4.10 Functional Categories and Formal Features

What precedes substantially revises the framework developed in chapters 1–3. But we have not yet subjected functional categories to the same minimalist critique. In the final section I would like to explore this question, a course that leads to another fairly substantial modification. Even more than before, I will speculate rather freely. The issues that arise are fundamental to the nature of CHL, having to do with the formal features that advance the computation (primarily strength, which drives overt operations that are reflected at the A-P interface) and the functional categories that consist primarily (sometimes entirely) of such features.

4.10.1 The Status of Agr

Functional categories have a central place in the conception of language we are investigating, primarily because of their presumed role in feature checking, which is what drives Attract/Move. We have considered four functional categories: T, C, D, and Agr. The first three have Interpretable features, providing “instructions” at either or both interface levels. Agr does not; it consists of –Interpretable formal features only. We therefore have fairly direct evidence from interface relations about T, C, and D, but not Agr. Unlike the other functional categories, Agr is present only for theory-internal reasons. We should therefore look more closely at two questions.

(181) a. Where does Agr appear?
    b. What is the feature constitution of Agr?

In section 4.2.2 we tentatively assumed that Agr lacks $\phi$-features, just as it (fairly clearly) lacks an independent Case-assigning feature, that being provided by the V or T that adjoins to it. If Agr indeed lacks $\phi$-features as well, we would expect that the $\phi$-features of a predicate Pred (verb or adjective) are added to Pred (optionally) as it is selected from the lexicon for the numeration. We had little warrant for the assumption about $\phi$-features, and so far it has had little effect on the analysis. But it becomes relevant as we attempt more careful answers to the questions of (181). I will continue to assume that the original assumption was correct, returning to the question at the end, after having narrowed significantly the range of considerations relevant to it.

We have evidence bearing on question (181a) when Agr is strong, so that the position is phonetically indicated by the overt categories that raise to it: V and T by adjunction, DP by substitution in [Spec, Agr]. The richest example is an MSC with object raising, as in the Icelandic TEC construction (160).
Here three pre-VP positions are required within IP for nominal expressions: expletive, subject, and object. One position is provided by T. We therefore have evidence for two noninterpretable functional categories, the ones we have been calling Agr (Agr$_S$ and Agr$_O$). In MSCs, Agr$_S$ is strong, providing a specifier and a position for V-raising above the domain of strong T: in effect, a "double EPP" configuration. Another VP-external position is provided between T and VP by strong Agr$_O$. That is the basic rationale behind the analyses just outlined. It accords with the general minimalist outlook, but the anomalous status of Agr raises questions.

The background issues have to do with the strong features of T and Agr, and what appears in the overt specifier positions they make available. In the I position, preceding all verb phrases (main or auxiliary), we have postulated two functional categories: T and Agr$_S$. In MSCs the specifier position of each is nominal, DP or NP; hence, the strong feature must at least be [nominal-], meaning satisfied by the nominal categorial feature [D] or [N]. At most one nominal can have its Case and ϕ-features checked in this position, which suggests that one of the two nominals must be the pure expletive Exp, a DP. Let us assume this to be the case, though it is not as yet established. The observed order is Exp-nominal rather than nominal-Exp, a fact yet to be explained.

The best case is for Agr$_O$ to have the same constitution as Agr$_S$. Since Agr$_S$ allows and perhaps requires a D-feature, the same should be true of Agr$_O$. Hence, both Agrs attract DPs: nominals that are definite or specific. As noted, it follows that expletive-associate constructions observe the definiteness effect and that object raising is restricted to definite (or specific) nominals. This is close enough to accurate to suggest that something of the sort may be happening.

Recall, however, that the definiteness effect for object raising is at best a strong tendency, and that for expletive constructions its status is unclear. It does not rule out any derivations, but rather describes how legitimate outputs are to be interpreted: either as expletive constructions with at most weakly existential implicatures, or as list constructions with strong existential interpretation (see notes 42, 44). We therefore have no strong reason to suppose that the associate cannot be a DP—only that if it is, a certain kind of interpretation is expected.

With strong features, Agr provides a position for T- or V-raising (adjunction) and DP-raising (substitution), so there is evidence that it appears in the numeration. If Agr has no strong feature, then PF considerations, at least, give no reason for it to be present at all, and LF considerations do not seem relevant. That suggests an answer to question (181a): Agr exists only when it has strong
features. Agr is nothing more than an indication of a position that must be occupied at once by overt operations.\textsuperscript{128} Substitution can be effected by Merge or Move. If by Merge, it is limited to expletives, for reasons already discussed.

Pursuit of question (181b) leads to a similar conclusion. The function of Agr is to provide a structural configuration in which features can be checked: Case and $\phi$-features, and categorial features ([V-] and [T-] by adjunction, [D-] by substitution). The Case-assigning feature is intrinsic to the heads (V, T) that raise to Agr for checking of DP in [Spec, Agr], so there is no reason to assign it to Agr as well. With regard to $\phi$-features, as already discussed, the matter is much less clear. If Agr has $\phi$-features, they are –Interpretable, but there might be empirical effects anyway, as noted earlier. Continuing tentatively to assume that $\phi$-features are (optionally) assigned to lexical items as they are drawn from the lexicon, we conclude that Agr consists only of the strong features that force raising.

Certain problems that arose in earlier versions now disappear. There is no need to deal with optionally strong Agr, or with the difference in strength of Agr$_S$ and Agr$_O$. Since Agr is strong, the first problem is just a matter of optional selection of an element (strength of F) from the lexicon for the numeration, the irreducible minimum; and difference in strength is inexpressible. There still remains, however, a conflict between the $\theta$-theoretic principle that transitive verbs have a $v$–VP structure and the assumption that overt object raising is internal to this construction (see note 81 and p. 316).

Let us turn to the properties that remain.

Since Agr consists solely of strong features, it cannot attract covert raising.\textsuperscript{129} We have so far assumed that Subj (subject) and Obj (object) raise to the checking domain of Agr, entering into a checking relation with features of T or V adjoined to Agr (technically, adjoined within Agr$_{\text{max}}$, the X$^0$ projection headed by Agr). But with weak Agr gone, covert raising must target T and V directly.\textsuperscript{130}

There is now no reason to postulate Agr$_O$ unless it induces overt raising of DP to [Spec, Agr$_O$]. What about Agr$_S$? It appears in MSCs, but lacks independent motivation elsewhere, as matters now stand. For languages of the French-English type, then, Agr is not in the lexicon (unless MSCs appear marginally, with extraposition). Agr therefore occurs in highly restricted ways.

The next question is to inquire into the justification for Agr with strong features. Let us first look at Agr$_O$, then turn to Agr$_S$.

We restrict attention now to transitive verb constructions, which we continue to assume to be of the form (182), ignoring [Spec, V] (the case of a complex internal domain).
V raises overtly to the light verb \( v \), forming the complex \( Vb = [\_, V 
abla] \). Assuming unergatives to be concealed accusatives, the only other VP construction is that of unaccusatives lacking the \( v \)-shell, not relevant here.

Suppose that a derivation has formed (182) and Agr is merged with it. Agr is a collection of strong features, either \([D-]\) or \([V-]\) or both. As noted, we need not postulate Agr\(_O\) except for object raising; it does not consist only of strong \([V-]\). Holmberg’s generalization states, in effect, that it cannot be just strong \([D-]\). Let us tentatively assume, then, that Agr\(_O\) is \{strong \([D-]\), strong \([V-]\)\}. The effect of adding Agr\(_O\) is to compel overt raising of DP to [Spec, Agr] and of Vb to Agr.

Consider the first property. There is a simple way to force overt DP-raising without the functional category Agr: namely, by adding to \( v \) itself a strong D-feature (or perhaps, the more neutral strong \([\text{nominal-}]\) feature) that requires overt substitution in the “outer Spec” of a multiple-Spec configuration. If Obj raises to this position to form a chain (Obj, \( i \)), it will be in the checking domain of V and therefore able to check its Case and (object agreement) \( \phi \)-features. Recall that Subj inserted by Merge in [Spec, \( v \)] is not in the checking domain of \( v \), because it does not head a nontrivial chain.\(^{131}\)

Object raising, then, takes place when the light verb \( v \) that heads the transitive verb construction (182) is assigned the strong feature as it is drawn from the lexicon and placed in the numeration; see section 4.2.2. The choice is arbitrary, forced, or unavailable as the language has optional, obligatory, or no overt object raising, respectively. Since Subj is not in the checking domain, as just noted, it does not check this strong feature, so an outer Spec must be constructed for that purpose. One way is raising of Obj; I hope to show that all others are excluded.

