Antecedent-Containment and Ellipsis

Chris Wilder

Max-Planck-Gesellschaft, Arbeitsgruppe Strukturelle Grammatik
Jägerstr. 10-11, D-10117 Berlin, Germany
chris@asg.ag-berlin.mpg.de

0. Abstract

This paper [1] examines old and new data relating to the Antecedent-Contained-Deletion (ACD) phenomenon in English. The primary theoretical interest of ACDs lies in the evidence they provide for the assumption of LF movement processes - A'-movement (Quantifier Raising - May 1985 et. al.), or A-movement of objects in English (Hornstein 1994). There is one class of examples involving 'wide scope' VP-ellipsis, as in *John said that more trees had died than Mary did*, that has so far been overlooked. It is argued here that these examples must be included in the paradigm to be accounted for. The extended paradigm is then shown to support the following claims:

(i) ACDs do not support the assumption of covert A-movement of objects in English, contra Hornstein (1994).

(ii) ACDs do provide evidence for a covert A'-movement operation of the type identified as QR by May (1985) and others.

(iii) In addition to the ban on containment of a VP-ellipsis site within its antecedent VP at LF, the account of ACDs must refer to an independent principle that prohibits containment at PF of a VP-ellipsis site within the terminal string corresponding to its antecedent VP.

(iv) Certain examples have been mis-identified as ACDs arising from VP-ellipsis, which actually involve two adjacent deletions, neither antecedent-contained in surface order, one due to Pseudogapping (ellipsis of subparts of VP), the other due to the independent process of Backward Deletion (a.k.a. 'non-coordinate' Right Node Raising).

Furthermore, properties of ACDs are uncovered which require claims in May (1985), Larson & May (1990) about constraints on QR to be reconsidered.

1. The ACD phenomenon

1.1 VP ellipsis and antecedent containment

The antecedent contained deletion phenomenon forms a subcase of the VP ellipsis phenomenon illustrated in (1-2):

(1) a. John met Mary, and Bill did, too.
b. Although John met Mary, Bill didn't.
c. Although John did, Bill didn't meet Mary.
d. John met everyone that Mary did.
e. John said that Mary did.

(2) a. John [$_{VP2}$ met Mary], and Bill did [$_{VP1}$ e] , too.

b. Although John [$_{VP2}$ met Mary], Bill didn't [$_{VP1}$ e] .

c. Although John did [$_{VP1}$ e], Bill didn't [$_{VP2}$ meet Mary] .

d. John [$_{VP2}$ met everyone that Mary did [$_{VP1}$ e]] .

e. John [$_{VP2}$ said that Mary did [$_{VP1}$ e]] .

A VP ellipsis site (VPE), such as VP1 in the examples in (2), can be viewed as an anaphoric element that depends on the content of an antecedent VP for its interpretation. A VPE may take as its antecedent a VP in the same sentence, which may be arbitrarily distant from the VPE; alternatively, the antecedent may be a VP in the external discourse, which need not be linguistically explicit (cf. Sag (1976), Williams (1977)). Embedded in some suitable discourse, all examples in (2) are well-formed, with VP1 taking its antecedent from that discourse. In examples (2a-c), VP1 may take a VP in the same structure (i.e. VP2) as its antecedent, with no issue of antecedent containment arising. The special case of antecedent contained deletion is represented by (2d,e): VP1 is contained within VP2 and VP2 is understood as the antecedent of VP1. The puzzle that arises is that while (2e) is ill-formed (the intended reading is simply impossible), (2d) is not.

In section 2., three accounts of the contrast between (1d,e) under the 'antecedent-contained' reading are reviewed: May (1985) and Larson & May (1990); Baltin (1987); and Hornstein (1994). All these accounts depend on the assumption that (2) holds at some level (whether as axiom or theorem):

(3) A VPE may not be contained in its antecedent

It may be thought that there are semantic reasons for the ill-formedness of structures violating (3), namely that there is no finite interpretation that can be assigned. In (2e), for instance, the content of VP2 ('what John did') is dependent on the content of VP1 ('what Mary did'), but since the content of VP1 is in turn dependent on what John did, a circular reading results - similar to the effect found in (presumably syntactically well-formed) sentences like (4) under the reading indicated (cf. Haik 1987 for discussion):

(4) (*) [her$_j$ brother]$_k$ likes [his$_k$ sister]$_j$

A purely syntactic account of VP ellipsis that makes no reference to interpretation is also possible. For example, it is clear that the generation or licensing of VPE-constructions is contingent on some notion of syntactic identity between the the VPE and its antecedent. (3) could be made to follow simply from the requirement that VPE must be identical with its antecedent, under a plausible definition of syntactic identity. [2] There is, for example, no way for VP1, properly contained within VP2, to be identical to VP2 in terms of depth of embedding: VP2 is necessarily "deeper" than VP1. Hence, it is possible that no interface representation containing an antecedent-contained VPE that meets the identity requirement will be generable, so that questions of interpretation do not even arise.

Be that as it may, (2d) clearly is well-formed and does have a finite interpretation. Accounts of how that reading come to be possible take the form of assuming that the example has (at the relevant level) a syntactic representation different from (2d), which avoids violating (3). Details depend on general assumptions about how VPE's and their interpretation are analyzed. There are two main competing approaches: choice between these is essentially orthogonal to the issues surrounding ACDs. [3] In the proposal of Williams (1977) (cf. also May 1985, Hornstein 1994), a VPE is generated as a node lacking internal structure, dominating a designated phonologically null

terminal - [$_{VP}$ e] - which occurs in S-structures that feed PF (whence the gap). Following S-structure and prior to LF, the content of the antecedent VP is 'copied in' by a syntactic operation: a copy of the antecedent VP is substituted for [$_{VP}$ e]. The major argument for this approach lies in its ability to account for a range of parallelism effects governing the interpretive possibilities available for VPEs.

The main alternative consists in assuming that the VPE gap in PF arises through phonological deletion, i.e. material generated as a normal VP is deleted under identity with material belonging to the antecedent in the PF-wing of the grammar. To ensure the correspondence between deletion in PF and interpretive dependence at LF, material so deleted has to have been marked in some way in the syntax already, with this marking indicating dependence on its antecedent with regard to interpretation at LF (cf. Wilder 1995a for discussion).

In the context of the Minimalist Program, the latter approach is more natural, avoiding two extra assumptions: (i) the LF-copy operation, which is not reducible to Chomsky's (1995) "Move" or "Merge"; (ii) the addition of the designated terminal to the class of lexical items (given the restrictive hypothesis that structure is only licensed via projection of items from the lexicon). The 'PF-deletion' operation required in this alternative is independently assumed in the Minimalist model as the mechanism for trace-gap creation. Additionally, in examples like (5) (cf. also (2d)), the presence at S-structure of a wh-phrase in Spec,CP is not naturally combined with the assumption that VP1 dominates its trace only after LF-copying of VP2:

(5) I know who Mary [$_{VP2}$ met *t*], but I don't know who John did [$_{VP1}$ e]

More importantly, the fact that interpretive parallelism effects observed in coordinative VPE-constructions like (1a) are independently attested in full structures in which a VP is spelled-out but marked as dependent by deaccenting, means that some mechanism to ensure parallelism must be assumed independently of VPE. This removes the argument in favour of adopting the LF-copy mechanism (cf. Tancredi 1992, Chomsky 1993, Fox 1994; Fox also gives further arguments against the LF-copy approach).

In the following, I shall assume a PF-deletion approach to VPEs to be correct, although nothing hangs on this decision; in reviewing previous accounts, discussion is framed in terms of the LF-copying approach where appropriate.

Assuming the LF-copy theory, and assuming the structure (2e) as input to the copy-operation, there is a syntactic reason for the fact that VP1 cannot take VP2 as its antecedent. The output of copying results in a structure (6b) which contains a second instance of VP1, requiring to be replaced by a copy of VP2. The derivation will thus slip into infinite regress, a well-formed LF cannot be derived:

(6) a. John [$_{VP2}$ said that Mary did [$_{VP1}$ e]]
b. John [$_{VP2}$ said that Mary did **[$_{VP2}$ say that Mary did [$_{VP1}$ e]]]**
c. *etc.*

Under the PF-deletion approach, and given the merger theory of structure generation (Chomsky 1993, 1995), the example can also be excluded as not derivable. Assume that VP1 and VP2 must be structurally identical, i.e. syntactic copies of each other according to some measure of identity, for a VPE-dependency to arise. Then no representation corresponding to (6a) can be generated: VP2 cannot be generated independently of VP1 and be structurally identical to VP1, since VP2 must also properly contain VP1.

The conclusion with respect to (2d.) is that the structure given cannot be the correct one. Assuming that the antecedent VP is copied into [$_{VP}$ e], the structure given will lead to the same problem of infinite regress just sketched in (6) for (2e.):

(7) a. John [$_{VP2}$ met everyone that Mary did [$_{VP1}$ e]]
b. John [$_{VP2}$ met everyone that Mary did **[$_{VP2}$ meet everyone that Mary did [$_{VP1}$ e]]]**
c. *etc.*

Alternatively, under the PF-deletion approach, the structure giving rise to VP-deletion cannot have been generated, if VP1 is properly contained in VP2.

In the classic approach to (2d) of May (1985), the structure given is assumed to be the correct one for S-structure. Therefore "S-structure cannot be the level for the copying operation". It is then proposed that the S-structure representation is altered in the LF-component in such a way that the VPE is no longer contained within its antecedent VP prior to the copying operation needed to derive a well-formed LF. The operation that achieves this is Quantifier Raising, which raises the object NP containing the VPE in the relative clause out of VP, adjoining it to IP:

(8) a. *SS* John met everyone that Mary did [$_{VP}$ e]
b. *QR* [$_{IP}$ **[$_{NP}$ everyone that Mary did [$_{VP}$ e]]** [$_{IP}$ John met ***t***]]
c. *Copy* [$_{IP}$ [$_{NP}$ everyone [$_{CP}$ *Op* that M. *Past* **[$_{VP}$ meet *t*]**]] [$_{IP}$ J. *Past* **[$_{VP}$ meet *t*]**]]

At the stage at which copying applies, the antecedent VP consists of the verb and the trace left by QR, and the dependent VP is no longer contained within its antecedent. Copying the antecedent VP into the VPE thus avoids regress. Additionally, this analysis yields a representation which (i) satisfies other well-formedness conditions - for example, the relative clause operator c-commands a trace (variable) which it can bind; and (ii) forms an accurate basis for capturing the actual intepretation of the example. From (8c), a relatively simple translation mechanism will yield a formula such as (9), which represents the relevant aspects of the meaning of (8a) accurately enough:

(9) $\forall x$ (met(m)(x) $\rightarrow$ met(j)(x))

1.2 Conditions on Antecedent-containment

The assumption (3) is generally assumed to hold at LF (May (1985), Hornstein (1994), etc.). One exception is Baltin (1987), who claims that (3) holds at S-structure. In the following, I make the claim (somewhat related to Baltin's) that reference must be made to two independent conditions - (3), holding at LF, and (10), holding at PF:

(10) A VP ellipsis site E may not be *linearly* contained in its antecedent string

While a VPE may, at PF, apparently either follow or precede the string of terminals corresponding to its LF-antecedent (cf. (2b) vs. (2c)), (10) states that a VPE may not be properly contained within that string: i.e. for any two terminals (α, ß) belonging to the string spelling out the antecedent of E, if α precedes E, then E must not precede ß. This must be excluded independently of (3). (10) is motivated (i) by some new evidence showing that the distribution of VPE is governed by generalizations over surface orders that cannot be reduced to (3) holding at LF (sections 3-5); (ii) a reinterpretation of some familiar facts (section 6).

1.3 Pseudogapping vs. Antecedent-containment

Two further assumptions underlying the discussion of ACDs are:

(11) a. the target of VPE is a single constituent
b. that constituent is VP

One problem that threatens the whole discussion is the possibility that one or both of these assumptions is false.

