How Sentences Grow in the Mind
Agreement and Selection in Efficient Minimalist Syntax
John Frampton and Sam Gutmann

The title above is intentionally provocative. In spite of the title, it is not our intention to propose a model of performance. The “growth of sentences” will be considered at a significant level of abstraction and the discussion will be limited to syntax. But we do want to make clear our methodological bias. It is commonplace that generative grammar answers the question of how language makes “infinite use of finite means.” But work in syntax often pays little attention to elucidating how finite syntactic means are, or even could be, brought to bear in generating grammatical sentences. Our goal in this paper is to be as explicit as possible about how finite means are employed to construct grammatical sentences.

For several years, we have been pursuing the idea that the architecture of the language faculty is such that the syntactic computations required to build well-formed sentences are very straightforward and efficient. We stress that the focus is syntactic computation at a certain level of abstraction. The model of syntax that we propose is not a performance model in any sense. In performance, many factors not considered here would have to be brought into play. We have nothing to say about how this should be accomplished. In spite of this, we think that it is possible to get some idea of how the syntactic model that we propose might be integrated into a theory of performance. Since performance is part of the empirical domain which theories of syntax should ultimately shed some light on, we take this to be support for the proposals that will be made.

Since the view of syntax which is developed in this paper touches on issues of the appropriate level of abstraction for grammatical theories, it is useful to outline the view that underlies the proposals which follow. “Level of abstraction,” as used here, indicates how far the theory is abstracted away from a theory of performance. In spite of the fact that syntactic theories are inevitably set at a significant level of abstraction, we do assume two desiderata for syntactic theory. They are meant to enunciate a particular research program, not to give an a priori statement of a truth that all research programs must or should adhere to. Like all methodological assumptions, they are to be judged by the theory they lead to. We will offer little argument in their favor at the outset.

The first desideratum is that the theory should be usable by a theory of performance. Of course, in the absence of a theory of performance, it is not easy to judge what is usable and what is not. Nevertheless, the extent to which a syntactic theory is plausibly usable by, or incorporable in, a theory of performance argues in its favor, in our view. The second (related) desideratum is that the theory should account for how, from the syntactic side, the language faculty makes “infinite use of finite means.” We take such an account to be the defining property of generative syntax, viewed as the study of the syntactic aspects of the language faculty, which is finite.

Syntactic theories which are based on free generation and filtering completely ignore the question of how syntax knowledge might be brought to bear in performance. It is not easy to see how this might be accomplished if syntactic knowledge is limited to filters on representations. Versions of the Minimalist Program which rely on numerations face a similar problem. Numerations simply transfer the question of how successful derivations can be constructed to the question of how successful numerations can be built. But the latter question is part of syntactic theory and cannot be avoided.
In GB theory and earlier theories, the job of syntax was to transform a given D-structure into an S-structure; an input-output problem. X-theory generated D-structures and syntax operated on these deep structures. The question of computational efficiency or plausibility for an input-output problem is well-defined because the task is well-defined. Since Chomsky (1994), with the elimination of D-structure, the issue of computational efficiency is much more subtle. For a given input-output problem, the notion of computational efficiency is well-defined, but in the absence of such a problem, how is “efficiency” to be understood? Chomsky’s notion of a “numeration” provides some grounding for considerations of efficiency, since it recreates a kind of input-output problem. But it suffers from the considerable problem, fatal in our view, that it offers no guidance about how viable numerations can be assembled. Most numerations will not lead to a well-formed output.

We take the view that sentences grow, partially determined by syntax but partially nondeterministically; choice of lexical items at various points in the derivation is nondeterministic. Any discussion of efficiency in lexical choice belongs to a theory of pragmatics and we will have nothing to say about it. In spite of the difficulty in posing the question of efficiency outside the context of a clearly defined input-output problem, we will identify certain characteristics of theories of the syntactic computation as fatal to the efficient growth of sentences and propose a syntax in which sentences grow efficiently.

Besides the nondeterministic choice of lexical items, the proposed syntax consists of two operations: merger driven by selection (where selection is understood in the broadest sense of specifying what heads a given head must or can immediately dominate) and agreement (which may be accompanied by movement or merger). We take the view that merger is driven by selection (rather than free, with the semantic interface/component discarding ill-formed structures) because of our demand that syntax be efficient. There is also empirical support for this view: selectional restrictions appear to be highly local. If they were imposed by the semantic interface/component after the full syntactic tree was handed over, there would be no reason to expect local as opposed to global constraints. Section 1 contains a discussion of efficient sentence growth, and the role of selection is elaborated in Section 8.

The core of syntax, however, is agreement, and the bulk of this paper (Sections 2-7) is devoted to detailing a theory of the mechanisms of agreement, including feature sharing (Frampton and Gutmann 2000) and the distinction between phi and delta agreement (Frampton and Gutmann 1999), and explaining how the theory operates, both in “ordinary” cases and in challenging examples involving expletives and Icelandic quirky case.

1. The interplay of local syntax and interface conditions

Our intention is to work out a version of minimalist syntax for which it is clear how the syntactic theory might be incorporated into a performance theory. Before discussing the details of the proposed architecture, it is useful to make clear what we see as the crucial problem in syntactic computation, from the standpoint of efficient computation.

We will assume a bare phrase structure framework in which morphemes are chosen in the course of the derivation. A derivation proceeds by operations/steps of various kinds:

1. choice of a morpheme;
2. merger driven by argument structure or selection;
3. agreement, which may or may not be accompanied by merger or movement.
We will say that an operation acts on a state and transforms it to a new state. Some of the states that the growth of a sentence traverses may consist of multiple disconnected representations. In the derivation of *the woman saw the man*, for example, the DP $\alpha = [\text{the woman}]$ may be formed before the phrase $\beta = [v^* [\text{see} [\text{the man}]]]$ is formed. Here, $v^*$ is the transitive verbalizer. At some point in the derivation, the state of the computation consists of two disconnected phrases. The word “state” is used to emphasize the nature of the intermediate points in syntactic derivations. It is used also to keep contact with the idea that our view of syntactic derivations is that they model (at a more or less considerable degree of abstractness) the sequence of states that the brain goes through in producing a sentence.

Syntax has two components: Local Narrow Syntax (LNS) regulates the possible transitions, and the Interface Conditions (IC) determine whether a state is a well-formed terminal representation (wfttr). Some of the kinds of conditions that have been proposed, by one researcher or another, for LNS and IC are listed in (1).

(1) 1. Local Narrow Syntax (LNS)
   a. conditions on what can merge with what (Θ-Criterion, subcategorization, selection in general);
   b. conditions on what can agree with what;
   c. conditions under which agreement induces movement;
   d. locality conditions on agreement and movement; and
   e. the architecture of the cycle.

2. Interface Conditions
   a. a prohibition against uninterpretable features;
   b. Case filter;
   c. Θ-Criterion;
   d. various conditions on chains; and
   e. a requirement that well-formed terminal state consists of a single phrase of a certain type

One can think of the growth of a sentence as the successful step by step extension of a path through a labyrinth of possibilities determined by the LNS, ending at a representation which satisfies interface conditions. The path starts at the empty state and, since it is successful, ends at a wfttr. It could be that some states which the LNS can reach are doomed, with no subsequent sequence of operations admitted by the LNS leading to a wfttr. Doomed states are indicated by a the gravestone symbol $\text{RIP}$ in Figure 1 on the next page. The key to the efficient growth of sentences is the avoidance of doomed states. Doomed states require backtracking, the computation must be undone and returned to an earlier state. This introduces inefficiency and complexity into the computation.

The state underlying (2), for example, is doomed:

(2) it to seem John to be here

The subject of the embedded clause does not have case. Under standard assumptions on the locality conditions on agreement and the changes of state permitted by the LNS, there is no sequence of operations under which the subject can ever get case.

The state is doomed. Successful growth must either avoid such doomed states, or the grammar must somehow be provided with a “doom detector” so that the hopeless pro-
cess can be stopped, backtracking undertaken, and the derivation sent off down a new path through the labyrinth. We take the strong point of view that the growth of sentences depends on the LNS being powerful enough to avoid transitions (without lookahead) to doomed states. We began the exploration of this idea in Frampton and Gutmann (2002), calling it “crashproof syntax”.

