Some Aspects of Movement and Deletion
Pauline Jacobson
Brown University

One of the basic assumptions in most current grammatical the-
ories is that, for a given derivational stage, a node can be di-
rectly dominated by only one node. As a consequence of this,
what I will call the conventional theory of movement assumes that
for a sentence like:

1. John seems to grow

there is no stage at which <u>John</u> is both the subject of <u>seem</u> and
the subject of <u>grow</u>. By the conventional theory I mean also one
which maintains that, at the output of a rule moving some NP, no
NP remains in the position from which that NP is moved. I will
call a conventional theory of deletion one in which, for a sen-
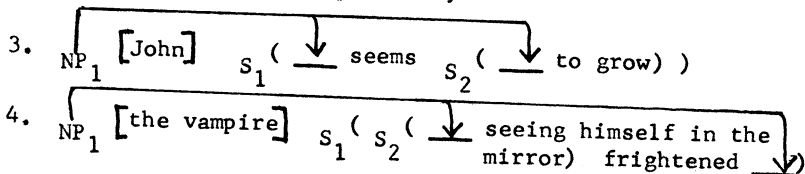tence like:

2. Seeing himself in the mirror frightened the vampire

there are underlyingly two occurrences of <u>the vampire</u> and where
there is a stage at which the subject of <u>see</u> is removed, leaving
no subject in the lower clause.

This paper primarily explores some consequences of a theory
which does not assume that a node can be directly dominated by
only one node; I will refer to this as the Multi-Domination (MD)
theory. I will also briefly consider a theory which does not per-
mit multi-domination, but where at the output of a movement or
deletion rule some NP remains in the position of the moved or
deleted NP. (An example of this with respect to movement is a
trace theory such as that outlined in Fiengo (1974)). This paper
is not meant as an argument for either of these theories, but
simply as a brief investigation and comparison of some of their
properties.

1. Some Aspects of a Multi-Domination Theory

By a MD theory I mean any theory which permits a node to be
directly dominated by more than one node at a single derivational
stage. More specifically, then, a MD theory of movement would be
one in which a sentence like (1) has a level of representation in
which there is a single NP node (dominating <u>John</u>) which is the
subject of <u>seem</u> and the subject of <u>grow</u>. Similarly, a MD theory
of deletion would posit a level of representation for (2) in
which there is a single node (dominating <u>the vampire</u>) which is
in both clauses. Thus I will represent the relevant levels of
(1) and (2) as (3) and (4) respectively:

3. NP_1 [John] S_1 ($\underline{\quad\downarrow\quad}$ seems S_2 ($\underline{\quad\downarrow\quad}$ to grow))

4. NP_1 [the vampire] S_1 (S_2 ($\underline{\quad\downarrow\quad}$ seeing himself in the
mirror) frightened $\underline{\quad\downarrow\quad}$)

No particular claim is intended by this notation, and I will, as much as possible, consider these treatments of movement and deletion independently of any larger theories. Sampson (1975) has proposed a MD theory for deletions; Postal (in recent talks at Harvard and Brown) has also suggested this treatment of deletion, and Lakoff (1975) has proposed a theory incorporating a MD treatment of both movement and deletion. However, the latter two theories make a number of other claims which are not relevant here; for example, neither include derivations in the standard sense. I will consider a theory which is, in all other respects, like a standard derivational theory. Thus I will assume that there are at least two levels of representation for (1) - an underlying one in which John occurs only in the lower clause, and a later level whose representation is (3). (2), on the other hand, would have (4) as its underlying representation. Moreover, my remarks will be made within a non-relational framework, though the basic claim of a MD theory is neutral between a relational and a non-relational framework.

The MD theory could maintain that (3) and (4) are surface structures, or it could posit a later level at which NP_1 is removed from the lower clause. I will assume a theory in which there is no representation where the lower clause has no subject; thus (3) and (4) are surface structures. Yet clearly more needs to be said, for it must be predicted that John is not pronounced in both clauses. Let us assume for now that there is a principle roughly as follows:

5. If there is an occurrence of NP_a in S_1 such that S_1 is the first S-node up dominating NP_a and NP_a is dominated by S_2 where S_1 dominates S_2, then NP_a is not pronounced in S_2 (i.e., an NP is pronounced only in its highest clause).

