morphological affixation possibilities and leads to a simplified account of the similarities between them and adjectival passives.

References

Department of Cognitive Science
School of Social Sciences
University of California
Irvine, California 92717

Antecedent Containment or Vacuous Movement: Reply to Baltin

Richard K. Larson, Robert May

Sentences like (1a–c) have been widely analyzed as involving an ellipsis site ([VP e]) that is located within its antecedent at surface form (Bouton (1970), Sag (1976), May (1985), Haük (1987), Larson (1987), Fiengo and May (forthcoming)):

(1) a. John [VP saw everyone that you did [VP e]].
   b. Eunice would [VP agree to anything I would [VP e]].
   c. Could Max [VP do something you couldn’t [VP e]]?

Such antecedent-contained deletions have been taken to provide evidence for a rule of Quantifier Raising (QR), which “disentangles” the elided phrase and its antecedent at a level of Logical Form (LF), thus allowing the relation between them to be stated in a noncircular way.

Baltin (1987) has proposed an alternative account of examples like (1a–c) in which antecedent-contained deletions do not exist as such at surface form; rather, the empty category uniformly occurs within a clause that has undergone string-vacuous movement (VM). No appeal to QR is made on this view, nor to any properties characterizing this rule. In this note we critically examine Baltin’s proposal, observing a variety of points on which the QR and VM analyses make different predictions. We show that in each

We are grateful to Mark Baltin, Noam Chomsky, Chris Collins, James Higginbotham, Ivan Sag, and an anonymous LI reviewer for comments on this work. Portions of this paper were presented at the 1988 annual meeting of the Linguistic Society of America, New Orleans, Louisiana.
case the predictions of the former are confirmed whereas those of the latter are falsified. We go on to argue that the VM analysis is not only empirically inadequate but also internally inconsistent. As it turns out, the type of derivation proposed by Baltin does not resolve antecedent containment but actually induces it at LF. The VM analysis is thus ruled out by the principles that rule out antecedent containment generally.

1. The QR and VM Analyses

The QR analysis of antecedent-contained deletion (ACD) sentences has been formulated in a number of ways. For concreteness, we adopt the version in May (1985), which derives from Sag (1976). This account assigns example (1a) the surface form (2a):

(2) a. \[ [Ip \text{ John PAST } [vp \text{ see } [np \text{ everyone that you PAST } [vp \text{ e}]]]] \]

Here the matrix VP contains an empty VP whose content it supplies. Given their containment relation, this content cannot be supplied at S-Structure without infinite regress, since reconstructing the VP headed by see would only bring along another occurrence of \([vp \text{ e}]\), which would itself require reconstruction, and so on. According to May (1985), this situation is resolved by QR at the level of LF. QR adjoins the object NP everyone that you did e to the clause, thus extracting \([vp \text{ e}]\) from VP:

(2) b. \[ [ip[np, everyone that you PAST [vp e]] [ip \text{ John PAST } [vp \text{ see } [np, t]]]] \]

This now permits noncircular reconstruction of \([vp \text{ e}]\), yielding a form that correctly represents the interpretation of this example:¹

(2) c. \[ [ip[np, everyone that you PAST [vp see [np, t]]] [ip \text{ John PAST } [vp see [np, t]]]] \]

The VM analysis of (1a) proceeds quite differently from that just sketched. Baltin (1987) assumes that although \([vp \text{ e}]\) is dominated by its ultimate reconstruction source at D-Structure, containment is broken up at S-Structure by string-vacuous extraposition of the relative clause. This results in the relative (CP) being attached to VP, as in (3a). For reasons we will return to, Baltin assumes that extraposition does not create traces or that the latter may delete. Hence, the final surface form is as in (3b), where \([vp \text{ e}]\) stands outside the VP that is its antecedent:

(3) a. \[ [ip \text{ John PAST } [vp[vp \text{ see } [np \text{ everyone } [cp, \text{ e}]]] [cp, \text{ that you PAST } [vp \text{ e}] ]] ] \]

(3) b. \[ [ip \text{ John PAST } [vp[vp \text{ see } [np \text{ everyone] } [cp \text{ that you PAST } [vp \text{ e}] ]] ] \]

¹ In Larson (1987) and Fiengo and May (forthcoming) it is assumed that a QR-ed quantifier can adjoin to any maximal projection and not just IP; in Larson (1987) it is furthermore proposed that reconstruction of VP (and, more generally, of XP) occurs at the point where the moved quantificational phrase has adjoined to VP (to XP).
2. Three Differences

The QR and VM analyses differ sharply on the surface form of ACD examples, the process responsible for breaking up antecedent containment, and the level at which this process applies. This yields a variety of divergent predictions about the form and interpretation of ACD sentences.