We begin by considering a case-assignment problem which under early Minimalist Theory proposals made the immediate detection of doomed states impossible.

2. The problem of case assignment to participles

Explaining the mechanism of case assignment in examples like (3) has been central to the development of syntactic theory for many years.

(3)  a. Civilians(3pl,nom) are believed(pl,nom) to have been injured(pl,nom).
     b. They believed civilians(3pl,acc) to have been injured(pl,acc).
     c. There are believed(pl,nom) to have been injured(pl,nom) civilians(3pl,nom).
     d. They believed there to have been injured(pl,acc) civilians(3pl,acc).

For ease and clarity of presentation, assume here a variant of English which allows expletives constructions as needed in (3) and displays participle agreement for number and case as in some Germanic languages (Icelandic, for example).
The history is much too long and involved to review, but some high spots are worth highlighting. In the late 80s, the prevailing view was that in the movement cases (3a,b), participial number agreement was done cyclically as the nominal raised through the specifier of a participial head. In the highest position, a "case-marking position," case was assigned and percolated down the chain to the participial heads. For (3a), for example, the derivation looked like (4):

\[
\begin{array}{cccccccc}
\text{T} & \ldots & \text{Prt} & \ldots & \text{Prt} & \ldots & \text{civilians}
\end{array}
\]

For the expletive cases, (4c,d), since the results appear to be so similar to the raising cases, appeal was made to some kind of abstract chain formation or coindexing. Number agreement percolated up this abstract chain and case percolated down it.

From the standpoint of efficient computation, the early proposals of Chomsky’s Minimalist Program (henceforth MP) were mixed. On the one hand, the bare phrase theory of opened up the possibility of an efficient algorithm deriving wfrs. On the other, the reliance on numerations partially begged the question since there did not appear to be any algorithm for constructing successful numerations. A reliance on comparison of derivations presented further obstacles to efficient computation of wfrs. Frampton and Gutmann (1999) overcame these difficulties, but case assignment remained a problem.

The feature-checking proposals of Chomsky 1995, Chapter 4, are difficult to reconcile with efficient computation, since the LNS can allow derivations to progress deep into a cul-de-sac. Under those proposals, the nominal is inserted with a case feature, which must be checked in the course of the derivation. According to this proposal, both (5a) and (5b) would be valid intermediate representations.

(5) a. believe [ there to have been injured(pl,nom) civilians(3pl,nom) ]
   b. believe [ there to have been injured(pl,acc) civilians(3pl,acc) ]

The problem from the standpoint of avoiding doomed states is that the states in (5) are transformed to doomed states if the “wrong” case checker is merged into the derivation. The mode of operation of the theory is simply to exclude those derivations by an output filter which bars unchecked features or somehow cancels a derivation if cases clash. Feature checking produces massive overgeneration and filtering rather than anything that could be called the growth of wfrs.

Chomsky’s “Derivation by Phase” (2000, henceforth DBP) elaborates a theory of agreement which is much more suited to efficient computation. First, the agreement mechanism is feature valuation, not feature checking. Second, the participial head is assigned case (i.e. has its case feature valued) directly by the accusative case assigning head, not by agreement with the nominal. These assumptions allow the problem discussed above in relation to (5) to be overcome. In the derivation of (3), for example, the nominal *civilians* enters the derivation with a valued \( \phi \)-feature and an unvalued case feature, while the participial head enters with an unvalued number feature and an unvalued case feature. Prt probes and values its \( \phi \)-feature in the configuration:

\[
\begin{array}{cccc}
\text{Prt} & \text{injure} & \text{civilians} \\
\phi[] & \phi[3pl] \\
\text{Case[]} & \text{Case[]}
\end{array}
\]
Later in the derivation, the transitive preverb (associated with believe) probes in the configuration:

(7) \( v_{tr} \) believe ... Prt injure civilians

\[ \phi[\ ] \quad \phi[\text{pl}] \quad \phi[\text{3pl}] \]

\[ \text{Case[\ ]} \quad \text{Case[\ ]} \]

Chomsky proposed that, in spite of the apparent intervention effect, \( v_{tr} \) can see past Prt and agree with civilians, valuing its case feature Acc in the process. The circumvention of the intervention effect relies on the observation that participles are inflected only for number, not for person. In Chomsky’s terminology, they do not carry “full \( \phi \)-features.” See DBP for the details. While ingenious, this solution is not sufficiently general. It does not extend, for example, to (3b) above. See Frampton, et al (2000) for a detailed discussion of the problems.

The next section proposes a modification of the DBP theory of agreement that permits a satisfactory solution to the case problem. The leading idea is that the \( \phi \)-features on Prt and the nominal in (6) become identical in the course of the derivation. Not simply identically valued, but identical. Agreement with the \( \phi \)-feature associated with Prt is simultaneously agreement with the \( \phi \)-feature associated with the nominal because the same \( \phi \)-feature is associated with both heads. Under this theory, there is no intervention which must be overcome.

3. Feature Sharing

For the present, we consider only the features which play a role in the DBP case system: number, person, and case features. This feature system will be modified later, but these three features from the DBP system are sufficient to illustrate the idea of feature sharing.

We begin with Num and, for the sake of simplicity, consider a language in which there is only singular and plural number. Features are usually thought of as properties. If the syntactic terminal \( \alpha \) “has” the feature Plural, we understand this as meaning that the plurality value of \( \alpha \) is plural. Essentially, this supposes the existence of predicates PLURAL and SINGULAR which are true or false of syntactic terminals. The introduction of the idea of an unvalued number feature complicates this picture. One approach is to extend the idea above and introduce a third predicate UNVALUED-NUMBER. Another alternative, which we pursue, is to consider number features to be objects, not properties. We suppose that there is a class of objects called number features. The predicates SINGULAR and PLURAL are true or false of number features, not directly of syntactic terminals. There is no core predicate UNVALUED-NUMBER. We can say that a syntactic terminal has an unvalued number feature if it “has” (i.e. contains) a number feature for which neither SINGULAR nor PLURAL is true.

If a syntactic terminal A with an unvalued number feature probes and agrees with a syntactic terminal B for number, the operation can be (and usually is) thought of as (8). The change is minimal, simply valuing the number feature of A.

(8) Num ... Num ... Agree Num ... Num ...

\[ \text{A} \quad \text{B} \quad \text{A} \quad \text{B} \]
Since we assume that number features are objects, not properties, there is another way to view the agreement process.

We pursue this approach to agreement. Agree induces feature sharing, with matching features coalescing into a single shared feature, which is valued if either of the coalescing features is valued.\textsuperscript{3}

If this idea is extended to all of the features which enter into agreement relations (at least number, person, and case in DBP), the agreement process induces operations like that shown below:

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 (10b) above could just as well have been depicted as (11). 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 that T is a nominative case assigner. Agreement with T will result in valuing a Case feature. (Henceforth we will omit the ellipsis marks which indicate features not entering into the agreement process. It should be assumed that such features may be present.)

Application of Agree in (12) 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 (12), 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 (10) and (12) now established, a more compact notation will be employed, which is intended to summarize both (11) and (12).

\begin{equation}
T \quad A \quad B
\end{equation}

\begin{align*}
&\text{Per}^*\quad\quad\quad \text{Per}[3] \\
&\text{Num}^*\quad\quad\quad \text{Num}^*\quad\quad\quad \text{Num}[pl] \\
&\text{Case}^*\quad\quad\quad \text{Case}^*
\end{align*}

The diagram (13) 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. Note that there is no deletion of features, only coalescence of features. Note also that the representation (13) has the potential to be misleading if the representational convention is not understood. The three Num features connected by horizontal lines in (13), for example, represent a single shared feature, not three different features related in some way.

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:

\begin{equation}
T \quad A \quad B
\end{equation}

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

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. Structural case morphology 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 take this to its logical minimalist conclusion and remove structural case features from the syntax entirely. Structural case is assigned in the mapping from syntax to morphology. A head which a nominal shares its $\phi$-feature with determines the structural case of the nominal if the head is a case-assigning head. Under this view of case assignment, traditional formulations of the Case Filter are untenable. A replacement will appear in the next section. If a nominal is not assigned case in this way, and is not assigned inherent case, and there is no default mechanism for assigning case, and if the morphology rejects nominals which have not be assigned a case feature, then the derivation containing that nominal will fail at the morphological interface. This is not sufficient, however, to capture the full range of Case Filter effects. The notion of inactive element plays a major role in the DBP theory of locality. Inactivity is determined in that theory by valuation of a case feature. This is not an option if case features are eliminated. This will have major implications for the theory of locality which will be developed later.

