Features in Minimalist Syntax

David Adger and Peter Svenonius
*Queen Mary, University of London**
and *CASTL, University of Tromsø*†

February 5, 2009

1 Introduction

This chapter outlines a number of major issues concerning features in Minimalist syntax. Our purpose is neither to survey the field nor to provide a particular theory of features in minimalism. It is rather to delineate what we take to be the core conceptual issues that the notion of feature raises within minimalist approaches to (transformational) grammar.

We begin by pointing out the different perspectives taken on the notion of feature by Minimalism and the unification-based feature grammar tradition. We then clarify the notions of category and feature, taking category to have essentially a positional definition, while feature is defined as a property of a category that sub-classifies it. With this in hand, we distinguish two kinds of features in minimalist syntax: those which play a role primarily at the interfaces with sounds and meaning, and those whose function is primarily syntax-internal. We then explore the way that features create configurations which can be compositionally interpreted, highlighting their role in dependency formation, in constituent construction and in displacement effects. Finally, we consider the role that interface features play at the interfaces with sound and meaning.

1.1 Some ontological clarifications

An important issue that needs to be clarified when discussing features is their theoretical status, since they are used rather differently in different frameworks in formal linguistics. In certain unification-based frameworks, such as HPSG, features are used as part of a description language for grammatical theory:

> "Intuitively, a feature structure is just an information-bearing object that describes or represents another thing by specifying *values*

*Adger gratefully acknowledges the support of a Leverhulme Major Research Fellowship.
†Svenonius has benefitted from conversations about this subject matter with colleagues at CASTL at the University of Tromsø, in particular Gillian Ramchand and Michal Starke.

for various *attributes* of the described thing; we think of the feature structure as providing partial information about the thing described" Pollard and Sag (1987, 28)

From this perspective, a rich feature theory is a reasonable thing to posit, since the feature theory does not constrain the objects of the theory, but merely describes them (cf. King 1994 for discussion). Some other theory is required to constrain the linguistic objects themselves (for example the type hierarchy in HPSG).

The alternative view assumed in minimalist work is that the features are properties of syntactic atoms and hence are directly objects of the theory: a feature [plural] for example is used analogously to chemists' use of H for the real world thing hydrogen. From this perspective it is crucial to say what the possible feature structures are such that the properties of the features allow them to enter into relationships with other features, analogously to saying what the properties of atoms are such that they can enter into relationships with other atoms.

From this viewpoint, the constraints on the feature theory are substantive and amount to constraining the theory itself, something which is not true when features are seen as a description language. This means that entertaining alternative hypotheses about feature structures is tantamount to entertaining alternative theories. The minimalist framework can be seen as a set of guidelines which constrain the general hypothesis space within which these various theories can be entertained. Of course, it is important to be explicit about what the theories themselves are so as to be able to evaluate their empirical adequacy and their theoretical parsimony.

1.2 Some terminological clarifications

Generative grammar traditionally makes a distinction between the notions of category and feature, a distinction which stems from the immediate constituent analysis advocated by the American Structuralists (Bloomfield 1933, Wells 1947), formally implemented as a phrase structure grammar (PSG) (Chomsky 1957) (cf. Chomsky 1965, 75ff on the inadequacy of immediate constituent systems in dealing with certain cross-classification issues in syntax). The categories used in PSGs represent classes of elements that (i) occur in complementary distribution and (ii) are restricted to certain positions within constructions (e.g. Bloomfield 1926): hence we have the categories of Noun (e.g. *cat*, *sincerity* ...), Complementizer (*that*, *if*, *0* ...), Adjective (*cold*, *final*, *Scottish*), etc. These categories often need to be sub-classified (animate count Noun (*cat*), question complementizer (*if*), gradable Adjective (*cold*), etc.). In the structuralist and generativist tradition, this sub-classification is not correlated with positions in constructions. From this perspective, categories are essentially defined syntagmatically, while sub-categories are paradigmatic. The distinction has its roots in the Aristotelian conceptions of substance and form, which Greek and medieval grammarians used to understand parts of speech (categories) and inflections (featural specifications) (cf. e.g. Lyons 1968).

Following an observation by G. H. Matthews, Chomsky (1965, 79–80) notes that extending PSGs by adding further production rules to capture subcategories (as was done in Chomsky 1957) loses the possibility of capturing generalizations that cross-cut the subcategories. Take, for example, a set of of PS-rules like those in (1) (Chomsky 1965, 80):

(1) N → Proper
N → Common
Proper → Pr-Human
Proper → Pr-non-Human
Common → C-Human
Common → C-non-Human

In such a system, the symbols "Pr-Human" and "C-Human" are unrelated as they are atomic category labels. However, this obviously poses problems when we want to capture generalizations about, for example, "human" nouns. To solve this problem, Chomsky proposes an extension of PSGs which allows the categories to bear distinctive features (an idea imported from work in phonology, Jakobson et al. 1951). The distinctive features used in *Aspects* are mainly binary, but other possibilities are also considered (see, for example the discussion on pp 170ff, and associated notes). We take up the question of what kinds of values features can have in more detail below. An important consequence of the introduction of features is that the extended phrase structure system no longer models the kind of taxonomic theory of linguistic structure defended by the American Structuralists, since the features add an extra cross-classificatory dimension.

In *Aspects*, Chomsky suggests that there may be no category-feature distinction at all, the parts of speech N, V, A, P being simply another set of features alongside Animate, Past, etc. (Chomsky 1965, 207 ff.). He gives lexical entries such as those in (2) in which category labels have no special status.

(2) a. *sincerity*, [+N, −Count, +Abstract]
b. *boy*, [+N, −Count, +Common, +Animate, +Human]

But in practice, the Aristotelian distinction was not abandoned; thoughout the Extended Standard Theory and Government-Binding periods, various kinds of rules were commonly formulated to single out the features N, V, A, and P, for example X-bar theory and government. When COMP and INFL and D were eventually integrated as functional categories, it was by treating them like N, V, A, and P, subject to the same rules of projection and so on, and distinct from features like Past and Wh. In fact, the original arguments for treating COMP and INFL as projections went back to positional arguments in work by den Besten (1981), Platzack (1983), Emonds (1978) and others, in an essentially structuralist sense: the verb appeared in more than one place, therefore an additional category was needed. Similar arguments were raised for other functional categories such as Agr (Pollock 1989) and D (e.g. Taraldsen 1991).

The category/feature distinction is also commonly assumed within Minimalism, although it is little discussed. Despite the elimination of a base component, the syntagmatic ordering of expressions in constituent structure must still be captured somehow, and one might take only a subset of features to be relevant to that ordering, in which case the notion of category feature is still present (e.g. if both C and T can 'have' ϕ features, but it is their C and T features and not their ϕ features which determine their first merge position; cf. Chomsky and Lasnik 1993, 526). In Minimalist grammars, this syntagmatic ordering is typically dealt with via a hierarchy (or sequence) of functional categories (an extended projection, in Grimshaw's 1991 sense; a functional sequence, in Starke's 2004 terms—we will adopt Starke's term in what follows). To the extent that this hierarchy implies a special class of features which are visible to the constituent forming operations of language (e.g. the Edge features of Chomsky 2008), the notion of category re-emerges.

There is one branch of minimalist work, namely cartography (e.g. Cinque 1994, 1999; Rizzi 1997) which actively investigates the hypothesis that the number of features which are 'categorial' in that they project phrase structure is quite large. This work reopens the possibility suggested by Chomsky in *Aspects*, of eliminating the category-feature distinction, a possibility which arises perennially (for example, Kayne 2005a,b discusses the possibility that each syntactic node bears at most one syntactically active feature). However, if every feature projects, then the cross-classification problems originally noted by Matthews and Chomsky must be addressed, something which is yet to be done.