2.1. The Form of the Subordinate Clause

The VM analysis predicts that relative clauses in ACD sentences should look like extraposed relatives internally. The QR account predicts that they will look like in situ relatives. We can test this difference with the distribution of subordinators. It has been widely observed that whereas in situ relatives allow a wh-form, a complementizer, or no “clause introducer” at all (4)–(5), the last possibility is excluded with their extraposed counterparts (6)–(7):

(4) The man who [that $\emptyset$] Mary asked about finally showed up.

(5) I visited a man who [that $\emptyset$] John mentioned recently.

(6) The man finally showed up who [that $\emptyset$] Mary asked about.

(7) I visited a man recently who [that $\emptyset$] John mentioned.

Stowell (1981) suggests that this follows from the adjunct position of extraposed relatives, which prevents the trace of a deleted complementizer from being properly governed, in violation of the Empty Category Principle (ECP) (Chomsky (1981)).

Given this result, the QR analysis predicts that the full range of Comp possibilities will show up in ACD examples: wh, that, and null. By contrast, the VM analysis predicts that the null forms will be excluded. As (8)–(9) illustrate, the full range of forms is observed:

\[ 2 \]

A number of speakers judge ACD examples to be less acceptable with wh-forms than with that and null forms. This phenomenon may be related to what Carlson (1977) observes in relation to what he calls “amount relatives.” In examples like (i), where the relative clause variable ranges over quantities or amounts, a significant preference for non-wh forms is found:

(i) Ted saw all of Alice who [that $\emptyset$] he wanted to.

Whatever is behind this result, we note that its presence underscores once again the basic difference between the ACD phenomenon (which favors $\emptyset$ and that) and relative clause extraposition (which favors that and which).
(8) I visited everyone \(\{\text{who} \atop \text{that}\}\) you did e.

(9) John would reject any suggestion \(\{\text{which}\atop \text{that}\}\) Mary would.

These facts thus support the QR theory over the VM theory.

2.2. The Position of the Subordinate Clause

The VM analysis takes the relative clauses in ACD examples to be right-adjointed elements. The QR analysis takes them to be in situ phrases. Baltin (1987) cites data that appear to support the former view. He observes that ACD examples with a relative clause in sentence-internal position are often rather weak:

(10) a. John will find everyone that Bill finds easy to work with, easy to work with as well.

b. ??John will find everyone that Bill does ____ easy to work with.

(11) a. John will make every student that Bill makes take an extra exam, take an extra exam as well.

b. ??John will make every student that Bill does ____ take an extra exam.

Baltin takes these facts to favor a VM analysis, assuming that right adjunction for the relative clause containing [vp e] requires a right-peripheral position.

As it turns out, this assessment of examples like (10)–(11) and their bearing on the VM and QR theories is defective in two important ways. First, the characterization of the data is incomplete; the facts are more complex than the simple contrasts in (10)–(11) suggest. Second, given the possibility of multiple vacuous movement, the VM analysis turns out to predict only a right-adjointed position for ACD relatives, and not a right-peripheral one. Hence, the extent to which examples like (10)–(11) actually support the VM analysis is by no means clear.

2.2.1. Two Classes of Non-right-peripheral ACD. Close inspection of constructions like (10)–(11) reveals that they divide into two distinct classes: a class that yields ill-formedness absolutely, and a class for which judgments of ill-formedness vary according to the length of the material intervening between the gap and right periphery.

The first class involves sentences in which the quantified NP is in a position that is not properly governed. The central case is where NP is the subject of a tensed clausal complement. In all such examples, attempted ACD yields strong ill-formedness:

(12) a. I expect (that) [everyone you expect will visit Mary] will visit Mary.

b. *I expect (that) everyone you do ____ will visit Mary.

(13) a. I find (that) [everyone you find is qualified] is qualified.

b. *I find (that) everyone you do ____ is qualified.
(14) a. John believed (that) [everyone you believed was a genius] was a genius.
   b. *John believed (that) everyone you did ____ was a genius.
(15) a. I predicted (that) [no one you predicted has been a liar] has been a liar.
   b. *I predicted (that) no one you did ____ has been a liar.

