Balázs Surányi

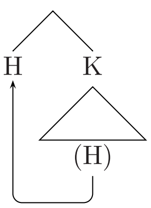
Head movement qua *Root Merger**

1 Introduction

It has been a recurrent theme in the recent minimalist literature that head movement sticks out in a typology of movements as exceptional, and hence its status in the computational system itself is questionable. The ultimate source of the exceptionality of head movement is the assumption that head movement involves adjunction.[1]

The present paper argues that it is possible to retain the descriptively beneficial aspects of head movement which head movement has been motivated by, but to do away with the unwanted complications. According to the view advocated here, head movement is not adjunction, but, in terms of generalized transformations, root merger. This means that under the right conditions a head H can be moved out of the current phrase marker K, merging H with K and projecting H into HP, with K a complement of H, as below.

(1)



I demonstrate that this view of head movement stays clear of the problems associated with head movement *qua* adjunction. I argue further that head movement *qua* root merger is driven by cyclic spellout (Epstein et al. 1998; Uriagereka 1999; Chomsky 2000, 2001), and results in checking the c-selectional feature of some component of H—hence, it is fully compatible with the Last Resort character of the computational system. On the present proposal the symmetry of head movement

* The present paper is a short version of a write-up of my talk at GLOW 46 in Lund, Sweden in April 2003, and an invited lecture at Yeungnam University, Korea in August 2003. I thank the audiences for their questions and comments. I gratefully acknowledge the support of project grant OTKA No. TS 40 705 and the Békésy György scholarship. The discussion is simplified here for reasons of space limitations; see Surányi (2002, 2004a) for more details.

[1] Recent reactions to the problematic nature of head movement in minimalist theory follow two markedly different paths: according to Chomsky's (2000) suggestion, head movement is to be relocated to the PF branch of the computation, while proposals have been made (e.g., Sportiche 1998; Mahajan 2000, 2001; Koopman & Szabolcsi 2000) to reanalyze head movement effects as resulting from remnant XP movement. While these are viable treatments of various displacement patterns involving apparently word-level elements, neither of them can be maintained as a general approach to head movement phenomena, as argued in Surányi (2004a).

and phrasal movement is more pervasive than in the standard account, and the remaining idiosyncrasies of head movement follow from what it is driven by.

2 Head movement *qua* adjunction in checking theory

Head movement is described in standard minimalism as an adjunction operation, moving a lower head element to adjoin to a higher head category (cf. Baker 1988; Chomsky 1993, 1995)

(2) $[_{YP}\ [_{Y}\ X_i\ Y]\ [_{XP}\ X_i\ ]\ ]$

Two major problems that have been noted for this structural description of head movement are (i) and (ii) below (cf. Brody 1997; Mahajan 2000, 2001).

(i) It requires a complication of the definition of c-command to allow X to c-command out of Y if something like a c-command condition (or Proper Binding Condition) is to hold of movement relations. This complication is typically phrased in terms of a distinction introduced to hold between containment vs. dominance or between segment vs. category.[2]

(ii) Head movement apparently defies the Extension Condition on Generalized Transformations (i.e., it is counter-cyclic).[3]

These are two often acknowledged problems related to adjoining head movement. In fact, adjoining head movement faces further severe complications. To these I turn next.

(iii) Head movement qua adjunction, as conceived under the checking theory of movement, creates a disjunction in the definition of the checking domain: functional head H is checked either by an adjoining head, or by a specifier.

In fact there are two aspects to this problem. The first one is conceptual: disjunctive definitions are inelegant. The disjunction itself is not overt in the definition of checking domain; it is concealed in its negative definition: the checking domain of H is the set of positions that form minimal residue of the domain of H but do not belong to the complement domain of H (Chomsky 1993). The checking domain then is heterogeneous, though this problem is technically circumvented through a non-naturally complex and *negative* definition. Further, the notions that figure in this definition (residue of H, complement domain of H) do not have any role in the theory independent of this very definition: another reason for concern.

[2] Arguably, this problem does not arise if the c-command condition holds of the Agree relation.

[3] Covert movement has ceased to be a post-Spell Out operation, and is no longer counter-cyclic, given the elimination of the bifurcation of the overt and covert cycles (cf. Groat & O'Neil 1996; Bobaljik 2002; Pesetsky 1998; Uriagereka 1999; Chomsky 2000, 2001; see also Brody 1995). Chomsky (2000) replaces the Extension Condition with the Least Tampering Condition, which demands that c-command relations previously established in the derivation must be preserved after movement. As I argue in Surányi (2004a), the Least Tampering Condition faces an overgeneration problem: it is a weakening of the Extension Condition that allows not only head-adjunction, but also operations that are clearly never attested.

What makes these definitional problems all the more prominent is the fact that what the definition defines is probably *the* central structural relation in syntax; a relation that on minimalist grounds (i.e., if language is of optimal design) is expected to be simple.

The same incongruity is preserved in recent modifications of the checking mechanism (which takes place under Agree), where the non-homogeneity of the structural domain licensing checking (valuing/deletion) has two facets. First, it applies to the generalised EPP feature itself, which, once again, can be eliminated *either* by head-adjunction *or* by phrasal merger/displacement into a specifier position. Second, a feature can be checked either under Agree for movement (essentially, c-command, confined to a local domain by the Phase Impenetrability Condition), or via (first) Merge into specifier position in case of (phrasal) expletives.