To illustrate the mechanism being proposed, consider the full derivation of (3b). The
intervention problem is straightforwardly bypassed. 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 \( ⟨ \) on their left.\(^5\) The Tense of raising infinitives is denoted by \( T_r \). We assume that \( T_r \) initially has an unvalued \( φ \)-feature.

\[ (15) \] They believed civilians(3pl,acc) to have been injured(pl,acc).

1. Prt injure civilians
   \[ \phi^* \square \phi \]
2. \( T_r \) have -en be Prt injure civilians
   \[ \phi^* \square \phi^* \square \phi \]
3. (civilians\(_j\) \( T_r \) have -en be Prt injure \( t_j \)
   \[ \phi \square \phi^* \square \phi \]
4. (they \( v_{\text{ins}} \) believe (civilians\(_j\) \( T_r \) have -en be Prt injure \( t_j \)
   \[ \phi \square \phi^* \square \phi \]

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

On line 3 above, movement was carried out. Because of feature sharing, however, this plays no role in the derivation. We could just as well have left the nominal in place, as in (16). Linking (in the sense of sharing features), which persists, identifies potential landing sites. Placement in surface form can be determined at the point of spellout on the basis of shared features.\(^5\)

\[ (16) \] 1. \( T_r \) have -en be Prt injure civilians
   \[ \phi^* \square \phi \]
2. (they \( v_{\text{ins}} \) believe \( T_r \) have -en be Prt injure civilians
   \[ \phi^* \square \phi^* \square \phi \]

4. \( δ \)-agreement

The previous section 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. We will not, however, pursue the \( DBP \) framework of assumptions. Instead, we return to the theory elaborated in Frampton and Gutmann (1999, henceforth \( CC \)) and use the idea of feature sharing to simplify it.

Much of the complexity of \( CC \) came from the insistence that the theory adequately account for expletive insertion and for quirky case phenomena. The \( DBP \) framework drives the formation of NP-chains (we use the term descriptively) by \( φ \)-feature agreement, with the EPP determining which instances of \( φ \)-agreement trigger movement. In \( CC \), we took the position 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.\(^6\)

Instead, we will suppose that two features are relevant to the case/agreement system. The first is just \( φ \), which comes in valued and unvalued varieties. Possible values are the usual person, number, and gender (or noun class) values. The second will be called \( δ \). 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. In many cases, \( \delta \)-agreement is entirely parallel to \( \phi \)-agreement. But in pure expletive constructions (i.e. \textit{there}-type expletive constructions) and quirky case constructions, the \( \delta \)-agreement structure is different than the \( \phi \)-agreement structure.

Before we advance a tentative proposal for what values \( \delta \) assumes, it is worth emphasizing 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 a 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.\textsuperscript{7} We speculate that this morphology reflects the values of the \( \delta \)-feature.

\( \delta^\ast \) and \( \phi^\ast \) will be used as compact notations for unvalued \( \delta \) and \( \phi \)-features. Suppose for a start that nominals enter the syntax with valued \( \delta \) and \( \phi \)-features and the structural case assigners initially have \( \delta^\ast \) and \( \phi^\ast \), 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 (\textit{if} 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, it merges with the phrase that it selects. The unvalued features of that head are then satisfied by applications of Attract. We call the head which is the locus for applications of Attract the \textit{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.\textsuperscript{8}

As in CC, Attract finds targets of attraction by first conducting a top-down search from an unvalued feature of the pivot, seeking a feature which matches this feature. If a matching feature 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” (i.e. outside the current representation) 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 to value the unvalued pivot feature 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 (17a). The derivation is represented by the diagram (17b). Recall that (17b) 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.

(17) 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 (17). As proposed above, 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 out 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 δ-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 Tr (raising T). We assume that Tr has both an unvalued δ-feature and an unvalued φ-feature. Tr must therefore be distinguished from tensed T other than by its δ and φ-features.

(18) T ⟨Mary vtrs expect Tr ⟨Bill vtrs like Sally
        δ∗——δ δ∗——δ δ∗——δ δ∗——δ
        φ∗——φ φ∗——φ φ∗——φ

Consider this derivation at the point just after Select has introduced vtrs into the computation, with its selected VP complement and selected subject.

(19) ⟨Mary vtrs expect Tr ⟨Bill vtrs like Sally
        δ δ∗——δ δ∗——δ δ∗——δ
        φ φ∗——φ φ∗——φ

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

(20) We expect them(acc,pl) to-have been seen(acc,pl).

\[
\ldots \quad v_{trs} \quad \text{expect} \quad T_r \quad \text{to-have} \quad \text{Prt} \quad \text{see} \quad \text{them} \\
\delta^* \quad \delta^* \quad \delta \quad \delta \\
\phi^* \quad \phi^* \quad \phi^* \quad \phi
\]

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.

5. Chains, the Chain Condition, and Locality

Feature sharing, triggered by Attract, links heads via shared \( \delta \) and \( \phi \)-features. It builds structures like (21), which is one of the linked sets of heads created by feature sharing in (20).

(21) \( v_{trs} \quad T_r \quad \text{Prt} \quad \text{people} \\
\phi^* \quad \phi^* \quad \phi^* \quad \phi
\)

We call the set of four linked heads in (21) 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.\(^9\) 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, \( v_{trs} \), \ldots). 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
13

in Icelandic, which we discuss in a later section, makes this untenable. López (2002) 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 (22), from Kiswahili.\(^{10}\)

(22) Juma a-ta-kuwa a-me-pika chakula

$Juma \{3s, NC1\}$-future-$be \{3s, NC1\}$-perfective-$cook$ $food$

‘Juma will have cooked food.’

The diagram (23) shows the agreement structure we assign to (22).

(23) $T \text{ be } v_{asp} \langle Juma \text{ v}_\text{trs} \text{ cook } \text{ food} \rangle$

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

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

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, but $T_r$ 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.

5.1. 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 (perhaps integrated into the mechanism of agreement), and many variations on these themes. All three devices are employed in DBP. 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 (1981) and the early development of GB-theory.

Consider, for example, (24).

(24) *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 (25).

(25) *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 (25). Combining these two requirements led to the condition:

(26) The head, and only the head, of an A-chain is in a case position.
We will shortly propose that a version of (26) plays a key role in minimalist syntax.

Rather than rule out examples like (24) and (25) 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 applied at various points in 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 interface 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. They were part of LNS. Rather than a derivational constraint, we will assume that (27) is part of IC, the post-syntactic interface conditions. Its effect is to mandate a one-to-one correspondence between δ-features and TE-heads as an interpretability condition at the output interface.

(27) 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 IC is simply an ill-formed derivation. Although IC cannot directly determine the course of a derivation, we will argue that it does play a central role in determining the design of LNS and an important role in shaping the syntactic lexicon. The assumption that syntax is computationally efficient implies that the computation avoids doomed states, ones that will violate (27) at the interface, for example. This places restrictions on LNS. It also restricts the syntactic lexicon, eliminating heads from the syntactic lexicon which have a feature structure or selectional properties which lead to doomed derivations.

In addition to the locality which is implicit in the mechanics of the Attract algorithm, we also assume (28). It is a reincarnation of the “Tensed-S Condition” of early GB theory.

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

C-Complement Opacity is part of Local Narrow Syntax.

Clearly, (28) is related to Chomsky’s notion of a “phase.” Chomsky (2000, p. 106) proposed that locality is in part the consequence of cyclic spellout. Under this theory, the complements of certain heads, phasal heads, are transferred to phonology (“spelled out”) at certain points in the derivation. Spellout syntactically isolates this material from
higher material. In spite of its conceptual appeal, there is reason to doubt the theory of locality which Chomsky's theory of phases leads to. Legate (2003) shows that the evidence adduced to establish that transitive preverbs are phasal equally establishes that participial preverbs are phasal. But case-marking in expletive constructions in which the object of a passive must be case marked in a configuration in which multiple participial heads intervene between the object and the case assigner is incompatible with Prt being phasal. Consider, for example:

(29)  a. There is believed to have been killed many people.
    b. [Prt believe to have been [Prt kill many people]].

