Syntax and Symmetry

Michael Brody
UCL, LI-HAS

1. Mirror Theory (MT) methodology

A plane, like the surface of an infinitely large lake has infinitely many symmetries. Rotation with any center and any angle, reflection with any mirror line, any translation results in the same plane. The set of symmetries of concentric circles, like circular waves created by a stone thrown in the lake, is a proper subset of the symmetries of the quiescent infinite lake-surface. Here rotation is symmetric only if its center is the stone's entry point, reflection only with a mirror line that traverses this entry point. No translation or other rotation or reflection is symmetric. Many symmetries of the plane are broken. We instinctively look for an explanation not of the symmetries but of symmetry breaking when we explain the concentric waves with reference to a dropped stone or a jumping fish. (The example comes from Stuart 1995.)

In general it is the departures from symmetry rather than the symmetries that are in need of explanation. Hence (validly) eliminating asymmetry from the theory is ipso facto making the theory more explanatory. This was the thinking behind earlier work in MT (Brody 1997, 2000) which went some way towards reintroducing symmetry into morpho-syntactic structures. However, if such an approach to syntax is taken seriously to its logical conclusion, it is in apparent opposition to all minimalist type asymmetrical structure building (including now also MT), which invariably involves a(nti)symmetrical concepts like labelled merge, spec-head, head-complement, probe-goal and c-command.

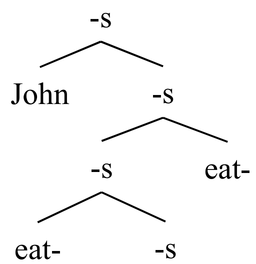
2. MT content

MT reduced the bare phrase structure (Chomsky 1994) tree in (1), where I indicated both the word (X^0)-internal and the word-external structure, to the simpler (2), which involves only a morphological (*eat*-, -*s*) and a (symmetric) syntactic (*John*, -*s*) spec-head relation. The syntactic head-complement relation between -*s* and *eat*- was taken to be a (symmetrically) directionally mirrored morphological spec-head relation.

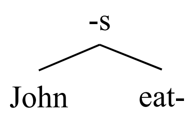
Given that word-internal spec-head relations were thus expressed in the form of syntactic complementation relation, the syntactic difference between heads and phrases (minimal and maximal projections, - whether absolutely or contextually defined) was eliminatable: instead of the redundant (circled) projection line in (3) a single representant of the morpheme -*s* serving both as head and as phrase suffices in (2). MT appeared to eliminate the need for phrasal projection. (Note that bare phrase structure trees eliminated only the absolute distinction between a head and phrase, but not the need for the projection line.)

STIL – Studies in Linguistics.6:167-177
Special Issue on "Directionality of Phrase Structure Building" ISSN: 2281-3128

(1)



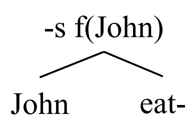
(2)



3. Problems with MT

Let us first observe that the spec-head relation (or spec-x relation to be slightly more neutral, since the head in MT corresponds to both the head and the phrase of the minimalist approach), the basic primitive relation of MT, while construable in various ways, is inherently necessarily not symmetric, – a crucial problem for our present more radical approach. Secondly let us note a dissimilarity: the morphological spec-x relation is a pure configurational relation of concatenation, while the syntactic spec-x relation appears to involve feature sharing. A more precise representation of (2) is (2'), where some feature *f(John)* of the syntactic spec (*John*) shows up on the head *-s*.

