Syntactic Features

Peter Svenonius*

Completed in 2017; published in April 2019
in *Oxford Research Encyclopedia of Linguistics*

Summary

Syntactic features are formal properties of syntactic objects which determine how they behave with respect to syntactic constraints and operations (such as selection, licensing, agreement, and movement). Syntactic features can be contrasted with properties which are purely phonological, morphological or semantic, but many features are relevant both to syntax and morphology, or to syntax and semantics, or to all three components.

The formal theory of syntactic features builds on the theory of phonological features, and normally takes morphosyntactic features (those expressed in morphology) to be the central case, with other, possibly more abstract features being modeled on the morphosyntactic ones.

Many aspects of the formal nature of syntactic features are currently unresolved. Some traditions (such as HPSG) make use of rich feature structures as an analytic tool, while others (such as Minimalism) pursue simplicity in feature structures in the interest of descriptive restrictiveness. Nevertheless, features are essential to all explicit analyses.

*CASTL, University of Tromsø – The Arctic University of Norway. Thanks to Antonio Fábregas, Martin Krämer, Tarald Taraldsen, Gillian Ramchand, Terje Lohndal, David Adger, Thomas Graf, Daniel Harbour, and two anonymous reviewers, for useful feedback on earlier stages of this work.

Contents

1 Traditions of features in linguistic description

The theory of features in syntax derives from earlier work in morphology and phonology, especially the foundational work in the first half of the twentieth century by Jakobson and Trubetzkoy and others (e.g., Jakobson 1990 [1942], Jakobson et al. 1951; see Clements and Hume 1995, Halle et al. 2000 for discussion).

In phonology the term *feature* is normally used in the restrictive sense of "distinctive features" of phonemes—all and only the properties that are necessary to uniquely distinguish each item in the phoneme inventory of a language. These properties are normally assumed to be phonetically (or articulatorily) grounded. They define natural classes. For example a distinctive feature distinguishes /p/ from /b/ in English, as demonstrated by the existence of minimal pairs such as *pray* and *bray*. In contrast, no distinctive

feature distinguishes the aspirated /p/ ([p^{h}]) in *pat* from the unaspirated one in *spat* or *ape*.

Suprasegmental properties such as stress and tone are not definitional of phonemes—even in a language with lexical stress, the /a/ in the stressed syllable in a word like Russian *muká* 'torture' is not a distinct phoneme from the /a/ in the unstressed syllable in *múka* 'flour' (Jakobson 1990 [1942]); and similarly for tone. On the whole, they behave separately from the segmental properties embodied by the system of distinctive features for phonemes. As a result, stress and tone contrasts are not normally treated as part of the same feature system as the one defining the phonemes, and indeed they show distinct characteristics from those properties which are (see Hyman 2010 and Clements et al. 2010 for discussion of the issue of whether tone is a feature, and whether tones are differentiated by features).

Similarly, in syntax it is common and reasonable to restrict the term 'feature' to properties which share certain characteristics. By analogy to distinctive features being the minimal necessary distinguishing properties of phonemes, we might take distinctive features in syntax to be the minimal necessary distinguishing properties of syntactic heads. This would entail that properties of phrases such as indices, bar-level, and c-command domains are not features (though note that many linguistic traditions do not recognize this restriction; phrasal features are discussed further in §2.5).

If features are always properties of heads, then this has the attractive consequence that all features might be contained within the 'lexicon' or inventory of syntactic building-blocks, and that none are introduced into syntactic structures by any other route than the combination of those building blocks, though features might be copied or spread by syntax. In that case, when a feature appears to be a property of a phrase, it must be inherited from some head. For example, in (1), nominative case appears to be a property of the whole noun phrase, as it appears on every part of the noun phrase, but it could have spread throughout the noun phrase from a single head.

(1) nga-la-wan-da ngarrku-wa kang-ka
1-PL-POSS-NOM *strong*-NOM *language*-NOM
'our strong language' (Kayardild (Australian), Evans 1995:235)

Some head could bear the nominative feature as an 'inherent' or lexical property.[1]

[1]This could be true in a number of different ways. For example, if nouns combine

To take an example of a syntactic property that would normally not be analyzed as a feature, consider indices. An index is relevant to the syntax, for example a pronoun cannot be coindexed with (bound by) another noun phrase within the same clause, so that in *Kim saw him*, the object must not be bound by the subject.