2 Structures of feature systems

The question of how features are structured itself splits into two: (i) what is the internal structure of a feature and (ii) how do features come together to make larger structures? Our purpose in this section is to lay out the range of possibilities for what a theory of features might be under the basic assumption that the features are to be interpreted as primitives of the system. We begin with what we see as the simplest system that can be entertained and then investigate systems of increasing complexity, evaluating them on the basis of various minimalist assumptions.

2.1 Privativity

We begin with what we call PRIVATIVE features. A privative feature is a feature with no properties beyond its distinctiveness from other features; and a privative feature system is one in which all features are privative. Since the feature has no properties, two linguistic structures will be different from each other with respect to a privative feature just in case the feature is present in one and absent in the other. In this system features are simply defined as a list:

(3) An atomic symbol drawn from the set F = {A, B, C, D, E, ...} is a feature

For example, one could represent the difference between *cat* and *cats* as:

(4) *cat* [N]; *cats* [N, plural]

Here we have two privative features, [N] and [plural] and we can construct two different linguistic objects by maintaining the presence of [N] and allowing the presence of [plural] to vary.

Privativity places constraints on the kinds of operations that can be defined on features. If all features in a system are privative, then rules referring to features treat all features alike (except insofar as they may invoke specific features individually). This is perhaps too restrictive, since rules may need to refer to types or classes of features, implying some organization in the feature system.

One way to organize features is in terms of a functional sequence; the functional sequence C–T–*v*–V is sometimes interpreted as a constraint on external merge, and can be understood to encode cooccurrence restrictions from top to bottom (e.g. if C is present then so is T, though not vice-versa).

Another kind of organization of privative features is a 'geometry,' for example, person, number, and gender features can be grouped under a single node for ϕ which is relevant to agreement (e.g. Harley and Ritter 2002, Béjar 2004, drawing on feature theory in phonology, cf. Clements 1985). Such a geometry is normally understood to encode distributional implications from bottom to top, e.g. [speaker] and [addressee] both imply [person], so if [person] dominates them in the geometry, then a rule can be formulated that refers to both by invoking the [person] feature; and if third person is the absence of speaker and addressee features, then the [person] feature covers all persons.[1] If ϕ dominates a node which includes number and gender features but which excludes person features (Harley and Ritter's 2002 INDIVIDUATION), then participial and adjectival agreement can be understood as referring to this node (cf. Chomsky's 2000 use of the notion 'ϕ-completeness').

One theoretical issue here is to what extent the geometry must be stipulated: does such a system require us to posit a syntax-external module to organize the features, or is it possible to derive the properties of the geometry from the syntax of the structure or the semantics of the features (Harbour 2007)?

So it is possible in a privative system to distinguish linguistic objects from each other, but it is difficult to see how linguistic objects can enter into dependency relations in such a system. To see this, consider an agreement relation

[1] In our sense of the term privative, the existence of a sequence or geometry does not make a system non-privative if the only way in which it classes features is in the hierarchy itself. However, if for example the nodes [person], [number] and [gender] have a different status from their dependents [speaker], [plural], and [feminine], then that implies that there are classes, and the system is non-privative.

between a determiner and a noun in a pattern like the following:

(5) a. this mortal coil
b. *this mortal coils
c. these mortal coils
d. *these mortal coil

If the feature [plural] is privative, and has no properties beyond presence or absence, then there is no way within the feature system itself to capture the agreement relation: we need to have a separate rule that stipulates that a plural feature on an N must be copied to D, or match with a plural feature on a D. However, as Adger (in press) notes, there is no obvious way to state this within Minimalism because of Inclusiveness: aside from lexical items, the framework only allows very general syntactic operations, and interface conditions. Maintaining Inclusiveness implies that there must be some syntax-internal feature matching operation. However, any such feature matching operation assumes that features have some property (the property of requiring matching) which means that the system is no longer privative, in the strictest sense.[2] The other general syntactic operations behave similarly: they cannot make reference to classes of features, in a privative system, and so cannot pick out those features that should be involved in selection or movement from those that should not. For these reasons, a purely privative system is inadequate for human language, which displays syntactic dependencies. Either we need to abandon Inclusiveness, or we need to increase the descriptive capacity of the feature theory to something more powerful than a strictly privative system.

One response to this argument might be to take agreement effects to arise because of a movement operation; the formulation of *checking domains* in Chomsky (1993) essentially ensured that all feature-checking occurred under movement or external merge, and Chomsky's (1995, 262) notion of feature movement provides a way of partially unifying agreement and movement. But the unification is only partial: an agreement chain spells out in more than one place, unlike the chain formed by phrasal movement, and an agreement target is never syntactically complex, unlike the landing site for movement, which can be. Thus, even if agreement is modelled as feature movement, we still need to distinguish two circumstances, something which is (apparently) not possible in a strictly privative feature system.

An alternative response might be to deny that agreement is fundamentally syntactic, and to say that interface constraints play a role in enforcing agreement, as in Dowty and Jacobson's (1989) analysis of agreement as a semantic phenomenon. However, there are cases of irreducibly syntactic agreement, for

[2]Sticking to a fully privative system, an account could technically be constructed for this simple example, e.g. by stipulating a rule spelling out D as *these* when a [plural] feature is present in the tree (which in turn spells out as affixal *-s*); this system could be privative as long as no feature other than D agrees, and no feature other than [plural] triggers agreement; but the account would not generalize, for example to subject-verb agreement or to adjectival agreement in other languages, and any such extension would require the creation of a feature class, resulting in a non-privative system.

example the appearance of case and gender agreement morphology on various dependents of a noun phrase (see Svenonius 2007 for further discussion of this issue). These are furthermore not reducible to a morphological component, as they are constrained by syntactic conditions of locality. Hence, at our current level of understanding, mainstream minimalist work is correct in assuming that there is some kind of syntactic agreement, and the feature system of natural language cannot be entirely privative.

2.2 Feature classes

Every feature system so far proposed for natural languages assumes at least tacitly that features are organized in some way, for example if the categories N, V, A, P, C, T, and D are visible to Merge but other features are not (category features), or if N, V, and A assign thematic roles but other features do not (lexical features), or if C, T, and V are organized in one functional hierarchy and P, D, and N in another (different extended projections), or if only the features NOM, ACC, DAT, and GEN satisfy the Case Filter (case features), or if the features PLURAL, PARTICIPANT, and SPEAKER are copied in an agreement process (ϕ features) but certain other features are not.

Any property or rule which applies to the members of a subset of features defines a CLASS of features; such a system is descriptively distinct from, and more powerful than, a system which arranges privative features in an implicational geometry. For example, N, V, A, P, C, T, and D can be members of a feature class CATEGORY; NOM, ACC, DAT and GEN can be members of a feature class CASE; and SPEAKER can be a member of the feature class PERSON.[3] Below we discuss a different conception of attribute-value matrices, one which is even more descriptively powerful in that it allows the notion of 'valuation' to be stated.

A system with classes of features is not privative in our terms, since features have the property of belonging to a class which is itself 'active' in the grammar (that is, there are rules or principles which refer to the class). When, in the literature, a lexical entry is represented something like: *me* [D, ACC, SPEAKER], the notation is silent as to whether the feature system is privative or not; in practice, feature classes are almost always assumed.

