Chapter one

On the Status of Representations and Derivations

Michael Brody

1 Representations and derivations – the status of the mixed theory

1.1 Restrictiveness and duplication

As set out in earlier work, elegant syntax (ES) differs from the minimalist framework in several important respects.[1] I shall elaborate here some remarks made earlier on those features of this approach that relate to the so-called representational–derivational issue. I argued that since chain and move express the same type of relation, a theory that contains both concepts is redundant, and, therefore, at least in the setting of ES, wrong.[2] As has been also noted repeatedly, the issue is more general: there is a redundancy built into the architecture of theories that assume that both representations and derivations play a role in the competence theory of narrow syntax.[3]

Let us note first that a general conceptual argument from simplicity in favour of a pure (representational or derivational) theory against a mixed one is weak or nonexistent. This is because it is in principle possible that derivational and representational principles are both necessary in syntax and that they hold in different domains, and/or are distinguished also by other independently needed principles and properties – i.e. cluster in a modular fashion. Such clustering of properties with chains in one module and move in another does not seem to obtain in narrow syntax (the lexicon to LF-interface mapping), but this does not seem to be a necessary state of affairs, but rather an empirical fact about language. It may be that in wider domains, like the theory of mind, for example, both derivational and representational components will be necessary. The important point here is that the argument from redundancy against mixed theories of narrow syntax, to be discussed below, is not purely conceptual but is ultimately empirically based.[4]

Consider then representations and derivations in narrow syntax.[5] In principle there are two possibilities here (ignoring now logically possible but apparently nonexistent mixed situations that involve both possibilities in a modular fashion). Either derivational and representational accounts of the lexicon to LF relation are (a) empirically distinguishable, or (b) they are not. Although it may have been sometimes argued that both of these situations obtain, it is obvious that these two states of affairs are incompatible.

I return to (b) in the next subsection. Let us consider first the situation where we take (I think correctly, see section 1.2 below) the representational and the derivational theory to be empirically distinguishable. When the argument against the mixed theory was initially put forward there were essentially no attempts to construct analyses that relied on the existence of both derivations and representations. Given the lack of such arguments one obviously opts for either a fully derivational or a fully representational theory on general grounds of restrictiveness.

While there may now be some contributions in the literature that postulate both representations/chains and derivations/move and exploit one or another assumed (typically stipulated) difference between these pairs, as far as I am aware there are essentially no strong arguments for postulating both concept-pairs as part of narrow syntax.[6] Nobody has attempted to show that the results achieved in the less restrictive framework, that apparently involves systematic duplications (a property that is strange even in a minimalist setting, let alone ES), cannot be restated in a nonmixed system that avoids redundancy and lack of restrictiveness. There are also no attempts to argue that the assumed advantages outweigh the considerable burden of weakening the grammar. It is clear that even if focused arguments existed for the claim that both derivations and representations must exist side by side within the competence system of the language faculty and largely duplicate each other, these would have to be treated with extreme caution, since they would amount to a proposal to adopt a less restrictive grammar.[7] Everything else being equal, there are clearly more analytical possibilities in a theory that has both representations and derivations with differing properties than in a system that only has one of these concepts.

I shall refer to these considerations as the argument from restrictiveness against the mixed theory of narrow syntax. Let me summarize this argument. Suppose that representations/chains on the one hand and derivations/move on the other have different properties. (This seems to be the case.) Then it's an empirical question which notion(-sets) are the right ones. Having both would weaken the theory in the sense of increasing the analytic options available (see note 7), hence very strong arguments would be needed to maintain that both concept-sets are part of the competence theory of syntax. No strong argument appears to exist. Further, in addition to the problem of the unmotivated lack of restrictiveness, we would also have the problem of the unmotivated systematic (representational–derivational) duplications.

1.2 *Principles of I-language*

Suppose then, as is sometimes suggested, that arguments for a mixed theory are lacking because the issue they would address is effectively meaningless.

Representations and derivations are just notational variants, they are simply different approaches to expressing the same notions and the same generalizations. Suppose that there were no empirical differences to distinguish the derivational and the representational views.

But on such an assumption a mixed theory like standard minimalism only becomes more strange. Putting aside the uninteresting case where notational variance means synonymy – two names for the same concept – let us look at the situation where we take derivation/move and representation/chain to be two different aspects of, or two different ways of looking at, the same phenomena. Consider first a situation in physics that might be somewhat similar. The famous double slit experiment of quantum theory can be interpreted either in terms of probability waves or in terms of a particle being able to traverse multiple trajectories before hitting a target.[8] The two interpretations apparently do not result in distinguishable empirical predictions. (This is the case now, and may or may not remain so in the future). Assuming this fact, it would be a strange theory that postulates both multiple trajectories and probability waves, say mapping one into the other. It would be much like a theory whose ontology is committed to two entities, the evening star and the morning star, in the context of the assumption that ultimately they are empirically indistinguishable. The standard minimalist framework mapping derivations into representations appears to be equally curious – especially so when viewed from the perspective of ES, which rules out in principle the option of attributing syntactic redundancy to the effect of selection or to evolutionary accidents.[9]

To repeat, on the assumption that representation/chain and derivation/move are just notational variants (i.e. no empirical evidence distinguishes them), they are either just different names for the same notions or perhaps different but (at least currently) not empirically distinguishable notions. So one could suggest that the choice between them is not real, that one of them is just a way of looking at the other. In such a situation it may be reasonable to look for some deeper notions that subsume the two competing ones. But it seems mistaken to conclude from the assumption that, say, move captures the properties of chain, that both chain and move are part of the grammar. If we talk about (some module of) I-language, and say that y is part of it, hence a real object and furthermore that x is just an aspect of y, a way of looking at or treating y, this does not then seem to entail postulating x as a distinct element of the mind. Further evidence would be necessary for that, but by hypothesis this would be unavailable if the two notions cannot be distinguished empirically. I shall refer to this consideration below as the argument from I-language ontology. So this argument is meant to establish that the mixed theory cannot be defended even on the (empirically dubious) grounds of derivations and representations being notational variants. But the main argument against mixed theories remains the consideration based on restrictiveness and duplication: there is relatively little evidence for distinguishing derivations and representations, and not surprisingly there is essentially no serious evidence for adopting both.

2 Representations or derivations

2.1 Derivational theories and weak representationality

Suppose then that the rejection of mixed derivational–representational theories, mainly on grounds of empirically and conceptually unmotivated lack of restrictiveness, is correct. Next comes the related but distinct and secondary issue of whether syntactic theory is better thought of as purely derivational (PDT) or purely representational (PRT). By a PRT of narrow syntax (or LF) I understand a system that generates the interface level in the mathematical sense of generation. This consists of a set of constraints or principles that determine well-formedness. We could assume that, essentially as in the standard minimalist framework, these constraints can only include bare output conditions and a definition of possible LF structures (that bare output conditions constrain further).

