note that, as seen in (91), the best cases of swiping involve sprouting and some speakers find swiping with an adjunct antecedent PP slightly degraded (indicated by "?") regardless of islands.

- (110) a. *John wants to hire [someone who fixes cars with something],
but I don't know **with what**.
b. (?)John wants to hire [someone who fixes cars with something],
but I don't know **what with**.
- (111) a. *John will be mad [if Mary dances with a certain guy], but I don't know **with who**.
b. (?)John will be mad [if Mary dances with a certain guy], but I don't know **who with**.

The above asymmetry has an implication for the structure of swiping. Note that all the examples in (110) and (111) involve sluicing of adjunct PPs (*with*-PPs). If the ECP analysis above is on the right track, swiping cannot have the representation that leaves an adjunct PP trace, as illustrated in (112)b. If it does, the adjunct PP trace inside an island should violate the ECP, in the same way as (112)a.

- (112) a. I don't know [**with what**]_i [~~John wants to hire [someone who fixes cars *t_i]]].
b. I don't know [**what with**]_i [~~John wants to hire [someone who fixes cars *t_i]]].~~~~

The lack of the ECP effect in swiping indicates that an adjunct PP trace should not be left in swiping.

4.2.4. Previous analyses of swiping

I will briefly review two of the recent analyses of swiping to make the above point. I will summarize the analyses of swiping by Merchant (2002) and van Craenenbroeck (2004), and point out that both analyses predict that swiping should leave an offending adjunct trace, as illustrated in (112)b, contrary to fact.

4.2.4.1. Merchant (2002)

Merchant (2002) analyzes swiping as pied-piped wh-movement followed by head-movement plus IP-deletion. Under his analysis, the derivation of swiping is as illustrated in (113).

- (113) John was dancing, but I don't know who with.
- | | | |
|----|---|---------------------|
| a. | [CP [PP with who] [IP John was dancing t _{PP}]] | pied-piping |
| | ↑ | |
| b. | [CP [PP who_D+with_P t _{who}] [IP John was dancing t _{PP}]] | head-inversion (PF) |
| | ↑ | |
| c. | [CP [PP who_D+with_P t _{who}] [IP John was dancing t_{PP}]]] | IP-deletion (PF) |

First, the PP *with who* moves to [Spec, CP] by wh-movement in overt syntax. Then, the wh-phrase *who*, which is a D head, head-adjoints to the preposition by head-movement at PF. Finally, the IP is deleted as in ordinary sluicing.

This analysis has several advantages. First, it accounts for the fact that only simple *wh*-elements can be swiped (e.g. *who*) and not complex *wh*-elements (e.g. *which person*), as shown in (114).

(114) John was dancing, but I don't remember **who with**/***which person with**.

He assumes that simple wh-elements are heads while complex wh-elements are phrases. Only simple wh-phrases allow head-movement in (113)b because phrases cannot be adjoined to heads. Second, his analysis also explains the fact that swiping cannot have a complement PP antecedent by positing the Givenness Condition (See 4.1.1. for details).

However, this analysis does not predict absence of the ECP violation in swiping out of a propositional island. The derivation in (115) leaves an adjunct trace inside an island, which should violate the ECP; it is not antecedent governed because it is inside the island, and not lexically governed, either, being an adjunct PP.

(115) I don't know [**what+with** t_{what}]_I [~~John wants to hire [someone who fixes cars t_i]].~~

4.2.4.2. Van Craenenbroeck (2004)

Van Craenenbroeck (2004) proposes a double-CP analysis for swiping. He claims that the CP of a swiped clause has a double-layer structure: one CP (CP1) selecting another (CP2) as its complement. He accounts for swiping as P-stranding at the Spec position of the inner CP layer, i.e. [Spec, CP2], as shown in (116).

(116) John was dancing, but I don't know who with.

- | | | |
|----|---|-------------|
| a. | [_{CP2} [_{PP} with who] [_{C2} [_{IP} John was talking t_{PP}]]] | pied-piping |
| b. | [_{CP1} who [_{C1} [_{CP2} [_{PP} with t_{who}] [_{C2} [_{IP} John was dancing t_{PP}]]]]] | wh-movement |
| c. | [_{CP1} who [_{C1} [_{CP2} [_{PP} with t_{who}] [_{C2} [_{IP} John was dancing t_{PP}]]]]] | IP-deletion |

The derivation proceeds in the following way. First, the first C (C2) head is merged and the PP *with who* moves to [Spec, CP2]. Second, the second C (C1) head is merged, attracting the wh-phrase. The preposition *with* is stranded in [Spec, CP2]. Then the IP is deleted. The surface form *who with* is thus obtained.