The second class involves sentences in which NP is governed by the verb forming part of the reconstruction source. This includes examples where NP is a simple object complement, and examples where it is subject of a small clause or exceptional Case marking (ECM) structure. In such cases acceptability appears to hinge on the length of the material following the deletion site. Thus, whereas examples like (10b) and (11b), with extended material, are unacceptable or awkward, many shorter examples are judged to be quite good: ³

(16) a. I gave everyone that you gave two dollars, two dollars as well.
   b. I gave everyone that you did ____ two dollars.
(17) a. Tommy put everything he could put in his mouth, in his mouth.
   b. Tommy put everything he could ____ in his mouth.
(18) a. Max considers everyone that you consider smart, smart as well.
   b. Max considers everyone that you do ____ smart.
(19) a. Alice painted every barn that you painted red, red as well.
   b. ?Alice painted every barn that you did ____ red.
(20) a. Mo expects no one that you expect to visit Mary, to visit Mary as well.
   b. ??Mo expects no one that you do ____ to visit Mary.

The separation between these two classes is quite sharp, as illustrated by examples that differ only in the finiteness of their complement:

(21) a. *I expect (that) everyone you do ____ will visit Mary.
   b. ?I expect everyone you do ____ to visit Mary.
(22) a. *I find (that) everyone you do ____ is qualified.
   b. ?I find everyone you do ____ to be qualified.
(23) a. *John believed (that) everyone you did ____ was a genius.
   b. ?John believed everyone you did ____ to be a genius.
(24) a. *I predicted (that) no one you did ____ has been a liar.
   b. ?I predicted no one you did ____ to be a liar.

For us, the first member of each pair in (21)–(24) is completely unacceptable, whereas the second member is merely awkward at worst.

These results are problematic for the VM analysis. If the latter permits only ACD examples with a right-peripheral relative, then it will flatly exclude acceptable cases like

³ (17) and similar well-formed “internal” examples are noted in Carlson (1977); (18) was pointed out to the first author by R. Kayne as a well-formed case of ACD.
the (b) examples of (16)–(24). Furthermore, the VM analysis yields no clue to why the (a) examples of (21)–(24) should be judged so much worse than their untensed counterparts, since the source of ill-formedness is putatively the same in both cases.

By contrast, we believe the QR account allows a plausible approach to these data. We suggest that ill-formedness in ACD examples of the first class reflects familiar constraints on quantifier extraction in finite versus nonfinite clauses. As discussed by Borkin (1984) and May (1985), whereas quantified subjects can be given scope out of infinitives, this is not generally possible with tensed complements. Thus, whereas (25a) permits a wide-scope reading for everyone vis-à-vis someone and believe, according to which for each person x there is someone who believes x is a genius, (25b) permits only a narrow-scope reading for everyone, according to which there is some person who believes genius to be a universal characteristic:

(25) a. Someone believes everyone to be a genius.
   b. Someone believes (that) everyone is a genius.

Recall now that [VP e] can receive a proper reconstruction in the (b) examples of (21)–(24) only if QR is permitted to extract the quantified subject of a tensed embedded clause. (23a), for instance, requires the LF representation in (26):

(26) [IP[NP, everyone you PAST [VP e]] [IP John PAST believe [CP e, was a genius]]]

However, if quantified subjects are clause-bound in finite sentences, such LF representations will be impossible, and reconstruction will be blocked. We thus rule out our first class.

In contrast to the first class of ACD examples, we suggest that the (b) examples of (16)–(24) are fundamentally grammatical and that their length-sensitive unacceptability reflects performance limitations. Note that in all such cases, material for the reconstructed VP must be recovered from both the left and the right syntactic context of the empty category [VP e]. Supposing that the various “leveled” representations for a sentence are computed together “on line” as the sentence is processed, it follows that (27a) and (27b) will differ importantly at the point where [VP e] is encountered. Specifically, (27a) will require the resolution of [VP e] to be “held” until further VP material is received, whereas (27b) will not. This is rather analogous to what occurs with center-embedded examples like (27c), where interpretation of the inner relatives must be held pending the contribution of a larger domain. Here, as in (27a), the result is reduced acceptability:

(27) a. I consider [everyone you do [VP e]] smart.
   b. I consider smart [everyone you do [VP e]].
   c. ??The rat [the cat [the dog worried] chased] ate the cheese.
2.2.2. Persuade versus Believe. The proposal made above appears to be compromised at first by a contrast noted in Baltin (1987). Baltin observes that pairs like (28a,b) involving persuade and believe differ in acceptability, with the first, for Baltin, noticeably better than the second:

(28) a. I persuaded everyone you did [VP e] to be polite.
   b. I believed everyone you did [VP e] to be polite.