In fact, there is more to this problem than the conceptual incongruity. This disjunction in the checking domain in turn renders phrasal movement dependent on head movement to the same functional projection, as well as interdependent head and phrasal movement to the same functional projection—although possible to encode—impossible to explain. The first scenario is illustrated by overt or covert head-inversion to a functional head H which occurs only if some operator moves to the specifier of head H.

(3) $[_{\text{HP}}$ OP H^0 [

In fact, the interrelations of two movements are currently encoded by positing two, in principle independent formal features on the attracting functional head, one attracting a phrase into the specifier position, and one attracting a head to a head-adjoined position. Then, the generalization that the phrasal movement only occurs if head H bearing the operator feature is moved up is inexpressible. The second scenario is illustrated by agreement projections. As Chomsky (1995) suggests, they are merely projected to "house" the checking relation of a head and a DP. The problem that Chomsky raises is that the Agr head itself is never interpretable. Three syntactic elements bear agreement morphemes: the verb, the agreeing DP and the Agr functional head. Only one of these agreement morphemes is semantically interpretable, two are not. Moreover, to match the verb with Agr, one is forced to posit *two* morphemes in the Agr head, one attracting the verb, and one the DP. This means that there are altogether four morphemes, out of which only one happens to be interpretable. Within this system, Chomsky (1995) concludes, Agr projections had better not exist; he proposes that they don't. However, many researchers insist that there is solid empirical evidence that such projections do exist (cf. e.g., Belletti 2001 and references therein).

(iv) The locality of head movement is unmatched in the domain of phrasal movement: the locality of head movement is significantly stricter inasmuch as head movement cannot skip *any* c-commanding head position, cf. the Head Movement Constraint (HMC).[4]

The restriction that head movement cannot "excorporate" plays a crucial role in accounting for HMC effects. The "no excorporation" restriction means that although phrasal movement is successive cyclic, there is no successive cyclic head movement. Nevertheless, as Brody (1997) points out, in current theory, the "no excoporation" restriction is not properly derived from an independent source, and remains stipulative (cf. the spurious reference to a "WI" component in Chomsky 1993). Baker (1988) suggests that the excorporation prohibition follows on the assumption that word-internal traces cannot exist. It is not clear why this should hold. On the one hand, in not (or not fully) lexicalist approaches like Baker's, if separate heads can come together in syntax to form a word, why can they not separate again: what word will be sent to morphology ought to be determined by the final output of syntax, not some intermediate representation. On the other hand, on the lexicalist approach of checking theory, if words are not created in syntax, but prior to syntax, then word-internal traces will not arise to begin with.

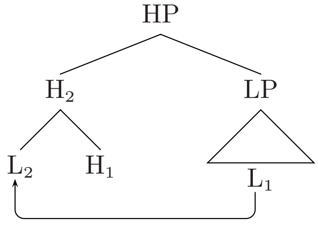
But even if the prohibition against excorporation was derived in some principled way, it would still not be clear why a functional head H cannot attract a head B which is further down in the hierarchy than the immediately next lower head A, if Relativized Minimality/MLC/Attract is sensitive to (classes of) features.

(4) [H [A [B

If the closest head that has the required feature is not the immediately next A, but the less close B, then B should be able to be attracted, unless further conditions are added. The asymmetry of head movement and phrasal movement appears to be a deviation from optimal design.

(v) Head movement *qua* adjunction also incurs complications with regard to the Uniformity Condition on chains (a descendant of Structure Preservation) (Chomsky 1993). This is because even if it is ensured that the moved head itself does not project, strictly speaking, a head-chain is not uniform in a Bare Phrase Structure theory building on relational definitions of projection levels. The lower link L_1 of a head chain projects, hence it is non-maximal (in fact, minimal), while the higher link L_2 does not project, hence it is maximal.

(5)



[4] Apparent long head movement phenomena are assumed here to be analysed as involving XP-movement.

(vi) Head movement as attraction to another head poses complications with respect to the strong/weak distinction. For instance, unless we make additional assumptions, it is unclear why, in some Germanic languages (including Mainland Scandinavian and some embedded contexts in German) clausal inflectional heads appear to alternate between being weak and being strong. They are weak and hence do not attract the verb overtly in embedded contexts (cf. (6b)), while they do so—given that head movement must proceed cyclically through Infl—in case there is a strong C above them, cf. (6a) ((V) stands for the landing site of covert V-movement).

(6) a. $[_{CP}\ [_{C}\ [_{T}\ [\ V\ ]\ T\ ]\ C\ ]\ \ldots\ [_{TP}\ SU\ [_{T}\ [\ V\ ]\ T\ ]\ \ldots\ [_{VP}\ [_{V}\ V\ ]\ ]]]$
b. $[_{CP}\ [_{C}\ C\ ]\ \ldots\ [_{TP}\ SU\ [_{T}\ [(V)]\ T\ ]\ \ldots\ [_{VP}\ [_{V}\ V]\ ]]]$

In cases like this, where an element H_n raises up to a higher head H_1, successive cyclically through the intervening head positions H_2, $H_3 \ldots H_{n-1}$ in one context, but stays in situ in another context where H_1 is saturated in some other way (or filled by some other element), the intervening heads H_2, $H_3 \ldots H_{n-1}$ must uniformly be "strong" in the first type of context, and uniformly "weak" in the second. This is frequently the case despite the fact that there may not be any other, independent difference, say in terms of finiteness or otherwise, between the two occurrences of the H_2, $H_3 \ldots H_{n-1}$ series. This patterning can be modelled technically by assuming that the high head H1 itself selects a "strong" H_2, which in turn selects a "strong" H_3 and so on down to H_{n-1} in the first case, whereas H_1 selects a "weak" H_2, which in turn selects a "weak" H_3 and so on down to H_{n-1} in the second case. In essence, such a technical (in terms of Chomsky 2001, "engineering") solution introduces two options in the lexical entries of H_1, $H_2 \ldots H_{n-1}$.

