Essex Research Reports in Linguistics 37 (2001)

Defective Intervention Effects, Die!

Martin Atkinson

The purpose of this discussion is to seek to clarify a number of issues surrounding the Minimal Link Condition (MLC) of Chomsky (1995) in the context of the innovations introduced in Chomsky (1998, 1999). More challengingly, it examines critically a set of observations that Chomsky has used to establish the importance of *defective intervention effects* in syntactic computation, and argues that modifications in Chomsky's assumptions that are more or less independently justified obviate the need to recognise such phenomena. It is proposed that from a conceptual perspective, this is an unqualified good thing! Defective intervention effects require us to suppose that deleted syntactic objects can enjoy an afterlife, whereby they can continue to effect syntactic processes after their demise. Furthermore, it appears necessary to suppose that in certain circumstances, the demise is genuine, with deleted objects being totally inert for the functioning of subsequent syntactic operations. This requires Chomsky to explore conditions governing the *annulment* of defective intervention effects. As a consequence, *two* stages of explanation become necessary to account for why a deleted object behaves in a manner we might expect: a first stage, offering reasons for why it might *not* behave in this manner, and a second stage proposing an account of why these reasons are not operative in particular cases.

The modifications proposed are, I will suggest, consistent with the fundamental principles underlying the Minimalist Program, a matter of some importance, as I do not intend what follows to be seen as attempting to undermine this approach to syntactic theory. Accordingly, the paper should be read as concerned, first and foremost, as an attempt to sketch an alternative implementation of minimalist ideas.

The strategy pursued throughout is explicitly limited to consideration of examples utilised by Chomsky (1998, 1999). For these examples, I have attempted to work with the assumptions that Chomsky himself relies on, seeking to assess the detailed consequences of these assumptions for the examples in question. In so far as what emerge as weaknesses lead me to suggest modifications, I have not attempted to justify these modifications in a wider empirical context. That would be a considerable undertaking and must await another occasion.

After Section 1, a brief introduction, Section 2 focuses on Chomsky's (1998) introduction and illustration of the need to recognise defective intervention effects. In Chomsky (1999) what we find is not so much further examples of the notion, but a set of cases where the phenomenon appears *not* to arise where we might expect it to. Of course, if there are no defective intervention effects, these cases – *defective* defective intervention

effects, as it were - cease to be puzzling, but it may nonetheless prove instructive to try to unravel and assess the details of the argumentation, and Section 3 seeks to do this. Section 4 offers some concluding remarks.[1]

1. Introduction

Chomsky (1995, 297) embeds a statement of the MLC in the definition of the operation *attract* in (1):

(1) K *attracts* F if F is the closest feature that can enter into a checking relation with a sublabel of K.

In (1), 'closest' is to be interpreted in terms of whatever metric proves appropriate for measuring syntactic distance (cf. Chomsky *op. cit.*, 299), and it is this part of the definition in (1) that expresses the MLC - if G can enter into a checking relation with a sublabel of K but is not as close to K as is F, then K will not attract G. A major advantage in regarding the MLC as part of the definition of attract, is that derivations containing violations of the MLC are not even constructed. As a consequence, computations of economy (at least in this regard) are entirely 'local,' with the most economical choice being forced at the point of application of an operation, and the need to compute economy over *sets* of derivations, with the most economical being preferred, is removed. With this innovation, some of the concerns concerning computational complexity expressed, for instance, by Johnson and Lappin (1997) do not arise.

The notions of attraction and checking appearing in (1) are replaced in Chomsky (1998, 1999), by a complex of relations and operations, including Match, Agree and Move, and an obvious first task is to formulate a version of (1) within this new framework. Two remarks on

---

[1] Chomsky (2001) became available after this paper was completed. There, without acknowledging any of the difficulties I discuss below, Chomsky offers a different approach to defective intervention effects. This appears to depend on adopting Multiple Agree in something like the sense of Hiraiwa (2001). Suppose that some appropriate matching criterion hold between α and both β and γ in the configuration in (i):

(i) $\alpha \ldots \beta \ldots \gamma$

Then, we have Agree (α, β) and Agree (α, γ) with the presence of β not blocking the latter operation *because it is* α *that renders* β *inactive*. This is to be contrasted with (ii), where matching is again satisfied between α and both β and γ, but where β is now rendered inactive by α' distinct from α:

(ii) $\alpha \ldots \alpha' \ldots \beta \ldots \gamma$

In these circumstances, inactive β will count as an intervener and block Agree (α, γ). Careful assessment of this proposal is likely to be of considerable interest, but is not possible here.

(1) indicate the direction in which we should proceed. First, (1) was intended to embrace cases of overt attraction (category movement) and covert attraction (feature movement). Second, the content of (1) is fundamentally concerned with *covert* attraction, which, for Chomsky (1995), is regarded as the basic operation linking structural positions. Put simply, the fact that overt category movement satisfies (1) is parasitic on the fact that an individual feature or set of features within the category does so, with the movement of overt phonological material being a consequence of factors having nothing to do with the motivation for the operation of attraction itself. In the more recent framework, *Agree* is the operation that takes over the empirical burden of feature attraction in Chomsky (1995), so the obvious step to take is to incorporate the MLC into a definition of Agree.

To avoid unnecessary complications, let us suppose that the appropriate metric for computing 'closeness' is based entirely on c-command. We might, then, propose (2), where > designates c-command:

(2) Agree (α, β) if and only if:

(i) $\alpha > \beta$;

(ii) Match (α, β);

(iii) there is no γ such that Match (α, γ) and $\alpha > \gamma > \beta$.

Here, the MLC is expressed by (iii), and, introducing terminology, we might say that γ falsifying (iii) *intervenes* between α and β and serves to block Agree (α, β). From this perspective, γ falsifying (iii) in this way gives rise to an *intervention effect*.

While the sense of (2) should be readily apparent, it is arguably not the most perspicuous formulation of what we are looking for. Note that Agree is properly viewed as an *operation* which, when it applies to α and β occurring in a particular syntactic object, produces a modified syntactic object and (2) tells us nothing about the nature of this operation.[2] What (2) does is set out the conditions under which Agree can apply, but a moment's reflection indicates that these can be specified more naturally as conditions on the *relation* Match.

[2] Proper consideration of this matter raises a number of complex issues. For instance, Chomsky wishes to maintain that if Agree (α, β) takes place, then, supposing further conditions are satisfied, the uninterpretable features in α are (a) deleted with respect to narrow syntax; (b) valued for the purposes of phonology. Furthermore, deleted features can remain in the derivation in a way determined by the architecture of phases. This appears to require that the output of Agree can comprise *two* types of feature. I shall return to these issues in Section 3.

Suppose, using the terminology of Chomsky (1998, 1999), that α is a set of uninterpretable features comprising a probe. We can propose (3), where the relevant features of β (those identical to features in α) comprise a goal:[3]

(3) Match (α, β) if and only if:

(i) α and β are active;

(ii) $\alpha > \beta$;

(iii) $\alpha \subseteq \beta$;

(iv) there is no γ such that $\alpha \subseteq \gamma$ and $\alpha > \gamma > \beta$.

Here, (i) requires that included in both α and β are uninterpretable features and (iii) is intended to express the proposition that matching cannot be 'partial' with only some features of α being identical to those of β, the remainder perhaps being identified with those of some more remote goal. Taking account of the fact that satisfaction of Match is a necessary condition for the operation Agree, this has the consequence that it is not possible for α to agree with more than one other object. As Chomsky (1998, 40) puts it, focusing on the core case in which the ϕ-set of a finite T acts as a probe: 'We take deletion to be a "one fell swoop" operation, dealing with the ϕ-set as a unit. Its features cannot selectively delete: either all or none. The ϕ-features of T do not agree with different NPs, for example.' While (iii) has this desirable consequence for Agree, I do not believe that it has ever been explicitly adopted by Chomsky, and in what follows, it will play a significant role. Finally, (iv) is a statement of

---

[3] It is important for Chomsky that identity of features abstracts away from feature values. Thus, the *unvalued* feature [person] is identical to the *valued* feature [3person]. It is this notion of identity that is crucial to the subset relations appearing in (iii) and (iv). So, for instance, {[person], [number]} ⊆ {[3person, Sgnumber]}. As (3) is worded, it is not appropriate to strengthen the improper subset relation in (iii) and (iv) to set identity. This is because in the central example of the ϕ-set of T seeking a matching goal, an appropriate goal will contain an uninterpretable Case feature that renders it active. As (3i) indicates, the requirement to be active is a condition on β, so the uninterpretable Case feature must be regarded as belonging to β. As a token of this feature does not belong to α, it follows that here at least, we need $\alpha \subseteq \beta$. It is, of course, possible to contemplate a definition of Match that utilises set identity, as in (a):

(a) Match (α, β) if and only if:

(i) α is active, $\beta \subseteq \delta$ and δ is active;

(ii) $\alpha > \beta$;

(iii) $\alpha = \beta$;

(iv) there is no γ such that $\alpha = \gamma$ and $\alpha > \gamma > \beta$.

the MLC, and it contains one noteworthy feature; unlike α and β, γ is *not* required to be active.[4] A minimally distinct formulation to (3) is (4):

(4) Match (α, β) if and only if:

(i) α and β are active;

(ii) $\alpha > \beta$;

(iii) $\alpha \subseteq \beta$;

(iv) there is no *active* γ such that $\alpha \subseteq \gamma$ and $\alpha > \gamma > \beta$.