But an index is not a feature, by the same logic that stress is not a feature in phonology; indices do not constitute minimal distinguishing characteristics of elements in the inventory of basic feature-bearing units (lexical items or heads, for syntax; phonemes or segments, for phonology). In support of this division, it appears that indices do not have the same characteristics as those features which do reside on heads in the lexicon, for example features can often be uninterpretable in a way that indices cannot (see §2.3 on (un)interpretability; see Kuhn 2016 and Schlenker 2016 for a discussion of whether American Sign Language 'loci' (locations of signs) are indices or features).

This essay consists of three main parts. Section one is an overview of some basic traditional concerns. Section two lays out formal properties of features. Section three discusses the taxonomic use of features to identify all and only the syntactic categories. There is also a section four in which some remaining open issues are identified.

1.1 Modularity

Since this essay is about syntactic features, it is appropriate to take a moment to reflect on what it means for something to be syntactic. As a matter of convenience, we can call features syntactic if they have some relevance to syntax, whether or not they also are relevant to morphology or to semantics. What are excluded, then, are purely morphological features and purely semantic ones (the term FORMAL FEATURES is often used in the syntactic literature to refer to syntactic features in this sense, e.g., Chomsky 1995).

An example of a purely morphological feature is declension class. Declension class is relevant to allomorph selection but not to syntax. Icelandic can be used to illustrate this; in Icelandic, some nouns belong to 'strong' declensions, which determines the endings they show for case and number. In the

with case values lexically and presyntactically; or if there are abstract case heads in noun phrases; or if nominative case comes originally from a finite head in the clause. On the other hand, it could also be false: nominative case might be assigned by a syntactic rule, and not be a property of any head in the lexicon.

strong declensions, the genitive singular is almost always different from the accusative and dative singulars (e.g., *barn* 'child NOM/ACC SG,' *barni* 'child DAT SG,' *barns* 'child GEN SG; *bók* 'book NOM/ACC/DAT SG,' *bókar* 'book GEN SG'). But in the weak declensions, the genitive singular is always identical to the accusative and dative singulars (e.g., *hana* 'rooster ACC/DAT/GEN SG,' *peru* 'pear ACC/DAT/GEN SG'). Thus, there are morphological generalizations concerning the declension classes which motivate the positing of features, but these features have no syntactic significance. The form of a determiner, for example, is sensitive to the gender of a noun it combines with but not to its declension class.

A purely semantic feature is one which is relevant to semantic interpretation but with no bearing on syntax. A candidate for such a feature in English might be JUVENILE, which would be a feature of the concepts expressed by *calf, kitten, puppy, lamb, piglet, foal, fawn, gosling, cub*, and so on, distinguishing them from such words as *cow, cat, dog, sheep, pig, horse, deer, goose*, and *bear*, but with no syntactic significance.

In contrast to morphology, it is not clear how much a featural analysis contributes to an understanding of semantic concepts, when those semantic concepts are not invoked by some grammatical rule. For the purposes of this article any question about how to analyze nonsyntactic aspects of semantics can be set aside. More vexing is the question of how to decide when a constraint is nonsyntactic, to which this article returns in §1.3 and §1.4.

1.2 Morphosyntactic features

The traditional understanding of syntactic features starts from a consideration of morphosyntactic features, features which have morphological expression and syntactically influenced distribution, such as agreement and case (see Adger and Harbour 2008 for a useful and thorough overview of the 'ϕ' features which are central to agreement).

In agreement phenomena, a TARGET of agreement, for example a verb or modifier, shows morphological marking corresponding to certain properties of a CONTROLLER, for example a noun phrase. This is plain to see in a language like Chichewa (a Bantu language spoken in Malawi and neighboring countries), where overt noun class marking appears both on the controller and the target (example from Mchombo 1998:503).

(2) M-kângo s-ú-ná-ngo-wá-phwány-á ma-ûngu
3-lion NEG-*3SM*-PAST-*just-6OM-smash*-FV *6-pumpkin*
'The lion did not just smash them, the pumpkins'

The noun meaning 'lion' belongs to class 3, as expressed by the prefix *m-* on the noun (glossed simply 3), and controls subject agreement on the verb in the form of a class 3 subject marker *u-* (glossed 3SM). Had the subject been first person singular, the subject marker on the verb would have been *ndi-*.