(vii) Finally, in a checking theory of head movement to functional projections, the functional heads that we identify as landing sites for head movement are typically phonologically empty. This may also be considered a drawback of the standard approach, if an important motivation of a functional head is the phonologically overt occurrence of that head. If a syntactic head must systematically be phonologically empty, this weakens the motivation of positing that head to begin with.

3 Head movement as Root Merger and structure building in strictly derivational syntax

3.1 Head movement as Root Merger

I now examine a possible response to the above problems which retains the descriptively beneficial aspects of head movement (which head movement has been motivated by), but which at the same time does away with the unwanted complications.

The alternative is to treat head movement as uniformly involving root (re-) merger. In minimalist terms of generalised transformations, under the right conditions a head H can be moved out of the current root phrase marker K, merging H with K and projecting HP, as in (1) above (the Root Merger Hypothesis, RMH).

This movement can be referred to as "substitution" instead of adjunction (in terms of a now anachronistic bi-partitioned typology of movements).[5] Re-merge, just as with phrasal movement, is recursive, i.e., head movement in these terms can be successive cyclic.

Chomsky (1993, 1995) argues extensively that moved phrases cannot project, given that that would violate the Uniformity Condition.[6] However, the Uniformity Condition does not exclude *heads* re-projecting after movement. This is because H is both non-projected and projecting in both links of the head chain in (1). What rules that option out according to Chomsky (1995 : 257) is that it is apparently not in conformity with Last Resort. I address this issue in **§3.4**.

For the moment, let me tentatively adopt (1), fleshing out the merits of the proposal. Below I demonstrate how the RMH resolves the problems for the treatment of head movement as adjunction above.

3.2 What the RMH buys

I address the enumerated complications in the order they were presented in **§3**.

(i) As far as the c-command condition is concerned, the moved head evidently c-commands its trace position. No definitional problems arise.

(ii) The Extension Condition is also trivially satisfied: the moved head extends the root. Head movement is no longer exceptional (and it does not require relinquishing the Extension Condition in favour of the problematic Least Tampering Condition; see note 3).

(iii) The hidden disjunction in the negative definition of checking domain is also dispensed with: given that there is no head already existing prior to head movement, one of the two configurations of the local checking relation ceases to exist. Then, the checking configuration in principle can be defined directly — a welcome consequence. I discuss the issue of checking and trigger further in **§3.3**. I take up the matter of interdependent head and phrasal movements in **§3.4**.

(iv) The effect of the HMC, i.e., the strictly local nature of head movement, in principle can be derived on this account if it can be shown that once external merge of a new head N occurs, a lower head H cannot be re-merged with the root. I defer this issue as well until **§3.3**.

(v) Given that the moved head projects, as I pointed out before, the Uniformity Condition is conformed to. No movement occurs into a head, hence a non-uniform chain which is maximal upstairs and non-maximal downstairs cannot come about. I believe that this is a step in the right direction in that nothing like the Uniformity Condition is expected to be part of syntax on minimalist assumptions: the Uniformity Condition is difficult to motivate as a bare output condition. Consider the analogy of the c-command condition on movement (aka Proper Binding Condition). Such a restriction ceases to be a condition per se in minimalism, as its effect falls out from the mechanism of the computational system of generalized

[5] Ackema et al. (1993) utilized V movement as substitution, due to reasons and motivations largely independent of those presented here.

[6] Though not all researchers share Chomsky's position, cf. e.g., Starke (2001).

transformations based on Merge. Similarly, the Uniformity Condition should have the status of a descriptive generalization: its effect should be the consequence of how permissible elementary operations of the computational system are defined. On the present account, one unwanted non-uniform chain configuration is eliminated, an important step towards removing the Uniformity Condition as such—fulfilling minimalist expectations.

This result is welcome. However, a caveat concerns the other relevant possibility of a non-uniform chain created through head movement. In the standard adjunction treatment, the other relevant restriction is that it cannot be the moved head that projects. In contrast, what needs to be guaranteed on the present account is that if head H moves and merges with the root, the target, i.e., the root, should not project. That this can indeed be guaranteed will be shown in **§3.3**.

(vi) The matter of once "strong", other times "weak" inflectional features will be discussed in **§3.5**.

(vii) Finally, recall the problem faced by the adjunction theory of head movement in checking theory related to morphophonological emptiness: the attracting inflectional heads on such a theory are invariably empty. This, I suggested, works towards undermining the very existence of those functional heads themselves. Now it should be clear that on the present account of head movement, there is an explanation for why those attracting inflectional heads are uniformly empty: this is because they do not exist prior to head movement at all.