(2')



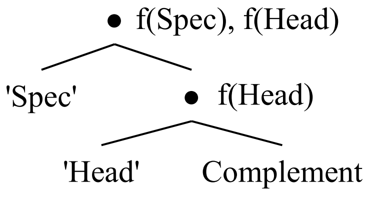
This observation by itself is just that, a fact. It leads however to a problem with two interrelated aspects both for MT and for the approach that I am currently pursuing. Spec-x agreement introduces a redundancy, it causes a feature to be repeated on another node. With the elimination of the projection line MT had some success in eliminating such redundancies, so the reoccurence of a similar redundancy elsewhere is unwelcome news. In fact the problem seems worse. Spec-x feature sharing does not appear to be a concept genuinely different from phrasal (feature-)projection: in both cases a (set of) feature(s) of a node shows up on another (immediately dominating) node. There is little evidence to show that standardly assumed differences between the two relations are more real than apparent. If so, then the elimination of projection in MT is at best partial: the feature duplication relation underlying this notion shows up elsewhere.

To repeat, from our current perspective we have at least three problems for MT:
(a) spec-of-x is itself an asymmetric notion
(b) spec-x agreement still introduces redundancy
(c) spec-x feature sharing does not appear to be a concept genuinely different from (feature-)projection, and if so, then the elimination of projection is incomplete

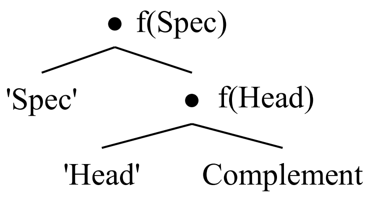
4. Starke (2001)

Michal Starke proposes the "doubly filled nothing" restriction (DFN), according to this an XP projection contains either a spec and a complement, or a head and a complement but never both a spec and a head. His proposal can perhaps be phrased more perspicuously by saying that the (X'-projectional) feature(s) of a head H only show up on the node immediately dominating H, they do not percolate further to the node dominating the spec. So according to his hypothesis, instead of the standard tree in (3), in accordance with the DFN we would have the one in (4).

(3)



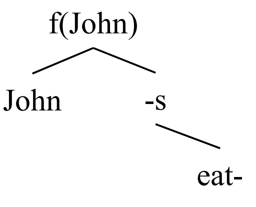
(4)



The diagram in (4) throws into relief the identity of the spec-x feature sharing and that of phrasal feature projection, – a point made in the previous section. Starke considers the spec in (4) as just another head that projects and considers the notion of spec and spec-head relation thereby eliminated from the theory. We can indeed say spec's are heads since they project. But with equal justification we could interpret (4) by considering the head to be a spec since it is in spec-x agreement with the head. The choice appears to be no more than that of perspective and terminology. If we take the latter terminology then it is obvious that Starke's approach is not as different from MT as it might seem at first. Both approaches can be validly thought of as taking the spec-x relation as primitive. (The two theories interpret the x-complement relation differently. MT takes x-complement to be an alternative (morphological) type of spec-x relation while in Starke's framework the head-complement relation appears to be implicitely composed of two different primitives: a spec-x (projection) and an x-complement relation.)

Starke's theory adds also the DFN restriction, which also in MT terms entails that in a spec-x configuration x can only contain a subset of the features of the spec and no other features. So adopting the DFN in MT would result in expanding (2)/(2') as (5):

(5)



More importantly for our current discussion, Starke's approach does not fare better than standard MT with respect to the problems listed in (a),(b),(c) in section 3. above. I adopt here Starke's terminology of head-complement relations instead of the spec-x based wording to make this clearer. First, instead of the asymmetry of the spec-x relation, in Starke's framework we have the asymmetry of the head-complement relation. Secondly, just like spec-x agreement introduces redundant features into the representation, so does the X'-projection of features of the head in Starke's approach (the exact same phenomenon by a different name, if we are correct). And thirdly, if spec-x feature sharing is not distinct from (X'-feature-)projection, then Starke's elimination of the 'spec of' relation is incomplete: in the guise of projection/label it remains as a subpart of the head-complement relation. To repeat, corresponding exactly to the MT problems in (a), (b) and (c) in section 3, Starke's theory raises the following questions:

(a') head-complement an asymmetric concept
(b') projection (of features of the head) still introduces redundancy
(c') if spec-x feature sharing is the same as (X'-feature-)projection, then elimination of the 'spec of' relation is incomplete (in the guise of projection/label it remains as a subpart of the head-complement relation)

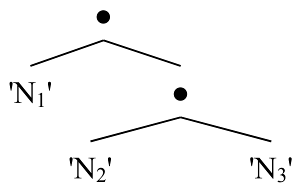
5. Symmetric sisterhood/Merge

It appears then, that our problems are caused by the mechanism that copies a set of feature(s) of a node to an immediately dominating node, sometimes called spec-head agreement, sometimes X'-theoretic head- feature-projection. Labelling is another common way to refer to the same concept. The logical and natural move is then to eliminate labeling, adopting the proposal of Collins (2002). Suppose we adopt symmetric sisterhood/Merge, the minimal assumption, and the obvious one given our general methodology. Clearly, an unordered set is a simpler, and therefore more preferable concept than an ordered one.

This makes all approaches which use the spec-head and/or the head-complement relation in syntactic stucture building impossible, – almost all current approaches, including of course the standard minimalist system, Starke's approach and MT. (Note that the lack of a syntactic head-complement relation suggests strongly that neither s- nor c-selection is syntactic, these presumably belong to the semantic and morphological component, respectively.)

It is not a new assumption that the order of functional heads (complements) is interpretively constrained. If true, then the interpretive constraints that entail this ordering may make syntactic labelling (alias spec-x projection/agreement, alias X'-featural projection/agreement) redundant in the general case. This is particularly easy to see in Starke's framework where all non-terminal nodes immediately dominate a head and a complement. In a configuration like (6) for example, either node N_1 or N_2 projects/labels the immediately dominating node. Suppose N_1 is or contains (in a way accessible for projection) $head_1$, N_2 $head_2$, and N_3 $head_3$. Suppose further that the (interpretively determined) order of these heads is: $head_1 > head_2 > head_3$. It follows, that in (5) $head_2$ must be higher than $head_3$, hence $head_2$ and not $head_3$ projects/labels the immediately dominating node. Thus there is no need to indicate by a label that $head_2$ is head of ($head_2$, $head_3$) and $head_3$ is complement of $head_2$.

(6)



Given more standard assumptions about X'-theory, certain additional complications are introduced if the features of the head typically project also higher than the immediately dominating node and hence some nodes carry the features of both the spec and the head. Let us reasonably assume that the issues that arise here can be resolved, perhaps via the adoption of Starke's DFN and assume with Collins that there is no need to syntactically label in general.

But there are at least two major new problems that result from this step. First, symmetrical merge creates an apparently unsurmountable obstacle to linearization. All theories of linearization exploit one way or another the asymmetry introduced by the labeling mechanism of merge, either directly or indirectly by making reference to concepts like head, spec, complement or projection levels. Members of unordered sets apparently cannot be ordered in a linguistically relevant fashion. Secondly, given labelling/projection/spec–x agreement, all constraints that involve c-command can refer to the domination relation instead; as argued first in the context of MT, Brody 2000, – now see also Chomsky 2005. It seems obviously necessary to have a linearization algorithm for syntactic structures and it would clearly be desirable to retain the result that c-command reduces to simple domination.

6. Interpretive labeling – semantic order

Suppose that labelling for c-command is interpretive, and takes place in the semantic component. Then the same must be true more generally of c-command/domination constraints. Schlenker (2004) convincingly argues that principle C should be derived from a Gricean maxim of minimization whose effect is to eliminate what he analyses as pragmatically redundant restrictors, providing independent evidence that c-command asymmetries may indeed be interpretive in nature.

