In our paper *Cyclic Computation* (1999), henceforth CC, considerable progress was made in simplifying syntactic computation, within the general framework of the Minimalist Program (MP). Comparison of derivations was eliminated from the theory, along with the need for numerations. Chomsky’s *Derivation by Phrase* (2000), henceforth DBP, proceeded along partially overlapping lines. In DBP, Chomsky introduced the important idea of valued versus unvalued features and shifted from the idea of feature checking to the idea of feature valuation. We adopt this idea and proceed by first reviewing the main ideas of the feature valuation analysis, then suggest that feature valuation is best seen as feature sharing. Rather than a valued feature giving a value to an unvalued counterpart, agreement will be realized as the sharing of a single feature between two syntactic terminals. With the mechanism of feature sharing in hand, we will then be in a position to revisit the analysis in CC and dramatically simplify both the analysis and the computation which it specifies.

This work grew out of the authors’ attempt to resolve certain inconsistencies in the DBP theory that were uncovered collectively in the course of close reading and discussion of DBP by a group consisting of Julie Legate, Charles Yang, and the authors. The theory of feature sharing developed here is quite independent (formally, not in spirit) from DBP, but readers who are interested in pursuing the DBP internal problems which led to the theory are encouraged to consult Frampton, et al (2000), which details some of the problems the group uncovered.

1. Feature Valuation

In CC, we concluded that Icelandic sentences like those in (1) showed that case on arguments is assigned when they enter the syntactic computation, and that it is checked or verified in the course of the derivation. The alternate (and traditional) view is that they enter the computation without a case feature, but that they are assigned one in the course of the derivation. It seemed to us, on the basis of Icelandic examples with past participle agreement like (1), that the latter idea was untenable. It required an appeal to ideas of case transmission via chains, ideas the MP was trying to eliminate.
(1)  

a. She(nom) was said(nom) to be popular.

b. They believe her(acc) to be said(acc) to be popular.

We reasoned that her in (1b) must already have a case feature at the point in the derivation where it establishes a relation with the participle, so that the participle could acquire its case at that point in the derivation. In short, her must already have accusative case before it enters into a relation with the higher accusative case assigner.

Chomsky’s introduction of valued and unvalued features in DBP allows a different approach to (1) and a way to reclaim the traditional idea that the structural case of arguments is established derivationally. He proposes that they enter the derivation with an unvalued case feature, which is valued in the course of the derivation. Case assignment works as shown in (2). A functional head (typically T or v∗, the preverb associated with transitivity) which bears “full φ-features” agrees with a nominal (X below). The agreement operation values the φ-features of the case assigner and values the Case feature of the nominal appropriate to the type of the case assigner. Nominative case assignment to a 3pl nominal is illustrated below.

\[
\begin{array}{cccc}
T & X & \xrightarrow{\text{Agree}} & T & X \\
\text{Per[\_]} & \text{Per[3]} & \to & \text{Per[3]} & \text{Per[3]} \\
\text{Num[\_]} & \text{Num[pl]} & \to & \text{Num[pl]} & \text{Num[pl]} \\
\text{Case[\_]} & \text{Case[\_]} & \to & \text{Case[nom]} & \text{Case[nom]} \\
\end{array}
\]

Chomsky’s idea for the participle agreement examples (1) is that the case of the participle is not determined by agreement with the nominal, but directly by agreement with the case assigning head. The idea is promising, but there are serious technical problems. Intervention problems, in particular, pose difficulties. 3

There is a subtle problem in getting feature valuation to do the work of feature erasure. Chomsky assumes that expletive there comes from the lexicon with an unvalued person feature. This feature must be valued in the course of the derivation. A typical derivation proceeds as in (3). First, an expletive is attracted. The attraction is on the basis of matching Person features, but neither is valued by the operation. Then, full agreement between T and women is carried out.
(3) There were women in the room.

1. T be women in the room
   Per[ ] Per[3]
   Num[ ] Num[pl]
   Case[ ]

2. there T be women in the room
   Per[ ] Per[3]
   Num[ ] Num[pl]
   Case[ ]

3. there T be women in the room
   Per[ ] Per[3]
   Num[pl] Num[pl]
   Case[nom]

How is the Person feature of the expletive valued at this point in the derivation? T is now fully satisfied, so that it is no longer probes for a matching Person feature.

Chomsky (p.c.) suggested that the relation between the Person feature of T and the Person feature of the expletive persists long enough for valuation of the Person feature of the expletive to be valued by valuation of the Person feature of T. Essentially, the computation must remember enough derivational history to carry this out. The theory of feature sharing, which we now proceed to develop, can be looked at as a theory of derivational memory. It accomplishes what is needed for (3), and much more.

2. Feature Sharing

We will pursue the intuition suggested by the remark at the end of the last section. How could the relation established by Agree be recorded? We need to consider carefully what features consist of. Features are often thought of as atomic symbols, something like numbers. From that point of view, one syntactic terminal node might be \{1, 3, 4\} and another \{1, 5\}. The 1 in one node does not have to be distinguished from the 1 in the other node. If Chomsky’s idea of valued features is correct, this cannot be the correct picture. Suppose we have two distinct terminal nodes:

(4) \{Num, \ldots \}, \{Num, \ldots \}

Can the two symbols “Num” in (4) denote the same thing? It could be the case, for example, that one Num feature has the value Plural and the other the value Singular. So it is certainly true that, in some cases, instances of features must be distinguished. Rather than (4), we need:
(5) \{\text{Num}_1, \ldots \}, \{\text{Num}_2, \ldots \}

We can now ask if it is ever the case that two different syntactic terminal nodes include the same feature. We propose that the answer is yes, and that this is the phenomenon of syntactic agreement. Consider (6) and suppose that Agree applies to the pair of nodes.

(6) \{\text{Num}_1, \text{Case}_2, \ldots \}, \{\text{Per}_3, \text{Num}_4, \text{Case}_5, \ldots \}

What is the result? In Chomsky’s DBP, Agree leaves the features of each terminal node, but values those unvalued features which match a valued feature. Instead, suppose that Agree induces feature sharing, so that matching features coalesce into a single shared feature, which is valued if either of the coalescing features is valued.\(^4\)

So (6) produces:

(7) \{\text{Num}_6, \text{Case}_7, \ldots \}, \{\text{Per}_3, \text{Num}_6, \text{Case}_7, \ldots \}

The value of \text{Num}_6 is the coalescence of the values of \text{Num}_1 and \text{Num}_4. The value of \text{Case}_7 is the coalescence of the values of \text{Case}_2 and \text{Case}_5. New indices were chosen, but index 6, for example, could just as well have been 1 or 4. The choice of index is not a substantive question, assuming that it is suitably distinguished. If the two coalescing features are both valued to start with, it is not clear that the result is coherent. But this will never arise, because Agree is driven by an unvalued feature.

A picture will make the idea clearer. Agree takes (8a) into (8b), assuming that none of the features indicated by the ellipsis marks match.

There are two things in particular which should be noted. First, the unvalued features coalesce in exactly the same way that a valued feature coalesces with an unvalued feature. Agreement is feature sharing, independent of value. Agree is blind to feature values. Second, the linear order on the feature tier is of no hierarchical or linear significance. The representation is symbolic, with the features under each node label organized into a set. The representation (8b) above could just as well have been depicted as (9). They are variant diagrams of identical syntactic representations.
Now consider agreement between the syntactic terminal nodes A and B above and a third terminal node T. Assume also that T is a nominative case assigner. Agreement with T will result in valuing a Case feature. (We omit the ellipsis marks which indicated features not entering into the agreement process. It should be assumed that such features are generally present.)

Application of Agree in (10) is straightforward. Some features are now shared between three syntactic terminals. Note that feature sharing has a major impact on how intervention is thought about. In (10), it makes no sense to think of the number feature of A intervening between the number feature of T and the number feature of B. This will be crucial in what follows.

With a mental picture of (8) and (10) now established, a more compact notation will be employed, which is intended to summarize both (9) and (10).

The diagram (11) is intended to convey considerably more information than simply the final representation. Asterisked features are those which were initially unvalued on the head they appear underneath. In the final representation, of course, all the Num features have coalesced into a single shared feature, which is valued. Sharing is indicated by the connectors joining features.