Such a structural definition could, for example, run along the following lines: a representation (tree) consists exclusively of nodes ($n_1, \ldots, n_n$) and the immediate domination relation such that each node (except the root) must be immediately dominated by some other unique node (ensuring the connectedness of the tree). (This sketchy definition is not intended as an actual proposal, but simply as an indication of the form a representational definition of LF could take. Under various theories, various elaborations will be necessary and various aspects of this definition may follow from elsewhere. (See, for example, Brody (1997b, 2000a) and Abels (2000) for more extensive discussions of two versions of a particular approach along the above lines.)

We could proceed further by defining constituents recursively in terms of immediate domination. Alternatively, as suggested in note 5, we could take domination as the primitive notion and assume that x immediately dominates y iff x dominates y and there is no z such that x dominates z and z dominates y. A constituent will then be a subtree that contains every node a given node n reflexively dominates. (Note, however, that in the context of the theory in the works just cited, it may not be necessary to define constituents for the purposes of narrow syntax at all. If linking of chain members, binding, etc. are taken to be matters of interpretation – a natural and empirically motivated assumption – then constituents might be visible/created only in the interpretive components.)

I assume further, though not crucially, that the question of how to assemble as opposed to constrain (or generate, in the sense of "specify") the representation falls outside of the competence theory of grammar and is part of how the linguistic competence system is used – most plausibly it corresponds to the theory of parsing and sentence production.

A PDT is an ordered series of operations with input and output, where the input may only consist of terminals and the outputs of some other operations.[10] The following three-way distinction will be useful: (i) a derivational theory is nonrepresentational if the derivational operations create opaque objects whose internal elements and composition is not accessible to any further rule or operation; (ii) a derivational theory is weakly representational if derivational stages are transparent in the sense that material already assembled can be accessed by later principles (i.e. the derivational stages are representations); finally (iii) a derivational theory is strongly representational if it is weakly representational and there are constraints on the representations (weak sense) generated.

It is clear that derivational theories must be at least weakly representational. Take an object z, the result of merging x and y. At some later step move can only apply to y if z is a transparent rather than an opaque object since otherwise y would not be accessible or even visible for this operation. Notice that even if move is reduced to merge and an interpretive linking operation (as in the theory of distributed chains, Brody 1998b, 1999a), the same conclusion would still hold: the interpretive link between x and y could not be established if z was opaque. The derivational theory therefore is at the same time a (weakly) representational theory with multiple (weakly) representational stages instead of just one at the interface.[11]

So there can be no derivational theories that are fully nonrepresentational. The derivational theory will always be a mixed one to some extent. It would also seem to be almost necessarily a multirepresentational theory. One might think that this sort of weak representationality does not matter, since the spirit of the theory remains derivational. I can see two problems with this sort skepticism about the argument. First, weakly representational derivational theories are clearly mixed theories and the I-language ontology argument above in section 1.2 applies to them just as much as to any other empirically unmotivated mixed theory. The fact that all derivational theories must be mixed then appears to already provide a good reason for rejecting derivational theories of all kinds.

Secondly, consider the suggestion that weak representationality does not matter, because the crucial difference between the representational and derivational view is that the latter is not strongly representational, there are no representational constraints on the structures that the derivation assembles, hence these structures (although weakly representational) are still not levels of representation in some more important sense. But given that derivational theories are at least weakly representational, a derivational operation must have an input and an output both of which are at least weakly representational. Hence a derivational operation involves, or is equivalent to, a set of representation-pairs: a set of possible input–output pairs (in fact, representation n-tuples in the general case, since in principle there can be more than one input or output). The operation can thus equivalently be thought of as a member of a (partially?) ordered set of multirepresentational constraints. We can understand a weakly representational derivational theory as having an ordered set of such multirepresentational constraints.

It should be clear then, that the distinction between weakly and strongly representational derivational theory, despite appearances, does not really have to do with the derivational–representational distinction. What the distinction between weakly and strongly derivational theory really concerns is the question of whether there are constraints that are additional to those captured by the postulated derivational steps (whether we view these latter as representational or derivational constraints) and bare output conditions. Currently the restrictive working hypothesis of many linguists working in this domain is that there are no such additional constraints. But the answer to this question may be either negative or positive, both on the representational and on the derivational view.

Consider current "derivational" theory with the operation merge, some applications of which are a suboperation of move. The input of merge is any two well-formed representation WR and WR′ (built from terminals and subtrees by merge) and the output WR″ is WR augmented by WR′ in a way that merge specifies. Thus in general merge is a tri-representational constraint. Where merge is a subpart of move, it applies to an element WR′ of a tree WR and augments WR with a proper subpart of WR, WR′. What merge specifies is that WR and WR′ will be sisters in WR″ and furthermore WR″ inherits its label from WR or WR′ (in the case of move, always from WR for reasons independent of merge).

Thinking of the derivational approach as a multilevel representational theory, we see that this constraint is essentially equivalent to the requirement that at every level L a (sub)tree ST″ is well formed iff (a) it immediately dominates two well-formed subtrees ST and ST′ [whose correspondents are present at L–1] each composed of terminals and other subtrees (in the case of move, ST′ is properly dominated by ST [at L–1]); and (b) ST″ carries the label of ST or ST′ (always ST in the case of move). Given this background, the question of whether there are any syntactic constraints that are additional to the structural definition of possible LF representations (whether in terms of merge and move or their representational equivalents, or in terms of different notions) has little to do with representationality or derivationality of the system. We expect, mostly on the grounds of the (at least partial) empirical explanatory success of theories heading in this direction, and on the basis of considerations of theoretical elegance, that there aren't any. But if there are, they can be stated either in derivational or representational terms. Note in particular that a constraint on a single representation can always be phrased as a bi- or tri-representational constraint with no restriction on the input(s), or with placing parts of the condition on the input as in fact happens in the case of merge and move. (Note in this connection also, that the square bracketed level statements above seem unnecessary, as expected from the viewpoint of the single level representational theory.)

Thus the essence of representationality appears to be weak representationality. Strong representationality does not seem to add a property that genuinely distinguishes between derivational and representational approaches. The distinction between weak and strong representationality in fact pertains only to the irrelevant, though otherwise important, issue of whether the elegant theory that assumes only a (hopefully trivial) structural statement and bare output conditions can be maintained. If it is true that the core concept of representationality is weak representationality, then of course, having shown that derivational theories must be weakly representational, the question of whether we should adopt derivational theories of narrow syntax again reduces to whether we should adopt mixed theories in this domain. As we have seen in the previous section, this we should probably not do.

2.2 Restrictiveness again

So current (apparently pure) derivational theory is equivalent to a restricted multirepresentational theory that has only such conditions on representations that can be stated as conditions that hold on two adjacent levels. As we have seen, it is in fact not clear that this really is a restriction with respect to a multilevel representational theory, since a single-level condition could be equivalent to a bi-level condition where the input may be any structure. The real difference between derivational and representational approaches is different. The representational theory is a single-level theory: all representations/derivations except the "final" representation, LF, are eliminated – so conditions can only hold here. This is clearly one obvious way to constrain the multirepresentational theory: assume the existence of only a single representation, the one corresponding to the final output of the derivational system. Henceforth I refer by representational theory unambiguously to the single-level representational approach. To emphasize the representational properties of derivational theories I shall use the term "multirepresentational."