Schlenker proposes that "As a sentence is processed, top-down ..." "Each time a pronoun or an R- expression which denotes *d* is processed in a context *c*, its sister is evaluated with respect to $c \wedge d$, which is the context *c* to which *d* has been added. In other words processing an R-expression has the effect of making it 'super-salient' for the expressions that are contained within its sister. He argues that: "super-salient entities" (whether extra-sententially (linguistically or not) or intra-sententially provided by the above procedure) by the Gricean minimization principle "must be denoted using a pronoun, unless some special pragmatic effect is obtained by using a full description."

An important aspect of Schlenker's approach from our perspective is the fact that the interpretive top-down processing of the sentence results in a linearly ordered "sequence of evaluation", where if x precedes y in a sequence of evaluation then x is super-salient for y. But instead of adding an R-expression to a structure independent sequence of evaluation for elements contained in its sister node, we can add (the relevant features of) the R-expression to the node immediately dominating it and consider the node dominating the R-expression to be a member of the sequence. In other words the sequence of evaluation for a category x is always an (ordered) sequence of interpretively annotated (labelled) nodes starting from the root of the tree and containing all nodes dominating x.

Note that there seems to be no need to restrict super-salience to binding theory: we may assume that other restrictions that in standard terms involved c-command make reference to this notion. In particular we may assume that chains, which I have argued are interpretive constructs (eg. Brody 1998), are similarly restricted: each chain-member must be super-salient for the next. C-command type asymmetries can continue to reduce to domination.

7. Interpretive labeling - linear order

Symmetric syntax is not compatible with mirror theoretical structures where a head dominates its complement. Nothing prevents however the assumption that interpretive labelling, which as we have seen appears to play a crucial semantic role is also made use of in the morpho-phonological component. Let us make the natural assumption that labeling in the morpho-phonological component is by (morpho-) phonological, features: the PF features of a bound morpheme head *x* are inserted in the node immediately dominating *x*. This effectively recreates mirror theoretical representations, where the morphemes of a word are adjacent elements; – each (morpheme of a given word) being immediately dominated by (being the complement of) the linearly following one.

Such representations make a straightforward ordering algorithm possible, whose core principle is optimally simple:

(7) if *x* immediately dominates (ID) *y* then *y* precedes *x*

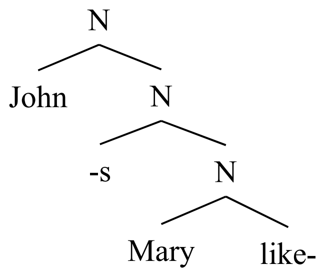
We assume that where ID is a word-external relation (7) orders nodes of the tree, and where ID is a word- internal one (7) orders morphemes. Let us also additionally adopt (8) for expiciteness:

(8) a. ID entails adjacency (of morphemes word-internally, of nodes/domains word-externally)
b. all nodes must be ordered
c. words are spelt out in their designated (highest?) node

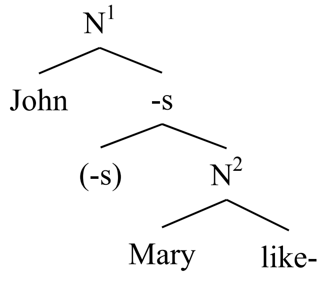
(8a) can be thought of as a version of the no crossing branches requirement, explicit or implicit in all ordering theories. (8b) makes the order linear. (8c) establishes the link between morpho-phonological material and syntactic positions in a way that is apparently much simpler than standard head movement (see Brody 2000 for discussion).

Consider for example (9), a simplified core sentential structure. Given morpho-phonological interpretive labeling the PF features of the bound morpheme -s are copied to the immediately dominating node, resulting in the structure in (9'). As the PF features will be ignored in their source position by spellout (I indicate this by parentheses in the diagram) the relation between (9) and (9') is in fact PF movement. I number the N nodes in (9') for ease of reference only.