Notice that, with respect to deletion, this principle makes the apparently correct prediction that, in terms of the conventional theory, a deletion site never asymmetrically commands the deletion controller. (It has often been noted that there are cases where neither the deletion site nor the controller command the other. (5) does not cover these cases, and so it might appear that extra principles are needed here. However, in the next section I will show that at least one such case is accounted for in the MD theory by (5)).

The discussion above only concerns cases in which a node is in two clauses at a single derivational stage; a MD theory could also maintain that an NP can be in two positions within a clause. Thus, for example, this theory would posit (7) as both the underlying and the surface structure of (6):[1]

6. John washed

7. $\mathsf{NP_1}$ [John] ($\downarrow$ washed $\downarrow$)
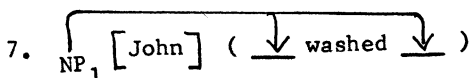
Similarly, a MD theory of movement would predict that in the surface structure of (8):

8. John was killed by Mary

John is both the subject and the object of kill. Here again some principle is needed to predict that John is not pronounced in both positions; it could be assumed that there is a principle which can be characterized very roughly as follows:[2]

9. If NP_a is in a position X and a position Y where X precedes and commands Y, then NP_a is pronounced only in position X.
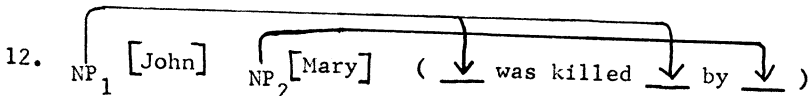
There is here an interesting parallel between a MD treatment of a sentence like (8) and the trace theory proposed in Fiengo (1974). The latter theory claims that a moved NP leaves a trace in the position from which it is moved and that the trace is present in surface structure unless some other NP moves into that position. Consider a MD theory which makes the following standard assumption:

10. For each position, only one node can be in that position at a single derivational stage.[3]

Both the MD theory, then, and a trace theory predict that John (or a trace created by the movement of John) is the object of kill in surface structure, while both theories predict that Mary (or a trace created by the movement of Mary) is not the surface subject of kill. Thus the surface structure of (8) in a trace theory is:

11. John was killed t by Mary

while in the MD theory it is:

12. NP_1 [John] NP_2[Mary] (↓ was killed ↓ by ↓)

2. An Argument for a MD Treatment of Deletion

One advantage of the MD treatment of deletion concerns a class of cases in which a deletion controller acts as though it were in the position of the deletion site with respect to various processes; such an example is a sentence discussed by Akmajian (1972) as a reply to Postal and Ross (1971). Thus Postal and Ross argued that, given the assumption that a deletion site must be commanded by the controller, a sentence like:

13. Getting herself into college is hard for me to imagine Betsy being willing to consider

must be derived by a rule moving the get-clause into subject position rather than from a structure in which this clause is the underlying subject of hard and deletes the object of consider. Here the subject of the get-clause is deleted by Betsy; if this

clause is the underlying object of <u>consider</u> then <u>Betsy</u> commands
the deletion site at some stage of the derivation. If, on the
other hand, (13) were derived from a structure in which the <u>get</u>-
clause is the underlying subject of <u>hard,</u> then there is no deri-
vational stage at which <u>Betsy</u> commands the deleted subject in the
<u>get</u>-clause.

While a movement analysis can account for (13), Akmajian
points out that it cannot be maintained that the <u>get</u>-clause is
not the underlying subject of <u>too crazy</u> in a sentence like:

> 14. Getting herself into college is just too crazy for me
>   to imagine Betsy being willing to consider

Yet here too the subject of the <u>get</u>-clause is deleted. Thus the
representation for (14) in the conventional theory is, roughly:

> 15. S_2(B. get herself into college) is too crazy for me
>
> S_3(I imagine S_4(B. willing S_5(B. consider
>
> S_6(B. get herself into college))))

While the subject of S_6 is commanded by another occurrence of
<u>Betsy</u>, the subject of S_2 is not.