The noun meaning 'pumpkin' (*dzungu* in the singular) is class 5, but is used here in the plural, which is class 6, hence prefixed with *ma-*; the corresponding object marker on the verb is the class 6 object marker *wá-* (glossed 6OM). Had the object been first person singular, *ndi-* would have occupied the position immediately before the verb stem instead.[2]

The properties controlling agreement need not be overtly expressed on the controller, as in the North Sámi example in (3), where dual agreement on the verb is obligatory when the subject refers to two individuated people.

(3) Mu vánhemat bođiiga.
*my parent.*PL *came.*DU
'My parents came' (North Sámi (Finno-Ugric), Nielsen 1926:302)

In North Sámi, nouns show a singular-plural distinction (cf. *vánhen* 'parent' ∼ *vánhemat* 'parents'), but no dual (though pronouns have a dual form: *mu* 'my,' *munno* 'our DU (belonging to us two),' *min* 'our PL').[3]

The Chichewa and North Sámi examples, despite their differences, are both typical of agreement phenomena in that noun phrase arguments control agreement on predicates in features such as person, number, and gender, the so-called ϕ-features. Number and gender agreement may also appear on modifiers of nouns, along with case. The 'noun classes' of Bantu languages are a conflation of number with gender in the technical sense, because they control agreement of this kind (Corbett 1991). Gender is a syntactic feature (a morphosyntactic one), because it is copied or shared across syntactic structures by agreement, subject to syntactic constraints (see also §1.4 for

[2]In the gloss, FV stands for 'final vowel.'

[3]Toivonen (2007) discusses the similar pattern in Inari Saami. In cases like this, the issue arises of whether the features expressed on the target are abstractly present in the featural make-up of the controller; if not, then agreement cannot consist only in the simple copying of features from controller to target.

more discussion of the syntactic nature of gender).[4]

The features of the controller are logically prior to those on the target (Zwicky (1986) calls them 'inherent'), in the sense that they do not depend on the target for their presence, while the features on the target are derived through agreement (Zwicky (1986) calls them 'imposed'; Booij (1996) calls them 'contextual'). The relationship between features of the target and controller forms the basis for models in which features drive syntax, such as the Minimalist program (see for example Stabler (1997) and Collins and Stabler (2016) for formalizations).

1.3 Morphosemantic features

The term 'morphosemantic' feature is sometimes used for features which are morphologically expressed and have a consistent connection to meaning but have no syntactic impact (e.g. Corbett 2012:49). Corbett illustrates this with the Maltese dual; in contrast to North Sámi, some nouns in Maltese have a morphologically distinct form with dual semantics, but again in contrast to North Sámi, there are no distinctly dual agreeing forms of predicates or modifiers, hence dual in Maltese appears not to have syntactic significance. Another example might be the diminutive *-let* in English *piglet, eaglet, booklet, cutlet, playlet, ringlet*, etc. It is not clear whether semantic content of this kind warrants the positing of features at all; the Maltese dual and the English diminutive might simply be bound morphemes with non-featurally organized semantic content.

Consider in this context whether the opposition between present and past tense in a language like English is morphosyntactic or morphosemantic. Certainly the syntactic distribution of finite verb forms is different from that of nonfinite verb forms, so finiteness is a morphosyntactic feature in English, but the question of whether present and past are visible to the syntax is subtler.

In general, the syntactic distribution of present and past verb forms is the same, if we assume, as is standard, that the acceptability difference between examples like (4a) and (4b) is semantic and not syntactic (hence the crosshatch, rather than a star).

[4]Bobaljik (2008) argues that case and agreement features are inserted late in a derivation, and calls them postsyntactic; but the distribution of case and agreement is controlled by syntactic categories, configurations, and principles, so they are morphosyntactic and distinct from purely morphological phenomena such as declension class.

(4) a. I leave tomorrow.
b. #I left tomorrow.

However, the phenomenon of sequence of tense (Ross 1967) suggests that the past tense is a syntactic feature in English, because the distribution of morphologically past forms is not constant across contexts, but is sensitive to syntactic structure. Consider the sentence in (5).

(5) I thought I left tomorrow (but it turns out that my flight is today).