In the derivation of (29a), the case-assigner for the nominal many people still has not been introduced into the derivation at the point that (29b) is formed. If participles establish phases, the nominal is buried inside multiple phases.

The locality conditions we propose are fairly weak; only (28) 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:

(30) Select, (Attract)*, Select, (Attract)*, Select, (Attract)*, ...

The applications of Attract in (30) are largely automatic, with little optionality. What determines whether or not a successful derivation is constructed are the applications of Select, in particular, the properties of the lexical choices that are made at each step. In order to illustrate this, consider the derivation of (31a), at the point in the derivation given in (31b).

(31)  a. It seems that Sally likes Mary.
    b. T ⟨Sally 𝑣_{trs} like Mary

It is crucial that at this point in the derivation that the TP (31b) be embedded as a C-complement. If not, continuing the derivation of (31a) leads to (32a) and then (32b).

(32)  a. T seem T ⟨Sally 𝑣_{trs} like Mary
    b. T seem T ⟨Sally 𝑣_{trs} like Mary
Instead of expletive insertion, the matrix T agrees into the embedded clause and a Chain Condition violation occurs: the lower T, a TE-head, is forced into a δ-chain with another TE-head.

If (31b) is embedded as a C-complement, the derivation continues to (33a) and (33b) results. C-Complement Opacity provides the needed isolation between the embedded clause and the matrix TE-head.

\[
\begin{align*}
\text{(33)} & \quad \text{a. } T \quad \text{seem} \quad C \quad T \quad \langle \text{Sally} \rangle \quad v_{\text{trs}} \quad \text{like} \quad \text{Mary} \\
& \quad \delta^* \quad \delta^* \quad \delta \quad \delta^* \quad \delta \\
& \quad \phi^* \quad \phi^* \quad \phi \quad \phi^* \quad \phi \\
\text{b. } & \quad T \quad \text{it} \quad T \quad \langle \text{Sally} \rangle \quad v_{\text{trs}} \quad \text{like} \quad \text{Mary} \\
& \quad \delta \quad \delta^* \quad \delta \quad \delta^* \quad \delta \\
& \quad \phi \quad \phi^* \quad \phi \quad \phi^* \quad \phi \\
\end{align*}
\]

We will return later to discuss expletive insertion in more detail. Here we simply note the fact that the C in (33a) prevents the matrix T from valuing its agreement features (i.e. δ* and φ*) 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 (33b).

Note that *seem* does not always select a CP complement. The clausal complement of *seem* in (34) is not a CP.

\[
\text{(34)} \quad \text{Mary seemed t to like Sally.}
\]

For the raising verb *seem*, the choice of CP or TP is straightforward. If the embedded clause is headed by T, 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 is that only C selects finite T.

5.2. 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 to an ill-formed chain in the output. The complementizer was needed to protect a δ-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 δ-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 δ and φ-features (i.e. assigns accusative case), but does not select a subject. A derivation would start out along the lines of (35a), then proceed to (35b) after clausal Tense is merged. The continuation is forced and yields the hopeless (35c). 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 δ-chain.
It is clear that unless the preverb has a subject with a $\delta$-feature, the derivation will be ill-formed. The preverb will not end up as the head of its $\delta$-chain, but the Chain Condition demands that a TE-head must be the head of a $\delta$-chain. A subject $\delta$-feature is needed to shield the $\delta$-feature of the preverb from attraction by T. This is just the empirical fact which Burzio’s Generalization describes.

Compare (35) with the case where the preverb has a subject with a $\delta$-feature.

(36) Sally(nom) grows tomatoes.

A TE-preverb necessarily has a $\phi$-feature, because the Chain Condition demands that it head a $\delta$-chain. This preverb must then have a subject with a $\delta$-feature.12

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 part of IC, with no role in LNS. As we proceed, we will see that other crucial facts about $\delta$-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.

5.3. 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. Similarly, our proposal here that the Chain Condition, part of IC, plays a central role in syntax threatens the idea of “derivational simplicity.” In principle, the characteristics of IC could be such that 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 true. We have already seen examples. The fact that only C selects finite T prevents a Chain Condition violation, as does the requirement that \( v_{\text{trs}} \) selects a subject. We return to this question in Section 8.

6. Expletives

There is a cluster of interpretive properties related to information structure: new information, focus, nonspecificity, and indefiniteness. In what follows, we will call this simply focus. Two types of configurational interpretation of focus have been widely discussed in the literature. One type is movement to the specifier of a focus head, with interpretation tied to the specifier position of this head. Another type, discussed most prominently by Diesing (1992), ties a focus interpretation to certain VP internal positions. Movement to Spec[\( X \)] for some focus head ties movement to the interpretive properties of the target position. The DBP analysis of object shift in Icelandic (OS), which derives from Holmberg (1999), proposes a significant innovation in syntactic theory by allowing movement possibilities to be tied directly to the interpretive properties of the source position of the movement rather than the target position.

Since the network of assumptions on which it is based differs in a number of key respects from the theory developed here, particularly because phases and movement to the edge of phases play such a large role in the DBP analysis, we will not attempt to review the details of the analysis here. Instead, we try to extract the conceptual core of the analysis and then use it to give an account of there-type expletive constructions. The core idea is that nondeterminism can be introduced into the syntactic computation on a language specific basis by allowing an alternative to the standard syntax. Choice of the alternative must have interpretive consequences. We can think of it as being licensed by these consequences for interpretation. Icelandic assigns a focus interpretation to VP-internal objects, provided the verb has raised out of the vP. Icelandic has a syntactic option for avoiding this focus interpretation. Chomsky proposes that the option is to treat the transitive preverb as if it had an EPP feature. The position of object shift is higher that Spec[\( v_{\text{trs}} \)], but Chomsky assumes that Spec[\( v_{\text{trs}} \)] is a necessary escape hatch on the way to the higher position that the shifted object ends up in. Employing this option carries with it interpretive implications. It signals that a VP internal focus interpretation is being avoided. Choice of the syntactic option is therefore anomalous if the verb has not raised, because if it has not raised there is no focus interpretation to avoid. This is Holmberg’s Generalization.

If we consider there-type expletive constructions in English, the complementarity is striking. In these constructions, exceptional syntax, non-movement to Spec[T], is associated with a focus interpretation. Icelandic licenses exceptional movement to Spec[v] in order to avoid a focus interpretation. English licenses exceptional non-movement to T with in order to force a focus interpretation. Under the plausible assumption that the expletive itself is irrelevant to semantic interpretation, it must be the case that the focus (indefinite/nonspecific) interpretation is associated with particular configurations. In English, a focus interpreta-
tion is associated with the predicate internal subject position of certain predicates. Corresponding to exceptional agreement-induced movement to Spec[v] in Icelandic, English has exceptional non-movement to T. In Icelandic, the exception is licensed to avoid a focus interpretation, while in English it is licensed to avoiding the possibility of a non-focus interpretation. A non-focus interpretation of the nominal is possible in (37a), but not in (37b).

(37)  a. Someone T be [t in the room].
  b. There T be [someone in the room].

The next step is to identify the nonstandard derivational option that is allowed in English in order to ensure that the relevant nominal remains predicate internal. There are two possibilities. It could be that EPP-forced movement associated with δ-agreement is annulled and that there is inserted strictly as a phonological filler of Spec[T], with no agreement with T. One might imagine that the derivation of there is someone in the room is:

(38)  T be someone in the room → T be someone in the room
      \[\phi^* \phi \delta^* \delta \phi^* \phi \delta^* \delta \] → \{there T be someone in the room
      \[\phi^* \phi \delta^* \delta \]

The expletive in (38) cannot be inserted with a δ-feature, because a Chain Condition violation would result. δ-chains must be headed by a TE-head. The possibility of there with no δ-feature, however, is a fatal flaw in an analysis along these lines. There would be nothing to stop the derivation of there seems there to be someone in the room, which must be ruled out. At the point in the derivation given in (39), the same motivation that licensed non-movement and there-insertion in the embedded clause (keeping someone in a focus position) would license it in the higher clause.

