Some Milestones in the Development of Binding Theory

Howard Lasnik

University of Maryland

lasnik@umd.edu

The original theories of anaphoric relations were transformational. Anaphoric pronouns were derived from full NPs; reflexives were derived from pronouns, as were missing subjects of certain infinitivals (EQUI).

(1)   "... there are actually good reasons for introducing [pronouns] transformationally as substituends for *NP*'s of the proper kind... [a] transformation ... replaces an animate NP by "he.""   Chomsky (1955, p.502)

(2)   "In §109.6 we found that such sentences as "John wanted him to come" (= 311) are introduced by transformation. If we investigate these sentences in more detail, we discover that there are certain restrictions on the occurrence of pronouns. Alongside of 311 we have 325 but not 326:
    325   (a) *I wanted him to try*
           (b) *I wanted you to try*
           (c) *I wanted to try*
    326   *I wanted me to try*
           The only way to avoid a special restricting statement on the level **P** is to add a mapping Φ_x^P that carries 326 into 325c:
    327   Φ_x^P carries *I-want-I-to ⌒try* into *I-want-to ⌒try*"   LSLT p.519

(3)   *You wanted you to try

Third person is slightly trickier:

(4)   328   (a) *he wanted him to try*
               (b) *he wanted to try*
    The simplest way to handle this situation appears to be to set up two distinct elements *he* and *he** corresponding to the element *he^w* of **W**, *he* being an element just like *I* and *you* (which accounts for 328b), and *he** being an ordinary proper noun (which accounts for 328a, just as we have "he wanted John to try"). The establishment of this pair of homonyms on the level **P** is further supported by its usefulness for other purposes, as we will see below. We thus replace 327 by the more general characterization

    329   Φ_x^P carries $\begin{Bmatrix} I \\ you \\ he \end{Bmatrix}$ *-want-* $\begin{Bmatrix} I \\ you \\ he \end{Bmatrix}$ *-to ⌒try*   into
    $\begin{Bmatrix} I \\ you \\ he \end{Bmatrix}$ *-want -*   *to ⌒try*

Thus, bound anaphoric and freely referring ones are different lexical items. For instance, (5) is ambiguous, rather than vague.

(5)   John thinks he is clever

(6)   The setting up of two elements he and he* also
        has support in the referential use of these words. Whereas *I* and *you*
        have an unambiguous reference in sentences like "John said that I
        would come," "John said that you would come," in "John said that he
        would come" the reference is ambiguous. In our terms, if "he" in this
        sentence is derived from the syntactic element *he* (the pronoun), the
        reference is to *John*, as it is in 330b, where *he* becomes Ø; if it is derived
        from the syntactic element *he** (the proper noun), the reference is to
        a second person, as in 330a. While the results of our grammatical
        analysis naturally tell us nothing about reference, this syntactic discussion
        does provide the means for an adequate description of reference in some
        semantic description of English.   p.525

(7)   We also have
        330   (a) *John wanted him to try*
                (b) *John wanted to try*
                (c) *John wanted me to try*
                The simplest way to account for this is to revise Φ_x^P, replacing
        329 by

        331   Φ_x^P carries $\left\{ \begin{array}{c} I \\ you \\ NP_y \end{array} \right\}$ -want- $\left\{ \begin{array}{c} I \\ you \\ he \end{array} \right\}$ -to⌢try   into

                $\left\{ \begin{array}{c} I \\ you \\ NP_y \end{array} \right\}$ -want - to⌢try $(NP_y \neq I, you)$

                Φ_x^P applied to "John wanted him* to try" gives 330a; i.e., it does
        not apply. Applied to "John wanted him to try," Φ_x^P gives 330b.   pp.519-520

What about reflexives?

(8) T_{self} as we have stated it is only a special case of a quite general
        transformation. Note that "him," "you," etc. in 351 are each the object
        of the complex verb "persuade-to try" (an instance of *Verb-Complement*).
        But it is true in general that the object of a verb undergoes the transformation
        T_{self}, even in kernel sentences. Thus we have
        355   *he saw himself*, etc.
        The transformation T_{self} must be reformulated, then, to hold more
        generally of *NP-V_T-NP*, and it must be given as a mapping, since it is
        obligatory even for kernel sentences. We replace 352 by 356, defining
        the component Φ_0^P of Φ^P

356   Φ_x^{P} is determined by $(Q_{self}, \delta_{self})$, where

$$Q_{self} = \left\{ \left(\left\{ \begin{array}{c} I \\ you \\ NP \end{array} \right\} VP_A \frown V_T, \left\{ \begin{array}{c} I \\ you \\ he \end{array} \right\} \right) \right\} \quad (NP \neq I, you)$$

$\delta_{self} \leftrightarrow (U, U, U, self)$ p.530

Locality of the 'clause-mate' type was also first hinted at in LSLT:

(9)   The investigation of "self" and its distribution will, I think, turn out to be of some importance for the study of transformational structure. The occurrence in a sentence of *X-self*, for some pronoun *X*, indicates a special relation between this element and some noun or some other pronoun in the sentence. To give a general rule concerning all such cases directly for all sentences may be quite difficult. But we may be able to show that whenever this relation exists between two positions in a complex sentence *Z*, the elements filling these places have some fixed relation in the kernel structures from which *Z* is derived (e.g., they may be subject and object). If so, the distribution of *self* may be statable simply in terms of this kernel sentence relation. The simplification thus introduced can be an important support for transformational analysis, in particular cases.   p.534

Chomsky (1965) has interesting discussion of reflexivization, referencing Lees and Klima (1963). In this discussion, **referential indices** are introduced as part of the theory, to further explicate the notion of 'identity' relevant to deletion operations (of which pronominalization and reflexivization are two):

(10)   "As an additional illustration, consider the reflexivization operation (see Lees and Klima, 1963, for a detailed discussion). It has frequently been observed that in a sentence such as "John hurt John" or "the boy hurt the boy," the two phonetically identical Noun Phrases are necessarily interpreted as differing in reference; sameness of reference requires reflexivization of the second Noun Phrase (this is also true of pronominalization). Various attempts have been made to build an account of this into the syntactic component, but none has been very convincing. The availability of lexical features suggests a new approach that might be explored. Suppose that certain lexical items are designated as "referential" and that by a general convention, each occurrence of a referential item is assigned a marker, say, an integer, as a feature. The reflexivization rule can be formulated as an erasure operation that uses one Noun Phrase to delete

another. As in the case of relativization ... the erasure
leaves a residue, in particular, the feature [±Human] and it introduces
the new phonetic element *self*. Thus when applied to "I
hurt I," the first Noun Phrase is used to delete the second, finally
giving, "I hurt myself." But by the recoverability condition on
deletion, the reflexivization rule (similarly, the pronominalization
rule) will apply only when the integers assigned to the two items
are the same." pp.145-146

(11) "The semantic component will then interpret two
referential items as having the same reference just in case they
are strictly identical - in particular, in case they have been
assigned the same integer in the deep structure. This gives the
right answer in many cases, but there are interesting problems
that arise when the referential items are plural, and of
course there are problems in specifying the notion "referential"
properly." p.146

Clause-mate was the standard locality configuration for the first couple of generations:

(12) Lees and Klima (1963):
A. Reflexive Rule
X - *Nom* - Y -*Nom'* - Z $\rightarrow$ X - *Nom* - Y -*Nom'* + *Self* - Z
where *Nom* = *Nom'* = a nominal, and where *Nom* and *Nom'* are within the same simplex sentence
B. Pronoun Rule
X - *Nom* - Y -*Nom'* - Z $\rightarrow$ X - *Nom* - Y -*Nom'* + *Pron* - Z
where *Nom* = *Nom'*, and where *Nom* is in a matrix sentence while *Nom'* is in a constituent sentence embedded within that matrix sentence

(13)a   I see myself
   b   *I see me
(14)a   I told John to protect me
   b   *I told John to protect myself

(15) Reciprocal Rule
X - $N+Pl$ - Y -$N'+Pl$ - Z $\rightarrow$ X - $N+Pl$ - Y -$N'+Pl$ + *Recip* - Z
where $N = N'$ and they are within the same simplex, and where N is a noun, Pl is the plural morpheme, and *Recip* is the reciprocal morpheme

(16) John and Mary frighten one another
(17)a *They forced the king to help one another
   b *They forced the king to help themselves

Chomsky (1965) alluded to problems with plurals. Some of the problems raised by plurals are first discussed by Postal (1966):

(18)a *I like us
    b *We like me   Postal (1966)

(19)   The Inclusion Constraint (TIC)   Postal (1969)
(20)a *We were proud of me
    b *John and I were proud of me

(21)a *Max and Mary$_i$ criticized her$_i$
    b *Max$_i$ criticized Mary and him$_i$
    c *Max and I criticized me   Grinder and Postal (1971)

(22)a I think we will win
    b We think I will win
    c Max and Mary$_i$ think she$_i$ will win
(23)   TIC is defined over "Clause Mates".   Grinder and Postal (1971)

(24)   "It is apparent ... that TIC is a restriction on the possible syntactic reflection of certain
        Semantic Representations. It is not a constraint on the well-formedness of Semantic
        Representations per se. This is particularly clear in cases like the following, where the
        Clause Mate structure over which TIC is defined is a result of transformational
        operations, in this case the rule RAISING:"
(25)a   Joe$_h$ believes that we$_{h,i}$ are under surveillance by the FBI
    b *Joe$_h$ believes us$_{h,i}$ to be under surveillance by the FBI

(26)a   We believe that I am under surveillance by the FBI
    b *We believe me to be under surveillance by the FBI

The Chomsky (1965) take on reflexivization and clause-mates:
(27)   "Notice, incidentally, that the reflexivization rule does not always
        apply (though pronominalization does) even when the two
        Nouns are strictly identical and hence coreferential. Thus we
        have "I kept it near me" alongside of "I aimed it at myself,"
        and so on. The difference is that in the first, but not the second,
        the repeated Noun is in a Sentence-Complement to the Verb.
        Thus "I kept it near me" has a deep structure of the form
        "I - kept - it - # S #," where S dominates "it is near me." But
        "I aimed it at myself" has a deep structure of the form "I -
        aimed - it - at me" (there is no underlying sentence "it is at
        me"). The reflexivization rule does not apply to a repeated N
        dominated by an occurrence of S that does not dominate the
        "antecedent" occurrence of N."   p.146

The 'Insertion Prohibition'
(28)   "This particular remark about
        English is, apparently, a consequence of a more general condition

on transformations, namely that no morphological material (in this case, *self*) can be introduced into a configuration dominated by S once the cycle of transformational rules has already completed its application to this configuration (though items can still be extracted from this constituent of a larger "matrix structure," in the next cycle of transformational rules)."   p.146

The introduction of command

Langacker (1969) explored the following question:

(29)   "Under what circumstances can a definite noun phrase NP^a be used to pronominalize an identical noun phrase NP^p?"   p.160

Interestingly, Langacker does not commit himself to a transformational approach:

(30)   "The analysis to be proposed is essentially neutral with respect to various alternative ways of representing pronominalization in a generative grammar. For example, it makes no difference for our purposes whether pronominalization is effected by a single transformational rule or whether a series of operations is involved. Likewise, it makes no difference whether pronouns are derived by reducing fully specified underlying noun phrases or whether they are present in deep structure (although we will adopt the former alternative for purposes of exposition)."   p.161

Langacker begins the discussion with an observation similar to the generalizations embodied in (12) above:

(31)   "Reflexivization may be regarded as a special form of pronominalization, the form that occurs when NP^a and NP^p are constituents of the same simple sentence."   p.162

First structural constraint on reflexivization:

(32) NP^p cannot precede NP^a

(33)a   **Those men** outsmarted **themselves**
     b   ***Themselves** outsmarted **those men**
(34)a   Penelope told **Peter** about **himself**
     b   *Penelope told **himself** about **Peter**

Langacker indicates that this precedence constraint ultimately follows from a more general constraint governing all types of pronominalization. More on this later.

Complementarity between pronouns and reflexives

(35)   (12) simply stipulates the complementarity. The reflexivization and pronominalization transformations apply in complementary domains.

(36)   Langacker tries to do better:

(37)   "By ordering the obligatory reflexivization rule before the regular pronominalization rule, we avoid having to restrict pronominalization so that it cannot apply when NP^a and NP^p are in the same simple sentence."   p.163

We need to block pronominalization from turning (38) into (39).

(38)   **Those women** admire **those women**
(39)   *__Those women__ admire **them**

If reflexivization comes first, it obligatorily turns (38) into (40), so (39) can never arise

(40)   **Those women** admire **themselves**

(41)   However, "it seems best to consider reflexivization and pronominalization as two variants of the same process ... There are two reasons for this. First, the two operations are very similar, formally and intuitively. Second, the constraint we will formulate for pronominalization will serve without modification to rule out reflexive sentences such as [(33)b] and [(34)b]."   p.163

A potential way to capture the complementarity without ordering two separate rules:

(42)   "... reflexivization can be viewed as the result of regular pronominalization plus some other operation (such as *self* insertion)."   p.163

Linear order also seems to be relevant when NP^a and NP^p are in separate clauses:

(43)a **Ralph** is much more intelligent than **he** looks
    b **This woman** hates the man who wronged **her**
    c I will give **these kittens** to the girl who wants **them**
    d Tell **that man** that **he** can't go in there
    e **The artist** who painted these pictures must work while **he** is sleeping

(44)a *__He__ is much more intelligent than **Ralph** looks
    b *__She__ hates the man who wronged **this woman**
    c *I will give **them** to the girl who wants **these kittens**
    d *Tell **him** that **that man** can't go in there
    e *__He__ must work while **the artist who painted these pictures** is sleeping

However, sometimes a pronoun may precede its antecedent:

(45)a The woman who is to marry **him** will visit **Ralph** tomorrow
    b The man who wronged **her** is hated by **this woman**
    c The girl who wants **them** will receive **these kittens** as a present from me

(46)   "We might hypothesize ... that pronominalization is possible just in case NP^a is higher in the tree than NP^p, meaning that NP^a is dominated by fewer S-nodes than NP^p."   p.164

However, "there are cases where the pronoun is higher in the tree than its antecedent":

(47)a   The woman who is to marry **Ralph** will visit **him** tomorrow
   b   The man who wronged **this woman** is hated by **her**
   c   The girl who wants **these kittens** will receive **them** as a present from me

(48)   "Therefore, neither a constraint based on linear ordering nor a constraint based on relative depth of embedding is sufficient to handle pronominalization ..."   p.165

(49)   "At the same time, however, both of these factors seem to be pertinent in some way, which suggests that the restrictions on pronominalization can be stated when they are considered simultaneously."

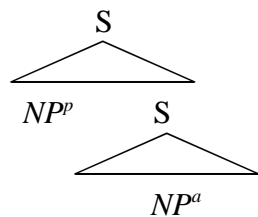Four possibilities for the interaction of the two factors

(50) NP^a precedes NP^p and is higher in the tree than NP^p:

   Ralph is much more intelligent than **he** looks

```
            S
          /   \
        /       \
     NP^a        S
               /   \
             /       \
            NP^p
```
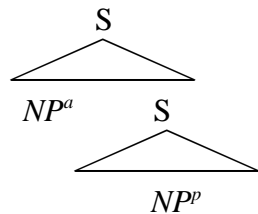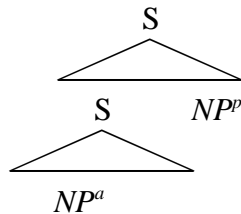
(51) NP^p precedes NP^a and is higher in the tree than NP^a:

   *** He** is much more intelligent than **Ralph** looks

```
            S
          /   \
        /       \
     NP^p        S
               /   \
             /       \
            NP^a
```
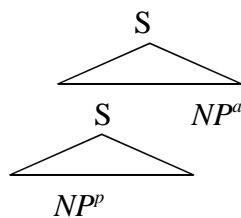
(52) NP^a precedes NP^p but NP^p is higher in the tree than NP^a:

The woman who is to marry **Ralph** will visit **him** tomorrow

```
              S
         ／￣￣＼＼
       S        NP^p
    ／￣￣＼
   NP^a
```

(53) NP^p precedes NP^a but NP^a is higher in the tree than NP^p:

The woman who is to marry **him** will visit **Ralph** tomorrow

```
              S
         ／￣￣＼＼
       S        NP^a
    ／￣￣＼
   NP^p
```

(54)   "It would seem, then, that NP^a can be used to pronominalize NP^p unless NP^p both precedes NP^a and is in a higher echelon of embedding than NP^a."   p.166

But "The constraint is too strong":

(55)a   The girl who loved **him** thinks that the woman who killed **Peter** is a fink
    b   After **she** left so hurriedly, it was discovered that Peter had insulted the man who was escorting **the pretty blond**

```
                 S
         ／￣￣￣￣＼＼
       S             S
    ／￣＼          ／￣＼
   NP^p           S
               ／￣＼
              NP^a
```
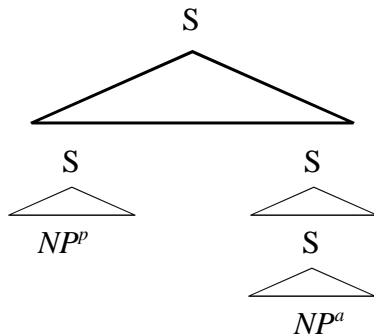
(56)   "In [(51)], the S-node that most directly dominates NP^p also dominates NP^a. In [(55)], however, the S-node most directly dominating NP^p does not dominate NP^a."   p.167
(57)   "We will say that a node A 'commands' another node B if (1) neither A nor B dominates the other; and (2) the S-node that most immediately dominates A also dominates B.