The question that immediately arises is: which of (3) or (4) is correct (if either)? A syntactic object (feature complex) is inactive if it contains no uninterpretable features, and a consequence of inactive status is that such a complex is *not* available as a goal – it is *defective* in one sense of this term.[5] This becomes particularly important in the system of Chomsky (1998, 1999), which regards the ultimate deletion of features as determined by the structure of phases. From this perspective, an uninterpretable feature occurring in phase P_n is merely 'marked for deletion' (Pesetsky and Torrego, 2000) until P_{n+1} is reached, and it remains accessible to the computation in some sense to be made precise throughout this period: it enjoys the afterlife referred to above.[6] If we have cases of inactive γ blocking Match in this way, it is appropriate to refer to these as cases of *defective intervention effects*, so the choice between (3) and (4) comes down to the question: do defective intervention effects exist?

2. The Case for Defective Intervention Effects

Chomsky's (1998, 38-9) answer to the question posed at the end of the introduction is affirmative. He introduces 'defective intervention constraints' as 'occurring in [$\alpha > \beta > \Gamma$], …

---

[4] There are further issues raised by formulations such as (3) that I shall not pursue here. For instance, (3ii) requires that the relation of c-command must be definable for subsets of the features comprising the nodes in familiar phrase structure representations. I don't believe any interesting substantive issues arise in this connection, but prudence would see the details worked out and presented. For the suggestion that c-command can be readily extended to features and feature sets, see Chomsky (1998, 32).

[5] It is important to clearly distinguish this notion of defectiveness from that employed by Chomsky in his discussions of infinitival T in raising and ECM constructions. In this latter context, defectiveness is identified with having an incomplete (relative to finite T and infinitival T in control structures) ϕ-set. The idea of defectiveness we are concerned with here is quite different.

[6] As is well known, Chomsky (1995) also contains a treatment of the removal of uninterpretable features from the computation that distinguishes between the *deletion* and *erasure* of checked features. The precise relationship between the deletion/erasure and marked for deletion/deletion contrasts will not be pursued here. As the subsequent footnote makes clear, however, these contrasts cannot be identified.

where β and Γ match the probe α, but β is inactive so that the effects of matching are blocked.' Before turning to cases, let us briefly consider one consequence of this answer. For simplicity, suppose that a *single* uninterpretable feature F is what makes β active when it enters the computation. We further suppose that this feature is marked for deletion by some operation, and continues with this 'intermediate' status until the appropriate phase is reached. Between the derivational stage at which it is marked for deletion and the phase at which it is actually deleted, the *very same feature* is not accessible as a goal, but is accessible for the MLC, as the latter is understood in (3iv).[7] While this characterisation goes some way towards satisfying the call for precision signalled above, I believe that this dual notion of accessibility is *prima facie* uncomfortable, and provides an initial reason for harbouring suspicions about the notion of defective intervention.[8]

Chomsky (1998) contains a number of examples that are intended to illustrate the necessity for allowing an intermediate syntactic object to be inactive while still giving rise to intervention effects, and thereby favouring (3) over (4). It is appropriate to consider these in turn. One is provided by (5) (p. 45, Chomsky's (47iv)):

(5) *there seem [$_{SU}$ several people] are [$_{PRED}$ friends of yours]

The details of how predicate nominals are analysed in terms of Chomsky's assumptions are not clear, but this is not important for the argument sketched here, as the general intentions are apparent enough. What he says about (5) is that it 'illustrates a defective intervention

---

[7] This is sufficient to establish a distinction between the deletion/erasure contrast of Chomsky (1995) and the marked for deletion/deletion contrast In Chomsky (1995), it was not possible for a single feature to be simultaneously erased (invisible at LF and invisible to the computation) and deleted (invisible at LF but visible to the computation). What we have here, however, appears to allow a feature to be accessible/inaccessible to the computation *in certain respects*.

[8] A possibility to bear in mind is that of treating the MLC as an 'interpretive' principle operative at the LF-interface, an option that Chomsky (1999, 22) appears to be attracted by in his discussion of Object Shift. If we were to follow this route, it could be maintained that being marked for deletion insulates an object from operations in the narrow syntactic computation, where the goal credentials of an object are vital, but leaves it visible to LF-interpretive processes (note that the distinction still cannot be identified with that between deletion and erasure, as the latter concerned visibility/invisibility *in the computation*). Without wishing to suggest that such an account could not be developed, it should be noted that this strategy loses the determinacy gained by building the MLC into the definition of Attract, Agree or Match, i.e. we would effectively extract (iv) from (3, 4) and treat it as an independent LF-condition. This, in turn, would allow the narrow syntax to produce derivations that were subsequently ruled out at LF, introducing the need for backtracking. In his web-published commentary on Chomsky (1999), Uriagereka (n.d.) appears to suggest that backtracking in a 'narrow derivational window' need not compromise the system. This may be the case, but without a detailed framework for assessing the consequences of computational complexity, the non-existence of backtracking procedures of any kind retains its attractiveness. I am grateful to Norio Nasu for a remark that prompted these observations.

effect' as 'SU is visible (barring PRED as goal) but inactive, unable to establish agreement with matrix T.' As far as the details of the derivation go, we need to consider the point at which matrix T is merged to give (6):

(6) T seem [several people are friends of yours]

Reconstructing Chomsky's argument, the ϕ-set of T seeks a goal, and while the ϕ-set of *people* is a superset of the probe's ϕ-set, it does not qualify as a goal because *several people* is inactive, its Case feature having been deleted in the subordinate clause. Despite this status, the ϕ-set of *people* can block the satisfaction of Match between the ϕ-sets of T and *friends*, i.e. the ϕ-set of *people* constitutes a defective intervener. Accordingly, we have neither Match (T, *people*) because of the inactive status of *people*, nor Match (T, *friends*) because of the role of the ϕ-set of *people* as a defective intervener, and the ϕ-set of T cannot be deleted and valued from within its domain.[9] Nor is the structure saved by merging *there* with (6), since it is assumed that *there* contains only a [person] feature and, being ϕ-incomplete, it is not able to delete the ϕ-set of T (see Chomsky (1998, 40; 1999, 4). Thus, the derivation continues to contain uninterpretable features and does not converge.

Now, it is obvious what Chomsky is trying to establish here, but it is equally clear that the example doesn't serve these goals. This is because if it is to illustrate a defective intervention effect, the defective intervener must intervene between a probe and an *active* goal (cf. 3i, 4i). But (7) indicates very straightforwardly that *friends* cannot function as an active goal in (6):

(7) several people are friends of yours

We must suppose that in the derivation of (7) all uninterpretable features are deleted. But (7) is just the subordinate clause in (5), so presumably, *friends* contains no uninterpretable features at the relevant stage of the derivation of (5), i.e. it is no more active than *people* is. This way of looking at things yields the conclusion that we have neither Match (T, *people*) nor Match (T, *friends*) in (6) because neither *people* nor *friends* fulfils the conditions on β in (3i) (or 4i). But this is all we need in order to account for the non-convergent status of (5). If merged *there* does not delete the uninterpretable ϕ-features of matrix T and if there is no active lower goal for this T to agree with, the derivation will not converge. We need say no

[9] From now on, I shall use expressions such as Match (T, *people*) rather than a more cumbersome expression indicating that it is only designated features of T and *people* that enter into the relation Match.

more. It appears, then, that (5) provides no evidence for the need to recognise defective intervention effects.

In a footnote to his discussion of (5), Chomsky cites (8), maintaining that this 'perennial trouble maker … falls into place if the (undeleted) [person] feature of embedded *there* bars association of matrix T to [*three men*].' (Chomsky's text has *a man*, an obvious typographical error):[10]

(8) *there seem there to be three men in the room

To understand the issue here, we need to first consider the intermediate stage of the derivation in (9):

(9) there to be three men in the room

Chomsky's assumptions demand that infinitival T here is ϕ-incomplete, containing only uninterpretable [person]. After Agree operates between this [person] feature and the complete ϕ-set of *men*, the former is marked for deletion, but remains accessible to the computation. When *there* is merged in Spec, T, thereby satisfying T's EPP feature, its [person] feature acts as a probe, identifying the marked for deletion [person] feature of T as a goal. [11] Agree now functions between *there* and infinitival T, and, as infinitival T is not ϕ-complete, the [person]

---

[10] Note that the fact that Chomsky's own discusssion refers to the *undeleted* [person] feature of *there* immediately indicates that *defective* intervention effects cannot be at issue here. If the [person] feature of *there* is undeleted, *there* is active and a candidate to function as a goal.