2.3 Second order features

In our description of feature classes, in the previous section, we discussed cases in which a given feature either belongs or does not belong to a given class, e.g. ACCUSATIVE is a CASE feature. Any such class can be stated as a property

[3] To see that such a system is descriptively more powerful than a standard rooted directed graph geometry of features, consider that any such graph can be represented as a set of sets of features, where each node in the graph corresponds to the set of all nodes that it (reflexively) dominates; the reverse, however, is not true, e.g. a feature system which consists of the classes {A, B} and {B, C}. To take an example from the literature, in Chomsky (2000), C and v are phase heads, while v and V are θ-assigners.

of features, and is sometimes referred to as a feature (for example, one might state the Case Filter as requiring that a DP must bear a CASE feature). For the sake of clarity we introduce the term SECOND ORDER FEATURE for any property or feature which defines a class of features. In the cases discussed above, a given feature belongs or does not belong to a given class, and the property of belonging to such a class can be called a FIXED SECOND ORDER FEATURE. For example, the category feature T may be said to have the fixed second order feature of EPP in a language, *qua* feature type rather than feature token (we explain below why EPP is a second order feature). There will then be a class of features which have this property. Such a second order feature could be universally fixed, or it could be parameterized; the parametric association of second order with first order features is a common way of conceptualizing parametric variation in syntax.

Work in minimalist syntax also assumes a number of properties that a given feature in a given language may have in some cases or not have in others, that is *qua* feature token, rather than feature type. For example, the category feature T might have the second order feature of uninterpretability or not (see below for more on uninterpretability as a second order feature), with certain tokens of T in linguistic representations being uninterpretable and others not (see, for example, Pesetsky and Torrego's 2001 analysis of nominative case). We will call such properties VARIABLE SECOND ORDER FEATURES.

The most common example of a variable second order feature in linguistic literature outside minimalism is the minus sign, normally interpreted as negation.[4] In such a system, the union of features [A] and [B] is [A, B], and the union of [A] and [−B] is [A, −B], but the union of features [A] and [−A] is impossible or empty (but see Harbour 2007 who takes this combination to be equivalent to uninterpretability). In the absence of negation, such incompatibilities must be ruled out by other factors. In practice, the minus sign is not widely used in minimalist syntax (as opposed to morphology); we discuss the related use of uninterpretability below.

Variable second order features are used in minimalism to capture dependency relations. This idea has multiple incarnations in various instantations of minimalism: strong versus weak features (Chomsky 1993); interpretable versus uninterpretable features (Chomsky 1995); features with the 'EPP' property and features which lack this property (Chomsky 2000); valued versus unvalued features (Chomsky 2001); etc.[5] Each of these properties itself can be construed as a feature associated with another feature.

In this kind of system a first-order feature has some structure, which implements the association of the second order feature with the first (the second order feature is essentially a property of the first order feature); for example, we can define what constitutes a possible feature as follows:

[4] In contrast, the plus sign is often used simply to flag something as a feature, without any assumption being made that there are minus values.

[5] The EPP (*Extended Projection Principle*, Chomsky 1982, 10) was originally thought of as a fixed second order feature of Infl, but was recast as a variable second order feature, in our terms, in Chomsky (2000, 102).

(6) a. An atomic symbol drawn from the set F = {A, B, C, D, E, ...} is a first-order feature
b. An atomic symbol drawn from the set P = {X, Y, Z, ...} is a second order feature
c. Where p∈P and f∈F, then f_p is a feature

For example, if T has the variable second order feature EPP, as represented in T_{EPP}, then EPP is an instruction to the syntax to merge a DP with T (assuming that T's complement is determined by the functional sequence, this DP will have to be a specifier).

Alternatively, if the requirement that a clause have a subject is described as involving a strong D feature on T, then what that means is that the T node consists at least of a bundle of a categorial feature T and another categorial feature D with a variable second order feature, the property of strength: [T, D_{STRONG}, ...]. The property of strength can be interpreted as an instruction to create a dependency relationship with a matching feature and upon achieving this, to delete (allowing the rules for dependency formation ensure that the matching DP must be moved into the checking domain).

A third possibility is that 'having a feature' is a variable second-order property. This is distinct from the notion of feature classes, if e.g. T can have a D feature (T_D) and D can have a T feature (D_T) (cf. Pesetsky and Torrego 2001). The grammar then states how to interpret features which are subordinated to other features (we return below to the notion of valuation).

Although the content of the second order features has had various incarnations in different implementations of minimalism, the core notion of second order feature has remained constant (although largely unrecognized).

There are two possibilities for interpretation of second order features: they have interpretations at the interfaces or they have interpretations via operations within the syntactic system itself. Both possibilities have been considered. For example, Chomsky (1993) proposes that strength should be thought of as uninterpretability at the SM interface, while Chomsky (1995) takes a strong feature to be one that must be checked as soon as possible after it as been Merged. The first hypothesis connects the property to the external systems while the second maintains that it is internal to the syntax proper, a distinction we discuss further in §3.

A further bifurcation in the notion of strength has to do with whether the strong feature can be satisfied by movement into the specifier of the category bearing the feature or to the category itself. For example, Chomsky (1993) proposes that the T head may bear strong features that cause its specifier to be filled by a DP, as well as strong features that cause the T^0 category to be filled with a verb.

The EPP PROPERTY of a feature replaces the notion of strength in Chomsky (2001). It is entirely formal, simply requiring that some syntactic unit be Merged as the specifier of the category whose feature bears this property. It is more general than strength, as it does not specify any properties of the element to be Merged, and hence it is also less restrictive.

Strength and EPP are essentially ways of ordering and localizing the elements that bear matching features, but more fundamental than this is the notion of feature matching itself. That is, what property of a feature entails that it must match another feature? The property of (UN)INTERPRETABILITY (Chomsky 1995, 277 ff.) is used as the driving force behind the establishment of syntactic dependency in many minimalist systems. The idea is that uninterpretability forces feature matching and any uninterpretable feature which has been matched is deleted. Feature matching is constrained by the structure of the derivation: uninterpretable features can be thought of as triggering a search of their sister (and whatever it dominates), the search terminating when a matching feature is found, or when some other barrier to the search is encountered (e.g. a phase boundary).

This interpretable/uninterpretable asymmetry in feature-feature relations is rather natural in a derivational system, since the uninterpretable features are those that drive the derivation, while the interpretable ones are those that are used, in the final representation, to connect with the semantic systems (or the phonological ones). Brody (1995) points out that an alternative 'bare' checking theory is more natural in a representational system: in such a system features are interpreted where they can be, and the interpretation of matching features, if those features are in an appropriate (c-command and locality respecting) syntactic relation (a chain), collapses to the relevant chain-position. Frampton and Gutmann (2000, 2001) develop a (derivational) model of agreement as feature-sharing which has this property as well, as do Adger and Ramchand (2005), in their interface principle *Interpret Once under Agree*. On these models, there are uninterpretable instances of features, but arguably no syntactic features which never have an interpretation (though Chomsky has suggested that case is such a feature, cf. e.g. Chomsky (1995, 278–279); see also Svenonius 2007 for discussion).

If matching and movement always co-occurred, then we could reduce everything to a single property which would simply create a local configuration between two matching features. However, this seems to be empirically incorrect, as we see feature matching (apparently) without overt movement:

(7) a. *There seems to be many men in the garden.
    b. *There seem to be a man in the garden.

Of course, if elements of a movement chain can be phonologically realized in either the lowest or highest chain position (Groat and O'Neil 1996), then one might take movement to always occur when feature matching takes place, but in some cases the moved element is spelled out in situ. For evidence that in fact we need to distinguish the feature matching case from the covert movement case, see Pesetsky (2000). An alternative takes there to be only overt movement, whose impact on linear order is disguised by other movements (Kayne, 1998).

2.4 Valuation

The second order feature of VALUATION is more complicated than that of interpretability, at least as it is usually used, in that a feature is not merely valued or unvalued, but rather it is valued *as something*, so valuation is strictly speaking a function rather than a property. It is not enough in such a system to simply list 'valued' in P of (6-b).

