

Noam Chomsky: “Derivation by Phases”  
(To appear in M. Kenstowicz ed., Ken Hale: A Life in Language, MIT Press, 2000)

James Yoon  
Linguistics, UIUC

\* The comments below are based on a manuscript version that was circulated in 1999. I understand that a later version, published in MITOPL, differs from the manuscript version slightly. May be some of the questions raised here find satisfactory solutions in the later revised versions.

1. Methodological Guidelines (‘First Principles’):

Chomsky proposes the following methodological guidelines. These are the ‘first principles’ that serve as the conceptual underpinnings of the theory to be developed.

Strong Minimalist Thesis (SMT):

Language is an optimal solution to interface (sound and meaning) conditions.

Strong Uniformity Thesis (SUT):

“In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.”

2. Architecture – Optimal and Not-so-optimal properties:

Given the methodological guidelines/first principles, especially SMT, some features of grammars appear to be ‘optimal’, while others don’t. The optimal properties may be thought of as ‘perfect’ solutions to the requirements imposed by the interfaces, while the not-so-optimal properties may be ‘imperfections’, given the first principles. Of course, the first principles are valid only up to empirical verification and it may turn out that on closer inspection the ‘imperfections’ are not what they appear to be.

Optimal properties (“perfections”):

- Interface levels, PF and LF.
- Inclusiveness Condition (lexical resource-driven computation).
- (Binary) Merge - an indispensable operation for any recursive system.

Not-so-optimal properties (“apparent imperfections”):

- Labels: distinctness of labels and label-determination algorithm, i.e., the question of whether Label should be admitted as a syntactic operation. (Note: In classical Categorical Grammar, constituents are not distinguished in terms of their Labels, as you may recall)
- Existence of syntactic levels other than interface levels.
- Derived tree-geometric syntactic relations.

Sister (defined directly from Merge)

Immediately Contain (defined directly from Merge, as it builds a rooted structure)

Contain = I-C (I-C)

Identity = SIS (SIS) (this seems to be a perverse way of defining identity!)

C-command = SIS (I-C(I-C ...))

Note that Precedence cannot be derived from the closure of Sister and Immediate Contain. Thus, Precedence cannot be a property of Narrow Syntax, while Dominance and Command can. Interesting reasoning. A similar reasoning is going to be extended to Move. Movement is not a primitive operation but one that is composed out of the more basic operations (Merge, Agree, Pied-pipe).

#### On ‘first principles’ and ‘epistemological priority’:

You may recall that in LGB, Chomsky argued against admitting Grammatical Relations as syntactic primitives on grounds that they do not meet the criterion of “epistemological priority”. The idea is that while GRs may still be referred to in syntax, they exist only as derived syntactic notions. As many have pointed out, the problem with this argument was that he never defined what “epistemological priority” was. Things are different in Minimalism. The First Principles, rightly or wrongly, serve the role of defining what is epistemologically prior. Given the first principles and the assumption that derived syntactic relations arise only by composing basic relations, GRs defined in terms of Immediate Containment/Dominance and Labels may still fall within the realm of narrow syntax. In contrast, Precedence cannot be referred to in narrow syntax, as it cannot be so defined. This means that Kayne’s (1994) LCA must be a principle of linearization (relevant in ‘converting’ narrow syntax to PF, an interface principle), rather than a narrow syntactic principle.

- Non (interface)-interpretable features:

Phonological features (need to be eliminated/converted to ‘interface-interpretable’ form)

Uninterpretable Formal features (need to be eliminated/converted to ‘interface interpretable’ form)

Interpretability is an interface concept, but features are *lexically categorized* with respect to Interpretability, since otherwise we couldn’t have Inclusiveness. As uF (uninterpretable formal) features are not sanctioned at LF, by hypothesis, they have to be eliminated in narrow syntax (to get an LF that conforms to “Full Interpretation”).

However, these features are often realized phonologically, so their “elimination” must allow for ‘conversion’ to phonetic features at PF. The conversion to PF is not information-preserving and may violate Inclusiveness and Interpretability (see below on this).

### 3. Uninterpretable Formal Features and Agree as an additional operation of the Computational System:

#### Observation:

Uninterpretable Formal (uF) features enter into **Agreement** (construed broadly to include Case-checking as well as traditional agreement) with “matching” Interpretable formal (iF) features.

- (i) Agreement is the relation that eliminates or converts uF features to ‘interface-interpretable’ form.
- (ii) If so, *Agree is necessary*, in addition to Merge (and whatever else we may need from among the “not so optimal” relations/concepts in section 2).
- (iii) The uF features subserve the phenomenon of ‘displacement’, though displacement itself may have an interface-interpretable function (i.e., “surface semantic effects”, such as Topicalization).
- (iv) It may then be that uF features and Agree are not real imperfections after all.

The ‘pairing’ of uF and iF features is important, and allows Chomsky to maintain a version of what he calls the “**Interpretability Condition**” (p.113 in the version in Step by Step) in MI (“Minimalist Inquiries” paper). The idea behind this condition is that the features that are used by grammatical systems are all interpretable features. However, if Interpretability holds throughout the system, there should be no such thing as a ‘purely syntactic’ feature. (Sounds like functionalism, doesn’t it?) The uF features look *prima facie* like they are such features, but in reality they are just iF features lodged on the ‘wrong’ category, we can maintain a weakened version of this condition.

Obviously, even this much seems to be an idealization. If you think about it, grammatical Gender on nominals, where it is supposedly interpretable, is not the same as real-world gender. That is, grammatical categories differ from real-world distinctions and conceptual categories. Interpretability then must be understood in the following way - grammatical features that are *inherent* to a category are interpretable; those that are *relational* to it are not.

It appears that even the weaker version of Interpretability is counterexemplified by things like the EPP feature and CASE feature, as these do not seem to be paired with iF features. However, we shall see subsequently that both of these features can be made consistent with Interpretability. The form of Interpretability we can maintain then is what I would call **Weak Interpretability**.

### 4. Agree:

Agree (here as in MI) replaces the older idea of covert category/covert feature-movement/Attract. Unlike the older ideas, however, movement is divorced from the Agree relation, though as we shall see, you do want principles pertaining to Agreement to play an indirect role in movement. Agree is similar to (long-distance) head government in GB theory (see Raposo & Uriagereka 1991 LI paper on long-distance Case-assignment. See also “percolated government” in Den Besten’s early work for Nominative objects).

The basic intuition about this relation seems to be the following.

- (i) Agree is not symmetric, but an **asymmetric**, relation between a **Probe** and a **Goal**. The Probe has uF features while the Goal has ‘matching’ iF features. Agree eliminates the uF features on the Probe by ‘valuing’ them and ‘removing them’/‘spelling them out phonologically’. Valuing is **unification** of the values of the features.
- (ii) uF features have **attributes with empty values** in their feature structures, i.e., [F: \_ ], while iF features have attributes with specified values, [F:x]. Nice. This is a very clever characterization of uF features and succinctly characterizes the difference between relational (uF) and inherent (iF) features.
- (iii) Conditions on the Probe-Goal relation:
  1. Probe and Goal must be sufficiently **local** to each other (*Locality*).
  2. Probe and Goal must be ‘matched’ in terms of attributes of features that enter into Agreement. However, match need not be identity. Specifically, the attributes on the Probe must **subsume** those on the Goal. This allows the Probe to have a subset of the attributes on the Goal. Defective Probes do not block Probing by a non-local complete Probe (*Maximize Match*).

If this is all that there is to Agreement, we have a problem. This is because the Goal phrase is a completely passive partner, as it has no uF feature. It is hard to see how the Goal can be displaced to the Spec of the Probe, as it often is. (On the other hand, we may surmise that this is what is going on in concord, which presumably never involves displacement). You could try to solve this problem by saying that the Probe has a feature, EPP. However, we don’t want just any XP to move to the Spec. We want to ensure that the right type of phrase moves – i.e., a DP moves to fill the EPP needs of a Probe which has phi-features, but not, say, an Adverb.

#### 4.1. Case and EPP on Probe:

NC assumes that the Goal DP has an uF, namely, Case (or something like Case in the case of A-bar movement, such WH, TOP, FOC, etc., which triggers the relevant phrase to undergo the right type of movement. On A-bar movement, see section 10 of this commentary). He then states the definition of Agree so that **both** the Probe **and** the Goal are required to have uF features. You’d think that this makes the Goal DP a Probe and the Probe a Goal as far as Case is

concerned, so that there is a bottom-up Agreement as well as top-down Agreement, and in a sense it is correct (though I am not sure Chomsky would endorse this interpretation).

Case is assumed to have the following properties:

- (i) It is a *uF on DPs*. So, (referential) DPs have something like {PHI/AGR:[PER:x, NUM:y, GEN:z], [CASE: \_\_ ], ... } as their feature structure. Expletive DP (“there”) is assumed to have the feature structure {[PER:x], ... }. That is, expletives do not require Case (Notice what this does to the old Case Filter vs. Visibility Condition debate).
- (ii) Case is *valued/eliminated under Agreement with a Probe that has complete phi-features* (Person, Number, Gender), though CASE itself is “not matched”, as it is “not a property of the Probes”.

I think I understand the reason behind this decision (thanks to discussion with Jungmin Jo). The idea seems to be that if CASE were an attribute on the Probe (say, T) and also on the Goal (DP), you’d end up with a situation where you would have to assume that CASE on either the Probe and Goal is interpretable, under Weak Interpretability. However, this is not the case. If you say that CASE exists as an attribute only on the Goal, this problem can be avoided.

How does Chomsky work out the intuition that Case is valued only through a relationship with (what used to be called in GB theory) a Case-assigning Head? He does so by assuming that *the value of the CASE attribute on the Goal DP is a complete PHI/AGR-feature structure on the Probe*. That is, CASE takes a feature structure, rather than an atom, as its value.

To work out this intuition technically, we need to adopt complex (typed) feature structures. If you think about it, we need to bring in complex feature structures to model our intuitions about the ‘EPP feature’ as well. The value of the EPP feature (we can call it SPEC) is a full category (recall that the EPP feature used to be called not too long ago a D-feature?). Thus, ‘valuing’ the EPP feature is equivalent to *adding* an overt Spec to the Probe with the feature. In contrast to these two features, the values of the individual phi-features can be matched/unified independently of the others. This is because their values are atoms.

The distinction between simple feature structures (where attributes have atomic values) and complex feature structures (where attributes have a feature structure as value) is a familiar one from unification-based theories such as G/HPSG and LFG, and can be used profitably to model what Chomsky is suggesting in DBP. Schematically;

Simple features (atom-valued):

PER: {1, 2, 3}

GEN: {m, f, n}

NUM: {sg, pl}

Complex features (feature-structure/category valued):

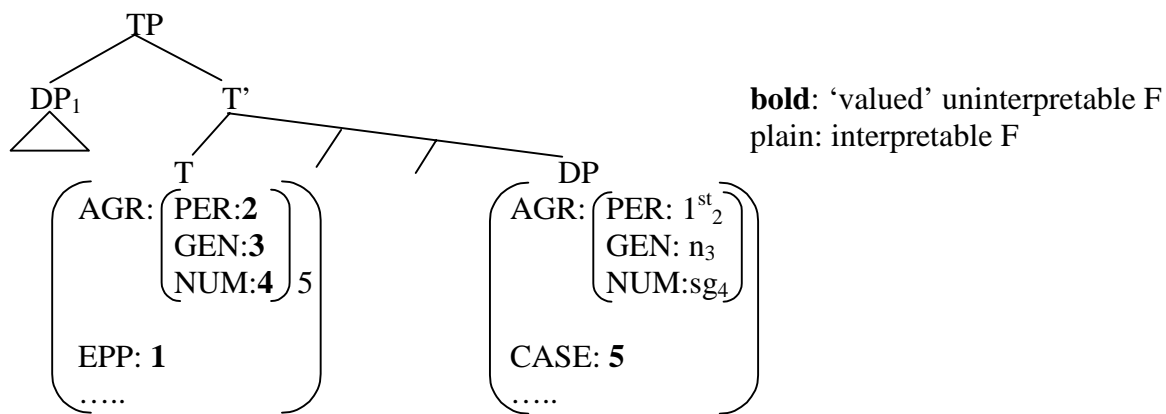
PHI/AGR: {feature structure whose attributes contain minimally PER, and may also contain GEN and NUM}

CASE: {complete phi-feature set/AGR}

EPP/SPEC: {non-empty set/list of categories}

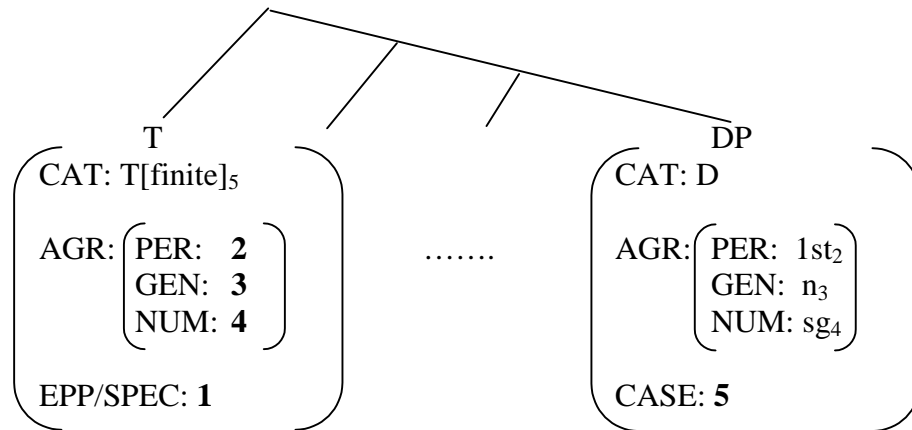
Notice that SUBCAT(egorization) (c-selection) would also have a set (or a list) of categories as value. If the EPP feature on a category takes as its value a set/list of cardinality  $n=1$ , you get a single SPEC; if it is valued for a set/list whose cardinality/length is  $n \geq 1$ , you get multiple SPECS.

The following diagram illustrates the mechanisms just developed. As stated earlier, valuing is unification of feature values.