[11] We appear to meet an inconsistency at this point. If an item's uninterpretable feature(s) are marked for deletion, we have seen that that item should not be able to function as a goal, although it may count as an intervener for defective intervention effects. However, for Chomsky's treatment of *there* as an X^0 able to act as a probe to go through (1998, 44), it must be the case that the marked for deletion [person] feature of non-finite T *can* act as a goal. It is easy to see this in the case of (i):

(i) there are three men in the room

To account for convergence in this case, Chomsky has to suppose that the [person] feature of *there* is deleted under Agree with ϕ-complete T. But, as the ϕ-set of T has already been marked for deletion under Agree with the ϕ-set of *men*, this account requires an item with its uninterpretable features marked for deletion to be able to function as a goal. The inconsistency can be removed by maintaining that being marked for deletion does not disqualify *all* uninterpretable features as goals, with ϕ-features being an exception to this requirement, but such a proposal does no more than restate the problem. In what follows, I shall not attempt to resolve this inconsistency.

feature of *there* is not marked for deletion under this operation. Moving on in the derivation, we arrive at (10):

(10) T seem there to be three men in the room

Here, the ϕ-set of finite T acts as a probe and Chomsky's comment suggests that the problem arises because the [person] feature of *there* serves to block Agree (T, *men*). However, note that the incomplete ϕ-set of *there* is *not* a subset of the complete ϕ-set of T. Accordingly, the conditions for Match, as specified in (3, 4) are not satisfied, and T is free (in fact, required) to probe further and identify the ϕ-set of *men* as a matching goal. From this, perspective, then, (8) does not even display an intervention effect, even less a *defective* intervention effect. Subsequent merger of *there* with (10) produces no new difficulties: the [person] feature of *there* can identify the marked for deletion ϕ-set of matrix T as a goal and be deleted under Agree (but see n10). However, *there* in the subordinate clause retains its [person] feature, and this provides us with an account of the non-convergent status of (8) without invoking intervention effects in any way. It seems, therefore, that the interpretation of Chomsky's assumptions adopted here yields a straightforward account of the non-convergent status of (8) without resort to intervention effects of any kind.

In the same footnote, Chomsky observes that his proposal for how to deal with (8) does not extend to providing an account for the relative acceptability of (11):

(11) there look as though there are three men in the room

Note that here the intermediate *there* is in the specifier position of *finite* T. Thus, its [person] feature will be marked for deletion under agreement with ϕ-complete T on standard assumptions. However, in terms of Chomsky's own discussion, this state of affairs should qualify intermediate *there* as a defective intervener at the later stage in the derivation when the ϕ-set of finite matrix T is seeking a matching goal; Match (T, *men*) should be blocked under (3) (although not under (4)). As the ϕ-set of matrix T cannot be deleted by merged ϕ-incomplete *there*, the status of (11) remains problematic. Observe again, however, that if Match requires the probe features to be a subset of the goal features, the [person] feature of intermediate *there* will not be a candidate goal for the ϕ-set of matrix T, irrespective of its active or inactive status. Thus, matrix T *will* be able to identify the ϕ-set of *men* as a goal, and the agreement manifested in (11) is immediately accounted for. It appears, then, that (11) has a straightforward account relying on no more than the requirement that goal features

constitute a superset of probe features, and Chomsky's suggestion (46) 'that the [person] barrier may be overridden in some manner' is beside the point.[12]

Further evidence cited by Chomsky (1998, 46) for the importance of defective intervention effects comes from the examples in (12) (his (48i, ii):

(12)a. *there were decided [$_{\alpha}$ PRO to stay with friends]

b. *XP T-seems that [$_{\alpha}$ it was told friends CP]

The important observation here for Chomsky is that the Case features of PRO and *it* are deleted in α with the consequence that they are inactive in later stages of the computation. However, they remain effective as defective interveners, blocking the operation Agree between the ϕ-features of matrix T and *friends*.[13] As a consequence, in (12a), the uninterpretable features of matrix T remain in the derivation and, as they are not deleted by merger of ϕ-incomplete *there* in Spec, T, the derivation crashes. For (12b), the account is similar: merger of *there* in the XP position does not delete the ϕ-features of T, raising of *it* from the subordinate clause is not possible, as *it* is inactive as a goal, and raising of *friends* is not possible because of the defective intervention effect induced by *it*. Thus, there is no way to fill the XP position in (12b) so as to delete the ϕ-features of matrix T, and the derivation fails to converge.

The reasoning introduced in connection with (8) and (11) is not applicable to (12a, b) if we suppose that PRO and *it* are both ϕ-complete. This will qualify them as candidates for matching the ϕ-set of matrix T, so it looks as if the only way to block Agree (T, *friends*) in these cases is to invoke defective intervention. However, for (12a), this follows only if we adopt Chomsky's (1998, 1999) proposal that phases are 'propositional' and defined on *v** and C. In this structure, α is CP and thereby constitutes a phase, but the matrix clause does not contain a phase-defining *v**, so moving up the derivation from α, the next phase we come to is matrix CP. Linking Spell-Out to phasal structure via the Phase Impenetrability Condition (PIC) of Chomsky (1999, 10) ensures that PRO is still accessible to the computation at the point at which matrix T is merged, and the difficulty arises.

---

[12] To temper enthusiasm, it is necessary to draw attention to the markedly less good status (for me) of the 'singular version' of (11):

(i)?? there looks as though there's a man in the room

[13] As Andrew Radford has observed, in this example, *friends* will itself be inactive at the relevant stage of the derivation on standard assumptions, its Case feature having been deleted by *with*. As a consequence, the reasoning that follows regarding (12a) is not strictly necessary to establish its failure to demonstrate the necessity of defective intervention effects..

As noted in Chomsky (1998, 20-21), an alternative view of phases has on occasions been considered. This is the view that they should be defined on the basis of convergence, with a syntactic object that contains no uninterpretable features being immediately removed from the computation and made available to the interfaces. He argues against this view, but in Atkinson (2000) I maintained that this rejection was based on a failure to distinguish two roles for phases in the system. On the one hand, they appear to be necessary to constrain access of the computational system to the numeration, and from this point of view, Chomsky's insistence that phases cannot be identified with convergent portions of derivations appears to be entirely justified (cf. Chomsky, 1998, 21). Additionally, however, they have a derivational role, determining the stages in a derivation at which material can be sent to the interfaces. Chomsky's position presupposes that the notion of phase needed for these two purposes is identical, but this may be incorrect. Specifically, I argued that juxtaposition of the copy theory of movement with the orthodox, successive cyclic account of *wh*-movement and the position on phases Chomsky favours leads to inconsistency. This inconsistency does not arise if phases, *for the purposes of structuring the derivation*, are defined in terms of convergence. In short, I suggested that it might be appropriate to distinguish *lexical selection phase* and *derivational phase*, with only the former being defined on the basis of *v** and C.[14]

Suppose, then, that we reconsider (12a), taking phases to be defined in terms of convergence. The difficulty with it immediately fall into place. The CP α is convergent, so nothing in α is available as a potential goal for the uninterpretable ϕ-set of matrix T at the point at which it enters the derivation. Merger of ϕ-incomplete *there* in Spec, T will not serve to delete this ϕ-set and the only way to 'save' this derivation would be to have a token of ϕ-complete *it* available in the numeration to give (13):

(13) It was decided [PRO to stay with friends]

Unfortunately, (12b) does not yield to the same form of reasoning. In this construction, α is not convergent, as the Case feature of *friends* remains undeleted at this point in the derivation, and it follows that if the convergence view of phases is to be exploited here, it must be supplemented in some way. One route that might be worth exploring is to consider how the system might 'test' for convergence and the possibility that such testing may lead to false

---

[14] The empirical support for this conclusion was reinforced by the suggestion that phases defined in terms of convergence provide a rather natural instantiation of the notion of 'good design.' A syntactic object resulting from a convergent derivation is 'complete' as far as any derivation in which it is embedded is concerned (setting aside the role of defective intervention effects), and it therefore makes design sense for the computation to not proceed with the potential for accessing this object, a potential that will never be realised. It is also noteworthy that the intuition that phases should be 'relatively independent in terms of interface properties' (Chomsky, 1998, 20) resonates with the idea that they be defined in terms of convergence.

positives. It seems reasonable to suppose that such 'testing' should be bounded in some way, and a natural way to construe such bounding is in terms of the 'edge' of a constituent. Successive cyclic *wh*-movement can be used to illustrate this idea. A *wh*-item that is to undergo further movement must retain an uninterpretable feature in an intermediate Spec, C. Thus, a search of the 'edge' of this CP will be sufficient to conclude that it does not converge, and it will remain accessible to later stages of the computation. Now, suppose that the system *always* takes a decision on convergence on the basis of such an 'edge' search. In the case of (12b), it will conclude (wrongly) that α *is* convergent and thereby ensure that no part of α is accessible to later stages of the derivation. Of course, in this case, the conclusion is incorrect, but this may be a price of requiring that the system acknowledge economy constraints such as one restricting depth of search in tests of convergence.