(58) NP^a **may pronominalize** NP^p **unless (1)** NP^p **precedes** NP^a**; and (2)** NP^p **commands**
         NP^a**.**

Langacker observes that the "constraint also works for reflexivization":
(59)a   **Those men** outsmarted **themselves**
    b   *__Themselves__ outsmarted **those men**
(60)   "[(59)b] is prevented because (1) NP^p (**themselves**) precedes NP^a (**those men**); and (2)
         NP^p commands NP^a (since the lowest S-node that dominates **themselves** also dominates
         those men). Thus the major constraint that must be imposed on reflexivization follows as
         a special case from the general restriction that must be imposed on pronominalization."
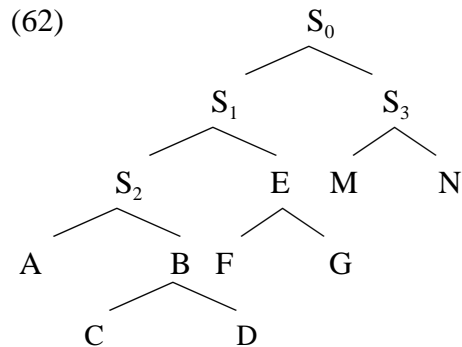   p.167

Early version of c-command - 'in construction with'

Klima (1964) proposed this notion in his investigations of scope of negation. Two classic early
discussions of anaphora - Ross (1967a) and Langacker (1969) - both explicitly argued that this is
the **wrong** relation for anaphora.

(61)   Node A of a phrase-marker is in construction with node B
         if B is dominated by the node which immediately dominates A.

<This is Ross's presentation, also given by Langacker. Klima's itself was a bit confusingly stated,
but as far as I can tell, it was actually the converse of (61). Following tradition, I will use (61).>

Ross observes that in construction with is stronger (i.e., more restrictive) than command.

(62)

(63)   "... E, F, and G command S_2, A, B, C, and D, but ... only E is in construction with these
         latter five nodes."   p.204

Ross's argument that in construction with is inappropriate for anaphora is based on (64).

(64)   I talked to Winston about himself

The source of this would be (65).

(65)   I talked to Winston about Winston
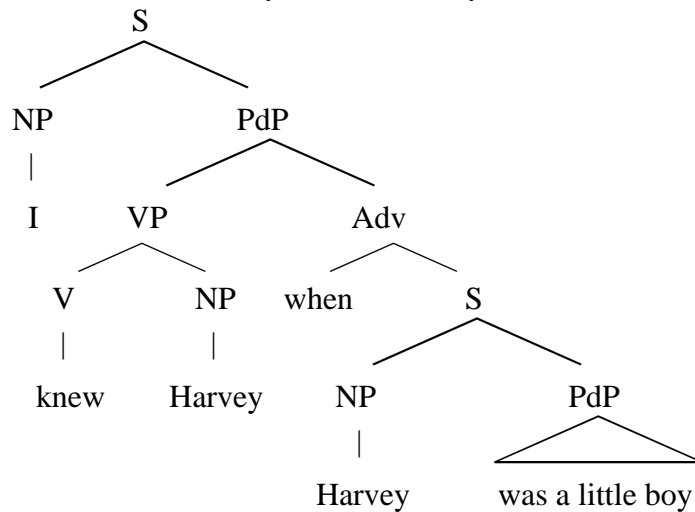But while the first *Winston* commands the second, it is not in construction with it (because of the PP dominating the first *Winston*).   <To this day, this remains a big problem in the theory of anaphora.>

Langacker's argument is that the major constraint on pronominalization cannot be correctly stated in terms of in construction with, but can in terms of command:
(66)a   I knew Harvey when he was a little boy
      b   *I knew him when Harvey was a little boy
(67)

```
                         S
              _____|_____
             |                       |
            NP                      PdP
             |                 _____|_____
             |                |             |
             I               VP            Adv
                          ___|___       ____|____
                         |       |     |         |
                         V      NP   when         S
                         |       |           _____|_____
                        knew   Harvey       |           |
                                           NP          PdP
                                            |          /\
                                         Harvey   was a little boy
```

(68)   "Since the second occurrence of **Harvey** in [(67)] follows the first but is *not* in
         construction with it, [(66)b] is allowed by [a constraint stated in terms of in construction
         with]. The first occurrence of **Harvey** both commands and precedes the second, however,
         so [(66)b] is excluded by the constraint we have adopted."   p.175

Command and clause-mates

(69)   "... if A were a member of a clause which did not
         contain B, then A would not command B, and conversely. To specify
         that two nodes command each other is to specify that each is dominated
         by the first node S above the other, and because of the formal properties
         of trees, these S nodes must be the same. That is, two nodes which
         command each other are in the same simplex sentence."   p.203

(70)   "... the English rule of Reflexivization ... can ... be formulated in terms
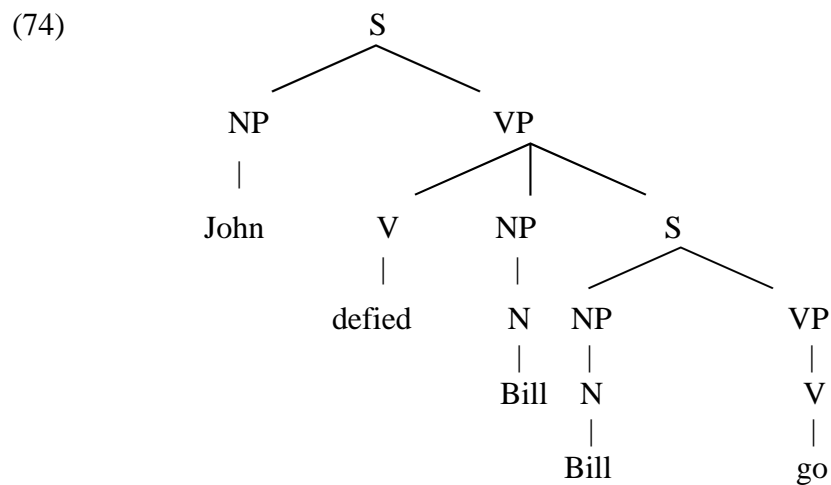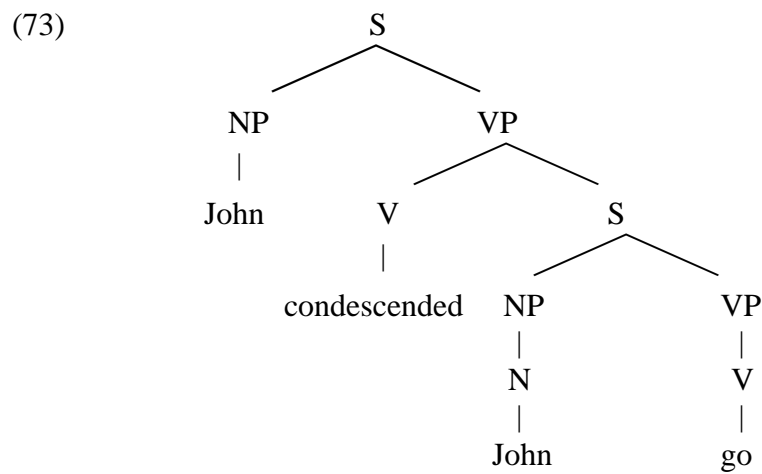   of double command."   pp.203-204

EQUI

Rosenbaum (1965) examined infinitival complements with missing subjects as in (71).

(71)a John condescended to go
   b John defied Bill to go

Rosenbaum is specifically concerned with the following:

(72)   ..."a sufficient condition for determining the deletion must specify which noun phrase in
        the main sentence must be identical to the initial noun phrase of the complement in order
        for the deletion to proceed."   p.21

(73)

```
                        S
                  _____|_____
                 NP            VP
                 |        _____|_____
                John     V             S
                         |        _____|_____
                    condescended  NP         VP
                                  |          |
                                  N          V
                                  |          |
                                 John        go
```

(74)

```
                        S
                  _____|_____
                 NP            VP
                 |        _____|_____
                John     V    NP         S
                         |    |     _____|_____
                      defied  N    NP         VP
                              |    |          |
                             Bill  N          V
                                   |          |
                                  Bill        go
```

(75)   Initial formulations of subject NP deletion rules for these two cases:

a. X NP$_1$ V [NP$_2$ VP]$_S$ Y
    1 2   3   4   5   6 $\Rightarrow$
    1 2   3 $\varnothing$ 5   6

b. X NP V NP$_1$ [NP$_2$ VP]$_S$ Y
    1 2 3 4   5   6   7 $\Rightarrow$
    1 2 3 4 $\varnothing$ 6   7

In the theory of the time, these rules couldn't be collapsed. Thus, as Rosenbaum observes, a generalization is being missed. As Rosenbaum sees it, the difficulty concerns the choice of controller of the deletion, matrix subject in (75)a, matrix object in (75)b. If that can be factored out, the separate processes can be united.

(76)   "The generalization that determines which of the two noun phrases in the
main sentence must be identical to the initial noun phrase of the complement
can be expressed in terms of a principle of minimal distance (henceforth
PMD). In the underlying phrase structure diagram [(74)], one observes that
the noun phrase in the main sentence which is relevant, i.e., the object noun
phrase, is also that noun phrase which is least distant from the initial noun
phrase of the complement. Distance here naturally can be defined in terms
of the underlying phrase structure itself by making reference to the number
of branches in the path which separates the NP nodes in the main sentence
from the initial NP node in the complement. Thus, for example, in diagram
[(74)] the number of branches separating the subject noun phrase of the main
sentence from the initial noun phrase of the complement is four. The number
of branches separating the object noun phrase of the main sentence, the
relevant noun phrase, from the initial noun phrase of the complement is three.
The principle correctly predicts that the object noun phrase of the main
sentence is the noun phrase which must be identical to the initial noun phrase
of the complement. Sentence [(71)a] reveals a special case of the principle. In
this instance, the subject noun phrase of the main sentence is the only noun
phrase in the main sentence and is, therefore, the least distant from the initial
noun phrase of the complement."   p.26

With PMD factored out, Rosenbaum states an EQUI rule covering both (75)a and b (and also covering backwards EQUI, which I have abstracted away from ("Seeing you there upset Bill"):

(77)   W (NP) X $\left\{ \begin{matrix} \text{for} \\ \text{POSS} \end{matrix} \right\}$ NP Y (NP) Z

      1   2   3   4   5   6   7   8 $\Rightarrow$

    (i) 5 is erased by 2

    (ii) 5 is erased by 7, where a constituent A is erased by a constituent B, e.g., A $\Rightarrow \varnothing$, just in
       case A and B meet the conditions imposed by the principle of minimal distance

Postal (1970) begins his discussion of EQUI with an observation important to the theory of
anaphora more generally, that there are three different types of coreference manifestation:

(78)a   Presupposed
   b   Asserted
   c   Inferred

(79)a   *Harry* analyzed *himself*
   b   *The one who analyzed Harry* was *Jack*
   c   *The boy* looked at *his male parent's only son*

"The user of [(79)a] presupposes that the NP are coreferents and asserts a certain kind of relation
between the tokens."   p.440

"In [(79)b] the user asserts an identity of reference between the NP, which are not presupposed
to be coreferential."   p.440

"Presupposed coreference involves processes which partially determine the grammatical form of
NP themselves. Assertive coreference does not. The NP in [(79)b] are not marked in any way
which indicates that they are coreferential."   p.440

The constraints and processes Postal discusses (and, in fact, the substance of all versions of BT)
are properties specifically of presupposed coreference.

(80)a   Harry wants __ to marry a millionaire
   b   Harry expects __ to win the election
   c   Harriet was unable __ to keep quiet
   d   Lucille is anxious __ to become immortal
   e   __ Finding out he had no liver didn't worry Flavnikoff
   f   __ Persuading Abe __ to live in a cave was not easy for me

The standard theory of such constructions at the time was based on work by Rosenbaum (1965),
Rosenbaum (1967). Rosenbaum had argued for a single uniform rule of deletion, one which
came to be called EQUI-NP DELETION (EQUI for short).

<The notion of coreference is tricky. It is not just, as Postal notes, that expressions like 'square circle' that don't refer in any possible world participate in it, but also that expressions that aren't even of the right type to be referential do also, as in Postal's ex.

(81)   Few criminals like being caught>

Postal argues that EQUI is (at least) a two part process, with the first part being pronominalization. He suggests a general principle from which this might follow:

(82)   Universal Deletion Constraint
        If a transformation T deletes an NP_a subject to the existence of a coreferent NP, NP_b, in the same structure, then at the point where T applies, NP_a must be pronominal.   p.489

Postal's major argument that EQUI depends on pronominalization is parallelism of limitations on the two phenomena.

(83)   "Certain otherwise ad hoc restrictions on EQUI are predictable from independently necessary constraints on PRONOMINALIZATION, if EQUI applies to structures only when they have already undergone PRONOMINALIZATION."   p.458

(84)   "... one can naturally think of deletion governed by coreference as equivalent to the existence of some general pronoun, call it Doom, which accidentally has the null phonological shape."   p.458

(85) $Doom_i$'s falling off the building injured $Harry_i$

        "The general form of our arguments showing that EQUI is dependent on PRONOMINALIZATION is ... the demonstration that ... restrictions on the distribution of *Doom* follow from restrictions on the distribution of ordinary coreferential pronouns:
        a. $*Doom_i$ ... A ... NP_i
        b. $*his_i$ ... A ... NP_i

Probably the most interesting argument is based on a phenomenon later called Weak Crossover. Postal attributes to S. Y. Kuroda the observation that backwards pronominalization is blocked in a class of case involving indefinite NP:

(86)a   If the man_i calls, you shouldn't talk to him_i
    b   If he_i calls, you shouldn't talk to the man_i
(87)a   If a (certain) man_i calls, you shouldn't talk to him_i
    b   *If he_i calls, you shouldn't talk to a (certain) man_i

To these, Postal adds

(88)   *The fact that he_i lost amused $somebody_i$ in the crowd
(89)   *The man who lost it_i needs to find $something_i$

Postal then claims that this constraint is mirrored by EQUI structures:

(90)a   *Finding out Greta was a vampire worried somebody
    b   *Discovering that their$_i$ daughters were pregnant worried some old ladies$_i$
    c   *Kissing was fun for some kids

<These data are controversial. Examples like these are often called good, part of a phenomenon called "PRO gate" by Higginbotham (1980). In a footnote, Postal acknowledges dialectal differences with respect to this patterning.>

More examples from Postal:

(91)a   *The woman who loved him$_i$ impoverished every businessman$_i$
    b   Every businessman$_i$ was impoverished by the woman who loved him$_i$
(92)a   *Finding that out worried each businessman
    b   *Discovering that God was dead upset every priest in the world
    c   *Falling off the building killed someone

Postal arrives at a multi-part analysis of EQUI, designed to capture the parallelism with pronominalization noted just above and to resolve a concomitant inconsistency with respect to cyclicity:
(93)   First, Postal argues that pronominalization is postcyclical, because it must follow WH-
        movement, which, according to Postal is postcyclical.