He proposes that the inner CP layer (CP2) represents a focus projection. Therefore, the preposition in [Spec, CP2] in (116) needs to bear focus. This has two good consequences. First, it accounts for the fact that swiping with an antecedent PP is degraded (recall the paradigm in (91)). Antecedent-less swiping such as (91)a/b is better because the swiped P must bear focus and it needs to be 'new' information rather than 'old/given'. (He adopts Merchant's (2002) assumption that the adjunct PP in (91)c does not count as 'given' because it is outside of the lowest VP segment, which is an antecedent of the IP-deletion. See 4.1.1.)

- (91) a. John fixed it, but I don't remember what with.
 b. John was talking, but I don't remember who to.
 c. (?)John fixed it *with something*, but I don't remember what with.
 d. *John talked *to someone*, but I don't remember who to.

Second, it explains the prosodic pattern in swiping. It is observed that, in swiping, the preposition always gets a prosodic stress and not the wh-phrase, as in (92).

of the wh-phrase could be an instance of sideward movement (Nunes 1995, 1999), if it occurs from a subtree (the PP) before it is adjoined to the IP.³⁰ Then the IP is deleted, leaving the hanging position PP behind.

This derivation generates the desired swiping word order, without involving an adjunct PP trace. As a result, the absence of the ECP violation in (112)b is expected. In the derivation of (119), there is no trace of the adjunct PP, and the only trace left by the wh-movement is an argument PP of *what*, which is head-governed by the preposition *with*.

(119) I don't know **what**₁ [~~John~~ [~~wants to hire~~ [~~someone who fixes cars~~ ~~PP~~]]] ^ [PP with **t**₁]

This way, the asymmetry between pied-piped sluicing and swiping is correctly captured under this analysis.

This analysis can accommodate the data set accounted for by Merchant (2002) and van Craenenbroeck (2004). First, by keeping Merchant's Givenness Condition (93) intact (see 4.1.), the licensing condition of swiping illustrated in (91). Second, the fact that the swiped P is focused ((92)) is explained by assuming that rightward movement is induced by some focus requirement. This assumption goes in accordance with the generalization we built in 4.1.3. that string-vacuous movement is allowed only when the PP is focused. The PP shift in (119) is string-vacuous, and such movement is induced by focus. Finally, the fact that only simple wh-elements allow swiping will be accounted for by a slight modification of the derivation in (118). I will return to this in 4.2.6.2.

4.2.6. More on PP shift analysis

There are some potential problems for the derivation in (118). I will refer to some of them in this section and offer potential solutions to them.

4.2.6.1. Right Roof Constraint

First, the PP shift in the derivation of (119) seems to violate Right Roof Constraint in (120).

(120) *Right Roof Constraint* (Ross 1967)

No element may be moved rightward out of the next higher S node.

Under the PP shift analysis, shifted PP in (119) needs to hang from the higher IP, as the shifted PP escapes the deletion of the higher IP.

(81) I don't know **what**₁ [~~IP~~ John [~~wants to hire~~ [~~someone who~~ [~~IP~~ fixes cars ~~PP~~]]]] ^ [PP with **t**₁].

This PP shift crosses the lower IP, which leads to violation of Right Roof Constraint.³¹

³⁰ Recall from Section 3 that we assume that adjuncts and complement PPs that undergo PP shift make subtrees before attachment to the VP/IP, which enables sideward movement from inside a shifted PP. This indicates that movement such as PP shift is a two-part operation: first, the PP is 'detached' from the original position and makes a subtree, and second, it is adjoined to the higher phrase.

³¹ Kim (1997) argues that swiping is sensitive to Right Roof Constraint, using the following data.

We could tentatively assume that this Right Roof Constraint is also a PF constraint and it is repaired under sluicing.³² If this is the case, the implication is that PF amelioration by sluicing not only saves an illicit representation (e.g. traces of movement that cross islands), but also saves an illicit operation, that would have caused a PF-violation if there were no sluicing. The PP shift that violates Right Roof Constraint in (119) does not leave a trace. Therefore, the representation of (119) itself cannot be excluded by any theory of trace distribution. As a consequence, Right Roof Constraint is a constraint on operations, not on traces. If this is ameliorated by sluicing, that means the operation is allowed only when the structure that the illicit PP shift crosses is to be later eliminated by PF-deletion.

4.2.6.2. *Derived position island*

The derivation of in (119) also seems to violate derived position island in (121) (Wexler and Culicover 1980; Merchant 2001).

(121) *Derived position island*

Wh-movement out of a moved element is not allowed.

In (119), wh-movement of *who* occurs out of a PP which had undergone PP shift. Such derivation is normally excluded, as illustrated in (122).