This is all fine and dandy except for the following problems. One problem is that since the DP on which the CASE attribute exists already has PHI/AGR, we need to know why it can't value the CASE attribute internally. We must stipulate that it is only a PHI/AGR on a Probe category that can value CASE on the Goal.

One way to solve this problem is to assume that the feature structure of the Probe T is as follows and posit that CASE on DP is valued by the CAT specification of the Probe – specifically, by a finite T, transitive V, etc. This makes Nom Case an uninterpretable T feature (as in Pesetsky & Torrego 2000). A desirable consequence of this move is that CASE is now distributed in a manner consistent with Weak Interpretability. Chomsky's analysis shown earlier actually does not solve the violation of Weak Interpretability. CASE on Goal DP is uninterpretable, but it is valued by a complex PHI/AGR feature structure on the Probe. However, the PHI/AGR feature structure the Probe is itself uninterpretable. CAT features, if we admit them into narrow syntax in some form (as atomic features or as derived features, a la Marantz 1997) are surely interpretable, so under the revision suggested here, CASE is valued by an interpretable feature, as desired.



We haven't determined yet whether the EPP/SPEC feature conforms to Weak Interpretability. Recall that Weak Interpretability demands that an uF feature on the Probe be paired with an iF feature on the Goal and vice versa. The EPP feature can be thought of as conforming to Weak Interpretability if we think of it in the following way -- the interpretable counterpart of the uF EPP/SPEC feature is the Goal phrase itself. This still leaves as a problem instances where the EPP feature is satisfied by an expletive. A possibility worth considering is that the EPP feature is a phonological feature, and hence not subject to Weak Interpretability (cf. Anders Holmberg 2000, "Scandinavian Stylistic Fronting: How Any Category Can Become An Expletive," LI 31.3. See also section 9, 10 of this commentary.).

- (iii) Case 'selects' the chunk that is to be moved. If Case is part of the feature structure of the Head of an XP, then presumably it is the entire XP that is 'selected' as the chunk to be moved. There seems to be some notion of "percolation" of features assumed here. I don't know whether the operation of **Label**, assuming it is needed in narrow syntax, suffices to model the percolation, or we need a separate operation **Pied-Pipe** (as Chomsky seems to assume).

Resolving this issue is important, as it bears on whether Move qualifies as a legitimate syntactic operation composed out of the basic operations. Chomsky assumes that it is, and that Move is composed out of Agree, Merge, and what he calls 'Pied-pipe', but he never mentions whether Pied-pipe is itself a basic operation of syntax. This last operation is introduced through the backdoor, so to speak. If 'Pied-pipe' is nothing other than Label, then Move can indeed be built out of the basic operations. (However, his use of Pied-pipe in later sections of the paper preclude it being unified with Label, I think).

For our purposes, let's assume that Label is what determines the XP chunk to be moved. I.e., Move targets the maximal structure that is Labeled by the lexical item that has a CASE attribute (assuming CASE is a feature that enters into Label). Targeting of the maximal structure may not even have to be stipulated, since in 'bare phrase structure' theory (Chomsky 1995), Label of X and that of the XP of which X is the Head are *token-identical*, i.e., the same thing. So, you cannot move X or X' without XP if what is being attracted is the Label (as I will argue is the case with EPP

below. EPP is valued by a Label, rather than phi-features). Incidentally, this means that Head Movement, if it exists, cannot be valued by Labels.

- (iv) It seems that a category with EPP and PHI-features (T, for example) will be able to attract phrases with PHI-features, namely, DPs. Otherwise, movement of any category should be sufficient to satisfy the EPP needs of the Probe. I don't know what guarantees this. It used to be easy when EPP feature was considered an uninterpretable D-feature, but now we need to state the dependency in an indirect way. I will have more to say on this below.

The result of adding Case on the Goal implies that;

Active Feature Restriction on Agreement:

(3i) Probe and Goal must both be **active** for Agree to apply.

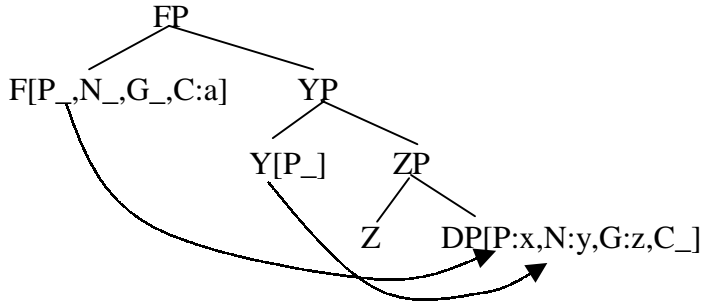
The status of this condition is going to be debated later in the paper. The issue has to do with things such as whether to allow one uF feature/attribute to value another uF feature/attribute in an Agreement configuration as long as there is an iF feature in an Agreement configuration with one of the uF features to value the uF feature, i.e., whether there is “*transitivity*” of Agreement (cf. Frampton et. al. 2000).

Of course, if you think about it, if ‘valuing’ is unification of values and unified values are *token-identical*, “transitivity” follows at once, though you might need to restrict such “transitive Agreement” to within a strong phase since Spell-out forgets earlier phases.

#### 4.2. Long-Distance Agreement, (Successive-cyclic) Movement and Locality:

Now that we've seen how Agreement (including Case) works, let's turn to issues of **locality**. Agreement is local, and so is the movement of XP to (the Spec of) a category that has the EPP feature. We need to examine how locality is ensured in the DBP system. The first approximation to locality is *local c-command*, but we want to build in “Relativized Minimality” effects into the definition of locality. (Note: some people, including NC, use the term **Minimal Link Condition** to refer to locality in general, but the term is misleading since Agree involves no movement and does not establish Chains whose ‘links’ can be referred to.)

Let's first consider the intervention effect of *phi-features* on Probes, specifically, the question of a closer Probe blocking the more remote Probe. (For ease of illustration, I am treating CASE as an attribute of the Probe, but you should understand the status of CASE on Probes in terms of the previous discussion).

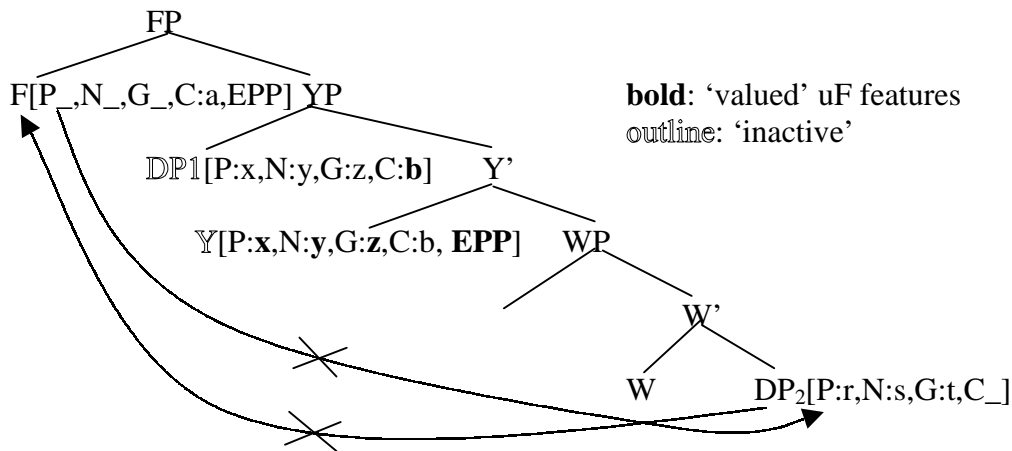


In the tree shown above, F is a **phi-complete** Probe, while Y, which is closer to DP than F, is a **phi-incomplete** Probe. Since the attributes of Y partially match (subsume) DP's attributes, Y can Agree with (the PER attribute of) DP, but Y cannot value DP's uF Case attribute. Therefore, DP is still active and can be accessed by the more remote F, which can value its CASE feature while the phi-features of DP value those on F. In this case, we have on the surface a **non-local agreement** between F and DP. The idea that only completely matching feature structures can give rise to blocking is called **Maximize Matching** in *DBP*.

Suppose Y has an EPP feature in addition to [PER]. Y should then be able to attract DP to its Spec. DP moved to SpYP is 'active', as its CASE attribute has not been valued. With DP in SpYP, F is now the local Probe that matches the attributes of DP. Therefore, F can Agree with DP and value the CASE feature of DP (with F's phi-features valued by those on DP). If F too has EPP, then DP moves subsequently to SpecFP to value the EPP feature. In this case, unlike the previous one, Agreement (as well as movement) is local.

What about the intervention effect induced by *EPP features* on Probes? If both F and Y have EPP features, it seems that you cannot bypass movement to SpYP and move directly to SpFP. This may be because F and Y are identical with respect to the EPP attribute. This indicates that the EPP feature must not be part of the phi-feature set of the Probe. I.e., [[EPP:\_\_\_], [PER:x, NUM:y, GEN:z], [CASE:\_\_\_]]. If it were, the relevant blocking could not take place under Maximize Matching. The separateness of EPP and PHI-features is what we were assuming all along.

Let us now turn our attention to the *locality effect induced by Goals*.



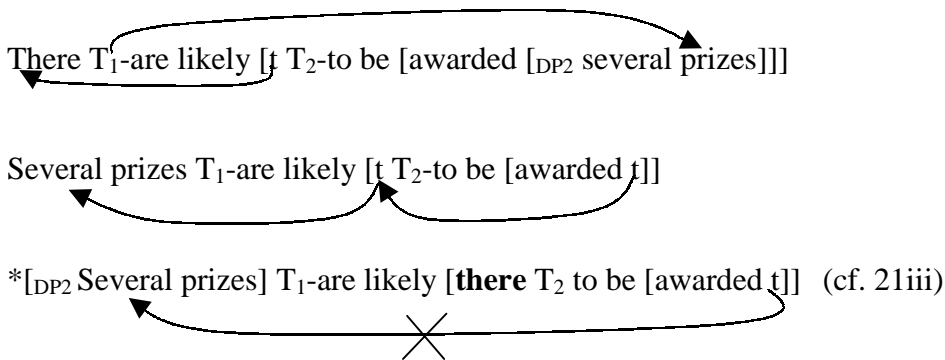
Suppose now that DP1 Agrees with Y and is attracted to SpYP (from within WP). DP1 then is **inactive**, as its uF Case has been valued. So is Y, as its uF features (phi-features and EPP feature) have been valued. However, DP2 is still active and so is F. The questions then are;

- (i) Can DP2 Agree with F? (i.e., phi-features of F Agree with those of DP2)
- (ii) Can DP2 Move to SpFP? (i.e., ‘Agree’ with the EPP feature of F)

Chomsky says “no” to both. DP1 (and/or Y?) blocks the **Agreement** of F and DP2, presumably because DP1 and DP2 have identical phi-feature attributes (equivalently for Y and F), despite the fact that DP2 and Y are both inactive. The idea that a checked feature/category can still figure in locality was called “**Defective Intervention Constraint**” in *MI*.

The same reasoning extends to **movement**. Since Y has an EPP feature, albeit a ‘valued’ one, it blocks the EPP feature on F from accessing DP2. Alternatively, we know that DP1 is the closer XP with phi-features to F than DP2. So, EPP on F could not attract DP2 over DP1, despite the fact that DP2 is inactive. This construal, from *MI*, requires that EPP cannot be satisfied by just anything. It has to be DP. Otherwise, the EPP feature would not “know” that a closer DP blocks a more remote DP from moving. This is why in earlier analyses, the EPP feature was thought to be valued by a category feature (N, D). In *DBP*, Chomsky wants to eliminate reference to categorial features (Why? Don’t ask – certainly you still need to refer to the categorial features of functional categories), so he suggests that EPP picks out (a category with) phi-features, even one with incomplete phi-features (such as the weak expletive “there”, which only has PER as its phi-feature).

**Expletives** raise an interesting problem. Suppose now that in a configuration like that shown above DP1 is an expletive, which has only [uPER], by hypothesis. Assume also that Y is a defective Probe (with [uPER], its [uEPP] valued by Merge of the expletive). Does DP1 block Agreement of F and DP2? It doesn’t, given long-distance Agreement in such constructions. Is the movement of DP2 to SpFP allowed, over DP1? It doesn’t seem to be, since otherwise we should get the third sentence below:



What we see here is that intervention effects for Agree and Move are different. This should hardly be news.

Chomsky in *DBP* appeals to the **Maximize Matching** principle mentioned earlier to rule out the third sentence. The idea is this. When the Probe  $T_1$  is looking down into its c-domain in the three sentences shown above, it finds “there” as the first thing that it could Agree with. However, Agreement with “there” is only partial, so  $T_1$  needs to access  $DP_2$  to get full Agreement, which it can given Maximize Matching. Therefore, you get non-local Agreement. In contrast, unlike

Agreement, “there” is the first thing down that could satisfy the EPP feature on T1. Therefore, DP1 “there”, not DP2, moves to SpT1.

#### 4.3. Checking [PER: \_\_ ] on the expletive: (Transitivity of Agreement)

There is a technical question of how to check [PER: \_\_] on “there”. In the first sentence above, T2 is a defective Probe, with {[PER: \_\_], [EPP: \_\_]}. When T2 Merges on top of VP, it can search its c-domain and find DP2, which has {[PER:3, NUM:pl, GEN:neut], [CASE: \_\_]}. DP2 will value T2 as [PER:3]. The EPP feature on T2 is valued by the Merge of the expletive “there”. The question now is what values [PER: \_\_] on the expletive.

In the above structure, there is a higher Probe, T1, but its phi-features are not valued either. T1 Probes down past the expletive (or Agrees with it vacuously? – Frampton et. al. 2000) and Agrees non-locally with DP2, which will value its phi-features. The question is whether this valuing has any effect on removing the uF PER feature on “there”. It doesn’t seem to unless we change the system.

One solution is to allow ‘there’ to be a Probe with respect to its uPER (while it is a Goal wrt the EPP feature of T2, being the right type of element to fulfill the EPP needs of T2). If so, ‘there’ can Probe down and find DP2 and value its PER feature. This is the solution suggested in *MI*. Another solution is to assume that the valued PER feature of T1 is able to value the uPER feature of “there”, as [PER:3]. As noted, this move would violate (3i), as the phi-features of T1 are no longer ‘active’. Yet another solution would be to allow T1 to Agree with the expletive vacuously. When T1 is valued by DP2, “there” is also valued automatically, as it Agrees with T1.