Suppose that an adverbial phrase Adv is adjoined to \( v^{\text{max}} \) and object raising crosses it, yielding the construction Obj–Adv–\( v^{\text{max}} \). That provides no reason to postulate an Agr position outside of \( v^{\text{max}} \); a strong feature need only be satisfied before a distinct higher category is created.\(^{132}\)
Overt object raising therefore seems to provide no compelling reason for assuming the existence of AgrO. The other property of AgrO is that it forces overt V-raising—actually to T outside of VP, so the effects are never directly visible. The motivation was theory-internal, but it disappears within the more restricted framework, as we will see. The property was a crucial part of the expression of Holmberg’s generalization that object raising is contingent on V-raising, but to introduce that consideration to justify postulation of AgrO is circular. For VP, at least, it seems that we should dispense with AgrO.

Consider adjectival constructions such as (55), repeated here. 

\[(183) \text{John is } [_{\text{AgrP}} t_1 \text{ Agr } [_{\text{AP}} t_2 \text{ intelligent}]]\]

We assumed that John is merged in the position of t2 in AP as [Subj, Adj] (subject of intelligent, in this case), raising to [Spec, Agr] for DP-adjective agreement, then on to matrix [Spec, I] for DP-verb agreement. Do we need a strong functional category (Agr) here to head the small clause complement of the copula? Assuming that [Subj, Adj] is analogous to specifier or complement of V—and more generally, that the complementarity of θ-theory and checking theory holds in this case as well as others—then agreement will not be checked in this position of merger. We assumed that the (– Interpretable) ϕ-features of the adjective Adj are checked by overt raising of its subject Subj to [Spec, Agr] and of Adj to Agr—the latter problematic, as mentioned, because Agr is weak in English (see note 51). We can now avoid that problem by eliminating Agr and adopting the analysis just proposed for overt object raising: Adj is assigned the feature strong [nominal-] as it is drawn from the lexicon, and [Subj, Adj] raises to the outer Spec required by the strong feature, entering the checking domain of Adj. In this case the derivation will converge only if the strong feature is selected, so the choice is in effect obligatory. Note that features of Subj cannot adjoin covertly to the Adj, as a review of the possible cases shows (on plausible assumptions).

We therefore eliminate AgrO in this case too, using simple mechanisms and overcoming an earlier problem about unexpected head raising. The structure of predicate adjectival constructions is not (183) but rather (184).

\[(184) \text{John is } [_{\text{AP}} t_1 \text{ } \{_{\text{A'}} t_2 \text{ intelligent} \}]]\]

For small clauses, we have something like the original assumptions of Stowell (1978) on which much of the work on the topic has been based, but consistent with other assumptions that we are now adopting.

If all of this turns out to be correct, we can eliminate AgrO from the lexical inventory entirely, for any language. Turning to Agrs, we need to consider only
MSCs, which have the surface order [Exp–V–Subj]. Our assumption so far is that the subject Subj is in [Spec, T] and the expletive in [Spec, AgrS], and that V has raised to AgrS. Suppose, instead, we follow the line of reasoning suggested for AgrO, eliminating Agr and adding an optional strong feature that assigns an outer Spec to T. The situation differs from the case of AgrO. [Spec, v] in (182) is required for independent θ-related reasons, so only one new Spec is required for object raising. In contrast, T requires no Spec, so we have to accommodate two Specs that are induced only by feature strength. Independently, we have to account for the fact that the order is not the expected (185a) but rather (185b), along with other observed properties.

(185) a. Exp [Subj [T^\text{ominax} XP]]
   b. Exp T^\text{ominax} Subj XP

MSCs appear only when the EPP holds. The question of their nature arises, then, only when T already has a strong [nominal-] feature, which is deleted when checked by DP or NP in [Spec, T]. Suppose that the derivation has reached the stage TP with T strong, and the numeration contains an unused expletive Exp. Then Exp can be inserted by Merge to satisfy the EPP, and we have an ordinary expletive-associate construction. The strong feature of T deletes and furthermore erases, since the derivation converges. Hence, overt MSCs exist only if T has a parameterized property of the kind discussed earlier (see below (58)), which allows a –Interpretable feature (in this case, the strong [nominal-] feature) to escape erasure when checked. If the option is selected, then there must be a multiple-Spec construction, with n + 1 specifiers if the option is exercised n times. In a language with the EPP but no MSCs, the strong nominal feature of T is introduced into the derivation with n = 0, hence erased when checked. In Icelandic, the descriptive facts indicate that n = 0 or n = 1; in the latter case, T has two Specs.

Let us see where this course leads, eliminating Agr from UG entirely—and, at least for our purposes here, keeping to functional categories with intrinsic properties that are manifested at the interface levels. The questions that arise are again rather delicate. Let us delay a direct investigation of them until some further groundwork is laid.

4.10.2 Core Concepts Reconsidered
To accommodate the change from an Agr-based to a multiple-Spec theory, we have to simplify the notions of equidistance and closeness that entered into the definition of Attract/Move. These were expressed in the principle (87), repeated here.
(186) β is closer to K than α if β c-commands α and is not in the minimal
domain of CH, where CH is the chain headed by γ, γ adjoined within
the zero-level projection H(K)^0_{max}.

But this no longer works: with the elimination of intervening heads, minimal
domains collapse. We therefore have to exclude nontrivial chains from the
account of equidistance, relying instead on the much more differentiated
analysis of features now available and the immobility of traces—that is, on
the fact that only the head of a chain can be “seen” by K seeking the closest
α to attract.

In the earlier formulation, the basic case is (85), repeated here in the more
general form (187) to accommodate adjunction of α to X as well as substitu-
tion of α in [Speck, X] (where X may already be the head of a complex zero-
level projection).

(187)  

When α raises, targeting X^{max}, it creates a new position τ(X), which may either
be [Spec, X] or adjoined to [Y–X] (= X^{0_{max}}); call τ(X) the target in either case.
The minimal domain of the chain CH = (Y, t) includes Spec_1 and ZP along
with τ(X) formed by raising of α, which is within ZP or is ZP. Crucially, Spec_1
is within the “neighborhood of X” that is ignored in determining whether α
is close enough to be attracted by X (technically, by its projection). That
assumption was necessary in order to allow α to cross Spec_1 to reach τ(X). In
a transitive verb construction, for example, it was assumed that X = Agr, Spec_1
= Subj, Y is the verbal element that adjoins to Agr, and Obj is within its ZP
complement. Obj has to raise to the checking domain of Agr for feature check-
ing either overtly or covertly, requiring that it be “as close” to the target as
Spec_1.

Most of this is now beside the point. We have eliminated Agr and its projec-
tion from the inventory of elements. For the case of overt object raising, the
structure formed is no longer (187) with X = Agr and τ(X)= [Spec, Agr], but
(188), with an extra Spec in YP.
Chapter 4

(188) 

\[
\begin{array}{c}
\text{Spec}_2 \\
\text{Spec}_1 \\
\text{Vb} \\
\text{YP} \\
\text{ZP}
\end{array}
\]

Vb is the verbal element (or its trace, if the complex has raised further to adjoin to T); Y' and YP are projections of the light verb v to which V has adjoined to form Vb; ZP = [t_v Obj], t_v the trace of V; and Spec_2 is the target τ(v) created by the raising operation. Spec_1 is Subj, and it is only necessary that it be no closer to the target Spec_2 than α in ZP. For this purpose, it suffices to simplify (186), keeping just to the trivial chain CH = H(K) (the head of K) and its minimal domain. We therefore restate (186) as (189).

(189) γ and β are equidistant from α if γ and β are in the same minimal domain.