Here, the interpretation of the tense in the embedded clause is not formally past, that is, it does not refer to departures prior to now; instead, the formal past tense expressed on the verb is imposed by the presence of a past tense in the matrix clause (for more discussion see Kratzer 1998, Stowell 2007, and references there). We could say that the past tense is not *interpreted* in the embedded clause (see §2.3 on interpretability of features). This is not motivated by the semantics of past tense alone (and not all languages with past tense show the sequence of tense phenomenon, Enç 1987).

Thus it seems that there is some evidence that past tense in English is a morphosyntactic feature, not just a morphosemantic distinction.[5]

1.4 Gender

In this section I use gender to illustrate some of the major issues raised above. Hockett (1958) defines gender as a system of noun classes which control agreement on associated words. Most grammatical gender systems have at their core an animate–inanimate distinction and/or a male-female distinction, and gender systems often reflect conceptual categories such as animals and plants (Corbett 1991).

An example of a language lacking gender on nouns is North Sámi. The grammar of North Sámi is sensitive to the distinction between animates and inanimates, for example the dual agreement illustrated in (3) is restricted to animate referents. But the sensitivity is not lexical; a noun which can refer to animates or inanimates controls agreement according to the properties of the referent, not of the noun.

[5]The conclusion is not incontrovertible, and some researchers continue to analyze sequence of tense as a semantic phenomenon. See §4 for some discussion of the trade-off between locating grammatical phenomena in the syntax and relegating them to an independent semantic component.

There is no evidence anywhere in the grammar of North Sámi of sensitivity to biological gender. There are lexical words for gender-specific concepts, such as *bárdni* 'boy' and *nieida* 'girl,' but there are no gender-based forms of function words, for example demonstratives agree with the noun in case, but have no gender-specific forms, as illustrated in (6).

(6) a. dat bárdni, dan bártni
that.NOM *boy*.NOM *that*.ACC *boy*.ACC
'that boy (NOM),' 'that boy (ACC)'
b. dat nieida, dan nieidda
that.NOM *girl*.NOM *that*.ACC *girl*.ACC
'that girl (NOM),' 'that girl (ACC)'

Nor are there any gender-specific pronouns, as illustrated in (7).

(7) Son jearai nuppádaššii sus.
s/he.LOC *asked second.time him/her*.LOC
'She/he asked him/her a second time'

Thus, though animacy matters for certain syntactic phenomena, there are no noun classes determining agreement phenomena, and therefore there is no gender in the technical sense. In this way North Sámi contrasts with Chichewa, as illustrated in (2), and also with many familiar Indo-European languages, including most of the Germanic, Romance, and Slavic languages, where nouns are partitioned into gender classes. Indo-European gender is significant for agreement phenomena, but typically not much else.

In Blackfoot, an Algonquian language, nouns are partitioned into gender classes, and the classes control agreement (Ritter 2014). That the classes are noun-based, and not referent-based, is clear from the fact that some nouns describing inanimates are exceptionally classed as animate (for example *is-toan* 'knife'). However, unlike the situation in the Indo-European languages, animacy in Blackfoot plays a central role in a variety of grammatical phenomena, interacting with argument selection, verb classification, and a system of obviation.

English clearly lacks grammatical gender of the German, Chichewa, or Blackfoot variety, having no gender agreement on noun modifiers.[6] However,

[6]It is a minor quirk of written English that some users write *blonde* of women and *blond* of men, but even in this written variety, this would not qualify as a gender in the Hockett sense, because it does not track noun classes but referents, as with North Sámi animacy.

unlike North Sámi, there is a three-way distinction in the pronominal system between male, female and inanimate forms. An interesting question for feature theory is whether these distinctions represent syntactic features, or whether they are purely semantic.

Pollard and Sag (1994:73) discuss the matter and suggest that the obligatory gender match between a bound pronominal form and its binder in examples like (8) has the qualities of a grammatical, rather than a semantic constraint.

(8) a. That dog is so ferocious, he even tried to bite himself/*itself.
b. That dog is so ferocious, it even tried to bite itself/*himself.

As with sequence of tense, the existence of 'uninterpretable' gender features in English seems to lend support to their treatment as syntactic features. Kratzer (1998) and Heim (2008) discuss cases in which a bound pronoun must be interpreted as if it lacked the usual semantic contribution of its ϕ features, as illustrated in (9) for gender.

(9) Only Mary thought that she wouldn't win.