A prediction of assuming that ‘there’ has PER valued in one of the manners sketched above is that the PER value of ‘there’ will differ depending on the phi-features of its associate. This is not necessarily a bad prediction.

Of course, we may question the wisdom of assigning PER as the only attribute of a weak expletive like “there” in the first place and seek alternative interpretations of the facts. For example, we may revert to the earlier assumptions (i) that “there” has no features except the categorical feature D, (ii) that EPP is valued by a full category DP, and (iii) that nonfinite T does not possess [uPER] but only the EPP feature. Under this set of assumptions, the problems noted above do not arise.

In relation to expletives, we need to consider how the ‘strong’ expletive “it” works – as in “**It** is obvious [that the world is round]”/“I consider **it** to be likely [that he will resign]”, etc. One thing that we presumably want to say is that the strong expletive “it” has a complete phi-feature set and may be an unvalued CASE feature. Now, if the expletive is inserted in SpTP to fill the EPP needs of T, we may wonder how its CASE gets valued, or for that matter, how the phi-features of T get valued, if Probing can only be down into the c-domain of the Probe. The strong expletive is Merged directly in SpecTP and does not raise there from a lower position (but see Moro 1997). Maybe we have to allow Agree between a Spec introduced by Merge and the Head as well. Or maybe all expletives lack CASE, including the strong expletive (in view of something like the Visibility Condition) and clauses have phi-features that value those on T.

I have to admit that while all this is interesting and relevant, I don’t know how precisely Chomsky resolves these questions in *DBP*. His exposition is very scanty – the relevant discussion is around examples (15) and (16). Frampton et. al. (2000) also reach slightly different conclusions.

More analyses:

[ $v_{[P: \_ N: \_ G: \_ C: d]}$  [ $VP$  expect [ $TP$  there $_{[PER]}$  T-to $_{[PER, +EPP]}$  [be awarded [ $P: a, N: b, G: c, C: \_$ ] several prizes]]]]]

- DP ‘several prizes’ Agrees with T-to, valuing PER of the latter. DP still active.
- ‘there’ inserted in SpT-to valuing EPP. T-to now inactive.
- DP’s CASE valued by  $v$ .  $v$ ’s phi-features valued by DP. DP inactive.  $v$  inactive.
- PER of ‘there’ valued by  $v$ ? Not valued?

[ $v_{[P: \_ N: \_ G: \_ C: d]}$  [expect [ $TP$  [ $P: a, N: b, G: c, C: \_$ ] several prizes] T-to $_{[PER, +EPP]}$  [be awarded  $t$ ]]]]]

- DP ‘several prizes’ Agrees with T-to, valuing the latter’s PER.
- DP moves to SpT-to, valuing its EPP feature. T-to now inactive.
- DP’s CASE valued by Agreement with  $v$ .  $v$ ’s phi-features valued by DP. Everything inactive.

5. Defective Probes:

The problem is that selection of “defective” TP is stipulated to be a lexical property in the system so far (as it has been since GB theory). A similar stipulation seems necessary for the “small  $v$ ”, a relative newcomer to the theoretical scene. Unaccusative and Passive VPs are assumed not to have “small  $v$ ” over them, which means T can select VP (defective selection) or  $vP$  (non-defective selection).

NC entertains some ideas as to whether a more principled theory of the selection of defective elements can be developed. He ends up not pursuing a possible way in which such a theory may be worked out (called **Alternative II**), except to retain the idea that T may be a substantive category. He adopts this assumption in the context of a more “conventional” theory of defective elements – called **Alternative I**, but he keeps coming back to ALT-II in footnotes. ALT-II models s-selection as involving the valuing/elimination of uF features. This we saw already is the way c-selection/EPP is/should be treated. Modeling selection as elimination of uF features is attractive and allows us to push further the intuition that syntactic derivation is driven by the need to eliminate/value/convert uF features (more on this below when we get to phases and the ‘scattering’ of interface representations).

(Actually, the intuition about selection/subcategorization in traditional generative grammar is perfectly consistent with the view of uF features as those with ‘unvalued’ attributes that are cancelled/valued by ‘matching’ categories. If so, selection is a way to eliminate uninterpretable features. E.g., a verb which s-selects a DP has an uninterpretable/unvalued feature which will be valued/canceled when it combines with a DP. The picture is a familiar one from Categorical Grammar. What Chomsky is trying to do here is different – he is inquiring into the reason as to why a given head selects particular dependents, e.g., why a particular uF selectional feature is associated with a particular head.)

Let's consider ALT-I first.

Alternative I ("conventional"):

(5) C selects  $T_{comp}$ ; V may select  $T_{def}$ , where  $V=ECM/Raising V$ .

Notice that under ALT-I, s-selection does not involve matching of uF phi-features. What (5) says is that  $T_{comp}$  is never directly selected by V and C may not select a  $T_{def}$  (what about "for-to"? May be we need to assume "to" is  $T_{comp}$  – cf. Pesetsky & Torrego 2000). We would have to say something similar about "small v" as well. T may select V (V has no phi-features), rather than "small v" (with phi-features). In this alternative, the selection of defective elements is treated as a matter of lexical idiosyncrasy. Therefore, NC suggests that a more principled way to deduce the selection of defective elements may be possible.

Alternative II:

(6) C is always phi-complete; T is phi-complete only when necessary.

That is, phi-completeness of the category s-selected is determined by the phi-completeness of the selecting Head. This actually does not completely remove lexical stipulation, it just restates the arbitrariness in terms of something more principled (?) – Agreement/matching.

There are two (or three) cases to consider.

C-T selection:

- C always has a (uF) complete phi-feature set, which is eliminated when it s-selects a TP with a complete (uF) phi-feature set. Cf. Pesetsky & Torrego (2000) on C bearing uT feature. Note that to make this work, you need to give up Condition (3i) and allow uF feature to value a matching uF feature as long as one of the uF features is valued by an iF feature (transitivity rears its head again).
- T has a complete (uF) phi-feature set only when necessary. Therefore TP without phi-features (or with incomplete phi-features) can be selected by V, which does not have a phi-feature set. T without phi-features (or with incomplete phi-features) can also select V, rather than  $v$  (which always has a full phi-feature, by hypothesis).
- NC says that under this alternative, it is "tempting" to associate phi-completeness and EPP. That is, only phi-complete Heads can have EPP. This will make the theory of 'strong' phases a bit more principled than it is now. C and  $v$  "head" strong phases, which are phases that allow movement to their Spec/Edge. The reason is that EPP is lodged on only phi-complete categories.

Under ALT-II, movement to SpCP is always possible, but then there would be no intermediate movement to SpTP of defective-infinitival TPs in Raising-ECM constructions. This would wreak havoc on all the standard analyses of successive-cyclic A-movement. Therefore, despite its conceptual attractiveness, NC does not adopt ALT-II and keeps working with ALT-I (but don't be surprised if later on he adopts it, because it is conceptually attractive).

v-V selection:

- Assume that “small v” always introduces an external argument and that it is (uF) phi-complete and may have an EPP feature.
- The analogy with C-T would then mean that V selected by such v should also be phi-complete.
- This means that v should never be selected by a phi-incomplete T, which is nice, if such a T can only embed unaccusative/raising verbs (recall these are VPs lacking phi-features, not vP).
- The complete uF phi-features of V are then checked by those of v via s-selection. Again, you need to allow uF features to check/value other uF features.
- Unless selected by v, V is defective, without (the full complement of? any?) phi-features and hence unable to assign/check Acc Case. Unaccusative/raising verbs fall in this class.
- This means that Acc checking would be from V<sub>comp</sub>, rather than v, for V with complete phi-feature set. (The question then becomes what does the phi-feature on v do?)
- Depending on where the EPP feature can be lodged, this means that there are four landing sites for movement – SpC, Sp<sub>v</sub>, SpT<sub>comp</sub>, and SpV<sub>comp</sub>. May be this is one too many.

Chomsky does not pursue Alternative II further, but seems to like the idea that the parallel between v-V and C-T and suggests that **T may be a substantive category**. He is going to use this idea to provide some “motivation” for why vP and CP act as cycles (strong phases).

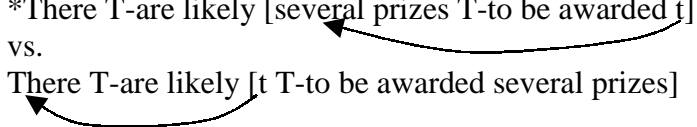
Small v and Psych verbs:

What about Psych verbs (transitive V with non-agentive subjects)? In *MI*, NC assumes that the Experiencer argument of “seem” (It seems **to Bill** [that ....]) may be introduced by a ‘small v’. Does this mean that in this usage “seem” has a small v but when it lacks the Exp argument, it lacks vP altogether? Seems so.

Assuming a “small v” here creates more problems than it solves, I think. For one, you cannot say that “small v” is correlated with Acc-Case checking, as these verbs don’t check Acc. Only the “small v” associated with Agents seems to check Acc-Case. If there is a reason for wanting “small v” with Psych verbs, it must have to do with phases. I haven’t figured out the crucial paradigm yet.

6. Deducing the Merge/Agree-over-Move preference:

\*There T-are likely [several prizes T-to be awarded t]  
vs.  
There T-are likely [t T-to be awarded several prizes]



The first is ruled out in favor of the second by the M-over-M Preference. There is one phase-cycle for the whole sentence (as we shall see) and an expletive exists in the Numeration. The idea is that SpT-*to* cannot be filled (to satisfy the EPP feature of T-*to*) by movement of “several prizes”, but only by insertion of the expletive, which subsequently undergoes movement to SpT-*are*. This is because Merge is ‘cheaper’ than movement. The preference is deduced since “Move” is an operation that is (almost entirely – not entirely, given Pied-Pipe) composed of the more basic operations – Agree/Pied-pipe/(Re)-Merge, while Merge is a simple operation. If less is more, then the Preference follows.

Not only is there a Merge-over-Move preference, there is/should be an Agree-over-Move preference. The effects earlier attributed to “Procrastinate” should follow from this (so he argues in *MI*), though it isn’t clear in the new conception how the intuition is implemented. Move is triggered only by EPP, which does not seem to have a principled relationship with phi-features that determine Agreement. Again, more work on the internal structure of the feature structures is needed to make this work.

## 7. Features, Lexicons, Lexical Arrays and Sub-Arrays:

NC assumes that there is a universal set of features provided by  $F_L$  (faculty of language) from which the Computational System (CS) constructs interface representations,  $EXP = \langle PF, LF \rangle$  for a given sentence. NC reasons that in principle, the CS could draw from the universal feature set directly to build EXP (Uriagereka in his commentary says that this is what Generative Semantics was like. Don’t know if I agree with this assessment), but as a matter of fact, CS doesn’t seem to work this way. The features which get into derivations are ‘filtered’ from the universal set F in several successive steps.

**F**: universal set of phonological, formal, and semantic-conceptual features



**F<sub>L</sub>**: Subset of F that figures in the lexicon of a particular language.



**LEX**: A set of lexical items LI constructed from F<sub>L</sub>.

The intermediate steps are justified in terms of their roles in reducing “computational complexity”.

FYI, in *MI*, the following constraints were suggested as the relevant complexity-related considerations. (p. 104, in the Lasnik Festschrift):

- a. Simple operations pre-empt more complex ones (Agree/Merge-over-Move).
- b. Search space for operations is limited (locality, as seen above).
- c. Access to the universal feature set F is restricted (as above).
- d. Computation is locally determined (no look-ahead/phases).

### 7.1. Nature of LEX:

- The set LEX is:
  - (i) A single collection, that is (most) LI's are “**signs**” expressing a pairing of PHON-FORM-SEM features.
  - (ii) Minimal, in that only non-redundant information is encoded. It is a set of **lexemes**, rather than phonological/morphosyntactic words.
  - (iii) PHON features are relevant to PF; SEM features are relevant to LF; and FORMAL features overlap with those in SEM (do we really want to say that FORM and SEM overlap? Recall the need to distinguish between conceptual/semantic categories and ‘corresponding’ syntactic-formal categories).
  - (iv) The subset of FORM features that does not overlap with SEM features is the uF formal features. These ‘drive’ the syntactic computation.
  
- Is LEX ‘distributed’, a la DM?
  - (i) Only when evidence warrants late insertion, as in inflectional morphology and suppletion. In such cases, the PHON part of the sign might be a place-holder, i.e., Q-FORM-SEM, with Q triggering the relevant phonological operations after late insertion.
  - (ii) Roots and predictable affixes may be signs (PHON-FORM-SEM complexes).
  
- Is LEX a “structured” collection of features?

I have been modeling LEX as a structured collection of features, using complex feature structures. Uriagereka in his commentary raises the interesting question of how the sub-structures of LEX got to be there if indeed LI's in LEX are a structured collection of features (as in familiar HPSG lexical entries). He mentions evolution as a possible answer, and some as yet unknown measure of optimality as another. I would be happy with a more mundane answer – namely, that there is a limited computation ‘prior to’ narrow syntax, in the Lexicon (though Marantz 1997 would disagree). Yet another possibility is that the structures in LEX reflect earlier syntax, as the thesis of lexicalization-grammaticalization would have it. I know that Chomsky is sympathetic to Marantz's position against the generative Lexicon, but Chomsky's ideas in *DBP* are compatible with a more traditional, un-distributed, view of the Lexicon, hence with the supposition that there could be pre-syntactic computation.

## 7.2. LEX-to-(Narrow) Syntax Mapping:

NC says that the phonological component can access  $F_L$ , even after the choice-determination of LEX. I can't figure out what this means, so I will ignore it. On the other hand, syntax doesn't. This means that LI's in LEX are mapped to syntax. How is this done?

- A choice of LI's from LEX, a Lexical Array (Numeration, a **multi-set**), is mapped to derivations, not the entire LEX directly.

