

Relating syntactic elements.
Remarks on Norbert Hornstein's "Movement and chains"

Michael Brody
Department of Phonetics and Linguistics
University College London
Gower Street WC1E 6BT
Tel. 44 171 380 7172

and
Institute of Linguistics, Hungarian Academy of Sciences
1014 Budapest, Szinhaz u. 5-9

e-mail m.brody@ling.ucl.ac.uk, brody@nytud.hu

Relating syntactic elements.

Remarks on Norbert Hornstein's "Movement and chains"

Abstract. There are two main arguments in M&C against chains, one based on the correlation of quantifier scope and binding, the other on the correlation between quantifier scope and thematic properties. Both rest on highly dubious background assumptions. Additionally even granting these assumptions both are flawed in similar ways. There are also numerous additional problems both with H's account of quantifier scope and control and with his arguments against chains.

If we take chains to be interpretively constructed, then multiple copies in chains may not differ from other multiple occurrences of lexical items (or structures constructed from lexical items) with respect to their origin. In all cases multiple occurrences are due to a (set of) element(s) being selected from the lexicon more than once. Multiple occurrences can then be interpreted as chains when they are in the (thematically) appropriate type of identity relation. This approach not only dispenses with syntax-internal chains as H justifiably desires, but also eliminates the additional unmotivated, arbitrary and redundant syntactic mechanism of movement that he (and the minimalist framework in general) assumes.

Keywords. movement, chains, quantifier raising, control, A-chains, generalized projection principle

8089 words

1. Some background

Norbert Hornstein's paper (M&C) in a recent issue of this journal (1998) "aims to provoke a debate" (p125). I will present below some considerations that I believe are relevant for evaluating the position outlined in M&C. M&C however does not consider the relative merits of alternative approaches.FN1

I argued in in earlier workFN2 (a) that movement and chains express the same concept and there are no good reasons to postulate both and (b) that there are reasons to prefer the concept of chain to that of move. (a) appears to be a simple and straightforward point with some straightforward consequences, which is nevertheless sometimes ignored and sometimes confused with (b). It is to H's credit that he is willing to call this quite central problem of the minimalist framework a problem, and makes an attempt to resolve it.

It is important to see that the major distinction is not between the position that eliminates move and the one that eliminates chains. This disagreement is quite secondary to the division between those researchers that are still willing to tolerate the (at least by now) conspicuous Occam's razor problem in the centre of a theory that claims to be minimalist and those who attempt to resolve it. Thus in this sense M&C is a contribution to carrying out the program of Brody 1995: the construction of a theory without representational-derivational duplications.

However, unlike the attempts to construct a fully representational or a fully derivational theory (for the latter see e.g. Epstein et al. 1998) H ignores the general problem implicit in the need to choose between movement and chains. Although he wishes to dispense with the concept of chain, he apparently wants to retain the notion of syntactic (LF) representation. But as has been noted (Brody 1995, see also Epstein et al 1998)FN3 in a derivational framework derivations carry the information attributed also to the (LF) representation, thus there are good reasons to be suspicious of a theory that postulates both syntactic derivations and a syntactic output (LF) representation. Arguments would be needed to show that the apparent violation of parsimony is only apparent, --but these are lacking in M&C (and elsewhereFN4). Thus even though H makes an attempt to eliminate the chain-move duplication, unlike others addressing the issue, he does not really try to resolve the dubious architectural redundancies of the standard minimalist framework. FN5

2. Comments on the introduction of M&C

The lighthearted style of H's introductory passages appears to lead him to phrase certain issues in ways that may easily mislead. For example H asks: "Do chains really exist? Are they empirically indispensable constructs? Are they theoretically attractive vehicles for describing and explaining the ineluctable fact that natural language grammars "displace" expressions?" (p.123). The word "displace" in quotation marks in the original cannot be intended as equivalent to the same word without quotation marks since on such a reading the questions H raises would be quite meaningless. If the "ineluctable fact" is that "grammars displace expressions" then we would just need to say that they need to displace expressions rather than express this in what would then be just a(nother) roundabout way, using the notion of chain.

Surely, the ineluctable fact is not that grammars displace expressions but rather that natural languages can relate elements, that in "who did you see" for example "who" and the object position/role of "see" is somehow related. We can then ask further if this relationship is optimally expressed by postulating a representation in which a chain links positions or a derivation in which a movement rule does or perhaps in some other partly or completely different way. So obviously H must be using "'displace' expressions" to mean the linking of positions/elements in this relatively pretheoretical sense. This makes the questions he asks potentially sensible ones.

But H immediately continues with the following statement: "Raising these questions betrays their likely answers." Clearly however raising the sensible question of how linking of elements should be expressed does not "betray [its] likely answer" in the least.FN6

Before examining H's main arguments, let me remark on two more points in his introduction. In the same paragraph in which a footnote clearly states "I have little to say here about the status of A'-chains", he concludes that "chains are superfluous constructs" (p.100). And then adds also: "Not only are chains superfluous, they are actually inadequate." Thus M&C in fact argues only against the existence of A-chains, but somehow wishes to conclude that no chains at all exist. So even if the two major arguments in the paper were correct, it would not remotely follow that chains in general do not exist, --in spite of repeated assertions in M&C to the contrary.