- (122) a. * $[_{CP} \text{Who}_i \text{ did } [_{IP} \text{you talk}] \text{ yesterday } [_{PP} \text{to } t_i]]?$
 b. *I don't know $[_{CP} \text{who}_i [_{IP} \text{you talked}] \text{ yesterday } [_{PP} \text{to } t_i]]$.

I speculate that, as discussed in Nakao and Yoshida (2006), the wh-movement out of a shifted PP in swiping is head-movement onto C head, rather than phrasal movement into $[_{\text{Spec}}, \text{CP}]$. It is not the same type of movement as (122), and presumably is not constrained by the derived position island constraint.³³ (123) illustrates the revised PP shift analysis of swiping.

-
- (i) Mary claimed that the opera was written in the 19th century, but we are not sure **who by**.
 a. *Long reading*
 *we are not sure **who** Mary claimed that the opera was written in the 19th century **by**.
 b. *Short reading*
 we are not sure **who** the opera was written in the 19th century **by**.

However, this is not a correct generalization, given that the following swiping sentence allows long-distance reading and swiping out of a propositional island (119) is possible.

- (ii) Mary claimed that the opera was written *by a certain composer*, but we are not sure **who by**. (Kim 1997)

It seems that swiping without an antecedent ((i)) obeys stricter locality conditions than swiping with an antecedent ((ii)). I have to leave the explanation for the contrast as an open issue.

³² Lasnik (2006) reaches the opposite conclusion, based on properties of English multiple sluicing.

³³ At this point, it is only stipulative to claim that head-movement to C is not constrained by derived position island. A possible principled explanation for this relies on the nature of derived position island. For example, many authors try to reduce derived position island under Chain Uniformity (Takahashi 1994; Ochi 1999; Lasnik and Park 2003; van Craenenbroeck 2004). Under such an analysis, the examples in (122) are ungrammatical because the wh-chain contains a non-uniform chain as in (i) (van Craenenbroeck 2004). The three links of the chain contains two different categories, DP and PP, which breaks the chain uniformity.

(123) John was dancing, but I don't know who with.

- a. [IP [IP John was dancing _PP] [**PP with who**]] PP shift
- b. [CP **who**₁+C⁰ [IP [IP John was dancing _PP] ^ [PP with **t**₁]]] head-movement of *who*
- c. [CP **who**₁+C⁰ [IP [~~IP John was dancing~~ _PP] ^ [PP with **t**₁]]] IP-deletion at PF

The wh-head *who* moves from the shifted PP and head-adjoints to C. This accounts for the fact that only minimal (head) type wh-elements but not phrasal type wh-elements undergo swiping ((114)), in the same way as Merchant's (2002) analysis.³⁴

This kind of derivation, however, should be unavailable for the non-slucied wh-constructions in (122). Otherwise, sentences in (122) would also be saved by the head-movement of *who* onto C head in the same way as (123). I claim that, in non-slucied wh-questions, the existence of I-to-C movement blocks the head-adjunction of wh-heads onto C. In (122)a, the auxiliary *did* is raised to C. I speculate that there is covert I-to-C movement in embedded questions such as (122)b. As a result, the C head in these examples is occupied and wh-heads cannot further attach to C. On the other hand, I-to-C movement never occurs in sluiced configurations (Merchant 2001; Lasnik 2001). (124) shows that the auxiliary *will* cannot survive IP-deletion, showing that the auxiliary is not raised to C in sluicing.

- (124) a. 'John will meet someone.' *Who₁ will [~~John meet t_i~~]?'
 b. *John will meet someone, but I don't know who₁ will [~~John meet t_i~~].

Since there is no auxiliary raising in sluicing, the C head is available for the wh-adjunction in (123). Thus, the derivation involving head-adjunction is only possible in swiping configurations.

4.3. Summary

In this section, I discussed two consequences of the copy-free analysis of PP shift related to swiping. First, if PP shift does not leave a copy, the fact that shifted complement PPs can be an antecedent of swiping is naturally accounted for; the shifted PP does not leave a copy and is not given within the VP, which conforms to the Givenness Condition of the licensing of swiping (Merchant 2002). Second, the copy-free analysis accounts for the fact that adjunct PP swiping across an island does not exhibit the status of an ECP violation; if swiping is derived by

(i) *... DP ... [PP P DP] ... [PP P DP] ... (non-uniform phrasal-chain: DP-PP-PP)

On the other hand, if the second movement involves head movement as I claim for (123), the chain would look like (ii). One can claim that a phrasal movement chain and a head movement chain cannot form one complex chain, unlike the complex phrasal chain in (i). If this line of analysis is on the right track, the lack of derived position island effect in head-movement would be naturally derived.