- LI's making up the Numeration are mapped to derivations yielding the interface representations EXP observing the Inclusiveness Condition (this means that Chain links are token-identical, for example).
- It is not the entire Numeration, but sub-arrays of the Numeration, which are accessed by the derivation. These sub-arrays define the **phases** of a derivation. Phases may be constructed in parallel (i.e., a finished phase not integrated into another phase) or consecutively (with a finished phase integrated into the next larger phase). However, for the derivation to converge all phases have to be eventually integrated and the root phase must converge.

All of these additional steps are argued to be necessary based on their presumed role in reducing computational complexity.

## 8. Phases:

### 8.1. What constitutes a phase?

**Phases** are:

- ‘Propositional’.
- Have the form  $[_{FP} F-[_{LP} L \dots]]$ .

Depending on one’s theory of functional categories, D-N,  $v^*$ -V, C-T (remember T is now a lexical category), etc. would be phases. Whether or not all three of them constitute phases (NC appears to think so) depends on what is meant by “propositional”. The idea is to find a common semantic characterization of phases-cycles. If DP is a phase, then “propositionality” is similar to the earlier idea of Complete Functional Complex (**CFC**) used in the 1986 Binding Theory. Of the phases some are strong -- CP (proposition + force) and  $v$ P (event structure – but only when  $v$  is Agentive) – while others are weak (DP?). Not all XPs are phases. TP and “weak” verbal projections (passive/unaccusative verbs with VP) do not constitute phases.

In *MI*, Chomsky argued that phases/cycles cannot be defined by ‘convergence’. The alternative referred to is one that says an XP that is convergent, i.e., with no outstanding unchecked features, defines a phase. This alternative, if it can be made to work, would be more interesting as it is less stipulative. A research topic, doubtless.

Roughly, a phase corresponds to the earlier notion of **cycle**. Cycle used to be S and NP in early TG. VP (old VP, now  $v$ P) was added as a cycle in Barriers. The present conception follows Barriers, except that VP is not a phase, as noted above. Another concept that is similar to phase is the notion of **Governing Category** or **Complete Functional Complex**, as noted.

So, given a Numeration, the question of how many phases there should be in deriving EXP is determined by inspecting how many tokens of phase-“heading” functional elements -- C,  $v^*$ , D -- there are in the Numeration. By assumption, there is at least one C (root C) in all Numerations.

**“Strong” phases** are:

- Phases whose ‘heads’ may have an EPP feature.

“Strong” phases are ones whose Head allows (successive) movement to (and through?) the Spec/Edge (including Head movement). Again, in the Barriers conception, CP and VP (now  $v$ P) allow movement to the Edge. DP is a cycle-phase, being ‘headed’ by a functional category dominating a lexical-root category. However, it is presumably not a strong phase. The reason may be as follows. While movement **to** the Edge of DP is fine (if Pied-piping involves movement), movement **through** the Edge is not. However, as we know, this is a dubious conclusion. Movement to and through the Edge of DP has been posited in many instances (cf. Szabolcsi on Hungarian DPs). Also, complement DP/NP allows extraction out of them to varying degrees, as is well-known.

“Independent support” for (strong) phases:

- They (the Edges of phases) are reconstruction sites. If reconstruction piggy-backs on movement (i.e., if reconstruction is to ‘trace’ positions), then this is evidence that movement can target strong phases. For  $\nu$ P, I don’t really know what the evidence for reconstruction is. The relevant evidence should consist of reconstruction to an Object-Shifted position.
- They have a “degree of phonetic independence”. I.e., CP can move, cleft, etc., but not TP.  $\nu$ P can move, but not VP (may be the evidence is the sort given in Huang 1993 where it is argued that a subject trace must be posited in a fronted “VP”. Huang assumed that AgrP is what is fronted. In the current version, what is fronted is  $\nu$ P).
- “PF and LF integrity” correlate.

Questions:

Do we need a tri-chotomy (non-phase vs. weak phase vs. strong phase) rather than a dichotomy? (phase vs. non-phase) That is, what is the difference between a non-phase XP and a weak phase XP? Are all XPs phases minimally? Chomsky has a cryptic answer in a footnote in *DBP* to the effect that you need the trichotomy, but I couldn’t follow his argument. If anyone can help me, I’d be grateful.

8.2 Phases and “Interpretation”/“Evaluation”:

Background:

In addition to being the modern equivalents of cycle, NC wants the derivation by phase model to have the consequence that the interface representations – PF and LF are “scattered” throughout the derivation. That is, they are built up incrementally in tandem with the building up of the syntactic structure. In this, NC is buying the leading ideas behind models of “multiple Spell-out”, such as that of Uriagereka.

This idea is intuitively plausible. If we assume that syntactic derivation is driven by the need to eliminate uF features (*selection* as well as *movement* and *agreement* can be modeled this way – remember ALT-II?), then in a sense, what the syntactic derivation is doing is trying to get to the ‘start’ symbol (actually, the ‘end’ symbol), the root C, from a given collection of LI’s by eliminating the uF features in the lexical array. (For those of you familiar with Syntactic Structures, you will see this is the reverse of a  $\Sigma$ -*derivation*). By the end of a successful, converging, syntactic derivation, there will be no uF features left in syntax except root C.

If you assume that LI’s are not ‘distributed’, as NC seems to assume, then the consequences of building up the syntactic structure to rid the structure of uF features will be the *simultaneous building up* of connected (may be tree-like) structures for phonology-phonetics and semantics. It is in this sense that the interfaces come to be **scattered**, and built up incrementally by the syntactic-formal part of the derivation. You can also see why Chomsky does not want the radical ‘distribution’ of features associated with LI’s. If they were distributed, the syntactic derivation cannot build the interfaces simultaneously.

If you are serious about not referring to purely syntactic structure (NC says the representational constraints consist only of ‘output conditions’ on interface representations, but none on the purely syntactic (partial) representations constructed by the derivation), you can say that the only representations that are built up by the derivation are the interface representations, LF and PF (cf. Stabler book). Since there is no purely syntactic object-representation that results from the derivation, there is nothing to refer to. You may also be able to see the family resemblance of this view of MP and ‘parallel’ or ‘synchronous’ TAG, or for that matter, the system of Jackendoff (1997 – but unlike this system, *DBP* is still “syntactocentric”), or the Combinatory Categorical Grammar model of Steedman (1996, 2000 – another “syntactocentric” model).

#### Implementation:

The way NC works out this intuition is interesting. He does not actually assume a **simultaneous** building of the interface representations. The interfaces in fact follow a step or two behind the syntactic derivation. Syntax finishes one strong phase and goes on to the next strong phase. Before syntax is done with the higher phase, the interface representations of the lower phase are constructed. I.e.,

#### Next Phase Interpretation:

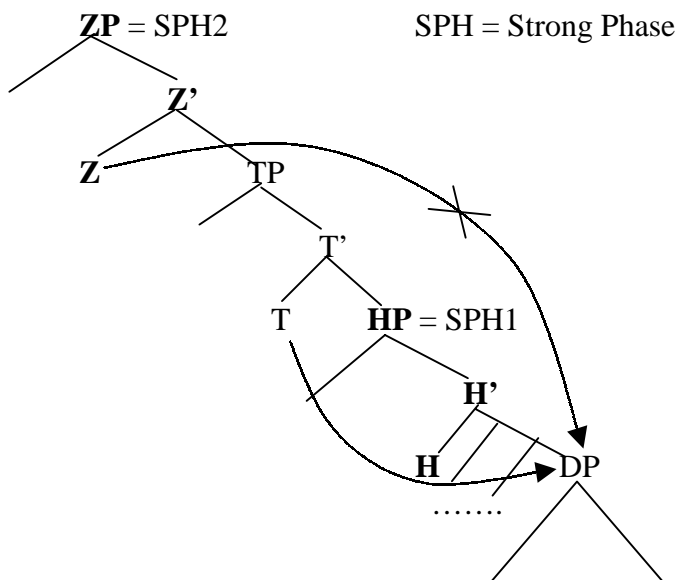
(10) Interpretation/evaluation for PH1 is at PH2, where both PH1 and PH2 are strong.

In order to make this possible, the opacity of a phase-cycle must be construed in such a way as to allow the operations at the higher phase to ‘peek’ at the elements in the lower phase. However, not everything within the lower phase should be accessible, since otherwise we are not going to have any cyclicity-phase effects. Thus, NC proposes the following:

#### Phase Impenetrability Condition:

(11) The (c-)domain of H is not accessible to operations at ZP, but only H and its Edge, where HP and ZP are strong phases, and Edge = {Head, Spec}.

Let’s see some examples to get a clearer idea of what is going on.



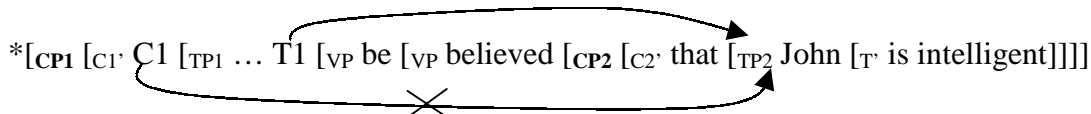
In the above diagram, a non-phase TP is located between HP and ZP. Assume that DP is the most deeply embedded constituent and is not a strong phase, say, the object within vP. If non-agentive verbs with quirky Case subjects have a small v, we want T to Agree with the DP embedded in vP (in fact, under VP within vP) and check Nominative Case. DP then must be accessible to T. How is this possible?

The intuition seems to be that TP is an **Extended Edge** of HP, to use a term that seems felicitous. Since  $\text{Edge}(\text{HP}) = \{\text{SpecH}, \text{H}\}$  and H can access DP, the idea seems to be that  $\text{Edge}(\text{TP}) = \{\text{SpT}, \text{T}\}$ , as **Extended Edge**(HP), can also access DP, though clearly DP is (contained within)  $\text{Domain}(\text{HP})$ .

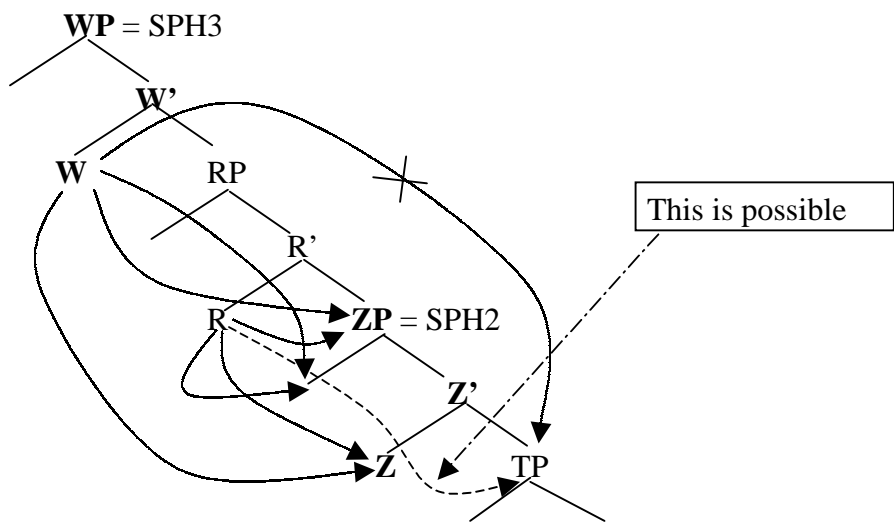
What about Z? Z will be able to access up to **Extended Edge**(HP), namely,  $\{\text{SpT}, \text{T}, \text{SpH}, \text{H}\}$ , but Z will not be able to access  $\text{Domain}(\text{HP})$  where DP is situated, unlike T. The reason Z cannot access DP is because DP is the  $\text{Domain}$  of a strong phase head H and Z is also a strong phase head. With T, which is not a phase head, the  $\text{Domain}$  of the strong phase head that it dominates must be accessible to it.

The following structure works similarly, as we can tell from the quote from note 28. R is the **Extended Edge** of Z and can access  $\text{Domain}(Z)$ , namely, TP.

“For A-movement, (3i – the provision that both Probe and Goal must be active for Agree) is not obviated; thus PIC(=11) prevents access to an inactive subject of CP (subject of CP?) (“John” below) from the next strong phase (that is, higher C1), but not from higher T1 with no intervening strong phase, as in;



with  $\text{Agree}(\text{T1}, \text{John})$  barred by (3i – as “John” is inactive within CP2) **but not by PIC...**”



You may ask yourself if there is any empirical reason why such long-distance access should be allowed. Maybe the reason has to do with the fact that A-movement of SpTP is possible in certain languages from the above structure. If the inactive nature of Goal is the only reason such movement is not attested in English (yielding the TSC effect), if the Goal can still remain active (say, it has two CASE attributes and only one attribute is valued within CP1), the higher T1 should be able to Agree with it. Since there is no intervening DP, the lower subject should also be able to move to the matrix subject position, yielding a sentence which violates TSC.

In Persian, such raising is possible (Ali Darzi's 1996 UIUC thesis). However, the raised DP does not Agree with the higher T, but only with the lower T. In other languages where such raising is reported (Ura's thesis), it seems you get double agreement though.

- A consequence of single-cycle derivation with scattered interface:

(13) Surface semantic effects are restricted to narrow syntax.

NC assumes, as will become evident, that while PF (and LF) is/(are) built in tandem with the syntactic derivation, there can be some displacement in PF, for example, the Thematicization/Extraction rule to be encountered later. Such displacement will not be driven by uF features, for one. However, since PF does not access the narrow syntactic derivation, any displacement in PF proper will not have any semantic consequences. We may wonder whether this allows for displacement in LF that is not triggered by uF features as well (QR). NC does not discuss this issue.