H also notes essentially correctly, that "a strong reading of inclusiveness [of Chomsky 1995] prohibits chains from being grammatical objects or LF interface elements as they are *not* constituents of the lexicon or simple rearrangements of lexical features" (p.101). But again the truth of the sentences that follow as if they described consequences, --does not follow. These statements just presuppose one of the possible answers to the issue at hand: "Chains are pure creatures of the computational

system. They emerge exactly via movement. They appear ...to be... supervenient constructs... if so there should be no grammatical principles stated over chains as such.." Syntax is not equivalent to grammar, and even if we accept inclusiveness for syntax, nothing follows with respect to the existence of chains in other parts of the grammar. Indeed in Brody 1996, 1998 I argued on partly independent grounds that chains are interpretive constructs that involve randomly relating elements at LF under identity and other standard chain conditions. I noted that marking chains at LF violates inclusiveness, and suggested that interpretive chain construction need not involve such marking.FN7

3. The argument from QR

H's major argument for the claim that "at the CI interface an A-chain has one and only one visible link" is based on his earlier reanalysis of QR (Hornstein 1996) as involving A-movement.FN8 He assumes that all but one chain link in each chain needs to be deleted at LF, and the position of the remaining chain link determines quantifier scope. For example for (1a) with the structure in (1b) there will be four interface configuration possibilities as in (2), -- deletion indicated by brackets:

- (1) a. Someone attended every seminar
 b. [_{AgRS} Someone [_{TP} Tns [_{AgRO} every seminar [_{VP} someone attended every seminar]]]]
 (2) a. [_{AgRS} Someone [_{TP} Tns [_{AgRO} every seminar [_{VP} (someone) attended (every seminar)]]]]
 b. [_{AgRS} Someone [_{TP} Tns [_{AgRO} (every seminar) [_{VP} (someone) attended every seminar]]]]
 c. [_{AgRS} (Someone) [_{TP} Tns [_{AgRO} (every seminar) [_{VP} someone attended every seminar]]]]
 d. [_{AgRS} (Someone) [_{TP} Tns [_{AgRO} every seminar [_{VP} someone attended (every seminar)]]]]

In (2a)-(2c) "someone" scopes over "every seminar" since the visible undeleted LF copy of the existential c-commands that of the universal. In (2d) scope relations are inverted, here the remaining copy of the universal c-commands the remaining copy of the existential.

There are many problems discussed in the literature that cast serious doubt on this treatment of quantifier scope cf. eg. Brody 1997a, Beghelli 1996, Kennedy 1997 Wilder 1995, 1997 FN9. For example it is not clear how the analysis is meant to cover quantifiers taking scope across adjunct or NP boundaries, or across various series of nonrestructuring infinitives, etc. For critical comments on H's reanalysis of covert A'-chains as A-movement see Beghelli 1996, Brody 1997a. H seems to consider it out of place to react to objections in M&C. Let us then also put the matter aside here and assume, I think counterfactually,

that the A-movement account of quantifier scope on which H bases his argument against chains is the correct one, --at least in the relevant respects.

H's argument runs like this. Raising constructions like (3b) also allow scope ambiguities like the one in (1)/(2) or (3a):

- (3) a. Someone is reviewing every report
- b. Someone seems (to Bill) to be reviewing every report

The ambiguity of (3b) disappears however when the matrix subject quantifier binds a matrix expression:

- (4) a. Someone_i seems to his_i boss to be reviewing every report
- b. Someone_i seems to himself_i to be reviewing every report

This follows if only a single chain link is visible at LF. This link is either in the matrix clause in which case it can bind the pronoun/anaphor but it has higher scope than the object in the lower clause whose highest position is the embedded spec-AgrO.

- (5) [Someone_i seems to himself_i [(someone) to be [every report [(someone) reviewing (every report)]]]]

If the remaining link of the subject chain is in the lower clause on the other hand, then it can have scope under the embedded object but cannot bind an element of the matrix clause:

- (6) [(Someone_i) seems to himself_i [(someone) to be [every report [someone reviewing (every report)]]]]

The desirable consequence that the binding of an element in the matrix prevents scoping under a quantified object of the embedded clause is lost however, according to H, if chains rather than links are interpreted at LF. The LF structure will then contain copies in all positions FN10 as in (7):

- (7) [Someone_i seems to himself_i [someone to be [every report [someone reviewing every report]]]]

"In [7], the 'someone' chain has 'his' in its scope since one of the links c-commands 'his', viz. the copy of 'someone' in the matrix spec-IP. ...Moreover, it should be possible to interpret 'someone' as under the scope of 'every report', since a link of the 'every report' chain c-commands a link of the 'someone' chain, viz. the head of the former chain c-commands the foot of the latter" (p105).

To see if this argument has force, let us consider first the following question. Does the assumption that chains are present at LF lead to the conclusion that a quantifier Q_1 that has a chain link both above and below another one Q_2 and thus can have scope either above or under Q_2 can also have scope *simultaneously*, ie. within the same interpretation, both above and under the other quantifier? For example "someone" can have scope either above or below the object in (1)/(2) or (3a), the two options corresponding to two distinct interpretations. But can "someone" have scope both above and below the other quantificational expression within a single interpretation? The answer is obviously negative, it would make no sense for Q_2 to be a function of Q_1 and at the same time for Q_1 to be a function of Q_2 , --this would lead to infinite regress. Thus there is no need to add any stipulations or to construe the notion of chains in a particular way to ensure that a quantifier with chain links both above and below another one cannot have scope simultaneously from both positions on the grounds of its chain structure.