Recall that a simple system of feature classes can be represented in terms of sets: the name of the set is the feature class and rules can generalize over all of its members. That representation does not straightforwardly lead to a notion of 'unvalued' feature.

Chomsky (2001) replaces the notion that uninterpretability drives feature-checking with the idea that unvalued features do so. We can state valuation in the following way: one class of features (the attributes) have the fixed second order feature that they can take another class of features as their values. Formally stated:

(8) a. a feature is an ordered pair ⟨Att, Val⟩ where
 b. Att is drawn from the set of attributes, {A, B, C, D, E, ...}
 c. and Val is drawn from the set of values, {a, b, c, ...}

In our terms, being able to have a value is a fixed second order feature, and having or lacking a value can be formally recognized as a variable second order feature, one which provides an instruction to the syntax.[6]

In this kind of system, the matching property that seems to be so important for capturing syntactic dependencies is built directly into the nature of the second order feature: identity of attribute is the precondition for matching, and the valued and unvalued features unify (specifically, in Adger in press, an unvalued feature is one which has the empty set as its value; in feature-checking, the empty set is replaced by non-empty values).

Once we allow such second order features, we can ask a number of questions about them, with the answers to these questions defining different theories of features. Are there multiple second order features in the system (e.g. both strength and interpretability) or only one? Relatedly, can individual features have multiple second order features or just one? Do only a subset of the features have second order features, or is this option open to all features (that is, are features organized into classes? If there are feature classes, are they TYPED, so that some features have some subset of the second order features and some others have some different subset? Even within this fairly minimal setup, there are many options.

For example, Adger (2003) has a rather rich system, where a feature can be interpretable or uninterpretable, valued or unvalued and weak or strong. The first property is used to establish syntactic dependencies (essentially agreement without movement), the second to capture the particular morphological category

[6]As noted above for second order features more generally, the property of having a feature as a value does not by itself entail feature classes, if e.g. T can have D as a value and vice-versa. However, in practice a system with valuation usually assumes feature classes.

associated with agreement and the third to ensure locality between the two features (that is, to trigger movement). The distinction between the first two properties is maintained because case features are taken to be uninterpretable even when they have a value. Pesetsky and Torrego (2007) also argue that there is a difference between a feature's interpretability and its status as being valued or unvalued.

There are further options within a system that structures features into attributes and values. Can both attributes and values themselves have second order features? Can values be drawn from the set of attributes, allowing recursion into the feature structure, as proposed by Kay (1979) and adopted into HPSG (Pollard and Sag 1994)? Can values be structured syntactic objects, as in GPSG-style SLASH features (Gazdar et al. 1985)? The general minimalist answer, insofar as the question is addressed, would be that this kind of complexity should be handled in the syntax, rather than in the structure of the lexical items. That is, one would like to adopt the *No Complex Values* hypothesis of Adger (in press).

A further question that arises is whether there are other second order features than the ones just discussed. For example, do certain features have the property that they can 'percolate' or be transmitted in ways other than standard projection (Chomsky 2007)? Clearly, the more restricted the options are, the simpler the resulting theory.

3 The interaction of features with syntax

It is not clear how many systems syntax interfaces with, but there are at least two, one concerned with perception and expression (S-M, or Sensory-motor) and the other with meaning; this latter might be a semantic module interfacing with other systems of thought, or it might be the systems of thought directly (the C-I systems, Conceptual-intentional). Following Svenonius (2007), we call features which play a role in both syntactic processes and phonological or semantic interpretation INTERFACE FEATURES; features which play a role only in syntax, we call SYNTAX-INTERNAL FEATURES. In these terms, we can ask whether the features visible to syntax are all interface features, or whether there are any syntax-internal features: that is, features which only play a role in conditioning the application of purely syntactic operations.

Within minimalist approaches to syntax, the syntactic operations are few and very general. We have already discussed the operation of feature matching (usually called Agree). In addition to Agree, there are two other core operations: Merge and Move. The function of Merge is to create larger syntactic units out of smaller ones, Merging two independent elements A and B to form C, which has A and B as immediate constituents; Move does the same thing, except that it draws B from within A. On this definition of Move, it is simply a variant of Merge (so we have External Merge, and Internal Merge, in Chomsky's 2004 terms; see also Starke 2001; Gärtner 2002). We consider these in turn.

3.1 Merge, External and Internal

Observationally, at least three instances of merge can be distinguished: extended projections, complements, and specifiers.[7] First, there is the merge of the extended projection, which follows a functional sequence of categorial features (C over T over *v* over V, for example). This is normally construed in terms of heads merging with complements (Brody 2000a being an exception, since non-morphological dependents in an extended projection are represented as specifiers). If the functional sequence is stated over categorial features, then this does not require additional selectional features to be posited.

Second, there is the merge of a selecting lexical category with an internal argument, for example, an adjective, noun or verb with a subcategorized complement. Since the categories involved vary considerably from lexical item to lexical item (e.g. different verbs may take finite or non-finite complements, DP or PP complements, or no complement at all), this merge falls under the descriptive heading of subcategorization or C-selection.

There are two ways to think about how features are relevant to C-selection. A widespread view is that there is little or no c-selection, and that complementation is determined by non-syntactic factors (see Borer 2005 for such a view). An alternative is to use the technology of features to implement C-selection, for example taking a verb which selects an object to bear a feature which implements this requirement (e.g. Chomsky 1965; Emonds 2000). Crucially, such features will have to be subject to a locality constraint on how they are checked, since C-selection is always local (Baltin 1989). Svenonius (1994) captures this by tying c-selection to head-movement, while Adger (2003) suggests that subcategorization features are always strong, and therefore always require local checking. Hallman (2004) invokes checking domains for selection defined in terms of mutual c-command.

A third instance of Merge is the merge of an argument into a specifier position. The conditions under which this takes place appear to be somewhat different from those governing complements, and hence the features involved may be different. For one thing, the specifier is apparently not merged until after the complement has been merged, judging from the fact that the specifier appears to asymmetrically c-command the complement. Additionally, there do not seem to be C-selectional relations between a verb and its subject (Chomsky 1965); that is, a verb does not subcategorize for the syntactic category of its subject. It is by now commonly assumed that subjects are not be introduced by the verb, but rather by some functional category (e.g. Kratzer 1996, Chomsky 1995); if a distinction between functional and lexical categories is maintained, then this asymmetry can be made to follow from the restriction of subcategorization features to lexical categories.

Internal Merge (Move) only takes place to specifier/head positions, and never to complement position, and so complement style C-selectional features are irrelevant to this case. It is possible that Internal Merge can be unified with the

[7]Setting aside adjunction; Kayne (1994) suggests that adjuncts and specifiers can be unified; see Kidwai (2000), Chomsky (2004) for a different view of adjunction.

third instance of External Merge mentioned above, i.e. Merge of an argument into a specifier.

This is where the second order features of strength/EPP come into play. As discussed above, these impose a requirement that some category containing the lower of the matched features is moved to some position local to the higher of the matched features. However, strength and EPP here behave differently, since an EPP feature is satisfied if any category is merged in the specifier, while strength requires movement of a projection of the matching category. The looser EPP formulation appears to be required for Icelandic *Stylistic Fronting* constructions (Holmberg, 2000), where any category can satisfy the requirement (the closest appropriate category moves). However, most cases of movement seem to target more specific features.

The question of whether internal merge can be unified fully with the introduction of arguments into specifiers is then partly a question of whether there are cases in which a probe (that is, a strong, uninterpretable, or unvalued feature driving movement) is specified as requiring internal versus external merge. In the canonical cases of the EPP, it can be satisfied by either. If natural language agreement is going to be unified with movement, then at least that will have to be done by somehow requiring internal merge. McCloskey (2002) argues in some detail that Irish complementizers distinguish whether a *pro* specifier is internally or externally merged.