Supposing that there is something to the above speculation, we can see that the status of (12b) is accounted for without the invocation of defective intervention effects. The problem is that α is incorrectly analysed as convergent, with the consequence that access to it is removed from the computation. Thus, when matrix T enters the derivation, it is not that its route to *friends* is blocked by inactive *it*; rather, there is no route to *friends*.[15]

It may be of some interest that the above speculation can be extended to the last example of defective intervention effects from Chomsky (1998) that I shall discuss here. This is the case of *wh*-islands, illustrated by (14), where *how* is to be construed with the subordinate clause.

(14)*how did you discover [which car John will fix]

Chomsky's recent work focuses on A-movement and does not contain a detailed account of A′-movement. However, in a brief paragraph (1998, 44-5), assuming parallels with his treatment of A-movement, he proposes that *wh*-expressions contain an uninterpretable [-wh] feature (cf. Case features in A-movement), and an interpretable feature, Q, matched by an uninterpretable feature Q in interrogative complementisers (cf. the role of ϕ-features in A-movement). In (14), the intermediate C acts as a probe, locating *which car* as a goal, Agree is triggered and (a) Q in the interrogative complementiser is deleted; (b) [wh-] in the *wh*-phrase is deleted; (c) *which car* is raised to Spec, C to satisfy the P-requirement of C. At this point, then, *which car*, having had its uninterpretable *wh*-feature deleted is inactive. Despite this, it can block matching between the uninterpretable Q feature of the matrix C and the

[15] For what it's worth, it is clear that the problem with (12b) resides *in* the clause *it was told friends CP*. The account I am considering locates the difficulty in this clause, and does not refer at all to failures to establish relationships between items in this clause and items higher in the structure. This might be seen as a small piece of evidence for the proposal, and is reminiscent of observations made by Epstein and Seeley (1999, 6).

interpretable Q feature of *how* in the subordinate clause. As a consequence, Agree (C, *how*) is not triggered, and it is impossible for *how* to raise to matrix Spec, C – the inactive *which car* serves to block Match (C, *how*).

It is worth noting that Chomsky's own view of phasal structure can be invoked to account for the properties of (14) without the need to recognise defective intervention effects. To see this, we merely need to note that, under PIC, the complement of the lower C will no longer be accessible to the computation at the matrix CP phase. So, while the 'edge' of this CP, including *which car* remains accessible to the matrix C, *how* is no longer accessible, and the question of whether *which car* intervenes to block Match (C, *how*) simply does not arise. Additionally, the complement of the lower C will not converge because of the presence of an uninterpretable [wh]-feature on *how*, and we can readily see that exactly the same consequence arises if phases are defined in terms of convergence with the testing for convergence being limited to 'edges.' Without the bounded search stipulation,, there are no phases in the derivation of (14) before we get to the matrix C, and as both *which car* and *how* remain accessible at this point, there would appear to be no alternative but to recognise defective intervention. However, if we invoke the strategy introduced above for dealing with (12b), we can see that the intermediate CP *will* constitute a phase that will be removed from the computation before the derivation moves on. Thus, The [wh-] feature of *how* will not be available as a potential goal for matrix C and the decision on phasehood will turn out to be erroneous, when LF discovers an uninterpretable [wh-] feature on *how*.

What I hope to have sketched so far, then, is an argument that the examples Chomsky (1998) cites for the importance of defective intervention effects can be accounted for without invoking this notion, so long as we maintain what looks like a reasonable constraint on Match and we adopt a convergence based view of phases, decisions on phasehood being 'risky,' because of the system's adherence to a bounded search.[16] I shall now turn to the complex discussion of defective intervention effects and their annulment in Chomsky (1999), with a view to assessing whether the assumptions I have adopted are sufficient to deal with the observations made there.

---

[16] There is one further case examined in Chomsky (1998, 47) that concerns the following (translated) example from Icelandic:

(i).*me (DAT) seem (pl) [t_{me} [John (DAT) to like horses (pl, NOM)]]

The issue here is that the matrix verb cannot agree with the embedded nominative, a fact that Chomsky describes as due to a 'defective intervention effect' with 'the ϕ-features of *John* block[ing] the T-associate relation between T-*seem* and nominative *horses*.' It is conceivable that the bounded search approach to testing for convergence might also play a role here, with the uninterpretable Case feature of *horses* outside the edge of the subordinate clause, which is therefore erroneously removed from the computation. Without detailed knowledge of the issues that might arise in Icelandic, there is little point in pursuing this speculation here.

3. The Annulment of Defective Intervention Effects

The first reference to defective intervention effects in Chomsky (1999, 12-13) appears in connection with his discussion of the examples in (15) (his (15ii, 16ii):

(15)a. there is likely to arrive a man
b. we expect there to arrive a man

For (15a), Chomsky considers the structure in (16):

(16) T be likely [there to arrive a man}

Rehearsing the analysis from Chomsky (1998), he maintains that the ϕ-set of matrix T acts as a probe, locating the [person] feature of *there* as a goal. Agree (T, *there*) operates, deleting the [person] feature of *there* and having the consequence that *there* is raised to Spec, T to satisfy T's EPP-feature. However, the ϕ-set of T is not deleted by ϕ-incomplete *there*, and is therefore free to probe again, identifying the ϕ-set of *man* as a goal. Agree (T, *man*) leads to the deletion and valuation of the ϕ-set of T and the Case feature of *man*, and the derivation converges. In connection with the operation of Agree (T, *man*), Chomsky observes that 'no intervention effect is induced by [the copy of *there*].' This remark merits some reflection.

The copy theory of movement requires that the raising of *there* to Spec, T leaves a copy of *there* in its initial position. Furthermore, since Agree (T, *there*) has the consequence that the [person] feature of *there* is deleted, this will also be a characteristic of the copy (Chomsky, 1998, 29) We are thus led to consider (17), where the tokens of *there* are inactive by virtue of containing no uninterpretable features (the subscripts are simply to enable unambiguous identification in the subsequent text, and I adopt the convention of prefixing uninterpretable features by *u* and striking through features marked for deletion):

(17)	there$_1$	T	be likely	[there$_2$	to arrive a man]
	[~~*u*person~~]	[*u*person]		[~~*u*person~~]	[3person]
		[*u*number]			[Sgnumber]

If defective intervention effects constitute a general phenomenon, we would predict that Agree (T, *man*) should be blocked by the intervening *there$_2$*. Clearly, this is incorrect, so steps must be taken to nullify the effects of *there$_2$* as a defective intervener.[17] To this end, Chomsky

[17] It is worth re-iterating remarks made in the introduction at this point. Chomsky has introduced the notion of defective intervention, and I have suggested that it is (a) conceptually suspect; (b)

makes two proposals. First, he offers (18), as a principle 'which has been suggested in various forms':

(18) maximise matching effects

Second, he considers (19), which he regards as 'conceptually plausible and empirically supported.'

(19) only the head of an A-chain (equivalently, the whole chain) blocks matching under the MLC

Applying these to (17), it is easy to see how they are intended to produce the desired outcome. Take (18) first. The natural way to construe this is as a constraint on Match: if α and β do not match maximally, then it is not the case that Match (α, β). Put like this, (18) looks like an informal version of (3ii, 4ii), the requirement that the features comprising a probe should be a (possibly improper) subset of an appropriate goal.[18] In (17), {[*u*person], [*u*number]} from T act as a probe. Locating *there*$_2$, this probe finds {[~~*u*person~~]}, but, as matching is not maximal, we do not have Match (T, *there*$_2$). A deeper search locates {[3person], [Sgnumber]} at *man*. Matching is maximal, so we have Match (T, *man*) and Agree (T, *man*) can operate.

Whereas the status of (18) is relational, voiding the defective intervention potential of *there*$_2$ because of its relationship to T, (19) is absolute. Because *there*$_2$ is not the head of an A-chain, it is disqualified as a defective intervener *tout court*. What we appear to have, then, is an overdetermination of the properties of (15a), with either (18) or (19) being sufficient to block the defective intervention status of *there*$_2$.

It is straightforward to see that the first of the above arguments is difficult to sustain. Recall that the raising of *there* to Spec, T is a consequence of Agree (T, *there*) operating in the configuration (16). But *there* in (16) no more provides a maximal match for the ϕ-set of T

---

not necessary for examples where it is invoked. We now meet a situation where Chomsky's own assumptions lead him to expect a defective intervention effect and the need to produce an account of why we don't have one. It would be much more straightforward to maintain that the reason we don't have one here is because defective intervention effects don't exist!