But if this is the case, then there are no reasons to expect that a quantifier Q_1 can bind from a chain position that is higher than Q_2 and simultaneously scope under Q_2 from another chain link position. If we take binding as is more or less standard to necessarily involve scope together with some kind of linking, say coindexation, then it is impossible for an element to bind from a position that is higher than its scope.FN11 Hence the impossibility of such a configuration trivially reduces to the impossibility of the simultaneous multiple scope configuration. Binding by Q_1 from a position higher than Q_2 entails scoping of Q_1 from a position higher than Q_2 . It is then impossible by the earlier reasoning for Q_1 to also scope under Q_2 .

In other words, even ignoring the problems with H's treatment of quantifier scope and granting all the assumptions of his theory, his attempted argument against chains will still remain flawed. The possibility the chainless grammar would exclude is necessarily ruled out anyway by independent considerations. H's evidence is therefore not an argument *for* but a potential argument *against* the proposed elimination of chains. This is because all other things being equal we would prefer a syntax that does not redundantly exclude structures or interpretations that the interpretive component will necessarily rule out in any case.FN12

4.Obligatory control

H's second argument is based on his treatment of obligatory control (OC) as A-movement and OC PRO as NP-trace in a chain/move relation that happens to involve more than one theta position. The approach to control that H advocates

also seems seriously problematic. H's arguments for his approach involve primarily on the one hand the similarities in antecedent choice between NP-trace and OC PRO (and local lexical anaphors) and on the other the difficulties in predicting the distribution of PRO. FN13

Similarities in antecedent choice (c-command, locality etc.) between OC PRO, NP-trace and local lexical anaphors have been known as H notes. (Cf. eg. Koster 1981, Manzini 1983 Sportiche 1985, Brody 1985). Additionally it is also clear and has also been generally understood that there are in principle two obvious ways to capture these similarities. Either the relevant instances of movement/chain relations are constrained by the interpretive construal rule relevant for (obligatory control) PRO and (local lexical) anaphors or (at least local lexical) anaphors and (OC) PRO are reduced to movement/chain relations.

Hornstein (1996/98) argues for the latter solution on the grounds that it captures the similarities in antecedent choice. But for quite some time, the problem has not been to find a way to capture the similarities but to decide which of the two available ways of doing so is right (if either). Since it is enough to consider OC PRO to be the same type of element as an NP-trace, say an anaphor, for binding-control theory (cf. eg. Brody 1985), it is not necessary to fully assimilate OC PRO (and local lexical anaphors) to chain/move relations. In order to establish a case for OC PRO being just an NP-trace, it would be necessary to argue that this way of capturing the similarities in antecedent choice is preferable to its alternative(s).

Beyond inconclusive general comments on the status of PRO in the minimalist framework (cf. e.g. FN13 above), H does not attempt to do this. Contrary evidence on the other hand is easy to recall. Take for example locality effects that generalize over non-obligatory control (NOC) and OC PRO. H considers NOC PRO to be an empty pronominal, but in fact this element has no exclusively pronominal properties (Brody 1985) and various anaphoric ones (Manzini 1983, Brody 1985). For example NOC PRO shows clear locality effects, as exemplified in (8) and (9):

(8) John thinks that Mary dislikes PRO teaching
herself/*himself/(?)oneself

(9) John told Mary how PRO to teach
herself/*himself/(?)oneself

As the (perhaps slightly marginal) possibility of the "arb" reading indicates these are NOC structures. (Compare the impossibility of "*John tried to wash oneself".) The effects in (8) and (9) not only exemplify the non-pronominal nature of NOC PRO, they also show that some appropriate version of the intervention constraint/relativized minimality/locality must constrain

antecedent choice for these elements. Call it constraint C. But if constraint C rules out the subject as the antecedent in (9), why should C not be able to do the same in the fully parallel OC example (10)?

(10) John persuaded Mary PRO to teach herself/*himself

H suggests that the minimal distance effect in (10), which prohibits the subject from serving as the antecedent, is due to the Minimal Link Condition (MLC) under his analysis in which the OC PRO in (10) is in fact an NP trace FN14. This looses the generalization that appears to link (9) and (10), since (9) contains a NOC PRO, for H not subject to move/chain. Notice that there is no question of trade-off between approaches here. The alternative approach according to which an interpretive constraint, a rule of construal, applies in both (9) and (10) can capture also the generalization that links (10) and NP-traces under the still quite standard assumption that NP-traces like PRO and lexical anaphors are also taken by the interpretive component to be dependent, anaphoric elements to which this constraint refers.

Hornstein (1996/98) and M&C claim also that assimilating OC PRO and NP-trace helps in accounting for OC PRO's distribution. According to H there remains nothing to account for: OC PRO is just an NP-trace its distribution will be exactly that of an NP trace.

But this apparent advantage is also spurious. Hornstein (1996/98) takes NOC PRO to occur only where OC PRO is not licenced, typically inside islands. As just noted, he proposes furthermore that NOC PRO is an empty pronominal, ie pro, licenced in English only in nontensed subject positions. Ignoring the obviously stipulative nature of this suggestion FN15, observe only, that if statement S is still needed to describe the fact that NOC PRO is restricted to nonfinite subject positions, then making S not to cover also OC PRO, which will then be restricted to the same contexts by different means FN16, is not by itself an achievement.