### 8.3. More Examples (up to second full paragraph of p.15 of the ms. version)

(15) (i) [ \_\_\_ T2-be likely [EXP T1-to [arrive a man]]]  
 (phi:\_) (per:\_) (phi:x)  
 (case:y) (case:\_)  
 (EPP:\_)

- The whole structure has only one phase, root CP.
- [per:\_] of EXP valued by that of T2 (though note here the problem of uF feature valuing another uF feature).
- [phi:\_] of T2 valued by [phi:x] of 'a man'.
- [EPP:\_] of T2 valued by movement of EXP.
- If T1 has EPP feature, then it must be valued by Merge of EXP.
- Spell-out of valued features, The structure converges as;

[There T2-is likely [ t-<sub>EXP</sub> T1-to [arrive a man]]]

(16) (i) [<sub>CP</sub> C [ we T1 [<sub>v\*P</sub> t-<sub>WE</sub> v\* [<sub>VP</sub> expect [<sub>TP</sub> EXP T2-to [<sub>VP</sub> arrive a man ]]]]]]]

phi:\_                      per:\_                      phi:y  
case:x                                           case:\_

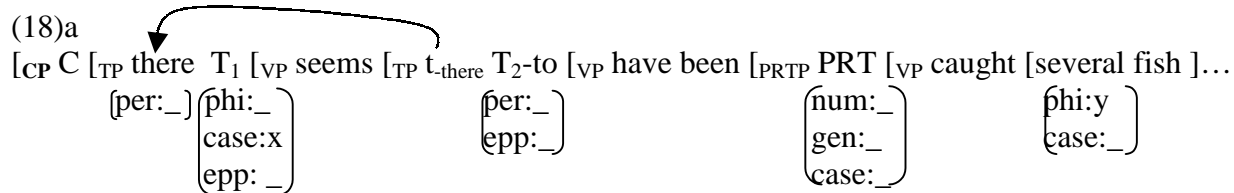
- There are two phases, CP and v\*P.
- Within the lower phase v\*P;
  - (i) phi-features of 'a man' valued by Agree with v\*.
  - (ii) phi-features of v\* valued by Agree with 'a man'
  - (iii) [per] of EXP valued by phi-feature of v\*.
  - (iv) case-feature of 'a man' valued by v\*
- The lower phase is convergent. Moving on to the higher phase:
  - (i) Subject 'we' Agrees with T1, valuing its phi-features.
  - (ii) T1 values the case-feature of 'we'.
  - (iii) EPP of T1 valued by raising of subject 'we'.
- The derivation is convergent, with no unchecked uF features (except root C). There is no need for operations at higher phase CP to 'peek' at elements on the Edge of the lower cycle v\*P.

#### Non-local Agree and Intervention (again):

NC proposes (17) as a possible reason for why EXP does not block the remote Agree relation between higher T and the DP in the c-domain of the EXP.

(17) Only the head of an A-Chain (equivalently, the whole Chain) block matching under the Minimal Link Condition.

The idea is that since only one link of the A-Chain = (there, t) c-commands T2 in (16), T2 is not 'more remote' than the A-Chain with regard to the DP 'a man'. However, this will not allow v\* to Agree with 'a man' in (16) unless the EXP raises (to Specv\*P). Therefore, NC proposes that (14 – Maximize Matching) is sufficient to allow non-local Agree. The EXP does not have a full set of phi-features and hence does not count as a closer element. (Notice here that NC is using the term MLC as a general name for locality, including locality of Agreement, though there is no movement and hence no 'link' in Agree – this is what I warned you about earlier).



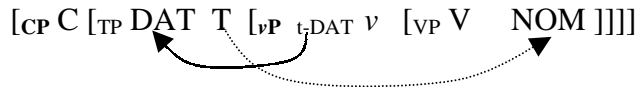
- There is one strong phase in the whole structure, root CP.
- Assume that participles project a functional category PrtP whose head PRT has the features [[NUM: \_\_, GEN: \_\_ ], [CASE: \_ ]]. That is, it has less than a full set of (unvalued) phi-features (realized overtly in some languages) as well as an unvalued Case feature. This last assumption helps explain participial Case agreement in relevant languages (Icelandic). Given the Strong Uniformity Thesis, we expect the same to hold in languages without the visible agreement.
- Proceeding bottom-up, we have the following:
  - (i) Num and Gen of PRT are valued by matching phi-features of DP.
  - (ii) Per of T2 valued either by Merge of EXP or by remote Agree with phi-features of DP. No intervention by PRT in the latter case since the phi-feature matrix is incomplete.
  - (iii) EPP of T2 valued by Merge of EXP. Per of EXP valued by the combination as well, or by Agree with T1.
  - (iv) Phi-features of T1 valued by Agree with DP, or partially with the valued phi-features of PRT. If the former, no intervention by EXP or PRT since both are defective. Case on PRT valued (indirectly, say NC) by Agree of T1 with DP.
  - (v) EPP feature of T1 valued by Move of EXP.
  - (vi) Case value of DP valued by Agree with T1.
- There are no outstanding features at the strong phase level. The derivation gets sent to PF for Spell-out.
- You can see that unlike the earlier conceptions of cyclicity (i.e., **Extension Condition**, as applied to Move and Merge in Chomsky 1995), in the present conception, there is quite a bit of *back-tracking* in derivations, especially for the Agree operation. Back-tracking (counter-cyclic) operations are allowed as long as they are within the same strong phase.
- We haven't seen examples that show the need to allow Edges of strong phases to be accessible at the next higher phase. Nor have we seen examples where a non-phase head contained between two strong phases can access the complement domain of the lower strong phase head. However, we can think of what the relevant examples would be. The first would be WH-movement.

Assuming that adjunction to VP in the Barriers system is movement to Spv\*P in the present system, a WH object DP, to get to SpCP, would have to get out to Edge(v\*P) first (by Object Shift, as we shall see). From here, it would be accessible from the next strong

phase head, C. Specifically, the (interpretable) [wh] feature on C could Agree with that on OBJ, valuing it. The OBJ can Move to SpecCP, if C also has an EPP-feature.



The second case would come from Nominative Objects with Quirky subjects (and also the scenario sketched in footnote 28 discussed earlier). I.e.,



T is a non-phase head situated between C and  $v$ , which are strong phase heads. However, you want it to Agree with the complement of V, giving it NOM Case. This is possible since T is the Extended Edge of  $v$ . Since  $v$  can access NOM, T can also access it. C, on the other hand, can access up to the Edge of  $v$ ,  $v$  itself and Spec $vP$ .  $vP$  is spelled out at CP.

#### Different conditions on Move vs. Agree locality – (21):

This was already discussed. The idea is that with regard to the EPP feature, the EXP is the first thing down that could satisfy it. So movement of lower DP is not allowed, given the Defective Intervention Constraints.

This is the point at which the exposition of the basic system ends. From this point on, ideas about PF-movement, Object Shift and EPP are explored. This is the logical point at which you may want to read other discussions, such as Frampton, et. al. 2000, before proceeding to the next part.

#### Summary:

Agree holds between two active elements X, Y where X c-commands Y and where the attributes of Y subsume those of X. Values of Agreeing features are unified.

Locality for Agree is closest c-command, except that non-local Agree is possible if a closer c-commander partially matches the attributes of the remote c-commander (Maximize Match).

Phi-features and EPP feature are in separate feature structures.

Once the head of a superordinate strong phase is introduced, the c-domain of the adjacent (lower) strong phase head is no longer accessible to syntactic operations at the higher phase, only the elements on its (Extended) Edge are.

(Partial) PF and LF (interpretation) of a lower adjacent phase must be constructed no later than the completion of the superordinate phase (i.e., as operations exit the higher phase).

## 9. PF-movement – Thematicization/Extraction (TH/EX): (p.15 - top of p.21)

In this part, Chomsky focuses on some word order peculiarities in unaccusative constructions in English and develops an interesting account of ‘PF displacement’. One of the manifestations of this PF-movement yields a structure similar to the output of Object Shift (OS) in other languages. OS is taken up in the next part of the paper (p.21-31).

### 9.1. Unexpected word order in unaccusative/transitive constructions in English:

- The DO in unaccusative constructions cannot surface in its ‘base-generated’ position. It seems that DO must be ‘extracted to the edge’ of VP **obligatorily** for such structures to be well-formed. Extraction could be to the right (as in extraposition) or to the left of the VP.

(22)

- (i) \*there came [several angry men] into the room.  
vs.  
There came into the room [several angry men]  
cf.  
\*there did [several angry men] come into the room

- (ii) \*there arrived [a strange package] in the mail.  
vs.  
There arrived in the mail [a strange package]  
cf.  
\*there did [a strange package] arrive in the mail

- (iii) \*there was placed [a large book] on the table  
vs.  
There was placed on the table [a large book]  
or;  
There was [a large book] placed on the table

- The DO cannot be extracted either. The extraction facts actually show two things. One, the DP cannot be generated in base position and be extracted by WH-movement, and two, even when the DO moves to a peripheral position on the edge of VP, WH-movement still cannot apply to the DO.

- (iv) \*[how many packages] did there arrive in the mail?  
(v) \*[how many packages] were there placed on the table?

- In other languages, the expected ‘base’ order surfaces (ask Paco for data from Romance). WH-movement of DO is also possible, as we see from Dutch.

(23) [hoeveel mensen] zijn er aangekomen? (Dutch)  
how.many men did there arrive

- Chomsky supposes that in fact, the obligatory extraction of DO may even occur in transitive constructions, if English has Transitive Expletive Constructions (TEC) as argued by Kayne.

(26)

There entered the room a strange man

vs.

\*there entered a strange man the room

\*there a strange man entered the room

cf.

\*there did a strange man enter the room

There will hit the stands a new journal

vs.

\*there will a new journal hit the stands

\*there will hit a new journal the stands

- As you can see, the TEC facts are not entirely parallel to unaccusatives. Plus, I don't know whether 'the room' and 'the stands' are within the VP or not. Be that as it may, let's assume;

(25) In unaccusative/transitive constructions (i.e., [<sub>VP</sub> v [<sub>VP</sub> V DO]] configuration), something must escape vP.

- There is an irrelevant complication here which is potentially confusing. I thought unaccusatives without an optional Experiencer did not have a vP. But in this section (cf. bottom of p.18), Chomsky says that unaccusatives have a vP, which is 'a weak phase'.  
I guess the reason he does this is to maintain (25) in a general form. If indeed the word order facts of TEC and those of unaccusatives are parallel, you want a single story for why the DO is obligatorily extracted. If you are going to implicate some property of v in the account (he seems to. See p.18 for the proposal that 'the phonological counterpart of EPP', which is lodged on v, whether strong or weak, is responsible for the movement), then you want the two to have similar structures.
- At any rate, what exactly is the status of this obligatory extraction? Chomsky calls the extraction in question **Thematization/Extraction** (TH/EX), TH for the leftward version and EX for the rightward version, I assume.
- He notes that the extraction in question does not "yield the usual surface-semantic effects (specificity, etc.)." (p.16). The assumption seems to be that displacement in narrow syntax (Subject Raising, Object Shift, etc.) can in fact have semantic-interpretive consequences. Since the extraction in question doesn't, it is hypothesized that;

(27) TH/EX is an operation of the phonological component.

## 9.2. Properties of TH/EX:

#1. The DO extracted by TH/EX (called EN in the text) cannot undergo WH-movement. Movement is barred from the base position or the derived position of the EN:

- (28) \*[how many packages] did there arrive in the mail?  
\*[how many men] did there enter the room?

#2. However, nominals other than EN can be WH-extracted.

- [to whom] was there a present[EN] given?  
?[at which airport] did there arrive three strange men[EN]?

#3. Nothing can be extracted out of the EN, whether in its moved position or the base position.

- (31) \*[what] are there [book about t](EN) being sold (in Boston these days)?  
(33) \*[who] was there delivered to your office [a picture of t](EN)?  
vs.  
?[who] did they deliver to your office [a picture of t]?

#4. However, EN is not opaque to other processes. It needs to be Case-marked and participates in Binding Theory and Absorption (probably holding on interface representations).

- (34) \*he thought there were [songs about John] being played on the radio  
they thought there were [songs about each other] being played on the radio  
who thought there were [songs about what] being played on the radio?

You should keep in mind an alternative account of these facts which has nothing to do with PF-movement. Namely, a ban on extraction from ‘internal constituents’.

## 9.3. TH/EX as PF-movement:

### Properties of PF-movement:

Given the properties observed above, Chomsky suggests that TH/EX is movement at PF. Let me try to explain a bit more. Remember that lexical items (LI's) are not ‘distributed’. This means that LI is a sign, {[PHON], [FORM], [SEM]}. Movement in narrow syntax, which consists of Agree-Pied-pipe-Merge, targets the relevant projection of LI and displaces the entire feature structure associated with the LI/phrase. PF-movement, on the other hand, displaces/eliminates only the PHON feature matrix of the LI/phrase, leaving the FORM and SEM features intact in the ‘base’ position.

The trace of syntactic movement (syntactic trace) is called ‘**inactive trace**’ in (36).

- (36)b. Inactive trace disallows Match (as well as Agree, etc.)

It is ‘inactive’ because the entire feature structure associated with the ‘trace’ is displaced by Move (more accurately, enters into multiple Merge relations). Thus, we can imagine why such a trace might not by itself (independently of its ‘antecedent’) establish syntactic relations, such as Agree or figure as a blocker in Maximize Match, as stated in (17).

The trace of PF-movement, on the other hand, is visible to Agree, since the relevant properties are still intact in the ‘base’, or trace position. However, the phonological trace (called ‘**trace**’ in 35. May be we should call it **active trace**) of EN is not accessible to Move. Nor can the EN move from the extracted position (NC doesn’t state this in 35, but it should be included).

- (35) i. TH/EX is an operation of the phonological component.
- ii. Traces (phonological traces) are inaccessible to Move, but accessible to some other operations.

Now, being unable to Move seems to be a common property of both ‘inactive’ (syntactic) and ‘active’ (phonological) traces -- as well as of the EN in the PF-extracted position. For the traces, Chomsky suggests that being a trace must make them ineligible for one (or more) of the operations involved in Move – Agree, Pied-pipe, and Merge. Phonological (‘active’) trace can Agree, while non-trace EC’s (PRO, pro) can be Merged, by assumption. It must then be that traces, active or otherwise, *cannot undergo Pied-pipe*.

- (36)a. EC disallows Pied-piping.

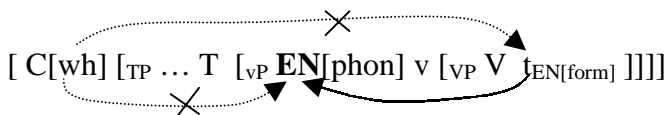
The reason being;

- (37)ii. Pied-piping requires phonological content (i.e., the presence of PHON feature).

