Cyclicity and the Approach the Probe principle

Phil Branigan

Memorial University

March 2013

Abstract. Consideration of the problematic relationship between the No Tampering Condition (NTC) and ‘tucking in’ and ‘feature inheritance’ constructions forces an interpretation of NTC in which only the complement of a probe is untamperable. The resulting merge space available for displaced phrases is constrained by the Approach the Probe Principle (APP), a particular locality constraint which brings a phrase as close as possible to the head which triggers movement. Empirical advantages of the APP are demonstrated in a range of constructions. The word order reversals which APP-constrained movement makes available are identified in French causatives, epistemic datives, and other applicative structures.

1. Ensuring the cyclic application of Merge

The idea that transformation operations must observe strict cyclity has been a mainstay of generative grammar since Chomsky (1965). This general claim has been applied particularly rigorously in recent minimalist analysis, and is embodied in the Extension Condition (Chomsky, 1995) and, more recently, the No Tampering Condition\(^1\) (NTC) (Chomsky, 2007, 2008):

"A natural requirement for efficient computation is a ‘no tampering condition’ NTC: Merge of X and Y leaves the two [syntactic objects] unchanged. If so, then Merge of X and Y can be taken to yield the set X, Y, the simplest possibility worth considering."

(Chomsky, 2008)

\(^1\)Watanabe (1995) actually anticipates the NTC; he observes that the Extension Condition can be derived if we require that movement operations should not disrupt the set of c-command relations already built into a phrase marker.
Given the NTC, Chomsky continues, the characterisation of syntactic movement can be minimal, perhaps even ‘perfect’. The head of the root may include an edge feature, which permits the root to undergo a Merge operation. (Languages may vary, apparently, in whether an edge feature remains active past the first Merge operation.) As such, when the semantics require that movement should occur, any phrase—internal or external—may Merge with the root in the manner permitted by the NTC. (Movement which falls under the EPP is slightly different, and not derived so directly.) Thus, for example, wh-movement occurs because C has an edge feature, and therefore a wh-phrase may Merge as a specifier to the root CP.

In short, the NTC plays two roles: one empirical, and one metatheoretical. It precludes counter-cyclic operations which are widely accepted as impossible, and simplifies the formulation of the basic operation of (narrow) syntactic movement. Both are evidently imperative elements of any interesting model.

Although the intuitive basis for the NTC is quite clear, the empirical grounds for it are less so. The NTC is immediately undermined, in fact, by two apparently widespread phenomena: ‘tucking-in’ operations and the sort of counter-cyclic movement which Chomsky (2008) and Richards (2007) treat in terms of ‘feature inheritance’. The former phenomenon is well established in Richard’s (1997) study of Bulgarian multiple wh-movement (and other similar constructions).

(1) Koj kogo vižda?
   who whom saw
   ‘Who saw whom?’ (Bulgarian)

As Richards shows, the derivation of questions like (1) involves two wh-movement operations. The first one displaces koj to make it a specifier for CP. The second then moves kogo and merges it as a lower specifier in the same CP. The latter operation clearly violates the NTC, at least if we interpret the NTC with its obvious intuitive sense.

The second class of problematic derivations emerges from Postal’s (1974) ‘raising to object’ cases, such as (2).
We believe John sincerely to be intelligent.

As Postal observed, the orientation of the adverb *sincerely* in such examples is toward the verb in the matrix clause, despite its position to the right of the embedded subject *John*. This word order therefore seems to show that *John* has raised out of its own clause into the matrix verb phrase. Lasnik (2001) offers a minimalist account of this type of data in which the subject of the infinitival complement undergoes NP-movement to a specifier position in the matrix VP. (More accurately, Lasnik proposes the subject raises to an AgrP projection which translates pretty directly into VP in more recently refined models, but for the purposes of this discussion, the difference does not seem to matter much.)

For Lasnik, this necessitates an EPP feature in V (or AgrP) which appears only when v assigns accusative Case. Satisfaction of the EPP feature then gives rise to the matrix vP structure in (3).

As Lasnik observes, this analysis works only with some connection between accusative Case features and the EPP feature in V. What is more, given Pesetsky’s (1992) “Agent/ECM correlation”, the connection between the properties of v and the subject of an ECM complement is still stronger
than Lasnik suggests. This correlation involves the fact that verbs with an agentive subject cannot normally participate in ECM. Thus, while (4a) is acceptable, (4b) is not.

(4)  a. Mary discovered Bill to have read the book.
    b. *John affirmed Mary to have entered the room.

In current terms, this pattern must concern the relationship between \( v \)—the head which is responsible for the subject \( \theta \)-role—and the ECM subject. Evidently, \( v \) may bear unvalued \( \phi/\text{Case} \) features in an ECM context only when it has specific thematic properties. The properties of the lower, lexical \( V \) are essentially irrelevant to understanding the correlation.

Chomsky (2008) points out that Lasnik’s analysis can be improved if it is the features of \( v \) which are responsible for the displacement of the downstairs subject. Then \( V \) itself need not coincidentally bear an EPP feature in exactly the circumstances in which \( v \) must assign accusative Case. But implementation of this idea requires a mechanism to enable a counter-cyclic Merge operation. Chomsky proposes that \( v \) may transfer its EPP feature (and others) to \( V \), which then attracts the lower subject to become its own specifier. But this means that movement to the specifier position of \( VP \) cannot take place until after \( v \) is present in the structure. Again, the derivation clearly violates the NTC, unless some artful loophole can be devised to evade this conclusion.