The derivational approach constrains the multirepresentational theory differently, in a way that does not resolve the problems of the mixed theory. The derivational representational duplication now translates as the duplication between the final representation and the relevant aspects of all representations generated that carry the same information. Sisterhood and projection is duplicated at multiple levels by the effects of merge and chain by those of move.[12] The derivational theory ignores the problems of duplication and lack of restrictiveness, but suggests a different restriction. In this approach constraints like merge and move (which, as we have seen, are effectively equivalent to multilevel representational constraints) are individuated and are crucially required to operate in a sequential manner.

Perhaps there are aspects in which the sequential derivational theory is more restrictive than the unilevel representational theory in an empirically motivated way. As far as I know, this has never been argued and there is little to indicate that this might be the case. On the other hand, there is immediate evidence of this type for the unilevel representational theory. It is more restrictive than existing derivational approaches since it disallows bleeding relations, which do not seem to occur in narrow syntax. In particular the effects of the cycle follow automatically from the representational nature of the theory. But the cycle (unlike an inviolable extension condition that current derivational approaches reject) is just an additional stipulation under the derivational system.

If there really were derivational components in syntax we would expect bleeding relations to occur with some regularity, and if syntax was fully derivational, as is frequently suggested, bleeding relations should be commonplace. Derivational systems are eminently suitable to express the situation where one operation bleeds another rule or constraint. Consider cases where lack of bleeding of some constraint C can be detected as the fact that ungrammatical sentences (ruled out by C on one derivation) do not become grammatical on a different one where the context for C would not arise. Take for example the well-known fact that the wh-island or the subject island constraint cannot be bled by a derivation that involves movement before the relevant configuration is created, as e.g. in (1) and (2):

(1) a. What did you wonder Mary bought (what) when ==>
    b. *What did you wonder when Mary bought (what) (when)

(2) a. Who was bought [a picture of (who)] ==>
    b. *Who was [a picture of (who)] bought ([a picture of (who)])

To deal with the descriptive problems, the usual restrictive assumption added to derivational framework has for a long time been the idea of the cycle in various incarnations. The derivations in (1) and (2) do not obey the cycle: cyclic application of all rules and constraints removes this empirical problem together with other similar ones. The solution is less than satisfactory if proposed as an explanation of the lack of bleeding in derivational frameworks. While the cycle may be a simple and attractive construct, nevertheless it is an additional stipulation that (as first observed in a somewhat different framework by Freidin 1978) appears to be unnecessary on the representational view. Until the cycle is independently motivated, the representational theory has the advantage of being more restrictive than the derivational theory in an empirically motivated way. The derivational approach can achieve the same degree of restrictiveness and empirical adequacy only by invoking an additional descriptive stipulation.[13]

Epstein et al. (1998) proposed that the cycle is a consequence of an appropriately defined notion of c-command, together with a PF ordering requirement. The intuitive idea is that a relation based on c-command must be defined between all terminals of the tree – to make possible the exhaustive ordering of the terminals at PF by Kayne's (1994) Linear Correspondence Axiom (LCA) – and c-command is defined in terms of merge (as holding in a particular way between the merged categories, see section 3.1 below). In a countercyclic operation applying to A, A will not therefore have this c-command-based relation established with higher nodes in the tree. Such operations will thus be impossible.

As noted in Brody (1997a), the account based on PF ordering does not rule out, however, all violations of the cycle. Since traces are invisible at PF and therefore do not need to be ordered, countercyclic movement or merger of A followed by cyclic raising of A is still incorrectly allowed. The approach allows also lowering rules if followed by cyclic raising – highlighting another aspect in which the derivational theory is less restrictive than the representational.

In the representational theory chains are neutral with respect to lowering, raising, and round trip (lowering followed by raising into the same position) derivations. These distinctions by now rather clearly seem empirically unmotivated. Although they could be stipulatively grafted onto a representational theory, the basic concepts of this approach, unlike that of the derivational theory, do not naturally provide for these unnecessary distinctions.

The reliance of Epstein et al.'s explanation of the cycle on the LCA is also questionable. The status of the LCA as an external stipulation on an otherwise overgenerating derivational system raises the same issues as the cycle. Surely we should prefer a theory in which the basic building blocks of hierarchical relations simply did not permit the types of structures that in standard frameworks we need the LCA to rule out. Brody (1997b, 2000a) presents a theory with this property, and recently Kayne also discussed the problematic nature of the externally stipulated LCA and argued for a partly similar approach (Kayne 2000).

In addition to these considerations there is an even more crucial problem with deriving the cycle from (an appropriately constructed) c-command: the notion of c-command has a complexity presented by its asymmetrical nature, so it is probably even more problematic than the cycle that it is called for to explain. See Brody (1997b, 2000a) and below, especially note 18.

3 C-command

3.1 Derivational definition

Epstein pointed out in an influential paper (1995) – see also Epstein et al. (1998) – that in the cyclic derivational framework of the minimalist approach, c-command can be defined as in (3):[14]

(3) x c-commands all and only the terms of a category y with which x was paired by merge or by move in the course of the derivation

He compared (3) with Reinhart's representational definition, which I restate in (4):

(4) x c-commands y iff
 a. the first branching node dominating x dominates y; and
 b. x does not dominate y; and
 c. x does not equal y.

Epstein claimed that the derivational definition in (3) answers certain questions concerning properties of the relation that are "unanswerable given the representational definition of c-command" (p. 19 in the MS). Before looking at this claim, notice that (4) can be made more easily comparable to (3) if it is restated as (5) in a form parallel to (3):[15]

(5) x c-commands all and only the terms of its sister

He suggests that (3) explains that (a) x appears to c-command whatever the *first* (and not fifth, nth etc.) branching node dominating x dominates, since "this is the projected node created by pairing of x and y . . ." Furthermore x does not c-command (b) the first branching node dominating x, (c) nodes dominated by x and (d) x itself – in each case the reason being that x was not paired with the category in question by merge or move during the derivation.

But the derivational definition in (3) appears to give us neither more nor less insight into why these properties characterize c-command than the representational definition in (5). We can say without any loss (or gain) in understanding that x appears to c-command whatever the *first* (and not fifth, nth etc.) branching node dominating x dominates, since "this is the node that dominates (all and only) the terms of x and those of its sister y." Similarly instead of saying that x does not c-command itself, the nodes dominating it and the nodes it dominates because x was not paired with these, we can say without any apparent loss of insight that x does not c-command these because these are not its sisters (since all and only sisters are paired).

The insight these alternative definitions give is limited. In the case of the representational version we might ask why sisterhood is relevant. Additionally we don't know why x c-commands the terms of its sister rather than, conversely, x's terms c-command x's sister. Or why does not x only c-command its sister or why all x's terms don't c-command all the terms of x's sister. The same questions arise for the derivational statement: here we may ask why derivational pairing is relevant – notice that pairing is not identical to c-command but only enters its definition. The other questions just asked in connection with the representational version also arise here: why a paired category c-commands the terms of its pair rather than conversely, or symmetrically (i.e. why x does not only c-command its pair or why terms of x do not c-command terms of x's pair).