We can go further in making the notation compact if we observe that the Case feature need not be indicated at all. The fact that A and B are case-marked is determined by the fact that they share a $\phi$-feature with a case assigner. In our view of agreement, this is the essence of the case-marking relation. In the diagram, both the Per and Num features of B are linked to T and the Num feature of A is linked to B. The diagram is therefore:
Eliminating Case as a feature removes an awkward asymmetry from the theory. Unvalued Person and Number features are valued by agreement with a valued counterpart. Case was valued by a different mechanism. What we are now proposing is that Case is not a syntactic feature at all. Case morphology (structural) is determined by the featural linking which is the residue of agreement. The original insight that agreement is the basis of structural case assignment is due to George and Kornfilt (1981). We have taken this to its logical minimalist conclusion and removed structural case features from the syntax entirely. Case is assigned in the mapping from the terminal representation in the syntactic derivation to morphology. A TE-head which a nominal shares its $\phi$-feature with determines its structural case. Since the notion of active versus inactive element has played a major role in the theory, and rests upon the presence or absence of a valued case feature, this view of case assignment will have significant implications.

We can now return to the problem of valuing the Person feature of the expletive in (13), which we discussed in the last section.

(13) There were women in the room.

a. there $\text{T}$ be women in the room

\[
\begin{align*}
\text{Num}^* & \quad \text{Num} \\
\text{Per}^* & \quad \text{Per}^* \\
\text{Per}^* & \quad \text{Per}
\end{align*}
\]

b. there $\text{T}$ be women in the room

\[
\begin{align*}
\text{Num}^* & \quad \text{Num} \\
\text{Per}^* & \quad \text{Per}^* \\
\text{Per}^* & \quad \text{Per}
\end{align*}
\]

First, the expletive is attracted (out of the Numeration, for Chomsky). The (unvalued) Person feature of the expletive coalesces with the unvalued Person feature of the attracting head. Note that although the ultimate “aim” of attraction is valuing unvalued features, the operation is carried out only on the basis of feature matching and is blind to whether or not the attracted feature is valued. Second, the unvalued Person and Number features of $\text{T}$ attract the Person and Number features of the nominal and coalesce with them. All features are valued at this point. There are only two distinct features remaining. It does not make sense to ask which two features remain. The algorithm specifies feature coalescence, not feature elimination.

We now turn to the more complex example (14), involving past participle agreement. Selected steps in the derivation are given. In order to allow the linear representation to indicate some of the hierarchical structure, specifiers are indicated by a left delimiter $\langle$ on their left. The Tense of raising infinitives is denoted by...
T, and Chomsky’s proposal that T has no person feature and a number feature which is initially unvalued is adopted (for the purposes of this illustration only).

(14) He expected us(acc,pl) to be killed(acc,pl).

1. \[ \begin{array}{c}
\text{Prt} & \text{kill} & \text{us} \\
\text{Num}^* & \text{Num} \\
\text{Per} \\
\end{array} \]

2. \[ \begin{array}{c}
\text{T} & \text{be} & \text{Prt} & \text{kill} & \text{us} \\
\text{Num}^* & \text{Num} \\
\text{Per}^* & \text{Per} \\
\end{array} \]

3. \[ \begin{array}{c}
\langle \text{us} & \text{T} & \text{be} & \text{Prt} & \text{kill} & \text{t} \rangle \\
\text{Num} & \text{Num}^* \\
\text{Per} & \text{Per}^* \\
\end{array} \]

4. \[ \begin{array}{c}
\langle \text{he} & \text{v} & \text{trs} & \text{expect} & \langle \text{us} & \text{T} & \text{be} & \text{Prt} & \text{kill} & \text{t} \rangle \\
\text{Num} & \text{Num}^* & \text{Num}^* \\
\text{Per} & \text{Per}^* & \text{Per}^* \\
\end{array} \]

5. \[ \begin{array}{c}
\langle \text{he} & \text{v} & \text{trs} & \text{expect} & \langle \text{us} & \text{T} & \text{be} & \text{Prt} & \text{kill} & \text{t} \rangle \\
\text{Num} & \text{Num}^* & \text{Num}^* \\
\text{Per} & \text{Per}^* & \text{Per}^* \\
\end{array} \]

The derivation finishes with merging the matrix T and further agreement.

On line 3 above, movement was carried out. We can see, in fact, that this plays no role in the derivation. We could just as well have left the nominal in place, as in (15). Linking (in the sense of sharing φ-features), which persists, identifies potential landing sites. Placement in surface form can be determined post-syntactically, on the basis of this linking.

(15) \[ \begin{array}{c}
\langle \text{he} & \text{v} & \text{trs} & \text{expect} & \langle \text{T} & \text{be} & \text{Prt} & \text{kill} & \text{us} \rangle \\
\text{Num} & \text{Num}^* & \text{Num}^* & \text{Num} \\
\text{Per} & \text{Per}^* & \text{Per}^* & \text{Per} \\
\end{array} \]

3. Refining the Cyclic Computation Account

The previous sections motivated the idea of feature sharing as a way to surmount a technical problem with DBP’s proposal for replacing the idea of feature checking and deletion with feature valuation. Our intention, however, is not to pursue the DBP framework of assumptions. If we go back to the framework constructed in CC, armed with Chomsky’s idea of feature values and our idea that agreement is feature sharing, a very elegant system emerges.
Much of the complexity of CC came from the insistence that the theory adequately account for quirky case phenomena. The DBP framework drives the formation of NP-chains (we use the term descriptively) by $\phi$-feature agreement, with the EPP determining which instances of $\phi$-agreement trigger movement. It seems to us that the formation of NP-chains headed by a quirky case element shares so many properties with the formation of NP-chains headed by a structurally case marked element that this view cannot be correct. The key diagnostic for quirky case-marked nominals is precisely the absence of agreement.

We will suppose that two features are relevant to the case/agreement system. The first is just $\phi$, which comes in valued and unvalued varieties. Possible values are the usual person, number, and gender (or noun class) values. The second, we will call $\delta$. It also comes in valued and unvalued varieties. It will be $\delta$-features, rather than $\phi$-features, that play the major role in specifying what movement occurs. At this point, the values which $\delta$ takes on remain obscure. Before we speculate on this point, however, it should be emphasized that the syntax is concerned only with a certain calculus of $\delta$ and $\phi$-features and the links they establish between heads. Actual feature values do not come into play until the syntactic structure is mapped to a morphological structure.

It will be clear from the role that $\delta$-features play in what follows that a natural candidate for $\delta$-feature values is some kind of definiteness marker. The verbal morphology of a few languages has morphological marking which indicates a definiteness/indefiniteness contrast for the object. Hungarian is one well-known example. Lieber (1992) discusses Kálmán’s (1965) analysis of the Uralic language Vogul, which has such morphology. We speculate that this morphology reflects the values of the $\delta$-feature. We will later build an analysis of there-type expletive constructions on the basis of the assumption that certain nominals (i.e. those that appear as the associate of an expletive) can enter the syntax with an unvalued $\delta$-feature. One considerable advantage of relating $\delta$-features to definiteness is that it offers a possible avenue to explain our otherwise stipulative assumption that only indefinites can have an unvalued $\delta$-feature.

We will use $\delta^*$ and $\phi^*$ as compact notations for unvalued $\delta$ and $\phi$-features. This assumption will be modified as we proceed, but suppose for a start that nominals come with valued $\delta$ and $\phi$-features and the structural case assigners come with $\delta^*$ and $\phi^*$, unvalued counterparts of $\delta$ and $\phi$. “Nominal” here is a somewhat imprecise designation for the phrases that enter the case/agreement system, arguments and nominal expletives (it in English). Following CC, we suppose that the derivation is built up by alternate applications of Select and Attract. Heads are introduced into the syntax by applications of Select, so that they never appear except with whatever phrases they select. After Select applies and introduces a head into the syntax, the unvalued features of that head are satisfied by applications of Attract. We call the head which is the locus for applications of Attract the pivot. All unvalued features of the pivot must be satisfied before the derivation can proceed to a new application of Select. The introduction of a pivot, combined with its Select and
Attract operations, constitutes a cycle.