(ii) ... D ... [PP P D] ... [PP P D]... (two uniform chains: D-D and PP-PP)

³⁴ One might argue that such derivation should be excluded by Head Movement Constraint (HMC) anyway. However, instances of long head-movement are reported in Roberts (1994), Lobeck (1995), Pan and Hu (2000), and Takahashi (2002) and the existence of HMC is doubted by Chomsky (1995). A question, however, remains open as to in what situations long head-movement is allowed.

rightward movement of a PP (Kim 1997; Nakao and Yoshida 2006) and it does not leave a copy, there is no adjunct trace left to violate the ECP.

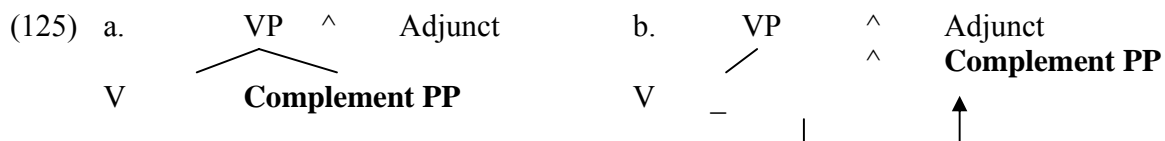
Section 5: Some theoretical considerations on copy-free movement

So far, we have seen empirical arguments to show that PP shift does not leave a copy (Sections 2 and 4). This section considers theoretical issues arising from the assumption that such copy-free movement exists.

5.1. X-bar theory and theta-roles

I have claimed that complement PPs as well as adjunct PPs undergo PP shift without leaving a copy/trace. If no copy of the shifted complement PP is left, that means there is no complement position in the VP at LF. Thus, the X-bar configuration I assumed in (7) is not preserved at LF. Once the complement PP is moved, the PP behaves like an adjunct (under my analysis in Section 3, as a detached subtree). Therefore, the structure only preserves the Projection Principle before PP shift ((125)a) but not after PP shift ((125)b).

(7) [XP [XP [XP Spec [X' X⁰ Complement]] Adjunct] Adjunct]



An obvious question that arises is whether the thematic relation between the V and the complement PP is preserved in (125)b. If theta-roles are assigned only by configurational relationships such as head-complement or head-spec, the thematic information is no longer available in PP shift, which is problematic. However, if theta-roles are features as many authors assume (Lasnik 1995; Bošković and Takahashi 1998; Bošković and Lasnik 1999; Hornstein 1999, 2001. See 3.3.4), we can assume that the theta-feature of the V and the complement PP is checked under the first Merge of the complement PP in (125)a, and erasing the complement position later on is not problematic for theta-checking purposes. If this assumption is correct, it suggests that we need derivations in syntax: theta-roles are checked in the course of the derivation and not assigned/checked in a theta-configuration in a certain representational level. A purely representational theory of syntax, which claims that all principles in syntax are defined in terms of representations, cannot be maintained.

Such an assumption might lead to departure from the strict minimalist idea that LF is the component where all and only features related to ‘meanings’ are interpreted. Since Chomsky (1995), it has been claimed that components other than LF and PF should be eliminated from the grammar. Because ‘sounds’ and ‘meanings’ are the only elements that have absolute necessity in language, the grammar should accommodate only two components that correspond to them. Information about theta-relationships is presumably part of ‘meanings’ you need to interpret from the sentence. If such information is unavailable at LF and it is rather satisfied in the course of derivation (by first Merge), LF is no longer the representation where *all* the meanings are

encoded.

There are at least two ways to solve the problem. The first one is to relax the definition of LF and say theta relationships do not need to be encoded at LF. LF contains other information related to meanings, such as scope and coreference relationships, but the core argument-predicate relationships are determined by Merging an argument to check theta-features. If this line of explanation is on the right track, the information about which predicate checked off theta-features with which arguments needs to go directly to the semantic representation without going via LF. Let us consider the derivation of (126)a. We can assume that each lexical item in the Numeration (126)b is born with a theta-feature such as [agent] and [patient]. In this example, *hit* is inherently born with agent and patient theta-features because of its property as a two-place predicate. There are no such deterministic theta-features for arguments, but the derivation crashes unless there is one argument with an agent theta-feature and another with a patient theta-feature: *John* and *Mary* in the below example³⁵. *Mary* Merges with the verb as a complement to check off the patient features ((126)c), and *John* checks the agent features by merging as a specifier of the verb (I omit the vP projection for simplicity). Note that this first-Merge information is not necessarily retained till LF, given the possibility of copy-free movement such as PP shift. Therefore, the information about ‘what checked which feature’ needs to be either (i) read from the Numeration or (ii) sent off to the semantic representation without being encoded at LF. The skeleton of the semantic representation (126)e is obtained from this information.³⁶