3.3 The Indirectly Driven Movement Hypothesis

On minimalist assumptions, movement is driven by legibility conditions at the interfaces. However, if (1) is the correct conceptualization of head movement, then there is no attracting head that head movement targets. The question in a minimalist setting then is: what drives head movement? I argue now that head movement is akin to Indirectly Feature-Driven Movements (IFM) of Chomsky (2000, 2001).

3.3.1 IFM and head movement

Chomsky assumes that syntax is a cyclic derivational structure building mechanism where the operation of Spell Out applies once per cycle. Effectively the same conception of cyclic, multiple Spell Out is put forward in Epstein et al. (1998) and Uriagereka (1999). Chomsky makes the particular proposal that the cycles that are relevant for Spell Out are to be identified with strong phases, essentially CP-s and vP-s with an external argument (perhaps DP-s and PP-s as well), i.e., "Spell Out is cyclic at the phase level" (Chomsky 2001 : 9). The proposal is that "interpretation/evaluation for [phase] PH1 is at the next *relevant* phase PH_2" (Chomsky 2001 : 10, (9)). Interpretation of PH1 cannot be at PH1, because that would terminate the derivation (and in all cases (except the root CP) would result in crash). Thus, there is a small delay, which Chomsky identifies as one "relevant" phase, where "relevant" is specified as strong, i.e., only strong phases count. The

Phase Impenetrability Condition (PIC) restricted to strong phases is a case of this slightly delayed interpretation property of the computational system.[7]

Elements that have yet unchecked (unvalued) offending features at the completion of a strong phase then must move to the phase edge if they are to be accessible for later Agree operation. The syntactic mechanism that implements this is non other but the optional assignment of an uninterpretable EPP- or P(eripheral)-feature of the head H of such phases.[8] "The two features are introduced to allow the general theory of movement to apply without change in this case" (Chomsky 2000 : 23, fn. 51). However, such optionally assigned uninterpretable features should give cause for minimalist concern. In fact, Indirectly Feature Driven movement (i.e., movement to intermediate phase edge positions) is determined locally. As Chomsky puts it, "Local determination is straightforward: [...] an uninterpretable feature in the domain at the phase level determines that the derivation will crash" [unless it is moved to the phase edge, *BS*] (Chomsky 2000 : 22). In other words, it can be locally determined that there are two alternatives: either movement of the offending element does not apply, in which case the derivation crashes at the next step, or the offending element is moved, in which case the derivation can continue. In this light, the introduction of the uninterpretable features on the phase head H does not appear to be strictly necessary: IFM is locally determined to be unavoidable.

This idea is confirmed by a look at the operation of QR in Chomsky's (2001) system. QR is not feature-driven, but must have an effect on the output (=,INT) (Chomsky here follows Fox's and Reinhart's relevant work).[9] Then movement is still a free operation, though it must result in an immediate/local effect on the output: (i) either elimination of an offending feature (by valuing it, or deleting it (the latter in the case of EPP)), or (ii) a difference in interpretation (QR). Now IFM technically results in the elimination of an uninterpretable EPP/P-feature on the head of a phase. However, the offending feature in the tail position of the IFM-chain is not valued. It appears that by moving the element that bears it to the relevant phase edge, the offending feature in the tail position of the IFM-chain is removed: it ceases to be offending, and the phase can safely go to Spell Out. If this is so, then the idea that IFM is determined locally to be unavoidable and hence the optional introduction of an uninterpretable EPP/P-feature is redundant can be implemented by formulating Last Resort for movement as below:

(7) **Last Resort**
A syntactic movement is licensed iff it results in the elimination of an offending feature from the Spell Out Domain.

[7] In Chomsky (2000), every phase (excepting its head and edge domain) becomes inaccessible for later syntactic operations once the next higher phase is completed, whether they are strong or not, i.e., the PIC applies at each phase level.

[8] Chomsky (2000) formulates this feature assignment as being carried out in the computational system; an equivalent treatment would assign these features already in the Lexical (Sub)Array.

[9] Surányi (2003, to appear) argues extensively against the Stowell-Beghelli-Szabolcsi approach, in which quantifiers move to check formal features, and presents an account of their data relying on non-feature-driven QR.

The Spell Out Domain is identified as the domain of a strong phase head by Chomsky (2000, 2001). Given that on the present account nothing moves to adjoin to heads, the head may be included in the Spell Out Domain, i.e., the still accessible domain can simply reduce to the edge itself. I propose that this is so.

The formulation in (7) builds on the idea that by moving the element to the edge of a strong phase, the offending feature is removed from the tail copy of the IMF chain. Last Resort entails that IFM can only involve displacement to the edge of a strong phase, and cannot apply strong phase internally, since that would not take the offending element out of the Spell Out domain. Other movements, which all result in the elimination of a feature off the moved element too (which includes elimination of the offending feature in the tail copy), conform to Last Resort without change.

With the above treatment of IFM in mind, assume now a situation where head H of a strong phase PH still bears an offending uninterpretable feature upon completion of the phase which no element internal to PH can satisfy. At that point, in principle there are two options: either a new head is merged, or H is moved and merged with PH. In the first case, The Spell Out Domain will contain an offending feature: that of H. In the second case, this feature will be removed from the Spell Out Domain itself. Then, such head movement will be licensed under Last Resort as a case of IFM.