Hence, γ = Spec_2 and β = Spec_1 are equidistant from α = Obj in the illustrative example just discussed.

We now define "close" for Attract/Move in the obvious way: if β c-commands α and τ is the target of raising, then

(190) β is closer to K than α unless β is in the same minimal domain as (a) τ or (b) α.

We thus have two cases to consider. We ask (case (190a)) whether β and τ are equidistant from α, and (case (190b)) whether β and α are equidistant from τ. If either is true, then β does not bar raising of α to τ. In case (190a), β and τ are in the minimal domain of H(K); and in case (190b), β and α are in the minimal domain of h, for some head h. In case (190a), β is in the "neighborhood" of H(K) that is ignored, in the sense of earlier exposition.

By case (190a), Obj within ZP in (188) is close enough to be attracted by Y' (= YP, at this point), since Spec_1 is in the minimal domain of H(Y') and is therefore not closer to Y' than Obj; Spec_1 and Spec_2 (=τ) are equidistant from Obj. Therefore, either Subj in Spec_1 or Obj (in ZP) can raise to the new outer Spec, Spec_2, required by the strong feature of v. Both Obj and Subj must raise for Case checking, and something must raise to check the Case feature of T (or of some higher category if T is a raising infinitival, as already discussed). By case (190b), overt object raising to Spec_2 does not prevent subject raising.
from Spec₁, because Spec₂ and Spec₁ are equidistant from any higher target; both are in the minimal domain of v. How about direct raising of Obj from within ZP, targeting T, crossing Subj and Spec₁? That is barred by the MLC, since Subj and Obj are not equidistant from T, given the v–VP analysis of transitives; they are in different minimal domains. We will return to a closer analysis, reviewing other options skirted here.

Consider the following counterargument. Suppose the language has the EPP and optional object raising: T requires [Spec, T] and v permits an outer Spec, Spec₂, beyond Subj in Spec₁ (both overt). Suppose that Obj raises to [Spec₂, v], then raises again to [Spec, T], satisfying the EPP. That much is permitted. Subj and T have not had Case features checked, but that can be overcome by covert raising of Subj, targeting T, which is also permitted. So the derivation converges, incorrectly. But this derivation is blocked by economy conditions. It involves three raising operations, and two would suffice for convergence: object raising followed by subject raising to [Spec, T] (in both cases, with two violations of Procrastinate, the minimal number with two strong features). So the unwanted series of steps, though permitted, is barred by economy considerations: shorter derivations block longer ones.

The computation is local: after raising the object, we choose the operation that will lead to the shortest convergent derivation: raising of Subj to [Spec, T]. We also have empirical support for the tentative assumption made earlier that shorter derivations, locally determined in this sense, block longer ones (see discussion of (114)).

Note that we have lost Holmberg’s generalization and other effects of V-raising on extension of chains; that is a consequence of excluding chains from the definition of “closeness.” Such generalizations, if valid, would now have to be stated in terms of a property of Vb in (188): it can have a second outer Spec only if it is a trace. There is no obvious reason why this should be so.

In any event, the earlier, more complex definition of equidistance and closeness is not necessary and in fact not possible. The notion of equidistance may still be needed, as in cases just reviewed and others, but it has narrower scope.

The conclusion that equidistance is still needed relies on a tacit assumption that could be challenged: that the strong feature of v must be satisfied by the outer Spec, Spec₂ of (188), not the inner Spec, Spec₁. All we know, however, is that some Spec of v is motivated by considerations of θ-theory (to host the external argument) and is therefore independent of the strength of v; the other Spec is present only to check the strength feature. But both Specs are within the minimal domain of v, so either is available for θ-marking of the external
argument of a transitive verb. Suppose we allow this possibility, so that the outer Spec can host the external argument. In that case we can drop the notion of equidistance entirely, simplifying (190) to the statement that $\beta$ is closer to the target K than $\alpha$ if $\beta$ c-commands $\alpha$. It follows, then, that Obj can only raise to the inner Spec, Spec₁ of (188), to check the strength feature and undergo overt Case marking. If overt object raising takes place, then Subj will be merged in the outer Spec to receive the external $\theta$-role provided by the configuration. With “closer than” restricted to c-command, only Subj in the outer Spec can be attracted by T (note that Subj always has features that will check sublabels of T). Therefore, Obj is frozen in place after overt object raising, and the conclusions reached above follow directly.¹³⁴

On these assumptions, it follows that Subj always c-commands Obj within IP. In particular, this is true in expletive constructions, whether or not object raising has taken place; that appears to be generally the case, with some unexplained exceptions (see Jonas and Bobaljik 1993). We also have a somewhat more natural account of agreement, with the inner Spec uniformly entering into the relation (the Spec $\theta$-position is not subject to it for reasons already discussed). It also should be the case that only the inner Spec in a multiple-Spec construction can be a binder (assuming that locality enters crucially into binding, in one of several possible ways) though control may be more free, as it often is (see section 1.4.2); that appears to be the case (Hiroyuki Ura, personal communication). Further questions arise when we turn to verbs with complex internal argument structure, which I am ignoring here.

Let us keep available these two options for the notion “closer than,” noting however that the one just sketched is simpler, hence to be preferred if tenable (and certainly if empirically supported). I will present the examples below on the assumption that object raising is to the outer Spec to ensure that the required consequences follow even under this more complex alternative; it is easy to check that the arguments run through (more simply, in some cases) if closeness reduces to c-command so that object raising is to the inner Spec and only the outer Spec is attracted by T.

We also have to settle some questions about adjunction that have been left open but become more prominent in this much more restrictive framework, covert adjunction being the most interesting case. Empirical evidence for covert operations and the structures they yield is harder to obtain than for their overt counterparts, but it exists, and conceptual arguments also carry us some distance, at least.

One reasonable guiding idea is that interpretive operations at the interface should be as simple as possible. Barring empirical evidence to the contrary,
we assume that the external systems are impoverished—a natural extension of minimalist intuitions to the language faculty more broadly, including the systems (possibly dedicated to language) at the “other side” of the interface. That means that the forms that reach the LF level must be as similar as typological variation permits—unique, if that is possible. These assumptions about the interface impose fairly restrictive conditions on application and ordering of operations, cutting down the variety of computation, always a welcome result for reasons already discussed. At the A-P interface, overt manifestation provides additional evidence. Such evidence is largely unavailable at the C-I interface, but the general conceptual considerations just reviewed carry some weight. We have been implicitly relying on them throughout: for example, to conclude that covert features adjoin to the head of a chain (specifically, a raised verb), not to the trace or optionally to either; see (94) and discussion.

The central problem about covert adjunction concerns the structure of $T^{0\text{max}}$ at LF. Consider first the richest case: a TEC with object raising (Icelandic). Putting aside the observed position of $T^{0\text{max}}$, we assume its form at LF to be (185a), repeated here, with YP an instance of (188).

\[(191) \text{Exp} [\text{Subj} [T^{0\text{max}} \text{YP}]]\]

Exp and Subj are specifiers of the T head of $T^{0\text{max}}$, which is formed by adjunction to T of Vb = $\left[\upsilon V \upsilon\right]$. In this case $T^{0\text{max}}$ is (192) and the complement YP is (193).

\[(192) \begin{array}{c} T \\ Vb & T \end{array}\]

\[(193) \begin{array}{c} & \upsilon^{\text{max}} \\ \text{Obj} & \upsilon' \\ & \upsilon' \\ & t_{\text{subj}} \\ Vb & VP \\ & t_V & t_{\text{Obj}} \end{array}\]

Here Obj and $t_{\text{subj}}$ are specifiers of the $\upsilon$ head of Vb.
Suppose V raises overtly and Obj does not (French, or optionally Icelandic). The complement of T differs from (193) in that it lacks the outer Spec occupied by Obj, which remains in the position \( t_{\text{Obj}} \). The formal features FF(Obj) raise to \( T^{0_{\text{max}}} \) for feature checking. Before this covert operation takes place, \( T^{0_{\text{max}}} \) is again (192). To maximize similarity to the LF output (192), FF(Obj) must adjoin to the complex form (192) itself, forming (194), not to the deeply embedded V within Vb, which actually contains the relevant features for checking.

\[
\begin{tikzpicture}
  \node {T} child {node {FF(Obj)} child {node {T}} child {node {Vb} edge from parent[draw=none]}};
\end{tikzpicture}
\]

(194)

The operation is permitted, since the features of V are sublabels of the target of adjunction; and it satisfies the conditions on “closest target” (\( t_{\text{Subj}} \) being “invisible,” as discussed) and formation of complex \( X^0 \)'s. Assuming this to be the general pattern, we conclude that adjunction is always to the maximal zero-level projection \( X^{0_{\text{max}}} \), never “internally” to one of its constituents—the simplest assumption in any event.