(94)a   *He$_i$ thinks Mary hates Bill$_i$
    b   The fact that he$_i$ lost worried Bill$_i$
    c   *He$_i$ visited some of the men who liked Bill$_i$
    d   Bill$_i$ visited some of the men who liked him$_i$

(95)a   Which of the men who liked him$_i$ did Bill$_i$ visit
    b   Which of the men who liked Bill$_i$ did he$_i$ visit

(96)   "Here, backwards pronominalization is possible out of the complex NP which has been
        fronted by WH Q MOVEMENT, just because this complex NP contains a subordinate
        clause in which the pronoun occurs. This follows from the general condition, if
        PRONOMINALIZATION is defined to apply after WH Q MOVEMENT. With the
        reverse ordering, however, the facts in sentences like [(95)] become exceptional, since
        the general condition predicts the illformedness of backwards pronominalization in
        structures like:
        [(97)]   Bill$_i$ visited wh some of the men who liked Bill$_i$"   p.454

<In a slightly different guise - 'reconstruction' - this issue is still alive 3-1/2 decades later. Much more below.>

Now, Postal's argument that WH-movement is postcyclical (the 'Preposition Orphan Argument'):

(98)   In English, a preposition may generally either be pied-piped or stranded:
(99)a   To whom were you speaking
    b   Who were you speaking to
(100)   The same is true for long distance movement:
(101)a   Who did you think Bill wanted Mary to talk to
    b   To whom did you think Bill wanted Mary to talk

(102)   "... the assumption of cyclical nature for WH Q MOVEMENT would mean that the trip
        to sentence initial position would have to be accompanied *in several jumps*, the *wh*-NP
        moving each time to the front of the next highest sentence. Under this assumption,
        however, there is no way to explain why the preposition cannot be left behind at any of
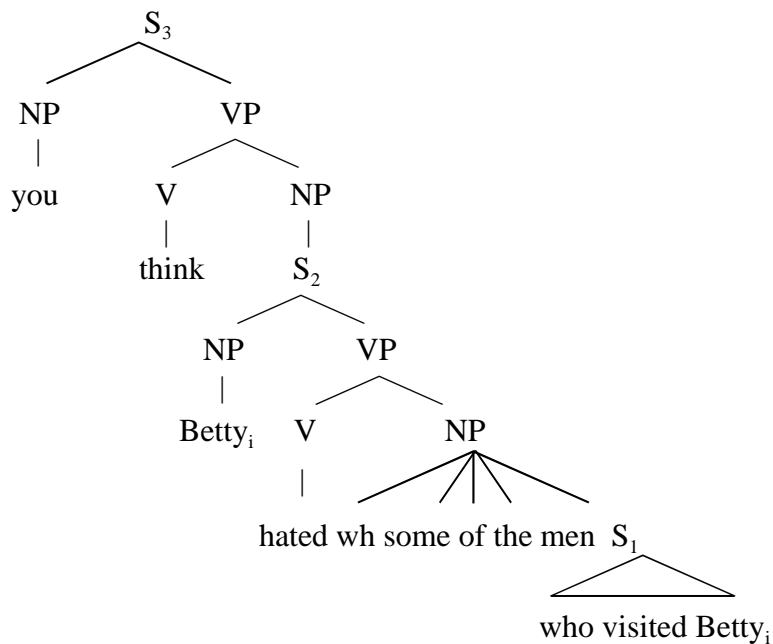   these positions."   p.455

(103)a *Who did you think Bill wanted to Mary to talk
    b   *Who did you think to Bill wanted Mary to talk
    c   *Who did to you think Bill wanted Mary to talk

(104)   "But these facts follow automatically from stating the rule as a movement of the *wh*-NP
        with optional accompaniment of the preposition, if the movement is carried out in a
        single swoop. But this can only be done if the rule is noncyclical."   p.455

<Very interesting argument, but it seems to confuse two related but distinct notions - cyclic rules
and successive cyclic rules. At any rate, see Chomsky (1973) for a response to this argument.>

Postal also offers a direct argument that pronominalization is not cyclic:

(105)

(106) Suppose pronominalization were cyclic. On the S_2 cycle, only forward pronominalization
       will be possible, and it will be obligatory. Thus, there will be no way to derive
(107) Which of the men who visited Betty$_i$ do you think she$_i$ hated

Thus, pronominalization is postcyclical. And since EQUI depends on pronominalization, it too is
postcyclical. But Postal also reports an argument, attributed to George Lakoff, that EQUI is
cyclic:
(108) "... clearly cyclical rules like RAISING and PASSIVE distort phrase structures in such a
       way that the defining configurations for the correct application of EQUI will not be met
       after these cyclical rules have applied. In short, these rules move the controller NP in
       such a way that it no longer stands in a generally specifiable relation to the NP whose
       deletion it determines."   pp.485-486

(109) Harry wants to be seen by Lucille trying to kiss Betty
(110) [$_{S4}$ Harry$_i$ want [$_{S3}$ Lucille see [$_{S2}$ Harry$_i$ try [$_{S1}$ Harry$_i$ kiss Betty]]]]
              a   b   c

(111) "Here the *c* occurrence of *Harry$_i$* must be deleted under EQUI by identity to the *b*
       occurrence. Yet this latter will undergo RAISING and PASSIVE on the S_3 cycle. After
       this, the *b* occurrence is no longer in a position to determine the deletion of the *c*
       occurrence. This naturally leads to the suggestion that EQUI apply in this case on the S_2
       cycle, before the *b* occurrence of *Harry$_i$* is raised and passivized on the S_3 cycle. Since,
       moreover, EQUI must apply again, deleting the *b* occurrence in terms of the *a*
       occurrence, which can only happen after the *b* occurrence has become a derived subject,
       the argument for cyclical application of EQUI becomes apparently overwhelming."
       p.486

Resolution of the apparent paradox:
       Lakoff's argument doesn't show that "the *deletion* must be carried out cyclically but
       rather only that those NP which are ultimately to be deleted *must be marked for this fact
       cyclically.* That is, suppose in [(109)] on the S_2 cycle the *c* occurrence of *Harry$_i$* is
       marked with some feature, let us call it [+Doom], a feature assigned by a rule to a
       particular NP, NP$_a$, just in case:
    [(112)]a NP$_a$ is the subject NP of a complement sentence S_i;
          b there is a coreferent of NP$_a$ in the immediately higher sentence S_{i+1}."   p.486

(113) DOOM MARKING (cyclical)
       WH Q MOVEMENT, WH REL MOVEMENT (last- or postcyclical)
       PRONOMINALIZATION   (last- or postcyclical)
       DOOM ERASURE

Interpretive (i.e., non-transformational) theories of pronouns, reflexives, and control

(114) Jackendoff (1972), further developing ideas of Jackendoff (1969) and Dougherty (1969), explores "the alternative theory that pronouns and reflexives are present in deep structure and that their antecedents are determined in the semantic component.
       In this approach, which I will call the interpretive theory, noun phrases are unmarked for coreference relations in deep structure. Rules of semantic interpretation establish relations between pairs of noun phrases, marking them coreferential or noncoreferential with each other."   p.108

(115) " Instead of being produced by transformations, pronouns and reflexives will be generated by the base component as lexical items, marked with the feature [+pro], but like other noun phrases, unmarked for reference. Reflexives, in addition, will be marked with the feature [+refl]."   p.108

  Motivation for an interpretive approach:

(116) "If a transformation produces pronouns, it must be able to make use of coreference relations. However, transformations generally cannot mention coreference relations: there is no rule, for example, that preposes a noun phrase if it is coreferential with some other noun phrase. Thus in stating pronominalization and reflexivization transformations that refer to intended coreference, we are implicitly granting transformations power which they do not in general possess. As part of the consistent effort to reduce the power of transformations, we will construct a theory of grammar in which no transformation can use coreference as a criterion of application."   pp.108-109

(117) Possible response to this argument: As in Chomsky (1965), factor out of any particular transformations a general condition on 'recoverability of deletion'. As Chomsky notes, "... this identity condition need never be stated in the grammar, since it is a general condition on the functioning of grammars."   p.225

(118) "A pronominalization transformation changes a fully specified noun phrase into a pronoun. This is the kind of drastic change in the form of lexical material we would like to eliminate from transformations under the constraints of the Extended Lexical Hypothesis ..."   p.109

(119) "Interpretive rules that mark coreference can as easily mark a pair of noun phrases noncoreferential as coreferential ..." [while a transformational approach provides no obvious way to do this]   p.109

<Much of modern binding theory assumes just such a parallelism between coreference and noncoreference specification.>

Another claimed advantage centers around the 'Bach-Peters Paradox' of Bach (1970), whereby in

a transformational account an infinite deep structure is apparently needed for sentences like (120).

(120)   The man who deserves it will get the prize he wants

(121)   "If the underlying structure of pronouns is a fully specified NP identical with its coreferent, both *it* and *he* in [(120)] must have infinite deep structures:

        (4.2) The man [who deserves the prize [which the man [who ...] wants]] will get the prize [which the man [who deserves the prize [which ...] wants]]"   p.110

(122)   "If the reference of pronouns is determined by a rule of semantic interpretation, the deep structure of [(120)] contains the pronouns themselves, so there is no recursion. Furthermore, in the process of semantic interpretation, a pronoun need not be replaced with a duplicate of the noun phrase with which it is coreferential (which would again bring up the problem of recursion), but rather it may just be marked coreferential with another noun phrase."   p.110

Jackendoff's final argument is based on 'anaphoric epithets' (which J. calls 'pronominal epithets'):

(123)a I wanted Charlie to help me, but the bastard wouldn't do it
      b Irving was besieged by a horde of bills and the poor guy couldn't pay them

(124)   "There are many noun phrases such as *the bum*, *the bastard*, and *the poor guy* that can be used coreferentially with another noun phrase if they are reduced in stress. These 'pronominal epithets' can occur in some subset of the environments in which pronominalization is possible, and they function semantically more or less as specialized pronouns. We would obviously miss a generalization if we did not handle them by a rule of the same kind as pronominalization, hopefully a rule that could collapse with pronominalization."   pp.110-111

(125)   "In a transformational framework, however, the generalization cannot be captured. The pronominalization transformation changes NPs into pronouns. In a consistent treatment, in certain contexts an NP could be turned into a pronominal epithet instead. But then which pronominal epithet should the NP be changed into? The meaning is obviously changed if we substitute an epithet for a pronoun or one epithet for another."   p.111

(126)   "In an interpretive framework, we can mark epithets as special lexical items which may function as pronouns in certain contexts of the pronominalization rule, adding their lexical meaning to the intended attributes of the person they refer to. Thus no change need be made in the nature of the interpretive theory in order to include these cases."

<An important question: Is it necessary to treat anaphoric epithets as pronouns to capture their interpretive properties? We will see that Jackendoff's overall theory requires this, but for reasons that will be questioned later.>

Jackendoff's theory

Three core properties of the theory:

(127) "First, coreference is an exclusively semantic property that cannot be referred to by transformations."   p.111
(128) "Second, coreference is an aspect of semantic interpretation that has nothing to do with the functional structure of the sentence."
(129) "Third, coreference is formalized in the present approach as a binary relation holding between two NPs (or their semantic readings). Three or more NPs can be understood as mutually coreferential only if they have been marked pairwise coreferential."

The technical guts of the theory:

(130) "... we will express coreference relations explicitly in a *table of coreference*. Each entry in the table will consist of a pair of noun phrases and one of the relations *coreferential* or *noncoreferential*. In a complete semantic interpretation, the table will contain an entry for each pair of noun phrases in the sentence. The table will be built up one entry at a time by the application of semantic rules of coreference."   p.111

The completed table is subject to well-formedness conditions, including:
(131) Consistency Condition
       If the table of coreference marks two NPs coreferential, those NPs must in fact be able to describe the same individual.   p.112

(132) John washed himself
(133)a John washed him
     b John washed John
     c John washed Bill

Reflexives

(134) Reflexivization, first approximation (abstracting away from environment statement)
       (α is a variable to be replaced by + or -, but both occurrences of α must be replaced with the same sign on any particular application of the rule.)

       Enter in the table:

       NP^1 α coref $\begin{bmatrix} NP^2 \\ \alpha \text{ reflexive} \end{bmatrix}$ in the environment ...
       OBLIGATORY

       "[(134)] says that in the proper contexts for reflexivization, NP^2 is coreferential with NP^1 if and only if it is reflexive."   p.112

Table entries for (132)-(133):

(135)   *John* coref *himself*
(136)a *John* -coref *him*
    b *John* -coref *John*
    c *John* -coref *Bill*

(137)   Bill told John about himself
(138)   Convention on the application of (134):
    "... if two possible environments for interpretation of a reflexive crop up at once, either reading is possible."   p.113

(139)   *The boy shot herself

(140)   *the boy*   coref   *herself*

(141)   This "can be rejected by the Consistency Condition [(131)]. *The boy* describes a male individual, *herself* a female. Obviously, then, they cannot be used to describe the same individual. Yet the table of coreference requires that they do. The conclusion must be that the sentence is deviant."   p.114

(142)   *Himself was sick
(143)   (Requirement that reflexives have antecedents)
    For every reflexive R in a sentence, there must exist an entry in the table of coreference of the form X coref R, where X is some other NP in the sentence.   p.115

(144)   *Himself shot himself
(145)   *himself* coref *himself*

(146)   "Though both NPs appear in the table, only one appears on the right-hand side of an entry, as required by the condition. The covert assumption is that there is a formal distinction between the left-hand side of the table, containing antecedents (NP^1 in the reflexive rule) and the right-hand side, containing corresponding anaphoric expressions (NP^2 in the rule)."   p.115

(147)   "Condition [(143)] differs from the Consistency Condition in that it appears to be a purely formal requirement on the table. It is not connected with conceptualization of the world since it deals only with the occurrence of the feature *refl* in the table."
    p.115

Pronouns

(148)	John looked at him	noncoreferential
(149)	John said that he was sick	ambiguous
(150)	John said that John was sick	noncoreferential
(151)	John said that Bill was sick	noncoreferential

(152)   "Pronouns are at most optionally coreferential with another NP. Any two nonpronominal noun phrases in the same sentence, morphologically identical or not, are always noncoreferential."   p.116

<This assumption about the facts is in large part responsible for Jackendoff's claim that anaphoric epithets are a kind of pronoun. However the assumption is not obviously correct, as seen in examples like
(153)   John's students admire John
More on this later.>

(154)   Pronominalization rule, first approximation
Enter in the table:

$$NP^1 \ \text{coref} \ \begin{bmatrix} NP^2 \\ + \text{pro} \end{bmatrix} \ \text{in the environment ...}$$

OPTIONAL

Crucially, in addition to this coreference rule, we need a noncoreference rule:

(155)   Noncoreferentiality rule
If for any NP^1 and NP^2 in a sentence, there is no entry in the table $NP^1 \ \pm\text{coref} \ NP^2$, enter in the table $NP^1 \ \text{-coref} \ NP^2$.
OBLIGATORY

(156)   "This rule says that any noun phrases that have not yet been related by a rule of coreference are noncoreferential. The application of [(155)] ensures that every pair of NPs appears in the table."   p.116

How the rules interact to give the right result for (148):

(157)   "In [(148)], reflexivization obligatorily applies, entering *John* -coref *him* in the table. If pronominalization were then to apply, it would enter *John* coref *him* in the table. But then the reading of the sentence would be anomalous, since the table would assert that *John* and *him* both do and do not refer to the same individual. The Consistency Condition rejects this interpretation."   p.116

Analysis of (149) ("John said that he was sick")

(158)   "In [(149)], reflexivization does not apply <Of course, this depends on the environment of the rule.> , but pronominalization may apply, entering *John* coref *him* in the table. If pronominalization does not apply, the table is empty when the noncoreferentiality rule is applicable, so the rule enters *John* -coref *him* in the table, producing the other possible reading of the sentence."   p.116

Analysis of (150) ("John said that John was sick") and (151) ("John said that Bill was sick"):

(159)   "In [(150)] and [(151)], only the noncoreferentiality rule is applicable, so the two NPs are marked distinct."   p.116

(160)   John said that Bill had shot him

(161)   Reflexivization enters *Bill* -coref *him* in the table.
(162)   Pronominalization applies between *John* and *him*.
(163)   Noncoreferentiality rule applies between *John* and *Bill*.

(164)   *Bill* -coref *him*
          John coref *him*
          John -coref *Bill*

Or, if we choose not to apply pronominalization at all, noncoreferentiality applies between *John* and *him* yielding:

(165)   *Bill*   -coref *him*
          John -coref *him*
          John -coref *Bill*

(166)   "... it should be noticed that no ordering has been established between reflexivization and pronominalization... The non-coreferentiality rule must, however, follow both of these rules."   p.117

Environment for Pronominalization

The transformational predecessors of Jackendoff's interpretive rules were essentially correct about the structural environments of the rules (and, in fact, the ordering of the rules vis à vis movement transformations). So the interpretive rules could be ordered just where the transformations had been, and the environment of the transformations is replicated. As for ordering, Jackendoff ultimately concludes (for theory internal reasons (pp.129-131)) that Pronominalization is cyclic, following all transformations on each cycle <a very important foreshadowing of Multiple Spell Out>, and that the Noncoreferentiality rule is last cyclic, following all transformations on the final cycle of the derivation.
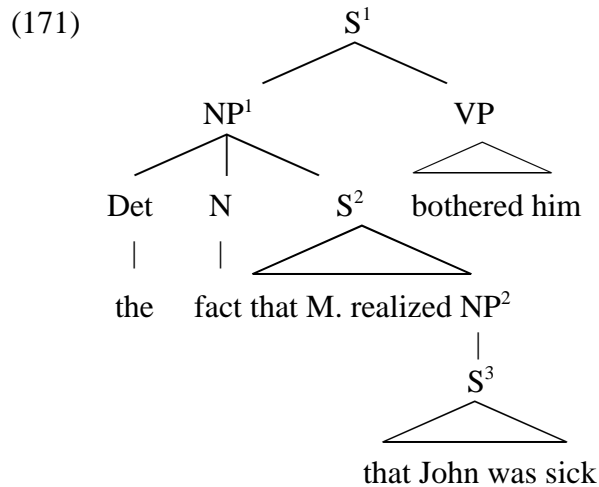
(167)   Pronominalization rule
        Enter in the table:

        NP^1 coref $\begin{bmatrix} NP^2 \\ + pro \end{bmatrix}$ unless NP^2 both precedes and commands NP^1

        OPTIONAL

(168)a   Jake left town after he robbed the bank
     b *He left town after Jake robbed the bank
     c   After Jake robbed the bank, he left town
     d   After he robbed the bank, Jake left town

(169)   "With the antecedent to the left ('forward pronominalization'), pronominalization is
        generally possible. With the antecedent to the right ('backward pronominalization'), the
        pronoun must not command the antecedent."