This contrast appears difficult to explain on a processing account since the material intervening between [VP e] and the right periphery is the same in both cases. On the other hand, according to Baltin (1987), the difference is correctly predicted under the VM theory. Briefly, in the case of persuade the material to the right of everyone comprises two constituents (a relative and an infinitival complement), each of which may extrapose to the periphery of VP. This allows a non-antecedent-contained S-Structure representation corresponding to (28a):\(^4\)

(29) a. [IP
       [NP
           I
           I [PAST]
           VP
           VP
           CP CP\(_i\)
           PRO to be polite
       V NP CP\(_i\) you did [VP e]
       persuade everyone e]

On the other hand, in the case of believe the material to the right of everyone comprises two constituents, of which I' cannot be extraposed. Hence, (28b) has only an antecedent-contained S-Structure representation, which Baltin assumes to be disallowed:

\(^4\) For simplicity, the trace of relative clause extraposition has been deleted, following Baltin (1987); however, see section 3 for arguments that such traces must actually be present.
When the contrasts in (28) and the tree in (29a) are examined carefully, it turns out that their consequences do not actually pose a problem for the QR theory. Indeed, perhaps surprisingly, these data actually provide further support for the QR analysis over the VM analysis. To see why, consider (29a) again. Note that since the matrix VP contains a CP-trace bound by PRO to be polite, when it is copied into [VP e], the latter will also contain a CP-trace bound by PRO to be polite. This means that (28a) is predicted to be synonymous with (30a). But this prediction is not correct. On its most natural reading, (28a) is not synonymous with (30a). Rather, it is synonymous with (30b):

(30a) a. I persuaded [everyone that you persuaded to be polite] [to be polite].

b. I persuaded [everyone that you persuaded] [to be polite].

(30a) and (30b) differ crucially in the class of individuals claimed to have been influenced. In (30b) it is the general class of individuals-subject-to-persuasion. In (30a) it is the specific class of individuals-who-can-be-persuaded-to-be-polite. We judge (28a) to be ambiguous between (30a,b) but to strongly prefer the reading (30b). More precisely, we judge (28a) on its nonpreferred reading (30a) to be equally as acceptable as (28b).

The upshot of these remarks is thus the following. The QR theory predicts (28a,b) to be accessible to the same degree when relativized to equivalent readings. By contrast, the VM theory predicts the example with persuade to be grammatical and the example with believe to be ungrammatical. Reflection on (28)–(30) shows that the QR theory’s prediction is the correct one, despite initial impressions; (28a,b) are in fact equally available when relativized to equivalent readings. This result is simply obscured by the presence of an additional reading for (28a) that is not available for (28b). This reading is
responsible for the greater apparent acceptability of the former; however, this reading is not the one predicted by the VM theory through (29a). Absence of the expected difference in acceptability for (28a,b) thus makes these sentences a problem for the VM account and not evidence in support of it.\(^5\)

2.2.3. The Problem of Multiple Vacuous Movement. Although Baltin (1987) cites examples like (10)–(11) in support of the VM analysis, there is in reality a significant gap between data and conclusion. This gap arises from the possibility of moving several categories vacuously in one and the same derivation.

Baltin (1987) observes that ACD examples are possible not only with full relatives but also with free relatives as in (31a) and (32a). Noting that these are not plausibly analyzed in terms of relative clause extraposition, he suggests that they involve vacuous “Heavy NP Shift,” as in (31b) and (32b), which also adjoins a category rightward:\(^6\)

(31) a. I said whatever you did.
   b. I said e \[\text{NP whatever you did} \[\text{VP e}\]]

(32) a. Max visits whoever he can.
   b. Max visits e \[\text{NP whoever he can} \[\text{VP e}\]]

By analogy, we may surmise that Baltin would analyze “free comparatives” such as (33a) and (34a) (discussed in Larson (1987)) as involving vacuous rightward AP Shift, as in (33b) and (34b):

(33) a. John will grow however tall his father did.
   b. John will grow e \[\text{AP however tall his father did} \[\text{VP e}\]]

(34) a. Bill’s party will last however long Fred’s does.
   b. Bill’s party will last e \[\text{AP however long Fred’s does} \[\text{VP e}\]]

It follows that the VM theory must invoke a full range of vacuous movements including at least vacuous CP movement (relative and comparative clause extraposition), vacuous PP movement (PP extraposition), vacuous NP movement (Heavy NP Shift), and vacuous AP movement (“Heavy AP Shift”).