Again, let's rehearse also a well known piece of contrary evidence. Recall that in principles and parameters theory, the different government requirements on PRO and trace entailed their different distribution, eg. the ungrammaticality of (11) as opposed to the grammaticality of (12c). Note that "attempt" like "believe" (12a, b) obligatorily assigns accusative Case (13a) and can passivize (13b), but unlike "believe" (15), it can also (obligatorily) control (14):

(11) *John was attempted t to leave

(12) a. John believed *(this)

b. This was believed t

c. John was believed t to have left

- (13) a. John attempted *(this)
- b. This was attempted t
- (14) John attempted PRO to leave
- (15) *John believed PRO to have left

It is not clear why (11) is ungrammatical if the distribution of obligatory control PRO and NP trace is in principle the same and PRO is just an NP trace in a chain that happens to involve multiple theta positions. In other words, what is then the difference between "attempt" and "believe" that makes them behave differently from the other in raising-control contexts ((11)/(14) vs (12c)/(15))?

Another old problem is this: if PRO is just an NP-trace then why is (16) ungrammatical?

- (16) *John hit t

And if this has to do with a Case theoretical violation (the impossibility of movement from Case position), as is sometimes suggested FN17, then why is there no verb like "hit", say "HIT" such that it assigns both subject and object theta roles but no Case to its object. This fell under the part of Burzio's generalization that stipulated that subject theta role is dependent on the existence of Case for the object if this is present. This generalization was itself a consequence of chain theory: a Caseless object must form a chain with the subject. So if non chain-root positions cannot be thematic (cf. Chomsky 1982, Brody 1995, 1998), then if the object is Caseless the subject position will be a non chain root position and must be nonthematic. Hence no verb like "HIT" can exist (cf. Chomsky 1982, Brody 1993). This derivation of the relevant part of Burzio's generalization is lost if chains can have multiple theta positions. Accordingly, Hornstein 1996/98 suggests that "behave"-type verbs in English are like HIT, (17a) is to be analyzed as involving an object to subject A-chain with two theta roles. (He suggests extending the analysis to the local lexical anaphor cases in (17b,c), although he does not explain in the paper how he intends to deal with the difficulties this move would create, some of which he notes, --cf. FN13 above).

- (17) a. John behaved t
- b. John behaved himself
- c. John hit himself

Furthermore he proposes that in (18) the ECM verb "expect" has the same properties, --it also has the option of not assigning an accusative even though it has a thematic subject position. But firstly this would incorrectly predict that "*John expects" is grammatical, --in fact ECM verbs like HIT do not seem to exist. This fact seems difficult to explain in H's approach but follows immediately if no verb like HIT exists (cf. Chomsky 1982,

Brody 1993) and objectless "behave" is intransitive, interpreted roughly as "self-behave". Secondly, even assuming the existence of HIT type ECM verbs, the question would still remain why we never find this verb where the subject-complement chain forced by the lack of accusative would involve only a single theta position? Why is a structure like (19a), where "expect" is a hypothetical verb with no accusative and a subject theta position, always ungrammatical? Why can't a structure like (19a) ever express what (19b) does, why can't (a verb that fits into) this construction exist?

(18) John expects to like Mary

(19) a. *John expects t to seem/to be obvious that S

b. John expects it to seem/to be obvious that S

Stipulative answers are easy to find (although H appears to ignore the problem and does not provide any), but these compare rather unfavorably with the solution where the impossibility of (19a) is simply another consequence of the prohibition of chains with a non-root theta position (Cf. Chomsky 1982 or the Generalized Projection Principle (GPP) in Brody 1995. The GPP is ultimately a consequence of more general quasi-semantic considerations, cf. Brody 1998).
FN18

In sum, H appears to provide no real grounds for questioning what one might regard as one of the major insights of the period between standard theory and the principles and parameters framework: the isolation of the thematic and locality properties of movement/chain relations and their differentiation from those of the construal rules. In spite of serious doubts, for the sake of argument let us again proceed nevertheless on the assumption that H's theory of control is tenable.

5. The argument from control

H's argument against chains based on his theory of control exploits the well known contrast between (20) and (21):

(20) Someone seems (to John) to be reviewing every report

(21) Someone hoped to review every report

The scope ambiguity discussed earlier in connection with raising structures like (20) disappears in the control structure in (21). In (21) "someone" must have scope over "every report". H attributes the difference between raising and control to the fact that within his system of assumptions all but one copy in the chain must delete by LF. As we have seen in the raising case in (20) he assumes that this may be either a copy of "someone" that is higher than the LF position of "every report" or a copy that is lower than this QP. In contrast in the control structure a

copy of "someone" that is lower than the matrix theta position cannot be the LF copy on the assumption that this copy does not carry the matrix theta role and at LF all theta roles must be expressed. A copy not lower than the matrix theta position must remain, which carries the subject theta roles of both the matrix and the embedded predicates.

There are difficulties with considering this account as an argument against chains. Notice first that this evidence is part of the standard argumentation for the move/chain vs. control distinction, from which the different properties follow. (Control involves no chain, hence no possibility of reconstruction for inverse scope.) Hence H's discussion of the material seems better interpreted as an attempted defence of his alternative approach to control against one more piece of standard counterevidence.

H makes no attempt to argue that the account provided in terms of his approach is superior in any way to the standard account, he only argues that under his assumptions about control involving A-chains, the facts can be made to follow if chains are eliminated. But even if H's dubious theory of control is adopted, there will be no argument here for eliminating chains from the theory. Theta assignment inevitably involves some type of linking, let us take it to involve theta (role/position/assigner)-binding as is often suggested (cf. esp. Williams 1994). (To what extent this is exactly the same kind of binding as for example the binding of an anaphor is an interesting question, but largely irrelevant here.) Given the understanding of the relationship between scope and binding discussed above in section 3, nothing further needs to be said to rule out the impossible reconstructed interpretation in (21): the higher theta position cannot be bound by an element E from a position higher than the scope position of E. The account rests on assumptions about the interpretive component that appear to be independently necessary. The fact that H's chainless grammar also excludes the relevant reading is again not an advantage but a potential liability. Pruning multiple (chain-)copies at LF results in a redundant account of the data, a redundancy that can be avoided if multiple copies are present at LF. Thus again we appear to end up with a potential argument *against*, rather *for* H's chainless grammar.