Of course, it is always possible that a specific and elegant way to interpret the NTC might be be found which would render the derivation of (3) unproblematic. But absent any such interpretation, the conclusion to be drawn at present seems to be that the NTC is too restrictive. What is more, the two constructions which undermine the NTC do so in different directions. *Tucking in* operations involve an NTC violation in which the structure higher than the probe is tampered with; internal object shift requires tampering with something inside the first projection of the probe. And since, under Chomsky’s approach, the mechanism which leads to movement into a lower specifier position is accessed only for A-movement, the prospects for unifying these two problematic cases are not encouraging.
A third empirical problem for the NTC (as formulated above) concerns the role played by head-movement in the derivation. Chomsky takes the explicit position that head-movement must not occur in the narrow syntactic derivation, given the NTC (or the Extension Condition), and this conclusion seems inescapable. But Branigan (2010) provides a number of case studies in which either head-movement or phrasal movement may occur to satisfy the featural requirements of a probe. What is more, phrasal movement which alternates with head-movement (and sometimes follows head-movement) can involve the sort of movement which everyone would accept as a narrow syntactic operation. Thus, for example, in the embedded Norwegian question in (5a), the wh-phrase must be followed by the *som complementiser, which is forbidden in (5b).

(5)  

(a) Vi vet hvem *(som) snakker med Marit.  
   we know who talks with Mary

(b) Vi vet hvem (*som) Marit snakker med.  
   we know who Marit talks with

(Taraldsen, 1986)

Branigan shows that this pattern is derived as follows. The complementiser *som originates in the Fin position, and the subject—whether it is a wh-phrase or not—is attracted to become the specifier to Fin, as in (6).

(6)  

\[
\text{ForceP} \\
\text{Force} \quad \text{FinP} \\
\text{subject} \quad \text{Fin'} \\
\text{Fin} \quad \text{TP} \\
\text{som}
\]

Force must attract the closest target with a “force-marking” feature, which is true of both Fin and wh-phrases. When the subject is a wh-phrase, therefore, Force attracts it directly, and *som remains in situ. This produces the structure in (5a), wherein *som is pronounced. When the subject is
not a wh-phrase, however, Fin is the closest target for Force, and must therefore adjoin to Force. (Subsequent attraction of a more distant force-marking target ensures wh-movement of the non-subject wh-phrase in (5b) .) But *som* is not pronounceable in Norwegian (unlike Swedish) when adjoined to Force, so it cannot be pronounced in (5b) .

While we cannot exclude the possibility that some head-movement operations take place outside the narrow syntactic derivation, cases like this (and others discussed in Branigan (2010)) indicate that some head-movement is as narrowly syntactic as wh-movement. And as such, the narrow syntax must not be so restrictive as to exclude it, as Chomsky’s formulation of the NTC does.

In fact, under the most transparent interpretations, the NTC is violated even in contexts where a phrase simply moves to become the unique specifier of a probe, such as English wh-movement. If movement takes place to satisfy a requirement of a probe—as virtually any minimalist analysis will assume—then movement will never leave the original phrase marker containing the probe unchanged, because the motive for movement is to alter the properties of the probe itself.

To sum up these general observations, the NTC, which provides a strong conceptual basis for cyclic application of movement operations, appears empirically indefensible in Chomsky’s formulation, because it blocks a range of attested operations in which phrases raises to positions lower than the root, and because it ensures that movement of any sort cannot be driven directly by the features of the head of the root phrase.

On the other hand, it remains clear that many varieties of countercyclic operations must be excluded, so the NTC cannot be simply abandoned. Nor would we wish to abandon Chomsky’s conception of movement as a ‘perfect’ operation, however perfection is operationalised.² The best approach would appear to be to revise the NTC in a way which preserves the necessary empirical coverage and which adheres to the intuitive foundation which Chomsky establishes for this constraint. The formulation in (7) appears to qualify as a viable alternative.

²The Move-α operation of Government-and-Binding theory is arguable a perfect theory of movement, as well, because nothing beyond the fact of movement is required of its characterisation. The *provocation* operation in Branigan (2010) is comparable.
(7) *No Tampering Condition* (revised)

Given $\pi$, $X$, $Y$ and $Z$, where $\pi$ is the head of $Y$ and $Z$ is the complement of $\pi$, Merge of $X$ and $Y$ leaves $X$ and $Z$ unchanged.

In other words, rather than insisting that nothing be changed in the phrase which initiates movement, (7) allows the initiating probe to be changed, but precludes any change in the remaining portion of the phrase marker (the complement). Like Chomsky’s original formulation, (7) is a plausible third factor constraint, which minimizes derivational complexity.

The fact that lower specifiers to a probe and specifiers to a lower complement are very different types of Merge sites now ceases to be a problem. Both are allowed by the formulation in (7) without complication, simply because both positions are outside the portion of the phrase which is constrained by this general principle.

The immediate consequence of the NTC (revised)—henceforth NTC—is that the set of acceptable Merge sites now include every position outside of the original complement of the probe. This will include a new specifier for the probe, as always, but now it will include lower ‘tucked in’ specifiers, as well. It includes a new specifier for the complement, too, because the original complement is unaffected by being paired structurally with another phrase. The projection level of the original complement is altered from maximal to non-maximal, but the projection level is defined positionally and not in terms of content. And in the case of head movement, adjunction of an $X^0$ category to the probe will also be allowed, since this also leaves the complement unaffected.

(8) acceptable Merge sites in $X$ for $Y$

\[
\begin{align*}
a. & \quad X \quad Y \quad \pi \quad Z \\
b. & \quad X \quad Y \quad \pi \quad Z \\
c. & \quad X \quad Y \quad \pi \quad Z
\end{align*}
\]
But now a new question arises. If multiple Merge sites are available in principle, then how does the derivation determine which Merge site to utilise for any particular movement operation. In the case of head-movement, the solution is apparent, as the head-adjunction Merge site is the only appropriate one. (I return to this point below.) But for phrasal movement, the choice of a specifier for the probe or a specifier for the complement must still be made on some basis. The solution to this problem constitutes the remainder of this paper.