3.3.2 Phases: minimal delay and selection

Chomsky (2000, 2001) assumes that Spell Out applies only at a subset of phrases, viz. strong phases. Here I will follow Müller (2003) and adopt the view that each maximal projection is a phase in the derivation, in particular, a strong phase. Essentially the same view is proposed by Takahashi (1994), although in different terms.

In fact, Chomsky's identification of the class of Spell Out domains with the domain of C and strong v is both conceptually and empirically suspect. On Chomsky's definition, phases and only phases are assumed to be propositional and phonologically independent. One criticism that can be levelled at such a definition, is that these two definitional criteria do not appear to be either sufficiently precise or empirically accurate if they are expected to single out CP and vP.

On the other hand, if there is a conceptually necessary delay in semantic/phonological interpretation, then the delay up to strong phase is not optimal. The delay of the application of Spell Out is unavoidable: otherwise the built structure would be sent to interpretation systems and the derivation would be terminated. However, on minimalist assumptions (i.e., if syntax has optimal design), then this delay is expected to be minimal, incurring minimal operational complexity. That is, on conceptual grounds, an already built portion of structure is expected to be subjected to Spell Out at the earliest possible subsequent stage in the derivation. It is assumed that when a head is introduced in the derivation, it establishes multiple Agree relations simultaneously (cf. e.g., Boškovic 1999). That entails that all uninterpretable features of a head must be checked/eliminated immediately after merging the head itself. In consequence, the earliest possible stage for Spell Out cannot be arrived at before all uninterpretable features of the head are eliminated.

Then the operation of Spell Out can apply: the Spell Out Domain of the current phrase marker (i.e., the head merged with its complement) is sent to the interpretive components. In brief, if syntax has optimal design, then each phrase must be a strong phase. This is then the null assumption regarding the necessary delay of Spell Out; an assumption I term Minimal Delay.

Transposing the conclusions of the discussion of IFM in §**3.3.1**, IFM then applies to the edge of each phrase. I argued above that IFM subsumes movement of a head bearing an uninterpretable feature. I suggested that IFM of the current head along the lines of (1) is the only option if the current head is the head of a strong phase, which is to be sent to Spell Out. That, taken together with the assumption of Minimal Delay yields the consequence that on the completion of each phrase whose head H still has an uninterpretable feature that cannot be checked by an element from its complement domain will move up as outlined in (1). This will continue cyclically until all uninterpretable features of H are finally eliminated.

This mechanism does not allow a derivational move where K=/=H in (1). In other words: it is guaranteed that in (1) K = H. Put differently, no HMC-violating derivations are ever possible: the HMC effect is derived.

At this point an issue that I have consistently been agnostic about becomes significant: when a head moves up in the fashion of (7), why should the label of the newly created projection be H? Discussing labels, Chomsky (2000) points out that the determination of labels is straightforward in the computational system, given that each merge relation is asymmetric. Adjunction is asymmetric by definition (pair merge). In set merge, the relation of the two elements is also asymmetric, given that the relation is either that of selector and selected item, or checker and checkee (the EPP feature needs to be eliminated). On some analyses (cf. Svenonius 1994; Holmberg 2000), (c-)selection is also feature checking; similarly, arguments check theta features (cf. Boškovic & Takahashi 1998 and references therein). I will follow this approach here: all set merge (whether first or second) is driven by feature checking needs. This is formulated in (8):

(8) Merge is triggered by checking needs.

Labels are determined straightforwardly by the following simple descriptive principle:

(9) **Labelling principle**
The checked element (probe) projects.

If (c-)selection is checking, and if the complement B c-selected by head A is canonically sister to A, then a straightforward hypothesis is that the checking of c-selectional features occurs under sisterhood. In fact, that all checking occurs under sisterhood was proposed by Zwart (1992, 2003), and in a different setting, by Epstein et al. (1998), who formulate a notion of "derivational sisterhood", relying on derivational mutual c-command (which properly includes the ordinary notion of sisterhood). Let us assume this to be correct:

(10) Checking configuration is sisterhood.

In terms of deriving syntactic relations from operations of the computational system, this is the optimal situation: the elementary operation is Merge. Then (10) can be stated as (11):

(11) Checking occurs under Merge.

(8) and (11) together determine that when H in (1) is moved up and merged with K, either checking of H or checking of K must occur. I propose that the checking that takes place falls into the category of (c-)selection: H (c)-selects K. Therefore, H projects.

This proposal entails that H can move *qua* IFM in the fashion of (1) only if H (c-)selects K. But if K = H, as I have suggested, then how could H select itself? The answer is essentially that the H in the lower position (H_1) and H in the higher position (H_2) are not categorially identical, and H_1 selects H_2. This paradox is resolvable once the checking mechanism is scrutinized within derivational syntax.[10]

In what follows, I build on Chomsky's (1993; 1995) checking theory, but propose to modify his strong lexicalist position: inflectional affixes, say, of a verb are treated as already being on the verb when it combines with its object (as in checking theory), but their presence is not only morphological, but also syntactic. Recall that in checking theory it is assumed that heads enter syntax fully inflected (i.e., a strong lexicalist position). Further, the morphologically inflected heads also carry syntactic checking features, whose order is the exact mirror image of the (relevant partial) hierarchy of extended projections above the lexical projection. Out of these ordered checking features, it is always the currently outermost feature that can enter checking; that feature is syntactically active. This duplication and the stipulation of a mirror image ordering serves to derive the Mirror Principle, expressed schematically in example (12) (or rather, Mirror Generalization; see Brody 1997 for relevant discussion).