Continuing, note that given the well-formedness of examples like (35a), the VM analysis is obliged to permit multiple vacuous movements in a single derivation. In view of the interior position of the relative clause containing the empty VP, (35a) must be analyzed as involving a double instance of vacuous movement, where both relatives are adjoined to VP, as in (35b):

\(^5\) It should be pointed out that the natural reading of (28a)—the one synonymous with (30b)—is not presently predicted under either the QR or the VM analysis and thus represents a challenge for both. Nonetheless, on the range of data for which both theories make a definite prediction, the data in (28) clearly support the QR analysis over the VM analysis.

\(^6\) Larson (forthcoming) proposes an alternative account of the “NP Shift” phenomenon in which the relevant examples derive, not by rightward movement of an NP, but rather by leftward raising of a predicate. This analysis appears compatible with the QR theory of ACD.
(35) a. I interviewed everyone you did [VP e] that wanted a job.
b. I [VP [VP interviewed [NP [everyone e]] e]] [CP, you did [VP e]]
[CP that wanted a job]].

Observe now that given the inventory of possible vacuous movement rules, and
the possibility of multiple vacuous movements in a single derivation, it will actually
be possible to generate various non-right-peripheral ACD examples under the VM ac-
count. For example, without principles that specifically exclude it, Baltin’s (10b) (re-
peated below) can be given the analysis in (36), where the small clause subject has
undergone vacuous NP Shift and the small clause predicate has undergone vacuous AP
Shift, resulting in a doubly adjoined structure similar to (35b):

(10) b. *John will find everyone that Bill does ___ easy to work with.
(36) John will [VP [VP find [SC t, t]] [NP, everyone that Bill does [VP e]]] [AP, easy
to work with]].

Thus, despite their recruitment in support of the VM analysis, data like those in
(10)–(11) do not clearly follow from the VM proposal. As it stands, this proposal does
not predict a right-peripheral site for vacuously moved elements, but only a right-adjoined
one.

2.3. The Interpretation of [VP e]
The QR and VM analyses make sharply different predictions about the interpretation of
the empty VP node in antecedent-contained deletion structures. Under the former,
[VP e] is expected to take various possible VP antecedents depending on the scope of
the quantified NP containing it. Under the latter, [VP e] is predicted to take only a single
possible antecedent owing to the Generalized Subjacency constraint of Baltin (1987):

*Generalized Subjacency
In the structure [a . . . [β . . . Y . . . ] . . . ] . . . X, no rule may move an element
Y to position X if α and β are maximal projections.

This principle bounds rightward movement, permitting an extraposed relative to adjoin
no higher than the IP or VP immediately dominating its associated NP. As a result,
[VP e] can take only the VP in its immediately containing clause as antecedent.7

2.3.1. ACD and Scope. Data involving ACD in intensional contexts appear to support
the QR theory. Consider (37):

(37) John wants to visit every city in Italy.

As is well known, such examples permit a de dicto and a de re reading of the quantified

7 It should be noted that Generalized Subjacency does not actually provide a sufficient account of the
boundedness of rightward movement even in Baltin’s own analysis. In addition to Generalized Subjacency,
an independent (and ad hoc) stipulation is required to the effect that rightward adjunctions do not iterate. For
an alternative account of this phenomenon, see Guéron and May (1984).
NP. These readings correspond to the two scopes NP may assume vis-à-vis the inten-
sional verb want. On the former, narrow-scope reading, John has a single, general desire
to visit Italian municipalities. On the latter, wide-scope reading, John has a desire to
visit x, for each specific Italian city x.

Consider now (38a–c), parallel to (37), where the first involves antecedent-contained
deletion:

(38) a. John wants to visit every city you do e.
b. John wants to visit every city you visit.
c. John wants to visit every city you want to visit.

Observe that (38a) is ambiguous with respect to how [vp e] is understood. More spe-
cifically, note that it shares the de dicto reading of (38b), according to which John has
a general desire to visit all cities that you visit. It shares the de re reading of (38b), on
which John has a desire to visit x, for each specific city x that you also visit. And it
shares the de re reading of (38c), on which John has a desire to visit x, for each specific
city x that you also want to visit.