Consider (11b). If the deletion were only to target the verb, then the dependent element would not be contained in its antecedent. The trace of the relative operator would then be a "remant" not contained within the deletion site:

(12) John [$_{V2}$ **met**] everyone [*Op* that Mary did [$_{V1}$ *meet*] *t*]]

The pseudo-gapping construction (13) provides evidence that such deletion of a verb alone, leaving its object as remnant, is possible:

(13)a. ? John met Bill and Mary did, Paul
b. John [$_{V2}$ **met**] Bill and Mary did [$_{V1}$ *meet*] Paul

Hence, (2d)/(7) are potentially irrelevant to the discussion of antecedent-containment.

Furthermore, the pseudo-gapping construction shows that discontinuous deletion sites (discontinuous in surface order at least) are possible. Consider:

(14)a. ? John sent the book to Bill and Mary did, the papers.
b. John [$_{V2}$ **sent**] the book [$_{PP2}$ **to Bill**] and Mary did [$_{V1}$ *send*] the papers [$_{PP2}$ *to Bill*]

Unless it can be shown that such discontinuous deletions form a constituent at some level, e.g. LF, the possibility must be countenanced that apparent VP-deletions are actually composed of adjacent deletions of smaller constituents. So (1a) (=15a) may involve adjacent V-deletion and NP-deletion sites, as in (15b):

(15)a. John met Mary, and Bill did, too.
b. John [$_{V2}$ **met**] [$_{NP1}$ **Mary**] Bill did [$_{V1}$ *meet*] [$_{NP1}$ *Mary*], too.

Hence, more complex examples of apparent ACDs that avoid the objection posed by (13) may turn out to consist of several deletion dependencies, none antecedent contained. Instead of involving an ACD (16b), (16a) may be analysed by analogy with (15b), as involving two adjacent, independent deletions, neither antecedent-contained:

(16)a. John met Mary everywhere that Bill did.
b. John [$_{VP2}$ **met Mary** everywhere that Bill did [$_{VP1}$ *meet Mary t*]]
c. John [$_{V2}$ **met**] [$_{NP1}$ **Mary**] everywhere that Bill did [$_{V1}$ *meet*] [$_{NP1}$ *Mary*] *t*.

I put this issue aside here, returning to it in section 6.1.

In section 2., I review some previous accounts of the ACD construction, highlighting crucial examples and other facts which these sought to account for. In section 3., these accounts are confronted with new data which motivate the approach to be defended here.

2. ACDs 1985-1995

2.1 May (1985)

May's classic analysis of ACDs, according to which hierarchical containment configurations holding at S-structure can be resolved at LF by QR, prior to LF-copying, supplies one of the major arguments for the existence of the quantifier raising operation in covert syntax.

One argument that May used to support the QR approach to ACDs was that only quantified expressions license ACDs in relative clauses modifying them. Thus (17), with a name modified by a nonrestrictive relative, is deviant under the antecedent-contained reading; this is predicted, since a name, not being a quantified expression, does not undergo QR (May 1985:12):

(17) *? John accused Mary, who Peter did [$_{VP}$ e]

The object NP remains in VP at LF, hence the VPE cannot take the VP headed by *accused* as its antecedent without inducing infinite regress.

This argument is put in doubt by the fact that (17) improves significantly if *too* is added, as observed by Wyngaerd & Zwart (1991):

(18) ? John accused Mary, who Peter did [$_{VP}$ e], too.

The pattern can be reproduced for VPE in coordinate contexts (19), suggesting that the cause of the deviance of (17) was incorrectly diagnosed by May, and that it is due to some completely independent factor (cf. also Hornstein 1994):

(19)a. *? John accused Mary, and Peter did.
b. John accused Mary, and Peter did, too.

The well-formedness of (18) does not, however, necessarily threaten the QR approach to ACDs in restrictive relatives. Lasnik (1995) notes some differences between ACDs in nonrestrictive relatives and those in restrictive relatives, including the contrast (21):

(21)a. John stood near everyone Mary did [$_{VP}$ e] .
b. * John stood near Sue, who Mary did [$_{VP}$ e] (too).

Lasnik suggests that non-restrictive cases involve a pseudogapping derivation of the type discussed in section 1.3. Then, (18) is to be analyzed as in (22a), and (21b) as in (22b):

(22)a. John [$_{V°}$ **accused**] Mary, who$_j$ Peter did [$_{V°}$ *accuse*] t_j too.
b. (*) John [$_{V°}$ **stood**] [$_{P°}$ **near**] Sue, who$_j$ Peter did [$_{V°}$ *stand*] [$_{P°}$ *near*] t_j too.

The deviance of (22b) can then be related to the impossibility of deletion of the V+P combination (*stand near*) in the coordinate structure (23a): [4]

(23)a. * John stood near Sue and Peter did, Mary
b. John stood near Sue and Peter did, near Mary

If the pseudogapping approach generalizes to all non-restrictive examples, then these are not ACDs at all, and the well-formedness of (18) is irrelevant. At the same time, the well-formedness of (21a), contrasting with (21b) and (22b), tells against the possibility of generalizing a pseudogapping analysis to all ACDs in restrictive relative clauses. The contrast can be handled by assuming a ellipsis of the whole VP including a stranded preposition (near) is possible in (21a), because QR breaks up the antecedent containment configuration prior to LF; while QR could not break the antecedent containment arising from VP-ellipsis in (21b).

2.2 Baltin (1987)

Baltin (1987) constructs an argument against the QR account of the resolution of ACD's from the observation that the account "overgenerates": readings are predicted to be possible that do not in fact exist.

One such example is (24a). The VPE in the relative clause may only take the embedded VP as its antecedent (24b), although the QR account predicts that the higher VP headed by *thought* should be a possible antecedent as well (i.e. that (24a) should be ambiguous between (24b) and (24c)):

(24) a. Who thought that Fred read how many of the books that Bill did?
b. = Who thought that Fred read how many of the books that Bill *read*?
c. ≠ Who thought that Fred read how many of the books that Bill *thought that Fred read* ?

As well as applying to quantificational expressions such as *everyone*, Quantifier Raising is assumed to apply in multiple interrogatives to wh-phrases located "in situ" at S-Structure, such as the the wh-phrase headed by *how (many)* in (24). This wh-phrase raises covertly to a position from which it takes scope as a wh-operator, namely the specifier of the root CP, where it adjoins to the wh-phrase *who* in the standard analysis. It is crucial to the QR account of the resolution of ACD's, that the whole quantifier phrase is raised by QR, including the relative clause containing the VPE (or at least, that pied-piping of the relative clause is possible). So, assuming that the relative clause in (24) is "pied-piped" by wh-raising at LF, the complex specifier of the root CP will contain the VPE at LF, as shown in (25a):

(25) a. [who$_j$ [**how many of the books that Bill did [$_{VP}$ e]]$_k$**]

b. [SPEC C° [$_{IP}$ t_j [$_{VP1}$ thought [that Fred [$_{VP2}$ read t_k]]]]]

The VPE is contained neither in VP1 nor in VP2 at LF (25b). Hence, both VP1 and VP2 should be available as alternative antecedents to VPE. But factually, only VP2 can antecede the VPE - cf. (24).

In the following, a reading such as (24b), where a VPE takes only the first VP up as its antecedent, is termed a narrow scope reading; (24c), where a higher VP is intended as antecedent, is a wide scope reading. There is no independent reason why the VPE in the LF-representation (25) should not take VP1 instead of VP2 as its antecedent. In (26a), where the relevant wh-phrase has raised overtly, both narrow and wide scope readings are possible:

(26) a. How many of the books that Bill did, did Mary think that Fred read ?
b. = How many of the books that Bill *read*, did Mary think that Fred read ?
c. = How many of the books that Bill *thought that Fred read*, did Mary think that Fred read ?

Thus, there is a mismatch between the scope taken by a scope-bearing element E and the scope of a VPE contained within the phrase headed by E. While QR at LF may assign the quantifier wide scope, the VPE may only get a narrow scope reading. Baltin calls this restriction on ACD's the Boundedness restriction.

In his analysis of the ACD construction, Baltin makes a different assumption concerning the filter on antecedent-containment:

(27) There is no well-formed sentence in which a VPE is contained in its antecedent in the S-structure representation.

If (27) is true, then QR in the LF-component will not suffice as a mechanism by which antecedent-containment is resolved. Even if the application of QR should result in a VPE no longer being

contained in its antecedent at LF, this would happen, as it were, 'too late' in the derivation to save the sentence, which will aleady be marked as ungrammatical (in violation of (27)) prior to the QR operation. [5]

Baltin proposes that a different movement operation is responsible for resolving antecedent-containment, namely Relative Clause Extraposition (RCE). RCE is assumed to apply in the overt syntax, extracting a relative clause from the NP it modifies and right-adjoining it to the VP (or some higher projection). Supposing that RCE has applied - string-vacuously - to (1d), the S-structure representation looks like (28a), and (27) is satisfied. QR may then apply to yield (28b), on the basis of which VP1 may take VP2 as its antecedent without incurring regress (28c):

(28) a. John [$_{VP3}$ [$_{VP2}$ met everyone] [that Mary did [$_{VP1}$ e]]
b. [**everyone** [John [$_{VP3}$ [$_{VP2}$ met *t*] [that Mary did [$_{VP1}$ e]]
b. [everyone$_j$ [John [$_{VP3}$ [**$_{VP2}$ met t$_j$**] [that Mary did [$_{VP1}$ *meet* t_j]]

Given (27), QR is clearly insufficient to account for ACD-resolution, even if it is involved in the derivation of relevant examples. The hypothesis that RCE resolves containment, on the other hand, is not only compatible with (27), but also provides a means for accounting for the Boundedness Restriction on ACD-resolution.

Extraposition is subject to a limitation known as 'Right Roof Constraint' which prohibits it from moving a constituent out of the minimal sentence (CP) containing it. The Right Roof Constraint is usually illustrated by pointing to word order facts such as (29):

(29) a. Bill will say that John kissed everyone that he met tomorrow.
b. * Bill will say that John kissed everyone tomorrow that he met.

The relative clause, which modifies the object of the embedded clause, may not be placed to the right of the adverbial modifying the matrix VP, a linear position that corresponds on standard assumptions to a hierarchical position higher than the matrix adverbial, hence outside of the embedded clause. Given this interpretation of RCE and the Right Roof constraint, the impossibility of wide scope interpretation in (24) can be explained. Since RCE can only raise the relative clause prior to S-structure to a position within the clause containing the object *how many of the books*, the VPE inside it is no longer contained in the minimal VP headed by *read*, but is still contained within the higher VP headed by *thought*:

(30) Who [$_{VP4}$thought that Fred [$_{VP3}$[$_{VP2}$read how many of the books][that Bill did [[$_{VP1}$e]]]

Assuming that no post-S-structure operation affects the relative hierarchical positions of the VPs in (30), there is exactly one VP which does not contain VP1 at LF and which may function as its "antecedent", namely VP2. (QR may, in fact must, raise the phrase *how many...*, to ensure that at LF, VP1 contains a trace that can be bound as a variable by that phrase). In this way, Baltin is able to derive the Boundedness Restriction on ACDs.

2.3 Larson & May (1990)

In their reply to Baltin (1987), Larson and May (1990) give a series of counterarguments to the assumption (27), and to the proposal that all relative clauses containing VPEs in the ACD construction have undergone (possibly string-vacuous) extraposition.