[18] There are, of course, other construals that could be contemplated. For instance, if we suppose that the subset requirement is not satisfied anywhere within the domain of the probe, it might be suggested that an appropriate goal is whatever provides the *largest number* of features that can be identified with features of the probe. However, it is not difficult to see problems for such a construal. For instance, it is not guaranteed to provide a *single* goal, as there could be candidate goals in the domain matching the probe to the same extent. Furthermore, it requires exhaustive search of the domain before it can safely identify a goal. The subset condition in (3ii, 4ii) gives rise to neither of these difficulties. The closest β in the domain of α such that $\alpha \subseteq \beta$ constitutes *the unique* goal for probe α. If there is no β satisfying the subset requirement, α has no accessible goal.

than does *there*$_2$ in (17). Accordingly, if Chomsky is to be consistent, he must suppose that we do not have Match (T, *there*) in (16). What we have is Match (T, *man*), and the prediction that (20) is the next stage in the derivation:[19]

(20) a man T be likely [there to arrive a man]

The issues raised by (19) are slightly more complex. This is because whereas the straightforward interpretation of (18) is that it imposes an obligatory constraint on Match, (19), as worded, is *permissive* in the sense that it allows an object that is not the head of an A-chain to be transparent to Match. It does not, however, *require* this transparency. Now, it is not clear to me whether Chomsky intends this permissive interpretation of (19). If he intends something stronger, a wording along the lines of (21) would be more appropriate:

(21) the head of an A-chain (equivalently the whole chain) blocks matching under the MLC *and* an object that is not the head of an A-chain is transparent to matching.

Given this, we first note that *there* in (16) is no more the head of an A-chain than is *there*$_2$ in (17), since a standard requirement is that an A-chain contains a θ-position. If (21) is what is intended, it should be clear that it gives rise to exactly the same difficulties as (18). However, (19) does not require *there* in (16) to be transparent to Match, and (19) is therefore compatible with Agree (T, *there*) operating in (16), as Chomsky's account demands. Equally, for (17), (19) permits *there*$_2$ to be transparent for matching purposes, so Agree (T, *man*) is a legitimate operation at this stage in the derivation.

Although the above suggests that the imposition of (19) is consistent with Chomsky's account of (15a), there is one noteworthy complication. As (19) is permissive, it allows *there* in (16) to be transparent to matching. Thus, we can have Agree (T, *man*) and derive (20) as the next stage in the derivation. As observed in n17, this does not require that the system recognise **a man is likely there to arrive* as convergent, as a consequence of Agree (T, *man*) is the deletion of T's φ-set. So this φ-set will not act as a probe again, with the result that the [person] feature of *there* remains undeleted. Thus, while (19) permits two derivations to be

---

[19] Note that we do not go so far as to predict that (i) is well-formed:

(i) *a man is likely there to arrive

This is because the [person] feature of *there* will remain undeleted in (20), so the derivation will not converge. A possibility that will be considered further below is that different conditions might be imposed on Match depending on whether a potential goal is active or not. According to this way of looking of things, it might be possible to maintain that matching obtains between T and *there* in (16) but not between T and *there*$_2$ in (17). The relevant discussion concerns (28) and (29) below.

constructed on the basis of the lexical array underlying (15a), only one of these is convergent and it corresponds to the correct structure. By contrast, (18) and (21) appear to require that only the non-convergent derivation is available.

Is it possible to offer an account of (15a) that overcomes the difficulties with (18) and (21) and offers an alternative to adopting (19)? An obvious candidate for such an account is to subscribe to what Chomsky (1999) refers to as Alternative (II) (see also Epstein and Seeley, 1999). According to this view, there is no A-movement in the derivation of (15a), and *there* is merged directly in Spec, T. At the relevant stage of the derivation, we have (22):

(22) T be likely [to arrive a man]

The uninterpretable φ-set of T probes and locates the φ-set of *man* as a matching goal. This matching is maximal in the sense that the subset condition of (3ii, 4ii) is satisfied. Agree operates, deleting and valuing the φ-features of T along with the Case feature of *man*. Since *there* is present in the lexical array, the principle of Merge over Move, requires that the EPP feature of T is satisfied by pure Merge of *there*. The [person] feature of *there* is deleted under Agree (*there*, T) and the derivation converges.[20] On this account, there are no potential defective interveners, and therefore no requirement that their status be annulled. Even without defective interveners, the principle in (18) (interpreted as 3ii, 4ii) plays a role, but not one that leads to inconsistency. However, the lack of copies of moved objects means that the account does not rely on (19) or (21) at all.

Similar considerations arise for the ECM construction in (15b). Chomsky focuses on the representation in (23):

(23) we [$_{v*P}$ *v** - expect [there to arrive a man]]

Of this representation, he says (12): 'Agree holds of (*v**, [*there*]), deleting the [person] feature of [*there*] but leaving *v** intact so that Agree holds of (*v**, *man*).' As Chomsky leaves open the question of whether *there* raises to Spec, *v** or to some other position in *v**P, the issue of copies of moved items giving rise to defective intervention does not necessarily arise. However, even with the passage cited here, it is easy to see the difficulty with Chomsky's account in the light of (18) or (21). The match between the φ-set of *v** and the [*u*person] feature of *there* is not maximal, so *there* should simply be bypassed in *v**'s search for a goal, which yields Match (*v**, *man*). Agree (*v**, *man*) has the consequence that the φ-set of *v** is

[20] It is worth reiterating that, as it stands, this account still embraces the inconsistency raised in n10 of allowing a marked for deletion feature, in this case the [person] feature of matrix T, to function as a goal.

deleted so that *v** cannot subsequently act as a probe. The [person] feature of *there* remains undeleted, and no convergent derivation can be constructed. Equally, for (21), *there* is not the head of an A-chain, so *v**'s search for a goal should immediately proceed to *man* with the noted undesirable consequences.

As far as the role of (19) with respect to (23) is concerned, parallel considerations to those discussed for (16) arise. The convergent derivation is available, as (19) permits Match (*v**, *there*). Additionally, the non-convergent derivation required by adherence to (18) or (21) is produced if the permission licensed by (19) is not exercised and the system regards *there* as transparent for matching.

For an alternative that avoids these difficulties, the obvious suggestion is to suppose that *v** has an EPP-feature that is satisfied by pure Merge of *there*. Thus, we first consider a stage of the derivation schematised in (24):

(24) [$_{v*P}$ we *v**-expect [to arrive a man]]

Match (*v**, *man*) is the only option under (3ii, 4ii), and the φ-set of *v** is deleted along with the Case feature of *man*, both of which are also valued as a consequence of Agree. Then, *there* is merged as an outer Spec, *v** to satisfy *v**'s EPP feature, and its [*u*person] feature is deleted under Agree (*there*, *v**) (but see n18). Subsequent merger of matrix T yields (25):

(25) T [$_{v*P}$ there [we [*v** - expect [to arrive a man]]]]

In (25), the complete ϕ-set of T locates the ϕ-set of *we* as a goal, skipping the intermediate *there*, which does not satisfy (3ii, 4ii). Raising of *we* to satisfy the EPP feature of matrix T along with raising of the verbal complex gives us (15b).

As noted already, reliance on either (18) or (19) is thought by Chomsky to yield the desired effects for (15a, b).[21] The next pair of examples considered, he maintains, enable us to discriminate between these two principles. The examples are (p. 13):

(26)a. there seem to have been caught several fish

b. we expect there to have been caught several fish

[21] In fact Chomsky observes that (19) has a role in the derivation of (15b) only if the raising of *there* to Spec, *v** or 'to a position within the *v** complement' (37) takes place in (23). This seems to be mistaken, since even if *there* remains *in situ* in (23), it does not constitute the head of an A-chain, and (19) will permit the ϕ-set of *v** to regard it as transparent in its search for a matching goal.

The important additional property that shows up in these examples is that of agreement (morphologically marked in Icelandic, for example) between the participle *caught* and *fish*. Focusing on (26a), Chomsky suggests that we first consider the stage of the derivation in (27), where '… PRT is a light verb distinct from *v**' (p. 37):

(27) [$_{\alpha}$ PRT [catch [$_{DO}$ several fish]]

As PRT is assumed to be adjectival, Chomsky further proposes that its ϕ-set 'consist[s] of (unvalued) number, gender and Case, but not person' (p. 14).[22] The ϕ-set of PRT acts as a probe, locating the interpretable ϕ-set of *fish* as a goal. Since the latter set is complete, the [*u*number] and [*u*gender] features of PRT are marked for deletion and valued, the latter operation ensuring the correct morphological agreement in a language where this is overt. However, since neither of the Case features of PRT and *fish* are themselves valued, they are incapable of deleting and valuing each other. Accordingly, when α is completed, both PRT and *fish* remain active in the derivation.
Chomsky next considers (28):

(28) T	seem [there	to have been [$_{\alpha}$PRT	[catch [$_{DO}$ several fish]]]]
[*u*person]	[*u*person]	[~~*u*number~~]	[3person]
[*u*number]		[~~*u*gender~~]	[Sgnumber]
[*u*gender]		[*u*Case]	[?gender]
			[*u*Case]

Generalising the discussion from (15a), he supposes that the φ-set of matrix T acts as a probe and locates the [*u*person] feature of *there* as a goal. The consequences are deletion of the [*u*person] feature of *there* and raising of *there* to Spec, T to satisfy T's EPP feature (not included in the representations):[23]

---

[22] I assume that it is a slip to regard Case as belonging to the ϕ-*set* of PRT. More accurately, we should say that its uninterpretable features contain the unvalued ϕ-features [number] and [gender] and a Case feature.