Consider an English-type language with overt raising of Subj but not V or Obj. To achieve maximum impoverishment at the interface, we want \( T^{0_{\text{max}}} \) at LF to be as similar as possible to (194)—in fact, identical except that in place of Vb it has FF(Vb), since V-raising is covert. FF(Obj) therefore cannot raise to V or to the verbal complex Vb before Vb adjoins to T; if it did, the structure formed would be quite different from (194). After covert V-raising, FF(Obj) adjoins to \( T^{0_{\text{max}}} \), again forming (194) at LF (with FF(Vb) in place of Vb). The ordering is forced by bare output conditions, if the conjecture about poverty of interface interpretation is correct.

Suppose the language lacks overt raising of either Subj or Obj, so that both FF(Subj) and FF(Obj) raise covertly to TP. The poverty-of-interpretation conjecture requires that Vb raise to T before either subject or object raising; we thus have (192) once again, as desired, whether V-raising is overt (as in a VSO language) or not (in which case Vb in (192) is replaced by FF(Vb)). TP now attracts Subj, which is closer than Obj, forming a structure identical to (194), except that the features FF of (188) happen to be FF(Subj) rather than FF(Obj). But it is also necessary for FF(Obj) to raise. That is now possible, since the trace of Subj (unlike Subj itself) is not closer to TP, being inaccessible to Attract/Move. \( T^{0_{\text{max}}} \) therefore ends up as (195).
Recall that in a normal expletive construction, the strong D-feature of T is satisfied by an expletive rather than the raised Subj—the case of Merge that was assimilated to Attract/Move (see end of section 4.5). In this case T<sub>0 max</sub> is again (195), and [Spec, T] is occupied by the expletive. An expletive construction thus has a certain structural resemblance to a VSO configuration, as has been implicit throughout.

We conclude that adjunction is to the maximal X<sub>0</sub> projection X<sub>0 max</sub> and that heads raise before elements of their domains, conditions reminiscent of cyclicity. These are descriptive generalizations that we derive, not principles that we stipulate: they follow from the minimalist principle of poverty of interpretation at the interface. Simple and plausible assumptions suffice to guarantee virtually the same LF form for T<sub>0 max</sub> over the typological range we are considering. So far these conclusions are motivated only by the conceptual requirement of maximizing uniformity of LF outputs. They supplement the earlier conclusion that traces of A-movement never enter into Attract/Move, whether overt or covert.

Suppose that output conditions at the LF interface rule out (195) under the rather natural requirement that FF(Subj) must c-command FF(Obj) in T<sub>0 max</sub> if both are present. Hence, Obj has to adjoin to T<sub>0 max</sub> before Subj does. But that is impossible, since Subj is closer to T than Obj if both remain in situ. It follows, then, that at least one of Subj and Obj must raise overtly if the expression is to converge, a hypothesis that has been advanced several times. The requirement on T<sub>0 max</sub> generalizes the conclusion that Subj must c-command Obj overtly, which follows from the simplification of the notion “closer than” to just c-command, as already discussed.

We also have to settle some questions about the positions in which expletives can appear. The problems are not specific to this analysis; they arose before (more broadly, in fact), but were ignored. The basic descriptive fact is (196).

(196) Exp can only be in [Spec, T].
We have to determine why this is the case, a question that has a number of facets.

Suppose that Exp is merged in a θ-position, one of the possible violations of (196). That leads to a violation of the θ-Criterion, hence a deviant expression, a conclusion that suffices to allow us to dismiss the option here. Still, a factual question that has arisen several times before remains: is this a case of convergent gibberish or of a crashed derivation?

Suppose, say, that Exp is merged as subject of a transitive verb construction, yielding the VP (197), an instance of (182).

\[
\begin{array}{c}
\text{Exp} \\
\text{v'} \\
\text{v} \\
\text{VP} \\
\text{saw someone}
\end{array}
\]

We next merge strong T, which attracts Exp, yielding finally the overt form (198).

\[
\text{(198) there saw someone}
\]

Raising of Exp satisfies the strong feature of T (EPP), but not its Case feature. Covertly, FF(Obj) adjoins to T, providing a Case feature. But there are two heads that have to check a Case feature: T (nominative) and see (accusative). Hence, the derivation will crash, with one or the other not satisfied. Such a derivation could converge only if the language had a verb SEE, like see except that it assigns no Case. Since no issue of blocking convergent derivations seems to arise, we are left without a satisfactory answer to the question of the status of the analogue of (198) with SEE in place of see. It could be that the derivation crashes because of a violation of the θ-Criterion, so there can be no such verb as SEE; or it could be that the derivation converges and a language can have such a verb (perhaps English in fact does), though it appears only in deviant expressions violating the θ-Criterion. We have seen that an argument lacking a θ-role is illegitimate, violating FI and causing the derivation to crash; but the question of assignment of θ-role remains open.

Putting the option aside as either nonconvergent or gibberish, we consider only Exp merged in a non-θ-position. We thus keep to the earlier conclusion that the only position that Exp can enter either by Merge or Attract/Move
is one induced by a strong [nominal-] feature, hence [Spec, T] or [Spec₂, v] (the outer Spec) in (188). We also know that nothing can be raised to a θ-position. Hence, Exp never appears in a θ-position in the course of a derivation.

Can Exp be merged in the outer Spec of v, [Spec₂, v] in (188)? There are two cases to consider: Exp remains in this position, or it raises to [Spec, T]. The latter possibility is excluded, because the effect at both PF and LF is the same as merging Exp in [Spec, T] in the first place. Therefore, the economy principle (76) prevents selection of the strong [nominal-] feature of v for the numeration. The only remaining case, then, is merger of Exp in [Spec, v], where it remains at LF.

At LF, Exp is simply the categorial feature [D]. Any phonetic features would have been stripped away at Spell-Out, and we have seen that Exp has no other formal features. Lacking semantic features, Exp has to be eliminated at LF if the derivation is to converge: its D-feature is – Interpretable and must be deleted by some operation. Therefore, [D] must enter into a checking relation with some appropriate feature F. As we have just seen, T does not offer a checking domain to which Exp can raise, so F must raise to the checking domain of Exp, which means that F must adjoin to it. What is F? Independently, there is good reason to believe that the categorial feature [N] adjoins to [D] regularly, namely, in the D–NP construction (Longobardi 1994). The optimal assumption, then, would be that it is adjunction of the feature [N] to Exp that eliminates its (sole remaining) feature. The feature [D] of Exp cannot be erased in this configuration: that would eliminate the category completely, leaving us with an illegitimate syntactic object. Therefore, checking of the categorial feature of Exp (its entire content) deletes it but does not go on to erase it, by the general principles already discussed.