(170)   The fact that Mary realized John was sick bothered him

(171)

(172)   The fact that he realized John was sick bothered him

(173)

```
                        S¹
                ╱            ╲
            NP¹               VP
        ╱    |    ╲          ╱  ╲
     Det    N     S²      bothered him
      |      |    ╱  ╲
     the   fact that he realized NP²
                              |
                              S³
                            ╱   ╲
                    that John was sick
```
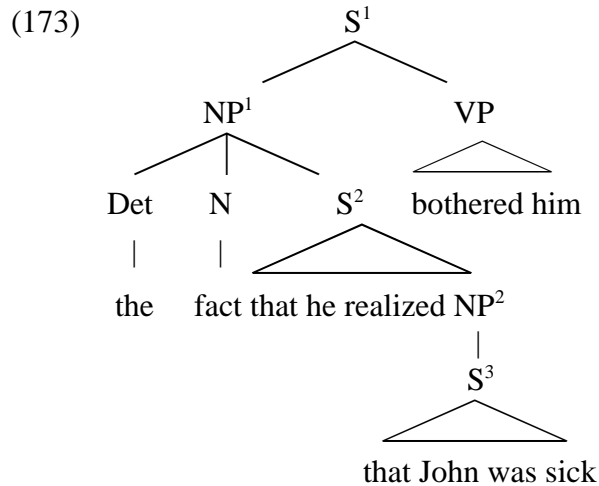
According to Jackendoff, (173) is surprising, since, when *he* and *him* are coreferential, forward pronominalization (between *John* and *him*) is evidently bad. Here is his account:

(174) Pronominalization can apply between *he* and *him* and also between *John* and *him*.

(175) Pronominalization cannot apply between *he* and *John*, since *he* precedes and commands *John*.

(176) Thus, the noncoreferentiality rule will apply to *he* and *John*.

(177) *he*   coref   *him*
       John coref   *him*
       he   -coref *John*

(178) "... this table is inconsistent: it asserts that *him* is the same individual as two distinct individuals. Therefore the reading must be rejected.

(179) This leaves open only the possibility of marking *him* coreferential with one or the other of the other two NPs, letting the noncoreferentiality rule to [sic] mark it noncoreferential from the other."   p.120

The last cyclic nature of the noncoreferentiality rule (155):

(180) Who that Mary knew do you think she visited

(181) [$_{S1}$ you think [$_{S2}$ she visited who [$_{S3}$ that Mary knew]]]

(182) "In [(181)], the environment for pronominalization does not appear on the S^2 cycle. However, after fronting on the S^1 cycle to form (180), the pronoun and NP can be marked coreferential.

(183) If [the noncoreferentiality rule (155)] were in the cycle ... it would apply obligatorily on the S^2 cycle of [(181)], to mark *she* and *Mary* distinct, since that pair is not yet entered in the table of coreference... But if we make the noncoreferentiality rule last-cyclic, pronominalization will have a second chance to apply to *she* and *Mary*, on the S^1 cycle, giving the correct reading.

(184)   Only if pronominalization does not apply at any point in the derivation ... do we want the conditions for the noncoreferentiality rule to be met." p.130

Environment for Reflexivization

Jackendoff's statement incorporates a version of the clause-mate and command requirements. Note that Jackendoff, following Chomsky (1970), takes S and NP to be the cyclic nodes.

(185)   Reflexivization

$$NP^1 \; \alpha \; coref \begin{bmatrix} NP^2 \\ \alpha \; reflexive \end{bmatrix}$$

if each of the following conditions holds:
(a) NP^2 has not yet appeared on the right-hand (i.e., anaphoric expression) side of the table;
(b) NP^2 is immediately dominated (except for a possible preposition [sic]) by VP or N';
(c) NP^1 is in the main clause of the present cycle;
(d) NP^2 does not both precede and command NP^1. p.136

(186)   *John thinks that himself is sick   [Out by (b), plus noncoref. rule. Condition (b) is necessary since J. has given up the simplex sentence condition.]
(187)a *The book which I gave John impressed himself
     b *I gave a book about John to himself   [These are out by (c) plus noncoref. rule.]

Thematic Hierarchy Condition

(188)   Thematic Hierarchy
        1. Agent
        2. Location, Source, Goal
        3. Theme

(189)   Thematic Hierarchy Condition on Reflexives
        A reflexive may not be higher on the Thematic Hierarchy than its antecedent
                                                                    p.148
(190)   *John was shaved by himself

John is Theme and *himself* is Agent, so THC is violated.

(191)   *Himself was shaved by John

THC is satisfied but the example "violates the structural condition on reflexivization, so it is much more crashingly ungrammatical."

(192)   John was touching the bookcase
John is ambiguous between Agent and Theme. *The bookcase* is Location (or possibly Goal).

(193)   John was touching himself

John is only Agent.   p.149

(194)   *John was being touched by himself

Satisfies THC on reflexivization, but violates another THC: In passives, the by-phrase must be higher on the hierarchy than the derived subject.

(195)   Bill strikes Harry as pompous

Bill is Theme and *Harry* is Goal.

(196)   ?I strike myself as pompous

Out by THC.

(197)   Bill regards Harry as pompous

Bill is Goal and *Harry* is Theme.

(198)   I regard myself as pompous

(199)   ?John appears to himself to have hit Bill

J. wants this to violate THC, even though *John* is an Agent and *himself* a Goal.

(200)   "It is reasonable that a raised NP such as the derived subject of *seem* [or *appear*] has no
        thematic relation at all with respect to its new clause, since thematic relations are related
        to deep structure grammatical relations ... Then, to prevent [(199)], we need only assume
        that lack of thematic relation counts as the lowest position on the Thematic Hierarchy."
                                                                                    p.154

(201)   John and Bill like each other
(202)   *John and Bill thought that Mary liked each other

(203)   Each other is a reflexive pronoun   p.170

(204)   John and Bill told me stories about themselves/each other
(205)   *John and Bill told me Mary's stories about themselves/each other

(206)   (?)John and Bill strike each other as pompous

Control

(207)a Bill hoped to be able to leave soon
      b Bill hoped for Sam to be able to leave soon
      c Bill hoped for Bill to be able to leave soon
      d Bill hoped for him to be able to leave soon

(208)   "We observe that only when the complement subject is absent on the surface is it interpreted as coreferential with the subject of the main clause... The rule apparently looks rather like reflexivization."   p.180

(209)   The complement subject "cannot be any normal pronoun or NP, as we can see from [(207)a-d]. One solution might involve a special pro-NP (call it DEL) which has no phonological interpretation and serves the sole semantic function of acting as a coreferential subject complement [sic]."   p.180

(210)   Complement Subject Rule, first approximation
        Enter in the table:
        NP^1 α coref NP^2 if NP^2 is the subject of a *for-to* or *poss-ing* complement, NP^1 is in the main clause of the present cycle, and NP^2 is alpha equal to DEL.
        OBLIGATORY

(211)   Additional well-formedness condition: Reject any reading with an uninterpreted DEL.

(212)   Bill hoped [for DEL to be able to leave soon]
        Cycle 1: nothing happens
        Cycle 2: (210) enters *Bill* coref *DEL* in the table

(213)   For all of (207)b-d, $NP^2 \neq$ DEL.

Problem with (210): "it requires a special pro-NP which is required nowhere else in the grammar
        ..."
                                                                p.181
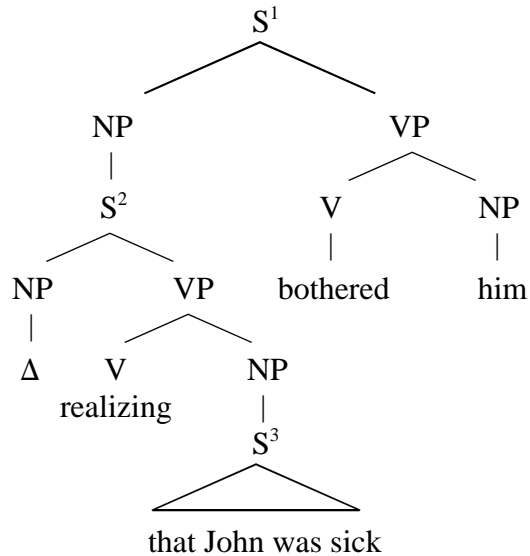Alternative (attributed by J. to Chomsky):
(214)   "Suppose that the phrase structure rules and lexical insertion are optional, so that potential deep structures can be generated containing nonterminal nodes ... symbolized by Δ ... these structures would be semantically ill-formed, since no semantic information would be available for empty nodes...   <a precursor of Full Interpretation?>
(215)   Suppose that some rules of semantic interpretation give readings to empty nodes ... under certain conditions. Then the semantic blocking of an empty node can be prevented, just in case it is interpreted by one of these rules."   p.181

Jackendoff's treatment of the famous example of Ross (1967b), where forward pronominalization seems to be blocked:

(216)   Realizing that John was sick bothered him   <Ross's actual ex. is "Realizing that Oscar was unpopular didn't disturb him".>

(217)

S^1
  NP   VP
  S^2 V   NP
  NP   VP   bothered him
  Δ   V   NP
      realizing S^3
             that John was sick   p.186

(218) -S^3 cycle: nothing
       -S^2 cycle: pronominalization can't apply between Δ and *John*, so they are eventually
       marked distinct by the noncoreferentiality rule (on the final cycle)
       -S^1 cycle: complement subject rule applies marking Δ coreferential with *him*
(219) If pronominalization also applies between *John* and *him*, a paradoxical reading results:
       John   coref *him*
       him   coref Δ
       Δ   -coref *John*
(220) The only possibility, then, is not to apply pronominalization between *John* and *him*.

Wasow (1972) - a (mostly) interpretive theory

For reflexives, Wasow adopts the transformational approach of Helke (1971):

(221) "... reflexives consist of a head noun <u>self</u> with a possessive pronominal determiner. The
       necessary agreement between this determiner and some other NP in the same clause is
       accounted for by a transformation which inserts pronominal copies of the NP's into empty
       determiner nodes."   Wasow p.17
(222) "The constraint that reflexives and their antecedents be in the same simplex sentence ... is
       accounted for by Chomsky's insertion prohibition (Chomsky (1965, p.146 <(28)
       above>)"   p.17

All the rest of Wasow's theory is interpretive. In particular
(223) "... there is no transformation of Pronominalization which replaces an NP by a definite
       pronoun when the NP meets certain conditions of identity with another NP."   p.21

Wasow's initial statement of the conditions on definite NP anaphora:

(224)   Given an NP and a definite pronoun in the same sentence, the NP may serve as the
        antecedent for the pronoun unless
        (a) the pronoun and the NP disagree in gender, person, and number;
        (b) the pronoun is to the left of the NP and the pronoun is not more deeply embedded
        than the NP   p.61
(225)   "...the best criterion for deciding the theoretical status of [(224)] is whether it obeys the
        various constraints which have been proposed for the different categories of rules.
        Judging by this criterion, anaphora rules seem to belong to none of the usual categories,
        for they obey none of the constraints on syntactic and semantic rules proposed in Ross
        (1967a) and [Chomsky (1973)]."   pp.79-80
(226)   "This conclusion is supported by the fact that anaphoric relations may hold between
        elements of different sentences. In this respect anaphora differs from all other syntactic
        and semantic pheonomena [sic] which have been studied by generative grammarians...
        One rather attractive   suggestion that has been made with respect to the theoretical status
        of anaphora rules is that they are not rules of any particular grammar, but part of
        linguistic theory... Although little work has been done on anaphora in languages other
        than English, the fact that no grammar book of any language ever includes instructions
        for associating pronouns and antecedents suggests that people learning foreign languages
        do not have much trouble with this aspect, which in turn suggests that the rule involved
        may be universal. (This was pointed out by Witten (1970))."   p.80

Strong Crossover

(227)a   <u>Who</u> said Mary kissed <u>him</u>
     b *<u>Who</u> did <u>he</u> say Mary kissed
(228)a   The man <u>who</u> said Mary kissed <u>him</u> was lying
     b *The man <u>who</u> <u>he</u> said Mary kissed was lying

We can't handle this by ordering the pronominal anaphora rule (224) before WH-movement,
because of:
(229)a *<u>He</u> finally married one of the girls <u>Bill</u> had been dating
     b   Which of the girls <u>Bill</u> had been dating did <u>he</u> finally marry   p.138

Wasow's solution (with the core idea -trace theory - attributed to Peter Culicover, and also to the
1971 manuscript version of Chomsky (1973)):
(230)   "WH-fronting could be formulated so that a phonetically null copy of the WH-word is
        left behind in its pre-fronting position. By the arguments of Chapter 1 §3.3, this copy
        would necessarily be anaphorically related to the WH-word. Hence, if the WH-word
        itself enters into an anaphoric relation with a pronoun, then by the Transitivity Condition
        (Chapter 1 ,§§3.2.1. & 3.2.2), the copy and the pronoun must enter into an anaphoric
        relation. If this relation is blocked, then the sentence is deviant. It therefore follows that
        WH-words' behavior with respect to anaphora should be just what it would be if they
        were in their pre-fronting positions."   p.139

Transitivity condition
(231)   If A, B, and C are three elements in a sentence such that an anaphoric relation holds
        between A and B and an anaphoric relation holds between B and C, then the sentence is
        marked ungrammatical unless an anaphoric relation holds between A and C.

(232)a   He said [Mary kissed someone]
      b Who did he say [Mary kissed Δ]

(233)   "Now, if <u>who</u> and <u>he</u> in [(232)b] are to be allowed to enter into an anaphoric relation, the
        Transitivity Condition requires that Δ and <u>he</u> also be anaphorically related. Consequently,
        the resultant sentence will be ungrammatical for the same reason that [(232)a] is."   p.139

Wasow acknowledges a potential problem, then proposes a solution that will seem familiar:

(234)   "... if , as the previous chapter suggested, empty nodes are to be treated by the anaphora
        rules just like pro-forms, then there is no reason why <u>he</u> may not serve as the antecedent
        for Δ, in which case the relevant anaphoric relations would always be possible ... So
        some means must be found to keep <u>he</u> from serving as the antecedent of Δ.
            This can be accomplished formally simply by marking the traces left behind by
                WH-fronting with some feature (e.g., [-pro]) which distinguishes them from
                pronouns."   pp.139-140

 "Conditions on Transformations" Chomsky (1973)

In this work, Chomsky developed a set of conditions on movement, and used them, directly or
indirectly, to handle restrictions (of the clause-mate sort) on anaphoric processes.

(235)   The dog is believed [__ to be hungry]
(236)   *The dog is believed [__ is hungry]

(237)   Tensed Sentence Condition (1st version)
                No rule can involve X, Y in the structure
                $...X...[_{\alpha}...Y...]...$
                where α is a tensed sentence

Reciprocals were treated with an each-movement transformation, a partial retreat from the
'interpretive' account of Jackendoff:

(238)a   The candidates each hated the other(s)
      b The candidates each expected [the other(s) to win]
      c The candidates each expected [that the other(s) would win]
(239)a   The candidates hated each other
      b The candidates expected [each other to win]
      c *The candidates expected [that each other would win]

(240)   Passive (i.e., NP preposing) and <u>each</u>-movement obey (237).

(241)   "...one rule that obviously does not satisfy the condition is Coreference Assignment..." as
        in (242).   p.228
(242)   John said that he would leave

<Much more on this shortly>

(243)   The candidates expected [__ to defeat each other]
(244)   *The men expected [the soldier to shoot each other]

(245)   Specified Subject Condition (1st version)
                No rule can involve X, Y in the structure
                ...X...[$_\alpha$...Z...-WYV...]...
                where Z is the specified subject of WYV in α
(246)   'Specified subject': "a subject NP that contains either lexical items or a pronoun that is not
        anaphoric"

(247)   In (244), *the soldier* is Z, blocking each-movement from *the candidates* (X) to *the
        other(s)* (Y).

(248)a   The men saw [$_{NP}$ pictures of each other]
     b *The men saw [$_{NP}$ John's pictures of each other]

(249)   NP counts as a relevant α (as already hinted by Jackendoff).