As in CC, Attract finds targets of attraction by first conducting a top-down search from the pivot, seeking a feature which matches the unvalued feature of the pivot. If a feature matching the unvalued feature of the pivot is found, the top-down search stops and the unvalued pivot feature and the feature which was found by the search are coalesced and a shared feature results. If no matching feature is found, or if one is found and coalescence has taken place, but the unvalued feature of the pivot remains unvalued because the feature it coalesced with is itself unvalued, then an “external search” is allowed in order to value the unvalued feature of the pivot. This will be discussed more fully when we consider expletives, which enter the derivation when a top-down search from the pivot fails and an external search is undertaken. In no case does the top-down search from the pivot proceed after a matching feature has been found, even if agreement with that feature is impossible for some reason, or if agreement does not succeed in valuing the unvalued feature of the pivot. This is the view of “intervention effects” which we proposed in CC.

To illustrate this view of the agreement process, consider first (16a). The derivation is represented by the diagram (16b). It is actually a (partial) representation of the derivation, not simply the final representation. There are no unvalued features in the final representation. The connections between the features and their initial status with respect to valuation indicates some of the derivational history. In the final representation, the connected features have coalesced into a single valued feature, shared by the relevant heads.

(16) a. Sally likes Mary.
   b. T ⟨ Sally vtrs likes Mary
      δ* — δ δ* — δ
      φ* — φ φ* — φ

There are several comments to make about (16). Movement has not been indicated, just agreement. Movement is determined by properties of the heads, on the basis of the feature sharing which agreement establishes. It is a striking advantage of realizing agreement as feature sharing that it succeeds both in simplifying the representation, by eliminating redundant features, and incorporating a partial history of the derivation into the representation. For the purposes of this paper, we leave open the question of whether movement takes place cyclically, or by phase, or not until the final representation is spelled out. Agreement, of course, is carried cyclically as the representation is built up by merger.

There is no direct indication of structural case, which is determined post-syntactically on the basis of feature sharing. We assume that the requirement that nominals are case-marked is a morphological requirement, not a syntactic requirement. Indeed, Icelandic assigns (default) nominative case post-syntactically to nominals which are neither inherently case-marked nor in an appropriate feature sharing relation with an appropriate head. We shall see shortly that alongside
of “case marking” in the standard sense, which has the usual morphological reflex, there is a strict syntactic requirement that nominals share a \( \delta \)-feature with a suitable head. This generally has no morphological reflex, but is crucial in the case/agreement system.

The next example illustrates both ECM and \( T_r \) (raising \( T \)). We assume that \( T_r \) has both an unvalued \( \delta \)-feature and an unvalued \( \phi \)-feature. \( T_r \) is distinguished from tensed \( T \) other than by its \( \delta \) and \( \phi \)-features.

\[
\begin{array}{cccc}
\delta^* & \delta & \delta^* & \delta^* \\
\phi^* & \phi & \phi^* & \phi^*
\end{array}
\]

Consider this derivation at the point just after Select has introduced \( v_{\text{trs}} \) into the computation, with its selected VP complement and selected subject.

\[
\begin{array}{cccc}
\delta & \delta^* & \delta^* & \delta \\
\phi & \phi^* & \phi^* & \phi
\end{array}
\]

The \( \delta \) and \( \phi \)-features of \( T_r \) and the nominal have coalesced, so that when \( v_{\text{trs}} \) looks down the tree for a \( \delta \) or \( \phi \)-feature to attract, it sees only these features, which are valued. It does not make sense to ask if they are features of \( T_r \) or features of the nominal. They are both.

If we consider an Icelandic counterpart, with the additional complexity of participle agreement, we see that feature sharing makes case assignment to the participle entirely natural. The \( \phi \)-feature which is shared by the nominal, the participle and \( T_r \) is attracted to \( v_{\text{trs}} \).

\[
\begin{array}{cccc}
\delta^* & \delta & \delta^* & \delta \\
\phi^* & \phi & \phi^* & \phi
\end{array}
\]

This analysis is a big improvement over CC, where examples of this kind forced us to conclude that case features must be present from the start, and checked by a higher case-checking head. Chomsky’s idea of feature valuation, coupled with feature sharing, allows a much simpler account. The account of this in DBP, however, runs into problems caused by intervention effects. See Frampton, et al (2000) for a detailed discussion of the DBP account.
3.1. Chains, the Chain Condition, and Locality

Feature sharing, triggered by Attract, links heads via shared $\delta$ and $\phi$-features. It builds structures like (20), which is one of the linked sets of heads created by feature sharing in (19).

\[ \begin{array}{cccc}
v_{\text{trs}} & T_r & \text{Prt} & \text{people} \\
\phi^* & \phi^* & \phi^* & \phi
\end{array} \]

We call the set of four linked heads in (20) a $\phi$-chain. An $F$-chain is the maximal set of heads containing a particular feature ($F$). An element of a chain is called its head if all the members of the chain are in its maximal projection. We will call $\delta$ and $\phi$-features, either valued or unvalued, nominal agreement features. This will usually be abbreviated to agreement features if the context makes it clear that nominal agreement is the issue. Note carefully that $\delta$-agreement is given equal status with $\phi$-agreement.

Details will follow, but in order to orient the reader, we begin with a paragraph which previews the role of “case assignment” in the theory we are proposing. There is a class of heads, which we call TE-heads, which plays a crucial role in making arguments visible to the interpretive system. They play a role with respect to both $\delta$-chains and $\phi$-chains. Every $\delta$-chain must be headed by a TE-head. This is a strict syntactic requirement. Some nominals, inherently case-marked, enter the derivation associated with a morphological case feature. Morphological case is assigned to other nominals in the post-syntactic mapping to morphology if they are in a $\phi$-chain with a TE-head, the particular morphological case being determined by the type of that TE-head. This is structural case assignment. There is no syntactic requirement that a nominal be in a $\phi$-chain with a TE-head. Other nominals, neither inherently case-marked nor structurally case-marked, can be assigned default case post-syntactically.

It would be highly desirable to be able to characterize the class of TE-heads independently, then go on to describe the role they play in the case/agreement system. Unfortunately, this is beyond our abilities, and we are forced to simply list the TE-heads (finite T, $v_{\text{trs}}$, ...). The prefix “TE” indicates that we have some hope that their characterization can be connected with the temporal/event structure.

Based on the observation that participles are inflected for number, but not for person, Chomsky identifies the case-assigning property with bearing both person and number features, which he calls “full $\phi$-features.” In our view, the behavior of quirky case nominals in Icelandic, which we discuss in a later section, makes this untenable. López (2000) gives another argument against the view that having full $\phi$-features is directly correlated with the ability to assign case. He observes that in some Bantu languages, agreement with the subject surfaces morphologically on both tense and an aspectual head which precedes the verb. Significantly, agreement is for both person and number, and noun class (NC). He gives (21), from Kiswahili.\(^7\)
The crucial point for the present discussion is that the perfective aspectual head has full $\phi$-features (both person and number), in Chomsky’s terminology, but is not a case assigner.

We therefore assume that the correlation between full $\phi$-features and the ability to assign case is not direct. This is important to what follows, since we assume that $T$ and $T_r$ both have (initially unvalued) both $\delta$ and $\phi$-features, where here (and henceforth) $T$ denotes finite $T$ and $T_r$ denotes raising $T$.

The fact that $T$ is a TE-head and $T_r$ is not must be due to the former’s finiteness, not simply the fact that it bears both $\delta$ and $\phi$-features. Consistent with their identical syntactic feature structure, it will become clear in the next section that in the theory we develop, $T$ and $T_r$ have identical syntactic behaviors. They are distinguished not by their syntactic behavior, but by their role in determining the interpretability of the chains they appear in.

### 3.1.1 The Chain Condition, Locality

Various devices have been proposed to constrain the locality of agreement in the development of the Minimalist Program: inactivity, phases, the Minimum Link Condition, and many variations on these themes. All these devices are aimed at ensuring that chains of only a particular kind are built. The concerns for chain well-formedness go back to Chomsky (1982) and the early development of GB-theory.

Consider, for example, (23).

(23) *Mary was believed t likes Sally.

At one point in the development of GB theory, this was ruled out by supposing that A-chains could contain only one “case position.” But that requirement failed to explain (24).

(24) *It seems Mary to be believed t likes Sally.

The additional requirement that the “head of an A-chain” must be in a case position was used to rule out (24). Combining these two requirements led to the condition:

(25) The head, and only the head, of an A-chain is in a case position.