This sentence is perfectly felicitous even in a context where some or all of the other people under consideration are male; it can imply that none of them thought that *he* wouldn't win. Thus, when the bound pronoun is interpreted as a variable, the usual presuppositional restrictions on *she* are absent. This suggests that gender features on pronouns can be 'uninterpretable' in English (see §2.3). English could then be said to have formal gender features on pronouns, though not on lexical nouns.

2 Formal properties of feature systems

There are many unresolved issues about how features are structured and organized. Rich feature systems can be developed as a tool for linguistic description and theory development, in which case arbitrary decisions can be made about their formal properties, and the systems can be made fully explicit (see for example Gazdar et al. 1985, Pollard and Sag 1994, and Stabler 1997 for explicit feature systems for natural language). On the other hand, if features are meant to more directly model cognitive aspects of linguistic competence, then their nature is an empirical question, though in that case any

fully explicit system for syntax is currently premature.[7] Here I will simply lay out some of the major issues and terms.

2.1 Negative values

An issue which has been discussed at some length in phonology is whether some or all features are binary or privative. Binary features encode oppositions directly, for example if /p/, /t/, and /k/ are [–voice] and /b/, /d/, and /g/ are [+voice]. This sense of opposition was central to the phonological theory of distinctive features as originally developed in works such as Jakobson (1990 [1942]) and Jakobson et al. (1951).

In that tradition, binarity is connected with markedness; one member of an opposition is "actively modified" (in Trubetzkoy's terms, cited in Jakobson and Waugh 1979; on p. 93 of the reprint in Jakobson 1987), while the other lacks that characteristic. Plus values are often taken to correlate with markedness, in which case the minus value of any feature represents the unmarked case. In some analyses, minus values are filled in by rule wherever plus values are absent (as in the "radical underspecification" proposal of Kiparsky 1982).

Sometimes it is argued that the system need not refer to minus values at all, in which case the oppositions can be cast in terms of privative features (Lombardi (1996) argues that this is true of cyclic phonology but not of postlexical phonology). Privative features can only be present or absent, for example /b/, /d/, and /g/ might have the privative feature [voice], while /p/, /t/, and /k/ lack it; this would mean that [voice] can spread or otherwise be affected by phonological rules, but voicelessness cannot. Alternatively, /p/, /t/, and /k/ might have the feature [spread glottis] (Iverson and Salmons 1995).

[7] Graf (2017) shows how features can be eliminated from grammars by stating constraints which include disjunctive lists of lexical items. For example instead of saying that *himself, herself, itself* and *themselves* share a feature [anaphor], Principle A can be stated to refer directly to the list of those words. Graf notes that in many cases, the feature-free versions of constraints are significantly more complicated to state. In other cases, the reverse is true: although all linguistic constraints can be stated entirely in terms of GPSG/HPSG-like feature structures, the results are often computationally very cumbersome. A theory of cognitive representations will have something to say about when featural representations have the most independent utility, for example in acquisition, processing, functional neuroanatomy, or other areas.

In a Jakobsonian binary system, [+voice] and [–voice] are inherently contradictory and cannot coexist on the same segment. In a privative system, there is nothing inherently contradictory between [voice] and [spread glottis]. An opposition like that of English would only require one or the other, and a language with both would be expected to have a four-way contrast (as in Hindi).

Embedded in a well-developed theory of distinctive features, a privative system is potentially more restrictive, since negative values cannot be referred to. However, any binary opposition can in practice be recast in terms of a privative one (see Blaho (2008), ch. 1 for an example of how OT constraint rankings can emulate reference to the absence of privative features). Restrictiveness is not inherent in the choice of privative versus binary features but is a property of the feature calculus seen as a whole.

Another way to frame the question of whether features are privative or binary would be in terms of acquisition: When a learner acquires a binary paradigmatic featural opposition (such as singular-plural, masculine-feminine, finite-nonfinite, wh-non-wh, definite-nondefinite, etc.), for a given domain of elements D, is that opposition normally (all else being equal) encoded in the learner's linguistic knowledge in terms of (i) a privative feature F associated with the marked members of the opposition, or (ii) a partition of D into two classes, +F and –F? There are two versions of (ii): (iia), the partition is represented as such in the lexical entries for elements in D (the usual assumption of binarity), and (iib), only the +F values are represented in the lexical entries, while –F values are filled in by rule (the radical underspecification alternative).