2. Approach the Probe

In Richards’ treatment of Slavic multiple wh-movement, tucking-in operations are driven and constrained by two distinct grammatical principles. The first is the general minimalist premise that syntactic movement—of phrases, at least—occurs in order to provide a specifier for the triggering head (Chomsky, 1995). The second is a novel principle that a phrase must move to a position as close as possible to the probe which attracts it. Richards does not provide a label for this latter constraint. In Richards (2001), he suggests that the effect follows from the Shortest Move principle, but he chooses not to isolate any particular precise implementation of Shortest Move. And since the term “Shortest Move” has been used to mean a number of things in the literature, it will be convenient to have a term to isolate the idea that phrases land close to the probe; let us refer to this as the Approach the Probe Principle (APP). Crucially, the APP must constrain movement operations rather than resulting representations, because the effect of tucking-in is always to leave the first specifier in a more distant position from the head than it was before. What is more, the APP does not itself trigger movement. It merely constrains the choice of landing site when a higher or lower specifier position within the same phrase are available options.

A formal characterisation of the APP requires that we pay close attention to the relevant structures, and in particular, to how specifiers are characterised. The minimalist literature offers two competing views: specifiers are sometimes defined as sisters to intermediate projections of a head (Chomsky and Lasnik, 1991); in other works, they are treated as phrases adjoined to a maximal projection (Kayne, 1995). But the Kaynean approach does not allow for the possibility of multiple
specifiers, and since I assume, following Richards, that these are real, it will be necessary to adopt the more conservative view that specifiers are not adjoined phrases.

Relative closeness can be defined in a few steps. We first define the relation of *immediate connection*:

(9) **Immediate connection** (definition)

Given two nodes, \(x\) and \(y\), \(x\) is immediately connected to \(y\) iff \(x\) immediately dominates \(y\) or \(y\) immediately dominates \(x\).

I take dominance to be a reflexive relation so that every node is also immediately connected to itself.

A less local notion of connection can be defined, as well.

(10) **Connection** (definition)

\(x\) is connected to \(y\) iff:

a. \(x\) is immediately connected to \(y\), or

b. \(x\) is immediately connected to \(z\) and \(z\) is connected to \(y\).

This recursive definition obviously connects every node in a phrase marker to every other node, which is of little value. But it also allows us to define a *minimal connection set*:

(11) **Minimal Connection set (MC)**:

The minimal connection set (MC) for \(x\) and \(y\) is the smallest set of nodes \(\{z: \text{either } z \text{ is connected to } x \text{ or } z \text{ is connected to } y\}\).

Where \(x = y\), the minimal connection set will be \(x\). Otherwise, the minimal connection set will consist of \(x\), \(y\), and the smallest group of nodes which must be traversed to move from \(x\) to \(y\) either upwards or downwards in a tree.

Now the relative distance between a probe and two competing potential merge sites can be
Cyclicity and the APP

determined as in (12).

\[ (12) \quad \text{Closer} \quad (definition) \]

\[ \mu \text{ is closer than } \mu' \text{ to } \pi \text{ iff } |MC(\pi, \mu)| < |MC(\pi, \mu')|. \]

The initial motivation for the APP comes from the relatively rare situation in which a single head takes on multiple specifiers, and for this case, it is evident that a principle like this is necessary. But once we include the APP as a module in the theory, it becomes evident quite quickly that it has other explanatory applications. The APP has force even in the analysis of phrasal movement which forms a single specifier, such as English wh-movement.

In embedded wh-questions in English, for example, wh-movement occurs because the C (=Force) probe attracts a single wh-phrase, and the landing site is the specifier of CP. In Chomsky’s (2008) terms, C bears the edge feature, so CP will Merge with something. Given the NTC (revised), there are three positions into which a displaced category could be Merged; two positions, if wh-movement is phrasal movement. A wh-movement could create a specifier in CP, or it could create a specifier in the Complement of C (possibly FinP).\(^3\)

Given the APP, the CP must be selected as a Merge site, because it is closer to C than the lower specifier position is. Specifically, suppose the structure in which wh-movement is to occur to be (13), where C is the probe and CP and TP are both potential Merge sites.

\[ (13) \quad \text{\begin{tikzpicture}[baseline=(current bounding box.center)]
    \node (C) at (0,0) {C};
    \node (CP) at (0,1) {CP};
    \node (TP) at (1,0) {TP};
    \draw (C) -- (CP);
    \draw (C) -- (TP);
    \end{tikzpicture} . . . \text{wh-phrase} . . . \]

The minimal connection set (MC) for C and CP will be \{C, CP\}. MC(C, TP) will be \{C, CP, TP\}. As \(|\{C, CP\}| < |\{C, CP, TP\}|\), CP is closer to C than TP is.

\(^3\)The effects are indistinguishable if a provocative feature rather than an edge feature is what drives movement.
In a Slavic multiple wh-question like (14), the same reasoning will apply throughout the movement of the first wh-phrase.

(14) Koj kogo vižda?
    who whom saw
   ‘Who saw whom?’

When C triggers movement of the second wh-phrase, the initial structure is (15), and a merge site for *kogo* must now be determined.

(15)

There are now three potential Merge sites to consider: CP, C', and TP, so the relative distance of each from C must be calculated.

(16) a. $MC(C, CP) = \{C, C', CP\}; |MC(C, CP)| = 3.$

b. $MC(C, C') = \{C, C'\}; |MC(C, C')| = 2.$

c. $MC(C, TP) = \{C, C', TP\}; |MC(C, TP)| = 3.$

Since $2 < 3$, C' is closer to C than either CP or TP is. Tucking-in is therefore the only way to satisfy the APP.