(12) affixation

$$\begin{tikzcd} \mathrm{V} \arrow[r, bend left] & v \arrow[r, bend left] \arrow[l, bend left] & \mathrm{T} \arrow[l, bend left] \end{tikzcd}$$

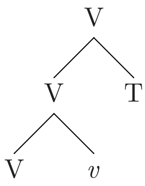
c-selection

I propose to eliminate both the duplication and the ordering stipulation by assuming that the inflectional affixes combine with the stem syntactically prior to the point where the fully inflected stem merges with another (independent) element.

I now show that this assumption makes it is possible to explain how H_1 can select H_2. The structure of the complex verbal head in a language like English is given in (13):

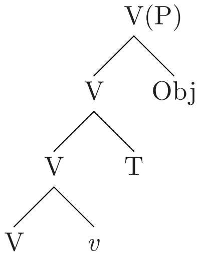
[10] Due to space limitations, the discussion that follows is simplified; see Surányi (2004a) for the details.

(13)



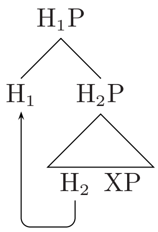
V checks its c-selectional feature against its complement, say an object DP, and because it was V whose feature was the probe in this case, it is V that projects, i.e., the label of the newly formed constituent will be V. This is represented in (14).

(14)



is built. Since its head still contains uninterpretable features, IFM of the head takes place. (1′) below is re-written from (1) to reflect the results of the preceding discussion (the phrasal projection level is used merely as a notational convenience; XP is complement to H_2). In the movement of the head in (14), H_2 is (13); i.e., the head corresponding to (13) within (14) will be merged with (14).

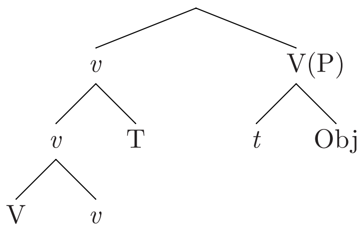
(15)



Chomsky (1995, 2000) demonstrates that labels can be determined by the asymmetry in the relation of the two merged elements. More recently, Hornstein and Uriagereka (2003) have argued that the labels themselves can change in the course of the derivation (binary quantifiers can covertly reproject after meeting their syntactic requirements).[11] I propose to adopt this view here. Since the V component in (14) has checked all its features, it ceases to be a probe. That entails that when the complex head is moved up, it is no longer labelled by V. It is then the next element, *v*, that will determine the label:

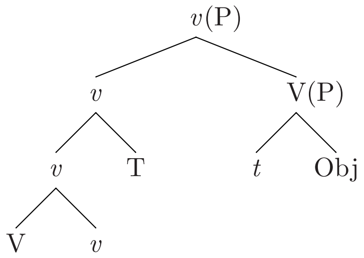
[11] A radical position is put forward by Collins (2003), who argues for the elimination of labels as such. I believe that the reprojection proposal that I am making here can be transposed to a system without labels.

(16)



Then, the c-selectional feature of *v* is checked now, and the label of the projection will then be the element whose feature was checked, i.e., the probe *v*, as in (16).

(17)



If *v* has a further uninterpretable feature, i.e., if there is an external argument, then a further merger takes place, still labelled by *v*.

Covert head movement in this model will need to be considered identical in all respects to overt head movement, except that the head will not be pronounced in the head link of the chain. That is, the overt/covert distinction is only a matter of which syntactic copy is assigned a phonological form (in line with Groat & O'Neil 1996; Bobaljik 2002; Pesetsky 1998; Brody 1995). This can be conceptualized as follows: if an element has checked off all its PF-uninterpretable features then it is ready to be spelled out, and therefore it will be sent to PF. An analogous assumption is made in Chomsky (2001):

> "The simplest assumption is that the phonological component spells out elements that undergo no further displacement, with no need for further specification [i.e., checking, *BS*]." (Chomsky 2001 : 10)

Covert movement occurs if although PF-uninterpretable features are already checked, there are still LF-uninterpretable features to saturate. Then, such an element will keep moving covertly (with its phonological matrix already stripped away) until it is fully checked. That is, for the present purposes covert head movement and overt head movement are treated exactly alike, except for the spell-out position in the chain.[12]

[12] An alternative is to adopt Lasnik's (1995) hybrid approach in which covert head movement corresponds to affixal, but syntactically self-standing heads undergoing a morphological Affix Hopping operation, while overt head movement involves syntactic head movement.

Having illustrated how the present assumptions about the nature of IFM, phases and selection combine to produce derivations, I return to the loose threads of **§3.2**. One central concern, the effect of the HMC has been pointed out to be derived in the proposed model on the assumption that Spell Out applies at the completion of each phrase.

Another issue concerned the definition of the checking domain, which is a heterogeneous set of position on the standard account. With the head-adjoined position gone from the checking domain, the basic disjunction between head-adjunction and specifier positions is dispensed with. The checking domain can be defined directly. The treatment I adopted takes checking to be realized as part of the elementary operation of Merge.