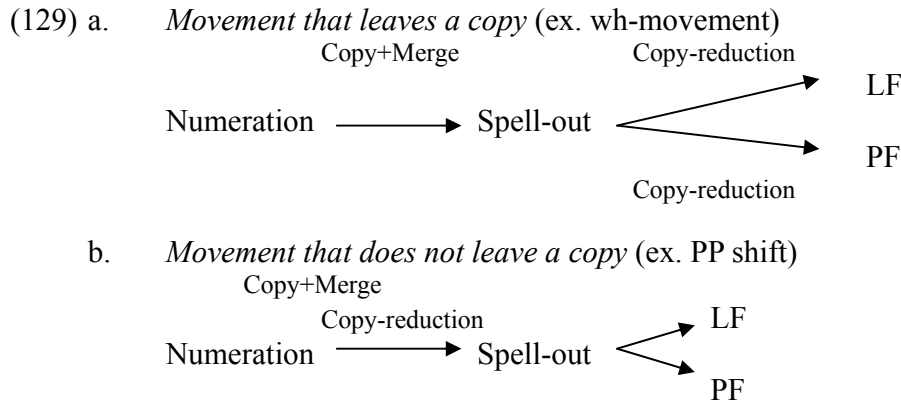
- (126) a. John talked to Mary.
b. N = {John₁[agent], talk₁[agent, patient], to, Mary[patient]}
c. [talked[agent, ~~patient~~] to Mary[~~patient~~]]
d. [John[~~agent~~] [talked[~~agent, patient~~] to Mary[~~patient~~]]
e. $\exists e$ [Agent(John, e) & talk(e) & Patient((to) Mary, e)]

Whether you choose (i) or (ii), you need to modify the minimalist assumption of the grammatical components in (127). If (i) is the case, the semantic representation should be constructed combining the theta-information in Numeration and other information at LF, as illustrated in (127)b. On the other hand, if (ii) is the case, the semantic representation should be constructed partially based on the syntactic derivation before spell-out as in (127)c.

³⁵ Alternatively, we could assume that an argument only has non-deterministic theta-features, whose interpretation changes according to where it is Merged (just [θ], rather than [agent] or [patient]). The choice does not affect the argument here.

³⁶ I do not commit to the issue of the preposition itself here and I will just assume the PP *to Mary* itself gets a patient theta-role of *talk* in this example. However, it might be the case that *Mary* gets a theta-role from the combination of the verb and the (displaced) P. In that case, representing the theta-feature inside the Numeration becomes a more complicated issue because the V and the P are separate lexical items in the Numeration.

point of PF and LF. If the copy reduction happens before spell-out, the information of the original copy does not reach PF or LF, as illustrated in (129)b. It leads to the same effect as movement having no copy. On the other hand, copy reduction occurs separately at LF and PF, as in (129)a, in movement that leaves a copy; thus, the copy to interpret at LF can be different from the copy to pronounce at PF, which causes reconstruction phenomena.



This way, we can still accommodate copy-free movement without treating ‘Move’ as a primitive grammatical operation.

However, I have not answered the important question as to *why* there are two types of movement. Specifically, I will need to make clear the following points.

- (130) a. What is the typological distinction between two movement categories?
 b. Why is there such a distinction? Why cannot copy-reduction happen before spell-out in (129)a and/or why cannot it happen after spell-out in (129)b?
 c. Is either of (129)a and b more economical than the other? Which one is the ‘default’ case (i.e. more economical application of copy-reduction) and which one is marked?

(130)a is an empirical question about the typology of movement. In this paper, I have shown that PP shift is a copy-free movement. However, what makes PP shift as such is yet to be defined. For example, PP shift is rightward movement of PPs, but it is not the case that all types of PP movement do not leave a copy; wh-movement of a PP, for example, shows reconstruction.

(131) [For which friend of himself₂]₁ did John₂ work t₁?

Also, it does not seem to be the case that all types of rightward movement do not leave a copy. For example, Phillips (1996) shows that Heavy NP shift shows binding reconstruction as in (132), unlike PP shift.

- (132) a. I described [the victim whose sight had been impaired by the explosion]₁ to himself₁.
 b. I described _ to himself₁ [the victim whose sight had been impaired by the explosion]₁

We need to widen the empirical coverage of the inquiry and clarify what exactly is the condition of copy-free movement. There might be other such movement except for ‘PP movement to the

right’.

If we find out what movement can be copy-free, (130)b is a question to be asked: why is such movement only allowed in that category of movement? We need to investigate what difference between, for example, wh-movement and PP shift is responsible for the distinction.