(39)  T seem \{there T be someone in the room
      \[\phi^* \phi \delta^* \delta \phi^* \phi \delta^* \delta \]

The result would be there seems there to be someone in the room.

Another approach seems more promising. Up to this point, we have assumed that when valued and unvalued features coalesce, the result assumes the value of the valued feature. Suppose that now, as an option, coalescence of δ* and δ can produce a shared unvalued δ-feature, rather than a shared valued δ-feature. Call this exceptional feature sharing (EFS). Like the Icelandic option, it is licensed by its effect on the resulting interpretation.

Consider now the derivation of there is someone in the room, which exploits EFS.

(40)  T be someone in the room → T be someone in the room
      \[\phi^* \phi \delta^* \delta \phi^* \phi \delta^* \delta \]

The (δ)* is used to indicate a formerly valued feature which becomes unvalued as a result of EFS.
Crucially, the $\delta$-feature of the pivot $T$ remains unvalued, so it probes again. Since it cannot be valued internally (and only if it cannot be valued internally), it can be valued by attracting an element directly from the lexicon. The element must be an expletive, because this is unselected merger. Assume that there has a valued (nonspecific) $\delta$-feature. The derivation above continues to:

\[(41) \begin{array}{c}
\langle \text{there} T \text{ be someone in the room} \\
\phi^* \quad \phi \\
\delta \quad \delta^* \quad (\delta)^* \\
\end{array} \]

$CC$ gave accounts of (42a) and (42b) without relying on numerations. These accounts are not valid under the assumption that it is EFS that is responsible for keeping the nominal in its base position.

\[(42) \begin{array}{ll}
a. & *\text{There seems someone to be in the room.} \\
b. & *\text{There seems there to be someone in the room.}
\end{array} \]

An EFS-based account is, however, straightforward. EFS is only licensed as an option to keep nominals in a position to which an indefinite/nonspecific interpretation is assigned. Under the entirely plausible assumption that it is only predicate internal positions which can be assigned the special interpretation, (42a,b) are immediately ruled out. EFS cannot be employed to keep a nominal in an embedded Spec[T] position and it cannot be employed to keep an expletive in this position.

A further advantage to basing expletive insertion on EFS licensed by a forced focus interpretation is that it gives an immediate account of why expletives are never associated with the object of a transitive verb. Since the standard syntax keeps objects in their base position, no exceptional syntax with the same result is licensed. Expletive insertion (of the there-type) can only occur as a reflex of employing EFS to prevent $\delta$-agreement with an EPP-head.

The next task is to work through a variety of expletive constructions and show that the predictions which are made about raising and $\phi$-agreement correspond to the empirical facts.

\[(43) \begin{array}{c}
\langle \text{there} T \text{r be some people in the room} \\
\phi^* \quad \phi \\
\delta \quad \delta^* \quad (\delta)^* \\
\end{array} \]

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.
In Icelandic, as opposed to English, there is a much wider array of surface positions/configurations which receive a focus interpretation. Consequently, there is a much wider array of expletive constructions. In addition to stage-level predicates, the nominals in the object position of unaccusatives and passives can take part in EFS. The subject of transitive preverbs can also take part in EFS.

The interaction of expletives and participle agreement in Icelandic proceeds without complication, illustrated in (44) and (45).

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

\[ \langle \text{there T be Prt see many people(pl,nom)} \rangle^{\delta^*} \phi^* \phi^* (\delta)^* \phi \]

(45) He expects there to have been seen(acc,pl) many people(acc,pl).

\[ \ldots \text{v}_{\text{EFS}} \text{ expect } \langle \text{there T to-have been Prt see many people} \rangle^{\delta^*} \phi^* \delta^* \phi^* (\delta)^* \phi \]

In some Icelandic sentences, there is both exceptional movement of an object out of the vP in order to avoid a focus interpretation and exceptional non-movement of a vP-internal subject of a transitive preverb in order to force a focus interpretation. Example (46a) is from Jonas (1996:37). She shows that the VP is (46b), with traces indicated by parentheses. The adverb never marks the left edge of the VP and the position of the predicate red shows that the subject has not been extraposed.

(46) a. There painted the house never any students red.
   b. \[ \text{any students (v) (paint) (the house) red} \]_{\text{VP}}

Exceptionally, the house has raised out of the VP in order to avoid a focus interpretation, with any students remaining in the VP (by means of EFS) in order to force a focus interpretation.

Although this explains the main outlines of the relationship of expletive insertion and focus interpretation, there are significant problems concerning optional intermediate positions which we note, but cannot pursue here. One problem is how to account for the variations in (47). The judgments are for the Icelandic equivalents. The one order which is acceptable in English is the one order which is unacceptable in Icelandic.

(47) a. There have many cars been sold at this auction.
   b. There have been sold many cars at this auction.
   c. *There have been many cars sold at this auction.

It is not at all clear how to account for the apparent syntactic optionality with no interpretive reflex. One possibility is that the core structure is (47b) and that a superficial surface transformation shifts the object to the left; obligatorily in English and optionally in Icelandic.15 The different landing site in Icelandic, directly to the right of T, might be the result of the fact that in Icelandic, there is a general resistance to material appearing between auxiliaries and the verb. Even adverbs are disallowed.
In concluding this section, it should be pointed out that in a system in which attract operations are never optional, and in which unvalued features drive attraction, the device of EFS is a minimal solution to the design problem of allowing a nominal to share its $\delta$-feature with T (or $T_r$) without triggering movement. It is a minimal solution in the sense that feature sharing and the entire agreement mechanism are already in place. The change in the syntactic system is minimal.

6.1. 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 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 example in which nominal expletive it raises is given in (48).

(48) It seems to be likely that the earth is flat.

1. \( T_r \) be likely \([\text{that the earth is flat}]\)

2. \( \langle \text{it} \ T_r \) be likely \([\text{that the earth is flat}]\)

3. \( T \) seems \( \langle \text{it} \ T_r \) be likely \([\text{that the earth is flat}]\)

4. \( \langle \text{it} \ T \) seems \( \langle \text{it} \ T_r \) be likely \([\text{that the earth is flat}]\)

7. Quirky Case in Icelandic

If a nominal shares its $\phi$-feature with a TE-head, that head determines the structural case of the nominal. Quirky case constructions in Icelandic, like (49), in which help assigns dative case to its object, do not manifest subject/verb agreement. The verb always appears with 3sg agreement.

(49) Us(dat,1pl) was(3sg) helped.

The most straightforward assumption is that the subject does not share its $\phi$-feature with T.

We 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. Unlike structural case, inherent case is determined by selection. In (49), for example, Icelandic help selects a dative complement. The nonsharability of the $\phi$-features of inherently case marked arguments is akin to a constraint against multiple case-assignment, familiar in many syntactic frameworks. The $\phi$-feature is visible to the syntax, even though it is not sharable. Visibility is important 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. This effect has important consequences for dative experiencer constructions, discussed below.

If we attempt to derive (49), however, it is clear a further assumption must is needed. If nothing else is assumed, a derivation of (49) fails early, as shown in (50). We denote the nonsharable \(\phi\)-feature of an inherently case-marked nominal by \(\phi_0\). The participial Prt in (50) cannot value its \(\phi^*\)-feature.

\[
\begin{align*}
(50) & \text{ Prt help us(dat)} \\
& \quad \phi^* \phi_0
\end{align*}
\]

Icelandic, however, has a mechanism for rescuing (50). 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 two last resort mechanisms for eliminating unvalued \(\phi^*\)-features. An unvalued feature of raising T is simply deleted. For other pivots, 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 nonsharable, just like the \(\phi\)-feature of an inherently case-marked nominal. The continuation of (50) is then:

\[
\begin{align*}
(51) & \text{ Prt help us(dat)} \rightarrow \text{T be Prt help us(dat)} \\
& \quad \phi^* \phi_{dflt} \phi_0 \phi_{dflt} \phi_{dflt} \phi_0
\end{align*}
\]

In (51), we assume that the matrix T does not value its \(\phi^*\)-feature by sharing with the lower \(\phi_{dflt}\). ECM constructions make it clear that \(\phi_{dflt}\) is nonsharable, just like the \(\phi\)-feature of quirky case nominals.