3.4 Interdependent phrasal and head movements

I come now to the issue of overt movement patterns where a head and one (or multiple) phrase(s) are raised to the same functional projection, and the two (sets of) movements are interdependent. One case in point is operator movement accompanied by verb or auxiliary inversion. For ease of demonstration, I will write "WH" for the operator(s) in question, and "Aux" for the inverted head, and will use English-type root *wh*-movement to exemplify, within a simplified C-T-v-V clausal hierarchy. Now in the basic order without operator movement, Aux is not inverted; it is in T. Hence in the present model its selectional feature is not "strong"/PF-uninterpretable; it will not be moved overtly. Consider the case where Aux has a strong operator feature [op]. This in the present approach will be a property of the C morpheme within the complex head Aux. Given that at the stage when TP is completed Aux still has a strong uninterpretable feature, the movement of Aux by IFM will be overt, i.e., overt T-to-C is triggered. The rest is as normal: [op] attracts WH to CP. This account captures the mutual dependence of the two overt movements to the same projection, and it predicts that whenever the head in T contains a strong [op] feature, T-to-C and movement to [Spec,CP] will both be overt. Note that, correctly, it does not make the prediction that overt operator movement cannot exist without head-inversion: that will continue to be treated in the standard manner, i.e., either adjunction is involved, or a phonetically empty operator head does the operator-attraction.[13]

Notice that the related problem of the triplication (or quadruplication) of agreement features associated with AgrP projections can be avoided now. There is one agreement morpheme that is part of the complex verbal/nominal/etc. head (corresponding to the Agr head in the standard accounts), and one agreement morpheme on the agreeing phrase, typically a DP. When the Agr morpheme is the labeller, AgrP is projected, attracting the DP. There are no more features postulated than absolutely necessary for agreement to take place: two.

[13] In Surányi (2004b), first presented in 2000, I argue that the syntax of second, postverbal focus in Hungarian true multiple focussing constructions can be accounted for if verb movement is treated in the manner proposed here.

3.5 Strong/weak transmutations

I now address the last remaining point raised in **§3.2** as a problem for the standard account. Recall from section 2 the paradox posed by heads H_2 that intervene between a low position H_3 that a head can occupy if a head position H_1 higher than H_2 is filled by some other element.

(18) [H_1 ... [H_2 ... [H_3 ...

The paradox was that H_2 must be "weak" in the former scenario, yet, it must be "strong" when H_1 is not lexically filled: in that case H_2 must be strong to attract the head occupying H_3 on its way to H_1. The problem was illustrated schematically by some common Germanic patterns, reproduced below; however, the same complication arises in a variety of pairs of syntactic constructions and is not specific to Germanic.

(19) a. [$_{\text{CP}}$ [$_{\text{C}}$ [$_{\text{T}}$ [V] T] C] ... [$_{\text{TP}}$ SU [$_{\text{T}}$ [V] T] ... [$_{\text{VP}}$ [$_{\text{V}}$ V]]]]
b. [$_{\text{CP}}$ [$_{\text{C}}$ C] ... [$_{\text{TP}}$ SU [$_{\text{T}}$ [(V)] T] ... [$_{\text{VP}}$ [$_{\text{V}}$ V]]]]

Taking (19) for concreteness, on the standard account, (19a) involves a strong C attracting the verb, while in (19b) C is strong and is saturated by some overt (or non-overt) lexical element other than the verb. (19a) translates into the present model as having a verbal head whose C morphological component is strong. Because the verbal head has this strong C component, it will keep raising by IFM all the way up to the point where this C component becomes the label, i.e., up to CP. There is no need to stipulate that any of the intervening elements, here intervening morphemes within the verbal form, is strong. As for (19b), there is no significant divergence from the standard treatment: C is a strong free morpheme other than the verb. As can be witnessed, the paradoxical strong/weak transmutation of intervening heads does not arise.

4 Conclusion

In this paper I hope to have substantiated the following two central points: (i) the adjunction treatment of head movement is significantly more problematic that is commonly acknowledged, and (ii) there is a viable alternative approach that maintains narrow syntactic head movement, assuming that head movement is to be treated as root merger, and it should be assimilated to IFM, ultimately a consequence of the heavily cyclic nature of a strictly derivational minimalist model. The approach I have proposed eliminates the last syntactic distinctions between head and phrase movement (non-extending, only adjunction configuration), while deriving the properties that make "head movement" distinct in the descriptive sense (cf. the HMC) by assuming the computational system to tolerate only a minimal delay in Spell Out.

REFERENCES

Ackema, Peter, Ad Neeleman and Fred Weerman. 1993. Deriving functional projections. In: A. Schafer (ed.). Proceedings of NELS 23. Amherst: GSLA. 17–31.

Baker, Mark. 1988. Incorporation. Chicago: Chicago University Press.

Belletti, Adriana. 2001. Agreement projections. In: Mark Baltin and Chris Collins (eds.). The Handbook of Contemporary Syntactic Theory. Malden, MA & Oxford: Blackwell. 483–510.

Bobaljik, Jonas. 2002. A-Chains at the PF-interface: copies and “covert” movement. Natural Language and Linguistic Theory 20 : 197–267.