[23] Note the tension here between this attitude to the consequences of Agree (T, *there*) and that described in the text for Agree (PRT, *fish*). As far as the latter is concerned, the Case feature of *fish* cannot value the Case feature of PRT, being itself unvalued, and this is linked to the non-deletion of the Case feature of PRT. For Agree (T, *there*), we again have matching of unvalued features ([*u*person]), but here there is no talk of lack of valuation and the familiar deletion consequences of Agree are assumed to occur. Of course, there is the difference that whereas case valuation of PRT is overt in some cases, *overt* valuation of the [person] feature of a pure expletive is not an option. Arguably, however, this is incidental, and should not be used to 'massage' the fundamental operations of the system.

(29) there$_1$ T seem [there$_2$ to have been [$_\alpha$ PRT [catch [$_{DO}$ several fish]]]]

[~~*u*person~~]	[*u*person]	[~~*u*person~~]	[~~*u*number~~]	[3person]
	[*u*number]		[~~*u*gender~~]	[Sgnumber]
	[*u*gender]		[*u*Case]	[?gender]
				[*u*Case]

The ϕ-set of matrix T now probes again. The first thing we note is that by reasoning identical to that invoked for his analysis of (15a), Chomsky maintains that *there$_2$* is transparent to the search for a goal – it does not provide a maximal match for the ϕ-set of matrix T, nor is it the head of an A-chain. Thus, as far as *there$_2$* is concerned, there is nothing to prevent the ϕ-set of matrix T identifying the ϕ-set of *fish* as a goal. But what about PRT?

Recall that the assumptions in play have PRT with [*u*number] and [*u*gender] features marked for deletion, but still active by virtue of its undeleted Case feature. Chomsky says (p. 14): ' … the probe T matches the still visible goal PRT, valuing its Case feature; and the probe matches the goal DO, valuing the Case feature of DO as well as its own features (since DO is ϕ-complete).' Thus, on the one hand, matrix T's search for a goal stops at PRT, and on the other, it bypasses PRT to locate DO. What justifies this second step for Chomsky?

In response to this question, he again refers to the voiding of intervention effects via (18) and (19). In this connection, he says (*ibid.*): 'Once again there is no intervention effect induced by [there], or in this case by PRT either. In the case of [(15a, b)] there were two possible reasons: principle [(19)] (assuming raising of [*there*] to within matrix VP) or principle [(18)], which requires maximal (probe, goal) effects. In [(29)], there is no raising of PRT, so we must resort to principle [(18)] … The [number/gender] features of the probe bypass [*there$_2$*] and its [person] feature bypasses PRT, allowing probe-DO match.' It appears, then, that Chomsky has an argument that (18), involving the maximisation of matching, is the operative principle throughout these examples.[24]

Not surprisingly, it is not difficult to see that (18) does not give rise to the desired effects in Chomsky's analysis of (28). The status of *there* in (28) and *there$_2$* in (29) is identical to what we saw in connection with (15a). Since *there$_2$* in (29) is a copy of *there$_1$*, it is unclear how a principle such as (18) can differentiate between them, and if *there$_2$* is obligatorily

[24] For the sake of not complicating the discussion further, I shall assume that Chomsky is correct in his assessment of the irrelevance of (19) for the analysis of the role of PRT in (28). However, it is easy to see that while his remark that there is no raising of PRT in this construction is correct, this does not impinge on the significance of (19). The principle in question refers explicitly to the property of being the head of an A-chain as what is responsible for the defective intervention potential of an object. But PRT is no more the head of an A-chain than is an expletive. Accordingly, on the permissive construal of (19) I have presented in the text, there is nothing to prevent PRT being transparent to matrix T's search for a goal.

bypassed by the goal-seeking ϕ-set of matrix T in (29), it is necessary that *there* is similarly bypassed in (28). Setting aside the role of PRT and assuming the search for a goal extends to *fish*, we predict the next stage in the derivation to be (30):

(30) several fish T seem [there to have been [caught several fish]]

As regards PRT, the situation at first sight appears to be similarly inconsistent, although there is one difference that might be regarded as important, a difference that, having being clarified, might also provide a way of dealing with the expletive problem raised in connection with (15a) and now generalised to (28). Recall that what Chomsky proposes is that the ϕ-set of T treats PRT in two different ways. On the one hand, with its undeleted Case feature, it constitutes a goal for this probe, a consequence of which is the deletion of this Case feature. With its Case feature now deleted, it can no longer function as a goal, but we might expect it to produce a defective intervention effect, blocking Match (T, *fish*). However, it doesn't do this, and Chomsky proposes to account for this inability by relying on (18).

At this point, the interpretation of (18) for Chomsky becomes crucial. To date, I have regarded it as an informal version of (3ii, 4ii), but if consistency is to be achieved, it is readily apparent that this can't be what he has in mind here. In an attempt to make sense of the proposal, I see no alternative but to painstakingly (or, perhaps, painsmakingly!) reproduce the various steps of the account. In (29), then, the ϕ-set of matrix T acts as a probe and locates *active* PRT. Under (3ii, 4ii), we do not have Match (T, PRT), but if matching is a prerequisite for the operation of Agree, and if the deletion and valuation of PRT's Case feature is a consequence of Agree, then it must be the case for Chomsky that the conditions on matching are satisfied here. An interim conclusion, then, is that Chomsky's position is not consistent with the conditions on Match formulated in (3ii, 4ii). Let's recognise this by saying that Match′ (T, PRT) in (29), where Match′ is whatever Chomsky has in mind.[25] With PRT now inactive, T's φ-set probes again and can apparently bypass PRT to yield Match′ (T, *fish*). But recall Chomsky's (1998, 38-39) original characterisation of defective intervention effects referred to earlier as 'occurring in [$\alpha > \beta > \Gamma$], … where β and Γ *match* the probe α, but Γ is inactive so that the effects of matching are blocked.' (my italics – MA). We have inactive PRT and Match′ (T, PRT), so PRT should count as a defective intervener. As it doesn't, and as the reason Chomsky gives for it not so doing refers to maximisation of matching, it would appear to follow that the matching referred to in this principle cannot be Match′. The only way I can make sense of this is to suppose that Chomsky is operating with two distinct

[25] Presumably, a minimal condition on Match′(α, β) is that there is $\delta \in \alpha$ such that $\delta \in \beta$, i.e. at least one feature in the probe also occurs in the goal. If, in addition, there must be $\eta \in \alpha$ such that $\eta \notin \beta$, we perhaps have all we need for a definition of Match′.

notions of matching. On the one hand, there is Match, satisfying (3ii, 4ii), and on the other, Match′. For a potential goal β that is active, Match′ (α, β) will be sufficient to trigger Agree (α, β); however, if β is inactive, Match′ (α, β) does not have this consequence (Match (α, β) is required), and α must probe again in a search for a more remote goal γ.

Before proceeding, let us check that the above reconstruction is consistent with what Chomsky appears to intend, at least as far as the goal status of the φ-sets of PRT and *fish* in (29) are concerned. The ϕ-set of matrix T acts as a probe and locates active PRT as a goal. PRT is active and Match' (T, PRT) is sufficient to trigger Agree; as a consequence the Case feature of PRT is deleted and valued, but, as the ϕ-set of PRT is incomplete, the ϕ-set of matrix T remains undeleted and probes again. For this process, PRT is inactive, so Match (T, PRT) is required if PRT is to block T's search for a goal. This condition is not satisfied, so T can bypass PRT and locate active *fish* as a goal. We have Match' (T, *fish*), so Agree (T, *fish*) operates, with the consequence that all relevant features are marked for deletion.[26] In short, it appears that we can preserve an intelligible role for (18) if we are prepared to countenance two notions of matching. If an active goal β is identified by a probe α, Match′ (α, β) is sufficient to trigger Agree (α, β). If, however, inactive β is identified in the search for a goal, Match′ (α, β) allows α to bypass β in its search for a goal. Only if Match (α, β) obtains, will β block a deeper search and, in these circumstances, count as a defective intervener.

With this much in place, we can return to (15a) and immediately see that similar considerations might be invoked here. We have (17), repeated as (31):

(31) there$_1$	T	be likely [there$_2$	to arrive	a man]
[~~*u*person~~]	[*u*person]	[~~*u*person~~]		[3person]
	[*u*number]			[Sgnumber]

This configuration is produced by Match′ being satisfied by the ϕ-set of matrix T and the [*u*person] feature of *there*, a situation that is possible because, with its [*u*person] feature undeleted, *there* is active. However, after the operation of Agree triggered by this token of Match′ being satisfied, *there* is no longer active and (18), as it is now being construed, requires that it be bypassed by matrix T's ϕ-set in its search for a goal. It is easy to see that these considerations can be straightforwardly extended to deal with the relations between T and *there* in (25) and (28).