The optimal assumption requires nothing new. Exp has no complement, but in the relevant formal respects, the head-complement relation that allows N-raising to D is the same as the [Spec-α] relation holding between Exp in Spec and the X′ projection of T. In a properly formed expletive construction, the formal features FF(A) of the associate A adjoin to matrix I (which we now take to be T⁰ max), checking Case and agreement and allowing matrix-type binding and control. The categorial feature [N] of A comes along as a free rider and is therefore in the right position to adjoin to Exp, forming [D N Exp]. The configuration so formed places [D] in a checking configuration with raised [N], as in the D-complement structure. Like D that takes a complement, expletive D has a strong [nominal-] feature, which attracts [N] — a residue of the earlier adjunction-to-expletive analysis.
Returning to (196), we recall that the sole remaining problem was to show that Exp does not appear at LF in Spec$_2$ of (188)—the object-raising position. We have to show, then, that if Exp is in that position, no [N] can raise to it.\footnote{137}

There is only one possibility for N-raising to Spec$_2$: namely, the categorial feature [N] of Subj in Spec$_1$.\footnote{138} The operation is permitted, so we have to show that the derivation will crash for other reasons. We already have such reasons. Subj must raise to the checking domain of T, leaving a trace in Spec$_1$. We concluded earlier (see (94)) that traces of A-movement are inaccessible to Attract/Move. Therefore, if raising of FF(Subj) has taken place (either as part of overt substitution in [Spec, T] or as part of covert adjunction), [N] cannot raise from the trace of FF(Subj) to adjoin to Spec$_2$. The only remaining possibility is that Subj remains in situ in Spec$_1$ with Exp in Spec$_2$ and [N] raises from Subj, adjoining to [D] in Spec$_1$ automatically carrying along FF[N]. But FF(Subj) must now raise to T$^\text{max}$, raising the trace of FF[N], an operation that should also be barred by (94). We see that there is good reason to interpret that constraint (which was purposely left a bit vague) quite strictly. If so, the restriction of Exp to [Spec, T] is explained (with the qualifications of the preceding notes).

These observations suggest that a still stricter interpretation of (94) might be warranted, strengthening the condition on argument chains, repeated as (199), to the provision (200).

(199) Only the head of CH can be attracted by K.
(200) $\alpha$ can be attracted by K only if it contains no trace.

The suggestion would have been unacceptable in earlier versions for a variety of reasons, but the objections largely dissipate if covert movement is only feature raising and some earlier suggestions have merit.\footnote{139} One immediate consequence is that overt countercyclic operations of the kind that motivated the extension condition are ruled out (see discussion of (138)). Nevertheless, (200) may be too strong even within the core computational system; we will return to this.

We see, then, that the descriptive observation (196) is well established on reasonable grounds.\footnote{140} It is possible that something similar is true of nonpure expletives of the it-type, which are associated with CPs with complementizers: that, for, or Q (the phenomenon of extraposition, however it is analyzed). They do not appear with control or raising infinitivals.

(201) a. *it is certain [PRO to be intelligent]
   b. *it is certain [John to seem t to be intelligent]
Possibly the overt complementizer head of the extraposed associate raises to the expletive, deleting it as in pure expletive constructions and thus satisfying FI. But see note 68.

The analysis of (196) raises in a sharper form an unsettled problem lingering from before. In discussing (168a–b)—essentially (202a–b)—we observed that it is not clear why both are allowed (assuming that they are, when MSCs are permitted).

(202) a. there seems [TP [someone] t [TP t_{Subj} to be in the room]]
   b. there seems [TP t_{Exp} [TP someone to be in the room]]

To explain the notations, TP in both cases is the complement of seem; t is the trace of seem; t_{Subj} is the trace of someone; and t_{Exp} is the trace of there. Exp occupies the outer Spec of the matrix MSC in (202a), and its trace occupies the outer Spec of the embedded MSC in (202b). The Subj someone is in [Spec, T] in both cases, in the matrix clause in (202a) and in the embedded clause in (202b). Thus, the matrix clause is an MSC in (202a) and the embedded clause is an MSC in (202b).

The earlier discussion entailed that (202b) blocks (202a) because at the common stage of the derivation when only the most deeply embedded T projection has been formed, insertion of Exp is more economical than raising of Subj to its Spec. The only permissible alternative to (202b), then, should be (203), with Subj remaining in the unraised associate position.

(203) there seems [IP t to be [someone in the room]]

We now have a further problem. We have just seen that the construction (204), which is rather similar to (202b), is not permitted by virtue of the economy principle (76).

(204) Exp T ... [IP t Spec_{2} [Vb XP']]

Here raising of Exp from the Spec determined by the strong feature of v is barred: the strong feature cannot appear in the numeration because it has no effect on PF or LF output. In (202b), however, raising of Exp from the extra Spec determined by the strong feature of embedded T is permitted, even though it seems to be barred by this condition and, in case (202a), by yet another economy condition: that the most economical step must be taken at each point in the derivation.

The problems unravel in the present framework. By the economy condition (76), a strong feature can enter the numeration if it has an effect on output—in this case, PF output, because only a pure expletive is involved. That suffices to bar (204): adding the strong feature to the v head of Vb has no PF effect. Turning to (202), we see that (202a) derives from adding an extra strong
feature to matrix T, and (202b) from adding an extra strong feature to embedded T—two different elements. In each case there is an effect on PF output. Suppose matrix T enters the numeration without a strong feature that allows an extra T. Embedded T is a different element in the numeration: if it lacks the strong feature that allows an extra Spec, then we derive (203); if it has this strong feature, we derive the distinct PF form (202b). Therefore, (76) is inapplicable. The same is true of the economy principle that forced selection of expletive over raising in the earlier theory. That selection is forced only if the derivation converges; and if the embedded T has the strong feature requiring MSC, the derivation will not converge unless raising precedes insertion of Exp, giving (202b).

We have considered various kinds of expletive constructions, including embedding constructions that bar superraising, MSCs of various kinds, and ECM constructions. There is a fairly complex array of options. Within this range of constructions, at least, the options are determined by elementary principles of economy: (1) add optional $\alpha$ to the numeration only if it has an effect at the interface; (2) at each stage of a derivation, apply the most economical operation that leads to convergence. So far, at least, the results look encouraging.

4.10.3. Empirical Expectations on Minimalist Assumptions

With these clarifications, let us turn to the questions delayed at the end of section 4.10.1. The more restricted framework imposed by strict adherence to minimalist assumptions eliminates mechanisms that previously barred unwanted derivations. We therefore face problems of two kinds: to show that (1) the right derivations are permitted, and (2) the doors have not been opened too wide. The specific line of argument is sometimes fairly intricate, but it's important to bear in mind that at root it is really very simple. The basic guiding idea is itself elementary: that the array of consequences is determined by strict application of minimalist principles (to be sure, construed in only one of the possible ways; see the introduction). To the extent that the conclusions are confirmed, we have evidence for a conception of the nature of language that is rather intriguing.

We may continue to limit attention to simple transitive verb constructions, taken to be of the form (182) before T is added to yield TP. It suffices to consider overt V-raising, which brings up harder problems; the covert-raising alternatives fall out at once if this case is handled.

The first problem that arises is that we are predicting the wrong order of elements for MSCs. As noted, the observed order is (205b) instead of the predicted (205a).
(205) a. Exp [Subj [T° max XP]]
    b. Exp T° max Subj XP

The best answer would be that the order really is (205a) throughout the N → λ computation. If the expletive is null, we do not know its position, though (205a) is expected by analogy to the overt case. In section 4.9 we noted the possibility that the expletive in MSCs is overt in order to satisfy the V-second property, which may belong to the phonological component. If that is the case, the observed order is formed by phonologic operations that are extraneous to the N → λ computation and may observe the usual constraints (V → C), but need not, as far as we know: T° max-adjunction to expletive or to TP, for example. Let us assume the best case and see where that leads. We thus take the order to be really (205a), irrespective of what is observed at the PF output.

T and V have intrinsic –Interpretable features that must be checked: for T, [(assign) Case] (nominative or null); and for V, its φ-features and [(assign) accusative Case]. In addition, the nonsubstantive categories T and v may (optionally) have a strong [nominal-] feature, which is also –Interpretable. All have to be erased in a checking relation established by Merge or Move, by substitution or adjunction. Optional features are chosen when needed for convergence, as little as possible, in accordance with the economy condition (76). Features are deleted when checked. They are furthermore erased when this is possible, apart from the parametric variation that permits MSCs; see discussions below (58) and (185). Erasure is possible when no illegitimate object is formed (detectable at once). Checking takes place in the order in which the relations are established. These are the optimal assumptions: we hope to show that they allow exactly the right array of empirical phenomena. We are concerned now only with strong features.