The question (130)c is closely related to this question. In Minimalist syntax, the notion of economy plays a huge role. In contrast to GB syntax, where you can apply any operation as many times as you like unless the derivation or the representation is filtered out by some rules of grammar, Minimalism allows you to apply operations only when necessary. A more economical derivation/representation is always preferred to less economical ones. For example, Chomsky (1991) claims that V to I movement is more economical when it is covert, and that overt V to I movement is only allowed when it is necessary because of a strong feature. Similarly, if there is a preference between two types of movement in (129), the less preferred one must have a reason not to choose the more economical strategy (it becomes the answer to (130)b). On the other hand, the preferred one does not require such explanation because it is the default case. To answer those three questions is the biggest challenge for the future research.

Section 6: Conclusion

This paper considered properties of PP shift and argued that there is a type of movement that apparently does not leave a copy behind. I suggested a way to accommodate such movement within the Copy Theory of Movement, which is standardly assumed in the Minimalist syntax theory. We have seen consequences of existence of such movement. The observation that complement PPs can move without a copy suggests that their theta-roles are satisfied derivationally, which in turn suggests that derivational approach to syntax is indispensable. Also, if the theta-roles are derivationally satisfied and the information about theta-checking relations is not encoded at LF, the traditional notion of LF as a locus of “all” semantic information might have to be modified. Although there are questions remaining as to what kind of movement is allowed to leave no copy, the accommodation of copy-less movement in minimalist syntax will broaden the empirical coverage of movement in language.

References

- Baltin, Mark (1987) “Do Antecedent-Contained Deletions Exist?,” *Linguistic Inquiry* 18: 279-295.
- Boeckx, Cedric (2000) “A Note on Contraction,” *Linguistic Inquiry* 31: 357-365.
- Boeckx, Cedric (2001) “Scope Reconstruction and A-movement,” *Natural Language & Linguistic Theory* 19: 503-548.
- Bošković, Željko and Daiko Takahashi (1998) “Scrambling and Last Resort,” *Linguistic Inquiry* 29: 347-366.
- Bošković, Željko and Howard Lasnik (1999) “How Strict Is the Cycle?,” *Linguistic Inquiry* 30: 691-703.
- Bresnan, Joan (1971) “Contraction and the Transformational Cycle,” Ms., MIT, Cambridge, MA.
- Chomsky, Noam (1976) “Conditions on Rules of Grammar,” in *Essays on Form and Interpretation*, North-Holland: New York.

- Chomsky, Noam (1981) *Lectures on Government and Binding*, Foris: Dordrecht.
- Chomsky, Noam (1991) "Some Notes on Economy of Derivation and Representation," *Principles and Parameters in Generative Grammar*, ed. by Robert Freidin, MIT Press: Cambridge, MA: 417-454. [Reprinted in Chomsky (1995): 129-166.]
- Chomsky, Noam (1993) "A Minimalist Program for Linguistic Theory," *The View from the Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Ken Hale and Samuel J. Keyser, MIT Press: Cambridge, MA: 1-52. [Reprinted in Chomsky (1995): 167-217.]
- Chomsky, Noam (1995) *The Minimalist Program*, MIT Press: Cambridge, MA.
- Chung, Sandra, William A. Ladusaw, and James McCloskey (1995) "Sluicing and Logical Form," *Natural Language Semantics* 3: 239-282.
- Craenenbroeck, Jeroen van (2004) *Ellipsis in Dutch Dialects*, Ph.D. dissertation, Leiden University, Leiden.
- Epstein, Samuel D. (1987) *Empty Categories and Their Antecedents*, Ph. D. dissertation, University of Connecticut, Storrs.
- Epstein, Samuel D. (1999) "Un-Principled Syntax," *Working Minimalism*, ed. by Samuel D. Epstein and Norbert Hornstein, MIT Press: Cambridge, MA: 317-345.
- Epstein, Samuel D. (2000) "Adjunction and Pronominal Variable Binding," *Essays in Syntactic Theory*, Routledge: London: 51-66.
- Ferro, Lisa (1994) "Interpreting Coreference in Sentences with Focus: A Discourse-Based Alternative to Weak Crossover," Ms. University of Connecticut.
- Fox, Danny (1999) "Reconstruction, Binding Theory and the Interpretation of Chains," *Linguistic Inquiry* 30: 157-196.
- Fox, Danny and Howard Lasnik (2003) "Successive-Cyclic Movement and Island Repair: The Difference between Sluicing and VP-Ellipsis," *Linguistic Inquiry* 34: 143-154.
- Fox, Danny and Jon Nissenbaum (1999) "Extrapolation and Scope: A Case for Overt QR," *Proceedings of WCCFL 18*, ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, Cascadia Press: Somerville, MA: 132-144.
- Fukui, Naoki and Margaret Speas (1986) "Specifiers and Projection," *MIT Working Papers in Linguistics* 8: 128-172.
- Hornstein, Norbert (1994) "An Argument for Minimalism: The Case of Antecedent-Contained Deletion," *Linguistic Inquiry* 25: 455-480.
- Hornstein, Norbert (1995) *Logical Form: From GB to Minimalism*, Blackwell: Cambridge, MA.
- Hornstein, Norbert (2001) *Move! A Minimalist Theory of Construal*, Blackwell, Oxford.
- Hornstein, Norbert (2003) "On Control," *Minimalist Syntax*, ed. by Randall Hendrick, Blackwell: Malden, MA: 6-81.
- Hornstein, Norbert, and Jairo Nunes (2006) "Some Thoughts on Adjunction," Ms., University of Maryland, College Park.
- Hornstein, Norbert, Jairo Nunes and Kleanthes K. Grohmann (2005) *Understanding Minimalism*, Cambridge University Press: Cambridge/ New York.
- Huang, C.-T. James (1982) *Logical Relations in Chinese and the Theory of Grammar*, Ph.D dissertation, MIT: Cambridge, MA.
- Johnson, Kyle and Satoshi Tomioka (1997) "Lowering and Mid-Size Clauses," *Reconstruction: Proceedings of the 1997 Tübingen Workshop*, ed. by Graham Katz, Shin-Sook Kim, and Heike Winhart, Universität Stuttgart and Universität Tübingen: 185-206.
- Kayne, Richard S. (1994) *The Antisymmetry of Syntax*, MIT Press: Cambridge, MA.
- Kayne, Richard S. (2002) "Pronouns and Their Antecedents," *Derivation and Explanation in the*