While Chomsky’s Extension Condition/NTC would always force movement to the specifier position for the probe, the APP is less absolute. Because it only requires the closest possible position to the probe, a more distant position will become available if the absolute closest position is an impossible merge site for some reason. For phrasal movement, then, if circumstances are such that a specifier for a triggering probe will not be tolerated, then we should expect the APP to
be satisfied by merge into the specifier position of the probe’s complement.

Returning to the analysis of Postal’s ECM examples like (17a), let us consider how the position of the lower subject might be ensured given the revised NTC. The structure of the matrix vP prior to movement of John can be assumed to be (17b).

(17) a. We believe John sincerely to be intelligent.

With Chomsky, we may identify the matrix v as a probe triggering movement. The NTC then restricts the set of available Merge sites to three nodes in the matrix clause: vP, v’, and VP. But both vP and v’ can be ignored by the APP if a Merge site higher than the v probe always produces an unsuccessful derivation. And for some reason, English does not allow vP to contain a specifier at the interfaces. (Of course, the experiencer argument of v is merged as a specifier to vP, but this is acceptable because it is always moved out later.) A possible explanation for this apparent fact might be phase theoretic; Chomsky (2008) proposes that φ/Case features must be interpreted as soon as feature valuation takes place, and this requires that a Case-marked DP remain in the interpreted domain of a Case-marking probe. Given the phasal status of Case-marking v in (17), the implication is that John cannot raise to the edge of the matrix vP, leaving only VP as an acceptable Merge site. Whether or not this is the best account for the restriction on Merge sites in English.
object shift, what matters for the matter at hand is simply that the vP and v′ positions appear to be excluded, making the matrix VP the closest available Merge site by default.

In short, nothing needs to be stipulated to ensure the subject in English ECM complements will always Merge as a specifier of the matrix VP.

(It follows that Chomsky’s feature inheritance operation is unnecessary for this particular construction. I leave it as an open question whether feature inheritance can be persuasively motivated on the basis of other constructions.)

3. Argument Inversion and the APP

3.1 Deriving inversion

The proposition that movement can target a lower specifier position when necessary, and subject to the APP, makes a clear prediction in certain complex situations. While the tucking-in operation preserves the initial relative ordering of the displaced specifiers, movement of multiple specifiers to a position below a probe should have the opposite effect, and produce a mirror image relative ordering. Consider a structure (18), where F is a movement-triggering probe with two particular marked properties. One is that F cannot tolerate a specifier. The second is that F may be valued multiple times, triggering movement each time. F is like English v in the first respect, and like Bulgarian C in the other.

(18)

Let α and β both be phrasal categories which match the unvalued features of the F probe. In this situation, F will displace α, the closer goal, and then β in turn. When the derivation must integrate α, it cannot be merged as a specifier for F, so by the APP, it must take the next best option and merge as a specifier for G, generating (19).
When $\beta$ must be merged, the same logic applies, but as $G$ already contains a specifier at that point, $\beta$ must either be tucked-in or merged outside $\alpha$. The Merge site is now determined as follows. The NTC restricts possible Merge sites for $\beta$ in (19) to FP and GP. If FP is ruled out, then GP remains as the only option. ($G'$ will be excluded by both the NTC and APP; NTC ensures that the complement of the F probe is inviolable, while APP prefers Merge sites with a smaller MC—here $|MC(F, GP)| = 2$, while $|MC(F, G')| = 3$.\footnote{The APP would play a role in this context if there was evidence that the NTC applies relative to the initial structure for a probe. In other words, if the complement prior to any movement triggered by the probe were the structural component which the NTC seeks to preserve.}

Thus, in this situation, (19) will give rise to (20), in which the base ordering of $\alpha$ and $\beta$ has been derivationally inverted.

### 3.2 Epistemic datives

This prediction is validated by a range of different constructions in vP in a number of languages. One such is the French epistemic dative construction documented by Ruwet (1982). Some examples (from Ruwet, translations mine) appear in (21).
This construction has a number of intriguing characteristics. Most immediately, the construction provides a meaning of possession, in some sense, which unites the indirect object and the direct object, but in a context in which the verb itself is separated from the possession relationship. This can happen, as in the (21) examples, because the verb itself lacks any possessive sense inherently, but it can also occur with verbs which normally impart a sense of a possessive relationship, but which become divorced from that relationship by the nuance of context, as in the (22) examples.

(22)  

(a) On lui attribue beaucoup de bonnes fortunes.  
one him/DAT attributes lots of good luck  
‘People suppose that he has a lot of luck.’

(b) Les Alliés prêtaient à Napoléon l’intention de battre en retraite.  
the Allies loaned to Napoleon the intention of beat in retreat  
‘The Allies supposed that Napoleon had the intention to beat a retreat.’

The same effect is actually marginally possible with some English double object verbs (Branigan, 1992), like (23).

(23)  

(a) We grant Brad a certain naive charm.

(b) I give him 10 years, no more.

In (23a), although the verb *grant* can be used as a verb of “giving”, in this context it takes on a purely epistemic meaning (“I grant that P”), and the sense that *Tom* and his *charm* are in a
possession relationship comes from the syntactic context more than from the verb. Similarly, in (23b), *give* here means something more like “predict” and what is predicted is the proposition that the person in question will have 10 years of something.

What is particularly significant about the French epistemic datives is that the possessor behaves as a “subject” with respect to binding conditions, as can be seen in (24).

   ‘I believe them to have respect for each other.’

   b. *Ils me croient du respect l’un pour l’autre.
   ‘I believe them to have respect for each other.’

   c. Je leur prête de mauvaises intentions l’un envers l’autre.
   ‘I suppose them to have bad intentions towards each other.’

   d. *Ils me prêtent de mauvaises intentions l’un envers l’autre.
   ‘I suppose them to have bad intentions towards each other.’