We will shortly propose that a version of (25) plays a key role in minimalist syntax.
Rather than rule out examples like (23) and (24) by conditions on chains, the thrust of the MP was to try to build the architecture of the theory in such a way that it was forced to produce chains of the kind that are found, and only this kind. The feature checking mechanism, in its original form, made it automatic that movement stopped when case was assigned. With the move to a theory of agreement based on attraction, the burden fell to the locality conditions mentioned above, a notion of “inactivity,” and the idea that certain phrases were frozen by spellout which operated during the course of the derivation (Chomsky’s notion of “phases”).

The approach which we will take is a partial return to the GB idea that conditions on chain well-formedness play a direct role in the syntax. Well-formed chains are not simply an epiphenomenon, but the direct expression of core output requirements. Well-formed chains are part of the ontology of objects recognized by post-syntactic processes. Ill-formed chains are not. But there is a major difference between the role that conditions on chains play in the syntactic architecture we propose and the role they played in GB-theory, particular in its Move-α variant. In GB-theory, chain conditions were imposed as derivational constraints, in the sense that operations which introduced violations of the conditions were blocked. Rather than a derivational constraint, we will assume that (26) is an output condition (i.e. filter). Its effect is to mandate a one-to-one correspondence between δ-features and TE-heads as an interpretability condition at the output interface.

(26) Chain Condition: Every δ-chain is headed by a TE-head and every TE-head is the head of a δ-chain.

A TE-head with both δ and φ-features will turn out to head both a δ-chain and a φ-chain, which may be distinct.

Since we assume a computational process in which each step in a derivation is locally determined, with no comparison of derivations, failure at the interface/output has no effect on the step-by-step derivational computation. A derivation which violates an output condition is simply an ill-formed derivation. An output condition in a derivational theory with no lookahead and no backtracking is simply a filter on outputs.

In addition to the locality which is implicit in the mechanics of the Attract algorithm, we also assume (27). It is a reincarnation of the “Tensed-S Condition” of early GB theory.

(27) C-Complement Opacity: After a C-cycle is completed, the complement of C is inaccessible to Attract.

We give (27) as a constraint on derivations, in order to highlight it. It could be directly and naturally incorporated into the Attract algorithm, which specifies how the search for a feature which matches an unvalued feature of the pivot takes place. Clearly, (27) is related to Chomsky’s notion of a “phase”. For reasons of brevity,
we do not speculate here on these connections and do not examine the question of whether or not (27) should be derived from an assumption of cyclic or phasal spellout of some kind.

The locality conditions are weak, only (27) and the locality that derives from the mechanics of the Attract algorithm. Much of the burden of ensuring that well-formed outputs are achieved therefore falls on the lexical choices that are made in the course of building the derivation. A derivation is built by a sequence of cycles. In each cycle, a head (called the pivot of the cycle) is introduced into the derivation and its selectional requirements are satisfied by merger. The unvalued features of the pivot then probe and coalesce with matching features. If unvalued features of the pivot remain, the derivation terminates unsuccessfully. If no unvalued features of the pivot remain, either a new pivot is introduced or the derivation is terminated successfully. We can symbolize the derivational process as:

(28) Select, (Attract)*, Select, (Attract)*, Select, (Attract)*, ...

The applications of Attract in (28) are automatic. What determines whether or not a successful derivation is constructed are the applications of Select, in particular, the lexical choices that are made at each step. In order to illustrate this, consider the derivation of (29a), at the point in the derivation given in (29b).

(29) a. It seems that Sally likes Mary.
    b. \[ T \langle Sally \ v_{trs} \ like \ Mary \]
      \[
      \begin{array}{llll}
      \delta^* & \delta & \delta^* & \delta \\
      \phi^* & \phi & \phi^* & \phi
      \end{array}
      \]

It is crucial that at this point in the derivation that the TP (29b) be embedded as a C-complement. If not, continuing the derivation of (29a) leads to (30a) and then (30b).

(30) a. T seem T \langle Sally \ v_{trs} \ like \ Mary \]
    \[
    \begin{array}{llllll}
    \delta^* & \delta^* & \delta & \delta^* & \delta \\
    \phi^* & \phi & \phi^* & \phi & \phi^* & \phi
    \end{array}
    \]

b. T seem T \langle Sally \ v_{trs} \ like \ Mary \]
    \[
    \begin{array}{llllllllllllllll}
    \delta^* & \delta^* & \delta & \delta^* & \delta & \delta^* & \delta \\
    \phi^* & \phi^* & \phi & \phi^* & \phi & \phi^* & \phi
    \end{array}
    \]

Instead of expletive insertion, the matrix T agrees into the embedded clause and a Chain Condition violation occurs: the lower T, a TE-head, does not head a \( \delta \)-chain.

If (29b) is embedded as a C-complement, the derivation continues to (31a) and (31b) results. C-Complement Opacity provides the needed isolation between the embedded clause and the matrix TE-head.
We will return later to discuss expletive insertion in more detail. Here we simply note the fact that the C-complement in (31a) prevents the matrix T from valuing its agreement features (i.e. $\delta^*$ and $\phi^*$) by agreement with or into its complement, so that the only way that they can be valued is by attracting an expletive from outside its complement, shown in (31b).

Note that seem does not always select a CP complement. The clausal complement of seem in (32) is not a CP.

(31) a. $\text{T} \hspace{1em} \text{seem} \hspace{1em} \text{C} \hspace{1em} \text{T(} \hspace{1em} \text{Sally} \hspace{1em} \text{v}_{\text{tr}} \hspace{1em} \text{like} \hspace{1em} \text{Mary}$

$\hspace{1em} \delta^* \hspace{1em} \delta^* \rightarrow \delta \hspace{1em} \delta^* \hspace{1em} \delta$

$\hspace{1em} \phi^* \hspace{1em} \phi^* \rightarrow \phi \hspace{1em} \phi^* \hspace{1em} \phi$

b. $\text{it} \hspace{1em} \text{T} \hspace{1em} \text{seem} \hspace{1em} \text{C} \hspace{1em} \text{T(} \hspace{1em} \text{Sally} \hspace{1em} \text{v}_{\text{tr}} \hspace{1em} \text{like} \hspace{1em} \text{Mary}$

$\hspace{1em} \delta - \delta^* \hspace{1em} \delta^* \rightarrow \delta \hspace{1em} \delta^* \hspace{1em} \delta$

$\hspace{1em} \phi - \phi^* \hspace{1em} \phi^* \rightarrow \phi \hspace{1em} \phi^* \hspace{1em} \phi$

For the raising verb seem, the choice of CP or TP is straightforward. If the embedded clause is headed by $T_r$, it cannot be a C-complement. If it is headed by a case-assigning $T$, then it must be a C-complement. The crucial point for selection by seem is that it does not select T.

3.1.2 Burzio’s Generalization

The demonstration above that a certain clausal complement must be a CP complement relied on the fact that the absence of a C would lead an ill-formed chain in the output. The complementizer was needed to protect a $\delta$-feature inside the embedded clause from fatal agreement. The empirical facts which are described by Burzio’s Generalization have precisely the same source. A TE-preverb requires a subject in order to protect its $\delta$-feature from fatal agreement with clausal Tense (either $T$ or $T_r$).

In order to see this, consider a typical example of the kind of preverb which Burzio’s Generalization (under a modern interpretation) identifies as absent from the syntactic vocabulary. Suppose, for example, that $v'$ has $\delta$ and $\phi$-features (i.e. assigns accusative case), but does not select a subject. A derivation would start out along the lines of (33a), then proceed to (33b) after clausal Tense is merged. The continuation is forced and yields the hopeless (33c). The continuation is forced because, as we will examine more closely in the next section, expletive insertion is not called on unless unvalued features of the pivot (the T which has just been introduced into the derivation in this case) cannot be valued by agreement into (or with) the complement of the pivot. It is hopeless because the Chain Condition requires $v'$ to be the head of its $\delta$-chain.
It is clear that unless the preverb has a subject with a δ-feature, the derivation will be ill-formed. The preverb will not end up as the head of its δ-chain, but the Chain Condition demands that a TE-head must be the head of a δ-chain. A subject δ-feature is needed to shield the δ-feature of the preverb from attraction by T. This is just the empirical fact which Burzio’s Generalization describes. 8

Compare (33) with the case where the preverb has a subject with a δ-feature.

(34) Sally(nom) grows tomatoes.