---

[26] I use Match′ (T, *fish*) here, since *fish* is assumed to be active. Accordingly, Match′ (T, *fish*) is all that is needed to trigger Agree, and the fact that we have Match (T, *fish*) is immaterial. This indicates that (18) does not *drive* a probe to seek a maximal match as such. Rather, it provides a license for a probe to bypass any nonactive object that does not provide a maximal match. In other words, only inactive objects that provide a maximal match for a probe will give rise to defective intervention effects.

It might be appropriate to attempt to summarise the above discussion in general terms. There are three configurations of which we need to take account, schematised in (32) (α active and able to function as a probe throughout):[27]

(32)a. $\alpha > \gamma > \beta$; γ and β both active, γ matches α but not maximally, β matches α maximally (i. e. Match′(α, γ) and Match (α, β)).

b. $\alpha > \gamma > \beta$; γ inactive and β active, γ matches α but not maximally, β matches α maximally. (i. e. Match′(α, γ) and Match (α, β)).

c. $\alpha > \gamma > \beta$; γ inactive and β active, γ matches α maximally, β matches α maximally. (i. e. Match (α, γ) and Match (α, β)).

For (32a), Match′ (α, γ) so Agree (α, γ) is triggered. Because matching is not maximal, the features that make α active will not be marked for deletion, and the application of Agree moves us to the situation in (32b). In (32b), a potential defective intervener γ requires Match (α, γ). We do not have Match (α, γ), so α proceeds to β. We have Match (α, β), so Agree (α, β) is triggered. This is the annulment of defective intervention effects via maximisation of matching. Finally, in (32c), γ is again inactive, but now matching between α and γ is maximal. Thus, we have Match (α, γ) and a defective intervention effect should result. It is important to be clear, however, that none of the examples considered so far from Chomsky (1999) concern this possibility, as they are entirely devoted to the interplay between the situations depicted in (32a) and (32b).

While it appears possible to see the assumptions operative in Chomsky (1999) as providing a consistent account of the examples he discusses, it is not easy to feel comfortable with the mechanisms he appears to be committed to. The notion of matching as involving a conceptually attractive subset relation as in (3ii, 4ii) with the consequence that agreement is a 'non-distributed' process should not be given up lightly, and this raises the question as to whether alternatives to Chomsky's proposals exist. Re-iterating a point made much earlier, it is important that any such alternative should seek to remain within the guidelines provided by

---

[27] In fact, the stipulation that β matches α maximally in the configurations in (32) is not necessary if my reconstruction of what Chomsky's account requires is correct. All that is required is Match′(α, β), and the only important role for maximisation concerns the relationship between α and γ. This, in turn, gives rise to a formal difficulty, in that whereas Match′ (α, β) should not exclude Match (α, β) – we do not wish to illegitimise a remote goal on the grounds that it offers a maximal match to a probe, we do want Match′ (α, γ) to exclude Match (α, γ) in (32a, b). This is because *only* a non-maximally matching, inactive element should be transparent to α's search for a goal. This formal difficulty can be overcome by defining MATCH as the disjunction of Match′ and Match and replacing Match by MATCH in each of the second clauses in (32). I shall not implement this suggestion in the text.

the minimalist approach. We have already seen that such alternatives can be contemplated for (15a, b), so we next consider the issues raised by (26a), repeated as (33):

(33) there are likely to be caught several fish

As far as the expletive is concerned, the natural (Alternative II) assumption is to again suppose that it enters the derivation at the final stage without movement, so we can focus on the properties of the participle. Recall that Chomsky supposes that this object has uninterpretable number, gender and Case features, with the first two being deleted and valued under agreement with *fish*. However, the Case feature of PRT is not dealt with in this way, as at the point at which Agree (PRT, *fish*) operates, the Case feature of *fish* has itself not been valued. The consequence is that Chomsky has to propose mechanisms whereby the Case features of PRT and *fish* are valued under *distinct* tokens of Agree, Agree (T, PRT) and Agree (T, *fish*). The complications we have been concerned with above are a direct consequence of this view that PRT and *fish* agree in Case only by virtue of each of them agreeing with T. This contrasts starkly with number and gender agreement between PRT and *fish*, which is 'direct.'

We can perhaps make some headway in formulating an alternative approach to such structures by reflecting on the nature of the operation Agree. Consider a core case, schematised in (34):

(34) $T_{finite\ ...}$ [subject ….

In such a configuration, the ϕ-set of T acts as a probe and locates the matching ϕ-set of subject as a goal. Agree is triggered and, as regards the uninterpretable Case feature of subject, *two* consequences follow. First, it is deleted as far as the narrow syntactic derivation is concerned, essential for LF-convergence; second it is valued as nominative by some feature of finite T, this valuation being what is of importance for the phonology.[28] It seems natural to suppose, therefore, that Agree has a dual output, a deleted feature that may remain accessible in the narrow syntax, and a valued feature that is visible only to the phonology.[29] The alternative of

---

[28] For the sake of concreteness, we can suppose that there is a morphophonological reflex of nominative case. If there is not, there will still be valuation of this Case feature under Agree, but its phonological consequences will not be overt.

[29] Another way to think about this is to regard Agree as single-valued, effectively *converting* an uninterpretable formal feature into a morphophonological feature. From this perspective, deletion in the narrow syntax is an all-or-none process, and there is no place for such an intermediate status as being 'marked for deletion.' I am grateful to Mike Jones for suggesting this way of looking at things, which strikes me as conceptually attractive. However, I shall not pursue it further here.

supposing that a single feature can be simultaneously deleted and valued is not immediately intelligible. Note further that however feature valuation is annotated, it would appear to provide an exception to Chomsky's Inclusiveness Condition, whereby the syntactic computation is constrained not to introduce new material in the course of its operation. We might suggest, therefore, that the process of uninterpretable feature valuation, while being a reflex of syntactic configurations, is properly conceptualised as belonging to the phonology. Thus, the relevant aspects of a token of Agree operating on the configuration in (34) could be represented as in (35), where F is the interpretable feature of finite T responsible for the evaluation of [*u*Case] as nominative:

(35) Agree ⟨[F, *u*Person, *u*Number], [αPerson, βNumber, *u*Case]⟩ =
⟨⟨[F, ~~*u*Person~~, ~~*u*Number~~], [αPerson, βNumber, ~~*u*Case~~]⟩,
⟨[F, αPerson, βNumber], [αPerson, βNumber, NomCase]⟩⟩

In (35), the first object in the ordered pair providing the output of Agree remains available to the syntactic computation; the second is to be regarded as part of the phonology.

With the above in mind, consider the configuration in (27), repeated as (36):

(36) [$_{\alpha}$ PRT [catch [$_{DO}$ several fish]]

Here, the ϕ-set of PRT acts as a probe and locates the ϕ-set of *fish* as a goal. Deletion and valuation of the ϕ-features of PRT is a consequence of Agree, but, Chomsky maintains, deletion and valuation of the Case feature of PRT does not occur, since the Case feature of *fish* is not itself valued at this stage of the derivation.[30] But note that the 'standard' conditions for deletion of the Case feature of PRT are satisfied, since *fish* is ϕ-complete.[31] Suppose, then, that this Case feature *is* deleted in the narrow syntax. What about its valuation? Clearly, this cannot be provided by the Case feature of *fish*, as Chomsky notes, but, if we are prepared to locate Case valuation in the phonology, there would appear to be no immediate objection to identifying the valuation of the Case feature of PRT with the valuation of the Case feature of

---

[30] Observe that description of the problem in these terms is suggestive. The 'natural' way to view the Case feature of the participle is as determined by the Case feature of the nominal. But Chomsky's account is not consistent with this naturalness.

[31] The scare quotes here are intended to indicate that the situation is standard in only one sense. For core cases of the deletion of Case features, the requirement is that the *uninterpretable* ϕ-set of a *probe* should be complete in order to delete the Case feature of a matching *goal*. Here, we are supposing that a complete, *interpretable* ϕ-set of a *goal* can delete a Case feature in a matching probe. For discussion of some complexities arising in connection with deletion, once proper account is taken of probe-goal asymmetries, see Atkinson (2000).

fish, at a point at which the latter has been determined. There is little point here in speculating on notations that might achieve this, but the important consequence of this way of looking at things is that the Case feature of PRT is deleted at the same stage of the syntactic derivation as are its ϕ-features. Accordingly, there is no requirement that the ϕ-set of matrix T in (26a) should acknowledge PRT as an intermediate goal. At the point at which T enters the derivation, PRT will be inactive and the only legitimate goal for T's ϕ-set will be the ϕ-set of *fish*, an object that meets the conditions on Match set out in (3ii, 4ii). Maximisation is given a unified and conceptually attractive interpretation with none of the complexities set out in (32).