- Minimalist Program*, ed. by Samuel D. Epstein and T. Daniel Seely, Brackwell: Oxford: 133-166.
- Kim, Jeong-Seok (1997) *Syntactic Focus Movement and Ellipsis: A Minimalist Approach*, Ph.D dissertation, University of Connecticut, Storrs.
- Kitagawa, Yoshihisa (1986) *Subjects in Japanese and English*, Ph.D. dissertation, University of Massachusetts, Amherst.
- Koopman, Hilda (1984) *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*, Foris Publications: Dordrecht.
- Koopman, Hilda and Dominique Sportiche (1991) "The Position of Subjects," *Lingua* 85: 211-258.
- Kupin, Joseph J. (1978) "A Motivated Alternative to Phrase Markers," *Linguistic Inquiry* 9 (2): 303-308.
- Kuroda, Shige-Yuki (1988) "Whether We Agree or Not: A Comparative Syntax and English and Japanese," *Linguisticae Investigationes* 12: 1-47.
- Lakoff, George and John R. Ross (1976) "Why You Can't Do So into the Sink," *Notes from the Linguistics Underground*, vol.7 of *Syntax and Semantics*, ed. by James McCawley, Academic Press: New York: 101-111.
- Larson, (1988) "On Double Object Construction," *Linguistic Inquiry* 19: 335-391.
- Lasnik, Howard (1984) "Lecture 5: Further Properties of Empty Categories," Handout from a lecture at Tokyo Linguistics Seminar: A series of lectures on current issues in the theories of syntax and logical form, Tokyo, August 1984.
- Lasnik, Howard (1995) "A Note on Pseudogapping," *MIT Working Papers in Linguistics 27: Papers on Minimalist Syntax*, MIT, Cambridge, MA: 143-163.
- Lasnik, Howard (1998) "Some Reconstruction Riddles," *University of Pennsylvania Working Papers in Linguistics 5*: 83-98.
- Lasnik, Howard (1999) "Chains of Arguments," *Working Minimalism*, ed. by Samuel Epstein and Norbert Hornstein, MIT Press: Cambridge, MA: 189-215.
- Lasnik, Howard (2001) "When Can You Save a Structure by Destroying It?" *Proceedings of the NELS 31*, GLSA publications: Amherst: 301-320.
- Lasnik, Howard (2005) "How to Evade Moving Violations," Handout from a lecture at LSA Summer Institute, MIT, Cambridge, MA, July 2005.
- Lasnik, Howard (2006) "Multiple Sluicing in English?," A talk at Linguistics Colloquium Series, University of Maryland, College Park, September 2006.
- Lasnik, Howard and Mamoru Saito (1984) "On the nature of proper government," *Linguistic Inquiry* 15: 235-289.
- Lasnik, Howard and Mamoru Saito (1992) *Move α : Conditions on Its Application and Output*, MIT Press: Cambridge, MA.
- Lasnik, Howard, and Myung-Kwan Park (2003) "The EPP and the Subject Condition under Sluicing," *Linguistic Inquiry* 34: 649-660.
- Lightfoot, David (1976) "Trace Theory and Twice-Moved NPs," *Linguistic Inquiry* 7: 535-571.
- Lobeck, Anne (1995) *Ellipsis: Functional Heads, Licensing, and Identification*, Oxford University Press: Oxford.
- Merchant, Jason (2001) *Syntax of Silence*, Oxford University Press: Oxford.
- Merchant, Jason (2002) "Swiping in Germanic," *Studies in Comparative Germanic Syntax*, ed. by Werner Abraham and Jan-Wouter Zwart, John Benjamins: Amsterdam: 295-321.
- May, Robert (1977) *The Grammar of Quantification*, Ph.D. dissertation, MIT, Cambridge, MA.
- Musolino, Julien, and Jeffery Lidz (In press) "Why Children Aren't Universally Successful with