There are, then, two approaches which could account for the
grammaticality of (14). The first is that the subject of S_2 can
be deleted because the subject of S_6 can be. In other words, S_2
acts as though it were in the position of S_6 - the sentence which
it deletes. The second approach would be to suppose that the fact
that the subject of S_6 can be deleted is irrelevant. Thus this
approach maintains that the deletability of the subject of S_2
can be characterized by considering only the relationship of the
deletion site within S_2 and the controller.

The crucial sentences which could decide between these two
approaches are difficult to construct. However, a parallel situ-
ation exists with Picture Noun Reflexives (PNRs) where the facts
are, fortunately, simpler. Thus Postal (1972) argues that a PNR
must be commanded by its antecedent; this would account for the
following contrasts:

> 16. a. That picture of himself indicated to John that he
>   was ugly
>   b. *That picture of himself indicated that John was
>   ugly
> 17. a. I gave that picture of himself to John
>   b. *I gave that picture of himself to the woman who
>   knew John

Yet again the antecedent need not command the PNR in a sentence
like:

> 18. That picture of himself is too ugly for me to give
>   to John

Again there are two possible explanations for this apparent
counterexample to Postal's claim. The first is that a PNR is

possible here because the subject in (18) deletes the object of
<u>give</u> and because a PNR can occur in the deleted object (as in
(17)). The second approach would maintain that Postal's claim is
incorrect and that the ability of a PNR to occur in the subject
of (18) can be characterized solely in terms of the relationship
of this NP and the occurrence of <u>John</u> in the lower clause.

There are two problems with the latter approach. First, not
any occurrence of <u>John</u> in the lower clause is sufficient to allow
a PNR in the matrix subject. Thus in contrast to (18) consider:

> 19. *That picture of himself is too ugly for me to give
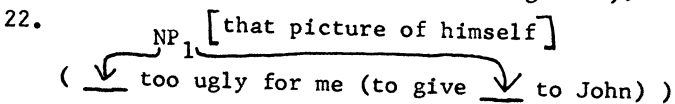>   to the woman who knew John

The contrast between (18) and (19) is clearly related to the con-
trast between (17a) and (17b). The ungrammaticality of (19) is
accounted for by the claim that a PNR can occur in the matrix
subject just in case it can occur in the NP deleted by this sub-
ject. Of course the claim that the ungrammaticality of (19) is
related to that of (17b) does not necessarily mean that (19) is
bad <u>because</u> (17b) is. But it is difficult to imagine a single
principle which accounts for both without reference to the rela-
tive positions of the deleted object in (19) and the PNR ante-
cedent.

Even more problematic for the claim that the conditions for
PNR can be described in terms of the relationship between the
subject in (18) and the occurrence of <u>John</u> in the lower clause
is the fact that PNR is impossible if there is no occurrence
of the picture noun phrase in the lower clause. In other words,
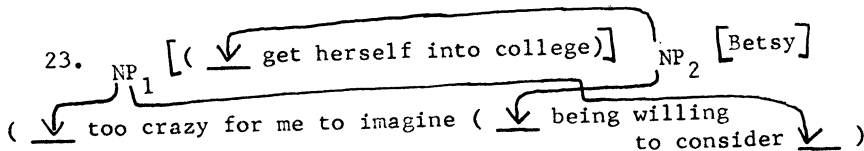consider the following contrast:

> 20. That picture of John is too ugly for me to give a
>   copy of the yearbook to him
> 21. *That picture of himself is too ugly for me to give
>   a copy of the yearbook to John

Notice that the positions of the PNR and the antecedent are iden-
tical in (18) and (21). Yet (21) is impossible.