Direct counterevidence to (27) comes in the form of examples involving a VPE both preceded and followed by terminals beloning to its antecedent at S-Structure. In (31a), the VPE is contained within a free relative that functions as the goal object in the double object construction; its antecedent includes the preceding verb and the following NP (the theme argument of *gave*): [6] ACDs

involving a VPE contained within the accusative subject in ECM constructions are also held to be acceptable (in disagreement with Baltin - cf. also Hornstein 1994:459):

(31)a. John **gave** [whoever he could _] **two dollars**
b. ? John **believes** [everyone I do _] **to be a genius**

Larson & May proceed to argue that QR alone makes the right distinctions among possible and impossible ACDs. In exactly those cases in which QR raises a constituent containing the VPE out of its antecedent VP before LF. The accusative subject in the infinitive complement (31b), for example, contrasts with the nominative subject of it finite counterpart, in that the latter may not host an ACD taking the VP headed by *believe* as its antecedent:

(32) * John **believes** [everyone I do _] **is a genius**

It is claimed that while the accusative subject may undergo QR to a position above the higher verb, the nominative subject is unable to leave its own clause. This claim correlates with the possibilities for scopal interactions between quantifiers: while the accusative subject in (33a) may take scope over the subject of the higher clause, the inverse scope reading is not possible in (33b):

(33)a. some student believes every teacher to be intelligent ok: $\exists\forall$ / $\forall\exists$
b. some student believes that every teacher is intelligent ok: $\exists\forall$ / * $\forall\exists$

2.4 Hornstein (1994)

In the context of the Minimalist framework of Chomsky (1993), Hornstein (1994) proposes what is essentially a variant of the QR account of ACD resolution. Following Larson & May (1990), Hornstein rejects (27), assuming that the sole filter on antecedent containment holds of LF-representations, and concluding that ACDs provide evidence for phrasal movement in the covert syntax. Disagreement arises as to the nature of that movement process. Instead of QR - movement to an A'-position - Hornstein argues that it is A-movement of objects at LF that resolves antecedent-containment in ACDs.

In Chomsky's (1993) proposal, direct objects in English, which stand in VP at S-structure ("Spell-Out"), must raise at LF to the specifier of a functional projection ("object agreement" - AGo) to check Case and agreement features. Hornstein seeks to establish that this assumption provides the basis for an account of ACD resolution without recourse to QR. The (simplified) derivation for (2d) is given in (34):

(34)a. John *Past* [$_{VP}$ met everyone that Mary did [$_{VP}$ e]]

b. [$_{IP}$ John *Past* [$_{AGoP}$ [**$_{DP}$ everyone that Mary did [$_{VP}$ e]**] AGo [$_{VP}$ *t* met *t*]]]

c. [$_{IP}$ John *Past* [$_{AGoP}$ [$_{DP}$ everyone that Mary did **[$_{VP}$ *t* met *t*]**] AGo **[$_{VP}$ *t* met *t*]**]]

Assuming an LF-copy approach to VPE, the Spell-Out representation is (34a). The direct object containing the VPE is in VP. Covert A-movement places the direct object DP in the specifier of AGo (34b). LF-copying may now substitute the main VP for the VPE in the relative clause of the raised object, no regress arises. [7] (Other technical assumptions on which this account depends are left aside here. [8]).

Hornstein cites conceptual and empirical reasons for rejecting the QR-based solution to the ACD issue. Firstly, within the Minimalist framework, given the hypothesis that all movement processes are driven by purely formal (morphological) rerquirements, it is not clear if QR even

exists. Aside from wh-movement, there is no apparent morphological requirement to which QR may be the response. Secondly, evidence independent of the ACD phenomenon that is used to support the postulation of QR - i.e. relative scope facts; in particular, scope inversion phenomena - does not require the assumption of QR (i.e. there are alternative semantic accounts for quantifier interpretation, including relative quantifier scope, that do not depend on qunatifier raising in the syntax).

Even if QR is adopted as the means for handling scope facts, this alone do not suffice to ensure that QR also provides for ACD-resolution. The latter depends on the further assumption that relative clauses are pied-piped under QR at LF. As Hornstein notes, the scope facts would be accounted for under the assumption that it is the quantifier alone that raises under QR, i.e. that further material is not pied-piped. That covert NP-movement, on the other hand, must pied-pipe the whole DP, Hornstein takes as given. [9]

Beyond this, there remain worries about the coextension of QR and ACD resolution. Appositive relatives modifying proper names in object position, for instance, do permit ACDs (cf. (18) in section 2.1 above), despite the assumption that names do not undergo QR. Baltin's Boundedness Restriction is another case where scope facts and ACD facts diverge, which remains unexplained in the QR approach.

The A-movement solution is claimed to capture all facts discussed so far, hence to be superior to previous accounts. Correct predictions are generated concering cases where QR targets and ACDs diverge: while names do not undergo QR, they do undergo LF-A-movement, like any other DP. Hence, the occurrence of ACDs in appositive relatives is expected. A-movement also differs from QR in that it is clause-bound; this fact can be used to explain the Boundedness Restriction observed by Baltin, without recourse to extraposition. Consider (24a), repeated here:

(24) a. Who thought that Fred read how many of the books that Bill did?

At LF, the phrase headed by *how many* undergoes LF A-movement to the AGo-projection immediately dominating *read* (35b). Even if there is an LF-operation of QR that raises *how many* to the specifier of the matrix CP (35a.) (thus fixing its scope), that operation does not pied-pipe the relative clause. The VPE thus remains inside IP2 at LF (35b.), and is unable to receive a wide scope reading (i.e. to take VP1 as its antecedent):

(35) a. [who$_i$ - **how many$_j$** C° [$_{IP1}$ t_i [$_{VP1}$ thought that IP$_2$]]]

b. [$_{IP2}$ Fred I° **[t_j of the books that Bill did [$_{VP}$ e]]$_k$** [$_{VP2}$ read t_k]]]]]

The only case where an ACD is able to take wide scope - i.e. where an VPE contained within an embedded clause is able to take the VP containing that clause as its antecedent - is where an A-movement process raises the DP containing the VPE out of the embedded clause. This happens in ECM constructions, in Chomsky's (1993) account. Consider (36): at S-structure, the embedded subject is contained within the clausal complement of *believe*:

(36) John **believes** [everyone I do _] **to be a genius**

At LF, however, this phrase raises to the AGo-projection immediately dominating *believe*, to check accusative Case (37). NP-movement brings the VPE out of the matrix VP, thus permitting that VP to act as antecedent:

(37) [$_{IP}$ John [$_{AGoP}$ **[$_{DP}$ everyone that I do [$_{VP}$ e]]** AGo [$_{VP}$ *t* believe [*t* to be a genius]]]]

The line is drawn, correctly, between accusative subjects (undergoing NP-movement at LF) and nominative subjects in finite complements to believe, which do not raise to an A-position outside their clause (36) vs. (32). Hence, the LF A-movement approach predicts VPEs which are contained within their antecedents at S-structure, to be permitted precisely where LF-A-movement undoes the containment configuration. Thus cases like (36), counterexamples to Baltin's extraposition account, fall out correctly.

While the A-movement proposal seems to make more accurate predictions than the QR-proposal with respect to wide/narrow scope readings and names, it may face other problems that arise from it being too restrictive in its choice of target. A VPE may give rise to an ACD configuration not only when contained in a relative modifier of a direct object, but also when contained in a PP-complement or an adverbial, categories which are not normally considered to be targets for A-movement: [10]

(38) a. John [talked [to everyone who Mary did _]]
b. John [recited his lines [in the same way that Bill did _]]

Hornstein suggests that the former case be handled in terms of A-movement of the DP out of PP into a higher agreement projection. In other words, such examples are used to motivate an extension of the LF-A-movement hypothesis to DPs in selected PPs. Examples involving adverbials at the right edge of VP (38b) are assumed not to constitute ACDs even at S-structure: it is assumed that the base-generated position is right-adjoined to (hence higher than) the antecedent VP.

Both of these proposals are problematic, but I do not dwell on them here, since there is a far more serious problem facing Hornstein's proposal, and the other approaches reviewed so far.

3. Wide scope ACDs

3.1 A new problem

As noted in the previous sections, a VPE in a relative clause modifying the subject of a finite complement may not take a wide scope reading (39a) (cf. (32) above). However, as Lasnik (1995) observes, citing R. Tiedeman, wide scope VP-ellipsis is possible in a relative modifying the subject of a finite complement clause - if the relative is extraposed (39b): [11]

(39) a. * John believes [everyone I do _] is a genius
b. ? John believes everyone is a genius [that I do _]

Similar pairs involving comparatives can also be constructed:

(40) a. * John claimed that more people than you did _ were stupid
b. John claimed that more people were stupid than you did _

The wide scope interpretation of the VPE in (39b) and (40b) is quite clear: the matrix VP is reconstructed into the VPE, so that they mean the same as the sentences (41) and (42) (with or without extraposition), respectively: [12]

(41) a. John believes everyone is a genius that I believe is a genius.
b. John believes everyone that I believe is a genius, is a genius.

(42) a. John claimed that more people were stupid than you claimed were stupid.
b. John claimed that more people than you claimed were stupid, were stupid.

With VPE, extraposition of the relative or comparative clause has a dramatic effect on acceptability, which none of the analyses discussed in the previous section is in a position to explain. The problem posed by the contrast is different in each case.

3.1.1 The QR account

For the QR account (May 1985), Larson & May (1990), the question raised is: why should extraposition of a modifier affect the ability of the head NP to undergo QR? In order for the QR account to be able to explain the wide scope reading of the VPE in (39b) and (40b), it must be assumed that the nominative can undergo QR, pied-piping its modifier, to a position above the matrix VP. [13] In order to account for the ungrammaticality of (39a) and (40a) under the wide scope reading, QR must be unable to raise the nominative to a position above the matrix VP. Either QR is prevented from applying, or an independent explanation must be found. Larson & May (1990:108) explicitly claim that the reason for the ill-formedness of (39a) is that the subject of a finite complement clause may not undergo QR into the matrix sentence ("quantified subjects are clause-bound in finite sentences"). Suppose ECP is the constraining principle. The application of ECP should be blind to whether or not a raised QNP contains a relative clause. How extraposition of a modifier could neutralize the ECP in this case, is completely mysterious. [14]

3.1.2 The extraposition account

For Baltin (1987), the problem is a different one. Given the S-structure constraint (27) of section 2.2 above, it appears at first glance that the paradigm (39)-(40) actually supports Baltin's approach. (39a) and (40a) clearly involve antecedent containment at S-structure, regardless of what happens at LF. It seems as if extraposition in (39b) and (40b) has the effect of removing the VPE from the containment configuration, exactly what Baltin claims. However, if extraposition obeys the Right Roof constraint, the relative clause in these examples can be adjoined no higher than the CP of the complement, where it will still be contained within the VPE-antecedent (the matrix VP), in violation of the S-structure contraint (27). In other words, the S-structure (43a), which obeys Right Roof, violates (27); while (43b) satisfies (27) but violates Right Roof:

(43)a. John [$_{VP}$ believes [$_{CP}$ [$_{CP}$ that everyone is a genius $_{CP}$] <u>that I do VP</u> $_{CP}$] $_{VP}$]
b. John [$_{VP}$ [$_{VP}$ believes [that everyone is a genius] $_{VP}$] <u>that I do VP</u> $_{VP}$]

The Right Roof Constraint plays a crucial role in deriving the Boundedness restriction discussed in connection with example (24) above. (39b) and (40b) are clear counterexamples to the Boundedness Restriction. Baltin assumed that the Boundedness restriction holds without exception. If it were the case that the Right Roof constraint were not operative in these examples, then (39b) could be assigned the S-structure (43b). Thus, the idea may seem worth exploring, that these sentences counterexemplify the Boundedness restriction for the very reason that extraposition is not constrained by the Right Roof Constraint. However, I do not follow this path here, for the following reasons, empirical and conceptual.