Boškovic, Željko. 1999. On multiple feature checking. In: Epstein & Hornstein (1999 : 159–187).

Boškovic, Željko and Daiko Takahashi. 1998. Scrambling and last resort. Linguistic Inquiry 29 : 347–366.

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

Brody, Michael. 1997. Mirror Theory. UCL Working Papers in Linguistics 9.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In: Kenneth Hale and Samuel J. Keyser (eds.). The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger. Cambridge, MA: MIT Press. 1–52.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In: Roger Martin, David Michaels and Juan Uriagereka (eds.). Step by Step. Essays in Honor of Howard Lasnik. Cambridge, MA: MIT Press. 89–155.

Chomsky, Noam. 2001. Derivation by phase. In: Michael Kenstowicz (ed.). Ken Hale: A Life in Language. Cambridge, MA: MIT Press. 1–52.

Collins, Chris. 2003. Eliminating labels. In: Epstein & Seely (2003 : 42–64).

Epstein, Samuel David, Eric Groat, Ruriko Kawashima and Hisatsugu Kitahara. 1998. A Derivational Approach to Syntactic Relations. Oxford: Oxford University Press.

Epstein, Samuel David and Norbert Hornstein (eds.). 1999. Working Minimalism. Cambridge, MA: MIT Press.

Epstein, Samuel David and Daniel Seely (eds.). 2003. Derivation and Explanation in the Minimalist Program. Cambridge, MA & Oxford: Blackwell.

Groat, Erich and John O’Neil. 1996. Spell-out at the LF-interface. In: Werner Abraham, Hoskuldur Thrainsson, C. Jan-Wouter Zwart and Samuel David Epstein (eds.). Minimal Ideas: Syntactic Studies in the Minimalist Framework. Amsterdam: John Benjamins. 113–139.

Holmberg, Anders. 2000. OV order in Finnish. In: Peter Svenonius (ed.). The Derivation of VO and OV. Amsterdam: John Benjamins. 123–152.

Hornstein, Norbert and Juan Uriagereka. 2003. Reprojections. In: Epstein & Seely (2003 : 106–132).

Koopman, Hilda and Anna Szabolcsi. 2000. Verbal Complexes. Cambridge, MA: MIT Press.

Lasnik, Howard. 1995. Verbal morphology. Syntactic structures meets the Minimalist Program. In: Hector Campos and Paula Kempchinsky (eds.). Evolution and Revolution in Linguistic Theory. Essays in honor of Carlos Otero. Washington, D.C.: Georgetown University Press. 251–275. (Reprinted in Lasnik 1999 : 97–119.)

Lasnik, Howard. 1999. Minimalist Analysis. Cambridge, MA & Oxford: Blackwell.

Mahajan, Anoop. 2000. Eliminating head-movement. The GLOW Newsletter 44 : 44–45.

Mahajan, Anoop. 2001. Word order and remnant VP movement. Ms. UCLA.

Müller, Goreon. 2003. Phrase impenetrability and *wh*-intervention. In: Arthur Stepanov, Gisbert Fanselow and Ralf Vogel (eds.). Minimality Effects in Syntax. Berlin: Mouton de Gruyter.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In: Pilar Barbosa, Paul Hagstrom, Martha McGinnis and David Pesetsky (eds.). Is the Best Good Enough? Cambridge, MA: MIT Press. 337–383.

Sportiche, Dominique. 1998. TBA. Ms., UCLA.

Starke, Michael. 2001. Move Dissolves into Merge: A theory of locality. Doctoral dissertation, University of Geneva, Geneva.

Surányi, Balázs. 2000. The left periphery in Hungarian: the division of labour between checking- and scope-driven movement. Talk at Peripheries conference. September 2000, York.

Surányi, Balázs. 2002. Multiple operator movements in Hungarian. Doctoral dissertation, LOT, Utrecht.

Surányi, Balázs. 2003. Quantifier interaction and differential scope-taking. Studies in Modern Grammar 34 : 31–70.

Surányi, Balázs. 2004a. Head movement and structure building in derivational syntax. Ms., Eötvös Loránd University, Budapest. (Paper presented at the 3rd Tools In Linguistic Theory conference, Budapest, May 2004.).

Surányi, Balázs. 2004b. The left periphery and Cyclic Spellout: the case of Hungarian. In: David Adger, Cecile de Cat and George Tsoulash (eds.). Peripheries. Syntactic Edges and their Effects. Dordrecht: Kluwer. 49–73.

Surányi, Balázs. to appear. Differential quantifier scope: Q-Raising versus Q-Feature Checking. In: Proceedings of CSSP 2003. Peter Lang.

Svenonius, Peter. 1994. C-selection as Feature-checking. Studia Linguistica 48 : 133–155.

Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut.

Uriagereka, Juan. 1999. Multiple Spell-Out. In: Epstein & Hornstein (1999 : 251–282).

Zwart, C. Jan-Wouter. 1992. Matching. In: Dicky Gilbers and Sietse Looyenga (eds.). Language and Cognition 2. Yearbook 1992 of the Research Group for Linguistic Theory and Knowledge Representation of the University of Groningen. Gronigen: University of Groningen. 349–361.

Zwart, C. Jan-Wouter. 2003. Agreement as sisterhood. Talk at the Comparative Germanic Syntax Workshop 18, 2003, Durham.