As the derivation proceeds, the first checking relation that can be established is by overt substitution in Spec of v. We have two proposals under consideration: (1) “closer than” is defined in terms of c-command alone and overt object raising can only be to the inner Spec; (2) “closer than” is defined in terms of c-command and equidistance, and the object may (perhaps must) raise to the outer Spec. Again, let us restrict attention to the more complex variant (2); under (1), no problem arises, as can readily be checked.

Suppose, then, that overt substitution is in the outer Spec of v: Spec2 of (188). In this case v has a strong [nominal-] feature. It cannot be checked by merged Subj, as we have seen, so there must be an extra Spec as in (188), more explicitly (206).
Again, Vb is the complex form \([_0 \nu \nu']\). Subj is in Spec\(_1\) for \(\theta\)-theoretic reasons unrelated to strength.

Spec\(_2\) must be filled overtly to remove the strong feature of \(\nu\) before a distinct larger category is formed: in this case, before merger with \(T\). This can be done by Merge or Attract/Move. We have already excluded Merge. An argument inserted in this position does not establish a checking relation (so that the strong feature is not checked) and also lacks a \(\theta\)-role, violating FI; the derivation crashes. Expletives cannot be merged in this position, as we have seen. The only option, then, is raising of either Obj or Subj. We want to allow only raising of Obj, which, as we have seen, then permits the derivation to converge by raising of Subj to the checking domain of \(T\), and only that way.

We have briefly (and incompletely) considered why raising of Subj to Spec\(_2\) of (206) is barred. Let us look at the possibilities more closely, to clarify the issues.

Suppose that Subj is raised to Spec\(_2\) in (206). It is in the checking domain of \(V\), and checking relations are established for Case and \(\phi\)-features. If features mismatch, then the derivation is canceled. If they match, then Subj receives accusative Case and object agreement, and the Case and \(\phi\)-features of \(V\) erase. The Case of Obj still has to be checked, and that will have to take place in the checking domain of \(T\). But unraised Obj cannot reach that position, as we have seen, because Subj is closer to \(T\). The trace of Subj in Spec\(_1\) is invisible to Attract/Move and therefore does not prevent raising of Obj, but Subj itself does.

The only possibility, then, is that Subj raises further, to the checking domain of \(T\). Now its trace in Spec\(_2\) is invisible, and FF(Obj) can raise to \(T\). The trace left in Spec\(_2\) deletes, and at LF the result is identical to the result of the derivation in which Spec\(_2\) was never constructed. The strong feature of \(\nu\) has no effect on the PF or LF output in this derivation and therefore cannot have been
selected for the numeration, by the economy principle (76). This option is therefore excluded. If Spec₂ exists at all, it must be formed by overt object raising.

As the derivation proceeds, the next checking relation that can be established is substitution in the first [Spec, T] that is formed (EPP), either by Merge or Attract/Move. For Merge, the only option is an expletive, which (if pure) establishes a checking relation only with the strong [nominal-] feature of T, requiring covert associate raising; raising of an expletive to this position works the same way. The only remaining case is raising of an argument to [Spec, T], necessarily Subj, as we have seen. Then its features enter into checking relations with the “most proximate” sublabels of the target, in the obvious sense: the φ-features of V (which are the only ones), and the Case feature of T. Vb cannot have a strong feature at this stage of the derivation, but a checking relation is established with the strong [nominal-] feature of T that forced the overt substitution. The sublabels in checking relations erase if matched by those of Subj; if there is a mismatch, the derivation is canceled.

Suppose either T lacks a strong [nominal-] feature or that feature has already been erased by substitution of Exp in Spec. Then FF(Subj) raises to adjoin to T⁻¹ max, forming (194), modified slightly here.

(207)

Checking proceeds exactly as in the [Spec, T] case, canceling the derivation if there is a mismatch, erasing – Interpretable features if there is a match. If Obj has raised overtly to [Spec, Vb], its features are checked there and undergo no covert raising. If Obj has not raised overtly, then FF(Obj) raises to T⁻¹ max, forming (195), repeated here.

(208)
The Case feature of T has already been erased by Subj, so FF(Obj) checks the Case feature of V, canceling the derivation unless its own Case feature is accusative. Nominative Case and subject agreement necessarily coincide.

We rely crucially here on several earlier assumptions, among them that Attract/Move “sees” only the head of an A-chain and that mismatch of features cancels the derivation. In discussing the latter topic, we considered the weaker principle (107) that allows the derivation to proceed after mismatch, possibly converging in some different way. Largely on conceptual grounds, we rejected that option in favor of (108), which cancels the derivation under mismatch: (108) is preferred because it reduces computational complexity. We now see that the decision was necessary on empirical grounds as well: if we were to allow the broader class of derivations, certain choices for the (optional) Case features of DP would have allowed the derivation to converge improperly, with accusative subject and nominative object. Once again, the principles chosen on conceptual grounds are empirically confirmed, as we would expect if the Minimalist Program is capturing something true and important about human language.

This is not an exhaustive review of possibilities, but it includes a range of cases that seem to be central and typical.

The problems that remain have to do primarily with MSCs. We have gone a considerable distance toward the conclusion that the structure (191), repeated here, is the only possible form for MSCs (with or without overt object raising).

(209) Exp [Subj [\text{T}^{\text{max}} \text{XP}]]

This structure is permitted, and we have seen that a number of unwanted possibilities are excluded. Some remain, however, and there is also a substantial conceptual problem: how do we account for the range of options?

We are concerned with the “double EPP” structure (210).

(210) [\text{TP} \text{XP} [\text{T} \text{YP} \text{T}']]

We want to restrict the options to (209). Let us break the problem down into several parts.

(211) a. XP and YP cannot both be Exp.
    b. YP cannot be Exp.
    c. XP or YP must be Exp.

If so, we have (209).

The conceptual question is (212).

(212) What optional choices can allow three options: (a) no Spec of T (VSO), (b) one Spec of T (EPP), (c) two Specs of T (MSC)?
Case (212c) is (210), which we hope to restrict to (209). We have to ask how the three options of (212) are specified, along with combinations of them: for example, Icelandic, with options (212b) and (212c).

Let us begin with (211a). The facts are straightforward: there are no structures of the form Exp–Exp, and furthermore, the empirical observation generalizes to a broader class of cases, including the perennial troublemaker (213).

(213) *there₁ is believed [there₂ to be a man in the room]

These structures are excluded if we take the strong [nominal-] feature of [D] to be [N-], not [nominal-] generally (including [N] and [D]). That is a reasonable move: it limits the possible cases to the ones we find in DPs generally, namely, N → D raising; and it has the consequence (plausible, if not unproblematic) that the associate of an expletive must be nonspecific, whether it is an ordinary expletive construction or an MSC with the associate in [Spec, T] (see pp. 315, 312). Let us adopt this assumption. Then Exp–Exp constructions are straightforwardly barred. Exp is a DP, so the strong feature of the outer Spec will not be checked by raising of the inner one: only DP–NP structures are possible, analogous to ordinary expletive-associate pairs. As for (213), the features of the associate a man can raise to embedded T in the normal way, with [N] then raising to there₂ to check its strong feature; but there₁ survives intact to LF, so the derivation crashes (raising of there₂ will not help, because it is a DP).

Since Exp–Exp is barred, to establish (211b) we have to show only that Argument–Exp is impossible in successive Specs. We know that at every point of the derivation, Exp is in the position [Spec, T] and that it must be in the Exp–Subj order of (209) at some stage of the derivation for its D-feature to be deleted by N-raising from Subj. What has to be shown, then, is that once (214a) is formed, Exp cannot raise to become [Spec, H], with a subsequent operation forming (214b) (XP the outer Spec of H).

(214) a. Exp-Subj
   b. [XP[HP Exp H']]  

The problematic derivation is barred straightforwardly. In (214a) the N-feature of Subj adjoins to Exp to delete [D]. But this operation is covert feature raising and cannot be followed by overt raising to Spec to form (214b).

It remains to establish (211c) and to answer the questions of (212). Let us begin by assuming (211c) to be true (we will return to this assumption). We now have to address (212); specifically, we have to determine what the choice of options must be. It is easy to make the distinction between no Spec and some Spec; that is a matter of availability of the strong [nominal-] feature for
The hard problem is to distinguish cases (212b) and (212c), each assuming that T has the strong [nominal-] feature. How can that distinction be expressed, within the limited resources available?