(250)a *I saw me
     b *I saw us
     c *We saw me
     d He saw him
     e The soldiers shot the officers (among them)

(251)   RI: "a rule of interpretation applying to the structure NP-V-NP (among others) [that]
        seeks to interpret the two NPs as nonintersecting in reference, and where this is impossi-
        ble... it assigns 'strangeness'."   p.241
<Chomsky correctly relates this to Postal's TIC mentioned above. Chomsky, though, calls
Postal's condition the "Unlike Person Constraint", a name that Postal never used as far as I can
tell>.

(252)a   We expect [them to visit me]
     b *We expect [me to visit them]
     c We believe [I may still win]

(253)   In (252)a, c, RI is blocked by SSC and TSC, respectively.

(254)   No rule can involve X, Y in the structure
$$...X...[_\alpha...Z...\text{-WYV}...]...$$
where (a) Z is the specified subject of WYV
    or (b) Y is in COMP and X is not in COMP
    or (c) Y is not in COMP and α is a tensed S

(255)   "...under the analysis proposed here there is no necessity for a rule raising the subject of
        an embedded sentence to the object position of the matrix sentence..."
<Only a finite clause, not just any clause, blocks RI, passive, and each-movement.>

(256)a *We persuaded Bill [PRO to kill each other]
     b We promised Bill [PRO to kill each other]

(257)a I (we) persuaded Bill [PRO to kill us]
     b *I (we) promised Bill [PRO to kill us]

Sometimes a null subject does count as a Z, blocking relations between X and Y.

(258)   Z is a specified subject with respect to X if it is not 'controlled' by (a category containing)
        X. (If Z is lexically specified, it is not controlled at all. PRO is controlled in the standard
        sense. Trace is controlled by its antecedent.)

(259)a *They appealed to John [PRO to like each other]
     b They appeared to John [$\underline{t}$ to like each other]   <$\underline{t}$ is trace, used by Chomsky for the
                                                                             first time here.>

(260)a We appealed to John [PRO to like us]
     b *We appeared to John [$\underline{t}$ to like us]

Given the generality of (254) (No rule can ...) the question arises of how long distance movement
is ever possible.

(261)a Who do you believe [John to have seen t]
     b Who do you believe [t saw Mary]
   Wh-Movement is argued here for the first time to be 'successive cyclic', involving movement
into the most local clause initial position ('COMP'), then from there into the next higher COMP,
and so on up the phrase structure tree. In addition, S', and not S, is taken to be the relevant
sentential domain. In both (261)a and b, then, the initial movement is into the lower COMP. This
correctly renders SSC ineffectual for (261)a. The first movement is entirely inside α. For the
second movement, from the lower to the higher COMP, movement is from a position outside the
domain of the lower subject to a position that is still inside the only potential higher α.
   Example (261)b raises more difficult problems. The initial movement, that into the lower
COMP, is once again in accord with the conditions. The movement into the matrix COMP would
seem to be in violation of TSC, however. The complement S', being tensed, is a relevant α.
Chomsky winds up simply stipulating that Wh-Movement from COMP (and to another COMP)

is exempt from the TSC: (254)b,c.

Thus, when Y is in COMP, TSC is void. Note that this exemption is not available for (248)b, since there is no intermediate COMP for Y, NP differing from S' in that regard. The necessity of including the complication in (254) suggests that (254) is irrelevant to Wh-Movement entirely. Freidin and Lasnik (1981) argue at some length that this is, in fact, the case <more on this below>. This would leave (248)b unexplained, but, as Chomsky observes, entirely apart from the SSC, definiteness (or perhaps specificity) is relevant to Wh extraction:

(262)   ?*Who did you see [the pictures of t]

Interestingly, one of the prior and subsequent major cases of anaphora, reflexivization, is almost entirely ignored. The only mentions (pp. 241 and 251) are the hint that the copying approach of Helke (1971) "...fits very well into the present framework."

Lasnik (1976)

(263)a   Oscar finally realized that he is unpopular
     b *Oscar finally realized that Oscar is unpopular
     c *He finally realized that Oscar is unpopular
     d   He finally realized that Mary is unpopular

(264)   An obligatory pronominalization transformation can rule out (263)b, but says nothing
        about (263)c, given that He is perfectly capable of referring to Oscar even when no
        pronominalization T is relevant, as in (263)d.

New empirical claim: In situations where backwards pronominalization is possible, so is forwards pronominalization, and further pronominalization need not obtain at all:

(265)   That Oscar/he is unpopular was finally realized by Oscar

(266)   Noncoreference rule - 1st approximation:
        If NP_1 precedes and commands NP_2, and NP_2 is not a pronoun, then NP_1 and NP_2 are
        noncoreferential. [A commands B if the minimal S node dominating A also dominates
        B.]   p.6

<Notice that the indices in the noncoreference rule are not 'referential indices'. Rather , they are just for ease of exposition. (266), like Chomsky's RI (251), makes no use of referential indices, in strong contrast to some of Chomsky's later (and earlier) theories.>

<See also Culicover (1976) for a coreference restriction based on precedence and Klima's 'in construction with'.>

(267)   Compare 'RI' of Chomsky (1973), which, "applying to the structure NP-V-NP (among
        others) seeks to interpret the   two NPs as nonintersecting in reference..."

(268) Recall the Jackendoff (1972) theory of anaphora:
    -Optionally label NP^1 and NP^2 [+Pro] <u>+coref</u> unless NP^2 both precedes and commands NP^1.
    -Obligatorily label any two NPs <u>-coref</u> if they have not been labelled <u>+coref</u>.

(269) The coreference rule in (268) is evidently superfluous for cases where it is empirically correct. Further, it incorrectly bans coreference between any two NPs if neither is a pronoun.

(270)   After I hit <u>John</u>, <u>the sissy</u> screamed

(271)a *<u>John</u> realizes that <u>the sissy</u> is going to lose
    b *<u>He</u> realizes that <u>the sissy</u> is going to lose
    c *<u>The sissy</u> realizes that <u>the sissy</u> is going to lose

(272) Anaphoric epithets behave like names rather than like pronouns.

"Split antecedence"

(273) After John talked to Mary, they left the room

(274) The problem of 'split antecedents' disappears (at least for referential pronouns): <u>they</u> is free to refer to any plurality at all, as long as (266) is satisfied.

Still problematic though are split antecedent examples with quantificational antecedents, as in (275) from Lasnik (1991).

(275) Every violinist told some pianist that they should play a duet.

Transitivity effects

(276)   *The woman <u>he</u> loved told <u>him</u> that <u>John</u> was a jerk

(277) (266) subsumes the effects (for NPs) of Jackendoff's (1972) Transitivity Condition, adopted by Wasow (1972) in the following form: If A, B, and C are three elements in a sentence such that an anaphoric relation holds between A and B and an anaphoric relation holds between B and C, then the sentence is marked ungrammatical unless an anaphoric relation holds between A and C.

Command $\rightarrow$ kommand

(278)   <u>His</u>/<u>John's</u> mother loves <u>John</u>

(279)   *<u>He</u>/<u>John</u> loves <u>John's</u> mother

(280) Noncoreference rule - 2nd approximation:
       If NP_1 precedes and kommands NP_2, and NP_2 is not a pronoun, then NP_1 and NP_2 are noncoreferential.   p.15
       [A <u>kommands</u> B if the minimal cyclic node dominating A also dominates B.]

Inclusion effects

(281) They assume that Bob will talk to Tom

(282) Their parents told Mary to play with Susan

(283) Noncoreference rule - Final: If NP_1 precedes and kommands NP_2, and NP_2 is not a pronoun, then NP_1 and NP_2 are disjoint in reference.   p.16

Pronouns as bound variables

(284)   Everyone believes that he is unattractive

(285) *Everyone's mother believes that he is unattractive

(286)   Everyone sat down after he walked in

(287) *After he walked in, everyone sat down

(288) A pronoun can function as a variable bound by a quantifier only if the quantifier both precedes and kommands the pronoun.   p.18

(289)   Who thinks he can solve the problem

(290) *Who does he think can solve the problem

(291) The scope of a Wh operator is assigned cyclically. Thus, <u>he</u> cannot be in the scope of <u>Who</u> in (290).

(292) Harry believes he is intelligent, and Bill does too   [Sloppy OK]

(293) Harry's mother believes he is intelligent, and Bill's mother does too   [Sloppy *]

(294) Deletion under sloppy identity is only possible when the antecedent of the deleted pronoun both precedes and kommands the pronoun.   p.20

(295) 'Sloppy' identity is really full identity. The deleted VP contains a bound variable, as does the antecedent VP.

"Conditions on Rules of Grammar" (1976)

In Chomsky (1976), we find one rather significant departure from Chomsky (1973), namely, the conjecture that TSC and SSC are solely constraints on anaphoric relations of a particular type, rather than (also) on the operation of transformations. Reciprocals are now treated in terms of a reciprocal rule which 'assigns an appropriate sense to sentences of the form NP . . . each other',

as suggested in Jackendoff (1972) and developed in Fiengo and Lasnik (1973). Further, it is proposed that the relation between an NP that has undergone NP Movement and the position from which it moved is one of bound anaphora, as proposed by Fiengo (1974). As such, it would naturally fall under the same conditions as reciprocal interpretation. Chomsky thus tentatively concludes that "we can regard SSC, in such cases, as a condition on surface structures applying quite generally to anaphora (hence to the NP-trace relation), rather than a condition on transformations".

(296)a The men like each other
    b *The men want [John to like each other]

(297) Reciprocal interpretation assigns an appropriate sense to sentences of the form NP...<u>each other</u> (and is constrained by the conditions).

(298)a The men like them
    b The men want [John to like them]

(299) Disjoint reference (DR) assigns disjoint reference to a pair (NP, pronoun) (and is constrained by the conditions).

(300)a   John seems [<u>t</u> to like Bill]
    b *John seems [Bill to like <u>t</u>]

(301) "...the relation between NP and the trace that it controls [is] a special case of bound anaphora..." That relation is constrained by the conditions. The conditions thus are conditions on surface structure applying to anaphora. (I have illustrated SSC. TSC is the same.)

For examples like (302)-(303) , Chomsky continued to assume the (1973) definition of 'specified subject'. In retrospect, it is clear that such a complication was unnecessary. Rather, instead of the transparent subjects **not** counting as Z, they **would** count as X. That is, the PRO or trace could be the relevant antecedent.

(302)a *They appealed to John [PRO to like each other]
    b   They appeared to John [<u>t</u> to like each other]

(303)a   We appealed to John [PRO to like us]
    b *We appeared to John [<u>t</u> to like us]

(304) The rules of anaphora relate surface structures (enriched to include traces) to LF. Perhaps more generally, surface structure determines LF.

(305) John thought that Bill liked him

(306) (305) is not a problem as it does not involve a rule of sentence grammar at all.

(307) The problem, of course, is "He thought that Bill liked John", which crucially does involve
        a rule of sentence grammar, as evidenced immediately below.

(308)a Who said Mary kissed him
     b Who said he kissed Mary
     c Who did he say Mary kissed (Wasow's 'Strong Crossover')

(309)a John said Mary kissed him
     b John said he kissed Mary
     c He said Mary kissed John

(310) for which person x, he said Mary kissed x

(311) Taking a variable to function as a name, (308)c then reduces to (309)c. But then we
        clearly are dealing with a rule of sentence grammar - the one barring anaphoric
        connection here. And this rule does not obey the conditions.

Reinhart (1976)   c-command

(312) "The restrictions on anaphora must ... specify the structural conditions under which two
        given NP's cannot have an anaphoric relation (or corefer). The relation of precede-and-
        command ... has been supposed to capture these conditions. I will argue, however, ... that
        this relation plays no role in determining anaphora options and has, probably, no other
        linguistic relevance."   p.7

(313) "My concern in the discussion to follow is, primarily, to define the structural relations
        which restrict anaphora options. However, a proper specification of these structural
        relations has much broader consequences: I will argue that these relations define a more
        general notion of the syntactic domain of a given node and that anaphora rules are
        restricted to operate on two given NP's just in case one of these NP's is in the domain of
        the other."   p.7

(314) "Lasnik's observation enables us to restate the problem of coreference restriction: the
        problem is not specifying the conditions under which a pronoun can precede its
        antecedent or under which an NP cannot be coreferential with a pronoun to its left (as all
        previous discussions assumed), rather, specifying the conditions under which coreference
        between two definite NP's is not free.
(315)   Stated in terms of syntactic domains, the resulting picture is that given two NP's, the
        crucial question is whether either of them is in the domain of the other. If this is not the
        case, there are no restrictions on the coreference options of these two NP's - they can be
        coreferential (or noncoreferential) regardless of whether they are pronouns or full NP's,
        and regardless of their linear order, in case only one of them is a pronoun. But, if one of
        the two NP's in question is in the domain of the other, the coreference restriction permits
        coreference only if this NP is a pronoun, and again, regardless of whether the other NP,
        which is the head of the domain, is a pronoun or not."   p.17
(316) "... the terms antecedent and anaphor are superfluous, As we have seen, the problem is

not defining the structural relations between antecedents and anaphors, but rather defining the structural conditions which affect the coreference options of any two NP's." pp.50-51, fn.8

A conceptual problem with traditional precede and command:

(317) "One of the problems with the syntactic domain defined in terms of precede-and-command ... is that it makes very little sense, since the domains produced are quite arbitrary. If the domain of a node A is everything to its right which is commanded by it, then given a sentence S, we can cut at any arbitrary node and consider it together with all the nodes to its right which are dominated by S as constituting a domain.

(318) Given a sentence like <u>Ben introduced Max to Rosa in September</u>, for instance, the domains picked out are, first, the domain of the subject (the whole sentence) and the domain of the verb (the VP, including the PPs). Since these domains are also constituents, their linguistic relevance is obvious. However, the same definition also yields the domains [Max / to Rosa in September], [to Rosa / in September], and [Rosa / in September]. In the same way, given a sentence like <u>The cover of the book is lost</u> , [the book / is lost] is one of the domains produced. If such arbitrary chunks of the tree constitute a syntactic domain, it is hard to see what content the notion of domain could have."   p.22

<Note that the last mentioned of the problems goes away once command is extended from S to all cyclic nodes. Note too that the remaining problems persist under current views of binary branching.>

Empirical problems (some of them due to Lakoff (1968)):

Unexpected coreference

(319)a Near <u>him</u>, <u>Dan</u> saw a snake.
     b In <u>her</u> bed, <u>Zelda</u> spent her sweetest hours.
     c For <u>his</u> wife, <u>Ben</u> would give his life.
     d How obnoxious to <u>his</u> friends <u>Ben</u> is.
     e Fond of <u>his</u> wife though <u>Ben</u> is, I like her even more.
     f(I predicted that Rosa would quit her job and) quit <u>her</u> job <u>Rosa</u> finally did.

Unexpected noncoreference

(320)a *Near <u>Dan</u>, <u>he</u> saw a snake.
     b *In <u>Zelda's</u> bed, <u>she</u> spent her sweetest hours.
     c *For <u>Ben's</u> wife, <u>he</u> would give his life.
     d *How obnoxious to <u>Ben's</u> friends <u>he</u> is.
     e *Fond of <u>Ben's</u> wife though <u>he</u> is, I like her even more.
     f *(I predicted that Rosa would quit her job and) quit <u>Rosa's</u> job <u>she</u> finally did.

In all of (319), the pronoun precedes and commands the name, yet coreference is still possible. (Note, though, that in (319)b,c,e kommand does not obtain.) And in all of (320), the pronoun follows the name, yet coreference is not possible.

Reinhart rejects a rule ordering account of these facts (what would be essentially a 'reconstruction' analysis) based in part on examples like these (repeated from above):

(321)a   Jake left town after he robbed the bank
     b *He left town after Jake robbed the bank
     c   After Jake robbed the bank, he left town
     d   After he robbed the bank, Jake left town

Assuming (b) to be basic, ordering the coreference condition before preposing would incorrectly rule out (c). Similarly for:

(322)   Who that Mary knew do you think she visited

(323)   All these kinds of cases "... suggest that there must be some structural properties of the surface structure of these sentences which determine their coreference options. These properties cannot be captured by the relation of precede-and-command ... "   p.26

Precede and command replaced by c-command

(324)   C-command (1st version)   <The name attributed by Reinhart to Nick Clements>
        Node A c(onstituent)-commands node B if neither A nor B dominates the other and the first branching node which dominates A dominates B.