It is important to see that if the derivational account of c-command is to be taken as evidence in a strict sense for a derivational view, then the question of why derivational pairing is relevant to c-command cannot be answered by saying that derivational pairing is the only mechanism that establishes (purely) syntactic relations. The existence of derivations cannot be presupposed in an argument that wishes to establish precisely that. So this way of answering would beg the question: does the pairing relation have to be derivational?

Epstein suggests also that the fact that c-command makes reference to branching can be explained in a framework where "Structure Building (Merge and Move) consists of Pairing, hence it invariably generates binary *branching*." Again, this point is in fact neutral with respect to the issue of whether syntax should be constructed as a representational or derivational system. The assumption that pairing by merge and move is always binary is an additional assumption – there is nothing in the notion of concatenation that would force this operation to always be binary. The syntactic concatenation could in principle operate on any number of elements. This would allow also the unary operation alongside the binary, ternary etc. options. But just as the concatenation operation can be restricted to be binary, correspondingly, the branching of trees can be restricted to the binary option, ensuring the same result in representational terms: the elimination of nonbranching nodes (along with the elimination of other n-ary branching for n not = 2).

Additionally, Epstein argues that the representational definition of c-command is inconsistent with the independently motivated hypothesis of the invisibility of intermediate projections.[16] He considers the example of the category that is the sister to a VP-internal VP-spec subject – I will refer to this as V′. If V′ is invisible for the computation of c-command relations then the elements contained in it (the verb and its complement) will c-command the subject and also the categories the subject contains. This is undesirable. On the other hand, Epstein suggests that the situation is different if c-command relations are determined derivationally by (3). Then even under the assumption that the intermediate projection V′ can ultimately neither c-command nor be c-commanded (i.e. if its c-command relations established by (3) are eventually eliminated), the subject will still asymmetrically c-command the verb and its complement as required by Kayne's LCA. Notice that if V′ is fully visible to c-command relations then the subject and V′ will symmetrically c-command each other, creating problems for the antisymmetry hypothesis.

Given the assumption of antisymmetry, it seems necessary to assume that V′ or more generally intermediate projections (or lower adjunction segments) are visible for the computation of c-command relation, but cannot themselves c-command or be commanded. There is nothing, however, in this state of affairs that would be "incompatible" with a representational view.

Consider instead the weaker claim that this behavior of intermediate projections can be naturally attributed to the assumption that at the point in the derivation where a category becomes an intermediate projection (i.e. once it projects further), its c-command relations become invisible (it neither c-commands nor can it be c-commanded) but nevertheless during the earlier stage of the derivation it has already participated in determining c-command by other nodes (it counts for the calculation of c-command by these).

The problem with this line of argument is that the interpretation of "becoming invisible" is not antecedently given, it is not any more natural to understand invisibility as entailing only the loss of ability to c-command and be c-commanded than to understand it as the loss of any c-command related role (including the role in the calculation of c-command relations between other nodes). Thus again the advantage of the derivational approach is only apparent. The statement that intermediate nodes participate in the calculation of c-command relations by other nodes but they do not participate in c-command relations themselves is not improved upon by saying that this latter property arises at a point in a derivation where the nodes become intermediate nodes/project further.[17]

3.2 *Derivational explanation?*

The various definitions of c-command – as Epstein notes in connection with his cyclic derivational version – do not explain why c-command exists, they just state its properties. The question remains why certain – or perhaps all – syntactic relations are restricted by c-command. Why cannot categories establish the relation with any other category in the tree? And if the set categories with which a given element can establish a (relevant) relation is to be restricted, why is it restricted precisely in the way the definition of c-command states, rather than in one of the infinitely many other imaginable ways?

Epstein offered an explanation within the cyclic derivational framework he adopted. This is based on two assumptions, which he refers to as (a) the first law/the unconnected tree law and (b) the law of pre-existence. The unconnected tree law states that a syntactic relation can only hold between elements that are members of the same tree and excludes relations between elements of unconnected trees. "Derivationally construed," as in (6), it disallows relations between elements that at any point in the derivation were members of different unconnected subtrees.

(6) [Epstein's (27)] T_1 can enter into c-command (perhaps more generally, *syntactic*) relation with T_2 only if there exists *no derivational point* at which:
  (i) T_1 is a term of K_1 (not = T_1) and
  (ii) T_2 is a term of K_2 (not = T_2) and
  (iii) there is no K_3 such that K_1 and K_2 are terms of K_3.

Given the cycle, the condition in (6) prevents sideways c-command between two elements x and y. In all such configurations cyclicity allows only derivations in which two unconnected subtrees have been formed at some stage that properly contain x and y respectively.[18]

Notice that "derivationally construed" actually adds another assumption to the unconnected tree law, namely that lack of (c-command) relation at any derivational level freezes and cannot be overridden later:

(7) If there was no (c-command) relation at any given point in the derivation between terms x, y (both already merged into some subtree) there cannot be a relation later.

Statement (7) still allows x to have a relation to (c-command) y where y c-commands x, since in such a configuration no unconnected subtrees that contain both x and y have been formed.[19] Epstein excludes this configuration by his principle of derivational "pre-existence" (8), which disallows x c-commanding y on the grounds that y was not present when x was introduced.

(8) x cannot bear a relation to y when y is nonexistent.

Given the assumption that the lack of a relation at a derivational point cannot be remedied at a later stage, i.e. (7), (8) entails the exclusion of what we might call upward or reverse c-command.

On closer examination, the condition in (6) does not actually explain, however, the impossibility of sideways relations. The intuitive content of the condition is that two categories unconnected at any point in the derivation cannot enter into a (c-command) relation. But in fact all merged/moved categories were unconnected before merger, all can still c-command the appropriate nodes. In order to allow categories to c-command at all, it is necessary to add the stipulation in (6i,ii) that "K not = T," i.e. that the top node of an unconnected tree does not count as an unconnected element. But this means that "K not = T" in fact just encodes the difference between c-command and lack of it. In other words, it encodes the difference between the c-command domain of x being the local dominating node of x (K = T) and a nonlocal dominating node of x (K not = T) not constituting such a domain. So instead of an explanation we have only another way of stating the c-command configuration.

Epstein comments on the "K not = T" restriction by noting about the top nodes (to be related by merge/move) of the unconnected trees, i.e. about K_1 K_2, that "each equals a root node, neither has undergone Merge or Move, hence each is (like a lexical entry) not 'yet' a participant in syntactic relations."[20, 21] In other words, the two instances of the "K not = T" stipulation in (6i) and (6ii) can be exchanged for an additional fourth subclause as in (6′):[22]

(6′) T_1 can enter into c-command (perhaps more generally, *syntactic*) relation with T_2 only if there exists *no derivational point* at which:
  (i) T_1 is a term of K_1 and
  (ii) T_2 is a term of K_2 and
  (iii) there is no K_3 such that K_1 and K_2 are terms of K_3 and
  (iv) merge/move has already applied to T_1 and T_2.

The intuition (6′) appears to express is that two terms that are integrated into some subtree by merge/move cannot form a relation if at any point in the derivation *after they have been so integrated* they are unconnected, i.e. they are members of distinct subtrees. With the addition of (6′iv), (6′) states that if applying merge/move to two elements x, y does not result in a subtree of which both are terms, then x does not c-command y, that is, either x or y must have been merged with some tree that included the other. (Invert the conditional: if x c-commands y then merge/move applying to x and y must have resulted in a subtree that includes both.)