Recall also, that as noted above, even if H's argument went through for keeping only a single copy of the members of an A-chain at LF, nothing would follow for the status of chains in general in the theory. A'-chains must still be present at or (probably) beyond LF, and all members of these (post-)LF chains must apparently be present at LF (cf the references in FN10 above).

6. Additional problems

There are various additional problems with H's account of the raising control distinction. (As distinct from his argument against chains based on this account, --as we have seen, making his account of this distinction work is necessary but not sufficient to make his argument against chains work.) Consider the crucial assumption, hidden in H's presentation in the relevant section of M&C: thematic information can be inherited from a lower position in a chain but not from a higher one, --see Brody 1995, 1998 for discussion of this assumption and its relation to the GPP. Some principle with this effect in the control case is obviously needed to make H's account tenable. If thematic information could be inherited from a higher chain position, then the lower copy in the assumed control NP-chain could remain at LF in (20), since this copy would "express" both theta roles. This would result in the impossible reconstructed inverse scope reading.

H packages theta role transmission in chains with movement: theta role transmission is then upwards (to c-commanding positions) because movement (which carries the element to which the theta role was assigned) is upwards. Surprisingly in the following section of his paper H argues for a derivational theory that allows both raising and lowering and thus has the consequence that theta roles can also be inherited from a higher chain position by a lower one FN19. He realizes the problem he created in this way at the very end of the article, where he suggests that "the Infl of a control structure [(presumably) the Infl in the embedded clause MB] has a strong d-feature while this feature in raising structures is weak" (p.125). He assumes that the extension requirement which forces bottom-up derivations holds only in overt syntax. Hence control must involve a bottom-up move while raising is in fact either LF raising or LF lowering.

There are numerous questions H's suggestion raises. We would immediately loose the account of the impossibility imaginary verbs like H's "expect" appearing in a structure like (19a), where the prohibition of downward transmission of theta roles seems crucial. The assumption that the extension condition refers only to overt syntax is also highly problematic (Brody 1997a, see also Pesetsky 1998, Chomsky 1998). For some further problems see also FN19. But let me just note here the fact that given these proposals, OC PRO is not just an (intermediate) NP-trace any more that happens to find itself in a chain with two theta roles. It is an NP-trace that for unknown reasons must have been created in a component of syntax that is different from the component where ordinary (intermediate) NP-traces arise. Again ignoring the fact that the evidence for covert categorial movement is debatable at best (Brody 1995, 1997, 1998, in prep), observe that instead of null Case, the distribution of OC PRO is now ensured in terms of the

"strength of the d-feature of Infl". Since neither NP-trace nor OC PRO appears to occur in (real) Case marked positions, both null Case and strength of D-features is primarily needed to differentiate NP-trace from PRO. Without evidence that null Case and strength of d-feature have significantly different empirical consequences, the choice between them remains a terminological matter. We seem to have here nothing but a different name for a stipulation that for some reason has often been mistakenly taken to be a solution.

The situation is in fact even worse. Null Case gives way under H's approach to not one but two distinct stipulations: the licensing requirement for NOC PRO and the strength of the d-feature of Infl condition for OC PRO.

7. Conclusion

In sum there are two main arguments in M&C against chains, one based on the correlation of quantifier scope and binding, the other on the correlation between quantifier scope and thematic properties. Both rest on highly dubious background assumptions. Additionally even granting these assumptions both are flawed in similar ways. In both cases the observed correlation follows from independently necessary interpretive considerations, ensuring the correlation also syntax-internally would needlessly complicate the grammar by creating redundancy in the overall linguistic system. There are numerous additional problems both with H's account of quantifier scope and control and with his arguments against chains, --some of these I remarked on above.

In an approach like H's the syntactic computational system has the ability to create copies and also to delete them. It seems a rather arbitrary choice to attribute this power to the syntactic computation. Chains express a (thematically) special type of identity relation between chain members. Since it seems unpromising to try to reduce all semantic identity relations to movement, the natural place for expressing chain-identity is the (quasi-)semantic component where, unlike in syntax, the concept of identity will be automatically available. If this is correct, then resolving chain-identity within syntax in the special subcase of A-chains is likely to be wrong. Syntactic treatment of chain-relations is redundant given quasi-semantic mechanisms that we expect to be independently needed and have the same effect. FN20 Taking chains to be interpretively constructed was suggested in Starke 1998 and in the framework of perfect syntax in Brody 1996, 1998. In this framework multiple copies in chains do not differ from other multiple occurrences of lexical items (or structures constructed from lexical items) with respect to their origin. In all cases multiple occurrences are due to a (set of) element(s) being selected from the lexicon more than once. (Features of) multiple occurrences can then be

interpreted as being in the identity (or perhaps the slightly weaker "non-distinctness", --cf. Brody 1995, 1998) relation. If the multiple occurrences all involve a single (and most deeply embedded) theta position/role then this identity can qualify as "chain-identity". So this approach dispenses with syntax-internal chains (which would violate inclusiveness) as H justifiably desires: 'chain construction' is distributed between syntax (multiple copies) and interpretation (identity). The distributed theory of chains eliminates also the additional unmotivated arbitrary and redundant syntactic mechanisms of movement and deletion that H assumes. Syntax does not need to create chain-copies, the independently necessary lexical selection mechanism suffices. Chain copies do not need to be linked (and/or deleted) syntactically at LF, since they will be in the independently necessary interpretive identity relation.