Firstly, it is doubtful whether the Right Roof effect is cancelled (i.e. that the "roof" is extended) in these constructions: the word order test clearly indicates that the Right Roof is still in place - no matrix adverbial may separate the extraposed relative from the rest of the complement clause:

(44)a. * John said that everyone is a genius yesterday that I did
b. * John claimed that more people were stupid when I asked him, than you did

Secondly, the very nature of the approach - establishing the satisfaction of an S-structure constraint (Baltin's (27)) by recourse to a rightward movement analysis of extraposition - conflicts with leading hypotheses of current theory. Within the minimalist framework, S-structure (in the

sense of earlier models) has been abandoned as a level of representation, so that it becomes desirable to reformulate constraints apparently holding of S-structure either as interface conditions holding at LF or PF, or as properties of derivational operations. Assuming Kayne's (1994) Linear Correspondence Axiom (or Chomsky's (1995) reinterpretation of it), linear order (precedence) loses its status as a primitive, becoming a property of PF derived from hierarchical relations (asymmetric c-command). One consequence is that UG does not permit rightward adjunction, nor rightward movement (adjuncts precede their hosts, and moved categories precede their traces, after linearization). Hence, extraposition cannot be analyzed as rightward movement of CP; and whatever the explanation is for Right Roof constraint effects (or their absence), these cannot be attributed to a bounding condition on rightward movement since this does not exist. For present purposes, I assume that extraposition is "stranding" of CP under leftward movement of the "head" of the extraposed clause, as in Kayne (1994) or Wilder (1995b). In that case, the extraposed clauses in (39b) and (40b) are actually located within the VP of the complement clause, as in (45):

(45) a. John $[_{VP}$ believes $[_{CP}$ that [everyone]$_j$ $[_{VP}$ is a genius [t_j that I do VP]]]]
b. John $[_{VP}$ claimed $[_{CP}$ that [more people]$_j$ $[_{VP}$ were stupid [t_j than I did VP]]]]

3.1.3 The A-movement account

For Hornstein (1994), the problem posed by wide scope ACDs is as simple as it is unresolvable. Hornstein's main goal is to establish the ACD phenomenon as a source of independent evidence for the "minimalist" conception of covert A-movement of objects in English. To do this, he must make the A-movement account of ACDs into a persuasive alternative to the QR account. His main hope of success in this task lies in establishing that the possible wide scope ACD readings are coextensive with the domain of LF A-movement. One contribution to that goal is provided by the case of Baltin's boundedness restriction, unexpected under the QR account. But in addition, the A-movement account depends crucially on the contrast between (46a) and (46b):

(46) a. John believes everyone I do _ to be intelligent
b. * John believes everyone I do _ is intelligent
c. John believes everyone is intelligent that I do _

Adding (46c) (=39b) to the paradigm ruins the picture. (46c) should be as ungrammatical as (46b). It is hardly plausible to assume LF A-movement of nominatives into a higher clause; even less so, if that operation depends on the presence of a relative clause that is extraposed.

In fact, Hornstein is aware of the problem. In a footnote, he addresses the issue of wide-scope ACDs and seeks to deny the facts:

> "May (p.c.) notes cases that are problematic for the above claims:
>
> (vi) Max thinks the yacht is longer than Oscar does
>
> (vi), he says, has the interpretation '...longer than Oscar thinks it is' [...] These cases all strike me as unacceptable. However, it is possible that more is going on here than present analysis accounts for [...] if matrix ACD readings are regularly possible from null VPs in embedded clauses, then this would tell against the minimalist approach sketched here. I will proceed as if this is not the case." (Hornstein 1994:462, Fn. 9 - my emphasis).

I do not think these facts are deniable.

3.1.4 Further properties of wide scope ACDs

What is the extent of the problem? Wide scope ACDs are not freely constructible, one has to hunt them. As Hornstein notes (fn. 9), there are unclear restrictions on matrix predicates and tenses:

(47) a. ?? John has heard that everyone is a genius [that I have _]
b. ?? John said that Bill was reporting everything [Sam did _]
c. * John doubted that the yacht was as long as she did
d. * John insists that my yacht is flashier than you do

This may help to explain why they have escaped notice so far (in addition, Baltin's examples illustrating the Boundedness restriction (24) might also have served to distract attention from them). Nevertheless, there are enough good examples to establish that this is indeed a productive phenomenon which cannot be simply dismissed.

The availability of a narrow scope reading can make a wide scope reading barely, if at all, accessible. For (48a), the reading (48b) obscures (48c), for example:

(48) a. John thought that Mary bought more books than Bill did
b. John thought that Mary bought more books than Bill did *buy*
c. John thought that Mary bought more books than Bill did *think that Mary bought*

In constructing examples, it is useful to be aware of, and control for, this "local antecedent" effect. In (49), where the embedded subject in the antecedent does not match the subject of the VPE in animacy, but the matrix subject does, the wide reading is at least as accessible as the narrow reading:

(49) a. John thought that the fire destroyed more books than Bill did
b. John thought that the fire destroyed more books than Bill did *destroy*
c. John thought that the fire destroyed more books than Bill did *think that it destroyed*

Prosodic factors can be important. A wide scope reading can be facilitated by parallel accent on the subject of the VPE and the subject of antecedent VP, with the subject of the complement clause in the antecedent deaccented (typographically: deaccenting in small pitch, accenting with ´ on vowel). In fact, wide and narrow readings for the VPE can be disambiguated prosodically in this way (cf. also discussion of deaccenting and VPE in Tancredi (1992)):

(50) a. Jóhn thought that the fire destroyed more books than Bíll did (wide only)
b. John thought that the fíre destroyed more books than Bíll did (narrow only)

Prosodic manipulation cannot save the wide scope reading in Baltin's example (51) (= (24) in section 2.2), however: it can only exclude the narrow scope reading as well:

(51) * Whó thought that Bill read how many of the books that Máry did

Two further properties of wide scope ACDs are relevant to the proposal to be made in section 3.2. Firstly, they are only found in right-peripheral position (cf. Hornstein 1994: fn.9):

(52) a. John said that there were more people than Mary did _
b. John said that more people were there than Mary did _
c. * John said that more people than Mary did _were there

The function of the head of the relative or comparative containing the ACD (subject, object, predicate, etc) does not seem to matter, as long as the VPE is in final position; i.e. if the wide scope reading is possible at all, then only if the VPE is in final position. Extraposition does not seem necessary (53a), (54a) - although it may operate string-vacuously, as in Baltin's proposal - unless the relative or comparative is not in final position (53b), (54b):

(53) a. John thinks that this will cause more harm than Mary does
b. * John thinks that this will cause more harm than Mary does to the proposal
c. John thinks that this will cause more harm to the proposal than Mary does

(54) a. John thinks the situation has got worse than Mary does
b. * John thinks the situation has got worse than Mary does through this
c. John thinks the situation has got worse through this than Mary does

Secondly, wide scope ACDs are found in relative clauses and comparative clauses, but they are not licensed in adjunct clauses (cf. also Hornstein 1994). A sentence like (55) is ambiguous with respect to which verbal projection the temporal adjunct modifies:

(55) John said that Mary would arrive when Peter did.

Adding the wide - narrow scope dimension for the VPE (and ignoring the possibility for an external antecedent), we expect the sentence to have four possible readings (56). In fact, it only has three.

(56) a. [[John said that Mary would arrive] when Peter did *arrive*]
b. [[John said that Mary would arrive] when Peter did *say that Mary would arrive*]
c. John said that [[Mary would arrive] when Peter did *arrive*]
d. * John said that [[Mary would arrive] when Peter did *say that Mary would arrive*]

If the temporal clause modifies *said*, then the VPE may take either the embedded (56a) or the matrix VP (56b) as antecedent. If the temporal clause is understood as modifying *would arrive*, then the VPE may only take the embedded VP as antecedent (56c). The intended reading for (56d) can be paraphrased as (57):

(57) John said at time t, that Mary would arrive at time t', t' being such that Peter said (at time t'') that Mary would arrive at t'.

This reading is utterly impossible for (55) - although it represents a possible interpretation of (58), the 'undeleted' counterpart to (55):

(58) John said that Mary would arrive when Peter said that she would arrive.

3.2 A new proposal

The evidence from section 3.1 speaks strongly against the proposal of Hornstein (1994): LF A-movement is too restrictive to be able to account for limitations on wide scope readings of ACDs. It is consistent with an A'-movement (i.e. QR) account, assuming that restrictions on QR are relaxed to permit LF-extraction of certain phrases from finite complement clauses. But that account is then too liberal to account for the distribution of well-formed and ill-formed ACDs - in particular, the contrasts involving extraposition. These in turn motivate a proposal in the spirit of Baltin's. There is not only an LF condition barring antecedent containment, which is voided by QR in certain cases; there is also an independent surface condition, which is voided by extraposition in certain cases.

That is, in addition to the condition (59) (= (3) of section 1.1 above) holding at LF, there is a separate principle (60), holding at PF:

(59) A VP-deletion site may not be contained in its antecedent **constituent** at LF

(60) A VP-deletion site may not be contained in its antecedent **string** at PF

In view of the considerations on the status of extraposition mentioned above, the surface condition cannot be stated in terms of hierarchical containment (as Baltin's condition is). Extraposition only serves to ensure that the VPE is not contained within the string corresponding to its antecedent; it does not remove the VPE from the hierarchical constituent of its antecedent. Instead, the condition is formulated in terms of "string-containment". What (60) excludes is any PF-string in which two terminals α, ß, belonging to the antecedent VP, are linearized with respect to the VPE such that α precedes the ellipsis site and ß follows the ellipsis site:

(61) * α < E < ß , where α, ß are elements of the antecedent of E

If syntactic objects (including S-structure or "Spell-Out" trees) do not contain precedence information, then (60) cannot hold of S-structure. The locution "at PF" is however intended to be vague, referring to some stage within the derivation from syntactic component to the "PF interface", in the sense discussed in Chomsky (1995). The earliest stage at which (60) can be assumed to hold is the level at which representations are linearized. The problem with assuming that (60) holds at the interface is of course that it refers to a phonetically empty element, by hypothesis not "visible" at the interface. [15]

(59) and (60) are each necessary independently of the other. The impossibility of a wide scope reading for a VPE in an embedded adjunct clause provides evidence for the necessity of the ban on structural containment (59) - since the adjunct appears string finally, the ban on string containment (60) would be insufficient. The adjunct - relative clause asymmetry (a VPE in a final embedded relative clause can receive wide scope interpretation) supports the assumption of the QR-escape hatch (assuming that an extraposed relative clause is "reconstructed", so that it can be pied-piped by LF-movement). The extraposition / intraposition asymmetry found with embedded relative clauses then provides the evidence for the existence of the additional surface filter (60). This filter must be formulated in terms of linear containment; hierarchical containment would be too strong (unless extraposition is rightward movement and the Right-Roof Constraint is lifted in the relevant cases).

Assuming that adjunct clauses do not undergo QR, wide scope VP-deletion in a final embedded adjunct (56d) violates (59), but not (60). A wide scope VPE in an embedded intraposed relative (62) violates the surface filter (60), but not (59) (assuming QR). A wide scope in an embedded final relative (63) violates neither.

(62)a. John believes everyone [that I do _] is a genius

b. *PF* * John **believes** everyone that I do (*believe (x) is a genius*) **is a genius**
c. *LF* [everyone that I do (*believe (x) is a genius*)] [John **believes x is a genius**]

(63)a. John believes everyone is a genius [that I do _]

b. *PF* John **believes everyone is a genius** that I do (*believe (everyone) is a genius*)
c. *LF* [everyone that I do (*believe (x) is a genius*)] [John **believes x is a genius**]

Additional evidence for the analysis of (62) in terms of LF-convergence but PF-crash comes from typical native speaker reaction to such examples - they are held to be "interpretable, but the word order is wrong". This reaction contrasts strongly with the reaction to embedded adjunct examples where a narrow scope reading is independently excluded: the wide scope reading is felt to be simply impossible ("the example is uninterpretable"). Their status is comparable to that of cases like (64), under the ACD reading:

(64) * John said that Mary did.