Under standard assumptions, the fact that the dative “possessor” can bind the reciprocal *l’un l’autre* shows that the dative c-commands the accusative phrase. The fact that the possessor is the only acceptable antecedant shows that it is the only c-commanding antecedent within a certain clause-like domain. In that respect, these datives behave much like the subject of a small clause complement: (25).

   ‘Pierre and Paul believe Marie amourous the one of the other’

   ‘Pierre and Paul believe Marie in love with each other.’

The binding properties of epistemic datives lead inexorably to the conclusion that the structure of the verb phrase in these examples must be something like (26), where E might be either an Applicative head, or some other less familiar category which can be selected by the verbs in question.
EP evidently constitutes a binding domain.

In the structure (26), we should not expect A-movement of the possessum to a position outside vP ever to be possible, because there are two higher nominals which should intervene. Even if the subject were removed, as in a passive version of (26), the presence of the possessor should be sufficient to block A-movement of the possessum. And this expectation initially seems to be confirmed in examples like (27).

(27)  
(a) *Une maîtresse dans chaque port lui était cru.  
a mistress in every port him/DAT was believed  
‘He was believed to have a mistress in every port.’  

(b) *Beaucoup de charme lui est trouvé (par tout le monde).  
lots of charm him/DAT is found by all the world  
He is found to have lots of charm (by everyone).

But Ruwet attributes to Paul Postal the observation that the passives of some epistemic datives are bad even when no movement occurs. So (27a) is not improved if an expletive is used in the subject position. Compare the status of (28a) with the grammatical result of using an expletive in a passive formed from a regular double object verb.

(28)  
(a) *Il lui est cru une maîtresse dans chaque port.  
it him/DAT is believed a mistress in every port  

(b) Il lui a été prêté beaucoup de livres.  
it him/DAT has been loaned lots of books
‘There have been loaned lots of books to him.’

Given the status of (28a), the ungrammaticality of the (27) examples need not be attributed to the A-movement of the possessum.

When passive structures are formed from other verbs (especially those which have as their core meaning some notion of transfer of possession), it turns out that the epistemic dative possessums can undergo A-movement.

(29) a. Une issue fatale est prévue à cette maladie (par le médecin).
   an outcome fatal is predicted to this illness by the doctor
   ‘This illness is predicted to have a fatal outcome.’

   b. Beaucoup de bonnes fortunes lui sont attribuées à tort.
   lots of good luck him/DAT are attributed by mistake
   ‘He is thought to have a lot of luck wrongfully.’

   c. ?Une maîtresse dans chaque port lui est souvent imaginée.
   a mistress in every port him/DAT is often imagined
   ‘He is often imagined to have a mistress in every port.’

We face a seemingly paradoxical situation. Given the binding facts, the possessor must be higher than the possessum. Given the A-movement facts, the possessum must be the most accessible goal for a T probe. The only way to resolve the paradox is by allowing the derivation to reverse the order of possessor and possessum midway. The possessum must be able to raise past the possessor before T enters the derivation.

Given the structure (26), the necessary inversion results if \( v \) has the properties of F in (18). \( v \) presumably does not accept A-movement to its edge in French, as it does not in English. Once again, we may attribute this pattern to the phasal status of vP, together with the principle that Case features assigned by \( v \) must be interpreted within the phasal domain. We may account for the Case-marking of both dative and accusative objects by supposing that \( v \) is able to Agree and assign structural Case twice, as argued by Bobaljik and Branigan (2006). Assuming that Case model, \( v \) must assign the marked (dative) Case first, and the unmarked (accusative) Case second. Person-
Case Constraint effects then reflect feature conflicts located within a single active probe. And just as in English ECM (where \( v \) assigns structural Case only once), the \( \phi/\text{Case} \) features of \( v \) trigger movement of their target DP.

Given these properties of \( v \) in French epistemic dative structures, the structure in ((26)) must be transformed into (30) in order to satisfy both the APP and the featural requirements of its component parts. The binding relations are evidently established within EP before the \( v \) probe assigns Case.

\[
(30) \quad \begin{array}{c}
\text{vP} \\
\text{subject} \\
v \\
\text{VP} \\
\text{possessor}_i \\
\text{possessum}_j \\
\text{V} \\
\text{EP} \\
\text{t}_i \\
\text{E} \\
\text{t}_j
\end{array}
\]

After (30) has been generated, however, the possessum is actually closer to T than the possessor, so in a passive vP, T can attract the possessor without worrying about any dative intervener.

3.3 French faire causatives

A similar situation arises with faire (and laisser) causatives in French, where the surface word order in transitive causatives has accusative objects preceding dative “subjects”.

\[
(31) \quad \text{On fera discuter sa thèse à nos collègues.} \\
\text{one make-FUT discuss his thesis to our colleagues} \\
\text{‘We will have our colleagues discuss her thesis.’}
\]

---

5See also Adger and Harbour for development of this approach to the Person-Case Constraint.
The downstairs verb precedes both subject and object, of course, but the position of the verb is something of a red herring. Guasti (1993) shows that the lower verb raises up quite high into the matrix clause in this type of causative, and that it can even precede matrix subject-oriented quantifiers. Therefore I assume that the lower verb has raised up to the matrix v position.\(^6\)

Despite the surface word order, the dative argument in transitive faire causatives retains subject-like properties. They bear the external \(\theta\)-role normally assigned by \(v\) (of the downstairs verb) to its specifier. And like accusative downstairs subjects in intransitive causatives, dative subjects can bind reflexive \(se\).