Footnotes

* I am grateful to an anonymous reviewer for questions that resulted in clarifications at several points.

FN1 M&C is of course not an isolated case in this regard. Kennedy 1997 for example, to take a paper on a related topic, explicitly states in a footnote that " my specific purpose in this article is to compare the account of ACD proposed in Hornstein 1994 only with an account that relies on QR. I will not discuss alternative analyses..." (p.668), but then concludes the paper by stating that "the facts associated with ACD structures support a theory of the syntactic representation of quantification in which full quantificational DPs --both determiner and restriction--are raised to an adjoined position at LF" (p.686). In other words from arguments that establish that Hornstein's object shift account is inferior to the QR account, he concludes that the QR account is correct. But this conclusion clearly does not follow when there are other existing alternatives in the literature (that he cites but does not discuss).

FN2 Brody 1995, 1996, 1997a, 1998. The two issues in (a) and (b) in the text below are carefully distinguished already in Brody 1995.

FN3 Epstein et al. also do not seem to overtly recognize either that they work out a version of the "radically minimalist theory" in the sense of Brody 1995, -- a theory without representational-derivational duplications. For example in their note 14 on p.160 they compare my earlier work in which spellout applies to the (L)LF level to a theory that they now reject which postulates both derivations and an LF interface which is the input to spellout. The comparison is strange. The existence of (L)LF, hence (L)LF spellout is consistent with the representational view but not with the derivational one (see the text and FN5

below). So my earlier work, where I propose to eliminate derivational representational redundancies, is not compared with the material in Epstein et al. where they attempt to resolve the same issue in a different fashion. Instead it is compared with derivational work now abandoned by them, that inconsistently adopted an assumption that was part of my representational approach and which does not even address the central issue of redundancy. (Incidentally, they also misrepresent here my proposals concerning the distinction between "overt" and "covert" chain. I did not argue for a solution in terms of "high" vs "low" spellout in chains but for an approach whose derivational equivalent is Chomsky's later FF-movement proposal, --cf. Brody 1998 for discussion).

For critical remarks on aspects of Epsteinian purely derivational work see Brody 1997a and 1997b.

FN4 Heycock 1995 is a rare case where it is explicitly argued that both derivational and representational (LF) conditions are needed. For a critical discussion see Brody 1997b. The standard minimalist theory shares with the representational framework the insistence that all conditions hold at the interface, an assumption that is in near contradiction with the adoption of a mixed representational-derivational approach (Brody 1995).

FN5 Brody 1995 contains several specific arguments that H does not consider against a theory that postulates both a syntactic interface (LF) level and derivations. I do not address here the more difficult choice between the fully derivational and the fully representational frameworks. I believe that they are not necessarily notational variants and there are at least some general reasons for preferring a fully representational theory to a fully derivational one. One pertinent consideration is the lack of genuine syntactic feeding-bleeding relations, relations that we would expect to be commonplace if syntax was (non-notationally) derivational (Brody 1998). Another set of relevant issues involve the natural integration of the syntax module with other, apparently representational competence systems (cf.e.g. Jackendoff 1997) and apparently derivational (e.g. parsing) systems of the mind-brain.

FN6 For the discussion of a similar and similarly circular argument in Chomsky 1987 see Brody 1995 section 1.9.

FN7 I noted also that this view "necessitates a global recoverability link between post-LF interpretation and PF since the latter component needs the chain structure information to operate properly (e.g. for "trace-copy" deletion, --Brody 1998, section 3.5.). H does not spell out his assumptions concerning the interaction of his chain-less syntax with the PF component.

FN8 H talks throughout in terms of deleting chain links in spite of the fact that his argument is that chains do not exist. While this seems somewhat careless, it is probably not a serious problem: it seems easy to reformulate chains for his purposes as say a set of derivationally created copies together with their original.

FN9 Brody 1995, Kennedy 1997 and Wilder 1995, 1997 mostly discuss problems in connection with H's treatment of antecedent contained ellipsis structures, but the problems carry over straightforwardly to his analysis of quantifier scope.

FN10 In Brody 1995 it is explicitly argued in detail that A'-chains have copies in all chain member positions at LF. See now Kuno 1998, Safir 1998 who also adopt this assumption.

FN11 Note that this requirement allows binding by an element E from more than one position if none of these positions are higher than the scope of E as in (i) a modified version of an example mentioned by a reviewer:

(i) Every girl seemed to her parents t to adore herself

The claim in Brody 1995 that multiple copies in chains are present at LF with different subparts of the copies interpreted in different positions is of course also not in contradiction with this understanding of the relationship between scope and binding.

FN12 Fox 1997 also (correctly) refers to the requirement that a bound pronoun must be in the scope of the quantifier that binds it, as a "(virtual) tautology". The argument would of course carry over to an account that attempted to derive the facts discussed in this section from properties of derivations. Epstein et al (1998) (who incidentally, like Fox, quote not Hornstein but Lebeaux 1994/1995 for the data) provide an analysis that they consider derivational. This analysis includes a statement that makes the crucial distinction: an anaphor a must be interpreted as "coreferential with a category taking scope over a" (p.72). Given this stipulation, their conclusion that "we presented a derivational analysis of the scope asymmetry" (p.73) seems unwarranted. They in effect presented a stipulation that they made no attempt to explain. This stipulation then entails the data when embedded in a set of assumptions that can be couched either in derivational or in representational terminology.