Suppose there is a parameter that allows two choices: one Spec or two Specs. That proposal fails, because the two-Spec option permits Subj and Obj both to raise to [Spec, T], violating (211c): in the current highly restricted framework, nothing prevents raising of both Subj and Obj to the Specs of (210). Suppose we modify this approach, appealing to an economy principle that allows the two-Spec option only if the derivation would not converge without it, a variant of (76). That fails too, because it bars even the desired Exp–Subj structure (209); as we have seen, the derivation converges as an ordinary expletive construction if there is only a single Spec = Exp. A fortiori, the parameter cannot be of the preferred form: \( n \) Specs, with \( n = 0 \) (case (212a)), \( n = 1 \) (case (212b)), or \( n \geq 1 \) (with (212c) a special case).

A more promising idea is to attend to the fact that each occurrence of [Spec, T] arises from an operation, either Merge or Move. Merge is cost-less, so it can apply freely. But it can apply only once to form [Spec, T]: only Exp can be inserted in this non-\( \theta \)-position in a convergent derivation, and as we have just seen, double-Exp constructions are barred. Therefore, Attract/Move must be applied at least \( n \) times to form MSCs with \( n + 1 \) Specs (\( n \geq 0 \)), and \( n + 1 \) times if no Exp is available for insertion. Each such application violates Procrastinate, which suggests that the parameters involved in (212) should be framed in terms of such violations.

A violation of Procrastinate that is required for convergence is not an economy violation; one that is not required for convergence is an economy violation. To facilitate exposition, let us distinguish terminologically between the two (radically different) kinds of violation of Procrastinate, speaking of forced violations (for convergence) and unforced violations (true economy violations).

In a VSO structure (type (212a)) with T weak, there are no violations of Procrastinate. In an S–VP structure (type (212b)), there is at most a single forced violation: that is, if no Exp is available for insertion. In an MSC an unforced violation is tolerated: that is, a violation of Procrastinate not required for convergence.

We therefore have two options.

1. T may be strong or not.
2. T may or may not tolerate an unforced violation of Procrastinate.

Option (215a) is taken over from the Agr-T system: T may have the strong feature [nominal-] (EPP), or it may not (VSO). If option (215a) is selected but
not (215b), then there is one Spec but there can be no unforced violation of Procrastinate: [Spec, T] is Exp if it is available and a raised argument if it is not. If (215b) is selected as well, there are arbitrarily many Specs; that is still too weak a conclusion, so (215b) must be sharpened.

To clarify the issues, let us compare the Agr-T approach to MSCs with the one we are now exploring. (215b) is the counterpart to the parameter (216) of the analysis reviewed in section 4.9.

(216) T may or may not have Spec.

In the system of sections 4.1 – 4.9, [Spec, Agr] is always available and (212b) and (212c) (single and double EPP) differ in the value of the parameter (216). The right outcomes are guaranteed by the principle (217).

(217) Only one argument (namely, Subj) can raise from VP to [Spec, I].

But as we have seen on the basis of a more careful analysis of properties of features and the empirical phenomena that were to be explained, principle (217) is untenable: Subj, Obj (and other arguments) have Interpretable formal features and thus can be raised even after feature checking if the MLC is satisfied, as it is in the relevant cases once we restrict ourselves to minimalist principles, dispensing with apparatus that is conceptually unnecessary and empirically defective. The apparent generality of the earlier analysis is therefore spurious.

Suppose a language selects (215b) as an option and thus allows a violation of Procrastinate. This yields exactly the right double-subject constructions (Icelandic), but it also permits MSCs with more than two Specs. We therefore have to restrict (215b) to the minimal number of violations of Procrastinate—exactly one—revising (215) to (218).

(218) a. T is strong.

b. T tolerates a single unforced violation of Procrastinate.

A VSO language has a negative setting for (218a) ([negative (218a)]: T is weak. An EPP language lacking MSCs has [positive (218a)] and [negative (218b)]: T is strong but there is no double-Spec. An Icelandic-type language with optional double-subject MSCs has [positive (218a)] and [positive (218b)]. The framework thus covers the typological range we have so far considered.

A prettier picture would include another option, the possibility of more than a single unforced violation of Procrastinate. Suppose, then, we have another parameter that distinguishes two types of language.

(219) T tolerates arbitrarily many unforced violations of Procrastinate.
The languages we have so far considered have the negative value for (219): no more than one unforced violation is tolerated (and no forced violations, if the setting of (218a) disallows strong T). Could there be languages with [positive (219)], allowing arbitrarily many unforced violations of (219), which we may interpret as meaning maximal unforced violations? In such languages all arguments are extracted to a position outside of IP. The candidate that comes to mind is a language in which all arguments do appear outside of IP, their syntactic roles being indicated only by some relation to an element within the remaining complex word: pro with matching features, or perhaps trace with arguments appearing in MSCs. If so, then (219) has the flavor of Baker’s (1995) poly-synthesis parameter. It seems a possibility worth exploring. If it makes sense, then we would interpret the system of parameters to imply that unforced violations of Procrastinate can be disallowed, minimal (once), or maximal (always). Needless to say, we are now pretty far out on a limb.

Still unsettled is case (211c): the impossibility of a double-EPP construction with no Exp, the two occurrences of [Spec, T] being occupied by Subj and Obj (in either order) (an Icelandic-type language that allowed MSCs with Subj-Obj, for example). The proposed answer to (212) bars this possibility. The language would have to allow strong T and an unforced violation of Procrastinate ([positive (218a–b)]) but only minimally ([negative (219)]). That set of choices (Icelandic) disallows the double-unforced violation required to extract exactly two arguments to [Spec, T] positions.

The options (218b) and (219) can be formulated readily in terms of the mechanisms for multiple-Spec mentioned earlier (below (58)), namely, by allowing a feature to escape erasure after checking. The option in (218b) is rejected if that possibility is disallowed for the strong [nominal-] feature of T, chosen if the possibility is allowed. The option in (219) is rejected if the strong [nominal-] feature can escape erasure once, adopted if it must escape erasure as often as possible.

Consider the unaccusative construction (220), where Nom is the sole argument.

(220)

\[
\begin{array}{c}
T \\
\text{max} \\
\hline
\end{array}
\begin{array}{c}
\text{T} \\
VP \\
\hline
\end{array}
\begin{array}{c}
V \\
\text{Nom} \\
\hline
\end{array}
\]
Suppose that $T$ is strong. Then either Nom raises overtly to $[\text{Spec, } T]$ or Exp is inserted in that position. Exp cannot raise from a lower clause; that would be barred by the closer Nom. But suppose Exp is merged in $[\text{Spec, } \text{VP}]$ (a non-$\theta$-position), then raised. The outcome is legitimate, but the derivation is again barred: by the condition that shorter derivations bar longer ones from the same numeration. The conclusion adds further weight to the assumption that this is a true economy condition. Furthermore, we fill the last gap in establishing (196), the basic descriptive observation that restricts expletives to $[\text{Spec, } T]$. The observation follows (nonredundantly) from independently motivated economy conditions, as we would hope.

Nothing has been said so far about quantifier raising (QR). The status of this phenomenon has been the topic of much controversy, which I will not attempt to review. Suppose the operation exists. Since it is covert, it must be feature raising: a quantificational feature $[\text{quant}]$ raises to adjoin to some $X_{\text{max}}^\text{null}$ that is a potential host (presumably $T$ or $v$, which, we might assume, have optional affix features allowing them to host $[\text{quant}]$). Suppose so. $[\text{quant}]$ is Interpretable and therefore need not be checked. The affix feature of the functional categories is optional, therefore chosen if it “makes a difference,” assuming that the economy condition (76) holds for this case, which differs from ones we have so far considered. Since there is no PF effect, the result would be to allow QR when it leads to a distinct interpretation, in some sense that must be made precise, an idea proposed by Reinhart (1993) in a broader framework and developed with considerable empirical support by Fox (1994). It falls naturally into the framework already outlined.