4. Concluding Remarks

Returning to the question posed in the introduction regarding the formulations of the MLC in (3) and (4), I hope that the above discussion has taken some small steps towards making a case for the version in (4), whereby only an *active* object, one that is visible to core syntactic processes, is capable of blocking a matching relation between a probe and potential goal. The conceptual case for this conclusion is easy to make: if an object is invisible as a goal, then it should be invisible to the syntax in a quite general way, and subscription to intermediate states of visibility is to be avoided if at all possible.[32] However, Chomsky (1998) presents a series of empirical observations, suggesting that such an intermediate state is, indeed, necessary, and I have suggested that none of these examples yields the desired conclusion in a straightforward way. Necessarily, analysis of these cases has involved speculation about how Chomsky's assumptions might be modified so as to not require defective intervention effects, and while some of these speculations have a modicum of plausibility, e.g. the view of matching as a relation requiring that probe features be a subset of goal features, the suggestion that *derivationally* convergence is what matters for phases, others, most notably the idea that testing for convergence, by virtue of involving a bounded search, does not guarantee correct outcomes and that valuation should be regarded as 'phonological,' are undoubtedly more controversial.

---

[32] I should point out here that in this paper I have not attempted to defend the general proposition that once a feature has been deleted, it is invisible to *all* subsequent syntactic processes. Indeed, in adopting Chomsky's view that deleted ϕ-features of T remain available as potential goals for merged *there* and accepting some notion of phase, I have proposed analyses that are inconsistent with this view (while noting one major difficulty in n10). As far as phases go, the case for their conceptual attractiveness can stand, irrespective of whether this is supported by arguments relying on the 'marked for deletion' status of features. However, as Chomsky (1998, 13) has observed, conceptual observations must be supported by empirical argumentation if they are to carry any weight. It follows that even if the arguments presented here have any force, this will be only a small step towards establishing the more general proposition regarding the status of deleted elements.

Of the proposed modifications, the formulation of the subset requirement between probe and goal features in (3, 4) is perhaps the least tendentious. As observed already, Chomsky appears to have never formulated an explicit formulation on the matching relation, but it is necessary to acknowledge here that Pesetsky and Torrego (2000) do this, and to note that the condition on matching they propose appears to be *not* consistent with what has been assumed throughout this paper. Their condition (*ibid.*, 20) appears in (37):

(37) Match Condition

> If a head H enters an Agree relation with a *set of phrases* K, each syntactic feature of H must be present on some member of K (not necessarily with the same value, including value for EPP) (my italics – MA).

In a footnote, they observe (42): 'The word *some* is important, since a C bearing *u*Wh and *u*T may attract a phrase bearing T that lacks *wh* (e.g. TP) and may attract a phrase bearing *wh* that does not bear T (e.g. *how* or *with whom*) – in two distinct operations.' It would be inappropriate here to explore the full set of assumptions with which Pesetsky and Torrego are working. To appreciate the brief discussion that follows, it is merely necessary to note that they follow Chomsky in supposing that *wh*-movement is a consequence of an agreement relation obtaining between an uninterpretable feature of C (they refer to this as *u*Wh rather than Q) and an interpretable feature of a *wh*-item. However, they differ from Chomsky in their assumptions about the mechanisms involving nominative case. Specifically, they suppose that nominative case is an uninterpretable tense feature (*u*T) on D, and they further propose that C contains a token of this uninterpretable feature. Without any attempt at evaluation, let us simply apply these assumptions in the analysis of (38):

(38) what did Mary buy?

Pesetsky and Torrego follow Chomsky's analysis as far as TP is concerned. Using their notation in what should be a transparent way, at the point before C is merged, we have (39):

(39) [[Mary, ϕ, *u*T] [[T, ϕ] [buy [what, Wh]]]]

Here, there has been agreement between the ϕ-set of T and that of *Mary*, resulting in the former being 'marked for deletion,' and the raising of *Mary*. We next merge C to give (40):[33]

[33] The notation here is intended to represent the fact that , for Pesetsky and Torrego, EPP is a feature of a feature. Thus, there are two distinct EPP-features in (40), each requiring movement into the appropriate local domain in order to be satisfied.

(40) [C, ϕ, *u*T(+EPP), *u*Wh(+EPP)] [[Mary, ϕ, *u*T] [[T, ϕ] [buy [what, Wh]]]]

Now, consider the nature of the probe in C in (40). There are two uninterpretable features, and one option is to suggest that *collectively* they constitute a probe {*u*T(+EPP), *u*Wh(+EPP)}. If we see this probe as seeking a single goal, it is evident that there is nothing in the domain of C that satisfies the subset requirement. If, however, we contemplate the set union of T and *what*, we see that this includes {T, Wh}, so we can maintain that the subset relation is satisfied between our probe and the *set* comprising {T, *what*}. Thus, given (37), C is free to enter an Agree relation with this set and the consequences are deletion of *u*T in C with raising of T to C, and deletion of *u*Wh with raising of *what* to Spec, C.[34]

I would like to offer one observation with respect to this analysis. This concerns the issue of how we should identify a probe. We have already seen that Chomsky has proposed that a *set* of uninterpretable ϕ-features should be regarded as integral for this purpose (i.e. we should not regard [*u*person], [*u*number], etc. as constituting *independent* probes). But there doesn't seem to be any compelling reason to extend this thinking to include other uninterpretable features, and if we do not do this, the formulation of the Match Condition in (37) becomes unnecessary. In (40), there are *two* independent probes, *u*T(+EPP) and *u*Wh(+EPP). Each of these seeks a goal set of features that includes itself; the former identifies {T} in T (more strictly, TP. See n32) and the latter {Wh} in *what*. All other aspects of the analysis are preserved, and we do not have to jettison the condition on matching in (3, 4).[35]

Perhaps the most obvious conceptual reason for harbouring suspicions about the existence of defective intervention effects is their apparent non-existence in circumstances where they might be anticipated to play a role. To argue that the attempts in Chomsky (1999) to account for such 'defective' defective intervention effects are not sound, as I have tried to do in Section 3, is to argue implicitly against the existence of defective intervention effects themselves. If there are no defective intervention effects, there will be no circumstances in

---

[34] In order to account for why T rather than TP is raised, Pesetsky and Torrego formulate (p. 5) a definition of *distance* according to which TP and T are equidistant from C. They supplement this with a stipulation (p. 6) as part of what they refer to as the Head Movement Generalization. This stipulation is that when a head agrees with a feature of its complement, the *head* of the complement is raised rather than the complement itself.

[35] I believe that there are also some potentially serious empirical problems with (37) in the context of some of the analyses in Pesetsky and Torrego's paper, particularly their account of the *that*-omission asymmetry in sentential subjects. Interestingly, much of their discussion in the second half of the paper requires an interpretation of matching as a relation between a probe and a *single* object functioning as a goal. What my observation in the text illustrates is that even if this is the case, it does not necessarily prejudice their core analyses of subject-object asymmetries in matrix *wh*-questions. However, it would take another paper to examine this matter in detail.

which we need to account for their non-existence. If, furthermore, it proves possible to produce minimalist accounts of the structures underwriting these arguments without resorting to defective intervention at all, the case I have tried to make will be that much stronger. I have suggested that a juxtaposition of Chomsky's (1999) Alternative II along with careful consideration of the nature of Agree may be all that it necessary to deal adequately with the examples in question, and thereby avoid the complex considerations concerning matching that Chomsky's own position seems to demand.

References

Atkinson, M. 2000, Uninterpretable feature deletion and phases. *Essex Research Reports in Linguistics*: 34.

Chomsky , N. 1995, *The Minimalist Programme*. Cambridge, Mass.: MIT Press.

Chomsky, N., 1998, Minimalist inquiries: the framework. *MIT Occasional Papers in Linguistics*, 15. MIT, Department of Linguistics.

Chomsky, N., 1999, Derivation by Phase. *MIT Working Papers in Linguistics*, 18. MIT, Department of Linguistics.

Chomsky, N., 2001, Beyond Explanatory Adequacy, Ms. MIT.

Epstein, S and D. Seeley, 1999, SPEC-ifying the GF "subject," eliminating A-Chains and the EPP with a derivational model. Ms. University of Michigan.

Haraiwa, K., 2001, Multiple Agree and the defective intervention constraint in Japanese. In Proceedings of HUMIT. To appear in *MIT Working Papers in Linguistics.*

Johnson, D. and S. Lappin 1997, A critique of the Minimalist Program. *Linguistics and Philosophy* 21: 273-333.

Pesetsky, D and E. Torrego, 2000, T-to-C movement: causes and consequences. Ms. MIT and UMass.

Uriagereka, J. n.d. Commentary on Chomsky's 'Derivation by Phase.' Available at http://www.ling.umd.edu/Courses/Ling819/papers

Author's Address:

Department of Language & Linguistics
University of Essex
Colchester CO4 3SQ
United Kingdom

Email: *matkin@essex.ac.uk*