\[
\begin{align*}
(52) & \text{ You believed us(dat,1pl) to have been helped(nom/acc,sg).}
\end{align*}
\]

If defaulted \(\phi\)-features were sharable, we would expect (53a), with Prt in a \(\phi\)-chain headed by \(v_{trs}\). Structural accusative case would be assigned to the participle. Instead, we assume that \(\phi_{dflt}\) is not sharable, so (53b) results. The nominative case which appears on the participle is default nominative case, which we assume is assigned morphologically if no other case is assigned structurally (i.e. via \(\phi\)-agreement with a TE-head) or inherently.\(^{16}\)

\[
\begin{align*}
(53) & \quad \text{a. *You believed us(dat) to have been helped(sg-dflt,acc).} \\
& \quad \quad \text{... } v_{trs} \text{ believe } T_r \text{ to-have been Prt help us(dat)} \\
& \quad \quad \quad \phi^* \phi_{dflt} \phi_0
\end{align*}
\]

\[
\begin{align*}
(53) & \quad \text{b. You believed us(dat) to have been helped(sg-dflt,nom-dflt).} \\
& \quad \quad \text{... } v_{trs} \text{ believe } T_r \text{ to-have been Prt help us(dat)} \\
& \quad \quad \quad \phi^* \phi_{dflt} \phi_0
\end{align*}
\]
Expletive constructions are also possible:

(54) There was(3sg-dflt) helped(sg-dflt,nom-dflt) many people(3pl,dat).

\[ \begin{align*}
\text{there} & \quad \rightarrow \quad \text{be} \\
\delta & \quad \rightarrow \quad \delta^* \\
\phi^* & \rightarrow \phi_d \end{align*} \]

7.1. Dative Experiencers, Exceptional $\phi$-Valuation

The quirky surface subjects discussed in the last section enter the structure as selected verb complements. Icelandic also has surface subjects which enter the structure as selected inherently case marked specifiers of preverbs. The class of so-called dative-nominative verbs have an experiencer preverb $v_{\exp}$ which selects an inherently dative marked subject (the experiencer) and a VP-complement. Sentences like (55) pose no problems for the feature sharing analysis that has been developed.

(55) Okkur líkaði strákarnir.

\[ \begin{align*}
\text{us(dat,1pl)} & \quad \text{like(3sg)} \\
\phi_d & \rightarrow \phi_s \\
\phi & \rightarrow \phi_0 \\
\end{align*} \]

\[ \text{‘we like the boys’} \]

The structure is:

(56) $\begin{align*}
\text{T} & \langle \quad \text{us} \\
\delta^* & \rightarrow \delta \\
\phi^* & \rightarrow \phi_d \\
\phi & \rightarrow \phi_0 \\
\end{align*}$

We suppose that the experiencer $v_{\exp}$ has a $\delta^*$-feature, but differs from $v_{\trs}$ in not having a $\phi^*$-feature. The $\phi$-feature of the quirky subject is visible, but not sharable. Spellout displaces the dative subject to the pre-T position.

Default agreement on T, as in (56), is what the theory developed to this point leads us to expect. In fact, however, although default agreement is preferred by a few speakers and rejected by only a few, the majority of speakers appear to favor agreement if (and only if) the object is 3rd person. See the careful discussion in Sigurðsson (1996). Since the default is 3sg, this effect is only evident if the object is 3pl. In place of (56), most speakers prefer (57).

(57) Mér líkaði strákarnir.

\[ \begin{align*}
\text{me(dat,1sg)} & \quad \text{like(3pl)} \\
\phi_d & \rightarrow \phi_s \\
\phi & \rightarrow \phi_0 \\
\end{align*} \]

\[ \text{‘I like the boys’} \]

The connection between the object and T in (57), passing over/through the dative experiencer, is indicated by a dotted line in the representation below.

(58) $\begin{align*}
\text{T} & \langle \quad \text{me} \\
\delta^* & \rightarrow \delta \\
\phi^* & \rightarrow \phi_d \\
\phi & \rightarrow \phi_0 \\
\end{align*}$

We will maintain the view that the intervening $\phi$-features of the dative experiencer block $\phi$-sharing in (58), so that the object nominal must therefore get default nominative case since it is not in a case chain with T. We assume, however, that Icelandic has an
exceptional \( \phi \)-valuation (\( E\Phi V \)) mechanism which, under some conditions, can be used in place of default valuation as a last resort mechanism for rescuing structures with unvalued \( \phi \)-features which cannot be valued by the ordinary mechanism of feature sharing. The dotted link in (58) is not meant to indicate feature sharing, only the relation between the source and target of valuation. The relation shown by the dotted line cannot be feature sharing because if an infinitival version of (58) is embedded under an ECM verb, it is nominative case which appears on the object, not accusative case from the ECM verb. If the relation were feature sharing, accusative case would be expected since the nominal would enter a \( \phi \)-chain with the accusative case assigning preverb in the higher clause.

We view \( E\Phi V \) as a peripheral repair mechanism which Icelandic has grafted onto the core feature sharing mechanism, with extensive ideologically and/or dialectical restrictions on its range of application. Particularly because of the complexity introduced by the range of variation, space issues preclude a thorough discussion of these restrictions. But a rough sketch is possible.

\( E\Phi V \) applies in expletive variations of (58) as well.

\[
\begin{align*}
(59) & \quad \text{had liked(sg) someone(nom) these socks(nom)}
\end{align*}
\]

\[
\begin{array}{c}
\delta \quad \delta^* \quad (\delta^*)^* \quad \delta^* \quad \delta \quad \delta \\
\phi^* \leftarrow \cdots \phi_{\text{int}} \leftarrow \cdots \phi
\end{array}
\]

\( E\Phi V \) also applies long distance:

\[
(60) \quad \text{Jon seem(pl/sg) be believed(pl) like horses(pl)}
\]

‘Jon seems to be believed to like horses.’

\[
\begin{array}{c}
\delta^* \quad \delta^* \quad \delta^* \quad \delta \quad \delta \quad \delta \\
\phi^* \leftarrow \cdots \phi_{\text{int}} \leftarrow \cdots \phi
\end{array}
\]

Recall that we assume that the \( \phi^* \)-feature of \( T_r \) is removed by deletion if it is not valued by sharing.

\( E\Phi V \) is blocked by multiple dative interveners.

\[
(61) \quad \text{me seem(pl/sg) Jon be believed(pl) like horses(pl)}
\]

‘Jon seems to me to believed to like horses.’

\[
\begin{array}{c}
\delta^* \quad \delta^* \quad \delta^* \quad \delta^* \quad \delta \quad \delta \\
\phi^* \leftarrow \cdots \phi_{\text{int}} \leftarrow \cdots \phi
\end{array}
\]

In all of the examples considered to this point, the source \( \phi \)-features are purely nominal. That is, they are features of a nominal category and no other. This is not always the case. While the results are simple in case the source features are purely nominal, with \( E\Phi V \) applying if there is a single intervener and failing to apply if there are multiple interveners, \( \phi \)-features which are not purely nominal are much less robust sources of \( E\Phi V \). We start with
(62), in which the source $\phi$-feature is not purely nominal, but shared between a nominal and raising Tense.

(62) Manninum virðast/virðist hestarnir vera seinir
the-man(dat) seem(3pl/3sg) the-horses(3pl) be slow
‘It seems to the man that the horses are slow’

\[
\begin{array}{c}
T (\text{the-man } \nu_{\text{exp}} \text{ seem } [ T_r \text{ be horses slow } ] \\
\delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta \\
\phi^* \leftarrow \cdots : \phi_0 \leftarrow \cdots \left\lbrack \text{OPTIONAL} \right\rbrack \cdots \phi^* \overset{\delta}{} \phi
\end{array}
\]

In somewhat more complex examples with source features that are not purely nominal, EΦV is blocked altogether:

(63) Það *virðast/virðist einhverjum manni hestarnir vera seinir.
there seem(*pl.sg) some man(dat) the-horses(nom) be slow
‘It seems to some man that the horses are slow.’

\[
\begin{array}{c}
\langle \text{there } T (\text{some man } \nu_{\text{exp}} \text{ seem } [ T_r \text{ be horses slow } ] \\
\delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta \\
\phi^* \leftarrow \cdots : \phi_0 \leftarrow \cdots \left\lbrack \text{BLOCKED} \right\rbrack \cdots \phi^* \overset{\delta}{} \phi
\end{array}
\]

(64) Hverjum *virðast/virðist hestarnir vera seinir.
who(dat) seem(*pl/sg) the-horses(nom) be slow
‘To whom does it seem that the horses are slow.’

\[
\begin{array}{c}
C T (\text{who } \nu_{\text{exp}} \text{ seem } [ T_r \text{ be horses slow } ] \\
\delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta^* \overset{\delta}{} \delta \\
\phi^* \leftarrow \cdots : \phi_0 \leftarrow \cdots \left\lbrack \text{BLOCKED} \right\rbrack \cdots \phi^* \overset{\delta}{} \phi
\end{array}
\]

The zigzag connection between C and who in (64b) is meant to indicate only that there is some wh-relation, without specifying precisely how wh-chains are constructed.