There is one issue not yet addressed. What to do with "genuine" string contained ACDs like (65), which - apparently - violate (60)?

(65) a. John gave whoever he could _ two dollars
b. ? John finds everyone I do _ easy to work with.
c. ? John believes everyone I do _ to be a genius

These cases are examined in section 6. below, where it is proposed that they do not involve antecedent-containment at all, hence do not pose any problem for (60).

4. Hierarchical containment and QR

In this section, two questions concerning the LF-side of the account are addressed: (i) what type of constituents undergo QR out of finite complement clauses (i.e. license wide scope readings for VPEs)? (ii) how far do such constituents move (what is the landing site of "long" QR)?

4.1 What undergoes QR?

Taking wide-scope VP-deletion as an indicator for QR, it turns out that QR can affect argument DPs of various types which have been claimed by other authors to be QR targets. Additionally, compared predicates must be assumed to be able to undergo QR. The VPE in each case is contained within a relative clause modifying the head noun, or, in the case of compared noun phrases and adjectival predicates, within the comparative clause selected by the degree word (e.g. the *than*-clause selected by *more*, the *as*-clause selected by equative *as* or by *the same*, etc.):

The possibility for argument noun phrases that are modified by relative clauses also seems to depend on the nature of the determiner. Wide scope VPEs occur most felicitously in relatives modifying DPs headed by "strong" determiners (*every*, *each*, *most*, etc - cf. (66)); definites, especially in combinations such as *the same* N, or *the very* N (67); and partitive DPs headed by "weak" determiners (68) ("strong" and "weak" in the sense of Milsark (1977), Diesing (1992)). In each case, the extraposition-intraposition contrast is evident, even where the extraposed variant itself is only marginally acceptable:

(66) a. John thought that most people were there that Mary did.
b. * John thought that most people that Mary did were there.
c. John thought that each proposal should be accepted that Mary did.
d. * John thought that each proposal that Mary did should be accepted.

(67) a. John said that the (very) same problem would arise that/as Mary did
b. * John said that the (very) same problem that/as Mary did would arise

(68) a. John said that many/none of the problems would arise that Mary did
b. * John said that many/none of the problems that Mary did would arise

All these DP types have been claimed by May (1985) and/or Diesing (1992) to undergo QR.

Examples involving relatives whose DP is headed by a weak determiner that is not a partitive, are marginal, bordering on the unacceptable:

(69) ??? John said that a / one / many / few / no problem(s) would arise that Mary did

In her discussion of indefinites, Diesing (1992:Ch.3) distinguishes a presuppositional and a non-presuppositional reading of indefinite DPs headed by weak determiners, claiming that presuppositional indefinites undergo QR, while non-presuppostional indefinites do not. Partitive indefinites are invariably presuppositional. Using ACDs as a diagnostic, the contrast between the two readings comes out clearly in simple examples (Diesing 1992:71):

(70)a. ?? I read few books that you did.
b. I read few of the books that you did

Under certain circumstances, (70a) can be rendered acceptable - namely when the context facilitates a presuppositional reading of the DP (in which there a salient set of books over which the determiner ranges, as in the sole possible reading of (70b)). As Diesing notes, (70a) is "unquestionably ungrammatical in the case of the cardinal (nonpartitive) reading". [16]

Plausibly, the difficulty with (69) can be attributed to the difficulty in accessing the presuppositional reading for the DP that would accompany the QR needed to resolve the VPE, just as in (70a). Additional support for this speculation can be derived from considering the weak determiner *some*. As Milsark (1977) has pointed out, the non-presuppositional ('cardinal') reading of DPs headed by *some* typically involve an unstressed variant *sm*; while stressing the determiner triggers a presuppositional reading. A wide scope VPE associated with *sm* is impossible, its counterpart with stressed *some* is relatively good:

(71)a. * John said that sm problems would arise that Mary did
b. ? John said that SOME problems would arise that Mary did

Already on the basis of simple examples like (72a), Diesing is led to suggest that comparatives undergo QR. Cases involving compared DPs seem to yield the best examples of wide scope ACDs (72b):

(72)a. I read more books than you did.
b. John said that more problems would arise than Mary did.

Compared embedded predicates also license wide scope ACDs:

(73) John said that his yacht was longer than Bill did

This indicates that it is the comparative morpheme, as a functional head governing the predicative AP analogous to the way a determiner governs NP, that licenses QR of its phrase, pied-piping the predicate along with the *than*-clause.

It is even possible for a compared adverbial modifying a verb in a finite complement to license a wide scope VPE in its *than*-clause:

(74) John said that it would rain more often than Mary did

(74) has a reading in which *more often...* modifies the embedded verb *rain*, while the *than*-clause contains a VPE taking the matrix VP *said that...* as its antecedent; which contrasts with the impossibility for a non-compared embedded temporal adverbial to contain a wide scope VPE (cf. (56) above):

(75) John said that it would rain when Mary did

The contrast (74) vs. (75) provides further support for the hypothesis that the comparative morpheme licenses QR of its phrase.

One problem for the QR approach noted above concerns that apparent possibility for appositive relatives modifying names to contain ACDs, as in (76a) - the problem being that names are not QR targets. However, appositive relatives in complement clauses do not support wide scope ACDs. (76b), which would have the reading (76c) if the DP headed by the name *Peter* could undergo QR, is deviant:

(76)a. John met Peter, who Mary did, too.
b. * John said it would please Peter, who Mary did, too
c. John said it would please Peter, who Mary said it would please, too

In section 2.1 above, it was suggested that examples like (76a) may involve pseudo-gapping, hence not bear on the issue of antecedent containment and QR at all. That solution is further supported by the fact that (77b) - with the intended reading (77a) matches the ill-formedness of (76b):

(77)a. John said it would please Peter, and Mary said it would please Bill.
b. * John said it would please Peter, and Mary did, Bill.

4.2 What is the landing site of "long-distance" QR?

The existence of wide scope ACDs motivates the claim that QR can raise constituents out of finite complement clauses into the next clause up. If this is so, then it is expected that the raised constituent will enter into scopal interactions with other quantified expressions originating in the higher clause. This expectation is not met, however. The questions that arise have to do with the analysis of scopal interactions in general, which I do not attempt to solve here. Instead, the problem will simply be laid out and left for further research.

It is already clear for simple cases (not involving wide scope ACDs) that a quantified expression originating in a finite complement clause does not interact scopally with any quantified argument (subject, direct or indirect object) in the matrix clause. The embedded subject in (78) cannot take scope over an indefinite in the matrix:

(78)a. Someone believes that everyone is a genius ok: $\exists\forall$ / * $\forall\exists$
b. John told someone that everyone is a spy

Addition of an extraposed relative clause modifiying the embedded subject, and containing a wide scope VPE does not alter this fact - the phrase headed by *everyone* cannot take scope over *someone* in (79):

(79)a. Someone believes that everyone is a genius that Mary does ok: $\exists\forall$ / * $\forall\exists$
b. John told someone that everyone is a spy that Mary did

In order to escape antecedent containment at LF, some phrase containing the wide-scope VPE (by hypothesis, the DP headed by *everyone* in (79)) must minimally gain scope over the constituent

(V') containing the matrix verb and its complement (the finite clause). At the same time, it must be ensured that rules of interpretation do not permit *everyone* to take scope over *someone*. In the theory of May (1985), quantified expressions may only move to positions adjoined to VP or IP (leaving aside movement to Spec,CP). Supposing that the phrase headed by *everyone* is adjoined to VP, and that the matrix subject *someone* is higher up the tree (e.g. in Spec,IP or adjoined to IP), an LF like (80a) is derived:

(80) a. someone$_y$ [everyone$_x$ that Mary *believes* [*x is a genius*]] [y **believes** [**x is a genius**]]
b. $\exists y\ \forall x$ (believe(m) (genius(x))) $\rightarrow$ (believe(y) (genius(x)))
c. "There is a *y* such that for all *x*, if Mary believes *x* is a genius, then *y* believes *x* is a genius"

From (80a), the formula (80b) would be derivable, which correctly represents the interpretation of (79a).

The problem posed by (80a) is that in current theories of scopal interactions, the matrix subject *someone* would be able to "get into the scope" of the raised DP, either by a rule of syntax such as quantifier lowering, or a rule of interpretation. In May's theory, for example, once an expression is adjoined to some VP, rules of interpretation enable it to interact scopally with any other expression adjoined either to that VP or to the IP immediately dominating. The lowest adjunction site consistent with resolving antecedent containment in (79) would be the VP headed by *believes* (or *told*). If the phrase headed by *everyone* has adjoined to that VP, then nothing would prevent it from being able to take scope over *someone*.

It does not seem possible however, to maintain that the landing site of QR is some lower position from which such scopal interactions would not be possible. [17] If QR in (79a) were is restricted to the finite compement clause, an LF would be generated in which the phrase headed by *everyone* would be in the scope of the matrix verb *believe*. The reading expected would be (81):

(81) a. $\exists y$ believe(y) ($\forall x$ (believe(m) (genius(x))) $\rightarrow$ (genius(x)))
b. There is a *y* who believes that for all *x*, if Mary believes *x* is a genius, then *x* is a genius.

Quite apart from the fact that this assumption about the landing site of QR would lead to an antecedent-containment configuration at LF, the reading (81) - which is available for the 'undeleted' examples (82), is <u>not</u> available for (79a):

(82) a. Someone believes that everyone that Mary believes is a genius, is a genius.
b. Someone believes that everyone is a genius that Mary believes is a genius.

The facts about wide scope ACDs suggest instead that the account of scopal interactions may be in need of revision.

There is one further constraint on wide scope ACDs that may be relevant to the issue of absent scopal interactions. The antecedent to the VPE in a wide scope ACD is restricted to the first VP dominating the finite clause containing the quantifier phrase. While (83a,b,c) are all possible, (84) is not:

(83)a John thinks that more trees died than Mary does _
b. John thinks that more trees seem to have died than Mary does _
c. John thinks that it seems that more trees have died than Mary thinks it does _

(84) * John thinks that it seems that more trees died than Mary does _

In each of (83a,b,c), the constituent (*more trees...*) that undergoes QR is located in the clause below the VP-antecedent to the VPE in the comparative clause:

(85)a John [thinks [that more trees died than Mary does *think died*]]
b. John
[thinks [that more trees seem to have died than Mary does *think seem to have died*]]
c. John thinks that it
[seems [that more trees have died than Mary thinks it does *seem have died*]]

In (84), the constituent (*more trees...*) that undergoes QR is located in the second clause below the VP-antecedent to the VPE in the comparative clause:

(86) * John
[thinks [that it seems [that more trees died than Mary does *think that it seems died*]]]

This paradigm suggest that "long-distance" QR is strictly local, in the sense that it may not raise a constituent further than the VP immediately dominating the first finite clause containing it. If this is correct, then the other facts discussed in this section would fall out, if the theory of scopal interaction is designed in such a way that a phrase raised from a finite complement and adjoined to VP cannot interact scopally with the arguments of the verb heading that VP.

5. German comparatives

Cross-linguistic support for the account of wide-scope VP-ellipsis sketched here is not easy to come by, since many languages do not show VP-ellipsis in the English pattern at all, or only to a much more restricted extent. However, a similar contrast to the one discussed in section 3. is found in an elliptical comparative construction that is not confined to English.

Consider first the English examples (87)-(88).

(87) Peter thought about it more often than Mary

(88) a. Peter thought that more trees had died than Mary
b * Peter thought that more trees than Mary had died

Assuming that the complement to *than* in (87), superficially a bare DP, is underlyingly clausal, this clause undergoes ellipsis which deletes all daughters of I', leaving the subject as the sole remnant. (87) thus has an abstract representation like (89):

(89) Peter **thought about it** more **often** than Mary *thought about it x-often*

Assuming further that I'-ellipsis, like VP-ellipsis, is subject to the PF-constraint (60), it is clear that (60) is satisfied by (89).