One difference for acquisition between the two cases (i) and (ii) is the significance of D for the learner. In a privative system, non-F elements of D are not distinct from elements outside D, with respect to F-phenomena, which means that D has no independent significance for the system. In a binary system, D is the class of things which are specified for ±F, so identifying D is crucial to understanding the distribution of –F (unless the system makes no use of underspecification, so that every object is assumed to be either +F or –F).

If D is a feature motivated independently of the distribution of F, then we can say that F is dependent on D in a feature geometry (see §2.2).

In syntax, the absence of a well-developed theory of distinctive features makes it more difficult to argue for or against negative values. Privativity is sometimes considered to be conceptually more basic and hence taken to

be the null hypothesis (Adger and Svenonius 2011, Cowper and Hall 2014). For example, Harley and Ritter (2002) assume privativity in their analysis of ϕ features (discussed further in §2.2), and the case theory of Caha (2009) is based on a presumption of privative features. In contrast, Harbour (2011) argues that binary features must be deployed in the correct analysis of Kiowa noun classes and number interpretation (note that for Harbour, features are interpreted as functions and the minus value corresponds to negation, not necessarily the unmarked value; see also Harbour 2016 on binarity in person features).

Privative and binary systems are sometimes contrasted with 'multivalued' systems, where a feature is a composite of an attribute and a value, for example [person:speaker], [case:dative], or [number:dual]. Here, [person:__] is an attribute, and *speaker* and *addressee*, for example, would be values.

Chomsky (2000; 2001) suggests that an unvalued attribute such as [person:__] cannot be interpreted at the interface of syntax with semantics. If the syntactic computation is at some level a computation of semantically interpretable structures, then the need to value unvalued attributes can drive syntactic derivations, as discussed in §2.3.

If, on the other hand, [person] by itself has a syntactic interpretation even without a value, for example as third person, then it is a feature in its own right. If another feature such as [speaker] is subordinate to the feature [person], then that represents a hierarchical organization of (possibly privative) features, discussed in the next subsection.

2.2 Depth and breadth of featural organization

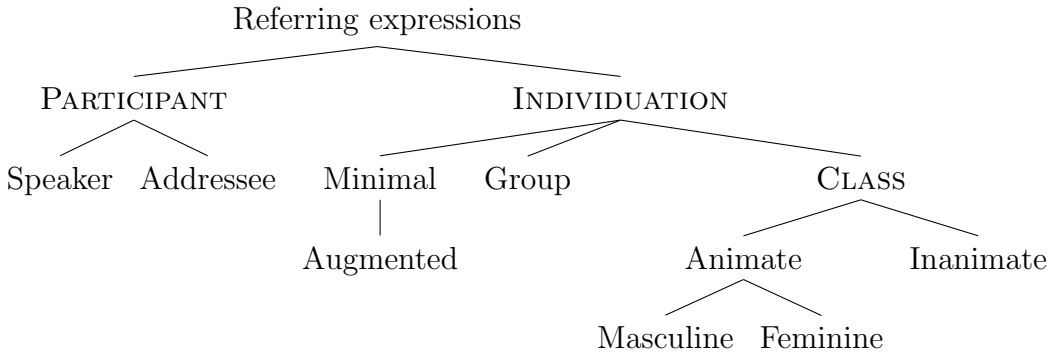
One set of unresolved issues involves how features are organized into classes. The organization may be 'deep' in the sense of having many levels, and it may be 'broad' in the sense of classes having many members.

A feature class is a grouping of features which behave together in some way. For example, person, number, and gender features are copied together under agreement, so behave as a class, dubbed ϕ-features by Chomsky (1981:330). In some cases the classes are treated as features in their own right, for example if ϕ has some syntactic interpretation, such as the property of being able to bear an index, regardless of what features it dominates.

This introduces a level of hierarchical organization among features, which is often represented in a graph, for example in the feature 'geometry' for ϕ of Harley and Ritter 2002, seen in (10). It includes features for person

(Speaker, Addressee), number (Minimal, Augmented, and Group) and gender (Inanimate, Animate, Masculine, Feminine).