T ⟨ Sally, v_1, grow, tomatoes, δ^* −−−−−−→ δ, φ^* −−−→ φ ⟩

TE-preverbs necessarily have a δ-feature, since the Chain Condition demands that they head δ-chains. We conclude that if there is a single TE-preverb, then it must have a subject with a δ-feature. If there are multiple preverbs, the condition that must be met in order to avoid an ill-formed chain in the output cannot be stated in terms of properties of individual preverbs. We cannot claim, for example, that a preverb which has a δ-feature necessarily has a subject with a δ-feature. It could be, for example, that one preverb introduces the external argument and the second assigns accusative case, as in (35).

(35) ⟨ Sub, v_1, v_2, V, Obj, δ, δ^* −−−−−−→ δ, φ, φ^* −−−→ φ ⟩

We can see from (35) that if multiple preverbs are possible, then it is some property of the complex of preverbs that must hold in order to ensure that a δ-feature of a TE-preverb is shielded from attraction by clausal Tense. Burzio’s Generalization is simply a statement of the fact that the δ-feature of a TE-preverb must be suitably shielded from attraction.

Burzio’s Generalization has been a major embarrassment to theories of case and A-chain formation. It has simply been an unexplained fact about the syntactic
vocabulary (the morphemes which the lexicon makes available to the syntax). The fact that we can derive Burzio’s Generalization (to the extent that it is stated precisely and is valid) from the Chain Condition is a major argument in favor of our idea that the Chain Condition is an output filter, with no role as a derivational constraint. As we proceed, we will see that other crucial facts about δ-chains also follow from the Chain Condition, so the Chain Condition is not simply an alternate way to stipulate Burzio’s Generalization. It appears to be a real explanation for the empirical facts which Burzio’s Generalization describes.

3.1.3 Local Avoidance of Output Violations

One of our aims in developing the kind of the computational architecture that we did in CC was to make the computation of derivations straightforward. We eliminated comparison of derivations because of the computational complexity this introduces. We avoided numerations for the same reason. For a given numeration, no derivational complexity is added. But most numerations will not lead to a successful derivation. How are “good numerations” chosen? Constructing derivations relative to a given numeration is a way to transfer complexity out of the derivational computation onto the problem of constructing numerations which allow successful derivations. Our proposal here that the Chain Condition, a filter on outputs, plays a central role in syntax threatens the idea of ‘derivational simplicity.” In principle, output filters could create a situation in which derivations themselves are simple and straightforward, given lexical choices, but the lexical choices (of functional material, in particular) which must be made in order to build a successful derivation cannot be made on the basis of local information. Essentially, this would reintroduce serious computational complexity into the system, somewhat like that introduced by comparison of derivations.

Fortunately, as far as we have been able to determine, the derivational choices that must be made can be done locally. It would be most desirable if the selectional properties of the syntactic vocabulary are sufficient to ensure well-formed chains. This may be close to being to true. We pursue the question in some detail in Frampton and Gutmann (2000). Here, we can give a few examples to indicate what we have in mind.

Suppose a language allows only a single preverb. It follows for reasons we discussed above that the case features of a preverb must be restricted to a subset of the case features of the subject of that preverb. Otherwise, the preverb will be useless, producing only ill-formed chains. Assuming that useless items are not acquired, it will therefore be a fact about the syntactic vocabulary of this language, which includes the selectional properties of the morphemes which are made available to the syntax, that preverbs will automatically fall into accord with Burzio’s Generalization.

Suppose we consider now a language in which there is a restriction to at most two preverbs. Suppose there is a preverb \( v_{trs} \), which assigns accusative case via unvalued \( \delta \) and \( \phi \)-features, but which does not select a subject. It follows, again
from the considerations already discussed, that this preverb can only be selected (if
a successful derivation is desired) by a second preverb which has no $\delta$ or $\phi$-feature,
but which selects a subject with $\delta$ and $\phi$-features. Again, the selectional properties
of the lexical items will ensure accord with our reinterpretation of Burzio’s Gen-
eralization. If a language allows multiple preverbs, the conditions become more
complex. But the important point is that it is a question of local bookkeeping, not
a crash (i.e. a result which is discarded by the output filter) at a remote interface.

3.2. Expletives

We started off by assuming that nominals must have valued $\delta$ and $\phi$-features when
they enter the derivation. Now suppose that a nominal can have an unvalued
$\delta$-feature. Consider the following,

(36) T be some people in the room

$\delta^* \quad \delta^* \quad \phi^* \quad \phi$

The $\delta$-feature of the pivot remains unvalued, even after agreement with the nominal
$\delta$-feature. If no further syntactic resources are available, the derivation fails.

Languages have various devices to value case features of the pivot if agreement
into the complement does not accomplish this. If the pivot cannot value its unvalued
features by agreement into its complement, it can look outside the complement. In
particular, it can carry out agreement with an external element whose case features
are valued and match the unvalued features of the pivot. If the pivot attracts an
external element, that element must be an expletive, because no theta role will be
assigned. A variety of external elements occur. A nominal expletive, *it* in English,
with valued $\delta$ and $\phi$-features, will simultaneously value both an unvalued $\delta$-feature
and an unvalued $\phi$-feature. A so-called “pure expletive”, *there* in English, with
only a valued $\delta$-feature, will value an unvalued $\delta$-feature. In Icelandic, it is possible
for the $\delta$-feature of the pivot to be valued and the $\phi$-feature unvalued. Icelandic
has default expletive $\phi$-agreement which values the $\phi$-feature.

Returning to English, we first consider expletive *there*, with a valued $\delta$-feature
and no $\phi$-feature. In order to keep the representations relatively simple, we will put
the expletive in the specifier position. But this is not crucial. Linear order will be
dictated by the Spellout computation on the basis of shared features. Continuing
the derivation of (36), we have:

(37) There are some people in the room.

$\{ \text{there} \quad T \text{ be some people in the room}$

$\delta \quad \delta^* \quad \phi^* \quad \phi$

The raising case is straightforward.
(38) There seem to be some people in the room.

a. \[ T_r \text{ be some people in the room} \]
   \[ \delta^* \quad \delta^* \]
   \[ \phi^* \quad \phi \]

b. \( \langle \text{there} \ T_r \text{ be some people in the room} \)
   \[ \delta \quad \delta^* \quad \delta^* \quad \delta^* \]
   \[ \phi^* \quad \phi^* \quad \phi \]

c. \( T \text{ seem} \langle \text{there} \ T_r \text{ be some people in the room} \)
   \[ \delta^* \quad \delta \quad \delta^* \quad \delta^* \quad \delta^* \]
   \[ \phi^* \quad \phi^* \quad \phi \]

Agreement with an expletive is forced in the \( T_r \) cycle in order to value the \( \delta \)-feature of \( T_r \). Recall that we assume that all pivot features must be valued cyclically, so the first pivot whose \( \delta \)-feature cannot be satisfied internally attracts an expletive.

If \textit{people} has an unvalued \( \delta \)-feature, (38) results. If its \( \delta \)-feature is valued, we get (39). There are no other options.

(39) Some people seem to be in the room.

\[ T \text{ seem} \ T_r \text{ be some people in the room} \]
\[ \delta^* \quad \delta^* \quad \delta \]
\[ \phi^* \quad \phi^* \quad \phi \]

No expletive is attracted in (39) because the unvalued \( \delta \)-feature can be satisfied internally, without going outside the structure to agree with an external expletive.

The two possibilities are (38) and (39), depending on whether the \( \delta \)-feature of the nominal is valued. There is no possibility of (40a) or (40b). Avoiding these has proved challenging, and is the main motivation for Chomsky’s introduction of the device of “numerations”.

(40) a. *There seem there to be some people in the room.

b. *There seem some people to be in the room.

The interaction of expletives and participle agreement in Icelandic proceeds without complication, illustrated in (41) and (42).

(41) There were(pl) seen(pl,nom) many people(pl,nom).