(32) a. Il a fait se rendre Jean à Paris.
he has made SELF render Jean to Paris
‘He made Jean go to Paris.’

b. Marie a fait s’acheter ces livres à Jean.
Marie has made SELF-buy these books to Jean
‘Marie made Jean buy himself these books’ (Rouveret and Vergnaud, 1980)

As with the epistemic datives, we must conclude (uncontroversially), that at some point in the derivation, the dative subject c-commands the rest of its downstairs verb phrase within a particular domain. Thus, minimally, the structure of e.g. (32b) should be (33) at an early stage.

\(^6\)XXX (in progress) argue that in this construction, as in several Germanic ones, the lower verb is displaced by the multiple head-movement process initially proposed by Collins (2002).
The inverted word order in the transitive causatives now emerges from essentially the same premises as in the epistemic dative constructions. Matrix \(v\) cannot accommodate A-movement, which is triggered twice when it assigns both dative and accusative Cases (in that order). Thus the dative subject is displaced from its original position as the specifier of the lower vP to become the specifier of the higher V. And the accusative object, which is displaced second, satisfies the APP and NTC by merging as a still higher specifier in matrix VP. The resulting structure is (34).\(^7\)

\(^7\)The structure shown does not incorporate the movement of the lower verb, which I have set aside.
The *faire* causative data presented by Rouveret and Vergnaud enables an additional argument which shows that movement of a goal is to a position lower than the probe. The relevant data begins in (35)–(36).

(35)  

(a) Jean fera aller Marie à Paris.  
Jean make-FUT go Marie to Paris  
‘Jean will make Marie go to Paris.’

(b) Jean y fera aller Marie.  
Jean there make-FUT go Marie  
‘Jean will make her go there.’

(c) Jean fera se rendre Marie à Paris.  
Jean make-FUT SELF render Marie to Paris  
‘Jean will make Marie go to Paris.’

(d) *Jean y fera se rendre Marie.  
Jean there make-FUT SELF render Marie  
‘Jean will make her go there.’

(36)  

(a) Marie a fait acheter ces livres à Jean à Paris.  
Marie has made buy these books to Jean to Paris  
‘Marie made Jean buy these books in Paris.’

(b) Marie les a fait acheter à Jean à Paris.  
Marie them has made buy to Jean to Paris  
‘Marie made Jean buy them in Paris.’

(c) Marie a fait y acheter ces livres à Jean.  
Marie has made there buy these books to Jean  
‘Marie made Jean buy these books there.’

(d) *Marie les a fait y acheter à Jean.  
Marie them has made there buy to Jean  
‘Marie made Jean buy them there.’

(37)  

(a) *Sa mère est arrivée à les faire s’acheter à la jeune fille.  
her mother is arrived to them make SELF buy to the young girl  
‘Her mother managed to get her to buy them for herself.’ (Kayne, 1975, p. 429)

(b) *La manque de scrupule les a fait s’adjuger aux députés.  
the lack of scruples them has made SELF award to the deputies  
‘Lack of scruples made the deputies appropriate them for themselves.’
What these contrasts show is that movement of clitic pronouns in this construction is subject to a intervention condition which is sensitive to clitic status of nominals. While clitics originating as complements in the downstairs verb phrase may raise into the matrix verb phrase (and beyond), they may not do so if there is another downstairs clitic which remains there. So in (35b) and (36b), \( y \) and \( les \), respectively, are permitted to appear as clitics in the root clause. But in (35d) the presence of \( se \) in the lower verb phrase blocks movement of \( y \) upwards, and \( y \) blocks movement of \( les \) in (36d).

A clitic attached to the downstairs verb does not block upwards movement of a clitic subject, however. And it makes no difference whether this is an accusative subject or a dative subject.

(38)  

\[
\begin{align*}
(38) \quad & a. \text{Cela les a fait se poser des questions.} \\
& \text{that them has made SELF pose of questions} \\
& \text{‘That made them ask themselves some questions.’}
\end{align*}
\]

\[
\begin{align*}
& b. \text{Elle leur fera se laver les mains.} \\
& \text{she them/DAT make-FUT SELF wash the hands} \\
& \text{‘She will make them wash their hands.} \quad \text{(Kayne, 1975, p. 298)}
\end{align*}
\]

This is as we would expect, if the downstairs clitic is attached below the specifier of vP, as in the ((33)) tree. An intervening clitic can intervene, after all, only if it is closer to the attracting probe. At the same time, an initial position higher than the lower \( v \) explains why clitics originating as downstairs subjects cannot be left downstairs. Clitics—at least those of the Romance variety—cannot lower to find a host. Since the clitic subjects in (38) must attach to some verb cluster, and since the only one which it can reach by raising is in the matrix clause, that is where it goes.

(39)  

\[
\begin{align*}
(39) \quad & *\text{Jean fera lui acheter ce livre.} \\
& \text{Jean make-FUT him/DAT buy this book} \\
& \text{‘Jean will make him buy this book.’}
\end{align*}
\]

But now consider exactly what is involved in the intervention effect in (36d). In order for the \( y \) clitic (attached to \( v \)) to prevent an upstairs probe from attracting \( les \), \( y \) must be situated higher
than *les* at the point in the derivation where this attraction might take place. In other words, *les* cannot move to the left too soon. Suppose, contrary to what I have been arguing, that objects in a *faire* causative were displaced to the left by forces internal to the downstairs complement. In other words, suppose there were some head F in ((36c)) which attracted *ces livres* past the subject and merged it as a specifier in FP, as in (40). (The movement of the lower verb to a higher position will still be a later operation.)