Thus these contrasts indicate that a PNR is possible in the
matrix subject just in case it is possible in the NP deleted by
this subject; the deletion controller acts like it is in the po-
sition of the deletion site. These facts fall out in the MD
theory, since the deletion controller <u>is</u> in the position of the
deletion site; this theory does not require two occurrences of
the NP. Thus, for example, the representation of (18) would be
roughly (22) (where irrelevant details are ignored):

> 22.
>   NP$_1$ [that picture of himself]
>   (too ugly for me (to give to John))

Here <u>John</u> commands the PNR. Similarly, the representation for
a sentence like (14) would be roughly:

23. NP$_1$ [(↓ get herself into college)] NP$_2$ [Betsy]

(↓ too crazy for me to imagine (↓ being willing to consider ↓))

Principle (5) predicts that Betsy is not pronounced in the get-clause since it occurs in the willing-clause, and the S dominating willing dominates the get-clause.

No parallel argument can be constructed for a MD treatment of movement. The fact that a moved NP acts as though it is in its pre-movement position with respect to certain processes can be accounted for in the conventional theory by the fact that there is a derivational stage at which that NP is in the pre-movement position.

3. Extending the MD Theory to Pronouns

Given a theory incorporating a MD treatment of deletions it would make sense to extend this treatment to pronouns; this has been proposed by Sampson and by Postal. Thus this theory would claim that, in a sentence like:

24. John$_i$ said that he$_i$ left

there is a single NP node (dominating John) which is the subject of say and the subject of leave. This means that principle (5), which predicts that an NP is pronounced only in its highest clause, must be abandoned. Rather, we could assume that there is some principle to the effect that an NP cannot be pronounced only in the lower clause; it could, however, be pronounced in both. Thus whether or not an NP is pronounced in both clauses would be governed by lexical items. Verbs which, in the conventional theory, govern Equi would permit an NP to be pronounced only in one clause. Such an extension would also handle cases in which a moved NP leaves a pronominal copy, allowing for a unified treatment of a sentence like (1) and a sentence like:

25. John seems like he grows

However, pronouns don't, in general, have the property of deletion sites discussed above. That is, the antecedent of a pronoun does not act as though it were in the position of the pronoun with respect to processes like PNR. Consider for example:

26. *That picture$_i$ of himself is too ugly for me to give to the woman who asked John for it$_i$

If the same kind of link held between pronouns and antecedents as holds between deletion sites and deletion controllers, then (26) should be grammatical since (27), where the picture noun phrase is in the position of the antecedent, is grammatical:

27. The woman asked John for that picture of himself

Postal (personal communication) has pointed out that the

claim that no pronouns have this property is incorrect; the head noun of a relative clause also acts as though it were in the underlying position of the relative pronoun with respect to PNR. Thus the following is possible, despite the fact that <u>John</u> does not command the PNR:

> 28. the picture of himself which I asked John to take

Here the head noun behaves as though it were the object of <u>take</u>, where this is the underlying position of the relative pronoun. (In a MD theory of movement this pronoun remains in this position in surface structure as well.) Again it is not any occurrence of <u>John</u> in the lower clause which will permit a PNR in the head noun. Thus while (28) is grammatical, (29), like (30), is not:

> 29. *the picture of himself which the woman who knew
>   John took
> 30. *The woman who knew John took the picture of himself

It appears, then, that (28) is grammatical because (31) is:

> 31. I asked John to take the picture of himself

In other words, a PNR is possible in the head noun just in case a PNR is possible in the underlying position of the relative pronoun.

Thus this suggests that the MD theory whould be extended to relative pronouns, but not to a pronoun like that in (26). But it is not clear why these cases should be handled differently.[4]

4. Some Problems for the MD Theory

We can now consider two constraints which pose problems for the MD theory; I will discuss these first with respect to movement and then with respect to deletion. The first problem concerns the constraint discussed in Ross (1967) which blocks (32d):

> 32. a. I gave the book to Mary
>   b. I gave Mary the book
>   c. I gave it to Mary
>   d. *I gave Mary it

The exact formulation of the constraint is not important here. We can assume that there is some constraint which has the effect of blocking structures in which material intervenes between a verb and an underlying direct object pronoun (or, a prepositionless pronoun).