Where the elliptical *than*-clause is associated with an embedded nominative, wide-scope interpretation is possible only when the comparative clause is extraposed (89a). (89b) only has the nonsensical interpretation "more trees than just Mary died", i.e. that Mary is a dead tree. The contrast can be accounted for if the examples (88) have representations as in (90).

(90) a. Peter
thought [that [more **trees** _] **had died** than Mary *thought that x-many trees had died*]

b. * Peter
thought [that [more **trees** than Mary *thought x-many trees had died*] **had died**]

(90b) violates the PF-constraint (60), while (90a) does not. (90a) also avoids antecedent-containment at LF, assuming that the extraposed elliptical *than*-clause is reconstructed into the comparative DP, yielding a constituent *more trees than Mary thought that x-many trees died* which then undergoes QR to a position above the matrix verb *thought*. Given that the same QR option is available for (90b), the ungrammaticality of this example can be taken as support for the hypothesis that I'-ellipsis is subject to (60).

Unlike VP-ellipsis, I'-ellipsis in comparative clauses is found in other languages, for example, German. The contrast in (88) can be reproduced for German, as illustrated in (91): [18]

(91) a. * weil Peter meinte, daß es öfter als ich geregnet hat
since Peter said that it more-often than I rained has

b. weil Peter meinte, daß es öfter geregnet hat als ich
"since Peter said that it rained more often than I (did)"

The German comparative morpheme selects the complementizer *als* which heads the comparative clause. [19] This clause may undergo I'-ellipsis, with the result that only the subject DP is overtly expressed. In simple clauses, in contrast to (91), extraposition of the *als*-clause (subject to the Right Roof constraint) is possible but not obligatory, regardless of whether it is elliptical or not, as shown in (92)-(93):

(92) a. ...weil er öfter als ich darüber nachgedacht hat
since he more-often than I there-over thought has
"since he has thought about this more often than I (have)"

b. ...weil er öfter darüber nachgedacht hat als ich

(93) a. ...weil er öfter, als ich darüber nachgedacht habe, darüber nachgedacht hat
"since he has thought about this more often than I have thought about it"

b. ...weil er öfter darüber nachgedacht hat, als ich darüber nachgedacht habe.

The examples (92) have the analysis (94). Neither (94a) nor (94b) violate (60):

(94) a. weil er [öfter als ich *x-oft darüber nachgedacht habe*] **darüber nachgedacht hat**
b. weil er [öfter _] **darüber nachgedacht hat** [als ich *x-oft darüber nachgedacht habe*]

By the same reasoning, the examples (91) are represented as (95):

(95) a. * weil Peter
meinte, daß es [**öfter** als ich *meinte daß es x-oft geregnet hat*] **geregnet hat**

b. weil Peter
meinte, daß es [**öfter** _] **geregnet hat** [als ich *meinte daß es x-oft geregnet hat*]

(95b) does not violate (60), but (95a) does. Assuming that in (95b), the extraposed clause reconstructs into the phrase headed by *öfter*, with the latter undergoing QR to a position above the matrix verb *meinte*, antecedent-containment is avoided at LF. Given that the second step, i.e. QR, is equally available to (95a), the ungrammaticality of (95a) provides support for the claim that (60) applies also in German. [20]

6. VP-ellipsis and string-containment

The account in the previous sections of the contrasts concerning wide scope ellipsis depended crucially on the PF-constraint (60). Now we return to the examples (96) (=65), which were put aside in section 3.2.

(96) a. John gave whoever he could _ two dollars
b. ? John finds everyone I do _ easy to work with.
c. ? John believes everyone I do _ to be a genius

The question is, given that constraint (60) excludes examples with an ellipsis site (properly) contained within the string corresponding to its antecedent in surface order, how do we account for good string-contained ACDs?

The idea I pursue here is that examples of this type are not ACD's at all. Rather, they involve two adjacent deletions, neither antecedent-contained in surface order, one due to Pseudogapping (ellipsis of subparts of VP), the other due to the independent process of Backward Deletion (BWD). The examples in (96) are therefore related to examples like (97), where part of VP in a relative clause undergoes deletion under identity with material belonging to the VP containing the object which that relative clause modifies:

(97) a. John gave [whoever Mary did _ two dollars] three dollars
b. John finds [everyone I do _ hard to work with] easy to work with.
c. John believes [everyone I do _ to be a fool] to be a genius

The claim is that (96) involves the same partial VP-Ellipsis as that in (97). In (96), this partial VPE interacts with Backward Deletion of the right-remnant of the embedded VP, under identity with material belonging to the matrix VP which follows the relative clause in surface order. The examples in (96) thus have abstract representations as in (98):

(98) a. John gave [whoever he could _ ~~two dollars~~] two dollars
b. John finds [everyone I do _ ~~easy to work with~~] easy to work with
c. John believes [everyone I do _ ~~to be a genius~~] to be a genius

Neither deletion site in (98) is properly contained within the string corresponding to its antecedent, in surface order.

Put another way, the claim is that the derivation of (96a) that involves VP-ellipsis (99a) is ill-formed, since it violates (60); while the derivation of the same string involving partial VP-ellipsis and BWD (99b) violates neither (60), nor any other constraint: [21]

(99) a. * John **gave** [whoever he could *give t two dollars*] **two dollars**
b. John **gave** [whoever he could *give t* ~~two dollars~~] two dollars

If this is correct, then the examples (96) have been mis-identified as ACDs arising from VP-ellipsis. What they actually indicate is that independent deletion processes may interact to give rise

to adjacent ellipsis sites. Hence, these examples have no direct bearing on the constraints governing antecedent-containment; instead, they relate to the question of how apparent VP-deletions are to be analyzed in the first place (cf. section 1.3).

To make this idea plausible, it is minimally necessary to show firstly, that in the good examples, the component deletions meet independently motivated constraints holding of each deletion process individually; and secondly, that individual violation of any such constraint leads to ungrammaticality, i.e. a bad derivation. The task is made more complex by the fact that for a given example, there may be more than one derivation to consider (cf. (99)).

In the next section, constraints on partial VPE are discussed. In section 6.2., constraints on BWD are introduced. Interactions between them are discussed in section 6.3.

6.1 Partial VPE (pseudo-gapping)

Attempts establish the nature of constraints on partial VPE are is hampered by the fact that "pseudogapping" examples are generally felt to substandard. This is especially true of examples where a verb is elided, stranding its direct object (the cases of interest here):

(100) ? John read the book and Mary did, the newspaper

Examples where the verb and its arguments (the traditional "V'-constituent") are elided, stranding peripheral adjuncts, tend to be far more acceptable:

(101)a. John read the book today and Mary did, yesterday
b. John put some money in the bank today and Mary did, yesterday.
c. ?* John put some money in the cashbox and Mary did, in the till.

However, the argument-adjunct contrast tends to disappear where remnants are finite clauses or PPs:

(102)a. John wrote that Bill was innocent and Mary did, that he was innocent.
b. John requested that Bill be sacked, and Mary did, that he be retained.

(103)a. If you'll explain this to the pupils, then I will, to their parents
b. John's staying with Mary and Bill is, with Sue

Occasionally, one encounters examples with stranded direct objects that are spotless ((104a) due to Sten Vikner), especially in comparatives:

(104)a. It is harder to persuade the Americans of this than it is, the English.
b. John started his paper at the same time as Mary did, her book.

Here, descriptive generalizations are proposed, drawing on comparative judgments on relevant example types - the aim being to establish where the line is to be drawn between grammatical examples of low acceptability, and genuinely impossible (ungrammatical) cases. The limits to pseudo-gapping so established can then be used in excluding otherwise possible VPE-BWD interactions. (I know of no theoretical account that accurately captures the distribution of generalizations given here).

There are two basic generalization at stake. The first is that no constituent of a VP may be deleted, unless the verb is itself deleted: the minimal VPE target consists of V alone (never of an object alone, for example). Hence the deletion of V is 'obligatory' (105a), the possibility for deletion of further constituents (105b) dependent on deletion of the verb. The second generalization is that pseudogapping may not target a string crossing a DP, PP or finite clause boundary (105c

(105) VPE targets:
a. V
b. optionally, DP, PP, finite CP
c. no subpart of DP, PP, finite CP

Hence, remnants tend to be complete arguments or adjuncts to the deleted verb. (106) illustrates ellipsis of a ditransitive V, together with its goal DP (106a), and its theme NP (106b):

(106)a. He gave them two dollars, and she did, three dollars (*give them...*)
b. ? He gave Peter two dollars, and she did, Mary (*give ... two dollars*)

In small clause and ECM constructions, the main verb may be deleted together with a subpart of its complement - the accusative subject (107a), or the accusative subject and the embedded verb (107b):

(107)a. ? He finds me stupid, and she does, smart (*finds me ...*)
b. ?? He believes them to like beans, and she does, spaghetti (*believe them to like...*)

The impossibility of deleting part of a finite clausal complement is illustrated in (108): [22]

(108)a. * I believe that John is a fool, and she does, is a genius (**believe that John...*)
b. * He believes that they like beans, and she does, spaghetti (**believe that they like..*)

A constraint which plays an important role in section 6.3 is that pseudogapping of a verb that is left-adjacent to a finite clause lacking a complementizer is impossible. This is so even in contexts where 'complementizer deletion' is otherwise possible: [23]

(109)a. * Mary believes (that) Peter is a fool and Bill does, he is a genius
b. Mary believes Peter is a fool and Bill does, that he is a genius

The ban on deletion of a complementizer following a gapped verb is independent of whether the complementizer of the complement of the antecedent verb is deleted (109b). A subcase of this generalization is illustrated in (110a). (110) shows that a verb whose complement contains a subject trace cannot undergo pseudogapping:

(110)a. * someone who he believes is a fool and (who) she does, is a genius.
b. * someone who he believes is a fool and (who) she does, that is a genius.

This fact follows from two conflicting requirements - the presence of the complementizer required by the constraint on pseudogapping induces a violation of whatever underlies the "that-trace filter". Where the latter is inoperative (object extraction), the complementizer may be overt, and pseudogapping is possible:

(111) someone who he believes is a fool and (who) she does, that we should hire.

Finally, recall from section 1.3 the question-mark raised, by the possibility for partial VPE, over the antecedent-contained status of ACDs. Given that discontinuous gaps can arise, it may be that apparent VP-gaps are actually a sequence of adjacent, smaller gaps. Then the antecedent for

each gap would be a correspondingly smaller constituent, possibly never containing its dependent gap (in surface structure).

An argument that genuine antecedent-containment arises at S-structure comes from considering wide scope ACDs. The fact that CP cannot be 'cut into' can be used to deny the possibility of analyzing the deleted string corresponding to CP as a collection of individual deletion sites. If the deletion site in (112a) is dismantled into its smallest independently motivated parts, the deleted CP is still a single deletion site (112b); the analysis (112c), for instance, not meeting the "major constituent" condition on VPE:

(112)a. John thinks everyone is intelligent that Mary does __
b. ... that Mary does [*think*] [*t is intelligent*]
c. * ... that Mary does [*think*] *t* [*is*] [*intelligent*]
d. John [$_{V2}$**thinks**]
[$_{CP2}$ everyone **is intelligent** that Mary does [$_{V1}$*think*] [$_{CP1}$ *t is intelligent*$_{CP1}$]$_{CP2}$]

Assuming the Right Roof Constraint to be valid, then that deletion site (CP) must be hierarchically contained in its antecedent (at S-structure) (112d).