The intuition we pursue is that EΦV fails in (63) and (64), but not in (62), because the former configurations are more complex in some way than the latter configuration. We make this precise in the following way. Recall that expletive insertion is a last resort mechanism for satisfying an unvalued $\delta$-feature. We interpret this to mean that it follows $\phi$-valuation in (63). At the point of $\phi$-valuation, the relevant structures of (62), (63), and (64), are (respectively):

(65) a. T (the-man . . .
\[
\delta^* \overset{\delta}{} \delta \\
\phi^* \phi_0
\]

b. T (some man . . .
\[
\delta^* \overset{\delta}{} (\delta)^* \\
\phi^* \phi_0
\]

+wh

In (65a), the intervener has its chain requirements satisfied. This is not the case in (65b), whose case chain is not yet well-formed because of the unvalued $\delta$-feature, or in (65c), which has a wh-feature that must enter a well-formed wh-chain. The generalization we draw from this is that EΦV whose source is not purely nominal is optionally possible, but only if the chain requirements of the intervener have already been satisfied.
Examples (60) and (61) are from Schütze (1997), who first noticed the contrast between the (near) impossibility of agreement in (61) and the (near) obligatory agreement in (60). His account is along different lines. Holmberg and Hróarsdóttir (2003) discuss (62–64) and provide an account of the contrast between the absence of agreement in (63) and (64) and the possibility of agreement in (62).17 Bobaljik (2004) pointed out the importance of distinguishing the cases in which the dative intervener is in the same clause as the source φ-features from those in which it is not. In our terms, this is the difference between a purely nominal source of EΦV and one that is not.

7.2. δ-Agreement and Structural Licensing

Examples like (62–64) above have implications for the grammar-internal mechanism which is responsible for the displacement of nominals at the surface, which we will call the EPP. The issue is the surface position of hestanir in the linearized output, to the left of vera. In the theory proposed here, mannínunum shares a δ-feature with “raising T” (Tr). Spellout treats Tense (both finite T and raising Tr) exceptionally, spelling out a nominal with a shared δ-feature in the pre-Tense position unless all such nominals have already been spelled out.18 This, we propose, is the core of the EPP. Crucially, it is δ-agreement, not φ-agreement, that is responsible for leftward displacement.

The following example (from Sigurðsson, 1991) makes the point even more sharply that δ-agreement is the issue, because neither the finite matrix T, nor the embedded raising Tr, nor veexp, are involved in φ-agreement with a nominal.

(66) a. Mundi þér virðast bátunum hafa verið bjargað?
   would(3sg) you(dat) seem the-boats(dat) have been rescued
   ‘Would it seem to you that the boats were rescued?’

   b. C T would (you veexp seem Tr to-have been rescued the-boats

   δ
   φ*φdfilt δ
   φ0
   δ
   δ

   We suggest that whatever particular form of the EPP is adopted, a satisfactory account of the surface location of bátunum in examples like (66) requires the recognition of some form of nominal agreement apart from φ-agreement (i.e. some version of what we call δ-agreement).

It has been known since Sigurðsson (1991) and Freiden and Sprouse (1991) that quirky subjects in Icelandic are subject to some structural licensing condition akin to structural case assignment. Note that in (51), (53b), and (54), although the quirky nominal is isolated from sentential φ-agreement, it participates normally in δ-agreement. In (66), both nominals are involved in δ-agreement, but neither is involved in φ-agreement. Every head which has a δ-feature is required to be in a chain with a TE-head, according to (27). It is this requirement of δ-agreement with a TE-head that is behind “structural licensing.”

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 (67), in Icelandic.

(67) *[them(dat) to have been helped] is likely.

They pointed out that, because the subject of the sentential subject in (67) is inherently case marked, the sentence cannot be ruled out as a violation of a case condition. Considerations
of this kind led to speculation that inherently case marked nominals in Icelandic must be assigned covert structural case in some way.

In order to account for (67), we must establish that \( \text{them} \) cannot share its \( \delta \)-feature with the matrix \( T \), the only available TE-head. If the embedded clause is not headed by a \( C \), we get:

\[
\begin{align*}
T & \text{ be likely } T_r \text{ to have been helped } \text{them} \\
& \delta^{*} \quad \delta^{*} \quad \delta^{*} \quad \delta
\end{align*}
\]

In this case, \( \text{them} \) raises and we get (69), not (67).

\[
(69) \quad \text{Them(dat) is likely to have been helped.}
\]

If, on the other hand, the embedded clause is headed by a \( C \), C-Complement Opacity blocks \( T/\text{them} \delta \)-agreement.\(^{19}\)

8. Selection: A local solution to a global problem

The previous sections have built up a theory of the structural case system. The heart of it is the feature-sharing theory of agreement, the Chain Condition, and the locality which results from intervention and C-Complement Opacity. The Chain Condition is part of IC (Interface Conditions). Feature sharing and locality are part of LNS. In this section we return to the concern of Section 1, growing derivations that do not get trapped in culs-de-sac. Here, we concern ourselves primarily with the syntactic issue, putting aside discussion of an analogous interpretive issue.

A leading idea of the Minimalist Program is that the LNS is an optimal solution to the problem of designing a computational system which produces outputs which are interpretable at the interface. In some of Chomsky’s writings, other factors influencing the design of LNS are suggested. One is that the computational system produce a rich array of structures, allowing a rich array of possible semantic interpretations. Another, of special interest here, is that the computational system be efficient.

There are several examples of how the design of LNS directly reflects properties of the IC. Feature valuation is the most transparent example. Chomsky has suggested that the source of syntactic movement is to be found in demands that the interpretive system makes on the relation of linear order and interpretation. The connection is not direct. The need for displacement, as a design problem, leads to the mechanism of unvalued features, feature valuation, and EPP properties. The mechanism is a solution to the design problem, exploiting the fact that valued features are already part of the system (it is supposed). Since unvalued features would be unintelligible at the output interface, they must be kept internal to the syntactic computational system by eliminating them in the course of the derivation. The design of the LNS directly reflects this. Features are valued cyclically. After a pivot is chosen, all its unvalued features must be valued before the cycle terminates. Note that the interface condition plays no role in the syntactic computation. The LNS never produces states with unvalued features, so the interface condition has no role to play as an output filter.

A second example, much more complex, is the EPP, the grammar-internal mechanisms responsible for nominal displacement to a “surface subject” position. Let us assume that the EPP reflects some kind of interface condition on surface order. Chomsky has suggested that it is a “theme/rheme” requirement of some kind. Assuming this, we ask how this output
condition is reflected in the design of LNS. The issue is subtle, because the LNS itself does not determine the surface position of underlying categories. Nevertheless, the underlying vocabulary and the LNS must furnish the means by which Spellout can position phrases in such a way that the interface condition is satisfied. One crucial property of the underlying vocabulary is that every Tense head has an $\delta^-$-feature. The design of LNS contributes by satisfying the IC by providing Spellout with structures with shared features which can be accessed to determine surface placement. The interface requirement is the root cause for some of the architecture of LNS and Spellout, but the interface requirement plays no direct role in the syntactic computation.