This example shows exactly the pattern of data expected under the QR theory. For
not only does it display ambiguity in the interpretation of [vp e], it also shows the specific
correlation of [vp e] interpretation and NP scope predicted on the QR account. Under
the latter, we expect a derivation in which NP remains within the embedded clause,
forcing [vp e] to take the lower VP as its antecedent, and forcing NP to be read de dicto.
In addition, we expect a derivation in which NP first adjoins in the lower clause, where
[vp e] is reconstructed, and in which NP then moves on to adjoin in the matrix clause
where it is read de re. Finally, we expect a derivation in which NP moves to the upper
clause (and hence is read de re) and in which [vp e] is also reconstructed in the upper
clause. This is the range of interpretations that (38a) shows; hence, these data strongly
support the basic claim of the QR theory that antecedent-contained deletion is crucially
tied to NP scope.

With respect to the de re construals described above, it should be pointed out that
there are certain independent factors affecting whether it is the embedded or matrix VP
that is reconstructed. Thus, contrast the examples in (39):

(39) a. John wants to visit a certain city that you do.
b. John wants to visit a certain city that you did.

We presume that the use of certain here forces de re readings for each of these sentences.
Note, however, that their most natural construals differ markedly. In (39a) the empty
VP is construed with matrix VP headed by want; in (39b) it is construed with the em-
bedded VP headed by visit. The relevant factor here, apparently, is the tense of the Infl
element preceding the ellipsis site: present tense induces a broad reconstruction, whereas
past tense induces a narrow reconstruction. Although it is not entirely clear to us why
tense should have just this effect, the importance of (39a,b) for present concerns is that
they show clearly that it is structurally possible for either VP to be reconstructed when
the quantifier phrase containing the ellipsis site has wide scope, since along the relevant structural parameters (39a) and (39b) are indistinguishable. When the quantifier phrase has narrow scope, on the other hand, only the embedded VP can be reconstructed, as predicted under the QR theory.⁸

Other cases that require broad scope for an embedded phrase show parallel properties; perhaps the clearest cases are to be found with multiple wh-constructions. Thus, the readings available in (40) parallel those of (39):

(40) a. Which student wants to visit which city that you do?
b. Which student wants to visit which city that you did?

Following standard proposals, we assume that the in situ wh-phrase undergoes LF movement, taking scope identical to that of the overtly moved wh-phrase. For the examples in (40), this implies that the embedded wh-phrases take scope in the matrix clause, the site of which student. Moreover, this movement licenses ACD, so that the empty VP can be construed, as above, with either the matrix VP (40a) or the embedded VP (40b). Again, this is precisely as expected under the QR account (modulo the tense constraint).

Baltin notes examples parallel to (40b) (see his example (14)), which appear to lack a matrix reconstruction for the empty VP. From this he concludes that the QR account predicts a nonexistent construal. The VM theory, on the other hand, purportedly accounts for these data by bounding extraposition to the embedded clause, hence making the matrix VP unavailable as a reconstruction source. Baltin’s conclusions are abrogated,

---

⁸ Mark Baltin has suggested to us that this scope argument may be compromised by the fact that extraposed relatives can escape infinitival complements to want. He points out that the extraposed relative in (i) must be viewed as positioned in the matrix clause on the reading where the temporal adverb modifies the matrix verb:

(i) I wanted to visit someone yesterday who I really liked.

The suggestion is that readings where [vp e] takes a matrix antecedent are ones in which the extraposed relative has a matrix attachment.

We make three points in response. First, it appears to us that examples parallel to (i) can be constructed that make the point opposite to Baltin’s. It seems to us that (iia) has a reading equivalent to (iib), where [vp e] takes the matrix VP as antecedent, but where the position and interpretation of the adverb compel us to view the relative as falling within the complement:

(ii) a. I want to visit a certain city you do [vp e] tomorrow.
b. I want to visit tomorrow a certain city you want to visit tomorrow.

The fact that extraposed relatives may apparently escape infinitives is thus insufficient to explain the facts of ACD.

Second, appeal to the extended extraposability of relatives in infinitives does not address the correlation noted above between the construal of [vp e] and the de dicto/de re construal of the NP to which the relative is attached.