6.2 Backward Deletion

The Backward Deletion (BWD) operation assumed here covers what is normally termed "Right Node Raising", most commonly found in coordinate structures (113a). Since we are dealing with noncoordinate structures, it is important to realize that the process also applies in noncoordinate structures, as in (113b) (from Williams 1990) and (113c):

(113)a. John talked to __) and really got to like my new friend from Kansas)
b. Anyone who talks to ___) really gets to like my new friend from Kansas)
c. ...go from the last town north __) to the first town south of that mosquito-infested river)

The conditions on BWD that are relevant below are: [24]

(114) a. the deletion site must be at right edge of its "domain"
b. the antecedent must be at right edge of its "domain"

There is no satisfactory account of what counts as a domain for BWD, that covers both coordinate and noncoordinate types. [25] Considering BWD in coordination, the domains mentioned in (114) must be conjuncts; moreover, these domains must be adjacent, with the conjunct containing the deletion site immediately preceding the conjunct containing its "antecedent" (hence the backward directionality). Assuming a view of phrase structure as proposed in Kayne (1994), including a Larsonian view of the internal structure of complex VPs, one can generalize from the coordinate to the noncoordinate cases by stating that the domain for the deletion site is a specifier, and the domain for its antecedent is the X'-sister of that specifier:

(115) ... [$_{XP}$ [**SPEC** ... ~~YP~~] [$_{X'}$ ... YP]] ...

Under the analysis of conjoined structures in which the conjunction is a head taking conjuncts as its complement and specifier, coordinate BWD configurations realize a specific variant of (115):

(116) [$_{\&P}$ [**SPEC=CP** ... ~~YP~~] [$_{\&'}$ and [$_{CP}$ ... YP]]

The non-coordinate cases also fit the scheme. (113b) involves the specifier of IP as the deletion domain, and I' as the antecedent domain (117a); (113c) might involve two PPs embedded in a Larsonian VP (117b):

(117)a. $[_{IP}\ [_{\textbf{SPEC=DP}}\ \ldots\ \sout{YP}\ ]\ [_{I'}\ \ldots\ \underline{YP}\ ]\ ]$

b. $\ldots\ [_{VP}\ [_{\textbf{SPEC=PP}}\ \ldots\ \sout{YP}\ ]\ [_{V'}\ \ldots\ [_{PP}\ \ldots\ \underline{YP}\ ]\ ]\ ]$

The cases we are interested in involve deletion in a relative clause modifiying a DP that is nonfinal in VP (118a). Under Larsonian assumptions, that relative clause forms the right edge of a complex DP in a specifier of an internal V'-constituent, as in (118b):

(118)a. John gave [whoever he could _ ~~two dollars~~] two dollars

$\ldots\ [_{VP}\ [_{\textbf{SPEC=DP}}\ \ldots\ [_{CP}\ \ldots\ \sout{YP}\ ]\ ]\ [_{V'}\ \ldots\ \underline{YP}\ ]\ ]\ \ldots$

In the following, whether (115) is correct or not, I assume that the right edge of the BWD domain must coincide with the right edge of the DP containing the relative clause. [26]

6.3 VPE / BWD interactions

In order for the account of "string contained" ACD's such as (96) as adjacent VPE and BWD deletions (98) to go through, it is necessary minimally to show independent word order rules interact with partial VPE and BWD to give the right results for good "containment" examples.

Beyond that, it is desirable to show that an independently motivated failure of partial VPE can be used to explain the ungrammaticality of a "string-contained" ACD; and likewise for BWD, to show that the ungrammaticality of a "string-contained" ACD can be attributed to solely to the violation of some constraint on BWD. The latter two points may of course be impossible to establish, insofar as the existence of a "bad" derivation may be obscured by a "good" derivation by another route, of the same string. Nonetheless, as will be shown below, all three goals can be secured: [27]

(I) For each "good" string-contained ACD, there is at least one derivation involving (partial) VPE, BWD, applying on independently well-formed surface word order
(II) If VPE fails, then the derivation fails
(III) If BWD fails, then the derivation fails

- *ad* (I):

All three examples (96a-c) have a derivation involving partial VPE and BWD and meeting requirements of both:

(119)a. John **gave** [whoever he could *give t* ~~two dollars~~] <u>two dollars</u>
b. John **finds** [everyone I do *find t* ~~easy to work with~~] <u>easy to work with</u>
c. John **believes** [everyone I do *believe t* ~~to be a genius~~] <u>to be a genius</u>

Partial VPE is independently established by the (relatively) acceptable examples (97) without BWD. BWD itself meets the requirements (114): the deletion target is right-peripheral in its domain (the DP containing the relative clause), and the antecedent to the deletion is right-peripheral in <u>its</u> domain - the matrix VP (strictly speaking, the V' sister of the DP containing the relative clause).

- *ad* (II):

For the original example (120) (=62), whose ungrammaticality was attributed to the PF condition (60) under the VPE-derivation (121a), it is now necessary also to exclude an alternative derivation, involving partial VPE interacting with BWD (121b):

(120) * John believes everyone I do, is a genius

(121)a. * John **believes** [everyone I do *believe t is a genius*] **is a genius**]
b. * John **believes** [everyone I do *believe* ~~is a genius~~] is a genius

In (121b), partial BWD targeting the string *is a genius* at the right edge of the relative clause, under identity with the same string at the right edge of the complement to the matrix *believe* should be possible.

In this derivation, however, partial VPE fails, for reasons discussed above: a verb whose complement clause contains a subject trace cannot undergo pseudogapping (109). Inserting a complementizer in the antecedent does nothing to save the example - the complementizer governed by *believe* in the relative clause cannot be included in a partial VPE, since VPE cannot delete "into" a finite clause: [28]

(122) * John **believes** [**that** [everyone I do *believe* [*that t* ~~is a genius~~]] is a genius]

Nor may the complementizer in the relative clause be left "undeleted" (to save pseudogapping) on pain of a *that*-trace violation:

(123)a. * John believes that everyone I do, that, is a genius
b. * John **believes that** [everyone I do *believe* that *t* ~~is a genius~~] is a genius

Hence, it can be concluded that an independently motivated constraint on partial VPE does real work in correctly excluding "string-contained" ACD's.

- *ad* (III):

To show that failure of BWD can lead to ungrammaticality in the realm of "string-contained" ACD's, it is sufficient to consider examples in which the part of the gap that is due to BWD (i.e. the part whose antecedent follows the gap) is itself not right-peripheral in the relative clause. Such examples are rather complex, and judgements subtle, but the contrasts are clear enough. Consider the paradigm (124):

(124)a. ? John believes [everyone I do _ to like spaghetti], to like beans
b. ? John believes [everyone I do_] to like beans
c. * John believes [everyone I do _ spaghetti], to like beans
d. ? John believes to like beans [everyone I do _ spaghetti]

(124a) involves partial VPE of *believe* in an ECM construction (cf. (97c) above). The same partial VPE is involved in (124b); the remainder of the gap (*to like beans*) is BWD-site, whose antecedent follows the gap, as in (124b'). The competing derivation (124b") which treats the whole gap as a VPE-site, violates string-containment (60):

(124)b'. John **believes** everyone I do *believe t* ~~to like beans~~, to like beans

b" * John **believes** everyone I do *believe t to like beans*, **to like beans** (**PF-containment*)

(124c) has no good derivation. Treating the deletion of *believes to like* as partial VPE as in (124c') - possible, given (107b) above - falls foul of the PF-condition (60). The alternative (124c") treats the gap as two deletions: pseudogapping of *believe*, as in (124a), and BWD of *to like*. The latter - crucially - violates the right-edge condition on BWD-sites (114a): the deletion is not right-peripheral in its domain (=the relative clause), since it is followed by non-deleted *spaghetti*:

(124)c'. * John **believes** everyone I do *believe to like* spaghetti, **to like** beans (**PF-containment*)
c". * John **believes** everyone I do *believe* ~~to like~~ spaghetti), <u>to like</u> beans) (**BWD*)

It is possible for the deletion site to be displaced to the right, by Heavy NP Shift of the accusative subject containing the relative clause, as in (124d). Crucially, now, the gap in the relative clause is no longer followed by part of its antecedent in surface order. Hence, there is a derivation involving partial VPE of the whole gap (possible, as in (107b)) which does not violate (60); and BWD is not necessary to derive any part of the gap:

(124) d'. John **believes to like** beans everyone I do *believe to like* spaghetti

Analogous paradigms are given in (125) and (126). The offending BWD-site in (125c) is *in the fridge*, the PP-complement of *put*; the element blocking right-peripherality of that BWD-site is the temproal adverb *at night*, distinct from the matrix *during the day* - hence necessarily part of the matrix. In (126c), BWD of *three dollars* is blocked by embedded *yesterday* - again, distinct from the matrix *today*, hence necessarily part of the relative clause. Reordering the gap to the right of the antecedent string in (125d) and (126d) again permits the gap to escape the conspiracy of the PF-constraint (60) and the BWD peripherality requirement (114a): [29]

(125)a. John puts [everything I do _ on the balcony at night] in the fridge during the day
b. John puts [everything I do _] in the fridge during the day
c. * John puts [everything I do _ at night] in the fridge during the day
d. John puts in the fridge during the day [everything I do _ at night]

(126)a. ? John gave [everyone that I did _ two dollars yesterday] three dollars today
b. ? John gave [everyone that I did _] three dollars today
c. * John gave [everyone that I did _ yesterday] three dollars today
d. John gave three dollars today [to everyone that I did _ yesterday]

From these paradigms, it can be concluded that constraints on BWD are independently needed to derive acceptability patterns among "string-contained" ACD's.

In sum, the contention that such examples do not constitute genuine ACD's finds subtle, hence strong, empirical support.

References

Baltin, M. 1987 Do antecedent-contained deletions exist? Linguistic Inquiry 18. 537-577.

Chomsky, N. 1993 A Minimalist Program for Linguistic Theory. In: K. Hale & S. Keyser (eds.) The View from Building 20. MIT Press, Cambridge MA. 1-53.

Chomsky, N. 1995 The Minimalist Program. MIT Press: Cambridge, Mass.

Chomsky, N. & H. Lasnik 1993 Principles and Parameters Theory. In: Jacobs, J., W. Sternefeld, A. von Stechow & T. Vennemann (eds) Syntax. An International Handbook., De Gruyter, Berlin. 506-569.

Dalrymple, M., S. Schieber & F. Pereira 1991 Ellipsis and Higher-Order Unification. Linguistics and Philosophy 14.
Diesing, M. 1992 Indefinites. MIT Press: Cambridge, Mass.
Fox, D. 1994 Quantifier Scope in VP Ellipsis. Ms. MIT
Guéron, J. & R. May 1984 Extraposition and Logical Form. Linguistic Inquiry 15, 1-31.
Haik, I. 1987 Bound VPs that Need to Be. Linguistics & Philosophy 10, 503-530.
Hornstein, N. 1994 An argument for Minimalism - the case of Antecedent-Contained Deletion. Linguistic Inquiry 25. 455-480.
Kayne, R. 1994 The Antisymmetry of Syntax. MIT Press: Cambridge, Mass.
Larson R. & R. May 1990 Antecedent containment or vacuous movement? Reply to Baltin. Linguistic Inquiry 21. 103-122.
Lasnik, H. 1995 Lecture Notes. Berlin.
May, R. 1985 Logical Form: Its Structure and Derivation. MIT Press: Cambridge, Mass.
Milsark, G. 1977 Towards an explanation of certain peculiarities of the existential construction in English. Linguistic Analysis 3. 1-29.
Sag, I. 1976 Deletion and Logical Form. PhD, MIT.
Stowell, T. 1981 The Origins of Phrase Structure. PhD, MIT.
Tancredi, C. 1992 Deletion, Deaccenting and Presupposition. PhD, MIT.
Vanden Wyngaerd, G. & J.-W. Zwart 1991 Reconstruction and Vehicle Change. In: A. van Kemenade (ed.) Linguistics in the Netherlands 1991.
Wilder, C. 1995a Some Properties of Ellipsis in Coordination. Geneva Generative Papers 2.
Wilder, C. 1995b Rightward Movement as Leftward Deletion. To appear in: Lütz, U. & J. Pafel (eds.) Extraposition and Extraction in German. John Benjamins: Amsterdam.
Williams, E. 1977 Discourse and Logical Form. Linguistic Inquiry 8, 101-139.
Williams, E. 1990 The ATB Theory of Parasitic Gaps. The Linguistic Review 6, 265-279.