Ross claimed that the constraint holds for surface structure. However, Wasow (1975) argues that the constraint is actually cyclic. His argument is based on the contrast between a sentence like (33b) and (33d):[5]

> 33. a. It would be hard to tell the children those stories
>   b. ?*The children would be hard to tell those stories
>   c. *It would be hard to tell the children them
>   d. *The children would be hard to tell them

Although (33b) is somewhat bad, (33d) is significantly worse.
Yet this contrast is not accounted for in the conventional theory
if the constraint holds only for surface structure. If the moved
NP is entirely removed from the lower clause, then nothing inter-
benes between <u>tell</u> and the pronoun in the surface structure of
(33d).

Given this, Wasow claimed that the constraint is cyclic. This
predicts that (33d) is bad just as (33c) is. On the <u>tell</u>-cycle
<u>the children</u> intervenes between the verb and the pronoun; this
NP is removed in (33d) only by the application of Tough Movement
on the <u>hard</u>-cycle. In the MD theory, on the other hand, the
ungrammaticality of (33d) can be accounted for by a surface con-
straint. Here the NP <u>the children</u> remains in its lower clause
position in surface structure.

But both a cyclic formulation of the constraint and the MD
account break down in cases where the pronoun rather than the
intervening NP is Tough-Moved:

    34. a. It would be hard to tell the children those stories
       b. Those stories would be hard to tell the children
       c. *It would be hard to tell the children them
       d. They would be hard to tell the children

Here there is no contrast between (34b) and (34d). But a cyclic
constraint predicts that (34d) should be bad, just as (34c) is.
On the <u>tell</u>-cycle the pronoun is in the lower clause and <u>the
children</u> intervenes between the verb and this pronoun.

The MD theory also predicts that (34d) is ungrammatical since,
in surface structure, <u>them</u> is in the lower **clause.** Thus the sur-
face structure of this sentence is, roughly:

35.   NP$_1$ [the children] NP$_2$ [them]

    (hard (to tell   ))
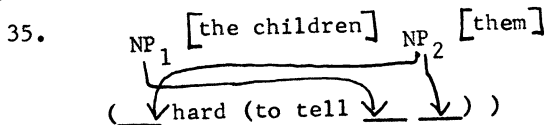
A similar problem is posed by the constraint which blocks
sentences like:

    36. *She$_i$ likes someone from Mary's$_i$ hometown

Postal (1970) notes that, in contrast to (36), a sentence in
which the object is moved by Wh-Fronting is grammatical:

    37. Who from Mary's$_i$ hometown does she$_i$ like?

The same contrast holds for cases involving Tough-Movement:

    38. *To talk to her$_i$ about Mary's$_i$ father would be hard
    39. Mary's$_i$ father would be hard to talk to her$_i$ about

Again the MD theory appears to predict that (37) and (39)
are ungrammatical, just as (36) and (38) are. The surface
structure of (39) would be, roughly:

40.

$$\text{NP}_1 \left[\text{Mary's mother}\right] \quad \text{NP}_2 \left[\text{her}\right]$$

(___ hard (to talk to ___ about ___))

Thus the surface structure of the <u>talk</u>-clause in (39) is identical to the surface structure of the <u>talk</u>-clause in (38).