\( \langle \text{there} \ T \text{ be Prt see} \text{ many people(pl,nom)} \)
\[ \delta \quad \delta^* \quad \delta^* \quad \delta^* \]
\[ \phi^* \quad \phi^* \quad \phi \]
He expects there to have been seen(acc,pl) many people(acc,pl).

\[ \ldots \text{v}_{\text{trs}} \text{ expect} \langle \text{there} \quad \text{T}_{r} \quad \text{to-have been} \quad \text{Prt} \quad \text{see} \quad \text{many people} \rangle \]

\[ \delta^{*} \quad \delta \quad \delta^{*} \quad \delta^{*} \quad \delta^{*} \quad \delta^{*} \]

\[ \phi^{*} \quad \phi^{*} \quad \phi^{*} \quad \phi^{*} \quad \phi^{*} \quad \phi \]

Unvalued $\delta$-features on nominal pivots: Although we suppose that a nominal can occur with an unvalued $\delta$-feature, we suppose that $\delta^{*}$-features on nominals can never be attractors. Although we stipulate this here, it may follow from the organization of the cyclic computation itself. The problems that arise if $\delta^{*}$-features on nominals can attract is clear from (43), the underlying structure of a transitive expletive construction. If $\delta^{*}$ on the subject attracts, the agreement shown in (43b) will result. It is doomed to failure since merger with T will lead to a $\delta$-chain containing both T and $v_{\text{trs}}$.

(43) \[ \langle \text{Sub} \quad v_{\text{trs}} \quad V \quad \text{Obj} \rightarrow \langle \text{Sub} \quad v_{\text{trs}} \quad V \quad \text{Obj} \rangle \]

\[ \delta^{*} \quad \delta^{*} \quad \delta \quad \delta^{*} \quad \delta^{*} \quad \delta \quad \delta^{*} \quad \delta^{*} \quad \delta \]

\[ \phi \quad \phi^{*} \quad \phi \quad \phi^{*} \quad \phi \]

If we recall how the computation is organized cyclically, then this problem evaporates. Heads select their arguments, then value their unvalued features. In (43a), $v_{\text{trs}}$ selects its subject and V-complement and initiates a cycle of attraction with its unvalued features as probes. The subject does not select, so its features never get the opportunity to probe.

More delicate are phrases like $[\text{pictures of Bill}]$. If pictures selects, we might expect:

(44) \[ \text{pictures of Bill} \rightarrow \text{pictures of Bill} \]

\[ \delta^{*} \quad \delta^{*} \rightarrow \delta \quad \delta^{*} \rightarrow \delta^{*} \rightarrow \delta \]

An account must be given for why (44) does not occur. We will assume that only functional heads probe. Non-functional heads are “non-cyclic,” in the sense that they can select, but the derivation proceeds without a cycle of attraction. They therefore evade the requirement that the cycle of attraction must value all of the unvalued features of the pivot.

In a system in which attract operations are never optional, and in which unvalued features drive attraction, the device of permitting unvalued $\delta$-features on nominals is a minimal solution to the design problem of allowing a nominal to share its $\delta$-feature with T (or $T_{r}$) without destroying the ability of that T to attract an expletive. It is a minimal solution in the sense that unvalued $\delta$-features and the entire agreement mechanism are already in place.
3.2.1 Nominal Expletives

The examples above exploited attraction of a pure expletive (*there* in English) to value the $\delta$-feature of a pivot whose $\phi$-feature was valued. If the pivot has an unvalued $\delta$ and $\phi$-features, a nominal expletive is used in English (and the Germanic languages generally) to simultaneously value both unvalued features of the pivot.

A raising example is given in (45a).

(45) It seems to be likely that the earth is flat.
   a. $\mathrm{T_r}$ be believed [that the earth is flat]
      \[\delta^* \phi^*\]
   b. $\langle \; \mathrm{it} \; \mathrm{T_r} \; \mathrm{be} \; \mathrm{likely} \; \rangle$ [that the earth is flat]
      \[\delta - \delta^* \phi - \phi^*\]

3.2.2 The Absence of Object Expletives

Why, for example, is (46) not generated?

(46) Sally does there like someone.

The issue is the possible continuations of (47), where the object has an unvalued $\delta$-feature.

(47) $\langle \; \mathrm{Sally} \; \mathrm{v_{trs}} \; \mathrm{like} \; \mathrm{someone} \; \rangle$
    \[\delta \delta^* \phi \phi^*\]

The pivot cannot value its $\delta$-feature by agreement into its complement. Earlier, we proposed that this allows an external search as a last resort. If a search for a target of agreement external to the complement of the pivot is undertaken in (47), the specifier of the pivot is the first element encountered. The pivot centered search will encounter the $\delta$-feature of Sally and (48) will result.

(48) $\langle \; \mathrm{Sally} \; \mathrm{v_{trs}} \; \mathrm{like} \; \mathrm{someone} \; \rangle$
    \[\delta \delta^* \phi \phi^*\]

This structure has no viable continuation, for reasons we discussed in relation to our earlier consideration of Burzio’s Generalization. The subject can no longer shield the case features of the preverb from a higher $\mathrm{T}$ or $\mathrm{T_r}$.
3.3. Quirky Case in Icelandic

The TE-head which a nominal shares its $\phi$-feature with determines the structural case of the nominal. Suppose that inherently case marked arguments in Icelandic have a $\phi$-feature, but that it cannot be shared, presumably because the case of the argument is already determined. This is essentially a constraint against multiple case-assignment, familiar in many syntactic frameworks. But suppose the $\phi$-feature is visible to the syntax, in spite of the fact that it cannot be shared. Visibility is relevant because we assume that $\phi^*$-probing searches for the first visible $\phi$-feature, and no further (i.e. no deeper). Inherently case-marked nominals will therefore block $\phi^*$-probing into their complements without themselves being the targets of $\phi^*$-attraction.

If we look at the simple quirky case example (49a) in Icelandic, it is clear that further assumptions must be made. If nothing else is assumed, an attempted derivation fails early, as shown in (49b). We denote the nonsharable $\phi$-feature of an inherently case-marked nominal by $\phi_0$. The participial Prt in (49) cannot value its $\phi^*$-feature.

(49) a. Us(dat) was(sg) helped.
   b. Prt help us(dat) $\delta$
      $\phi^*$ $\phi_0$

Icelandic, however, has a mechanism for rescuing (49b). As we saw with pure and nominal expletive insertion in the last section, the grammar provides some mechanisms for valuing unvalued features on the pivot, in case they have not been valued cyclically. Icelandic has a last resort mechanism for valuing a $\phi^*$-feature. A pure feature, $\phi$ with the default value 3s, is attracted to the pivot. In effect, there is expletive $\phi$-agreement. A $\phi$-feature which gets default valuation is non-sharable, like the $\phi$-feature of an inherently case-marked nominal. The representation (49) is transformed to (50a), and continues to (50b).

(50) a. Prt help us(dat) $\delta$
    $\phi^*$-$\phi_{dft}$ $\phi_0$

    b. T be Prt help us(dat) $\delta$
    $\phi^*$-$\phi_{dft}$ $\phi_{dft}$-$\phi_{dft}$ $\phi_0$

ECM constructions demonstrate why it must be the case that $\phi^*$-$\phi_{dft}$ is not sharable. Note that (51a) results, not (51b), which would result if defaulted $\phi^*$-features were sharable. Prt gets default nominative case in (51a).
(51) a. You believed us(dat) to have been helped(nom,dflt).

\[
\begin{array}{cccc}
  \delta^* & \delta^* & \delta \\
  \phi^* \phi_{\text{dft}} & \phi^* \phi_{\text{dft}} & \phi^* \phi_{\text{dft}} & \phi_0
\end{array}
\]

b. *You believed us(dat) to have been helped(acc,dflt).

\[
\begin{array}{cccc}
  \delta^* & \delta^* & \delta \\
  \phi^* & \phi^* & \phi^* \phi_{\text{dft}} & \phi_0
\end{array}
\]

Expletive constructions are also possible:

(52) There was(dflt) helped many people(3pl,dat).

\[
\begin{array}{cccc}
  \delta & \delta^* & \delta^* \\
  \phi^* \phi_{\text{dft}} & \phi^* \phi_{\text{dft}} & \phi^* \phi_{\text{dft}} & \phi_0
\end{array}
\]

3.3.1 Dative Experiencers

Not all quirky case constructions rely on default agreement. Dative-nominative verbs have an experiencer preverb \(v_{\text{exp}}\) which selects an inherently dative marked subject (the experiencer) and a VP-complement. Suppose that \(v_{\text{exp}}\) has a \(\delta^*\)-feature, but differs from \(v_{\text{trs}}\) in not having a \(\phi^*\)-feature.\(^9\) There is a dialectical split in Icelandic between (53a) and (53b).