If the discussion so far is on the right track, then a variety of apparent reasons for inclusion of Agr in the lexical inventory have been eliminated. The question of its existence is therefore narrowed, though not eliminated. Not all arguments for Agr have been considered. The discussion has been based on the assumption that Agr has no $\phi$-features—that these features are assigned to substantive lexical items as they are drawn from the lexicon. If Agr exists as the locus of $\phi$-features, it has an even more restricted role and unique status than before, with no apparent impact for the core computational processes; that seems dubious, at least. For the case of subject agreement, these apparent functions of Agr could perhaps be accommodated within the system just outlined by assimilating it with $T$: by assuming, that is, that as $T$ is drawn from the lexicon for the numeration, it too is optionally assigned $\phi$-features (as are nouns; and, I have so far assumed, verbs and adjectives). Note that this carries us back to something like the analysis that was conventional before Pollock’s (1989) highly productive split-I theory, though now considerably revised and
in quite a different setting. The agreement-based justification for Agr would therefore reduce to adjective and object agreement. To carry the matter further, it is necessary to look into a range of empirical questions that go well beyond the compass of this inquiry.

As matters stand here, it seems reasonable to conjecture that Agr does not exist and that \( \phi \)-features of a predicate \( P \), though –Interpretable, are like the Interpretable \( \phi \)-features of nouns in that they are part of \( P \) in the numeration, added optionally as \( P \) is selected from the lexicon.

### 4.11 Summary

Reviewing briefly, it seems that we may be able to eliminate the theory of phrase structure entirely, deriving its properties on principled grounds. Many consequences follow for the theories of movement and economy when these conclusions are combined with other minimalist assumptions. Taking the latter seriously, we are led to a fairly radical reformulation of the theory of the computational system that relates form and meaning, and to a sharpening and improvement of economy and other central notions.

We have been concerned with the computation \( C_{HL} \) that maps a numeration \( N \) selected from the lexicon to a pair of interface representations \( (\pi, \lambda) \), at PF and LF, respectively. At an arbitrary point in the \( N \rightarrow \lambda \), computation, the operation Spell-Out strips away phonological features, which enter the morphological component, are linearly ordered, and then are handed over to the operations that map them to \( \pi \). The \( N \rightarrow \lambda \), derivation proceeds in large measure independently of the “extraneous” requirement that language is manifested in sensorimotor systems: the sole effect may be the need to construct derivative chains involving categories to satisfy PF convergence (and, more broadly, the existence of Attract/Move in the first place). Apart from the mapping to PF, \( C_{HL} \) satisfies the conditions of uniformity and (virtually) inclusiveness. It consists of two operations, Merge and Attract/Move, which incorporates Merge. Economy conditions that are in large part readily computable select among convergent derivations.

Something like Merge is inescapable in any languagelike system, but the operation Attract/Move reflects peculiarities of human language, among them the morphology-driven “last resort” properties to which we have directed special attention.

Functional categories and their formal features occupy a central position in the workings of \( C_{HL} \). If the general approach of section 4.10 can be sustained, then the only functional categories are those with features that survive through the derivation and appear at the interfaces, where they are interpreted. Of the
functional categories we have considered, only T, C, and D remain. Strong features, which play a considerable role in overt manifestation and language variation, are narrowly limited in distribution. We have seen no reason to suppose that N or V, the basic substantive categories, have strong features. The strength property can be restricted, perhaps, to the nonsubstantive elements T and v that head the major projections within the clause, and to complementizers that serve as mood-force indicators (but see note 133).

The various refinements and simplifications along the way sharply reduce the problem of exponential blowup of economy calculations and overcome a variety of conceptual problems concerning Last Resort movement and "shortest move" (MLC). They also appear to be confirmed empirically in an interesting range of cases. Most important, they fall out naturally from minimalist assumptions, which require that we keep strictly to operations on features and local relations among them (categories only derivatively), taking into account the crucial distinction between ±Interpretable features that is determined by bare output conditions. Hence, we have confirmation that the inescapable conditions on language may be satisfied in something like a "best possible" way.

On minimalist assumptions, these seem to be the right kinds of moves, though doubtless not yet the right ones. Like earlier efforts guided by the same goals, they raise many questions and, if plausible, call for a good deal of rethinking of what has been assumed.

More generally, it seems not unreasonable to conjecture that language may approximate a "perfect" system in the sense described in the introduction. If this intuition is accurate, it would make good sense to press it to the limits to see what can be discovered about this curious and increasingly mysterious component of the human mind. To progress further along this road, we will have to find out the answers to a wide range of empirical questions that have come into view and that are sometimes formulable in fairly clear ways. We are left with hard and challenging problems of a new order of depth, and prospects for a theory of language with properties that are quite surprising.

Notes

On the background of this chapter, see the introduction. Thanks to participants in the MIT lecture-seminars of fall 1993 and 1994 for their many suggestions and criticisms in what was (as usual) a cooperative effort, inadequately acknowledged. Thanks particularly to Chris Collins, Sam Epstein, John Frampton, Sam Guttmann, and Howard Lasnik for comments on earlier versions that were the basis for Chomsky 1994a, and to Juan Uriagereka (among others) for improvements on a draft version.
1. The PF level itself is too primitive and unstructured to serve this purpose, but elements formed in the course of the mapping of syntactic objects to a PF representation might qualify.

2. For my own views on some of these issues, see Chomsky 1975b, 1992a,b, 1994b,c, 1995.

3. Recall that the ordering of operations is abstract, expressing postulated properties of the language faculty of the brain, with no temporal interpretation implied. In this respect, the terms output and input have a metaphorical flavor, though they may reflect substantive properties, if the derivational approach is correct.

4. The work in Optimality Theory mentioned earlier does not address such problems. In Prince and Smolensky 1993 there seems no barrier to the conclusion that all lexical inputs yield a single phonetic output, namely, whatever the optimal syllable might be (perhaps /ba/). That would be ruled out by Prince and McCarthy’s (1993) “containment condition” (suggested in passing by Prince and Smolensky 1993, 80) as a “non-obvious assumption” that they have “found essential”). But it is hard to see how this can be sustained in segmental phonology (as in the cases mentioned) without implausible assumptions about parsability, varying widely among languages. It seems likely that these approaches will have to postulate intervening levels within the computation to PF, raising the question of how they differ from rule-based approaches. They may well differ, at least in the domain of prosodic processes (which are hard to separate from segmental ones). At present it seems to me likely that Bromberger and Halle (1989) are correct in holding that phonology, unlike the rest of C_HL, is rule-based, perhaps apart from some specific subdomain.

5. In a subsequent paper, referring to the preceding note, McCarthy and Prince (1994) appear to accept the conclusion that in the theory developed in Prince and Smolensky 1993, there is a unique output for all inputs (for each language). Recognition of that conclusion they call “the fallacy of perfection.” They allege that comparable problems arise generally in the theory of language. If such problems did arise elsewhere (they surely do not, at least for the cases they mention), that would be a serious matter indeed, and recognition of it, no fallacy. They recognize the need to add input-output relations of some kind (“faithfulness”). Traditional approaches, dating back to Pāṇini and revived in generative phonology from the late 1940s, spell out “faithfulness” in terms of the notion “possible phonological rule” (embodying assumptions about natural processes) and economy considerations on the system of rules (evaluation metrics, markedness considerations, etc.). McCarthy and Prince (1994) propose that “faithfulness” be restricted to input-output conditions, but what they suggest seems to have no relevance to the standard problems (e.g., “identity between input and output,” a principle that is virtually never satisfied). The basic problem, long familiar, is the one mentioned earlier: crucial properties appear to hold not of input-output pairs but of intermediate stages, so that no input-output condition is formulable. Pending further clarification, we cannot ask how an approach in terms of Optimality Theory might differ from others, apart from what is suggested by particular cases studied.

6. I put aside further questions relating to interpretation of lexical items at LF (recall that their “descendants” at PF may not be identifiable there, being absorbed into instructions for the A-P systems that obliterate their identity). Such questions would carry us far afield into issues of language use (including the actual process of referring and the