An alternative motivation for movement has been proposed by Starke (2004). Starke points out that something must motivate the order of the merge of functional projections, and proposes that that is also what motivates movement to specifiers. The mechanism he suggests is the functional sequence, in effect using the functional sequence to replace the traditional spec-head relation. In Starke's system (see also Brody 2000a) the moved element satisfies whatever constraint would license Merge of an X^0 category in that position. This has the interesting consequence of collapsing the notions of specifier and head, at least for functional heads and their specifiers.

3.2 Agree

In minimalist papers from the nineties, feature-checking was generally held to be possible only in a very local configuration, called a *checking domain* (cf. Chomsky 1993); a distant goal would have to move into a local relationship with a probe in order to check its features. However, since features are assumed to drive movement, it seems that a relation must be established *prior* to the movement taking place; hence in Chomsky (2000, 2001), and subsequent papers, the checking mechanism, called AGREE, is assumed to be able to create a relationship between features at a distance. Phases are part of the theory of locality that constrains the Agree relation (though see Hornstein 2009 for arguments that Agree should be constrained to checking domains and distinguished from movement).

As already stated, Agree as it is usually conceived is a syntax-internal operation, without a specific interpretation at an interface, and so the second order feature of valuation could well be classified as a syntax-internal feature. On the other hand, if Agree is modelled as the matching of pairs of interpretable and uninterpretable features, and if interpretability is understood as interpretability at the interface, then the second order feature of (un)interpretability is an interface feature.

3.3 Licensing

There are various natural language phenomena which are discussed under the rubric of 'licensing,' for example anaphors, polarity items, deletion sites, traces, and various other elements are sometimes said to need licensing or to be licensed by certain configurations or elements. The general assumption underlying most work in the minimalist program is that these phenomena are either to be subsumed under the kinds of featural relations we have discussed here—for example as when Kayne (2002) casts binding theory in terms of movement, or when Zeijlstra (2004) analyzes negative concord in terms of the checking of uninterpretable negative features—or else are to be handled by different modules; the minimalist reluctance to posit additional syntactic modules (such as a binding module or a theta module) means that phenomena which are not reducible to known syntactic mechanisms are handled, outside of the syntax proper, by the semantic component, which is usually thought of as strictly interpretive (cf. Chomsky and Lasnik 1993 on binding as an interpretive procedure; phenomena like polarity licensing are regularly analyzed in semantic terms, cf. e.g. Giannakidou 1997 and references there).

4 Features and the interfaces

In this section, we discuss the connection of the formal features which are visible to syntax to the systems that syntax interfaces with.

4.1 Sensory-Motor

A given feature might consistently spell out in one way or another, for example the plural feature in nouns in English consistently spells out as some allomorph of *-s*, except in listed idiomatic morphological forms (such as *sheep*). This simply reflects the usual Saussurean arbitrary pairing of phonological content with syntactic and/or semantic content, and does not show anything different from the fact that *dog* has a pronunciation as well as a meaning.

In Chomsky (1995), following *Aspects*, this arbitrary pairing is listed in a lexicon that associates phonological and syntactic features. In a 'Lexicalist' model, the lexicon is the input to the computation. Alternatively, there is a distinction between the input symbols and the vocabulary items which replace them, a notion known as 'Late Insertion,' with complex morphological structures being built up syntactically and then associated with phonological forms via a computation which is part of the Spell-out operation (McCawley 1968; Halle and

Marantz 1993). This approach can be extended to non-concatenative meaning-sound linkages in a straightforward fashion. For example, if in a language like Igbo a possessor is marked with a high tone, we can assume that this represents the insertion of a morpheme whose phonological content is autosegmental. In none of these cases do we want to say that a phonology-syntax interface feature is at play.

Sometimes a non-segmental phonological property might be directly associated with a syntactic feature; for example the **L* H H%** contour associated with yes-no questions in English (Pierrehumbert 1980). If the pairing of phonological information and syntactic content is arbitrary, and the intonational contour could in principle have been associated with some other feature, then this can properly be thought of as another example of lexical insertion, no different in principle from the situation with *dog*.

More interestingly, features might represent instructions to the S-M system in ways distinct from the sound-content pairing of the sign, in which case they play a role in both modules and are potentially true interface features. One way in which features might connect to the S-M (Sensory-motor) systems is if they correspond to instructions to the spell-out procedure, for example in specifying that a head be spelled out to the right of its complement (for a head-final structure), or in specifying that a chain be spelled out at its foot (for LF movement). This is one way of characterizing the traditional overt vs covert movement dichotomy, cf. Groat and O'Neil (1996). If the system is so configured that only the highest link in a chain is spelled out, then all movement would be overt, but this seems to be empirically incorrect (at least to the extent that scope phenomena are handled in the syntax, rather than via independent semantic mechanisms, cf. e.g. Huang 1982; Fox 1999). It would seem then that some property of links in a chain is necessary to ensure the parameterization of overt vs covert movement.

Connected to the issue of overtness is the question, in derivational versions of minimalism, of whether operations take place after spellout. Chomsky (1995) suggested that LF movement could be modeled as the movement of features, while overt movement would be the pied-piping of additional material along with the attracted features. The overt versus covert parameter could then be a function either of the probe or of the goal, e.g. the probe could be specified to attract additional features, in the case of overt movement, or the goals could be specified as having features which percolate. Each case would involve a second order feature which could be modeled as syntax-internal.

Pesetsky (2000) argues that phrasal movement at LF must be distinguished from feature movement; if so, then an additional distinction needs to be drawn. Here again the notion of strength has been appealed to, in conjunction with a general principle of chain linearization. For example, Nunes (2004) argues that feature-checking determines the linearization of a chain: one copy checks features, while deletion of the other removes the unchecked features from the derivation.

In a strictly antisymmetric system of linearization with no LF movement, such as has been advocated by Kayne (1994, 1998), there would be no such features, and hence perhaps no features relating syntax to linearization, the linearization algorithms being sensitive only to non-featural information.

Brody (2000a,b) has proposed a 'mirror theory' which replaces head-movement while capturing various mirror-effects observed in morphology. In that theory, a functional sequence spells out as a morphological word at a position designated by a diacritic '*'; for example, French verb movement is represented by T*–*v*–V, while English would be T–*v**–V. In our terms, Brody's * is a (variable) second order feature (parametrically fixed) which is an interface feature. It is present in the syntax (as it is associated with syntactic features like *v*, not with vocabulary items) but simply instructs the phonological component where to spell out a word.

Outside of linearization, there are other possible syntax-phonology interface features. For example, scrambled elements are often phonologically destressed. If a single feature marked an element for movement and destressing, it would be a syntax-phonology interface feature (see, for example, Neeleman and Reinhart 1998). Alternatively, if sentential prosody is assigned by an algorithm sensitive only to structure, as has been suggested for example by Cinque (1993), then such features may not be needed.

It is sometimes assumed that a feature might flag a constituent for not spelling out, as with Merchant's (2001) E feature, which marks ellipsis sites. The alternative is again lexical, for instance to analyze ellipsis as a kind of zero-pronominal (see, for example, Lobeck 1995).

An interesting example of a potential interface feature is that observed in focus by stress. In some languages, including English, a word may be focused by stress, as witnessed in the difference between Rooth's (1985) examples (9-a) and (9-b), where capitals indicate focus stress.

(9) a. I only claimed that CARL like herring.
 b. I only claimed that Carl likes HERRING.