(325)   "The linguistic need for a relation like [(324)] has been observed before for problems other than coreference: [(324)] is the converse of the relation in construction with which was suggested by Klima (1964) to account for the scope of negation.
(326)   (In [(324)], node B is in construction with node A.)
(327)   It is also very close to the relation superiority suggested by Chomsky (1973), the difference being that superiority is asymmetric - nodes A and B cannot be superior to each other. Thus, sister nodes are excluded from the superiority relation, while definition [(324)] includes sisters (i.e., nodes that c-command each other).
(328)   The difference between the relations of command and of c-command is that while the first mentions cyclic nodes, the second does not - all branching nodes can be relevant to the determination of c-command relations."   p.32

(329)

```
                cy₁
              / | \
            A   B   C
               / \
              D   cy₂
                 / \
                E   F
```
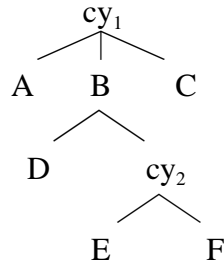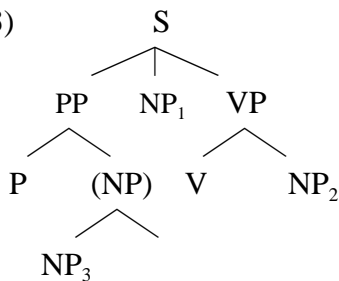
(330) "... node A both commands and c-commands all the other nodes in [(329)] - as is also true for node C - but node D, while commanding node C (since it is dominated by cy_1 which dominates C) does not c-command node C, since the first branching node dominating D, namely B, does not dominate C."   p.32
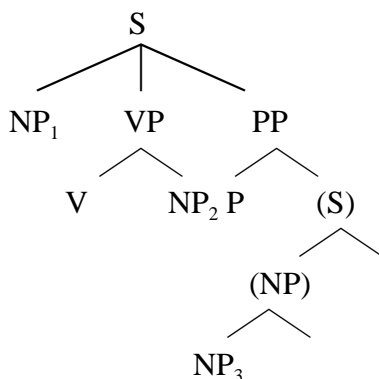
(331) The domain of a node A consists of A together with all and only the nodes c-commanded by A. (OR: The domain of a node A is the subtree dominated by the first branching node which dominates A.)

(332) "What the definition in [(331)] does, basically, is to select the minimal constituent of a node (α), namely all the nodes dominated by the node (β) immediately dominating this node (α)."   p.34

(333)

```
              S
           / | \
         PP  NP₁  VP
        / \       / \
       P  (NP)   V   NP₂
          / \
         NP₃
```

(334)

```
            S
          / | \
        NP₁  VP   PP
            / \   / \
           V  NP₂ P  (S)
                    / \
                  (NP)
                  / \
                NP₃
```

(335) "... in both trees NP_1 c-commands (and commands) NP_3, the only difference being that in [(333)] NP_3 precedes NP_1, while in [(334)] NP_3 follows NP_1. Therefore, by the c-command definition of domain, NP_3 is in the domain of NP_1 in both trees (in other words,

all the nodes in S are in the domain of the subject regardless of whether they precede or follow it). Hence, the coreference rule ... requires that in both trees, NP_3 must be a pronoun in order to be coreferential with NP_1. Applied to sentences with the structure of [(333)], this restriction blocks coreference in [(336)]"

(336)a *Near <u>Dan</u>, <u>he</u> saw a snake   <BUT cf. Near Dan, I think he saw a snake>
    b *Near <u>Dan</u>, <u>Dan</u> saw a snake

(337) "In a right-branching language it is often the case that the c-commanding node precedes the c-commanded node... This may account for how it has been possible for the precede-and-command rule to yield the right prediction in such an amazing number of cases and to nevertheless be the wrong rule. It may also help explain why the relation of precede is believed to play such a crucial role in the grammatical restrictions on coreference. Obviously, in an overwhelmingly large body of the language, forward pronominalization is the only grammatical option. Given the c-command relation, this fact is just an obvious result of application of the coreference rule to right-branching trees. This is not true for the relation <u>command</u>, where all the nodes dominated by the same S equally command each other. A rule stated in terms of command must therefore introduce the relation precede into the rule of coreference."   p.40
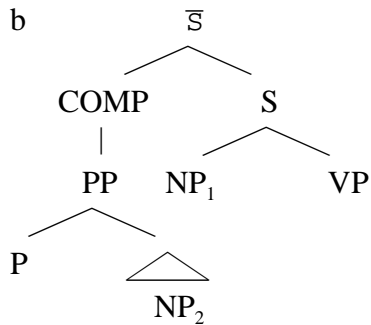
(338) "It is clear that the large correlation between the domains defined by precede-and-command and by c-command holds only for right-branching languages. The sharpest discrepancy between the domains picked up by the two definitions will show up in VOS languages (assuming that these languages have a VP). In such languages a preceding node would often be in the domain of a following node... the following examples from Malagasy (a VOS language with some evidence for a VP) suggest that this hypothesis should be considered. (The examples are from Ed Keenan; personal communication.)

(339)a namono   <u>azy</u> ny anadahin-d<u>Rakoto</u>
         hit / killed <u>him</u> the sister-of-<u>Rakoto</u>
         <u>Rakoto's</u> sister killed <u>him</u>.
      b *namono   ny anadahin-d<u>Rakoto</u> <u>izy</u>
         hit / killed the sister-of-<u>Rakoto</u> <u>he</u>
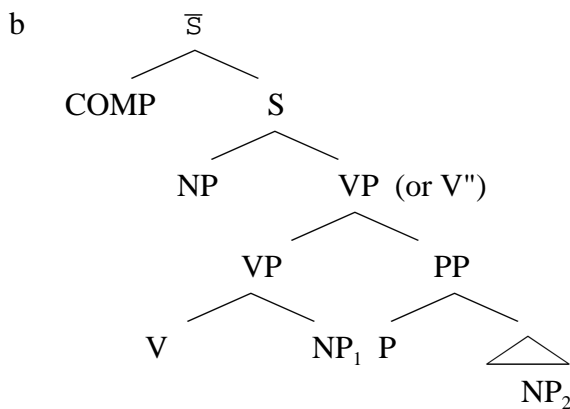   <u>he</u> killed <u>Rakoto's</u> sister.

 In [(339)a] the pronoun precedes and commands the antecedent, hence, by the precede-and-command restriction, the sentence should have been blocked. However, since the pronoun is in the VP, and, thus, does not c-command the antecedent (i.e. , it is not in the c-command domain of the antecedent), the c-command restriction correctly permits coreference. The sentence in [(339)b], on the other hand, does not violate the requirement of precede-and-command (since the antecedent precedes the pronoun), but coreference is, nevertheless, blocked. This is precisely the prediction of the c-command restriction since the pronoun c-commands the antecedent, although the antecedent precedes."   p.41

A modification of c-command

(340)a *In Ben's picture of <u>Rosa</u>, <u>she</u> found a scratch

b
```
                      S̄
                    /    \
                COMP      S
                 |       /  \
                PP   NP₁     VP
               /  \
              P    △
                   NP₂
```

(341)a *I met <u>him</u> in <u>Ben's</u> office

b
```
                    S̄
                  /    \
              COMP      S
                      /   \
                    NP     VP  (or V″)
                          /      \
                        VP        PP
                       /  \      /   \
                      V   NP₁   P     △
                                     NP₂
```

(342)   "The definition of c-command should ... be modified so that in the situation of the type of α over α, the higher α node should be considered as determining c-command relation." p.148

(343)   Modified c-command
        Node A c(onstituent)-commands node B iff the first branching node α_1 dominating A either dominates B or is immediately dominated by a node α_2 which dominates B and α_2 is of the same category type as α_1.

<Note the family resemblance to the category 'segmentation' of May (1985).>

Chomsky (1980) On Binding

Chomsky (1980) presents the first major revision towards modern Binding Theory. The notions 'bound' and 'free' are introduced, with their still standard definitions, as in (344)a,b:

(344)a An anaphor α is bound in β if there is a category c-commanding it and co-indexed with it
       in β.
    b Otherwise, α is free in β.
    c X c-commands Y if every (branching) category dominating X also dominates Y.

The SSC and TSC (actually renamed P(ropositional) I(s1and) C(ondition) in Chomsky (1976)),
are restated initially as (345):

(345) If α is an anaphor in the domain of the tense or the subject of β, β minimal, then α cannot
       be free in β, β = NP or S'.

(346) "The Opacity Condition as given here differs in several important respects from earlier
       formulations of the PIC and SSC. First, it is given as a condition on LF ... rather than as
       a condition on some collection of rules of grammar, including transformations, Control,
       Bound Anaphora, etc. Second, the Opacity Condition is not given as a "constraint on
       variables" relating two positions involved in some rule, but rather as a condition on the
       anaphor.
(347) Third, it is now unnecessary to introduce the notion "specified" in the analogue to the
       SSC. In [(348)], the reciprocal phrase is coindexed with PRO, and is therefore not free in
       S."   p.12

(348) I told them [$_s$ what PRO to give each other t]

(349) "This simplification overcomes a problem for the SSC noted by Lasnik."

(350) which men did Tom think that Bill believed t saw each other

(351) "Here *t* is the trace of *which men*, but if the Reciprocal rule associates *which men* and
       each other, it should be blocked by the SSC, since the specified subject *Bill* intervenes."

Elimination of 'redundancy' in the conditions

(352) *They told me [what I gave each other]

(353) "The structure [(352)] is blocked by the PIC, since *each other* is free in the domain of
       Tense, but also by the SSC, since it is free in the domain of the subject *I*. It has often
       proven a useful strategy to try to eliminate redundancies of this sort. We might do so in
       the present case by restricting the PIC to the subject of a tensed clause."   p.13

Hence, Chomsky proposed the Nominative Island Condition, a narrower version of PIC/TSC
involving only subjects of finite clauses.

(354) A nominative anaphor cannot be free in S' containing S.

Further, a longstanding empirical problem is purportedly solved:

(355)   They expected [that [[pictures of each other] would be on sale]]

TSC/PIC incorrectly excluded such an example, as *each other* does not have a binder inside the minimal finite clause containing it. But since *each other* is not nominative here, (354) allows the example. On the other hand, this is a classic instance of a position of lack of complementarity between anaphors and bound pronouns, so there is always at least half of a problem:

(356)   They expected that pictures of them would be on sale

Further, there is a bit of residual redundancy:

(357)   *The men think [that Mary said [that each other would win]]

The status of wh-trace

Based on an argument of Rizzi (1980), Chomsky suggests that a wh-trace is not an anaphor for Opacity, not just in Italian (as suggested by Rizzi), but in English, as well, the locality constraints on Wh-Movement given by Subjacency instead for the relevant cases. This now partially unifies the treatment of wh trace, given the argument from strong crossover, discussed above, that wh-trace has the behavior of a name. Surprisingly, though, Chomsky explicitly limits the suggestion that Wh trace is not an anaphor to Opacity (SSC) effects, and not to NIC effects. Thus, he proposes that the difference between the mildly ungrammatical Subjacency violation (358)a and the completely impossible (358)b is to be attributed to the fact that NIC as well as Subjacency is violated in the latter.

(358)a   ?What$_1$ [$_s$ did you wonder [$_{s'}$ who$_2$ [$_s$ t$_2$ saw t$_1$]]]
       b   *Who$_1$ [$_s$ did you wonder [$_{s'}$ what$_2$ [$_S$ t$_1$ saw t$_2$]]]

Taraldsen (1978), Pesetsky (1981), and Kayne (1980) present analyses along similar lines. But Freidin and Lasnik (1981) argue against any such treatment, based on strong crossover phenomena. It is true that the standard examples of crossover, such as those presented in Chomsky (1976), all involved object NPs (hence, were not inconsistent with the proposal that a nominative wh-trace is an anaphor). However, alongside the strong crossover examples presented by Chomsky, there are also examples where the trace is nominative. And contrary to what might be expected, such examples are no better than (358)b on the relevant interpretation:

(359)   *Who$_1$ did he$_1$ say t$_1$ kissed Mary?

Thus, wh-trace doesn't behave like an anaphor in NIC contexts any more than in SSC/Opacity contexts.

The OB technology

Index assignment:
(360)a First, movement involves obligatory co-indexing of moved category and its trace.
    b Then, the remaining NPs in a sentence are indexed in 'top to bottom' fashion: an index is assigned to NP only when all NPs that c-command it or dominate it have been indexed.

For anaphors, Chomsky proposes that 'referential indices' are assigned by 'rules of construal' guaranteeing co-indexation with an antecedent. This introduces substantial redundancy into the system, since anaphors that are not bound in the appropriate local domain will be excluded regardless. To the extent that all relevant phenomena can be described in these terms, we will have a strong argument for free indexing instead.

A referential index is an integer. Nonanaphors are also assigned referential indices, though not, of course, by the same algorithm as in the case of anaphors. As before, the algorithm operates top down. But this time, contra-indexing, rather than co-indexing, is what is forced. Each nonanaphoric NP that has not already received an index (by movement) is assigned an unused integer as its referential index. (In this case, as well, the stipulation that the index be a new one is ultimately unnecessary.) In addition to its referential index (an integer), each nonanaphor is also assigned what Chomsky calls an 'anaphoric index'.

(361) The anaphoric index of an NP is the set consisting of every integer that is the referential index of any NP c-commanding the NP in question. The (complete) index of a nonanaphor will be a pair (r, A) where r is the referential index and A the anaphoric index.

(362) $\text{John}_{(2, \{\})}$ told $\text{Bill}_{(3, \{2\})}$ about $\text{him}_{(4, \{2,3\})}$

(363) "We will interpret the anaphoric index $A = \{a_1, \ldots, a_n\}$ of α to mean that α is disjoint in reference from each NP with referential index a_i. Thus, *him* in [(362)] is disjoint in reference from *John* and *Bill*, and if *John* were to replace *him* (or *Bill*) in [(362)], the two occurrences of *John* would be disjoint in reference. Of course, "disjoint reference" in this context has to do with intended reference; actual reference is outside the scope of grammar. On coreference of lexical and pronominal NPs, see Lasnik (1976)."
  p.39

Notice that these indexing principles combined with this interpretive principle give the effects of both RI of Chomsky (1973) and the disjoint reference rule of Lasnik (1976). There is a conceptual difference, however. While the earlier analyses combined the syntactic and semantic aspects of disjoint reference into one rule, the proposal now under discussion modularly separates the two aspects. There are syntactic rules and well-formedness conditions determining what syntactic representations are allowed. Then there is an interpretive principle assigning a meaning to a syntactically well-formed structure.

The present analysis is not yet complete. It does correctly rule out all cases of overlap in reference discussed in Lasnik (1976). Further, it rules out all cases covered by RI. But it does not yet capture the fact that RI showed TSC and SSC effects. Thus, it incorrectly excludes a co-referential interpretation for *John* and *he* or *him* in (364):

(364)a John thinks he is clever
    b John thinks Mary likes him

Chomsky proposes to capture this property and the corresponding property of anaphor binding simultaneously. Specifically, he first proposes that the binding conditions (now NIC and Opacity) can be construed as deleting certain indices from the anaphoric index of a pronoun, thus in effect blocking certain cases of disjoint reference and permitting reference to be free. For an NP with no anaphoric index (i.e., an anaphor) the binding conditions operate on the referential index, altering it in a certain way rather similar to the way in which the anaphoric index is altered in the case of a pronoun. The binding conditions will thus affect the 'designated index' of an NP, where the designated index of a pronoun is its anaphoric index, and the designated index of an anaphor is its referential index (which is, of course, its only index).

(365)   Suppose that α has the designated index j and i is an integer such that i = j or i $\in$ j. Then α is free (i) in β if there is no γ in β with index i that c-commands α.

(366) $[j - i] = (j - i)$ if j is an integer and $\{j - \{i\}\}$ if j is a set.

(367)   Suppose that α has the designated index j and is free(i) in β (β = NP or $\bar{\bar{S}}$),
        where (a) α is nominative
          or (b) α is in the domain of the subject of β, β minimal.
        Then $j \rightarrow [j - i]$.

(368) NP_0 is not permitted in LF, where 0 is the referential index. This is the case of an inadmissible free variable, an anaphor that is not properly bound.   pp.40-41

Lectures on Government and Binding Chomsky (1981)

Chomsky begins by criticizing the On Binding theory of anaphora on several counts.

(369)   The OB system treated PRO as an anaphor. This was too weak in that it didn't entail that PRO occurs only in ungoverned positions, and too strong in that 'long distance control' as in (370) would be incorrectly excluded by SSC.
(370)   They thought I said [that [[PRO to feed each other] would be difficult]]

(371)   In OB there is a sort of redundancy between the theories of Case and binding. They both pick out the subject of infinitives as special, but by totally different means.

(372)   In OB the two configurations relevant to binding theory - subject of a finite clause and c-command domain of a subject - are in no way related.

(373)   The OB indexing conventions are complicated.

LGB approach

(374)   α is bound by β if and only if α and β are coindexed and β c-commands α.
(375)   α is free if and only if it is not bound.

(376)A   An anaphor is (A-)bound in its GC.
      B   A pronominal is (A-)free in its GC.
      C   An R-expression (fully lexical NP, or variable) is (A-) free.

(377)   α is a governing category for β if and only if α is the minimal category containing β, a
        governor of β, and a SUBJECT accessible to β.
(378)   SUBJECT = AGR in a finite clause; NP of S in an infinitival; NP of NP in an NP.