Since the two cases above involve movement, and since the arguments for the MD theory involve deletion (and relative pronouns) it might appear that the solution is to adopt a MD treatment of deletion and not movement. However, these problems extend to deletion cases as well. Thus, for example, consider:

41. The man from Mary's$_i$ hometown told her$_i$ that it would be hard to shave himself

If the deletion controller (<u>the man from Mary's hometown</u>) were in the position of the deletion site, then (41) should be bad, just as (42) is:

42. *John told her$_i$ that it would be hard for the man from Mary's$_i$ hometown to shave himself

The head noun of a relative clause also does not act as though it is in the underlying position of the relative pronoun with respect to this constraint. Thus (43) is grammatical, despite the fact that (44) is not:

43. the man from Mary's$_i$ hometown who she$_i$ likes
44. *She$_i$ likes the man from Mary's$_i$ hometown

There are, of course, ways to state these constraints to avoid the problem. For example, the constraint blocking (32d) could be roughly as follows:

45. If an underlying direct object pronoun is pronounced in a position X which follows the verb, then nothing can intervene between the verb and the pronoun

Similarly, the constraint blocking (36) could account for the contrast between (36) and (37) in terms of the fact that the NP <u>Mary</u> is pronounced in the relevant position in (36), while in (37) it is not.

However, there is no apparent reason why the position in which an NP is pronounced should be crucial here. If the position in which an NP is pronounced is relevant to the formulation of grammatical constraints, then a constraint like (45) is no more natural then a constraint like:

46. If an underlying direct object pronoun is in a position X which follows the verb, then nothing can be pronounced between the verb and the pronoun

(46) makes exactly the wrong predictions - it predicts that (33d) is grammatical and that (34d) is not. Yet this constraint is as plausible as (45). Moreover, if the position of pronunciation plays a role in grammatical processes, then the fact that a dele-

tion controller behaves as though it is in the position of the deletion site is not automatic in the MD theory. There is no apparent reason why the condition for a PNR, for example, is not that a PNR is possible if it is pronounced in a position which is commanded by the antecedent.

5. A Modification of the Conventional Theory

We can now briefly consider a modification of the conventional theory which does account for the cases discussed in Sec. 4. This modification claims that, at the output of a movement or a deletion rule, some NP remains in the position of the moved or deleted NP, but this NP has no internal composition. An example of this with respect to movement is a theory in which a moved NP leaves a trace; I will assume here a theory where a trace is neither a pronoun nor a full NP. (An alternative view is that what remains is simply an empty NP node; it follows from this that this NP is neither a pronoun nor a full NP.)

In this theory, the contrast between (33d) and (34d) is accounted for by a surface constraint. Thus in the surface structure of (33d) there is a trace intervening between the verb and the pronoun:

47. *the children would be hard to tell t them

and hence (33d) is blocked. In the surface structure of (34d), on the other hand, nothing intervenes between the verb and a pronoun, but rather between the verb and a trace:

48. they would be hard to tell the children t

Similarly, a surface constraint would block a sentence like (38) but not (39). The surface structure of (39) would be:

49. Mary's$_i$ father would be hard to talk to her$_i$ about t

Since the movement of Mary's father leaves only a trace there is, in the surface structure of (39), no occurrence of Mary in the lower clause. This NP does, however, occur in the lower clause in the surface structure of (38).

6. Conclusion

The problem for the MD theory posed by the grammaticality of (34d), (38) and (41) is that here a moved NP or a deletion controller does not behave as though it were in the pre-movement position or in the position of the deletion site in terms of its internal composition. Thus a theory in which there is simply a trace or an empty NP in these positions in surface structure accounts for the grammaticality of these sentences, as well as accounting for the ungrammaticality of (33d).

Yet neither the conventional theory nor this modification of it account for the deletion facts discussed in Sec. 2; these fall out in the MD theory. This suggests that this theory is worth exploring further; perhaps some principle can account for

the difference between the cases discussed in Sec. 2 and those in Sec. 4.

Footnotes

1. In the system developed by Lakoff (1975) multi-domination is possible only when an NP is in two different clauses. One problem with this theory is that it still requires a deletion process to account for a sentence like (6), while (2) is accounted for by multi-domination.