(9)



(9')



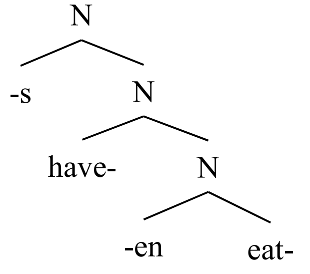
Let us determine first the word-internal order of the morphemes -*s* and *like*-. We might proceed either by ignoring the morphologically and phonologically inert N nodes or by assuming that they form null parts of words. For presentational purposes I assume the latter, as this avoids some, apparently technical, difficulties. In (9') N^1 ID the morpheme -*s*, which ID N^2, which ID the morpheme -*like*, so by (7) the word-internal order is: *like*-, N^2, -*s*, N^1.

Next, we turn to the word-external, 'syntactic' order of words/positions. Here it is necessary to first determine the structural location of the word *likes*, whose morphemes occupy more than one position in the tree. The standard mechanism for establishing the relation between word-internal morphemes and word-external positions is head-movement supplemented by the highly arbitrary strength features. As indicated in (8c), following a MT proposal I tentatively assume that words are invariably taken to be located in the highest position their elements occupy, thus in (9') likes is spelt out in the position of N^1.

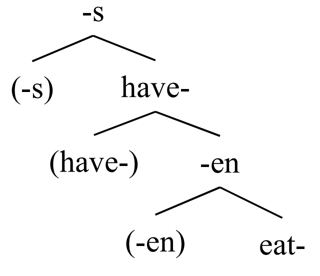
Using now (7) to determine word-external order, we see that *John* precedes N^1, since this node immediately dominates *John*. We have thus established that *likes*, spelt out in N^1, follows the subject. Furthermore, since all nodes must be ordered (8b), and the domain of any daughter of a given node N must be adjacent to N (8a), it follows that the other daughter of N^1, -*s*, must follow N^1. In fact, given the adjacency requirement (8a), everything dominated by -*s* will also necessarily follow N^1, and thus, by transitivity, will follow also *John*. As likes is spelt out in N^1, it precedes *Mary*, since the node *Mary* is dominated by -*s*; and -*s* and everything -*s* dominates follows N^1. Thus we have the correct word-external order: John precedes likes precedes Mary.

Consider the slightly more complex case in (10), (10') after labelling:

(10)



(10')



There is a word-internal relation between -*s* and *have*-, and between *eat*- and -*en*. The word-internal order, given the ID relations is *have*- precedes -*s* and *eat*- precedes -*en*. The former word is spelt out in the position of -*s*, the latter in that of -*en*. The ID relation between *have*- and -*en* is word-external, hence *eaten* spelt out in -*en* will precede the position of the *have*- node. But *eaten* in fact follows *has*. We can ensure this result if the *have*- node and its domain, hence also *eaten*, which is included in this domain, follows -*s*, the spell-out position of *has*. We can cause our principles entail that the *have*- node follows the -*s* node in (10') in exactly the same way in which they required the -*s* node follow N^1 in (9') if the (-*s*) node in (10') is visible and the (-*s*) – -*s* relation counts as word-external.

We must not take however the relation between *have*- and (*have*-) to be word-external in (10') since that would result in *have*- immediately dominating two elements with which it would be in a word-external relation. This would create a contradiction: both elements would have to precede *have*- (7) and both would have to be adjacent to it (8a,b). Capitalizing on the difference between the -*s* – *have*- (word-internal) and the *have*- -*en*- (word-external) relation, we might ensure that (-*s*) - -*s* but not (*have*-) – *have*- is a word-external relation using the rather natural assumption in (11). According to this, elements that form no word-external relations (i.e. are word-internal morphemes) must be terminals of the tree.

(11) A non-terminal node N must ID some element with which it is in a word-external relation

Since the relation between -*s* and *have*- is word-internal, (11) entails that the (-*s*) - -*s* relation is word-external. Given (11) or some equivalent device, the invisible (-*s*) node in (10') will have to precede -*s*, with which it forms a word-external relation. Hence the -*have* node and its domain must follow -*s*, by (8ab). So eaten, dominated by -*have* will follow -*s*, the node where has is spelt out; and we have derived the order: *has* precedes *eaten*.