(379)a *John$_i$ believes [(that) himself$_i$ is clever]
      b *They$_i$ believe [(that) each other$_i$ are clever]
      c *Mary$_i$ is believed [(that) t$_i$ is clever]

(380)a John$_i$ believes [himself$_i$ to be clever]
      b They$_i$ believe [each other$_i$ to be clever]
      c Mary$_i$ is believed [t$_i$ to be clever]

(381)a *John$_i$ believes [him$_i$ to be clever]
      b   John$_i$ believes [(that) he$_i$ is clever]

(382)a Mary believes [them to be clever]
      b *Mary believes [(that) them are clever]

(383)   Infinitivals (at least some of them) are not barriers to government, either for Case
        assignment (382) or for establishment of governing category (379)-(381).

(384)   *John$_i$ believes [Mary to like himself$_i$]
(385)   John$_i$ believes [Mary to like him$_i$]

(386)   *He$_i$ believes [(that) John$_i$ is clever]
(387)   *He$_i$ believes [Mary to like John$_i$]

(388)a *We$_i$ heard [their stories about each other$_i$]
      b We$_i$ heard [some stories about each other$_i$]

(389)   Does (378) successfully address (372)? Chomsky suggests that it does, in that SUBJECT
        of α is the most prominent nominal element of α, taking INFL (which contains AGR) as
        the head of S. [But notice it cannot be the head of NP that counts as SUBJECT of NP, or
        (388)b will be ruled out alongside (388)a.]

(390)   They$_1$ expected [that[[pictures of each other$_1$] would be on sale]]
(391)   They$_1$ expected [that[[pictures of each other$_1$]$_2$ AGR$_2$ would be on sale]]

(392)   μ is accessible to β iff β is in the c-command domain of μ and assignment to β of the
        index of μ would not violate (393).
(393)   *[$_\gamma$...δ...], where γ and δ bear the same index.

(394)   i.e., μ is accessible to β iff β is in the c-command domain of μ and μ is not coindexed
        with any category properly containing β.   [Kevin Kearney's formulation]

(395)   *John$_1$ thinks [that [himself$_1$ AGR$_1$ will win]]

(396)   They$_1$ think [it$_2$ AGR$_2$ is a pity [that pictures of each other$_1$ are hanging on the wall]$_2$]
(397)   *They$_1$ think [it$_2$ AGR$_2$ bothered each other$_1$ [that S]$_2$]

(398)   They$_1$ think [it$_2$ AGR$_2$ is a pity [that pictures of them$_1$ are hanging on the wall]$_2$]
(399)   They$_1$ expected [that[[pictures of them$_1$]$_2$ AGR$_2$ would be on sale]]
(400)   Problematically, anaphors and bound pronouns are <u>not</u> in full complementary
        distribution. Another example:
(401)a They read [each other's books]
    b They read [their books]

(402)   John tried [PRO to leave]   <Complement of *try* is S', a barrier to government.>

(403)a *I like PRO
    b *Susan spoke to PRO
    c *John believes [PRO to be intelligent]   <Complement of *believe* is S, not a barrier
                                                        to government.>
    d *John's belief [PRO to be intelligent]
    e *John believes [PRO is intelligent]

(404)   Proposal: PRO is a pronominal anaphor, hence, it must obey both (376)A and (376)B.
        That is, it must be both bound and free in its governing category. If it has a governing
        category, this is a contradiction, therefore it must have no GC. This (almost) entails that
        it must be ungoverned, the descriptive generalization covering (403). This deduction is
        standardly called the **PRO theorem**. It has the effect of permitting long distance control,
        by virtue of divorcing control from binding theory. See (369).
(405)   "...PRO is like overt pronouns in that it never has an antecedent within its clause or NP.
        PRO also resembles anaphors in that it has no intrinsic referential content but is either
        assigned reference by an antecedent or is indefinite in interpretation, lacking specific
        reference."

(406) *[Pictures of each other$_1$]$_2$ AGR$_2$ are on sale
(407) *[Pictures of PRO$_1$]$_2$ AGR$_2$ are on sale
(408) Addendum to (377): A root sentence is a GC for a governed element.

This elegant deduction of the central distributional property of PRO clearly relies on the fact that the definition of GC makes reference to government. Is there independent justification for this aspect of the definition? Chomsky considers a simplification whereby GC is changed to 'binding category' as follows:

(409) β is a *binding category for* α if and only if β is the minimal category containing α and a SUBJECT accessible to α.
(410)A   An anaphor is bound in its binding category
     B   A pronominal is free in its binding category

(411)   This modification has "no effects for elements that are governed since for such elements the governor will always be contained in the binding category."   p.220

(412)   John$_1$ believed [him$_1$ to be the winner]

Here the binding category for *him* is the entire sentence, since in the embedded clause there is no SUBJECT accessible to *him*. The mention of government is actually only significant for PRO. The PRO Theorem is, then, more like the PRO Principle.

The modular character of the theory

(413)   (376)A-C are purely syntactic: they filter out structures based solely on their formal properties. Do we need any associated semantics?

(414)   "John likes him" can't mean that John likes himself.
(415)   *John$_1$ likes him$_1$
(416)   John$_1$ likes him$_2$

(417)   If two NP's have distinct indices then...

(418)   They like him

(419)   ...then they are disjoint in reference.

(420) *We$_1$ like myself$_2$
(421)   →We$_1$ like myself$_1$ <Just a failure of agreement? No: "They like themselves" demands coreference.>

(422)   If two NPs have identical indices, then they are coreferential.

(423) We$_1$ think [I$_?$ will win]

(424) The problem is that NP's have at least three referential relations: disjointness; identity; overlap. But two numerical subscripts are either identical or distinct. See (373). (423) is the cost of addressing (373) in the way that LGB does, as discussed in Lasnik (1981).

A problem with expletives

(425) An expletive and its associated argument must be coindexed to establish the appropriate Case and agreement relation. But such a structure seems to violate Condition C:
(426) There$_i$ is a man$_i$ in the room
(427) There^i is a man^i in the room
(428) Binding theory cares only about <u>sub</u>scripts.

Binding conditions at S-structure

(429) [Which book that John$_i$ read]$_j$ did he$_i$ like $\underline{t}_j$
(430) *He$_i$ liked [every book that John$_i$ read]
(431) *Who thinks that he$_i$ read [which book that John$_i$ likes]

QR and LF wh-movement would incorrectly neutralize the Condition C difference between these examples.

(432) "...these examples provide prima facie evidence that the binding theory applies at S-structure, a conclusion that I will now adopt."   p.197

Barss (1986) draws the same conclusion for Condition A, based on examples like the following:

(433)a   John$_i$ wonders [[which picture of himself$_i$] [Mary showed to Susan]]
     b *John$_i$ wonders [who [showed which picture of himself$_i$ to Susan]]

(433)a shows that an anaphor within the embedded CP Spec can be licensed by an antecedent in the matrix subject position. Given this fact, the ungrammaticality of (433)b is surprising if anaphors can be licensed by virtue of their LF positions. (On the then standard theory, in LF, WH-phrase in situ, *which picture of himself*, moves to the embedded CP Spec position, where it takes scope. Thus, at LF, the configurational relation between *himself* and its antecedent is virtually identical in (433)a and (433)b. Hence, the ungrammaticality of (433)b shows that anaphors must be licensed at a level prior to LF, e.g., S-structure (and possibly at LF as well).

Knowledge of Language Chomsky (1986)

An alternative account of existential constructions, based on 'expletive replacement' driven by Full Interpretation:

(434) A man$_i$ is $\underline{t}_i$ in the room [where $\underline{t}$ is an A-movement trace, hence not a variable].

(435) Binding theory applies at LF, and <u>not</u> at S-structure.   [But cf. (431) and (433). Sorting

out this apparent contradiction became one of the major goals of minimalist theorizing about anaphora.]

den Dikken (1995) shows that Chomsky can't be correct **even for existential constructions**:

(436)   *There seem to each other$_i$ to be some applicants$_i$ eligible for the job

An account of non-complementarity between anaphors and bound pronouns

(437)a They$_i$ read [each other's$_i$ books]
     b They$_i$ read [their$_i$ books]

Following Huang (1983), Chomsky proposed to eliminate the total complementarity between anaphors and pronouns:

(438)   "...the relevant local domain is different in some respect for anaphors and pronominals ... this difference should fall out as an immediate consequence of the difference in their nature - namely, that anaphors must be bound whereas pronominals must be free - without any need to stipulate any further difference in the binding theory conditions for these two categories of expressions."   p.170

(439)   The following definitions and licensing conditions concern an expression E with indexing I. The indexing I and a pair (α,β) are <u>compatible</u> with respect to the binding theory if α satisfies the binding theory in the local domain β under the indexing I. A 'complete functional complex' (CFC) is a projection of a head including all grammatical functions compatible with that head.

(440)   I is <u>BT-compatible</u> with (α,β) if:
        (A) α is an anaphor and is bound in β under I
        (B) α is a pronominal and is free in β under I
        (C) α is an r-expression and is free in β under I

(441)   Licensing condition for a category α governed by a lexical category γ in the expression E with indexing I:
         For some β such that (i) or (ii), I is BT-compatible with (α,β):
        (i) α is an r-expression and (a) if α heads its chain or (b) otherwise
         (a) β = E
         (b) β is the domain of the head of the chain of α
        (ii) α is an anaphor or pronominal and β is the least CFC containing γ for which there is an indexing J BT-compatible with (α,β)

(442)   "...for α an anaphor or pronominal ... the licensing condition amounts to saying that the relevant governing category for α is the minimal one in which binding theory could have been satisfied under some indexing."   p.172

(443)   "... if α has no governor, then the licensing condition [(441)] is inapplicable."   p.172

This allows PRO in ungoverned positions. What blocks it from governed positions with a small GC for its pronominal aspect and a larger one for its anaphoric aspect?

(444) "If PRO is governed, then it is subject to the licensing condition (441). But this condition can never be satisfied, because whatever the indexing I, there can be no β such that I is BT-compatible with (PRO, β); as an anaphor, PRO would have to be bound in β under I, and as a pronominal it would have to be free in β under I. Therefore, PRO cannot be governed."   p.183

(445) The major instances of non-complementarity are now accommodated. But there is an apparent cost: we have lost the TSC/NIC:

(446)a *John$_i$ believes [(that) himself$_i$ is clever]
     b *They$_i$ believe [(that) each other$_i$ are clever]

(447) Chomsky's solution to this problem takes us full circle, all the way back to Chomsky (1973): the constraint on anaphora here should reduce to a constraint on movement (rather than vice versa). Chomsky suggests that the movement constraint (the ECP) at work in (448) is also relevant in (446) assuming that in LF, anaphors undergo movement to be in some appropriate very local relation with their antecedents.

(448) *Mary$_i$ is believed [(that) $\underline{t}_i$ is clever]

Note that this approach has the desirable effect of reducing some of the 'redundancy' in the treatment of (448). Given that it already violates ECP and the 'last resort' condition on A-movement, we would like it <u>not</u> to also violate Condition A.

Chomsky and Lasnik (1993) Towards a minimalist theory

A reanalysis of the distribution of PRO

A problem with the LGB treatment: argument PRO appears in non-Case-positions, necessitating a disjunctive version of the Visibility Condition:

(449)   A chain is visible for θ-marking if it contains a Case-position (necessarily, its head) or is headed by PRO

A new generalization: Like other arguments, PRO is not permitted to move from a Case-marked position, even to escape government:

(450)   (i) α to talk about β
        (ii) α to strike β [that the problems are insoluble]
        (iii) α to seem to β [that the problems are insoluble]

Suppose that (i) is a D-structure in the context "it is unfair --," with α = e and β = *John*. The last

resort condition bars raising of β to position α, yielding (451)i, because *John* is already visible for θ-marking without movement. Suppose β = PRO. On the assumptions now under consideration, PRO must raise to the position α to satisfy the nongovernment requirement. But that movement is impermissible, even though α is a legitimate position for PRO in other constructions, as in (451)iii:

(451) (i) *it is unfair [John to talk about t]
       (ii) *it is unfair [PRO to talk about t]
       (iii) it is unfair [PRO to talk about John]

(452) (i) It is rare for it to strike John [that the problems are insoluble]
       (ii) It is rare for it to seem to John [that the problems are insoluble]

(453) (i) *We want John to strike t [that the problems are insoluble]
       (ii) *We want John to seem to t [that the problems are insoluble]

(454) (i) *We want PRO to strike t [that the problems are insoluble]
       (ii) *We want PRO to seem to t [that the problems are insoluble]
<cf. (iii) *We want it to seem to PRO [that the problems are insoluble] >

PRO is now in an ungoverned position, heading a θ-marked chain. Hence all conditions are satisfied. But the constructions are radically ungrammatical, whatever the context.

Some principle requires that PRO behave like other arguments, moving from non-Case-positions and barred from moving from Case-positions. Suppose then PRO, like other arguments, has Case, but (at least sometimes) a Case different from the familiar ones: nominative, accusative, etc. From the point of view of interpretation, we might regard PRO as a "minimal" NP argument, lacking independent phonetic, referential or other properties. Accordingly, let us say that it is the sole NP that can bear null Case (though it may have other Cases as well, in nonstandard conditions that we will not review here).

It follows that the last resort principle applies to PRO exactly as it does to any argument: PRO is permitted to move from a non-Case-position to a position where its Case can be assigned or checked, and is not permitted to move from a Case-position. The Visibility Condition can now be simplified to (455):

(455) A chain is visible for θ-marking if it contains a Case-position

         Where, then, is null Case checked? Nominative Case is standardly checked in [SPEC, IP], where I involves the features of tense and agreement (T, AGR). It is thus a realization of a SPEC-head relation, with the head = INFL, the head of IP. It is natural, then, to take null Case to be a realization of the same relation where INFL lacks tense and agreement features: the minimal INFL checks null Case, and the minimal NP alone can bear it. More generally, we may assume that the infinitival element (with null agreement) and the head ING of gerundive nominals check null Case, so that PRO will appear in such constructions as:

(456)(i)   PRO to VP   (to be sick)
     (ii) PRO ING VP   (being sick)

Binding Theory without indices?

(457) (i)   If α is an anaphor, interpret it as coreferential with a c-commanding phrase in D
      (ii)   If α is a pronoun, interpret it as disjoint from every c-commanding phrase in D.
      (iii)   If α is an r-expression, interpret it as disjoint from every c-commanding phrase.

(457)ii is very reminiscent of RI of Chomsky (1973). And (457)iii is similarly reminiscent of the disjoint reference rule of Lasnik (1976).

Reconstruction

(458)   Which picture of himself does John like

Riemsdijk and Williams (1981) propose that Condition A is satisfied at NP-structure, a level of representation following all A-movement and preceding all A'-movement

(459)   [[which pictures of himself$_{i/j}$] did John$_i$ think [t' [Fred$_j$ liked t]]]   Barss (1986)

As Barss notes, when the antecedent is *John*, the NP-structure account fails (as would a deep structure interpretive procedure).

There are several pieces of machinery that would work:

(460)   Barss's proposal that all parts of a chain are accessible for determination of Condition A satisfaction.
(461)   The Belletti and Rizzi (1988) idea that Condition A can be satisfied anywhere in the course of the derivation.
(462)   The copy theory of movement, under which traces are all copies of the moved item. Chomsky (1993)   <already considered, but rejected, in Chomsky (1981, p.89)

Detailed discussion of reconstruction in Chomsky (1993).

(463) Arguments, of increasing strength, against an S-Structure condition:
(464)a The condition *can* apply at LF alone.
     b Furthermore, the condition sometimes *must* apply at LF.
     c Furthermore, the condition must *not* apply at S-Structure.
                                                                    Chomsky, p. 192

(465)a John$_1$ wondered [[which picture of himself$_{1/2}$][Bill$_2$ saw *t*]]
     b John$_1$ wondered [who$_2$ [*t* saw [which picture of himself$_{*1/2}$]]]
(466)a The students$_1$ asked [[what attitudes about each other$_{1/2}$][the teachers$_2$ had noticed *t*]]
     b The students$_1$ asked [who$_2$ [*t* had noticed [what attitudes about each other$_{*1/2}$]]]   p.205

(467) The bad readings of (465)b and (466)b are ruled out at LF, under the assumption that LF movement is not of the entire wh-phrase. Then no appeal to S-Structure is required.

(468) The $_2$ readings of all the examples implicate the 'copy theory of movement'.

(469) John wondered [[$_{wh}$which picture of himself]][Bill saw [[$_{wh}$which picture of himself]]]

(470) Then, by an LF "operation akin to QR" we have (471) or (472), depending on the size of the QRed item.