The explanation of the definition in (3) involves then breaking it up into two parts: x c-commands y if neither of the following two situations obtains: (a) there is no derivational point at which x, y have been integrated into unconnected structures and (b) there is no derivational point at which x is present/integrated but y is not. Clearly, we can bring the two parts of the account (6′) and (8) together again, since in both cases what is crucial is that there is a derivational point at which a (sub)tree exists into which x is integrated but y is not. But whether or not we make this improvement, the account provides no evidence for derivations, since it can again be easily restated in representational terms.

Instead of referring to a derivational point at which there is a (sub)tree into which x is integrated but y is not, we can say that x cannot c-command y if in the single syntactic representation there is a subtree which properly contains (i.e. contains but is not equal to) x but not y. Instead of rationalizing that all derivational stages must be checked for x–y connection and, where no c-command holds, there was one at which x was in a (sub)tree that did not contain y, we can presume that all subtrees in the representation must be checked for x–y connection and we have no c-command where we find one in which they are unconnected. (Note also that the representational version is in fact preferable, if the bottom-to-top derivation and the cycle have no independent motivation (see Brody 1997a and the text above), since the derivational account needs to assume these. Furthermore, the easy translatability of the account into noncyclic representational terms provides some additional evidence against these constructs.) But until we have an explanation of why a relation cannot be established at a later derivational stage that connects the relevant subtrees that were unconnected earlier (or, in representational terms, why the connection must hold in all subtrees), it will remain at the very least debatable for both the representational and the derivational versions to what extent the account explains and not just rephrases Reinhart's definition.

In contrast to the clear exposition of the nonexplanatory nature of the definition in (3) in Epstein's paper, this definition is itself sometimes taken to provide a sufficient explanation of c-command. Thus, for example, Groat (1995) states that while c-command is arbitrary as a representational definition, "it is explainable as a property of the derivation." Take a configuration like (9), where Z c-commands A B C, A,B does not c-command Z:

(9) Z+$[_C$ A B]

According to Groat this "follows straightforwardly if the relations formed by [merge] are in fact properties of the operation. Z is merged, hence Z is in relation with $[_C$ A B]. A B were not merged with Z, hence they are not in relation with Z."

But again, we need to decide if merge/move applies to trees or to categories. If the former, then in (9) Z merges with C, hence Z does not c-command A and B. If the latter, then say $[_Z$ D E] merges with $[_C$ A B], and D and E are incorrectly predicted to c-command A and B. In neither case do we get the desired result. We can, of course, stipulate c-command again, by saying, for example, that it is always a category that merges with a tree.

3.3 *Domination*

The core of the c-command problem is the arbitrary asymmetric conjunction in its definition: x c-commands y iff the following two conditions of somewhat different nature obtain: (a) there is a z that immediately dominates x, and (b) z dominates y. It is crucial, but unexplained, that the two subclauses make use of different notions of domination. None of the attempted explanations, some of which I reviewed in the previous section, are able to explain this asymmetry.[23] Consider a different approach (Brody 1997b, 1999a). Instead of trying to explain the strange properties of c-command, let us assume that no such strange properties exist because, despite appearances, no notion of c-command is part of syntax or more generally of the grammar. Cases where c-command appears to be useful are cases of accidental interplay between two (in principle unrelated) notions, one of which is domination.

How about the other notion? In standard frameworks this must sometimes be the specifier–head relation and sometimes the head–complement relation. I shall only consider here the specifier–head relation because in the ES representation provided by mirror theory (Brody 1997b, 1999b, 2000a) the head–complement relation reduces to domination. (In mirror theory heads and the associated "projected" phrases are not distinguished in the syntactic representation, hence c-command by a head H reduces to domination by H.)

Consider a typical condition that refers to c-command like, for example, principle C of the binding theory. Suppose that spec–head agreement has the effect of the head inheriting/sharing the referential/thematic features of its specifier. Then instead of requiring that an R-expression not be c-commanded by a coreferential category we can prohibit the configuration where the R-expression is dominated by an Agr node carrying the same reference.

Similarly, the requirement that each chain member c-command the next can be straightforwardly restated in terms of domination. Again I ignore head chains here, since in mirror theory their members will be in a strict domination relation with each other. Consider chains that are constructed on potentially larger structures (phrasal chains in standard terms). Assume that the members of these chains always occupy spec positions. Let us think of the heads associated via spec–head relations with the spec positions occupied by the chain members as themselves constituting a chain, call it r(estricted)-chain. (Note that an r-chain is a chain whose members are heads, but it has nothing to do with the head chains expressing the head–chain/movement relation. In mirror theory head chains in this latter sense reduce to morphology and do not exist narrow syntax internally.) It is the domination relation that must hold then between members of r-chains. Additionally and independently we require that r-chain members must have identical or nondistinct specifiers. This is natural since the heads participating in the chain are by virtue of that fact at least in some respects identical, so they will naturally require identical, or at least nondistinct (see Brody 1997b, 1998b), specs.[24]

4 Summary

The representational framework seems more restricted than the derivational one in that there are many derivations for a single representation, but not conversely. I argued on the empirical grounds of bleeding relations that some of the derivations need to be eliminated to reach descriptive adequacy. Additional asumptions are necessary in the derivational framework, which are not entailed by the hypothesis that syntax is derivational. As we have seen, the corresponding problems do not arise in the representational framework where the correct consequences follow directly from the representational nature of the system. Additionally I provided arguments against mixed derivational–representational theories of the kind where derivations and representations essentially duplicate each other's work. I showed that no observationally adequate pure derivational theory can exist, that on closer examination derivational theories are mixed theories with derivational–representational duplications, hence arguments against mixed theories hold also against apparently pure derivational theories.

In the second part of the paper I argued that the derivational explanation of the asymmetry in c-command is unsuccessful, hence no argument for a derivational approach can be based on it. I suggested that the explanation may be so difficult to find because this complex notion is epiphenomenal only and does not exist within the grammar. I put forward an alternative approach, developed in more detail elsewhere,[25] according to which syntactic principles thought to refer to c-command refer to simple domination instead. Other independently necessary principles of spec–head agreement ensure that reference to domination, instead of c-command, is sufficient.

Note finally that derivational explanations tend to assume that merge is a conceptually necessary part of the competence theory of syntax, and argue that given its inevitability, it should be taken as a basic concept that makes various other assumptions unnecessary. All that seems really unavoidable, however, is that consequence of merge, that lexical items must be related in some way, so that they form a (connected) syntactic representation. Other properties of merge, like its derivational (sequential) nature, the fact that it relates sisters directly, and also its projection and labeling properties, seem to be stipulative and arbitrary. Although many linguists are used to the notion of merge as an unanalyzed primitive, at least sequentiality, projection and labeling are curious additions to some basic relation R between lexical items, and together they appear to form a strange and arbitrary package. Furthermore these properties of R seem to be unnecessary and eliminable. They are in fact eliminated in mirror theory, the approach to LF hierarchical structure that ES subsumes.