6. Bare checking, – the elimination of probe-goal asymmetry

The minimalist checking theory Chomsky (1994) is asymmetric: a c-commanding uninterpreted (and unvalued) probe searches for a c-commanded interpreted goal (with a value) to inherit its value from.

As is well known there are cases where the c-commanding probe appears to be interpreted raising an apparently significant problem for this theory. A typical example is (12) where it would seem that an interpreted +*WH* probe searches for a goal *wh*-phrase.

(12) I wonder *who*(+*wh*)	C(+*WH*) +interpreted probe (?)	Bill saw *t*

Pesetsky and Torrego (2004) point out, that Chomsky needs to posit "two distinct features in C: an uninterpretable, unvalued feature *u*Wh with an EPP property (the feature that probes for a *wh*-goal); and a distinct, interpretable, valued feature *i*Q (the feature relevant to the interpretation of the clause). Correlations between clausal semantics and *wh*-type must be captured with mechanisms other than Agree" (p.7).

Their alternative proposal is to separate interpretability and valuation. In other words they assume that the +*WH* probe in (12) is interpretable but unvalued, and receives its value from an uninterpreted but valued *wh*-phrase goal.

We might wonder if the feature duplication of the earlier solution has not been simply exchanged for a different but perhaps equally unnecessary distinction between interpretability and valuation. Both duplications are made necessary by the assumption of asymmetry in the probe-goal configuration: that an uninterpreted (Chomsky) or unvalued (Pesetsky and Torrego) probe searches its c-command domain for a goal that provides interpretation or value for it. The suspicion that both duplications may be redundant is reinforced by the observation that it is an apparently unnecessary additional restriction that makes it necessary to invoke them.

Bare checking theory (Brody 1997b) simply requires all instances of features with a single semantic interpretation to be linked. Thus the interpretable +*WH* and the non-interpretable *wh*-feature of the *wh*-phrase must be (chain-)linked, there is no need to invoke an interpretation-independent notion of valuation. We do not need to talk about valuation, but if we wish to do so, we can allow the interpreted feature to value the non-interpreted one, – as expected on general grounds given the fact that valuation is semantically redundant. Without the ad hoc proviso that the structurally higher element must be non-interpretable or unvalued, we do not need to further duplicate our featural inventory invoking two sets of features for the analysis of such construction. See Brody 2000b for further discussion.

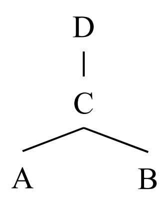
7. In a minimal symmetric syntax Move/Chain is undefinable

Returning to interpretive labelling, this copies a (set of) feature(s) to the immediate dominating node, where ID relation is understood in the standard way: *x* ID *y* iff *x* dominates *y* and there is no *z* such that *x* dominates *z* and *z* dominates *y*. In other words the ID relation (of interpretive labelling) implies a locality condition. The same locality condition is of course also implicit in the ID relation involved in the standard definition of Merge/sisterhood. Standard Merge creates a node that is in an ID relation to the merged categories. If A and B are sisters then there is a node that is in an ID relation with both.

We might ask if the duplication of invoking ID/locality in both syntax and interpretation is indeed necessary. Suppose A and A' are merged (sisters) and are (immediately) dominated by A"; similarly B and B' are (immediately) dominated by B". When we merge A" and B" creating say C, we take C to be the set whose members are (only) A" and B". Suppose however, that we understand Merge differently, so that merging the trees A" and B" will still result in C, but C is now defined as the set of all nodes it (non- reflexively) dominates and not directly by the ones it is in an ID relation with. In other words C dominates/contains A, A', A", B, B', and B" and C is defined as the set whose members are exactly these elements.