[1] This is a preliminary written version of material that was presented in colloquia in Berlin (July and November 1995) and Potsdam (November 1995). Thanks to those audiences, and also to the following individuals, for helpful discussions and comment: Manfred Bierwisch, Daniel Büring, Marcel Den Dikken, Gisbert Fanselow, Hans-Martin Gärtner, Gerhard Jäger, Peter Ludlow, Anoop Mahajan, Robert May, Tom Roeper, Michal Starke, Markus Steinbach, Anatoli Strigin, Ralf Vogel, Sten Vikner, Jan-Wouter Zwart. The paper was indirectly inspired by Howard Lasnik's lectures on Ellipsis and QR in Berlin in June 1995.

[2] Even the simplest examples such as (2) demonstrate that phonological identity is not required - *meet* is not identical with *met*.

[3] I leave a third proposal - that VPE is some kind of pronominal element with no internal structure at any syntactic level (cf. Dalrymple et. al. 1991) - out of consideration here.

[4] Although *near* can be stranded under wh-movement (i), the combination with pseudogapping in a relative clause (ii) is not good, for unclear reasons. Pied-piping of the preposition (iii) is much better:
(i) ? John stood near Sue, who Mary stood near, too
(ii) * John stood near Sue, who Mary did near too
(iii) ? John stood near Sue, near whom Mary did too

[5] Baltin does not comment on whether in addition to (27), there is a filter banning antecedent containment at LF. Nor does he make explicit whether he endorses QR and LF-copying as the mechanism behind VPE-interpretation in these examples; but it is difficult to see what else could have been intended.

[6] An analysis of (31a) in which both complements have shifted rightwards, vacating VP, as in (i), is precluded by independent considerations:
(i) John [$_{VP}$ gave t_j t_k] [whoever he could _]$_j$ [two dollars]$_k$
While RCE may apply to a goal NP, splitting off the relative clause from the NP-head as in (ii), free relatives are never split by RCE (iii). In argument positions, free relatives behave like NPs; in particular, in the "goal" position in the double object construction, neither undergoes rightward displacement of the Heavy NP Shift type (iv,v).
(ii) John gave [everyone _] two dollars who he met at the party
(iii) * John gave [whoever_]two dollars he met at the party
(iv) * John gave _ two dollars whoever he met at the party
(v) * John gave _ two dollars the boy next door
Hence, under standard assumptions, the VPE in (31) is contained within its antecedent VP at S-structure.

[7] Given the VP-internal subject hypothesis, the VP that is copied in (34c) contains an A-trace left by the main clause subject *John*, which must function in the relative clause as the A-trace of the subject *Mary*. It must therefore be assumed that NP-traces count as identical for purposes of VPE. This assumption is necessary in any case for the simplest cases: cf. (i), where the trace left by movement of John in the first conjunct must be assumed to be identical with the NP-trace forming a chain with Mary in the second conjunct:
(i) John was [kissed t] and Mary was _ too

[8] One issue concerns the interaction of V-movement with VPE in the Minimalist framework, where it is assumed that finite verb raise to the highest Infl-head (AGs) in LF. If it is assumed (i) that the chain of the verb must be included in the constituent that matches (is copied into) the VPE, and (ii) copying / matching follows all LF-movement, then the account must give up the assumption that finite V raises to AGs. Otherwise, the smallest constituent in the antecedent that matches (can be copied into) the VPE would be AGs', which by hypothesis contains AGoP, hence also the VPE itself, even after object raising. Hornstein suggests that English V does not raise higher than AGRo at LF, and that the target of "VP-Ellipsis" is in fact AGo', including the raised verb but excluding the raised object.

[9] Chomsky (1995:Ch4) has since developed an account of covert A-movement (object raising) that does not involve pied-piping of the whole object phrase.

[10] Example (31a), with the VPE contained within the goal NP object of give in a double object construction, can plausibly be handled via A-movement.

[11] (39b) is slightly degraded in Lasnik's judgement. I find the example perfect.

[12] There is no narrow scope reading for the VPE in these examples: *be + XP* (*is a genius / were stupid*) cannot reconstruct under *do / did.*

[13] It must be assumed in addition that the extraposed relative reconstructs into the subject prior to QR. See section 3.2.

[14] It has been proposed that QR may extend the domain of extraposition, such that the extraposed clause may cross a sentence boundary rightwards, in violation of the Right Roof Constraint, just in case the head of the extraposed clause itself raises the same distance leftwards - cf. Guéron & May (1984). But nowhere has it been proposed that extraposition may violate the right-roof constraint in order to extend the domain of QR.

[15] This discussion presupposes that (60) belongs to UG. One may speculate that the facts (60) is intended to capture may in fact not be linguistic at all, but rather reflect an extralinguistic condition, perhaps relating to the parsing mechanism.

[16] Thanks to Anoop Mahajan and H-M. Gärtner for directing me to this reference. It is important to realize that the non-presuppositional reading is not to be equated with the narrow scope reading in sentences with an additional quantifier. Rather, only in its presuppositional reading does a weak quantifier undergo QR, in which case it may take narrow or wide scope with respect to the second quantifier.

[17] I ignore here the possiblity that the raised DP adjoins to the matrix V'.

[18] Some speakers find (91b) degraded, but the contrast with (91a) seems to be robust.

[19] The precise categorization of *than / als* (e.g. clause-governing preposition or clausal head C) is not crucial to the argument.

[20] This account of the contrast in (88) (English), and between (94) and (95) (German) depends crucially on the assumption that the complements of *than* and *als* in the comparative is <u>always</u> clausal, with the bare DP in these cases being left behind by ellipsis. If there is the possibility for *than / als* to govern a bare DP rather than a clause in the syntax, then there is the possibility for an abstract representation of (88) or (95) without ellipsis, and the constrast goes unexplained. This assumption has further consequences regarding the notion of identity required for ellipsis. In (i), the DP *Peter*, in the complement of *than*, is interpreted as the subject of a VP - more correctly, an I'-constituent - whose antecedent is nonfinite. Given that the DP *Peter* is overt, it must be in a Case position (in terms of Chomsky&Lasnik (1993), not a position in which "null Case" is licensed). Hence this DP cannot be the subject of a non-finite IP (ii). It must be concluded that a finite I' can be deleted under identity with a finite I' (iii). The same goes for German (iv-vi):

(i) to eat more apples than Peter is not advisable
(ii) * [PRO **to eat more apples** than Peter *to eat x-many apples*] , is not advisable
(iii) [PRO **to eat more apples** than Peter *eats x-many apples*] , is not advisable
(iv) mehr Äpfel zu essen als Peter, ist nicht ratsam
more apples to eat than Peter is not advisable
(v) * [PRO **mehr Äpfel zu essen** als Peter *x-viele Äpfel zu essen*] , ist nicht ratsam
(vi) [PRO **mehr Äpfel zu essen** als Peter *x-viele Äpfel ißt*] , ist nicht ratsam

[21] VP-ellipsis sites, including partial VPE, are given in italics; their antecedents in bold. BWD-dependencies are indicated by crossing-through of items deleted under identity with underlined items.

[22] Notice that two independent (partial) VPE operations can coexist in a complex structure, even a verb-complement construction such as (iii):
(i) John will leave, if you come, and Mary will, if I do
(ii) John hates anyone who works for Labour and Mary does, anyone who does, for the Tories.
(iii) John believes that Mary likes beans, and Bill does, that Sue does, spaghetti
The last example does not counterexemplify the generalization in the text. The deletion of the lower finite verb "begins" a new (partial) VPE domain.

[23] This fact is probably related to Stowell's (1981) observation that a deleted complementizer has to be adjacent to its governing verb: cf. *Mary believes sincerely *(that) Peter is a fool.*

[24] See Wilder (1995a, 1995b) for fuller discussion of BWD.

[25] Williams (1990) suggests that noncoordinate RNR contexts are really conjoined structures - subject conjoined with predicate, etc. I reject this view.

[26] In (118a), the antecedent string exhausts the overt terminals of the antecedent domain (=V'). The converse situation - where the deletion site exhausts the deletion domain (=Spec), must probably be excluded. Kayne (1994), Wilder (1995b) assume that a deletion site may not c-command its "antecedent".

[27] If (I)-(III) are not met, we are faced with a problem of undergeneration, avoidance of which is a *sine qua non.* Independently of this, questions of overgeneration also arise. For example, Robert May (p.c.) asks why BWD cannot apply independently of partial VPE in examples like (i), pointing to the (near)-unacceptable status of (ii):
(i) John gave whoever he could, two dollars.
(ii) *? John gave whoever he could give, two dollars.
(iii) John gave [whoever he could give *t* ~~two dollars~~] <u>two dollars</u>
(iv) * John gave [whoever he could give *t* two dollars]
The derivation in (iv), with overt *two dollars* inside the relative clause, has the status of a projection principle violation; there is no deletion rule that can yield deletion of the second object of the matrix verb *gave*, hence it cannot have been projected into the syntax. The BWD derivation (iii), with overt *two dollars* in the matrix VP, requires a prosodic boundary between *give* and *two dollars*, indicated by the comma in (i). One factor contributing to the deviant status of (ii) may be the identity of the verbs, which makes partial VPE in (i) possible in the first place. There is some evidence from coordination that there is an economy-like principle which prefers derivations with deletion over derivations without ("if a deletion can apply, it must"); consider the paradigm (v-vii):
(v) John switched the light off and Mary the TV (*switched+off*)
(vi) * John switched the light off and Mary the TV off (*switched*)
(vii) ? John switched the light off and Mary the TV on (*switched*)
If non-identical verbs are chosen in the construction (ii), the example seems to improve somewhat:
(viii)?? John sent whoever he couldn't give, two dollars.
Independently of these considerations, the problem of demarcating the applicability of (non-coordinate) RNR is in any case unresolved, cf. Williams (1990), Wilder (1995a).

[28] In fact, in the version of this example that avoids the PF constraint (60) by extraposition, it makes no difference whether or not the antecedent contains overt *that*, although the appearance of overt that in the relative clause would cause a *that*-trace violation:
(i) John believes that everyone is intelligent that I do _
(ii) * John believes that everyone is intelligent that I believe that is intelligent
From this example alone it is impossible to decide whether (i) the 'overt' *that* is reconstructed into the VPE site, with the *that*-trace filter only applicable to overt *that*; or (ii) the 'deleted' complementizer is treated as nondistinct from the overt complementizer for purposes of VPE-identity. Other examples indicate the former. Overt verbs of "manner-of-speaking" such as *mutter*, for example, do not license the 'deleted' complementizer:
(iii) * John muttered everyone is intelligent

However, wide scope VPE, involving reconstruction of *mutter* together with its complement CP containing a subject trace, is possible:

(iv) John muttered that everyone was intelligent that I did.

(v) * John muttered that everyone was intelligent that I did mutter that was intelligent

(vi) John [**muttered** [**that** [everyone _] **was intelligent** [that I did *mutter that t was intelligent*]]]

(vi) shows that the that-trace effect can be voided in VPE contexts for the verbs in question. The *that*-trace effect is a phenomenon sensitive solely to the PF-status (overtly realized or not) of the complementizer.

29 The bare DP goal object of *give* * does not undergo Heavy NP Shift; hence the effect of reordering the impossible deletion site of (126c) is shown using the variant with *to* in (126d).