But the more important point here is different. Suppose that it could be shown that some, as yet undiscovered, inherent properties of derivations result in the observed

data. Given that this would be redundant with the conceptually (near-)necessary interpretive account, such a demonstration would again probably constitute a potential argument *against* and not *for* derivations.

FN13 Hornstein (1996) suggests also that PRO is a strange element in a minimalist setting where traces are copies and therefore the fact that his theory reduces PRO to NP-trace and *pro* is an advantage. It is not so clear however if there is a problem: if *pro*, the empty pronominal exists (which H also assumes), then why not PRO, an empty anaphor (cf. the text below and Brody 1985). Also if it turned out that PRO must be eliminated for some valid reason, there are still in principle various alternative options, --like FF A-chains Brody 1997a (see also Manzini and Roussou 1998 for a different execution that shares with Hornstein's work some of the problems raised here), or anaphoric Agr (Borer 1989).

The published version of Hornstein 1996, Hornstein 1998 appeared only after this paper was completed. The published version also includes the argument in the previous paragraph of this note and points out further problems with treating lexical anaphors in terms of movement. For reasons that are not obvious, H lists these problems as problems for a control module that seeks to specify the potential antecedents of OC PRO. Thus he appears to consider these difficulties as an argument against a theory that assimilates the antecedent choice properties of PRO to those of lexical anaphors. But the problems of reducing anaphora to movement are in fact problems not for the interpretive, construal rule approach but for his own. The generalizations extending over PRO and lexical anaphors need to be captured and indeed he attempts to assimilate not only PRO but at least programmatically also lexical anaphors to movement constructions.

FN14 On problems with the MLC see Brody 1997a.

FN15 Hornstein (1996/98) points out correctly the ad hoc nature of Chomsky and Lasnik's (1993) approach to the distribution of PRO as an element requiring "null Case" which is stipulated to be assigned/checked only in nontensed subject positions. He then goes on to propose his differently phrased but essentially identical "solution" for the distribution of NOC PRO.

FN16 Or essentially the same contexts. H suggests that "John behaved" can contain an OC PRO/NP-trace object, on this see the text below.

FN17 Including Hornstein 1996/98.

FN18 Chomsky has recently suggested a condition that would prevent merging arguments in non-theta positions. This

would rule out (19a). Notice however that postulating such a condition is an extremely dubious move since it effectively reconstructs one half of the concept of D-structure (a move that H also explicitly wishes to avoid). In late principles and parameters theory D-structure consisted of just two assumptions, (a) that all arguments are (merged) in a theta position and (b) that all theta positions are filled with an argument (in minimalist terms: all theta assigners are merged with an argument). But as argued in detail in Brody 1993 and Chomsky 1995, D-structure should not exist. In particular in a theory that postulates a level of LF (this includes both Chomsky's and Hornstein's work and mine) D-structure would only recapitulate interpretive properties that both empirically and conceptually belong to the LF level.

FN19 H is concerned here with the phenomenon that he calls "All For One Principle", according to which if a chain member checks Case, all copies of this element in the chain count as having had the relevant Case checked. H considers this a problem for his proposal to eliminate chains. His reasons are not entirely clear, since the principle can easily be translated into movement terminology. Given standard minimalist theory, for example the following formulation seems adequate for the purpose: a Case C on some element E counts as checked if C or a copy of C (on a copy of E), created by the computational system, is checked. Thus the all for one principle as such creates no problems for a derivational approach.

H however appears to assume that the principle cannot be derivationally reformulated and considers instead two alternative options. The first is taking Case to be interpretable which would neutralize the problem of unchecked Cases on copies since unchecked interpretable features can appear at LF. The second alternative assumes that derivations can include lowering applications of move from chain-top Case-checking positions. H tentatively opts for the latter approach. I believe that this approach raises many more problems than it solves (see the text for some comments on (the representational equivalent of) lowering and Brody 1995, 1998 for more discussion. Conversely I believe that there are strong independent reasons to take Case and indeed all syntactic features to be interpretable, see Brody 1997a. But I will not discuss H's two alternatives here for two reasons. First, as just noted the matter is a red herring in the context of the question of whether chains exist, --the effects of the all for one principle can be ensured both derivationally or representationally independently of the choice between H's alternatives. Additionally, H himself concludes the discussion of these alternatives like this: "It is fair to say that the arguments above are not entirely clear cut. They rely on questionable judgements and several ancillary assumptions of dubious veracity..." (p.122.).

H also comments on Chomsky's (1995) chain uniformity condition, and suggests that this should be restated to refer to copies. This is difficult to evaluate since H does not say what he means by copies. (Is the copy relation determined with reference to the syntactic derivation or the semantic interpretation? Are the pronouns in both "he seems (he) to have left" and in "He said he left" copies? if not why not etc.) Again I don't discuss H's comments on this condition in detail here since they are largely irrelevant to the matter at hand. The chain uniformity condition (if it existed) could also be stated both derivationally and representationally. For a different discussion of this condition see Brody 1998, for a theory that eliminates categorial projection altogether and therefore even the possibility of raising the chain uniformity problem see Brody 1997b.