(53) a. Jon(dat) like(sg) the books(nom,pl).

\[
\begin{array}{cccc}
  \delta^* & \delta & \delta^* \\
  \phi^* \phi_{\text{dft}} & \phi_0 & \phi
\end{array}
\]

b. Jon(dat) like(pl) the-books(nom,pl).

For many speakers, number agreement is impossible, as in (53a). Since we assume that Icelandic assigns default nominative case to nominals whose \(\phi\)-feature is not the feature of a TE-head and which does not bear inherent case, (53a) is what our assumptions predict. The intervening \(\phi\)-features of the dative experiencer block \(\phi^*\)-probing. The \(\phi\)-features of T must be valued by default agreement. The object nominal gets default nominative case.

Note that in (53b), T heads a \(\delta\)-chain which contains one argument and a \(\phi\)-chain which contains a different argument. Surface form (i.e. movement) is determined by the \(\delta\)-chain. This illustrates an important idea of CC. There is a division of labor in the mechanisms which are used to make agreement relations
visible in surface forms. Movement is used to display $\delta$-agreement, with little or no overt morphology, while morphology is used to display $\phi$-agreement.

For many other speakers, number agreement is allowed or even preferred (with varying strengths), as in (53b). For these speakers, there must be some mechanism for $\phi$-agreement which evades the intervention effect. One might imagine that feature sharing over $\phi_0$ is somehow permitted, as in (54).

\begin{align}
(54) & \quad \text{Jon(dat) like(pl) the books(nom,pl).} \\
 & \quad T \langle \text{Jon v}_{\exp} \text{ like the-books} \\
 & \quad \phi^* \rightarrow \phi_0 \rightarrow \phi
\end{align}

There are good reasons, however, to think that the mechanism of $\phi$-agreement in (53b) is not an exceptional variety of feature sharing, as in (54). If sharing across $\phi_0$ were possible, even exceptionally, we would expect (55). But only nominative case is permitted on books in (55) and similar examples, so that (55) is sharply ungrammatical in Icelandic.

\begin{align}
(55) & \quad *\text{We believe Jon(dat) to-like books(acc,pl).} \\
 & \quad T \langle \text{we v}_{\text{trs}} \text{ believe T}_r \langle \text{Jon v}_{\exp} \text{ to-like books} \\
 & \quad \phi^* \rightarrow \phi
\end{align}

We suppose that the exceptional mechanism is valuation, without sharing, the mechanism which DBP takes to be standard. In place of (55), this correctly predicts (56), where $\phi^+$ has been employed to denote the exceptionally valued $\phi$-feature (initially a $\phi^*$-feature) of $v_{\text{trs}}$. The values of $\phi^+$ agree (partially, we will see below) with the values of the $\phi$-feature on books.

\begin{align}
(56) & \quad \text{We believe Jon(dat) to-like books(nom,pl).} \\
 & \quad T \langle \text{we v}_{\text{trs}} \text{ believe T}_r \langle \text{Jon v}_{\exp} \text{ to-like books} \\
 & \quad \phi^* \rightarrow \phi
\end{align}

Crucially, the object does not come to share a $\phi$-feature with $v_{\text{trs}}$. It therefore is not assigned accusative case and must receive default nominative case.

The derivation of (53b) is therefore:

\begin{align}
(57) & \quad \text{Jon(dat) like(pl) the books(nom,pl).} \\
 & \quad T \langle \text{Jon(dat) v}_{\exp} \text{ like the-books(nom)} \\
 & \quad \phi^* \rightarrow \phi_0 \rightarrow \phi
\end{align}

24
Exceptional valuation, in those speakers who allow it at all, is very restricted. First, the valuing feature (i.e., the initially valued feature) must be a feature of a nominal. Second, the mechanism only values number, not person. Person receives default 3rd person valuation. Since the valuing feature must be a nominal feature, exceptional valuation does not iterate, as shown in (58). See Jónsson (1996), Schütze (1997) and Boeckx (1998) for discussion and alternate analyses.

(58) Mér vörðast/??vörðast Jóni líka bækur
me(dat) seems(sg/??pl) Jon(dat) to-like books(N,pl)
‘Jón seems to me to like books.’

\[
\begin{array}{ccccccc}
\delta & \delta^* & \delta^* & \delta^* & \delta^* & \delta^* & \delta \\
\phi^* \phi_{dft} & \phi_0 & \phi^+ & \phi_0 & \phi & \\
\end{array}
\]

Crucially, exceptional valuation of the matrix T in (58) is not possible, because the potential valuing feature for the \(\phi^\ast\)-feature of the matrix T is not a nominal feature.

Some further examples illustrate exceptional valuation in Icelandic. Jonas (1998) establishes that for many speakers agreement across a dative experiencer is possible, as in (59) and (60).

(59) There seem(pl) someone(dat) to-like books(nom,pl).
\[
\begin{array}{ccccccc}
\delta & \delta^* & \delta^* & \delta^* & \delta^* & \delta^* & \delta \\
\phi^* & \phi_0 & \phi^+ & \phi_0 & \phi & \\
\end{array}
\]

(60) There seems(pl) someone(dat) there to-have been read books(nom,pl).
\[
\begin{array}{ccccccc}
\delta & \delta^* & \delta & \delta^* & \delta & \delta^* & \delta^* \\
\phi^+ & \phi_0 & \phi^* & \phi^* & \phi & \\
\end{array}
\]

Note in (60) that although the valuer of \(\phi^\ast\) is a feature of \(T_r\), it is also a feature of a nominal (as well as a feature of Prt).

(61) Me(dat) seem(pl) the-books(nom,pl) to-have been read.
\[
\begin{array}{ccccccc}
\delta^* & \delta & \delta^* & \delta^* & \delta^* & \delta^* & \delta^* \\
\phi^+ & \phi_0 & \phi^* & \phi^* & \phi & \\
\end{array}
\]

‘The books seem to me to have been read.’
3.3.2 Structural Licensing

It has been known since Sigurdsson (1991) and Freiden and Sprouse (1991) that quirky subjects in Icelandic are subject to some structural licensing condition akin to structural case assignment. It has been sometimes argued that it was, in fact, covert structural case checking. In CC we attempted to accomplish structural licensing by building a case system parallel to structural case checking, but not triggering agreement. In the present framework, structural licensing is a consequence of the Chain Condition.

Consideration of the full range of examples which have been adduced to demonstrate structural licensing apart from case requirements is beyond the scope of this paper, but one example will suffice to show what is at stake. Freiden and Sprouse gave examples like (62), in Icelandic.

(62) *[me(dat) to have been helped] is surprising.

They pointed out that, because the subject of the sentential subject in (62) is inherently case marked, there does not appear to be any problem with its case, as there would be in the English equivalent (63).

(63) *[I/me to have been helped] is surprising.

Considerations of this kind led to speculation that inherently case marked nominals in Icelandic must be assigned covert structural case in some way. That is close to the conclusion we come to here. In our terms, the defect in (62) is that the lower $\delta$-chain is not headed by a TE-head, as required by the Chain Condition. The $\delta$-chains in (62) are shown in (64).

$$
\begin{array}{ll}
\text{T be surprising} & \text{C T_{r} to have helped me} \\
\delta^* & \delta \\
\text{δ} & \text{δ} \\
\end{array}
$$

There must be a C shielding the embedded clause; otherwise it would be a raising construction. The result is that the lower $\delta$-chain in (64) is ill-formed.

It is worth noting here that the view of raising CPs to subject implicit in (64), which connects it with $\delta$-agreement, can perhaps shed some light on the requirement that C be realized overtly in certain configurations. It could be that the contrast in (65a), compared with the lack of contrast in (65b), comes from the presence of a $\delta$-feature on C in (65a). No $\delta$-feature on C can be present in (65b). Otherwise, T would $\delta$-agree with C rather than resorting to an expletive, which is a last resort.

(65) a. That the world is round is surprising.
*The world is round is surprising.

$$
\begin{array}{ll}
\text{T be surprising} & \text{C the world is round} \\
\delta^* & \delta \\
\phi^* & \phi_{dft} \\
\end{array}
$$

26
b. It is surprising that the world is round.
   It is surprising the world is round.
   \[ \langle \text{it T be surprising C the world is round} \delta - \delta^* \phi - \phi^* \]

3.3.3 Nominal Versus Pure Expletive In Icelandic

Icelandic makes it clear that expletive \(\phi\)-agreement is a possibility, at least in some languages. The possibility of expletive agreement makes the choice between nominal and pure expletive attraction somewhat more subtle than the earlier discussion indicated. Consider the derivation of (66b), at the point when expletive insertion must be called on.