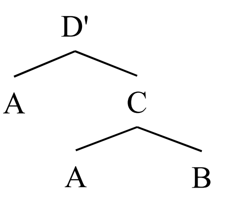
An interesting consequence of this apparently innocuous simplification for the sake of eliminating a central redundancy is that the trees in (13) - (17) cannot be distinguished. They all represent the same single structure, drawn in different ways. Thus (13) for example contains two non-terminal nodes D, and C where C is defined as the node that dominates A and B while D is the node that dominates A, B and C.

(13)



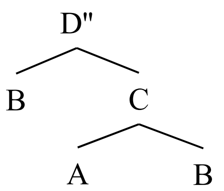
C: {A, B}
D: {C, A, B}

(14)



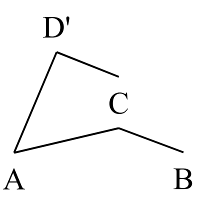
C: {A, B}
D': {C, A, B} = D

(15)



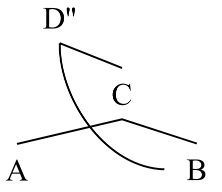
C: {A, B}
D": {C, A, B} = D' = D

(16)



D' = D

(17)



D" = D' = D

Given the standard set-theoretical assumption that the set {A, A} is the same set as {A}, the tree in (14) contains exactly the same terminal and nonterminal nodes as the one in (13). The same is true for the other examples. But (14) and (15) are standard representations of movement/chain formation, where in (14) A and in (15) B has moved. In other words, since (13), (14) and (15) are exactly the same tree, Move or Chain is undefinable in the syntactic system that dispenses with the locality condition of the ID relation. (16) and (17) are corresponding 'remerge' structures, that indicate that the conclusion carries over to these. See Brody 2005a for a discussion of this weaker definition of syntactic trees.

8. Summary

I argued for a symmetric syntax that uses unordered (labelless) sets, where each node is understood as the set of all nodes (non-reflexively) dominated by it. Chain construction is interpretive and free ("link alpha"), subject to constraints like super-salience and checking. (Bare) checking (doing much of the work of the movement rules of earlier theories) is also symmetric. There is a single kind of linear ordering provided by the domination relation that captures top to bottom antisymmetry of the tree and which is exploited by interpretive labelling. This is the common basis of both (morpho-) phonological linearization and (quasi-) semantic c-command phenomena.

References

Brody, Michael (1997a). Mirror theory. Ms. UCL, http://www.phon.ucl.ac.uk/home/misi/mirror_w5.pdf

Brody, Michael (1997b). Perfect chains. In L. Haegeman (ed.). Elements of Grammar. 139-167. Amsterdam: Kluwer.

Brody, Michael (1998) The minimalist program and a perfect syntax. Mind and Language 13:205-214.

Brody, Michael (2000). Mirror theory: Syntactic representation in Perfect Syntax. Linguistic Inquiry 31.1. 29-56

Brody, Michael (2005a). Move in syntax – logically necessary or undefinable in the best case? http://ling.auf.net/lingBuzz/000194

Brody, Michael (2005b). Bare checking – without probe-goal asymmetry. http://ling.auf.net/lingBuzz/000195

Chomsky, Noam (1994). Bare phrase structure. MIT Occasional Papers in Linguistics. Chomsky, Noam (2005). On Phases. Ms. MIT.

Collins,Chris (2002). Eliminating labels. In: Samuel Epstein and Daniel Seely eds. Derivation and Explanation in the Minimalist Program. Blackwell Publishers.

Pesetsky, David and Esther Torrego (2004). The syntax of valuation and the interpretability of features. Ms. MIT and Umass/Boston

Schlenker, Philippe (2004). Minimize restrictors! (Notes on definite descriptions, Condition C and epithets). http://ling.auf.net/lingBuzz/000220

Starke, Michal (2001). On the inexistance of specifiers and the nature of heads. Ms. University of Geneva. Stuart, Ian (1995). Nature's Numbers. Harper Collins Publishers, Inc.