FN20 Reuland 1997 argues in the Reinhartian spirit that principle B is essentially due to the preference for using an anaphor in the relevant structures. He suggests assimilating the relevant anaphor antecedent relations to movement/chain relations and proposes further that anaphor antecedent relations take precedence because they involve the computational system, C_{HL} , an "automatic" "preencoded operation". Reuland's theory can of course be restated without postulating C_{HL} , we could assume instead that the operation that links chain-members in the interpretive component is "preencoded" and takes precedence.

But as we have seen fully assimilating anaphor-antecedent relations to chains is problematic given the differences in thematic structure. Reuland assumes a local interpretation of the theta criterion. He suggests that in (i) for example the chain [Oscar, sich] can involve two theta positions since each one corresponds to a different predicate. He then assimilates also complex anaphors like English "himself" in a structure like (ii) to this pattern

- (i) Oscar voelde [sich wegglijden]
 Oscar felt himself slip away
- (ii) John likes [him+self]

But then why is (iii), with an A-chain ungrammatical? And if this is for Case theoretical reasons why does no verb "BELIEVE" exist as in (iv) where "BELIEVE" is like "believe" except it does not assign Case, --why does the relevant part of Burzio's generalization hold? (See the discussion of (16)-(19) above).

- (iii) *John believes t to have liked Mary
- (iv) *John BELIEVES t to have liked Mary

References

- BEGHELLI, F. 1995. *The phrase structure of quantifier scope*. Doctoral Dissertation. UCLA.
- BORER, H. 1989 Anaphoric Agr. in O. JAEGGLI and K. SAFIR eds. *The Null Subject Parameter*. Dordrecht: Kluwer.
- BRODY, M. 1985. On the complementary distribution of empty categories. *Linguistic Inquiry* 16:505-546.
- BRODY, M. 1993. Theta theory and arguments. *Linguistic Inquiry* 24:1-23.
- BRODY, M. 1995. *Lexico-Logical Form: A radically minimalist theory*. Cambridge, Mass.: MIT Press.
- BRODY, M. 1996. The minimalist program and a perfect syntax. ms. UCL. [to appear in *Mind and Language*.]
- BRODY, M. 1997a. Towards Perfect Chains. in L. HAEGEMAN ed. *Elements of Grammar, a Handbook of Syntax*, Kluwer. [First in TLP working papers 2.4. University of Budapest, 1995.]
- BRODY, M. 1997b, *Mirror Theory*. ms. UCL.
- BRODY, M. 1998. Projection and phrase structure. *Linguistic Inquiry* 29:367-398, [First in TLP working papers 2.4. University of Budapest, 1995.]
- BRODY, M. in prep. Why covert categorial movement is not necessary to resolve antecedent contained ellipsis.
- CHOMSKY, N. 1982. *Lectures on government and binding*. Dordrecht: Foris.
- CHOMSKY, N. 1987. Comments on reviews by Alexander George and Michael Brody. *Mind and Language* 2:178-197.
- CHOMSKY, N. 1995. *The Minimalist program*. Cambridge, Mass.: MIT Press.
- CHOMSKY, N. 1998. Minimalist inquiries. ms. MIT.
- CHOMSKY, N. and H. LASNIK 1993. The theory of principles and parameters. In J. Jacobs, A von Stechow, W sternefeld and T Vennemann eds. *Syntax: An international handbook of contemporary research*. Berlin: de Gruyter.
- EPSTEIN, S. ET AL 1998. *A Derivational Approach to Syntactic Relations*. OUP.
- FOX, D. 1997. Reconstruction, binding theory and the interpretation of chains. ms. MIT.
- HEYCOCK, C. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547-570.
- HORNSTEIN, N. 1996. Movement and control. ms. University of Maryland.
- HORNSTEIN, N. 1998. Movement and chains. *Syntax*. 1:99-127.
- HORNSTEIN, N. 1998. Movement and control. *Linguistic Inquiry* 30:69-96.
- JACKENDOFF, R. 1997. *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.
- KENNEDY, C. 1997. Antecedent contained deletion and the syntax of quantification. *Linguistic Inquiry* 28:662-688.
- KOSTER, J. 1981. Binding and control. ms. University of Tilburg.
- KUNO, S. 1998. Binding theory in the minimalist program. ms. Harvard University.

- LEBEAUX, D. 1994/1995. Where does the binding theory apply. In *University of Maryland Working Papers in Linguistics* 3. University of Maryland.
- MANZINI, R. 1983. On control and control theory. *Linguistic Inquiry* 14.3.
- MANZINI, R. and A. ROUSSOU 1997. A minimalist theory of movement and control. ms. Florence/UCL and Bangor.
- PESETSKY, D. 1998. Phrasal movement and its kin. ms. MIT.
- REULAND, E. 1997. Primitives of binding. UIL OTS Working Paper.
- SAFIR, K. 1998. Vehicle change and reconstruction in A'-chains. ms. Rutgers University.
- SPORTICHE, D. 1985. *Structural invariance and symmetry in syntax*. Doctoral Dissertation. MIT.
- STARKE, M. 1998. Locality. ms. University of Geneva.
- WILLIAMS, E. 1994. *Thematic Structure in Syntax*. Cambridge, Mass.:MIT Press.
- WILDER, C. 1995. Antecedent-containment and ellipsis. ms. Max-Planck-Gesellschaft, Berlin.
- WILDER, C. 1997. Phrasal movement in LF: de re readings, VP-ellipsis and binding. ms. Max-Planck-Gesellschaft, Berlin.