- Quantification,” *Linguistics*.
- Nakao Chizuru, Hajime Ono and Masaya Yoshida (2006) “When a Complement PP Goes Missing: A Study on the Licensing of Swiping,” to appear in *Proceedings of WCCFL 25*. Cascadilla Proceedings Project: Somerville, MA.
- Nakao Chizuru, and Masaya Yoshida (2006) “‘Not-So-Propositional’ Islands and Their Implications for Swiping,” to appear in *Proceedings of WECOL 2006*, University of California, Fresno (online publication).
- Nevins, Andrew and Pranav Anand (2003) “Some AGREEMENT Matters,” *Proceedings of WCCFL 22*, ed. by Gina Gardina and Mimura Tsujimura, Cascadilla Press: Somerville, MA: 101-114.
- Nunes, Jairo (1995) *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*, Ph.D. dissertation, University of Maryland, College Park.
- Nunes, Jairo (1999) “Linearization of Chains and Phonetic Realization of Chain Links,” *Working Minimalism*, ed. by Samuel D. Epstein and Norbert Hornstein, MIT Press: Cambridge, MA: 217-149.
- Nunes, Jairo (2004) *Linearization of Chains and Sideward Movement*, MIT Press: Cambridge, MA.
- Ochi, Masao (1999) “Some Consequences of Attract F,” *Lingua* 109: 81-107.
- Pan, Haihua and Jianhua Hu (2000) “Head Noun Movement, Focus, and Topicalization in Mandarin Chinese,” Ms., City University of Hong Kong.
- Pesetsky, David (1995) *Zero Syntax: Experiencers and Cascades*, MIT Press: Cambridge, MA.
- Phillips, Colin (1996) “Merge Right: An Approach to Constituency Conflicts,” *Proceedings of WCCFL 15*, ed. by Brian Agbayani and Sze-Wing Tang, CSLI Publications: Stanford, CA: 381-395.
- Reinhart, Tanya (1979) “The Syntactic Domain for Semantic Rules,” *Formal Semantics and Pragmatics for Natural Languages*, ed. by F. Guenther and S. J. Schmidt, Reidel: Dordrecht.
- Roberts, Ian (1994) “Two Types of Head Movement in Romance,” *Verb Movement*, ed. by Norbert Hornstein and David Lightfoot, Cambridge University Press: Cambridge: 207-242.
- Rosen, Carol (1976) “Guess What About?,” *Proceedings of NELS 6*, Montreal Working Papers in Linguistics: Montreal: 205-211.
- Ross, John (1967) *Constraints on Variables in Syntax*, Ph. D dissertation, MIT: Cambridge, MA.
- Ross, John (1969) “Guess Who?,” *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society: University of Chicago, Chicago: 252-286.
- Schwarzschild, Roger (1999) “GIVENness, AVOIDF, and Other Constraints on the Placement of Accent,” *Natural Language Semantics* 7: 141-177.
- Takahashi, Daiko (2002) “Determiner Raising and Scope Shift,” *Linguistic Inquiry* 33: 575-616.
- Tanaka, Hidekazu (2005) “Antecedent-Contained Deletion without Quantifier Raising,” *York Papers in Linguistics* 3. University of York: 173-184.
- Uriagereka, Juan (1999) “Multiple Spell-Out,” *Working Minimalism*, ed. by Samuel D. Epstein and Norbert Hornstein, MIT Press: Cambridge, MA: 251-282.
- Uriagereka, Juan (2006) *Syntactic Anchors (On Semantic Structuring)*, Cambridge University Press: Cambridge.
- Uriagereka, Juan (2002) “Pure Adjuncts,” *Actas del 11 Congreso de Gramática Generativa*, Universidad de Zaragoza.
- Yoshida, Masaya (2006) “Sometimes Smaller is Better: Sluicing, Gapping and Semantic Identity,” to appear in *Proceedings of NELS 36*, GLSA Publications: Amherst, MA.
- Wexler, Kenneth, and Peter W. Culicover (1980) *Formal Principles of Language Acquisition*,

MIT Press: Cambridge, MA.