(10)



This diagram encodes feature classes, for example putting all the gender features together under a class called Class, and grouping those together with number features in a class called Individuation. The claim is that features in a class behave together, for example when adjectival agreement copies gender and number without person. The geometry also encodes implications, for example in this tree, Masculine and Feminine are only possible if the feature Animate is present, and the feature Augmented is only possible in the presence of the feature Minimal.[8]

Harbour (2014; 2016) argues (contra Harley and Ritter) that the implications and groupings illustrated in geometries can and should be derived from properties of the features themselves, and therefore the geometry has no theoretical significance. In systems like that of HPSG (Pollard and Sag 1994), however, the organization of features into graphs directly represents implications and restrictions in the theory of feature types.

In cartography (see §3.2), a large set of category features is organized hierarchically, where the hierarchy represents a constraint on projection. This organization is deep, in the sense used here, but narrow, because there are very few categories at each level (normally only one, either privative or with a binary value). Caha's (2009) analysis of case, in this tradition, is based on a 'deep' hierarchical ordering of cases (instrumental over dative over genitive over accusative over nominative, for instance) but makes little or no use of

[8]In the analysis of number assumed by Harley and Ritter, Minimal in the absence of any other number features is singular, and in combination with Group gives dual number; and Augmented is added to Minimal to get trial and paucal numbers. See Harbour (2011) and Harbour (2014) for a more fully developed theory of number features.

feature classes.

The HPSG system of features is maximally deep and broad, since there are no limits on embedding of features; in fact, the feature system is fully recursive (Adger 2010, Adger and Svenonius 2011).

In Minimalist Grammars (e.g., Stabler 2011), features are often assumed to be arranged on heads in stacks, which determine the order of operations (selection features before licensing features, for example). As noted by Stabler (2013), equivalent effects can be achieved without stacks in a cartographic model where features are distributed across multiple heads; in effect, the hierarchy of functional heads is a stack of features.

2.3 Feature checking, interpretability and valuation

As noted in §1, features in one location in a syntactic structure interact with features in other locations, as when agreement features on a target covary with intrinsic features of a controller, or when licensing features on a case-assigner interact with the case-needing property of a noun phrase. In derivational terms, features are triggers for syntactic operations, or instructions to create certain kinds of syntactic dependencies (see Adger 2010 and Rizzi (2018) for clear statements to this effect, though the understanding goes back much further).

In one popular implementation, the syntactic relation among features is called 'checking': the features on the target have the property that they must be checked, which happens if matching features can be found within the appropriate domain (Chomsky 1995). Once checked, the features on the target are deleted. Deletion in this sense does not mean that the features are not morphologically expressed, only that they are not semantically interpreted, and do not interfere with semantic interpretation of the structures in which they are contained.

The essentials of this relationship provide the foundation for a formal theory of featural interactions in a broad range of syntactic domains. In Chomsky's (2000, 2001) probe-goal model of feature checking, the operation of feature checking is called Agree. The general term for the locus of features which need checking is PROBE, and the general term for the locus of the features which satisfy the needs of the probe is GOAL. A probe is said to 'search' a domain for a suitable goal or goals (see Hiraiwa 2005, Nevins (2007; 2011) on 'Multiple Agree,' where one probe interacts with multiple goals, but see also Haegeman and Lohndal 2010 and Frampton and Gutmann (2006) for

an alternative). The features involved in the relationship are assumed to be identical in one respect, since they must match in order for checking to take place; but in another respect they are distinct. Adger and Svenonius (2011) suggest the term 'second-order property' for properties that distinguish two instances of a feature; in these terms, a feature F on the probe has a second-order property that distinguishes it from the otherwise identical feature F on the goal.

HPSG similarly uses a basic featural relationship as a core syntactic device ("the central explanatory mechanism," Pollard and Sag 1994:19). In HPSG terms, the values of features at one node are *shared* with those of another node. Sharing is imposed in HPSG by TAGS which mark the parts of the feature structure that are unified (Shieber 1986, Pollard and Sag 1994). The relationship is not fundamentally asymmetrical; features at two tags must not conflict, but neither is specially marked as a probe.