(471) John wondered [[[$_{wh}$which picture of himself][$_{wh}$ t]]][Bill saw [[$_{wh}$which picture of himself][$_{wh}$ t]]]]

(472) John wondered [[which [$_{wh}$ t picture of himself]][Bill saw [which [$_{wh}$ t picture of himself]]]]

(473) With complementary deletion to produce an operator variable structure, we have:

(474) John wondered [[[$_{wh}$which picture of himself][$_{wh}$ t]]][Bill saw [[$_{wh}$ ~~which picture of himself~~][$_{wh}$ t]]]]
(475) John wondered [[which x, x a picture of himself][Bill saw x]]

(476) John wondered [[which [$_{wh}$ ~~t picture of himself~~]][Bill saw [~~which~~ [$_{wh}$ t picture of himself]]]]
(477) John wondered [[[which x][Bill saw x picture of himself]]]
(478)a In (474), <u>John</u> is the antecedent of <u>himself</u>.
     b In (476), <u>Bill</u> is the antecedent of <u>himself</u>.   p.206

(479) John wondered [[which picture of himself][Bill took t]]

(480) <u>Himself</u> in (479) can take <u>John</u> or <u>Bill</u> as antecedent, just as in the earlier examples, BUT only when <u>took</u> means 'pick up and walk away with'.
(481) When <u>took (pictures)</u> means 'photograph' (the 'idiomatic reading'), <u>Himself</u> can only take <u>Bill</u> as antecedent, according to Chomsky.

(482)a John wondered [[which x, x a picture of himself][Bill took x]]
     b John wondered [[[which x][Bill took x picture of himself]]]

(483) "Having abandoned D-Structure, we must assume that idiom interpretation takes place at LF ..."   p.207
(484) "Thus, *take ... picture* can be interpreted as 'photograph' only if the phrase is present as a unit at LF - that is, in (482)b, but not (482)a."

(485) This explains why in (482)a we can only have the nonidiomatic interpretation of *take*.

(486) The students$_1$ asked [[what attitudes about each other$_{*1/2}$][the teachers$_2$ had t]]
(487) Chomsky gives a parallel analysis here: <u>have</u> ... <u>attitudes</u> is a sort of idiom, so must be unified at LF.

(488) "The conclusions follow on the crucial assumption that Condition A *not* apply at S-Structure... If Condition A were to apply at S-Structure, *John* could be taken as antecedent of *himself* in [(479)] and the later LF processes would be free to choose either the idiomatic or the literal interpretation, however the reconstruction phenomena are handled ..." p.207

(489) "Thus, we have the strongest kind of argument against an S-Structure condition ... Condition A *cannot* apply at S-Structure." p.208

(490) But there is now a near contradiction with Chomsky's account of the Freidin-Lebeaux Condition C examples (to be discussed below), as Chomsky observes (p.208).

(491)a John$_1$ wondered [which picture of himself$_{1/2}$][Bill$_2$ saw t]
    b John$_1$ wondered [which picture of Tom$_2$][he$_{1/*2}$ liked t]

(492) In (491), the <u>of</u> phrase, being a complement, must reconstruct. This gives the right result for (491)b but not for (491)a.

(493) The $_2$ reading of (491)b is ruled out by the preference principle that Chomsky proposes for the Freidin-Lebeaux cases:
(494) "... try to minimize the restriction in the operator position ..." p.209

(495) Why "minimize the restriction"? Why not "maximize the restriction"? A speculation: When you minimize the restriction, you have QRed a smaller (and proper subpart) of what you would QR to maximize the restriction. Moving less is more economical than moving more (like the deduction of Procrastinate from economy).

(496) To allow the $_1$ reading of (491)a, we need it to be true that something makes the normally disfavored option necessary.

(497) That something is the LF cliticization approach to anaphora:

(498)a John self-wondered [which picture of t_{self}][NP saw [which picture of himself]]
    b John wondered [which picture of himself][NP self-saw [which picture of t_{self}]]

(499)a [[which picture of α] t]   α = t_{self} or *himself*
    b [which][t picture of α]

(500) If we select the syntactic option (498)a then we cannot select the interpretive option (499)b (with α = t_{self}).
(501) That option requires deletion of [t *picture of* t_{self}] in the operator position, which would break the chain (*self, t_{self}*), leaving the reflexive without a θ-role at LF.
(502) In short, if we take the antecedent of the reflexive to be *John*, then only the nonreconstructing option converges.

(503) several pictures were taken *t*

(504) the students asked [which pictures of each other] [*t′* were taken *t* by Mary]
(505) "One possibility is that the trace of the A-chain enters into the idiom interpretation (and, generally, into θ-marking), while the head of the chain functions in the usual way with regard to scope and other matters."   p.211

(506) the claim that John was asleep seems to him [$_{IP}$ *t* to be correct]
(507) "... if "reconstruction" is essentially a reflex of operator-variable constructions, it will hold only for Ā-chains, not for A-chains."   p.205

Condition C reconstruction

I. Condition C Complement/Adjunct Reconstruction Asymmetries (The 'Freidin-Lebeaux' Effect')

(508)a   Which report that John$_i$ revised did he$_i$ submit?
      b *Which report that John$_i$ was incompetent did he$_i$ submit?
                              Freidin (1986)
(509) a *He$_i$ believes the claim that John$_i$ is nice.
      b *He$_i$ likes the story that John$_i$ wrote.
      c *Whose claim that John$_i$ is nice did he$_i$ believe?
      d Which story that John$_i$ wrote did he$_i$ like?
                              Lebeaux (1988)
(510)a *Which claim that John$_i$ was asleep did he$_i$ later deny
      b Which claim that John$_i$ made did he$_i$ later deny
                              Munn (1994)
(511)a *Which claim [that John$_i$ was asleep] was he$_i$ willing to discuss
      b Which claim [that John$_i$ made] was he$_i$ willing to discuss
(512)a *The claim that John$_i$ is [sic] asleep, he$_i$ was willing to discuss
      b The claim that John$_i$ made, he was willing to discuss
                              Chomsky (1993)
(513)a *The claim that John$_i$ was asleep, he$_i$ won't discuss
      b The claim that John$_i$ made, he$_i$ won't discuss
                              Chomsky and Lasnik (1993)

(514)   The claim that John$_i$ was asleep seems to him$_i$ [$_{IP}$ *t* to be correct]
         Chomsky (1993)
(515)   *I seem to him$_i$ [*t* to like John$_i$]

(516)a The 'Extension Condition': structure must be built strictly cyclically.
      b Adjuncts are exempt from the Extension Condition; relative clauses are adjuncts.
      c "Reconstruction" is essentially a reflex of the formation of operator-variable constructions.

d An operator chain (a sequence of copies) undergoes complementary deletion.

e Condition C is an LF requirement.   Chomsky (1993)

(517)a [[Which claim][that John made]] was he willing to discuss ~~which claim~~   PF

   b [[Which [*t* ~~claim~~]][that John made]] was he willing to discuss [~~which~~ [*t* claim]]   LF

   c For which x that John made, he was willing to discuss x claim   Interpretation (?)

OR?

(518)a [[Which claim][that John made]] was he willing to discuss ~~which claim~~   PF

   b [[Which claim][*t*]][that John made]] was he willing to discuss [[~~which claim~~][*t*]] LF

   c For which x, x a claim that John made, he was willing to discuss x   Interpretation (?)

(519)a Which claim [that John was asleep] was he willing to discuss ~~which claim that John was asleep~~   PF

   b [Which [*t* ~~claim [that John was asleep]~~]] was he willing to discuss [~~which~~ [*t* claim that John was asleep]   LF

   c For which x, he was willing to discuss x claim that John was asleep   Interpretation (?)

BUT CRUCIALLY NOT

(520)a Which claim [that John was asleep] was he willing to discuss ~~which claim that John was asleep~~   PF

   b [[Which [*t* ~~claim~~]] [that John was asleep]]] was he willing to discuss [[~~which~~ [*t* claim]] ~~that John was asleep~~]]   LF

   c For which x that John was asleep, he was willing to discuss x claim   Interpretation (?)

OR

(521)a Which claim [that John was asleep] was he willing to discuss ~~which claim that John was asleep~~   PF

   b [[[Which claim][that John was asleep]]] *t*] was he willing to discuss [[~~which claim that John was asleep~~]*t*]   LF

   c For which x, x a claim that John was asleep, he was willing to discuss x
      Interpretation (?)

(522) "...preference principle for reconstruction: Do it when you can (i.e., try to minimize the restriction in the operator position)."   p.209

II. Concerns About the Generalization   Lasnik (1998)

(523) Which piece of evidence that John was guilty did he successfully refute?

(524) The widespread belief that John is incompetent, he deeply resents

(525) Whose argument that John was incorrect did you show him?

(526) How many arguments that John's theory was correct did he publish?

(527) This argument that John's theory is correct, he is now ready to publish.

(528) Which proof that Mary's theory is superior to John's did she present?

(529) Mary's attempt to hire John's student, he heartily endorsed.

(530) John's request to attend Mary's lecture, she immediately granted.

(531)a The claim that the director$_i$ was corrupt, he$_i$ was unwilling to discuss
      b That the director$_i$ was corrupt, everyone knew that he$_i$ would always be able to deny with
        a straight face   Postal (1997)

(532)a Whose allegation that John$_i$ was less than truthful did he$_i$ refute vehemently?
      b Whose claim that the Senator$_i$ had violated the campaign finance regulations did he$_i$
        dismiss as politically motivated?   Kuno (1997)

(533)a *Which claim that John$_i$ was asleep did he$_i$ later deny
      b   Which claim that John$_i$ made did he$_i$ later deny
                              Munn (1994)
(534)   Later than what, one might ask?

(535) *Whose claim that John$_i$ is nice did he$_i$ believe?   Lebeaux (1988)
(536) Susan: John is nice.
       Mary: John is nice.
    !John: I believe Susan but I don't believe Mary.

(537) Lydia Grebenyova's experiment (with UMD undergrads, 2004):
(538) Two claims have been made about John's arrest: that John was arrested yesterday and that
        John was arrested a week ago. John has a lawyer, whose name is Bill
(539) Which specific claim that John had been arrested did Bill deny
(540) Which specific claim that he had been arrested did John deny
(541) Which specific claim that John had been arrested did he deny
(542) 6 of 7 subjects accepted both (540) and (541) on the coreferential reading.

(543) What if the complement/relative asymmetry with WH-movement is illusory. How
        problematic is that for the theory?
(544)a (516)a vs. b is arguably just a stipulation, as is (516)c.
      b The status of (522) is unclear.
(545)   If anything, then, lack of that asymmetry might be a 'better' state of affairs. (The only
        mildly negative consequence, depending on your point of view, is that a potential
        argument for traces, i.e., copies, disappears.)

References

Bach, Emmon. 1970. Problominalization. *Linguistic Inquiry* 1: 121-122.
Barss, Andrew. 1986. *Chains and anaphoric dependence: On reconstruction and its implications*.
        Doctoral dissertation, MIT, Cambridge, Mass.
Belletti, Adriana and Luigi Rizzi. 1988. Psych-verbs and theta theory. *Natural Language and
        Linguistic Theory* 6: 291-352.
Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms. Harvard University,
        Cambridge, Mass. and MIT, Cambridge, Mass.*, .*[Revised 1956 version published in part
        by Plenum, New York, 1975; University of Chicago Press, Chicago, 1985].
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT press.
Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational*

grammar, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 184-221. Waltham, Mass.: Ginn and Co. [Reprinted in Noam Chomsky, *Studies on semantics in generative grammar*, 11-61. The Hague: Mouton, 1972].

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2: 303-351.

Chomsky, Noam. 1980. On binding. *Linguistic Inquiry* 11: 1-46.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, Mass.: MIT Press. [Reprinted in Noam Chomsky, *The minimalist program*, 167-217. Cambridge, Mass.: MIT Press, 1995].

Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, Vol. 1, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506-569. Berlin: Walter de Gruyter. [Reprinted in Noam Chomsky, *The minimalist program*, 13-127. Cambridge, Mass.: MIT Press, 1995].

Culicover, Peter. 1976. A constraint on coreferentiality. *Foundations of Language* 14: 109-118.

den Dikken, Marcel. 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26: 347-354.

Dougherty, Ray C. 1969. An interpretive theory of pronominal reference. *Foundations of Language* 5: 488-508.

Fiengo, Robert and Howard Lasnik. 1973. The logical structure of reciprocal sentences in English. *Foundations of Language* 9: 447-468. [Reprinted in Essays on anaphora, Howard Lasnik, 37-57. Dordrecht: Kluwer, 1989]

Freidin, Robert and Howard Lasnik. 1981. Disjoint reference and *wh*-trace. *Linguistic Inquiry* 12: 39-53. [Reprinted in Essays on anaphora, Howard Lasnik, 110-124. Dordrecht: Kluwer, 1989]

Grinder, John and Paul Postal. 1971. Missing antecedents. *Linguistic Inquiry* 2: 269-312.

Helke, Michael. 1971. *The grammar of English reflexives*. Doctoral dissertation, MIT, Cambridge, Mass.

Higginbotham, James. 1980. Pronouns and bound variables. *Linguistic Inquiry* 11: 679-708.

Huang, C. -T. James. 1983. A note on the binding theory. *Linguistic Inquiry* 14: 554-561.

Jackendoff, Ray. 1969. *Some rules of semantic interpretation for English*. Doctoral dissertation, MIT, Cambridge, Mass.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1980. Extensions of binding and Case-marking. *Linguistic Inquiry* 11: 75-96. [Reprinted in Kayne 1984]

Klima, Edward S. 1964. Negation in English. In *The structure of language: Readings in the philosophy of language*, ed. Jerry A. Fodor and Jerrold J. Katz, 246-323. Englewood Cliffs, N. J.: Prentice-Hall.

Lakoff, George. 1968. *Pronouns and reference*. Bloomington, In.: Indiana University. Distributed by Indiana University of Linguistics Club. [Reprinted in Syntax and Semantics Volume 7: Notes from the Linguistic Underground, ed James D. McCawley, 275-335 New York: Academic Press, 1976]

Langacker, Ronald W. 1969. On pronominalization and the chain of command. In *Modern Studies in English*, ed. David A. Reibel and Sanford A. Schane, 160-186. Englewood Cliffs, N. J.: Prentice-Hall.

Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2: 1-22. [Reprinted in Essays on anaphora, Howard Lasnik, 90-109. Dordrecht, Kluwer, 1989]

Lasnik, Howard. 1981. On two recent treatments of disjoint reference. *Journal of Linguistic Research* 1: 48-58. [Reprinted in Essays on anaphora, Howard Lasnik, 125-133. Dordrecht, Kluwer, 1989]

Lasnik, Howard. 1991. On the necessity of binding conditions. In *Principles and parameters in comparative grammar*, ed. Robert Freidin, 7-28. Cambridge, Mass.: MIT Press. [Reprinted in Essays on anaphora, Howard Lasnik, 149-167. Dordrecht: Kluwer, 1989].

Lasnik, Howard. 1998. Some reconstruction riddles. In *University of Pennsylvania Working Papers in Linguistics* 5(1), 83-98.

Lees, Robert B. and Edward S. Klima. 1963. Rules for English pronominalization. *Language* 39: 17-28.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1981. Complementizer-trace phenomena and the nominative island condition. *The Linguistic Review* 1: 297-343.

Postal, Paul M. 1966. A note on 'understood transitively'. *International Journal of American Linguistics* 32: 90-93.

Postal, Paul M. 1969. Review of A. McIntosh and M. A. K. Halliday, Papers in general descriptive and applied linguistics. *Foundations of Language* 5: 409-439.

Postal, Paul M. 1970. On coreferential complement subject deletion. *Linguistic Inquiry* 1: 439-500.

Reinhart, Tanya. 1976. *The syntactic domain of anaphora*. Doctoral dissertation, MIT, Cambridge, Mass.

Riemsdijk, Henk van and Edwin Williams. 1981. NP structure. *The Linguistic Review* 1: 171-217.

Rizzi, Luizi. 1980. Violations of the *wh*-island constraint and the Subjacency condition. *Journal of Italian Linguistics* 5: 157-195.

Rosenbaum, Peter. 1965. A principle governing deletion in English sentential complementation. RC 1519. Thomas J. Watson Research Center, IBM Corporation, Yorktown Heights. [Reprinted in R. Jacobs and P. Rosenbaum (1970) Readings in English transformational grammar]

Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, Mass.: MIT Press.

Ross, John Robert. 1967a. *Constraints on variables in syntax*. Doctoral dissertation, MIT, Cambridge, Mass. Published as *Infinite syntax!* Norwood, N.J.: Ablex (1986).

Ross, John Robert. 1967b. On the cyclic nature of English pronominalization. In *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday, 11 October 1966*, 1669-1682. The Hague: Mouton.

Taraldsen, Knut Tarald. 1978. On the NIC, vacuous application and the *that*-trace filter. Ms. MIT, Cambridge, Mass.

Wasow, Thomas. 1972. *Anaphoric relations in English*. Doctoral dissertation, MIT, Cambridge, Mass. [Revised version published as Thomas Wasow, *Anaphora in generative grammar*. Ghent: E. Story-Scientia, 1979]

Witten, Edward. 1970. Pronominalization: a handbook for secret agents. Ms. Brandeis University, Waltham, Mass.