2. As stated here, principle (9) makes little sense given the normal interpretation of precedence as a relation holding between nodes (not positions, since the notion position is not a primitive). This brings up some interesting formal problems that arise in a MD theory if precedence and dominance are relations between nodes. First, neither relation is irreflexive, nor are they asymmetric. Second, a principle like (9) cannot be formalized. As an approximation, we can consider the following two principles:

i. If NP_a precedes a node X and X precedes NP_a, then NP_a is pronounced to the left of X.

ii. If NP_a precedes a node X and X dominates NP_a, then NP_a is pronounced to the left of X.

Both (i) and (ii) account for (6) and (8); (i) predicts that <u>John</u> is pronounced to the left of the verb; (ii) predicts that <u>John</u> is pronounced to the left of the VP. But neither of these principles would account for a case in which, in a conventional theory, a deleted NP is a sister of and immediately preceded by the deletion controller. While such a situation might never arise in English it could arise in, for example, a VSO language. Thus if a VSO language allowed an object to be deleted by a subject (as in (6)), then neither (i) nor (ii) is adequate, since nothing intervenes between the subject position and the object position, and the subject does not precede a node which dominates the object.

Notice that a relational formulation of (9) avoids the problem; the principle can be stated as:

iii. If NP_a bears two relations X and Y to a V_x, where X is higher on the relational hierarchy than Y, then NP_a is pronounced as an X.

The problem does not arise here because grammatical relations, unlike positions, are primitives, and because the notion "higher on the hierarchy" holds between relations, not NP's. (An alternative relational theory would be one in which "higher on the hierarchy" were taken instead to be a relation between NP's, and where a notion like <u>subject</u> of a verb was defined as the NP bearing the highest relation to that verb. Here "higher on the hierarchy" is similar to "precedes" in a standard non-relational theory, and so the same formal problems will arise.)

This suggests then that a non-relational MD theory must include positions as primitives, where precedence would hold between positions.

3. Again a formal statement of (10) is difficult unless position is taken as a primitive. Here too (10) can easily be formulated in a relational theory:

    i. For any grammatical relation X, only one NP can bear the relation X to a V_x at a single derivational stage.

4. It might appear that the difference here has to do with an optional versus an obligatory coreference link. In other words, a relative pronoun and a deletion site like that in (18) must be coreferential to some particular NP. A pronoun like that in (26) need not have any particular NP as its antecedent; in fact it need not have any antecedent at all.

However, Stanley Peters has pointed out to me that this cannot be the salient property shared by relative pronouns and deletion sites as opposed to regular pronouns, since there are deletion sites which need not be controlled:

    i. The chickens are ready to eat
    ii. Those chickens are too ugly for me to eat

In both cases the deleted object of <u>eat</u> can be controlled by <u>the chickens</u>, but there is no obligatory coreference link here. Moreover, there are cases of deletion sites which must have controllers, but where there is more than one NP which could be the controller, as in:

    iii. John thought that watching that movie would annoy Mary

Here the subject of <u>watch</u> can be deleted either by <u>John</u> or by <u>Mary</u>.

5. I will be assuming that (33b) and (33d) are derived by Tough-Movement (not by a deletion rule).

References

Akmajian, A. (1972). "Getting Tough" in <u>Linguistic Inquiry</u> 3.

Fiengo, R. (1974). <u>Semantic Conditions on Surface Structure</u>, MIT Dissertation.

Lakoff, G. (1975). Unpublished memos on Dual Hierarchy Grammar.

Postal, P. (1970). "On Coreferential Complement Subject Deletion" in <u>Linguistic Inquiry</u> 1.

_____ (1972). "On Some Rules that are Not Successive Cyclic" in <u>Linguistic Inquiry</u> 3.

_____ and J.R. Ross (1971). "Tough-Movement Si, Tough-Deletion No" in Linguistic Inquiry 2.

Ross, J.R. (1967). <u>Constraints on Variables in Syntax</u>, MIT Dissertation.

Sampson, G. (1975). "The Single Mother Condition" in <u>Journal of Linguistics</u> 11.

Wasow, T. (1975). "Anaphoric Pronouns as Bound Variables" in <u>Language</u> 51.