Third, the fact that extraposed relatives escape infinitives appears to us to raise serious problems for Baltin’s general account of rightward movement; evidently, Baltin must either weaken his Generalized Subjacency condition to permit movement over two maximal projections when one is a nonfinite IP or else allow iterated rightward adjunctions of CP.

It is worth observing that the long-distance extraposability of relatives, apparently supported by (i), is problematic in other respects. For example, such relatives cannot be stranded by VP ellipsis, unlike other adjuncts:

(iii) a. June wanted to leave because she was angry and Mary wanted to [vp e] because she was sick.
b. *John wants to visit every city that you visit and Bill wants to [vp e] that Mary visits.
however, by his failure to observe sentences like (40a), which complete the paradigm of reconstruction expected under the QR theory and which directly counterexemplify the VM theory.9

2.3.2. Reconstruction and Logical Form. We conclude this discussion of the interpretation of \([vp \ e]\) by pointing out that the VM analysis retains an important dependence on the rule of QR and hence on the level of LF, despite intimations to the contrary in Baltin (1987). This is because the extraposed structures provided by the VM theory cannot represent the meaning of ACD sentences correctly without QR. Consider the VM structure (41b), corresponding to (41a):

(41) a. John saw everyone that you did.

b. 

```
  IP
  /  \
 NP  I'
  /    \
 John I  VP
         /  \
 VP  CP
     /   \
 V   NP
   see  everyone
```

If reconstruction is made directly from this source, the result of copying VP will be (42), which is completely ill-formed:

(42) *John saw [everyone \([cp \ O, \ e]\) that you saw everyone].

---

9 Baltin's actual example is given in (i), which differs from the cases in the text both in being a mixed 'person-number' question and in having a tensed complement:

(i) Who thought that Fred read how many of the books that Bill did?

For many persons (including ourselves), such mixed questions are decidedly odd, and even when simple examples like (ii) are intuited as well-formed, speakers often cannot state what an appropriate answer would be:

(ii) ??Who read how many books that Bill did?

Controlling for this, we can reconstitute Baltin's example as in (iii); note that it now contrasts with (iiib) in the expected manner:

(iii) a. Which student thinks that Bill visited which city that you do?

b. Which student thinks that Bill visited which city that you did?
Intuitively, the source of this ill-formedness is clear: CP is a relative clause containing an operator (here the empty operator O) that must bind a variable. In (42) the null relative clause operator (O₀) fails to bind a variable after reconstruction; hence, the structure involves a form of “vacuous quantification,” which we take to be generally unavailable (following Chomsky (1986)).

Proper reconstruction can only be made at LF, after QR has extracted the quantified object:

Only now will VP-copying yield a VP with an internal trace for O₀ to bind, as required.

3. What Excludes the VM Analysis?

The data reviewed above argue strongly that the QR theory of ACD examples is correct. In view of this, it is useful to consider what principles exclude the VM analysis. We believe the correct answer is one given by May, recorded in Baltin (1987, sec. 4.5). May points out that the VM analysis has the surprising consequence of inducing the very antecedent-contained structures that it is supposed to eliminate.

We noted earlier that Baltin (1987) assumes that extraposition does not create traces. This assumption is in fact a vital one, as shown by the following variant of (41) in which a CP-trace appears:
As observed above, QR must apply if we are to get a form appropriate for interpretation. If QR moves the largest NP category in (44), the result is (45a), where the CP-trace fails to be properly bound:

(45) a.

Suppose then that QR extracts the smaller NP, resulting in (45b):
LF copying of the matrix VP into the empty VP then yields (45c), where antecedent containment continues to hold:

In (45c) \([\text{CP}_e, e]\) lies within its own antecedent \(\text{CP}_e\); hence, the structure is ill-formed. Because of this result, Baltin (1987) is compelled to assume that extraposition creates no traces. By way of justification, Baltin states that no principle of grammar requires
extraposition to leave a trace. And he asserts that no direct evidence exists for its doing so. Neither claim is correct, however. As it turns out, there is a large class of ACD examples for which an extraposition trace is required under the Projection Principle of Chomsky (1981). This is the class of comparative constructions (46) and the class of comparative-like constructions involving same and different (47).\footnote{Constructions involving same and different are discussed by Carlson (1977; 1987).}

(46) a. John grew as tall as Bill did.
    b. Max worded the letter more carefully than Eunice did.
    c. Lisa could have behaved less rudely than she did.
    d. Alice will surely find as many marbles as Felix will.
    e. Phil interviewed more generals than Mike did.