Acknowledgments

I am grateful to Noam Chomsky for a series of detailed and helpful e-mail exchanges relating to some of the points discussed. I'd like to thank also Klaus Abels, David Pesetsky, Sam Epstein, Daniel Seely, and Michal Starke for comments that I hope resulted in clarifications. As usual, no agreement on anything or transfer of responsibility is implied. Earlier prepublication versions of this manuscript appeared in University College London Working Papers (UCLWP), 12 and in Israel Association of Theoretical Linguistics (IATL), 8 (Tel Aviv University).

Notes

1 Brody (1997a,b, 1998a,b). Although in these works I referred to the framework of elegant syntax as perfect syntax, the operative sense of perfection was invariably that of theoretical elegance. Hence the change of terminology.

2 Brody (1995, 1997a,b, 1998b). As has been noted before, in Brody (1995), the argument against mixed theories (which include both the pure derivational and

the pure representational alternatives) and the argument for the representation option (as opposed to the derivational one) are clearly distinguished. A certain amount of confusion has been generated in subsequent literature by not always keeping these two points distinct.

3 See esp. Brody (1995, 2000a). For more recent discussions of the "architectural" duplication see Epstein et al. (1998) and Starke (2000). See also Hornstein (1999, 2000) and Brody (2000b, 2000c), among other related matters, for a discussion of a somewhat curious position that retains the architectural redundancy but wishes to eliminate the chain–move duplication. Hornstein (1999) also argues for eliminating chains rather than move, but on the basis of flawed arguments (Brody 2000b). He attempts unsuccessfully to defend one of these in Hornstein 1999. See Brody (2000c).

4 A possible argument against the approach I'm taking here might be that it focuses narrowly on LF. When we take the full theory of expressions generated by the grammar this seems to include a derivational component: a mapping from narrow syntax to PF. Therefore, one could argue, the overall theory of grammar would be simpler if the theory of the lexicon–LF relation was also derivational. But we seem to know too little about Spell Out for this argument to carry much force. First, it is not clear that the Spell Out component is indeed derivational (i.e. sequential) and not just a one-step mapping. Secondly, even if they are derivational, we do not know if the principles of Spell Out are different or similar to those of narrow syntax. The general idea of syntax being a generative and Spell Out an interpretive component would not make it unexpected that Spell Out principles have a different cluster of properties from the principles of narrow syntax. If this is the case that would make at least the intuitive simple version of the simplicity argument inapplicable. More complex versions – like, for example, that the same principles apply differently in the two domains (the differences being due to the different properties of the elements to which they apply) – may still hold. But again we seem to know too little about Spell Out to make any such point with more confidence than its opposite.

5 It is sometimes suggested that a representational approach simply translates a derivational approach and with the cost of involving richer set theoretical assumptions. It is not clear how the richness of the set theory involved is relevant to what is an empirical issue: which system is instantiated in the mind of the speaker. This is an empirical matter to which both empirical considerations and conceptual considerations of sharpening the concepts involved may be relevant, but the mathematical properties of the object postulated to exist will have to be whatever empirical research (with concepts adequate for the task) determines them to be. Once the set theoretical point is eliminated from the picture, as I think it should be, it is clear that a priori we do not know if the derivational theory is a (perhaps misleading) translation of the representational approach or conversely.

Recursivity of LF structures is also sometimes cited as entailing derivationality, at least in spirit. Note, however, that we can define LF without employing rules that reapply to their output. LF could be structurally characterized, for example, by some relation R (e.g. immediate domination) that all nodes have to enter (with special clauses for roots and terminals). For more discussion see e.g. Brody (2000a). It is expected that in the framework of ES no narrow syntactic principles will refer to the notion of constituent. That is, it would not be necessary to define constituents for narrow syntax, but only for the interpretive modules. But in any case the notion of constituent would not necessarily have to be defined recursively in terms of immediate domination. If the primitive notion is domination, then a node together with every node it dominates is a constituent – and immediate domination can also be defined nonrecursively in terms of domination without intervention.

Remnant movement is also sometimes taken to provide evidence for derivations. In a derivational theory the context for a simple statement of the c-command relation between the element moved out from a constituent to be moved itself (the remnant) is destroyed by the later step of moving the remnant. The problem, however, does not seem to arise in a representational theory with copies, given the assumption that c-command of traces inside chain members, like anaphoric connections in general, have to hold only with respect of a single chain member. This corresponds to the fact that in a derivational approach c-command has to hold only at one step of the derivation. That is, with X forming a chain with a copy (indicated here by "t") inside the bracketed chain-forming remnant ("X extracted from the remnant"), X needs to c-command only one of the relevant copies ("t's") inside the members of the chain formed from the two bracketed remnants in (i):

(i) [. . . t . . .] . . . . . X [. . . t . . .]

Another question raised by David Pesetsky (personal communication) about representationality and remnant movement concerns principles of Spell Out, in particular, what ensures that in (i) the copy in the position indicated by the first "t" is silent. Concentrating on two-member "overt" chains without resumption, this appears to follow from the cyclic derivational theory. Move involves copy and delete. If X remerges and deletes before the remnant does (as it must, given the cycle), the original position of X in the copied remnant will be empty.

Note first that cyclicity of syntax does not in fact follow; the same result is translatable to a grammar with a representational syntax, if Spell Out is cyclic. Spell Out might have the two rules of identity check of chain members (corresponding to the identity requirement of copy) and +silent marking of the lower member (corresponding to part of delete). If these apply cyclically then identity check for the remnant will cover also the +silent mark on the lower copy of X in the lower remnant in (i), hence the higher remnant will have silent X. But even cyclicity of Spell Out can also apparently quite straightforwardly be avoided if we assume that the identity requirement on chain members covers also the +silent marking quite generally, but only up to recoverability. Thus in a remnant chain, if one member properly contains a +silent element, so must the other. But in an ordinary nonremnant chain, where the lower member is silent marked, the requirement will not entail that the higher member is also silent, since this would violate recoverability.

6 Heycock (1995) was one early case where it was explicitly argued that both derivational and LF conditions are necessary. For critical discussion see Brody (1997b) and Fox (1999). To take a somewhat random choice from relatively recent work that assumes and attempts to argue for a mixed theory, take first Nunes (2000), who argues that Move should be decomposed into copy (C), merge (M), form chain (FC) and chain reduction (CR). In fact M is not different from the usual merge operation that puts together phrase structures, CR is a Spell Out issue and C need not be separate from selecting from the lexicon the same thing twice. (The difference between the relation linking the two pronouns in "He said he left" and "He was seen (he)" does not have to do with different lexical access, as is sometimes suggested.) It is plausible to attribute that to FC having applied (or being able to apply legitimately) to the two pronouns in the second but not in the first structure. So only FC remains. In other words, it is not clear that this approach really needs to be different from a representational account. It looks different, of course, for Nunes C applies as part of a derivation. That a derivation exists and that C is part of it are thus additional assumptions.

In support of the assumption of keeping copy and (re)selection from the lexicon distinct Nunes refers to Chomsky's (1995) argument from expletive construction where greater cost is assigned to move than to merge to rule out (i).