Chomsky actually states (36) for all EC’s and not just traces. Therefore, he has to say something fishy about lexical EC’s like PRO/pro being X<sup>0</sup> categories not needing Pied-piping (when they undergo Move). Actually, the implementation does not matter so much as the results he wants. If ‘active’ phonological trace could undergo Move, we are back to a system with **covert Move**, which Chomsky wants to eliminate in favor of Agree, so that’s why he’s fishing for some reason to block syntactic Move without displacement of the phonological matrix.

EN and syntactic Move:

How do we explain the fact that the EN in its base or derived position does not undergo syntactic Move, such as WH-movement? The reasons are as follows. WH-movement has to access the ‘active’ trace in the c-domain of vP. One reason may be that vP is a strong phase. If so, by Phase Impenetrability, C[wh] may not access the base position of EN.



This does not seem to be the right story, since vP in unaccusatives is not a strong phase according to Chomsky. If vP is not a strong phase, C[wh] can access the base position of EN

directly. How do we then prevent EN from undergoing WH-movement? The only way is to assume that TH/EX applies before C[wh] is Merged into the structure. This may be why Chomsky assumes that TH/EX applies **at** vP (bottom of p.18). We shall see that the assumption that TH/EX is obligatory at vP turns out to be problematic. We have to allow TH/EX to ‘wait’ at least until TP is Merged. So, may be PF-movement in a given strong phase must take place before the phase-determining head is Merged into the structure. I don’t know what guarantees this result.

What about the derived position of EN? EN that is PF-moved to the Edge of vP is accessible from C[wh]. However, it does not possess the requisite feature structure (i.e., FORM feature structure) and cannot undergo syntactic Move.

#### PHON matrix and syntactic Move:

Notice that in this view, narrow syntax, while not sensitive to the content of the PHON feature matrix, is sensitive to the presence or absence of the matrix in an LI. We may also consider the implications of PF-movement for the nature of LEX. We saw that narrow syntactic Move displaces a projection of an LI that is not distributed, while PF-movement can selectively target only the PHON feature structure of LI. This allows certain ‘distribution’ of the feature matrices, but covert Move in narrow syntax is still disallowed. Thus, radical ‘distribution’ of matrices is not countenanced.

#### 9.4. Reduced Relatives (??):

If the assumption that PF-movement does not iterate is valid, we seem to have a problem below:

- (38)i there are expected to be caught [many fish](EN via TH/EX)
- ii there are expected to be [many fish] caught ( “ )
- iii there are [many fish] expected to be caught (iterative TH/EX?)
- iv [many fish] are expected to be caught (TH/EX feeding A-move?)

Chomsky says that (38iv) is ‘unproblematic result of successive-cyclic A-movement with the intermediate stage (39)’.

(39) [\_\_ are expected [TP many fish T-to be [vP caught \_ ]]]

I am not sure that it is ‘unproblematic’. To derive (38iv) from (39), you must allow **either** A-movement **or** TH/EX to apply, but not both. That is, as long as the DO escapes the VP in unaccusative structures (may be transitive structures more generally, as in 25), TH/EX need not apply. Since A-movement has this effect, when A-movement applies, TH/EX becomes optional.

A problem arises if we assume that TH/EX applies obligatorily (at vP). Assume that TH/EX obligatorily applies at vP/VP, T can still access the ‘active trace’ of DO, as T is the Edge of vP (which may not even be a strong phase). T Agrees with the trace of DO, values its Case and has its own phi-features valued. Now, by assumption, the active trace of DO cannot undergo Move. The question therefore is how did ‘many fish’ get displaced to the Spec of T? Not by syntactic

Move. It must be by PF-movement, but then PF-movement is not supposed to iterate, by hypothesis. On the other hand, if TH/EX does not apply, this problem no longer arises.

A rather stipulatory alternative mentioned earlier that allows TH/EX to apply at TP, rather than vP also works empirically, as you can tell.

If TH/EX is optional within vP/VP, we must ensure that we still have an account of why the DO/EN does not undergo syntactic WH-movement. I.e.,

\*[how many packages] did there arrive in the mail?

\*[how many men] did there enter the room?

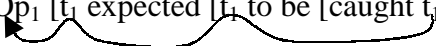
The reason we need to come back to this paradigm is the following. Since ‘there’ is in SpTP, we know that EN cannot syntactically Move to SpTP. However, if the EN can somehow **syntactically** Move to the Edge of vP, carrying its FORM feature matrix, then from there it should be accessible to C[wh] and undergo WH-movement. What could the relevant syntactic Movement be? It should be Object Shift (OS), if anything. It must be that OS is impossible in these sentences, as opposed to ‘real’ transitive sentences.

Another lingering problem is that if unaccusatives lack vP or have a vP that is not a strong phase, C[wh] should be able to directly access the DO/EN in its base position. Someone should write to Chomsky or write a squib.

#### Reduced relatives with ‘existential import’:

More problematic is (38iii), where the surface position of the EN nominal not an A-position but a position to filled by TH/EX movement. It seems that TH/EX has iterated, contrary to predictions. Chomsky says that (38iii) is a ‘true existential construction’ with a reduced relative clause. I.e.,

There are [[many fish]<sub>1</sub> [<sub>Op</sub>1 [<sub>t</sub>1 expected [<sub>t</sub>1 to be [caught <sub>t</sub>1 ]]]]]



As evidence for this analysis, Chomsky cites the following facts:

- #1. (38iii) unlike the other sentences in (38) has ‘existential import’ – cf. (40).
- #2. (38iii) is an island for extraction while others aren’t – (41iii vs. iv) (both are \* for me)

Whether or not you agree with this story, notice a couple of things about this alternative analysis.

First, in the putative reduced relative clause above, the DO/EN has in fact undergone syntactic WH-movement. The movement was fed by A-movements, as you can see. Again, we must assume that TH/EX is optional and in a possibly complementary distribution with A-movement (to SpTP, or with OS, as we shall see later).

Second, notice that (38iii) could be derived if A-movement can ‘feed’ TH/EX, under the assumption that the two movements in the lowest vP/VP are optional. The sentence arises from iterative TH/EX only under the assumption that the movement is obligatory, an assumption (which Chomsky seems to be making) that we have found to be problematic. The alternative derivation is shown below:

There are [<sub>VP</sub> many fish[phon] v [<sub>VP</sub> expected [<sub>TP</sub> t<sub>act</sub> T-to [<sub>VP</sub> [<sub>VP</sub> be caught t<sub>inact</sub> ]]]]]]

There is yet another analysis, where ‘many fish’ and ‘expected to be caught’ form a Small Clause which is selected by ‘be’. Facts discussed below make this alternative attractive.

Final comments on this section:

I have serious doubts about the existence of TH/EX, especially, the leftward version. It seems that unless you have a ‘be’, the constructions are ungrammatical.

- \*There arrived [a small package] in the mail (\* because EN is in base position)
- \*There did [small package] arrive in the mail (should be o.k. because TH/EX applied, but \*)
- \*There will [a small package] arrive in the mail (ditto)
- \*There has [a small package] arrived in the mail (ditto)

Leftward TH/EX in ‘transitive expletive constructions’ is also quite bad.

- ?There entered a room [a stranger].
- \*There will [a stranger] enter the room.
- \*There has [a stranger] entered the room.
- \*There did [a stranger] enter the room.
- cf.
- ?There was a room entered by a stranger (this certainly is a reduced rel. clause)

## 10. Object Shift, TH/EX, and EPP:

Leftward TH/EX in English looks similar to **Object Shift** (OS) in other languages, but given that there are no “surface semantic effects” associated with the former while the latter often has such effects, Chomsky concludes that the two are different – TH/EX is PF-movement while OS is narrow syntactic movement.

### Sidebar:

The assumption that narrow syntactic movement may give rise to ‘surface semantic effects’ is an idea that meshes with certain leading ideas in Construction Grammar. Chomsky has recently written things to the effect that A-movement to SpTP creates a configuration that expresses ‘agentive force’ (though I don’t see how ‘agentive force’ is a **surface** semantic effect. Theta-roles are supposed to be **deep** semantic phenomena, expressed configurationally). OS is often correlated with specificity, as many have pointed out. In Chomsky’s view, it isn’t that only [+specific] DPs undergo OS. Rather, it is the configuration (outer Spec of vP) that OS creates which is correlated with this interpretation. As a result, DPs which cannot be construed as specific may appear not to undergo OS.

### 10.1 Problem raised by OS:

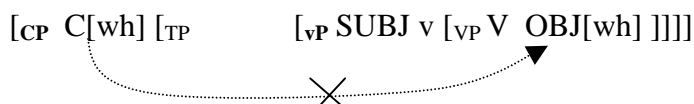
#### Prima facie problem:

Languages differ with respect to OS – some allow it (Icelandic, Scandinavian), while some don’t (English).

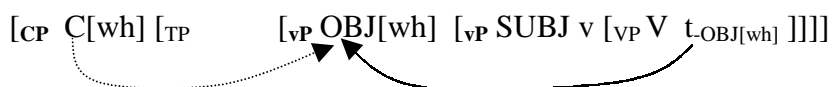
#### A restatement of the problem:

It’s not that there are languages which don’t allow OS. All languages allow OS, but in the apparently non-OS languages, the OS-ed DP must move on.

O.k. Let’s try to make the connection between the two. For Chomsky, the prima facie problem can be restated as above because WH-movement (of object DPs, at least) in **all** languages must proceed via OS. You may wonder why. Consider the following.



Given the **Phase Impenetrability Condition** (cyclicity), C[wh] can access only the Edge of vP/VP. Now, if OBJ stays in-situ, it can of course be Case-marked (by Agreement with v), but will not be able to undergo WH-movement to SpCP. Therefore, it needs to get out to the Edge of vP. Since the lower Spec is filled by the SUBJ, it goes to the outer Spec of vP, where it can be accessed by C[wh]:



Chomsky seems to think that this step is *none other than OS*.

If you think about it, though, there are reasons why you **don't** want to take this step to be OS. To see this, we have to compare the system with that in Barriers. What the PIC does is make  $vP$  a barrier for direct WH-movement of OBJ to SpCP. This is why anything that is moving out of  $vP$  must get out to the Edge of  $vP$  in order to get to SpCP – anything, including adjuncts! This is where the problem arises. Adjuncts don't undergo OS. So adjunct WH-movement out of  $vP$  cannot proceed via OS.

Why then does he want to say that OS mediates WH-movement out of  $vP$ ? It is because analyzing WH-movement in the system in *DBP* is not trivial, as we shall see below. A second reason is that he wants to employ the same system for explaining 'real' OS to account for the first stage of WH-movement (of OBJ) out of  $v^*P$ , as we shall see subsequently.

#### Analyzing WH-movement on the model of A-movement (as feature-driven movement):

As a first approximation, we need to endow  $v$  with (what Chomsky called in *MI*) a "P-feature" that Agrees with XP[wh] and displaces it to its Edge. The feature that Agrees with XP[wh] must be like phi-features. That is, it is presumably uninterpretable on the Probe but interpretable on the Goal. So, let's assume that WH-movement inducing Probes C and  $v$  have an uninterpretable/unvalued version of this feature – call it WH. Now, given the *DBP* system, the XP[wh] must have an interpretable/valued version of this feature. Assume that it does. It must also have something akin to the unvalued Case feature if it is to be 'active' (remember, only active phrases can undergo Agree or Move). It isn't easy to think of what this feature could be.

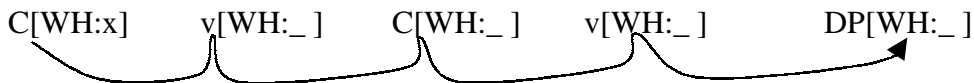
A possible line of attack is to capitalize on the intuition that wh-phrases are semantically indefinite quantifiers. However, unlike other indefinites, wh-phrases move. So, we may think that it is the wh-morphology on XP[wh] which is responsible for this movement. If so, this has to be the equivalent to Case which makes the phrase 'active'.

There are a number of questions we need to deal with before the preceding intuition can be made to work.

For example, what is the WH on the Probe Agreeing with? Is it Agreeing with the indefinite quantifier part of a wh-phrase or with the (uninterpretable) wh-morphology (we are now assuming) of the wh-phrase? Another question is the following. Is the WH feature on the Probe uninterpretable? The answer seems to be negative -- an interrogative C is not like the 'small  $v$ '. More generally, the C in whose Spec WH-movement finally rests is different from intervening Probes through which WH-movement takes place. In the terminology of GB, the former is a [+wh] Comp, while the latter are all [-wh] Comps. For [-wh] Comps, it makes sense to think of the WH feature on them as uninterpretable. However, on a [+wh] Comp, the feature seems to be interpretable. This is unlike the phi-feature set which is always uninterpretable on Probes, whether it is on a full phi-featured T (finite T) or on a defective T (non-finite T).

One possibility is to assume that the interpretable WH on a [+wh] C Agrees with the uninterpretable wh-morphology of the WH-phrase. Recall that in earlier comments about Case being valued by uninterpretable phi-features on T,  $v$ , I pointed out that this state of affairs leads to a violation of Weak Interpretability. I suggested therefore that Case should be valued by the interpretable Category feature of the Probe, rather than the phi-features. What I am suggesting here with respect to WH is in a similar vein. The uninterpretable wh-feature on Goals is checked by the interpretable Categorical feature (C[wh]) on the Probe.

If we want to make the intermediate Probes through which WH-movement passes behave like defective T's in A-movement, we would need to assume that they are not by themselves able to value the uninterpretable wh-feature of WH phrases but that they have an uninterpretable version of the feature on the C[wh] that partially Agrees with the WH phrase. By 'transitivity of Agreement', the intervening uninterpretable wh-features on Probes would be valued when that on the WH-phrase is valued by Agreement with the C[wh]. I.e.,



What about the Movement of WH phrases? For Move, the Probes C and v need EPP features. The EPP feature in question should be satisfiable by phrases other than DPs, so long as they are WH-phrases. Therefore, we must make it be sensitive to whether or not a phrase has WH-features (i.e., an A'-Spec, in the system of Rizzi 1990), since otherwise, DPs in intervening A-positions will block WH-movement, contra Relativized Minimality.