The interesting thing about this is that the focus must be represented in the syntax as well as in the phonology if one adopts the classic Y-model of syntax, where the only connections between semantics and phonology are through the syntax and the lexicon. However, if one denies this architecture, such focus can also been seen as a phonology-semantics feature.

One syntactic approach is a lexical treatment. There would be a functional head, FOC, which carries the semantic content of focusing whatever it is sister to (cf. Rooth 1985), and which would have an autosegmental phonological content. However, this would suggest that increased pitch and intensity is an autosegmental feature which could in principle be associated with any meaning, for example negation or past tense.

If, on the other hand, the focus feature is not a separate lexical item, then it would have to be a feature, present in syntax, which carries an instruction to the phonology (pronounce loud) and to the semantics (place in focus). Possibly, this semantic instruction does correspond to some syntactic operation (e.g. LF movement), but its insensitivity to syntactic islands (Anderson 1972) suggests not.

In sum, some syntactic features might have interpretations at the S-M interface, though there are no uncontroversial examples of this. In the case of linearization, the most likely examples involve different options in the spell-out of chains and of heads. Another class of cases involves non-pronunciation, and another concerns prosodic correspondences with information structure.

4.2 Meaning

At the other end of syntax is meaning, and here there is a lack of consensus regarding exactly how many modules there are and where the boundaries lie between them. In *Aspects*, some semantic features were tentatively posited in order to handle the selectional restrictions which render anomalous such sentences as #*The boy frightened sincerity* (*frighten* was taken to require a [+Animate] object). However, it became apparent that most such constraints are strictly semantic in that they are satisfied under paraphrases and not tied to individual lexical items.[8] The consensus since the early 70's has been that semantic selection is not part of syntax, but rather part of some semantic module or of the conceptual-intentional domain of thought. In general, then, there will be thousands of nominal roots in any language which are featurally indistinct as far as syntax is concerned.

However, the kinds of meanings which distinguish edibles from inedibles and draft animals from mounts are arguably different from those with which formal semantics is concerned, and it is there we find the language of logic, set theory, predicate calculus, and other tools. Notions such as quantification, negation, gradability, boundedness, telicity, plurality, and so on are part of this system. Let us use the term SEMANTICS for such formal representations, excluding vaguer, prototype-based meanings such as whatever distinguishes camels from reindeer or a joke from an insult (and leaving open the question of whether animacy is one or the other).

The question of how syntactic features relate to meaning can then be posed in the following way: how are first order formal features of syntax, and their second order features, related to semantic representations?

Suppose, for example, that we have developed a model in which the only second order features that a syntactic feature can have are those bearing on merge, agree, and spell-out, e.g. merge with something of feature X, or spell-out that something at the bottom of its chain. Then evidence that a given semantically interpretable feature is visible to syntax comes from data that shows that that semantically interpretable feature triggers merge, agree, or spell-out operations.

As an example, contrast a notion like 'dangerous' with a feature like *wh* or negation, which are visible to syntax while also clearly corresponding to semantic interpretations. First, we apparently never find a language in which

[8]Hence you can eat the result of a baking process, but you cannot eat the result of a syntactic transformation, although both objects are headed by the noun *result*; cf. Jackendoff (1972, 18).

dangerous is a syntactic feature, classing syntactic elements which behave in some consistent way together and correlating at the same time with a semantic interpretation of danger. This suggests that dangerous should be relegated to the conceptual-intentional realm along with whatever distinguishes camels from reindeer.[9]

Negation seems to be different from the feature dangerous. To see in what sense negation can be a syntactic feature, consider the rule of 'Neg-inversion' in English: if a negative element appears to the left of the subject, and scopes over the main predication, then the verb moves to second position.[10]

(10) a. I have at no time betrayed this principle.
b. At no time have I betrayed this principle.
c. *At no time I have betrayed this principle.

(11) a. I have never betrayed this principle.
b. Never have I betrayed this principle.
c. *Never I have betrayed this principle.

Thus, sentence negation does not require subject-auxiliary inversion, and subject auxiliary inversion does not require sentence negation (it occurs, for example, in yes-no questions, with *wh*-expressions in interrogatives, with the VP pro-form *so*, and expressions with *only*). However, there is a non-arbitrary connection between sentence level negation and subject-auxiliary inversion, in that any element which induces sentence negation and occurs to the left of the subject requires subject-auxiliary inversion.

The property of attracting a finite auxiliary is a second order feature, which is borne by certain features including one which is interpreted as sentence negation; thus we can say that some feature NEG has a consistent interpretation in the semantics as well as a consistent effect in the syntax. To the extent that other features behave the same way in the syntax, they are due to a fixed second order feature, and this might be parametrically fixed if other languages differ on this point.

Note that semantic interpretation alone does not seem to be enough, since for example *no more than three* and *at most three* are semantically equivalent, but only the one containing a morpheme with the negative feature can trigger

[9]One might imagine that noun classifications could be sensitive to such a feature, even if syntactic operations are not. The Australian language Dyirbal was made famous by George Lakoff's best-seller *Women, Fire, and Dangerous Things* for having a gender class in which the concepts in the title of the book were grouped together. Plaster and Polinsky (2007) argue that this is incorrect, and dangerous things is not the basis for a gender class in Dyirbal.

[10]Famous minimal pairs include the following, which show that if the fronted element is not interpreted as scoping over the main predication, it does not trigger inversion.

(i) a. With no job would Kim be happy.
b. With no job, Kim would be happy.

The first example can only mean that Kim would not be happy with any job, while the second can only mean that Kim would be happy if unemployed (the observation goes back to Klima 1964, but this example is based on a 1974 NELS paper by Mark Liberman; see Haegeman and Zanuttini 1991; Rizzi 1996 for syntactic treatments).

inversion.

(12) a. On no more than three occasions in history has there been a land bridge connecting Asia with North America.
b. *On at most three occasions in history has there been a land bridge connecting Asia with North America.
c. On at most three occasions in history, there has been a land bridge connecting Asia with North America.

This suggests that the expression *no more than* bears a formal feature, call it NEG, which is not present in the expression *at most*, despite the logical equivalence.

On the basis of phenomena like Neg-inversion, we posit a functional head with the following second order features: it attracts an XP with a NEG feature to its specifier, and it attracts a finite auxiliary (or has the * feature, in Brody's terms). Semantically, it has the result that the TP is interpreted as a denial. This does not appear to be a universally present functional head; Cinque (1999) concluded after a cross-linguistic survey that NEG was different from modality, tense, and aspect in not being ordered in a universal sequence.

In this particular case, it might well be that the second order features of this element in English are essentially accidental, a language-specific combination of properties. If so, then it should in principle be possible for a language to have a similar inversion rule for plural, or accusative, or some other syntactic feature.

All we can say is that in English NEG is visible to the syntax, since a syntactic head attracts it, and that it has a consistent semantic interpretation, as all elements with the NEG feature have something in common semantically.

4.3 Universal correlations

As noted, there are widely differing conceptions of where the interface lies between syntax and meaning. In a syntactically austere model, syntax consists only of such very basic operations as Merge and Agree and some variable aspects of linearization. In this case, syntax interfaces with a semantic module, where syntactic outputs are translated into semantic representations and such phenomena as quantification, telicity, scope, and so on are located. This module interfaces with CI.

In that case, the relationship of second order features with particular interface features such as NEG or INTERROGATIVE is accidental, and could at least in principle vary from language to language.

An example which suggests that this is on the right track comes from a comparison between English and certain Norwegian dialects: In English, *wh*-questions require a finite auxiliary in second position, a kind of V2 phenomenon, but topicalization does not require V2. In Norwegian, on the other hand, V2 is obligatory whenever any element is topicalized, but the V2 condition is relaxed, in some dialects, for *wh*-questions, where it is optional (see Vangsnes 2005 and references there).