(i) *There seems a man to have left

The strength of this argument is questionable since (i) may be excluded by independent reasons: for example, that no lexical element, expletive or not, is ever permitted in the infinitival subject position that follows *seem*-type predicates. On accounts that exploit this fact, assigning different cost to different derivations would become irrelevant.

In support of the assumption that derivation exists, Nunes cites also the following contrast:

(ii) "*Which book did you review this paper without reading?"
(iii) "Which book did you review without reading?"

This is supposed to motivate derivations on the grounds that *which book* moved sideward from an island in (iii) before it became an island and then to the front while a similar non-island-violating derivation is not possible in (ii). But there are no reasons why a largely similar alternative account could not be given in a representational vocabulary. In (ii) *which book* is separated from its trace (theta position) by an island. In (iii) it is not, since there is a trace in object position of the matrix clause. The trace in the island causes no violation if the wh-phrase needs only a single thematic trace to be subjacent to it (see e.g. Richards' (1997) principle of minimal compliance, a major and very interesting generalization of a proposal in Brody (1995), or this latter work for a somewhat different approach). All this seems straightforward, and makes no direct reference to parasitic chains. It is not clear why the derivational approach would be better. In fact for there to be an argument for derivations here, it would be necessary to argue that something along these representational lines can not be right, otherwise Nunes' account (and the derivational equipment it is supposed to motivate) is redundant and therefore undesirable.

Lechner (2000) proposes an interesting analysis of NP-comparatives, where an empty operator raises to an intermediate spec-C position and the AP moves into the matrix:

(iv) Mary met [young-er men]$_i$ [$_{CP}$Op$_j$ than Peter met [$_{DegP}$ [$_{AP}$ young men] Deg t$_j$]]

He suggests an argument for a mixed theory based on the following observation: "empty operators in spec-CP of the than-XP [do] not interfere with AP-movement" (p. 16). He observes that the two APs should not form a chain for thematic/semantic reasons, hence he suggests that these APs are linked by a move operation that applied countercyclically to avoid the island effect induced by the empty operator. Note that countercyclic operations seem to be (a) quite problematic (see section 2.2 below) and (b) they also seem to be beside the point if the relevant locality constraints (like on Lechner's assumptions of the thematic requirements) apply only to chains. Furthermore, no crossing problem will arise if the matrix AP and the empty operator are coindexed and the operator in turn is related to the whole degree phrase in the lower clause – as in other similar constructions analyzed in terms of empty operator movement since the late 1970s. Lechner provides arguments from principle C, etc., that the structure does not involve pure deletion only but movement/chain, but his evidence does not seem to distinguish between linking the AP to its matrix clause correspondent or linking it only to the operator at the edge of the embedded clause.

Pesetsky and Torrego (2000) provide an interesting and intricate new analysis of the *that*-t effect and various related matters. They argue for what they call "relativized extreme functionalism," which appears to be an approach essentially identical to Brody's (1997a) bare checking theory. (I think the colorful name they give is misleading: the issue involved in eliminating features that are in principle uninterpretable is one of restrictiveness and has little to do with functionalism.)

In bare checking theory all features must be interpreted in principle, but in a given sentence some occurrences of features may be in positions where their usual interpretation cannot be assigned to them, where interpreting them would not make sense. In such cases occurrences of features of type t (say *wh* for example) in position(s) where they cannot be interpreted will have to merge (presumably via the chain and the spec–head relation) with another feature of type t that is in a position where interpreting it would make semantic sense. Pesetsky and Torrego's approach is not completely identical to bare checking theory because they wish to retain the otherwise apparently dispensable operation of feature deletion (as it follows feature checking) in order to integrate into their system the anti-*that*-t effects in sentences with topicalization like:

(v) Mary said *(that) John she liked

However it is not clear if such sentences should or can be integrated with other data they analyze. Anti-*that* trace effects constitute a much less clear class of facts than *that*-t effects. Maybe a pause in cases like (v), where the matrix verb does not select for *that*, suffices, suggesting perhaps an approach in terms of parsing. Pesetsky and Torrego attempt to extend their theory to cover such facts, but at the cost of a set of otherwise unnecessary and ad hoc assumptions that in turn question the claim that these facts have genuinely been "integrated." It is necessary to reengineer their notion of locality into a less appealing form specifically to cover this case; it is necessary to retain the otherwise unnecessary operation of feature deletion; and it is even necessary to adopt a gamma-marking type mechanism, essentially identical to that of Chomsky (1999), that distinguishes deletion of a feature from the feature being marked for deletion – the latter carried part way through the derivation.

It seems fair to say that even if we assume that the anti-*that*-t effects must be treated syntax internally, Pesetsky and Torrego have not successfully integrated these into their theory. Assuming that anti-*that*-t effects need to be treated differently, all dubious theoretical adjustments and innovations just mentioned can be dispensed with. The argument for derivations that they consider to have provided then disappears, together with the curious gamma-marking type distinction between marking for deletion at one derivational stage and deleting at a later one. (Gamma marking for deletion as in Chomsky (1999) and Pesetsky and Torrego (2000) is clearly undesirable, and it is also dispensable in general since the deletion operation itself is in fact unnecessary – see e.g. Brody (1997a) on this latter point.)

7 To make the point of restrictiveness more concrete, recall for example that (as noted in Brody 1997a), Chomsky (1995) proposes a representational definition in addition to the derivational system of interface assembly (in effect an additional definition) of what counts as a well-formed syntactic object (cf. also Brody 1998a for some discussion). Or take the additional distinction he makes between deletion

(interface invisibility only) and erasure (essentially invisibility also for Move), where erasure occurs only if this would not violate the representational duplicate definition of well-formed syntactic object. Such duplications that exploit the derivational–representational duplication and distinctions that in turn might build on these additional duplications should probably have no place in a restrictive system of syntax and are indeed excluded in principle by avoiding the less restrictive mixed theory that makes them possible in the first place.

8 Remotely – and at least here irrelevantly – resembling syntactic chains.

9 See Brody (1995, 1997a,b, 1998b, 1999a, 2000b) and Epstein et al. (1998) for more discussion of the redundancy issue and related matters.

10 Actual PDTs and PRTs may have other restrictions, relating, for example, to the number of branches of nodes, etc.

11 If chain members are linked interpretively and at a single interface level, and furthermore the status of z can switch from opaque during the derivation to transparent at LF, then the theory may not be multirepresentational, but would still be mixed. The same conclusion seems to hold also for the various older and more novel multiply dominated single-element theories of chains, since the multiple positions of the relevant category need (also) to be interpretively linked. (Incidentally, this fact might render syntactic multiple domination unnecessary.)

12 In fact I argued that neither categorial projection (Brody 1997b, 2000a), nor the chain relation (Brody 1998b, 1999a) should exist narrow syntax internally, but I put these matters aside here.

13 Note that the examples in the text are not simply cases analyzed representationally that are translatable derivationally without any gain or loss in understanding – something that often seems to be the case with putative arguments for derivations. The examples here illustrate the point that there are several derivations for a single representation, some of which need to be stipulatively exluded by some principle that is not entailed by the derivational nature of the grammar. So it does not matter, for example, if in (1b) *when* in the lower spec-C is in the (intermediate) trace position of *what* or that there are two positions available here, one for each wh-phrase. The pure derivational theory that contains no traces/copies (if it did, it would encode earlier stages of the derivation into later representations) will not exlude (1b) without some auxiliary assumptions that prohibit the countercyclic derivation.