In fact, if you think about it, the intermediate stages of WH-movement may take place not for reasons of defective Agreement, as discussed above, but solely for reasons having to do with locality of Movement. That is, if the intervening strong phase heads did not have (an optional) EPP feature, the EPP on the highest C[wh] could not access the WH that will Move to its Spec.

If this is correct, and if we still want to maintain that the C[wh] and the WH-phrase that ultimately moves to it enter in an Agree relation, we have an extremely **non-local Agreement**. I.e.,



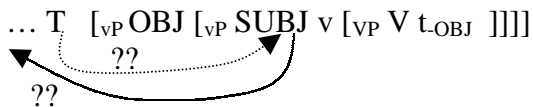
However, such non-local Agreement is not sanctioned in the system, as Agreement into the c-domain of a lower strong phase head is not allowed by the **Phase Impenetrability Condition**. This leaves us with two options – one, there is no Agreement between C[wh] and WH-phrase, so that WH-movement is like Topicalization, Object Shift, etc., and two, the intervening heads do have a 'weak' version of WH that Agrees with the WH on the Goal, but does not value it completely (as was suggested above). Neither of these alternatives is fully satisfactory. We may actually want to re-examine the question of whether the locality of Agreement (to include the above cases, treated in the literature so far as **Unselective Binding**) and the locality of Movement should both fall under the same Phase Impenetrability Condition.

Let's drop this interesting conjecture for now, and go along with Chomsky and assume that WH-movement of Objects at least proceeds via OS.

#### Apparent violation of Movement locality in OS:

Even without the aforementioned problem, OS raises a problem for Movement locality. Since Move is triggered by the EPP feature on Probes, and the metric of closeness is c-command, unless something else is said, an OS-ed OBJ should block T from Agreeing with the SUBJ in the lower Spec of vP (given the **Defective Intervention Constraint** according to which a full phi-featured DP is a blocker, whether or not the phi-feature set is valued/inactive). Neither should T

be able to trigger the Movement of the SUBJ in lower SpecvP over the OBJ to SpTP (since OBJ is a closer to SpTP. Again, by Defective Intervention, it does not matter that OBJ is inactive). I.e.,



(You might ask : how does OBJ Move over SUBJ in the first place ? The answer that is suggested is the **Merge-over-Move** Preference. That is, when v is Merged with VP, there is either a choice of Moving OBJ to its Spec or Merging SUBJ as its Spec. The idea is that the latter is cheaper. SUBJ does not block Agree(v, OBJ) as v does not c-command SUBJ. v has an extra EPP requirement beyond that satisfied by the Merger of SUBJ. Therefore, OBJ moves to the outer Spec in OS.)

An earlier stipulation allowed Move in such cases.

Equidistance Provision (Chomsky 1995):

(43) Terms of the minimal domain of H are equidistant from probe P.

That is, though OBJ c-commands SUBJ asymmetrically, they are both within the ‘minimal domain’ of v, where **MinDom** was defined as the Spec, Compl, and Adjuncts of a given head (cf. *MPLT* in Chomsky 1995).

Since the issue at hand has to do with elements on the Edge of H(=v), Chomsky suggests that this can be stated more narrowly as:

(44) Terms of the Edge of HP are equidistant from probe P.

However, in non-OS languages like English, there is an additional complication. Given that OBJ cannot ‘stay’ in its OS-ed position, the Equidistance provision seems to hold only if the *OBJ that is closer to T than SUBJ moves out of the OS-ed position*. Icelandic OS of Nominative OBJs with quirky SUBJs as (45) also reinforces this conclusion, where it seems that the OS-ed OBJ moves on beyond the Edge of vP. (We will come back to 45 later)

An apparent violation of cyclicity:

If derivations are strictly cyclic and bottom-up, you access T before you access C. This creates a problem. When T is accessing SUBJ in lower SpvP (for Agree or Move), the OS-ed OBJ is still in the outer SpvP. It is only when C is Merged above TP and C attracts the OS-ed OBJ to SpCP that the outer SpvP becomes phonologically null. In other words, T has to ‘know’ whether or not the OS-ed OBJ will Move to SpCP in order to Agree with and trigger the Move of SUBJ. This seems countercyclic, requiring *look-ahead*.

Two problems:

So, we have two problems – (i) why must the OS-ed OBJ move on beyond vP? (ii) how do we account for the apparent non-cyclicity of Agree(T, SUBJ)?

Regarding the first, Chomsky restates (44) as (46):

(46) The phonological edge of HP is accessible to probe P.

That is, in the following configuration, if XP has phonological content, SPEC is not accessible to the Probe P.

(47) [<sub>PP</sub> P [<sub>HP</sub> **XP** [<sub>H'</sub> SPEC [<sub>H'</sub> H YP ]]]

Looking ahead, we can see why this generalization would hold. Recall from the previous section that an **inactive trace** (trace of syntactic movement) does not admit Agree or Move (the latter as a consequence of lacking the ability to undergo Pied-pipe). Therefore, if XP is occupied by an inactive trace, SPEC is the closest element that could Agree with P, because Trace(XP) is not something that could figure in Agree or Move, by assumption. On the other hand, if XP is lexical, it is the closest thing that could Agree with P (or Move to SpP), and hence blocks Agree(P, SPEC). *There is thus no such thing as the Equidistance Principle.* (However, think about PF-movement. If PF-movement only displaces the PHON feature matrix, then the ‘active’ trace at the outer Spec of v\*P should block Agree(P, SPEC). So it isn’t clear to me how vacating the Edge of v\*P by DISL or cliticization will do any good.)

What about (ii)? **Next Phase Interpretation** will take care of the apparent non-cyclicity. Recall that the Edge of a lower strong phase belongs to the next strong phase for the purposes of Spell-out. Thus, the derivation at the Edge of vP cannot fail or succeed depending on the phonological properties of XP occupying the Edge. It is only at the next strong phase that this information becomes relevant. (46) is going to be evaluated *at the stage where the Edge of vP is spelt out*, namely, CP. If C has an EPP feature which attracts the XP at the Edge of vP, (46) is satisfied at CP. If not, it fails.

Intervention effect at the Edge – Move vs. Agree: (p. 23–24)

We have assumed with Chomsky that inactive trace is invisible to both Move and Agree. Chomsky wonders whether there might possibly be a distinction between Move and Agree with regard to the intervention effect of inactive trace (specifically, inactive trace blocks remote Agree, though no Move).

10.2 OS vs. non-OS languages:

Two possibilities :

**Possibility #1 :** [+OS] languages allow T to Agree with in-situ SUBJ even when OBJ is sitting on the outer Spec of vP (i.e., in its OS-ed position). [-OS] languages allow such access only if OBJ has vacated the Edge, a la (46), giving the appearance that OS has not applied. The parameter may have to do with “richness of T”.

**Possibility #2:** (46) holds in both types of languages. In [+OS] languages, the OS-ed OBJ can move on to a position that is lower than the surface position of the SUBJ, yielding SUBJ-V-OBJ order, while in [-OS] languages, there is no such word order option.

Some evidence for the second alternative:

- (i) In MSc, pronoun OBJs but not full NPs undergo OS. Possibly the OS-ed pronoun OBJ moves on to a higher position by something like cliticization and/or phonological movement (similar to TH/EX).
- (ii) In Icelandic, (45)/(53) might show a similar effect.

(45)/(53)

[many students.DAT]<sub>SUBJ</sub> find.PL<sub>VB</sub> [the computers.PL.NOM]<sub>DO</sub> [t<sub>SU</sub> t<sub>VB</sub> t<sub>DO</sub> not ugly.PL]

- The Nominative OBJ Agrees with T (Quirky SUBJ does not count as closer XP, perhaps due to its Quirky Case).
- Quirky SUBJ (generated as Spec of v) Moves to SpecTP by Attract Closest.
- Nom OBJ undergoes OS, attracted by v. Ends up on Edge of vP.
- OS-ed OBJ escapes Edge, by a phonological rule *DISL*.
- *DISL*, like TH/EX, cannot take OS-ed OBJ to SpTP (or higher positions).

### 10.3 Holmberg's Generalization:

Holmberg's Generalization (original form):

Object Shift occurs iff V raises out of VP (to T).

Problems raised by HG:

- (i) OS must 'look ahead' to determine whether or not V will raise to T (apparent countercyclicity).
- (ii) If WH-movement of OBJ in non-OS languages is OS, then there is a violation of HG with such movement as the verb does not raise out of VP/vP.
- (iii) OS seems sensitive to 'phonological adjacency', as follows:

- In MSc, if V raises to SpC/C (by 'verb topicalization'), even if an Aux occupies T blocking V-to-T, OS is still possible. The countercyclicity problem here is even worse. Shifted OBJ should be able to look ahead to the point where C is Merged to determine whether V will vacate the VP.

The diagram shows a syntactic structure:  $[[V]_C \quad [Aux]_T \quad [OBJ]_{VP} \quad [t_V \ X \ t_{OBJ}]]$ . An arrow points from the shifted object  $[OBJ]_{VP}$  to the verb  $[V]_C$ , and another arrow points from  $[OBJ]_{VP}$  to the auxiliary  $[Aux]_T$ , illustrating the 'look ahead' problem.

- Any element which intervenes between OBJ and the OS-position blocks the OS of pronouns in MSc.

HG restated (Holmberg 1999):

(54)

- (i) OS is a phonological operation that satisfies condition (ii) and is driven by semantic interpretation of the shifted object (new information, specificity/definiteness, focus, etc. – **INT** properties).

- (ii) OS cannot apply across a phonologically visible category asymmetrically commanding the object position except adjuncts.

Problems with the new HG:

- (i) It requires countercyclic operations.
- (ii) Unlike TH/EX, it does not behave like a rule of the phonological component, as it has ‘surface semantic’ consequences.
- (iii) Clause (54ii) does not hold in ‘invisible’ OS of WH-OBJ in English. Nor does it hold in ‘visible’ OS in Icelandic.
- (iv) The assumption that interpretation ‘drives’ operations is also problematic.
- (v) The rule does not have characteristics of phonological processes that take place under ‘adjacency’, such as T-V assimilation.

10.4 HG Reanalyzed – [+/-OS] Parameter:

General background:

What Chomsky is trying to do here is not just offer a reanalysis of OS, but in fact propose a way to account for **non-feature driven optional movements** which often have semantic consequences, such as Topicalization, Scrambling, and Object Shift. A-movement is a feature-driven movement in the sense of *MI* in that the moved DP has an uninterpretable feature that enters into Agreement with a Probe. This is why A-movement often ends up in the Spec of the Probe/Head that Agrees with the Goal DP, giving the appearance that the movement itself was ‘driven’ by the need to eliminate the uninterpretable Case feature on the DP.

Another movement that may be similar to A-movement is WH-movement, in the sense that the WH XP possibly has an uninterpretable “wh” feature that is like Case that is eliminated by Agreement with a Probe. However, in earlier comments, I pointed out some difficulties for this view of WH-movement.

Both of these movements are, in the words of Rizzi (1997), ‘**Criterion-based**’ movements, where some XP with feature P moves to some head with feature P’. Unlike these movements, we have optional movements like Topicalization, Scrambling, Focus movement, and Object Shift where (i) there does not seem to be an uninterpretable feature on the moving DP and, (ii) there is no obvious head that serves as the trigger of such movement. Rizzi (1997) actually proposes a system where these movements too fall under ‘Criterion-based’ movements, but at considerable conceptual and technical cost. For one, he has to posit optional projections of TopP and FocP into whose Specs XPs that are [+topic], [+focus] move to ‘check’ their features.

However, if you think about it, this is quite *ad hoc*. For one, in Italian at least, there is no overt head that is dedicated to expressing Topic and Focus. For another, the assumption that a semantic feature such as [+topic] drives movement does not sit well with a system where the movement operation itself is not triggered by semantics (i.e., there is no Greed).

Against this background, Chomsky is suggesting that these movements are better handled by imbuing certain heads with optional uninterpretable EPP features. These features drive the movement. However, there is no Agreement between the moved DP and the Probe that is responsible for the interpretation. A stipulation to the effect that the optional EPP feature can only be added if it has an effect on the output that is not available in the input (60 – Have An Effect on Output Condition, *HEOC*) ensures that the features will give rise to ‘surface semantic’

interpretations that are not available in the base. The actual interpretation, however, does not arise from the inherent properties of the moved XP, but rather from the *configuration* created by the movement being interpreted in one way or the other. This means that if an XP that does not admit the relevant interpretation moves, the derivation will be well-formed, but interpretation will fail – that is, the derivation will “converge as gibberish”.

Assumptions:

1. Semantic properties (**INT**) cannot ‘drive’ the movement. These properties are those read off the configurations/constructions that result from movement which involves ‘dumb’ formal features, i.e., an optional EPP feature on v\*. If a DP that fails to meet INT is moved, then the representation will be thrown out at the interface as deviant (i.e., the derivation ‘converges as gibberish’, as Chomsky used to say).
2. Assume that in a trivial A-chain (yes, we have these again, since covert Case-seeking movement is gone!), the single member can be assigned either a ‘base’ semantic interpretation **INT’**, or the ‘surface’ semantic interpretation **INT**.
3. In a non-trivial A-chain, the head position is assigned INT, while the base position is assigned INT’. I.e.,

$$C = \langle \underset{\text{INT}}{\text{DP}}, \dots, \underset{\text{INT}'}{\text{DP}} \rangle$$

The system: (61)

- (A) v\* is (optionally) assigned an EPP feature. (constrained by 60, HEOC)
- (B) The EPP position of v\* is assigned INT. (=57)
- (C) At the phonological border of v\*P, XP(OBJ?) is assigned INT’.
- (Yes = OS language; No = Non-OS language)

**Note 1:**

Phonological border = phonological (left) edge material, not to be confused with syntactic edge.

**Note 2:**

The configuration described in (C) can only arise if V raises out of v\*P. Therefore V raising is a prerequisite to OS. This is true even in languages where OS is possible otherwise. If V cannot raise out of v\*P, (C) is not relevant.

You may find this reasoning strange. I do. But Chomsky seems to adopt this to account in part for the original insight behind HG – that V must raise in order for OS to apply.