(40) \[
\text{Marie a fait} \ [\text{FP } \text{ces livres } F \ [\text{vP } \text{`a Jean } v \ y \ acheter ] \ldots]]
\]

This type of operation might account for the relative word order of subject and object without the APP and the idea that a displaced goal might be merged below its probe. But if F could displace an accusative goal in (40), then it should be equally able to displace an accusative goal in ((36d)), to produce an intermediate structure like (41).

(41) \[
\text{Marie a fait} \ [\text{FP } \text{les } F \ [\text{vP } \text{`a Jean } v \ y \ acheter ] \ldots]]
\]

But now the accusative clitic is located higher than the *y* adjoined to the downstairs *v*, and no intervention effect should be found. Since this is the wrong result, the conclusion must be drawn that there is no possibility of movement of the accusative object internal to the complement of *faire* before the upstairs *v* triggers movement. In other words, both ((36c)) and ((36d)) must leave the accusative object in its base position until the upstairs *v* assigns accusative Case. The structure for each will be (42).
When the upper v displaces *ces livres* in this structure, it merges as an outer specifier in the matrix VP. Because *les* is a clitic, however, displacement of *les* by upper v is blocked in (42) because the y clitic is closer, so the derivation of (36d) fails, as it should. Note that the derivation of such structures also fails if the matrix v does not engage with the lower object at all, even if it is attached as a clitic to the lower verb, as (43) shows.

(43) *Sa mère est arrivée à faire se les acheter à la jeune fille.*
her mother is arrived to make SELF them buy to the young girl

(Kayne, 1975)

Not only must v assign accusative Case in this context, but it must do so in a way which forces the clitic object to Merge at the matrix VP.

In these two French constructions, the reversal of underlying word order results because the v probe always triggers movement of both its dative goal and its accusative goal. If only the dative were displaced, no inverted order would be generated. But there appear as well to be languages in which the French pattern is exploited as one of two options. As described by Baker et al. (2012), Lubukusu causatives are much like French (verb incorporation aside), but are slightly less
restrictive. Alongside the theme-agent order in (44a), the agent-theme order in (44b) is possible (except where blocked by the Phase Edge Prominence Constraint).

\[(44)\]  
\[\begin{align*} 
\text{a. } & \text{Wafula a-nyw-esy-a} \quad \text{Wekesa ka-ma-lwa.} \\
& \text{Wafula Agr-Tns-drink-Caus-fv Wekesa Agr-Agr-beer}
\end{align*}\]

\[\begin{align*} 
\text{b. } & \text{Wafula a-nyw-esy-a} \quad \text{ka-ma-lwa Wekesa.} \\
& \text{Wafula Agr-Tns-drink-Caus-fv Agr-Agr-beer Wekesa}
\end{align*}\]

Baker et al. (2012) show that each order in (44) can feed A-movement operations, so that a passive causative can be formed with either the causee or the downstairs object as the final subject of the sentence. This means that the derivation must include some mechanism for optionally reversing the underlying order within the causative complement. Baker et al simply stipulate an optional A-movement internal to VP which violates Closest Move/Superiority in order to account for the reversed word order in (44b), but the APP offers a better approach. We need only Suppose the structure of both of these clauses includes the substructure (45) after the causative suffix is introduced.

\[(45)\]

The usual incorporation operations will raise the lower verb up to the causative suffix (Baker, 1988); our concern is with the objects. If the matrix \(v\) now simply displaces Wekesa in (45), then the APP will ensure that the causee will end up as the specifier in the upstairs VP. If the matrix
$v$ has the option of displacing *kamalwa* as well, then the object will remerge as a higher specifier in the upstairs VP, and the order in (44b) will result. In both cases, the uppermost specifier in VP should be accessible for subsequent A-movement, if attracted by a still higher probe.

### 3.4 “Leapfrogging”

The inversion mechanism which the APP makes available, and which allows Lubukusu passives to displace the lower object in a causative construction, resembles what McGinnis (1998) calls “leapfrogging”. And some of McGinnis’s own cases appear to show that it occurs in a monoclausal structures as well. (To this point, the focus has been on more complex structures, simply because there are more analytic possibilities in simple double object verb phrases, so it is more difficult to pin down what the base structure must be.) In Albanian, for example, standard quantifier-pronoun binding tests show that indirect objects assymetrically c-command direct objects. McGinnis illustrates this with the (46) data, from Massey (1992).

(46) a. Agimi ia dha secilit djalë pagën e tij.  
    Agim-NOM CL give each boy-DAT pay-ACC his  
    ‘Agim gave to each boy his pay.’

    b. *Agimi ia ktheu secilin liber autorit të tij.  
    Agim-NOM CL return each book-ACC author-DAT its  
    ‘Agim returned to its author each book.’

In passive sentences, however, where the direct object is the target of agreement, the reverse is true.

(47) a. Secili libër iu kthye autorit të tij.  
    each book-NOM CL kthye-NACT author-DAT its  
    ‘Each book was returned to its author.’

    b. *Secilit djalë iu dha paga i tij.  
    each boy-DAT CL give-NACT pay-NOM his  
    ‘Each boy was given his pay.’
The reversal of binding possibilities is to be expected if Albanian double object passives, like Lubukusu and French causatives, require \( v \) to displace both the dative and the accusative nominals in turn, and if this movement is followed by a merge operation which creates a VP (not vP) specifier. In that case, once again, the hierarchical ranking of accusative and dative arguments will be reversed from the base positions. Notice that active sentences in Albanian must not undergo the same process, despite the presence of both dative and accusative Case. What this implies is simply that the accusative Case assignment does not trigger movement in Albanian, but agreement of \( v \) with the lower nominal in a passive, when Case is not assignment, does require movement (presumably to render the lower object accessible to T). This then must be an option which the Case parameters established by universal grammar make available.