(13) a. There he has been.
b. Where has he been?

(14) a. Der har han vært.
there has he been
'There he has been'
b. Kor han har vært?
where he has been
'Where has he been?' (Tromsø Norwegian)

This suggests that the association of V-movement features with interrogative C but not with declarative C is 'accidental' in English; in our terms, what looks like a fixed second-order property in English appears to be a variable second-order property cross-linguistically.

Of course, in both languages, topics and *wh*-elements move, and it is argued that similar movement occur covertly in languages like Chinese and Japanese (Huang 1982; Watanabe 1993). Nevertheless, if syntax and semantics are strictly separated, then semantic properties such as being a quantifier should not have syntactic effects. If it turns out that quantifiers undergo movement, then it might be through 'accidental' placement of features, where the wrong arrangements lead to unusable structures.

On this view, being a quantifier in the semantics and taking scope in the syntax are correlated only functionally, in the sense that if an element X is interpreted as a quantifier and yet is not attracted to a scope position by some feature F in the syntax, then a derivation containing X will crash at the semantic interface, because it will be impossible to interpret X in its non-scopal position. Thus, a language will only be able to use X if it also has F.

However, there are many proposals in the literature for a tighter correspondence between syntax and semantics. For example, it is sometimes assumed that T universally carries an EPP feature, and that T maps onto tense semantically. It is quite typically assumed that *v* and V assign thematic roles but that C and T do not, and that C and T dominates *v* and V hierarchically. Similarly, if projections of N need case crosslinguistically while projections of V assign case, and N and V are semantically distinguishable, this is another universal correlation between syntax and semantics.

Work in Cartography, in particular Cinque (1994, 1999), posits a hierarchy of functional categories whose labels reflect their semantic content. In the functional hierarchy, for example, epistemic dominates deontic or root modality, tense dominates aspect, perfect dominates progressive, and so on. If each of the categories in the functional sequence otherwise has cross-linguistically variable second order features relating to merge, agree, and spell-out, then the only property which is universally connected to semantics is the functional hierarchy itself.

5 Conclusion

We have discussed a wide range of different phenomena which bear on the analysis of features in natural languages, and we have outlined what we take to be the different options for a minimalist theory of features. We have suggested a number of distinctions in feature theories: (i) the distinction between categories (which have positional motivation) and features (which have a cross-classificatory motivation); (ii) the distinction between first order and second order features and within this, the distinction between fixed and variable second order features; (iii) the distinction between interface features and module-internal features.

Second order features are motivated by the fact that languages have syntactic dependencies, and within a Minimalist system which embraces Inclusiveness, first order features must have some property which implements this dependency. Variable second order features, that is, properties which can be borne or not by individual instantiations of features in a given language, are the means of this implementation. Current minimalist theory takes the notion of valuedness as a variable second order feature which drives the formation of dependencies, as values are copied to the attribute lacking them.

Since fixed second order features have to do with the creation of dependencies for such relations as Merge, Move, and Agree, much of parametric variation can be thought of as residing in which first order features are associated with which fixed second order features in a given language (i.e which dependencies are syntactically active in a language).

Given these distinctions, various theoretical questions arise: can all featural distinctions be reduced to positional distinctions? How many second order features are there and what is their correct characterization? Can all types of first order features be associated with second order features (or, for example, is this restricted to, say, category features)?

A second major issue is the interaction of the notions of first and second order features with the notion of interface vs. syntax-internal features. Brody (1997, 143) posits the hypothesis of "Radical Interpretability," that syntax never makes use of elements which lack interface properties (either semantic/conceptual content or instructions concerning Spell-Out). The strongest version of this hypothesis would hold that all first order features have semantic interpretation (such as tense or negation or entity or quantity), and all second order features are interface properties (such as uninterpretability, or licensing overt Spell-Out).

Pesetsky and Torrego (2001, 364) discuss a slightly weaker notion under the rubric of "relativized extreme functionalism," namely that all grammatical features which are first order features, in our terms, have a semantic value (though some instantiations of such features may be uninterpreted).

Given the distinction we draw in this paper, one might take first order features to be interface features in the sense of having some consistent meaning representation, while the second order features would be only syntax-internal. Such a view would, we think, be compatible with relativized extreme functionalism, and would lead to the interface between syntax and the interpretive component being extremely minimal. The parametric data discovered in the cartographic tradition could be handled by allowing functional categories to have a simple format, associating certain second order properties with the first order features which are organized hierarchically by the functional sequence.

It is common in minimalist literature to assume many of the distinctions discussed above, without due care in questioning the addition of new first order or second order features into the system. It seems to us that a concentration on what might constitute a more minimalist theory of features is necessary, and we hope that by identifying and clarifying the issues here, we have brought such a theory one step closer.

References

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford University Press, Oxford.

Adger, David. in press. A minimalist theory of feature structure. In *Features: Perspectives on a Key Notion in Linguistics*, edited by A. Kibort and G. Corbett. Oxford University Press, Oxford.

Adger, David and Gillian Ramchand. 2005. Merge and move: *Wh*-dependencies revisited. *Linguistic Inquiry* 36 2: 161–193.

Anderson, Stephen R. 1972. How to get *even*. *Language* 48 4: 893–906.

Baltin, Mark R. 1989. Heads and projections. In *Alternative Conceptions of Phrase Structure*, edited by Mark R. Baltin and Anthony S. Kroch, pp. 1–16. University of Chicago Press, Chicago.

Béjar, Susana. 2004. *Phi Syntax*. Ph.D. thesis, University of Toronto, Toronto.

den Besten, Hans. 1981. On the interaction of root transformations and lexical deletive rules. *Groninger Arbeiten zur Germanistischen Linguistik* 20: 1–78.

Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2: 153–164.

Bloomfield, Leonard. 1933. *Language*. Holt Rinehart and Winston, New York.

Borer, Hagit. 2005. *The Normal Course of Events (Structuring Sense, vol. II)*. Oxford University Press, Oxford.

Brody, Michael. 1995. Hungarian focus and bare checking theory. In *Proceedings of the Göttingen Focus Workshop, 17. DGfS*, edited by I. Kohlhof, S. Winkler, and H.B. Drubig, vol. 69 of *Arbeitspapiere des Sonderforschungsbereichs 340*, pp. 197–210. University of Tübingen, Tübingen.

Brody, Michael. 1997. Perfect chains. In *Elements of Grammar*, edited by Liliane Haegeman, pp. 139–166. Kluwer, Dordrecht.

Brody, Michael. 2000a. Mirror theory: Syntactic representation in perfect syntax. *Linguistic Inquiry* 31 1: 29–56.

Brody, Michael. 2000b. Word order, restructuring, and mirror theory. In *The Derivation of VO and OV*, edited by Peter Svenonius, pp. 27–43. John Benjamins, Amsterdam.

Chomsky, Noam. 1957. *Syntactic Structures*. Mouton, 's-Gravenhage.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Ma.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge, Ma.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 1–52. MIT Press, Cambridge, Ma.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge, Ma.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step: Minimalist Essays in Honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, pp. 89–155. MIT Press, Cambridge, Ma.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, pp. 1–52. MIT Press, Cambridge, Ma.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and Beyond: The Cartography of Syntactic Structures, vol. 3*, edited by Adriana Belletti, pp. 104–131. Oxford University Press, New York.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, edited by Uli Sauerland and Hans-Martin Gärtner, pp. 1–18. Mouton de Gruyter, Berlin.

Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos Otero, and Maria-Luisa Zubizarreta. MIT Press, Cambridge, Ma.

Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An International Handbook of Contemporary Research*, edited by J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann, pp. 506–569. Mouton, Berlin.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24 2: 239–297.

Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne*, edited by Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi, and Raffaella Zanuttini, pp. 85–110. Georgetown University Press, Washington, DC.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2: 225–252.

Dowty, David and Pauline Jacobson. 1989. Agreement as a semantic phenomenon. In *Proceedings of the 5th Annual Eastern States Conference on Linguistics (ESCOL)*, pp. 95–108. Ohio State University, Columbus, Oh.

Emonds, Joseph. 1978. The verbal complex V-V′ in French. *Linguistic Inquiry* 9: 151–175.

Emonds, Joseph. 2000. *Lexicon and Grammar: The English Syntacticon*. Mouton, Berlin.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30 2: 157–96.

Frampton, John and Sam Gutmann. 2000. Agreement is feature sharing. Ms. Northeastern University.

Frampton, John and Sam Gutmann. 2001. How sentences grow in the mind: Efficient computation in minimalist syntax. Ms. Northeastern University.

Gärtner, Hans-Martin. 2002. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Akademie Verlag, Berlin.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, Ma.

Giannakidou, Anastasia. 1997. *The Landscape of Polarity Items*. Ph.D. thesis, Rijksuniversiteit Groningen.

Grimshaw, Jane. 1991. Extended projections. Ms. Brandeis University.

Groat, Erich M. and John O'Neil. 1996. Spell-out at the LF interface. In *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, edited by Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, pp. 113–139. Benjamins, Amsterdam.

Haegeman, Liliane and Raffaella Zanuttini. 1991. Negative heads and the neg criterion. *The Linguistic Review* 8 2-4: 233–251.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 111–176. MIT Press, Cambridge, Ma.

Hallman, Peter. 2004. Symmetry in structure building. *Syntax* 7 1: 79–100.

Harbour, Daniel. 2007. *Morphosemantic Number*. Springer, Dordrecht.

Harley, Heidi and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78 3: 482–526.

Holmberg, Anders. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31 3: 445–483.

Hornstein, Norbert. 2009. *A Theory of Syntax: Basic Operations and UG*. Cambridge University Press, Cambridge.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. thesis, MIT, Cambridge, Ma.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. No. 2 in Current Studies in Linguistics. MIT Press, Cambridge, Ma.

Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. 1951. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. MIT Press, Cambridge, Ma.

Kay, Martin. 1979. Functional Grammar. In *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society*, edited by C. Charello, pp. 142–158. University of California, Berkeley, Ca.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, Ma.

Kayne, Richard S. 1998. Overt vs. covert movement. *Syntax* 1 2: 128–191.

Kayne, Richard S. 2002. Pronouns and their antecedents. In *Derivation and Explanation in the Minimalist Program*, edited by Samuel David Epstein and Daniel T. Seely, pp. 133–166. Blackwell, Oxford.

Kayne, Richard S. 2005a. On the syntax of quantity in English. In *Movement and Silence*, edited by Richard S. Kayne. Oxford University Press, New York.

Kayne, Richard S. 2005b. Some notes on comparative syntax, with special reference to English and French. In *The Oxford Handbook of Comparative Syntax*, edited by Guglielmo Cinque and Richard S. Kayne, pp. 3–69. Oxford, New York.

Kidwai, Ayesha. 2000. *XP-Adjunction in Universal Grammar*. Oxford University Press, Oxford.

King, Paul John. 1994. An expanded logical formalism for head-driven phrase structure grammar. Arbeitspapiere des sfb 340, University of Tübingen.

Klima, Edward S. 1964. Negation in English. In *The Structure of Language: Readings in the Philosophy of Language*, edited by Jerry A. Fodor and Jerrold J. Katz, pp. 246–323. Prentice Hall, Englewood Cliffs, NJ.

Kratzer, Angelika. 1996. Severing the external argument from the verb. In *Phrase Structure and the Lexicon*, edited by Johann Rooryck and Laurie Zaring, pp. 109–137. Kluwer, Dordrecht.

Lobeck, Anne. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford University Press, New York.

Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge University Press, Cambridge.

McCawley, James D. 1968. Lexical insertion in a grammar without deep structure. In *Papers from the fourth regional meeting of the Chicago Linguistic Society*, edited by B. J. Darden, C.-J. N. Bailey, and A. Davidson, pp. 71–80. University of Chicago, Chicago.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and Explanation in the Minimalist Program*, edited by Samuel David Epstein and Daniel T. Seely, pp. 184–226. Blackwell, Oxford.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press, Oxford.

Neeleman, Ad and Tanya Reinhart. 1998. Scrambling and the PF Interface. In *The Projection of Arguments: Lexical and Compositional Factors*, pp. 309–353. CSLI, Stanford, Ca.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. MIT Press, Cambridge, Ma.

Pesetsky, David. 2000. *Phrasal movement and its Kin*. Linguistic inquiry monographs; 37. MIT Press, Cambridge, Ma.

Pesetsky, David and Esther Torrego. 2001. T to C movement: Causes and consequences. In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, pp. 355–426. MIT Press, MIT, Cambridge, Ma.

Pesetsky, David and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, edited by Simin Karimi, Vida Samiian, and Wendy K. Wilkins, pp. 262–294. John Benjamins, Amsterdam.

Pierrehumbert, Janet B. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. thesis, MIT, Cambridge, Ma.

Plaster, Keith and Maria Polinsky. 2007. Women are not dangerous things: Gender and categorization. In *Harvard Working Papers in Linguistics*, vol. 12. Harvard University, Cambridge, Ma.

Platzack, Christer. 1983. Germanic word order and the COMP/INFL parameter. *Working Papers in Scandinavian Syntax* 2.

Pollard, Carl and Ivan A. Sag. 1987. *Information-based Syntax and Semantics – vol. 1, Fundamentals*. CSLI, Stanford, Ca.

Pollard, Carl and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. CSLI and University of Chicago Press, Stanford and Chicago.

Pollock, Jean Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20 3: 365–424.

Rizzi, Luigi. 1996. Residual verb second and the Wh-criterion. In *Parameters and Functional Heads: Essays in Comparative Syntax*, edited by Adriana Belletti and Luigi Rizzi, pp. 63–90. Oxford University Press, New York.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar: Handbook in Generative Syntax*, edited by Liliane Haegeman, pp. 281–337. Kluwer, Dordrecht.

Rooth, Mats. 1985. *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst, Ma.

Starke, Michal. 2001. *Move Reduces to Merge: A Theory of Locality*. Ph.D. thesis, University of Geneva. Available at http://ling.auf.net/lingBuzz/000002.

Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *Structures and Beyond: The Cartography of Syntactic Structures, vol. 3*, edited by Adriana Belletti, pp. 252–268. Oxford University Press, New York.

Svenonius, Peter. 1994. C-selection as feature-checking. *Studia Linguistica* 48 2: 133–155.

Svenonius, Peter. 2007. Interpreting uninterpretable features. *Linguistic Analysis* 33 3-4: 375–413.

Taraldsen, Knut Tarald. 1991. Two arguments for functional heads. *Lingua* 84 2-3: 85–108.

Vangsnes, Øystein Alexander. 2005. Microparameters for Norwegian *wh*-grammars. *Linguistic Variation Yearbook* 5: 187–226.

Watanabe, Akira. 1993. *Agr-based Case Theory and its Interaction with the A-bar System*. Ph.D. thesis, MIT, Cambridge, Ma.

Wells, R. 1947. Immediate constituent analysis. *Language* 23: 81–117.

Zeijlstra, Hedde. 2004. *Sentential Negation and Negative Concord*. Ph.D. thesis, University of Amsterdam, Utrecht. Published by LOT.