Case 1: Non-OS language (not observing C) with V raising out of v\*P/VP:

- INT or INT’ can be assigned freely in the base position of OBJ.
- Under (60), EPP feature may be added to v\* only if it is necessary to guarantee an interpretation of the Specv\*P that is not INT or INT’. For example, an EPP feature may be added to guarantee WH-movement to Specv\*P, which by assumption is neither INT’ nor INT. (Note: this must mean that WH-movement to Specv\*P is ‘meaningful’ at v\*P,

though it never ‘stops’ there – see what this does to the question I raised earlier about implementing WH-movement in the *DBP* system)

- Otherwise, OBJ stays in-situ, whether or not it is at the phonological edge of v\*P (i.e., whether or not v\*-V has raised out of v\*P).

#### Case 2: OS language (MSc, Icelandic – observing C) with V raising:

- If OBJ is v\*P internal and not at the phonological edge (because an Aux raised to T, or some other non-adjunct category is at the edge), OBJ can be assigned either INT or INT’ as in non-OS languages. If OBJ moves to Spv\*P, it can only be for WH-movement, as described earlier. (Note: as should be obvious by now, if not already, this shows again how misleading it is to consider the first stage of WH-movement from within v\*P as Object Shift. Clearly, ‘Object Shift’ of WH-OBJ is possible in OS languages even when OS of non-WH OBJs is impossible!)
- If OBJ is at the phonological border of v\*P (say, there is no Aux and only adjunct intervenes between the edge and T), it is assigned INT’. If v\* has an EPP feature, OBJ moves to Spec, yielding a two-membered Chain where the head of the Chain (in Specv\*P) is assigned INT.
- If OBJ is something that does not admit INT interpretation (indefinite NPs, etc.), the derivation converges as gibberish.
- If OBJ is something that must have INT in addition to INT’ (pronouns – see discussion of (62i, ii) in the paper – again, another assumption slid in through the backdoor), not applying OS via an optional EPP feature on v\*P will again lead to a derivation that converges as gibberish.

#### Case 3: Non-OS language without V raising out of v\*P/VP:

- (61C) cannot hold, since the phonological border of v\*P is occupied by the unraised v\*-V.
- Therefore, the language is necessarily non-OS.

#### **Note:**

Under the reasoning thus far, you may ask why the optional EPP feature on v\* in such languages (English) could not be tied to INT, so that English has OS without V raising. Chomsky’s answer is the since INT or INT’ (or both?) is available by hypothesis in the base position of OBJ, there is no reason to induce INT interpretation by an optional EPP feature on v\*. Therefore, the feature can only be used to signal interpretation that is neither INT nor INT’, i.e., WH-movement.

Again, many unclarities remain. What if the base position of OBJ is assigned only INT’ but the OBJ is something that requires INT? Couldn’t we then assign EPP to v\* to signal INT? How do we ‘know’ in ‘assigning an EPP feature’ to v\*P that this feature will in fact have an effect on the output? Does the make-up of v\*P or the partial LF constructed thus far (if there is one, remember next phase interpretation?) guarantee that you will need an EPP feature? (A) of (61) does not seem to fit in nicely in a system that is supposedly ‘dumb’.

#### Case 4: OV languages:

- In OV languages (by which Chomsky means German and Dutch, not Japanese and Korean), if the language is [+OS], OS is always possible.

**Note:**

I don't understand what it means for OV languages to be [+/-OS]. I thought the parameter is (C) and that it didn't make sense to inquire about whether or not (C) holds unless V raising is possible. Apparently the correlation between V raising and (C) holds only in VO languages, but not in OV languages where regardless of V raising, the option exists to adopt (C). If the option is taken, OS happens to all OBJ's, as they are necessarily at the phonological edge of v\*P (the only things that could precede OBJ are by assumption adjuncts).

Case 5: Icelandic OS with SUBJ in-situ:

(63) there read<sub>T</sub> [<sub>VP</sub> **it** (never) [<sub>VP</sub> any students(SUBJ) [<sub>VP</sub> t<sub>OBJ</sub> ]]]

- This raises problems since the SUBJ is at the phonological border of vP but nevertheless OS has occurred.
- So, what's going on? Chomsky says that "principle (25)" that says something must escape transitive v\*P (since when did this become a 'principle'? Watch out for these backhanded maneuvers) guarantees, in accordance with (60), that v\*P must have an EPP feature which attracts OBJ.

**Note:**

It seems like an OT system, where (C), which is implicationaly related to the existence of V raising, can be overridden by 'principle (25)'.

Tentative conclusions:

- The reanalysis of HG confirms the basic outlines of the system so far.

(64) [<sub>CP</sub> C [SPEC T [<sub>v\*P</sub> XP [<sub>v\*P</sub> SUBJ v\* [<sub>VP</sub> V OBJ ]]]]]

- OS takes place at v\*P, but the conditions on it (61) are evaluated at the next phase CP. At CP, we can determine whether V raising has taken place (to T or C).
  - This is in accord with (10), next phase evaluation.
  - Spell-out takes place after evaluation at CP, namely, as the derivation exits CP.
- The basic intuition that OS involves both phonology and semantics is worked out through a simple parameter (61C) and a 'dumb' computational system where operations are not triggered by semantics.
  - There is no recourse to countercyclic operations violating the Extension Condition.

**Note:**

The last sentence must be the understatement of the day! The Extension Condition in its original incarnation stated that derivations must 'grow' in size with no backtracking. A

generalized version of this condition would state that at a given point in a derivation, all operations – structure-building or otherwise – can only look at a very local domain, namely, the information available in the Label of the **local tree** constructed by the derivational system up to that point, but not anything inside the local tree (*backtracking*) or anything that becomes available in subsequent stages of the derivation (*look-ahead*). Non-transformational theories such as GPSG-HPSG also have this view of locality, though in a strictly representational system. This is why, for instance, WH-movement is modeled as feature passing from one local tree to the next higher local tree in PSG frameworks.

The **Next Phase Interpretation** principle (10) actually allows a lot of backtracking, though it eliminates look-ahead. That is, the local domain that is available for inspection at a given stage of the derivation is quite large. This is done so that problems of ‘countercyclicity’, by which Chomsky means *the need for look-ahead in a strict cyclic system that conforms to strict Extension*, can be avoided and reinterpreted as *backtracking* to earlier stages of the derivation (actually, to the syntactic representation built by earlier derivational stages) within adjacent strong phases.

### 10.5 Speculations on v\*-V to T ‘assimilation’:

V-raising (including Affix Hopping) to T/C might be a rule of the phonological component, like TH/EX or DISL. This is not to say that all head movement is phonological. Perhaps, Bakerian incorporation is not. The reasons for supposing that inflectional V-raising is phonological are as follows.

- (i) If LF interface is uniform, languages do not differ at LF regardless of whether V raises or not. In other words, V and T and C might all be in their ‘base’ positions at LF universally, or it might be that at LF, V-T-C (and D-N) are brought together universally into “super-categories”.
- (ii) V-raising has no semantic consequences. This is typical of PF-movement.
- (iii) There are technical problems in guaranteeing which category moves to Head and which to Spec in case a (higher) Head has selectional features for both a Head and an XP. That is, Chomsky argues, why does XP end up in Spec and the lower head end up (adjoined to) the head? He says that if V-to-T is ‘phonological movement’, such problems do not arise.
- (iv) Verb second might also be PF-movement. In the old, single Spec system, there was a syntactic argument for Verb Second. But now, with multiple, iterating Specs, there is no syntactic basis for V2. (He says in a related footnote that the single Spec idea does not have logical priority over the multiple Spec idea. He says we should not confuse logic with history! Hmm. I thought that it was he who suggested that V2 could be naturally accommodated by the single SpecCP idea.)
- (v) There are other difficulties with head-movement, such as:
  - a. It is an adjunction rule, unlike other rules in narrow syntax.
  - b. It is countercyclic in ways that are not overcome by the DBP system. (Does anyone understand what this means?)
  - c. The raised head does not c-command the trace.

d. Head movement observes ‘somewhat different’ locality conditions.

All of these are unproblematic, he says, if (inflectional) head movement is a phonological process reflecting affixal properties. This is the line in Distributed Morphology (Bobaljik, Halle & Marantz), as you know, though in DM, there are those who want to posit both Head Movement as well as PF Merger for different types of morphosyntactic mismatches (Noyer, Embick).

Another way to eliminate Head Movement is to recast it as (remnant) phrasal movement (Koopman, Szabolcsi, Mahajan, following certain suggestions in Kayne 1994 about agglutinative morphology). One thing to note about this line of research is that a conceptual consideration these researchers rely on to ban Head Movement is the strict version of the Extension Condition. With the Extension Condition relaxed considerably, it is no longer a theoretical hurdle, though other problems remain, as noted above.

11. Sundry closing remarks:

Chains:

In MI, it is proposed that Chains should be construed as occurrences – i.e., the structural sister constituents with which an argument enters into syntactic relations.

John was killed

$CH(\text{John}) = \{K_1, K_2\}$

where;

$K_1 = T'$  (T-was killed),  $K_2 = V$  (killed)

Chomsky says there have been objections to this construal of Chains as occurrences (Epstein & Seely).

#1. An occurrence of argument includes  $X'$ , which is supposedly “invisible” to syntactic computation.

#2. ‘Chain’, which is defined as a set of occurrences derived by ‘multiple Merge’ of an identical element, is not a “legitimate syntactic object”, as the notion is defined in Chomsky (1995).

Chomsky says that these questions may not be ‘real’ questions, but simply “terminological artifacts”.

‘Invisibility of  $X'$ ’:

The supposed invisibility of  $X'$  rests on two assumptions – one, the supposition that  $X'$  is not interpreted at LF, and two, the assumption allows us to incorporate Kayne (1994). Chomsky says the first assumption is questionable at best, and the second, even if it were valid (it isn’t clear that

is it in the current system), could be handled by exempting (X', YP) from the asymmetric c-command relation.

He notes that you can avoid this problem altogether by defining 'occurrence' in terms of dominance (Immediately Contain), rather than Sisterhood. Recall that these are the 'free' relations that the indispensable syntactic operation Merge gives you, according to Chomsky. How would a dominance-based characterization of 'occurrence' work?

John T-was killed t  
 CH(John) = {K<sub>3</sub>, K<sub>4</sub>}  
 K<sub>3</sub> = TP, K<sub>4</sub> = VP

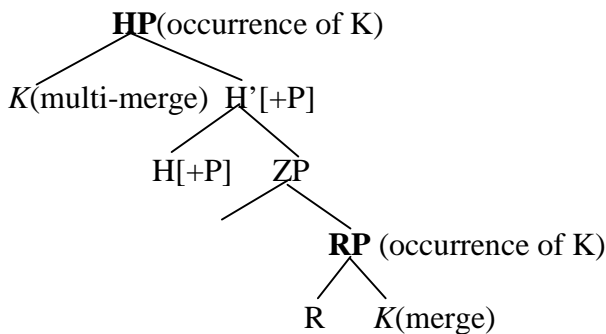
Chomsky says this technical implementation might be better than a Sister-based one as the notion "occurrence-of" becomes asymmetric. That is, while Sisterhood is a symmetric relation, Immediate-Dominance is not.

Status of Chains and Multiple Merge:

We first need to identify what the problem is. It has to do with apparent conditions/principles which refer to Chains. Now, if Chains are not legitimate syntactic objects, we could not have any principles that refer to them. Another problem with Chains is that while Move is a derivational operation, Chain is a representational concept. In a strictly derivational system, there shouldn't be principles which refer to a representational concept like Chain. (Note though, that with Next Phase Interpretation, the syntactic derivation can refer to properties of the partial representations in adjacent phases before these are shipped off to the interfaces).

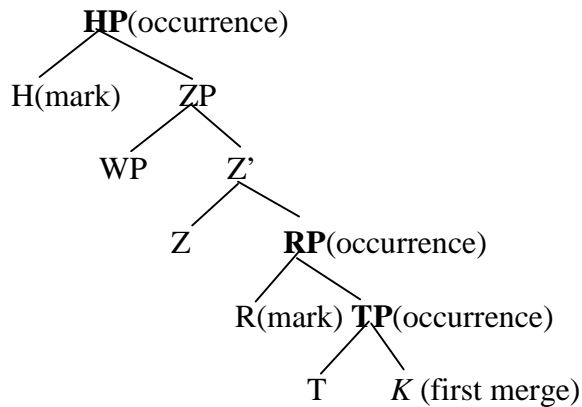
Chomsky says that principles which refer to Chains (Chain Uniformity Condition) can be expressed as conditions on the occurrences of the element entering into the Chain relation. Interpretative principles which refer to Chains (INT, INT' interpretation discussed earlier, reconstruction) can also be so stated.

Conception based on multiple Merge: A Head P has a property P(=EPP) which makes it the head of an XP that is an occurrence of K which (multiple)-Merges in its Spec. P(=EPP) is a selectional feature. H and K (in its first Merge position) Agree, while Pied-pipe determines the K that undergoes multiple Merge.



Re-interpretation: Move is not Agree/Pied-pipe/**Merge**, but Agree/Pied-pipe/**Mark**, where instead of multi-Merge, the Move operations 'Marks' the Probe H, and the HP that Immediately

Contains H is added as an ‘occurrence’ of K. The object K is Merged once, in its base position. I.e.,



‘Occurrence List’/Chain of  $K = \{ \langle K, HP \rangle, \langle K, RP \rangle, \langle K, TP \rangle \}$

Satisfaction of EPP is done by adding  $\langle K, HP \rangle$  and  $\langle K, RP \rangle$  to the ‘occurrence list’ under this alternative, and not by literally Moving (multi-Merging) K in the Spec of R and H. Rules of phonology spell-out K at its highest occurrence. ‘Mark’, like adding an EPP feature, is uninterpretable, but this is as it should be, since interpretable features do not drive computations.

In terms of the earlier discussion about complex features, what this means is that while in the earlier, conventional view, the EPP feature is valued by a complete category, in the proposed alternative, it is valued by something less than a full category, such as categorial features.

Notice that except for the supposition that K is Merged in the ‘base’ position, the reanalysis is similar to HPSG-type analysis of extraction where an extracted item shows up only once in the structure (at the top, rather than the base position) but enters into multiple ‘structure-sharing’ relations with intermediate constituents (heads and XPs) between the extracted site and the base position.