It is important to note that the mechanisms which the LNS adopts to generate outputs which satisfy the needs of the interpretive component are not perfect. These mechanisms are part of the LNS and Spellout in order to force a certain “theme/rheme structure.” But they overshoot the mark. From the standpoint of the interpretive component, is a man in the room, with no EPP displacement to the pre-Tense position, should be perfectly fine. It should be interpreted exactly as there is a man in the room. The EPP mechanisms are overly broad. Nevertheless, they constitute the device adopted by the grammar. The LNS is forced to introduce another mechanism, EFS and expletive insertion, in order to compensate for the effects of the overly broad EPP. This failure of a perfect fit between the interpretive system and LNS is what we have come to expect of the biology of interacting systems.

Our analysis of Burzio’s Generalization is along the same lines. A preverb which is inserted with an unvalued $\delta$-feature will inevitably cause a violation of the Chain Condition if it does not have a subject with a $\delta$-feature. Insertion of such a preverb will produce a doomed state. This is inconsistent with the assumption that the LNS efficiently calculates a wfrt. How are doomed states of this variety avoided? Like the local implementation of the EPP condition on surface order, it is implemented by granting certain properties to certain heads and making the LNS sensitive to those properties. The property is selection, of a subject in this case. The syntax is sensitive to selectional properties of heads and ensures that these properties are satisfied when the head is inserted into the syntax. The syntactic vocabulary is restricted to ensure that preverbs with a $\delta$-feature select a subject.

From this perspective, selection in the syntax is a minimal solution to the design problem of avoiding doomed states.

Because of C-Complement Opacity, complementizers must have selectional properties in order to avoid doomed states. Non-TE complementizers (in English, complementizers other than for) must select a TE Tense. Otherwise, the doomed state that Jack to be happy could be formed. It is doomed because C-Complement opacity prevents Jack from ever sharing its $\delta$-feature with a TE-head, so that a Chain Condition violation at the interface is inevitable. On the other hand, a TE complementizer must select a non-TE Tense head. Otherwise, for Jack is happy could be produced.

Although all of the cases in (70) have been discussed at one point or another, it is useful to review the status of each of these representations.

(70)  
  a. *Jack seems that (Jack) is happy.  
  b. *It seems that Jack to be happy.  
  c. *It seems [Jack to be happy ]IP

C-Complement Opacity implies that (70a) is underivable. Agreement is impossible into the complement of that, forcing it-insertion. Selection makes (70b) underivable, since C does
not select T. Since there is no C-Complement Opacity in (70c), matrix T will agree with Jack, so Jack will raise and there is no possibility of inserting a nominal expletive. Hence, (70c) is underivable.

Note that in each case in (70) the sentence is underivable in the framework we propose. Crucially, it is not the case that the sentence is ruled out by one or another interface condition. The derivations which might yield the sentences in (70) never reach the interface. In each case, a property of the LNS diverts the derivation away from a doomed state along a path that leads to a syntactically well-formed result. The LNS incorporates the IC into its mechanics, which is exactly what we should expect from a system that is designed to efficiently carry out derivations which lead to a well-formed terminal representation.

Finally, it is also important to note that a particular kind of redundancy is inherent in the way that the Minimalist Program is posed. The characteristics of the LNS are taken to follow (in a loose sense) from the properties of the interface, the assumption of efficient computation, and the assumption that the syntax is capable of producing a rich variety of syntactic structures (hence possible meanings). It should not be surprising that properties of LNS directly (or indirectly) reflect properties of IC, introducing a kind of redundancy. That is exactly what we should expect. So, at the cost of allowing inefficient computation, complement selection could be done away with and (70b) rejected at the interface. Similarly, subjectless δ-agreeing preverbs could be admitted into the syntactic vocabulary, and the structures they lead to rejected at the interface (as a Chain Condition violation). This would eliminate some redundancy, but the cost would be inefficiency in the computation.

We have suggested that the apparent redundancy between the properties of LNS and IC is only apparent. The properties of IC should be viewed as design conditions for the mechanisms of LNS. The fact that the properties of LNS mirror the properties of IC is a sign of good design, not an indication of theoretical redundancy. The shape that a theory is forced to assume if redundancy between LNS and IC is eliminated is clear: free derivation and IC filtering. This turns things on its head. In order to get such a system to work, the mechanisms which the LNS utilizes in order to meet interface requirements must be reinterpreted as interface conditions themselves. The EPP, for example, must be looked at as an interface requirement rather than an LNS mechanism for producing outputs which satisfy interface requirements. The properties of the interpretive component which the EPP is devoted to satisfying/exploiting have nothing directly to do with features. Feature valuation is only a means to an end.
Notes

1. Thanks to Eric Reuland and Michael Starke for their organization of the 2001 TILT conference at which this material was presented in its present form. Some of the material in this paper had appeared in an unpublished manuscript titled “Agreement is Feature Sharing” (2000) and some of it is based on Frampton and Gutmann (2002). The analysis of expletives presented at TILT 2001, however, is entirely new. Thanks also to the spirited TILT audience for their comments, as well as to an MIT “LingLunch” audience. Thanks to Cedric Boeckx, Noam Chomsky, Julie Legate, and Charles Yang for comments on various points in the analysis.

2. This may be too strong. It excludes derivational output constraints on operations, which require computation of the output, then (perhaps) one step backtracking if the output does not have the desired property. A weaker position, more easily defensible, therefore less interesting, is to allow one step backtracking: transitions to doomed states are possible, but the it must be immediately detectable that no continuation by transitions allowed by the LNS can lead to a wftr.

3. See López (2002) for a notion of “co-valued features” which has some similarity to the idea of feature sharing.

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

   \[ (A \; B \; C \; D \; \langle E \; F \; G \rangle) \quad \text{(i)} \]

   \[ (A \; \langle [B \; C] \; D \; E \rangle) \quad \text{(ii)} \]

5. The idea that there is a single position for a phrase \( \alpha \) plus a range of other potential Spec[\( X \)] positions indicated only by a relation between \( \alpha \) and various Xs is not unlike the DBP idea of “occurrences.” See Frampton, Gutmann, and Legate (forthcoming) for the analysis of Spellout for these structures.

6. It is possible that one could pursue the idea that quirky case elements fully agree, but for one reason or another the agreement is not overtly displayed. This does not seem to us to be a promising alternative.

7. 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. See also Collins (2002) for a similar view of the cycle.
9. Inherent case is determined by selection. There is no appeal to “case assignment”, feature sharing or case checking.


11. See Frampton, Gutmann, and Legate (forthcoming), which integrates C-Complement Opacity into a theory of Spellout.

12. 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 \( v_1 \) introduces the external argument and \( v_2 \) assigns accusative case, so that structures

\[
[\text{Sub } v_1 \; v_2 \; \text{V Obj}]
\]

would not run afoul of the Chain Condition.

13. Alternatively, the expletive could have already been introduced into the computation as an isolated phrase. It neither selects nor has probing features, so it can be introduced with no other effect than adding a disconnected phrase to the current state of the computation.

14. This recalls Collins’ (1997) idea of using a condition on chain extension to explain (70). Once the nominal moves out of the VP, there is no motivation for EFS, so the chain must extend.

15. Something along the lines of the DISL operation proposed in DBP (2001:30).

16. Icelandic has an array of default mechanisms which are used to support quirky subjects. Default agreement was discussed earlier. These are repair mechanisms which solve the problem of efficient computation in Icelandic. Without them, syntactic computations would enter doomed states.

17. They also point out that (i) is possible.

(i) Hverjum hafa hestarnir virst vera seinir?

who(dat) have(3pl) horses(3pl) seemed to-be slow

The subject of the embedded clause is displaced to the left and, surprisingly, there is plural agreement. In our view, this is not EΦV. The displacement is a spellout phenomenon and the agreement is post-spellout under adjacency. Post-spellout agreement makes sense because spellout, as we conceive it, linearizes morphemes, but does not carry out vocabulary insertion. See Frampton, Gutmann, and Legate (forthcoming) for the analysis.

18. Spellout in a higher position is also possible. Since linearization proceeds top-down (in our view), linearization in a higher position is simply prior linearization.

19. This furnishes an argument that CPs can participate in agreement/movement, but IPs cannot.

20. Preverbs with a δ-feature which do not select a subject are not banned by any principle of the grammar. But they will not persist in the lexicon because they are not usable.
References