Similarly in (2) it is not relevant that the subject island constraint apparently holds of subjects only. This is not a stipulation that is additional to what would be necessary to exlude the structure in a derivational framework. Derivationally, the assumption translates as the constraint holding only for extraction from subjects. This much is necessary so that the structure be excluded on the cyclic derivation, but does not suffice to rule out by itself the countercyclic derivation. On the representational approach, the representational statement of the subject island does not need to be similarly supplemented by (some equivalent of) the cycle.

Consider a different line of attack. On the representational approach we need to ensure that the trace/copy inside the subject is part of the chain that includes *who* in spec-C and the trace/copy inside the object. But again this is not an extra statement that would correspond to the stipulation of the cycle on the derivational view. If one A-position copy of *who* would be a trace and the other would not be, then the two copies of the subject-to-object chain of *pictures of who* would not satisfy the identity requirement on chain members that corresponds to the identity requirement of move, which is "copy (involving identity) and delete" on the derivational view. But properties of move in the derivational theory do not ensure the ungrammaticality of the countercyclic derivation, while given a representational approach, the corresponding properties of chain do.

14 Sections 3.1 and 3.2 correspond, with minor changes, to sections of Brody (1997b). For the purposes at hand "term" in (3) can be taken as a synonym of "constituent." Citations of and references to Epstein's work in these sections relate to Epstein (1995).

15 Or, if binary branching was not assumed then:

(5′) x c-commands all and only the terms of its sisters

Note that sisterhood is taken not to be reflexive in (5)/(5′).

16 In Brody (1998a), I argued that the best hypothesis to explain the invisibility of intermediate projections (for chain theory) is that they do not exist. See also note 12 for references to a later and stronger hypothesis ("telescope") that subsumes this one.

17 The problem of intermediate projections does not arise in the framework of mirror theory referred to in note 12, where no categorial projection exists.

18 Note that presupposing the cycle in the explanation of c-command and c-command in the explanation of the cycle (see section 2.2 above), as in Epstein et al. (1998), makes the explanation of these notions circular in addition to the other problems discussed in the text.

19 More precisely, no two unconnected subtrees have been formed that respectively properly include x and y.

20 More precisely, K_1 and K_2 have not yet undergone merge or the merge part of move.

21 Notice that "syntactic relation" here must mean: not yet part of the tree, and not as before, c-command.

22 Again, read "merge part of move" for "move" in (6′).

23 It is often suggested that c-command follows from the way semantics works but proponents of this view typically do not raise the question of why the semantics they assume has to work in the way that the strange asymmetry of the notion of c-command/scope comes into existence, why this relation must be what it is. So in effect such accounts often restate c-command in semantics but do not attempt to explain its surprising property. In fact as far as I am aware, all attempted explanations in syntax or semantics so far simply define c-command differently and stipulate the asymmetry differently rather than explain it.

24 In Brody (1999a), some empirical advantages of this view are sketched. Additionally, the substitution of domination for c-command may solve the antisymmetry problem of the well-motivated instances of c-command from the right (see Brody 1997b, 2000a; Brody and Szabolcsi 2000). The latter work discusses also the extension of this view, that semantic scope is similarly a matter of domination.

25 See Brody (1999a).

References

Abels, Klaus. 2000. "Move?" MS, University of Connecticut.

Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory*. Cambridge, MA: MIT Press.

Brody, Michael. 1997a. "Perfect chains." In Liliane Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer, pp. 139–67.

Brody, Michael. 1997b. "Mirror theory." MS, University College London. (Still earlier short version in *UCL Working Papers in Linguistics*, 9, University College London.)
Brody, Michael. 1998a. "Projection and phrase structure." *Linguistic Inquiry* 29, 367–98.
Brody, Michael. 1998b. "The minimalist program and a perfect syntax." *Mind and Language* 13, 205–14.
Brody, Michael. 1999a. "Chains in perfect syntax." Paper presented at the GLOW workshop on "Technical aspects of movement", Potsdam, April 1999.
Brody, Michael. 1999b. "Word order, restructuring and mirror theory." In Peter Svenonious (ed.) *Proceedings of the Tromso VO–OV Workshop*. 27–43. Amsterdam: John Benjamins.
Brody, Michael 2000a. "Mirror theory: syntactic representation in perfect syntax." *Linguistic Inquiry* 31.1.
Brody, Michael. 2000b. "Relating syntactic elements." *Syntax* 2, 210–26.
Brody, Michael. 2000c. "One more time." To appear in *Syntax* 4, 126–38.
Brody, Michael and Anna Szabolcsi. 2000. "Overt scope: a case study in Hungarian." MS, New York University and University College London.
Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.
Chomsky, Noam. 1999. "Derivation by phase." MS, Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. "Minimalist inquiries: The framework." In Roger Martin et al., eds., *Step-by-step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.
Epstein, Samuel. 1995. "Un-principled syntax and the derivation of syntactic relations." MS, Harvard University. Published version in *Working Minimalism*. 1999. Cambridge, MA: MIT Press.
Epstein, Samuel, et al. 1998. *A Derivational Approach to Syntactic Relations*. Oxford: OUP.
Fiengo, Robert and Robert May. 1994. *Indices and Identity*. Cambridge, MA: MIT Press.
Fox, Daniel. 1999. "Reconstruction, binding theory and the interpretation of chains." *Linguistic Inquiry* 30, 157–96.
Freidin, Robert. 1978. "Cyclicity and the theory of grammar." *Linguistic Inquiry* 9, 519–49.
Groat, Eric. 1995. GLOW abstract.
Heycock, Caroline. 1995. "Asymmetries in reconstruction." *Linguistic Inquiry* 26, 547–70.
Hornstein, Norbert. 1999. "Movement and chains." *Syntax* 1, 99–127.
Hornstein, Norbert. 2000. "On A-chains: A reply to Brody." *Syntax* 3, 129–43.
Kayne, Richard. 1994. *Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
Kayne, Richard. 2000. "Recent thoughts on antisymmetry." Paper presented at the Cortona Workshop on Antisymmetry, May 2000.
Kuno, Susumo. 1998. "Binding theory in the minimalist program." MS, Harvard University.
Lechner, Winfried. 2000. Class handouts, Blago Summer School, July–August 2000.
Nunes, Jairo. 2000. "Sideward movement." MS, University of Maryland.
Pesetsky, David and Esther Torrego. 2000. "T-to-C movement: Causes and consequences." In M. Kenstowicz, ed., *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press.
Reuland, Eric. 1997. "Primitives of binding." Utrecht Institute of Linguistics, working paper.
Richards, Norwin. 1997. "What moves where in which language." Doctoral dissertation. Cambridge MA: MIT.
Safir, Ken. 1998. "Vehicle change and reconstruction in A′-chains." MS, Rutgers University, NJ.
Starke, Michal. 2000. GLOW abstract.