A similar account can be given to another of McGinnis’ cases: British English dialects in which the lower of two objects is accessible for A-movement in passives. Thus, in (48b), *a book*, which presumably originates as the complement in an applicative phrase, ends up as the subject, while the higher object *Colin* remains inside the verb phrase. Again, the reversal of word order can be taken to reflect a double displacement of the two objects by T, with merger of each as specifiers for VP.

(48)  
a. Colin was given a book for his birthday.  
b. A book was given Colin for his birthday.

McGinnis obtains these results in Albanian, British English, and in other similar cases, by imposing her own condition on the how multiple specifier ordering is resolved. She proposes that a tucking-in ordering, which preserves the underlying order, is enforced when the attracting probe values the same type of feature in Agree. This is the case for Slavic multiple wh-movement. But if the probe values two distinct types of features, then the second specifier to be moved must become a higher specifier. Thus, for example, in McGinnis’ analysis of Japanese scrambling, a single head may value both \( \phi \)/Case features and a scrambling feature, and the scrambling feature will be satisfied by a phrase which raises to a higher specifier position.
Clearly this type of account will only be sufficient to the extent that a typology of features is available to ground it. Unless an independant metric of similarity is in place which allows us to judge when two features are to count as similar enough to produce a tucking-in order, the only way to identify what makes features “similar” is by looking to see whether they trigger a tucking-in result, or an inversion result. But this amounts to a circular logic, and must be abandoned. In contrast, the claim that the APP controls when tucking-in is preferred to inversion may be difficult to test, but it appears not to be circular, and is therefore preferable.

**Baker and Collins** (2006) discuss yet another set of cases which resemble the double object structures analysed by McGinnis, but in which the presence of an extra “linker” morpheme introduces an extra complication. An example is the Kinande sentence (49), where the linker y’ appears between the direct and the indirect object, and where the linker agrees in φ features with the preceding nominal.

(49) Mo-n-a-h-ere omukali y’-eritunda.
    AFF-1SS-T-give-EXT woman/1 Lk/1 fruit/5
    ‘I gave a fruit to a woman.’

As they show, the linker morpheme cannot simply be equated with an head position of ApplP, because the nominal which precedes it can originate as a higher or a lower element of the verb phrase. (Moreover, the applicative head is incorporated into the verb and transparently reflected in the verbal morphology.)

(50) a. Kambale a-seng-er-a omwami y’-ehilanga.
    Kambale 1S/T-pack-APPL-Fv chief/1 Lk/1 peanuts/19
    ‘Kambale packed peanuts for the chief.’

b. Kambale a-seng-er-a ehilanga hy’-omwami.
    Kambale 1S/T-pack-APPL-Fv peanuts/19 Lk/19 chief/1
    ‘Kambale packed peanuts for the chief.’

Baker and Collins simply accept that Lk may attract either a closer nominal or a more remote one, despite the evident challenge this poses for our understanding of locality constraints. Happily,
we can find a more principled way to describe the situation. The option of reversing the underlying word order can again be taken as an effect of the APP, if we simply assume that there is a movement-triggering head for which the displaced goals are merged only as specifiers for its complement. For example like (50), this requires a derivation for the verb phrase which includes a structure similar to (51).

(51)

```
LkP
  Lk  VP
    V  ApplP
      DP  Appl  DP
        omwami  ehilanga
```

We now need only characterise the Lk head as bearing $\phi$ features which may be valued either once or multiply, and which displace their goals. Lk itself can accomodate a specifier, but (as will be seen shortly) if it takes on a specifier immediately, the derivation will fail at a higher stage. So when Lk displaces a goal, the result must be integrated by a merge operation which targets VP. If Lk matches only once, then the result will be the structure (52a), and Lk will agree with $omwami$. If Lk matches a second time, then (52b) is generated, and Lk agrees with the second goal $ehilanga$.

In this respect, Kinande Lk behaves much like $v$ in British English passives.

---

Bruening and Rackowski (2001) document a similar agreement pattern in Wampanoag object agreement, where the second agreement target overwrites the effects of the first valuation.
The LkP category must be a complement to $v$, as Baker and Collins propose. And $v$ itself triggers movement with an unvalued $\phi$/Case feature, which it must match to the closest available nominal. As usual, $v$ cannot tolerate A-movement to its edge, so whatever type of object $v$ displaces will have to be merged as the specifier for its complement, which is LkP. Thus (52a) will give rise to (53a), and (52b) to (53b).
4. Conclusions

So now a fairly wide range of constructions have been identified which require movement which do not conform to Chomsky’s formulation of the NTC, but which satisfy the APP and the revised NTC. Many of these involve movement internal to VP but which is triggered by a valuation operation in the $v$ probe.

It is worth considering to what extent the mechanism of feature inheritance might be adapted to cover these same constructions. One might entertain the hypothesis that $v$ may sometimes transfer not just one, but two sets of $\phi$/Case features to the V head of its complements. Then V might be
able to attract both a dative and an accusative argument, just as Bulgarian C can attract multiple
wh-phrases. But then one would expect movement of two objects to take place in analogous fash-
on. In other words, a second object displaced by an inherited EPP feature would be expected to
tuck-in inside VP, preserving the original dative-accusative order. As we have seen, the opposite is
normally the case. In fact, once we consider data beyond the canonical one probe, one specifier sit-
uation, the patterns which emerge are consistently inconsistent with the strict Extension condition,
while they fit comfortably with the more nuanced furmulations suggested here.

References


versity of Chicago Press.


Baker, M. C., K. Safir, and J. Sikuku (2012). Sources of (a)symmetry in Bantu double object
constructions. Ms. Rutgers University and Moi University.


Press.

Bruening, B. and A. Rackowski (2001). Configurationality and object shift in Algonquian. Pro-
ceedings of WSCLA 5, 71–84.