(66) It is likely that the world is round.
    \( T \text{ be likely C that the world is round} \delta^* \phi^* \)

The subtlety is that default agreement opens up the possibility that a pure expletive could value the unvalued \(\delta\)-feature of \(T\) and expletive agreement the unvalued \(\phi\)-feature of \(T\). This would allow two different continuations of (66).

(67) a. it \( T \text{ be likely that the world is round} \delta - \delta^* \phi - \phi^* \)

b. there \( T \text{ be likely that the world is round} \delta \rightarrow \delta^* \phi^* - \phi_{dft} \)

The empirical fact appears to be that (67b) is impossible.\(^{10}\) What rules it out? The matter is not clearcut. We tentatively adopt the simple solution of supposing that if expletive valuation of both case features of \(T\) is required, the alternative of valuing them both simultaneously by attracting a nominal expletive is preferred to two different expletive attraction operations.

In line with what is an economy account of the determination of pure versus nominal expletive attraction, the expletive will never create a new case chain. An unvalued \(\delta\)-feature, with a valued \(\phi\)-feature, will attract only a pure expletive. Attraction of a nominal expletive, with \(\delta\)-feature sharing which augments an existing chain, would introduce a new \(\phi\)-chain into the representation.

This economy approach to pure versus nominal expletive choice might have to be reconsidered. There is a crucial Icelandic example which might resolve the issue, but we do not yet know the facts. The analysis above predicts that the expletive
(Exp) in (68a) is the nominal expletive. The structure at the point of expletive attraction is given in (68b). Both case features are unvalued, so a nominal expletive is used to value both agreement features simultaneously, as shown in (68c). If this analysis is confirmed, note that a $\delta$-chain containing two nominals results.

(68) a. $\text{Exp seems someone(dat) that the world is round.}$

b. $\text{T } \langle \text{someone } v_{\text{exp}} \text{ seem C } \ldots \phantom{\delta^*} \delta^* \delta^* \delta \delta \delta \delta \phi^* \phi_0 \phi \phi_0 \phi_0$

c. $\text{it T } \langle \text{someone } v_{\text{exp}} \text{ seem C } \ldots \phantom{\delta^*} \delta^* \delta^* \delta^* \delta \phi^* \phi_0 \phi_0$

4. Conclusion

We will end by highlighting some of the prominent characteristics of the theory we have elaborated, and then point to some outstanding issues that are raised.

The major difference between the proposals in this paper and much recent work is the prominent role played by chains in our analysis. This is largely due to the fact that feature sharing makes possible a more straightforward conception of chains than has been previously available. Case chains are formed by agreement, independently of any movement. Case chains in expletive-associate constructions, for example, are almost identical to their counterparts in corresponding non-expletive constructions. Divorcing case chains from movement makes it possible to give case chains a much more prominent role in the theory.

The Chain Condition is imposed as an output filter, not a derivational constraint. This means that if filtering of the output by the Chain Condition is to be avoided, it must be possible for local operations (primarily selection) to be designed in such a way that ill-formed chains can avoided.

The Chain Condition accounts for the examples like English (69a) and (69b) and Icelandic (69c), the phenomenon of structural licensing (non-licensing in this case) of quirky subjects.

(69) a. *It appears Mary to be believed t likes Sally.

b. *Mary is believed t likes Sally.

c. *It is surprising [ me to have been helped t].

Accounting for examples like those in (69) is not surprising. This is the traditional domain of application of conditions on chains. But surprisingly, the Chain Condition in the form which we have given it also predicts that the case and external
argument properties of preverbs have the characteristics that Burzio’s Generalization describes. It also predicts the absence of object expletives. Crucial to this is that there is no notion of inactivity. Attractors are voracious. They do not care if the features they attract have been valued or not, or whether attraction will destroy well-formed chains or otherwise make them impossible. Only the intervention of a Comp or appropriate intervening features will block attraction. Attraction can be viewed as an invaluable tool, because it is a mechanism to value features. But it is dangerous, requiring careful construction of representations so that its voracious nature does not destroy intelligibility of the output.

The outstanding issue that the theory developed here raises is the nature of the Chain Condition. Our terminology “TE-head” was meant to suggest some connection with temporal event structure. The Chain Condition suggests that what is interpreted is in some sense pairs which consist of an temporal/event head and an argument. It would support our conclusions if some basis for this could be found in temporal and event semantics. A related issue is giving some account for why the particular heads we call TE-heads are singled out as the source of case.

Notes

1. The impetus for this work was a careful study of DBP undertaken by a reading group consisting of Julie Legate, Charles Yang, and the authors. We gratefully acknowledge both the insights and the diligence of our collaborators in that effort, which illuminated some of the problems which this paper attempts to solve. We also acknowledge the attention which Noam Chomsky gave to the group’s concerns, and the helpfulness of his responses.

There are two major revisions of the April 2000 draft. First, in response to generally bad reviews of an earlier attempt to communicate the novelty of feature sharing in compact diagrams, a new pictorial format is used to represent derivations. Second, and more substantive, the section on Icelandic quirky case has been substantially revised and intervention effects are analyzed. We thank Jason Merchant, Cedric Boeckx, Luis López, and Gil Rappaport for very helpful written comments, and an MIT Ling-Lunch audience for their less formal but no less helpful comments. Thanks also to Dianne Jonas, Johannes Jónsson and Halldór Sigurðsson for help with Icelandic. A future major revision of the paper will do more justice to the comments we have received than the present draft does.

2. We have since learned that HPSG (see Pollard and Sag, 1994) has viewed agreement as feature sharing for a long time. That theory, however, is representational. Consequently, the entire derivational process by which features become valued, which is central to our analysis, plays no role in HPSG. The differences will be discussed in the final version of this paper along with connections to Yip, Maling, and Jackendoff (1987).


4. See López (2000) for a notion of “co-valued features” which has some similarity to the idea of feature sharing.

5. If there are no branching specifiers, this notation allows unambiguous representation of tree structure without using bracketing, as in (i) below. If there are branching specifiers, some bracketing is required, as in (ii). In the tree structures, the branching maximal projections are indicated
and their heads are underlined. In each case, the linear display uniquely determines the corresponding tree structure.

\[ i. \ (A \ B \ C \langle D \ (E \ F \ G) \ \text{Max}\ \langle B \ C \ D \ E \ \text{Max}\ ) \]

\[ ii. \ A \langle [B \ C] \ D \ E \ \text{Max}\ ) \]

6. According to Lieber, quoting from Kálmán:

Vogul verbs can occur in two different sorts of conjugations: the Indefinite (subjective) conjugation, and the Definite (objective) conjugation. The later is used when the object is “defined,” which is “1. when it is preceded by a demonstrative pronoun ... 2. when it has a possessive suffix ... 3. when the object is a personal pronoun ... 4. when it is already known or has been previously mentioned ... 5. when the object is in a subordinate clause” (Kálmán 1965, 53).


8. The Chain Condition, as we have formulated it, does not rule out:

\[ (i) \ T \langle \text{Sub} \ v \text{ Obj} \delta^* \delta \delta^* \delta \]

Can a preverb v select a subject with a \(\delta\)-feature, but no \(\phi\)-feature? If so, what structural case is assigned to the object above, T-case or v-case? We propose that expletive there has a \(\delta\)-feature, but no \(\phi\)-feature, but it may be that arguments must have a \(\phi\)-feature, making consideration of (i) moot. The discussion of inherently case-marked nominals in Section 3.4 is relevant. We propose there that inherently case-marked nominals do have \(\phi\)-features which would block (i), in spite of the fact that they do not enter into \(\phi\)-agreement relations. Precisely what is at stake is not entirely clear, but the issue should be noted.

9. We will not explore the possibility of v which has both \(\delta^*\) and \(\phi^*\)-features. This would allow a transitive verb with a quirky subject. There are several potential candidates in Icelandic, but if they do exist, they are rare.

10. This should be read cautiously. We have not yet conclusively pinned down the empirical facts on this point. The phonological form of the nominal and pure expletives is the same in Icelandic, when they occur overtly. They can only be distinguished because certain contexts force a null pure expletive.
References


López, Luis. 2000. Locality of Agreement. ms. (University of Illinois at Chicago)