(47) a. We visit the same people that/as you could.
    b. Brian couldn’t possibly take a different train than you do.

In comparatives, unlike relatives, the clausal complement is \textit{not} an optional modifier. Rather, it is a selected complement of the degree morpheme. This status is clear from the fact that the degree morphemes \textit{(more, \textit{-er}, and \textit{less} versus \textit{as})} and the nouns \textit{same} and \textit{different} govern complementizer form in the clausal complement \textit{(than versus as)}. And it is clear from the semantics of comparatives, in which the comparative complement functions as an argument of the degree word (see Klein (1988) and Larson (1988a) for recent discussion). Given this, it follows that for a large class of cases, the VM analysis will never yield an interpretable LF since it \textit{will invariably induce illicit antecedent containment at LF}.

To illustrate, (46a) receives the LF representation in (48a), with a CP-trace that is required under the Projection Principle:

(48) a.
Reconstruction of \([_{VP} e]\) now reintroduces the illicit structure of antecedent containment, resulting in an uninterpretable LF representation:

\[
\begin{array}{c}
(48) \ b. \\
\text{IP} \\
\text{AP} \\
\text{as tall} \\
\text{NP} \\
\text{John} \\
\text{I'} [\text{PAST}] \\
\text{VP} \\
\text{CP} \\
\text{as Bill PAST} \\
\text{VP} \\
\text{V} \\
\text{grew} \\
\text{AP} \\
\text{e} \\
\text{CP} \\
\text{e} \\
\text{VP} \\
\text{V} \\
\text{grew} \\
\text{AP} \\
\text{e} \\
\text{CP} \\
\text{e}
\end{array}
\]

Since extraposition traces cannot be eliminated for comparatives and comparative-like constructions involving \textit{same} and \textit{different}, it is worth considering whether they are in fact eliminable in the case of relative clauses. There is some evidence that they are not. We observed earlier that the VM analysis permits multiple rightward movements in one and the same derivation. Consider example (49a) in this light, which involves a relative clause (\textit{who I'd never met before}) modifying a “‘heavy NP” (\textit{a man in a pink hat}):

\[
(49) \ a. \ I \text{ saw a man in a pink hat who I'd never met before.}
\]

Let us ask what blocks vacuous extraposition followed by “‘Heavy NP Shift’” in this example, yielding the ill-formed (49b):

\[
(49) \ b. \ *I \text{ saw who I'd never met before a man in a pink hat.}
\]

If extraposed relatives leave traces, the answer is straightforward: (49b) will involve a
VP configuration in which CP fails to c-command and hence to properly bind its trace:  

(49) c.

```
  VP
   /\  
  /   \  
VP  CP  NP  
  /\   /\  
  V  NP  CP  
  see  who I'd never met before a man in a pink hat e
```

On the other hand, if extraposed relatives leave no trace, or if these traces delete, then the ill-formedness of (49b) is obscure. In fact, structure (49d) is analogous in relevant respects to the desired structure (43) discussed earlier:

(49) d.

```
  VP
   /\  
  /   \  
VP  CP  NP  
  /\   /\  
  V  NP  CP  
  see  who I'd never met before  
```

Hence, there seems to be evidence that extraposition does in fact leave traces even in the case of relative clauses.

Taking these points together, then, we have shown that there are indeed principled grounds for excluding vacuous movement as an account of ACD examples. Since extrapositions leave traces—undoubtedly so in the case of comparatives and arguably so in the case of relatives—a VM analysis will always induce at LF the very configuration that it was supposed to eliminate at S-Structure. The infinite regress that rules out S-

---

11 Under the definition of c-command introduced in May (1985), adjoined quantifiers have scope over all elements within the smallest maximal projection containing them, where one category X contains another Y iff all segments of X dominate Y (see May (1985) for details). We assume here that this definition applies only to adjoined quantificational elements and not, for example, to base-adjointed modifiers or adjoined clauses. Hence, CP fails to c-command its trace in (49c), even under May's extended definition of c-command. For a justification of this restriction, see Larson (1988b).
Structure reconstruction of antecedent-contained anaphors thus also rules out the VM analysis of Baltin (1987).

References


(Larson)
Department of Linguistics
State University of New York at Stony Brook
Stony Brook, New York 11794

(May)
School of Social Sciences
University of California
Irvine, California 92717