In the Minimalist Program, the relationship is conceived of as asymmetric, so that syntactic operations are driven by some need or deficiency of the probe, its second-order property. In Chomsky (1995:Ch. 4), it is proposed that the second-order property in question is uninterpretability at the interface with semantic interpretation (LF). Uninterpretable features (such as ϕ on a finite verb) must be checked, matched, and deleted with interpretable counterparts (ϕ on a noun phrase) or the derivation crashes.[9]

In Chomsky (2000; 2001), it is proposed that interpretability be equated with having a value; an uninterpretable feature is an attribute with no value, and feature checking is the copying of values from fully specified features. The possibility of features being morphologically expressed but semantically uninterpreted has figured centrally in several arguments, a couple of which came up in §1.3 (tense) and §1.4 (gender). Uninterpretability figures centrally in proposals about the nature of syntactic features by Pesetsky and Torrego (2001), Svenonius (2007), Zeijlstra (2014), and Smith (2015), among others.

Pesetsky and Torrego (2007) argue that interpretability and valuation cannot be conflated, as Chomsky suggested, but are independent second-order properties. For them, probing is always driven by the need to provide a value for an unvalued attribute, but in some cases the features on the probe are uninterpretable (as in agreement) and in other cases they are in-

[9]See Preminger (2014) for criticism of interface uninterpretability as an explanatory device for second-order features driving syntactic operations. He points out that 'last resort' or 'repair' mechanisms are often posited to avoid derivational crash when features cannot be checked.

terpretable (as in their analysis of English finite verbs agreeing in tense with an unpronounced T). Similarly, they argue, the features on a goal may be interpretable or uninterpretable.

Recent work in the Minimalist Program has further explored the configuration of Agree and other properties of the feature-checking relation. Koopman (2006) argues that Agree can be restricted to the spec-head configuration (a subset of sisterhood relations), granted certain other assumptions.

Baker (2008) argues that a probe can search 'upward,' to a c-commanding goal, subject to parametrization. Wurmbrand (2012), Zeijlstra (2012), and Bjorkman and Zeijlstra (to appear) suggest that probing is always upward; the c-commanded element in an Agree relation drives the operation. Rezac (2003) and Béjar and Rezac (2009) argue that Agree is cyclic, probing first downward, and then upward under certain conditions (as does Carstens 2016).

2.4 Strength

In the Minimalist Program, another second-order property of features is strength, the property which determines whether movement occurs (sometimes called the "EPP property," Chomsky 2000).

Movement is conceived as internal Merge, which means that the phrase containing the goal combines with the projection of the head bearing the probe, creating a new specifier. The specifier is linearized according to the usual principles of linearization, normally to the left of the head.

For example, in a language like English where the specifier of TP must be filled, T has a strong D feature. So T probes its domain for a D, and when D is found, it remerges with T and creates a specifier. This is then the subject. Parametric variation can be described in terms of strength; for example if a wh probe in C is strong in English, but weak in Chinese, then English will have overt wh movement to CP, and Chinese will not (Huang 1982; for Huang, 'covert' movement occurs after spell-out, while Chomsky 1995, p. 265, suggests that covert movement is the movement of features without accompanying structure, an idea pursued further in Bobaljik 2002).

2.5 Phrasal features

I suggested in §1 that features are intrinsically properties of heads, and can only be properties of phrases by some kind of extension, such as agree-

ment, percolation, or the like. However, there are empirical challenges to the strictest interpretation of that assumption. For example, a coordination of two singulars is often formally treated as a dual or plural, as in the following example from North Sámi.

(11) Mu eadni ja Máret-goaski leaba oappážat.
*my mother and Máret-aunt are.*DU *sisters*
'My mother and Aunt Máret are sisters' (North Sámi (Finno-Ugric), Nickel 1994:510)

It appears in such cases that the computation introduces a feature which is not lexically present in the atoms composing the agreement controller. This conclusion might be avoided if this case of agreement were treated semantically, that is, dual agreement is possible in case the referent consists of two specific animate individuals.

This might seem at first to be bolstered by the fact, already noted in §1.2, concerning (3), that lexical nouns in North Sámi don't have morphologically dual forms. However, there doesn't seem to be any implication that if agreement with coordinate singulars is dual or plural, then that agreement is semantically controlled rather than syntactically controlled. For example the English plural appears to be fairly thoroughly grammaticized on nouns (e.g., there are pluralia tantum such as *scissors*), and there, coordinations of singulars control plural agreement (*This and this are/*is ready*).

Computations are also necessary to determine the gender agreement controlled by coordinate structures, as discussed for Chichewa by Corbett and Mtenje (1987), whence the following examples. The plural of class 3 (subject marker *u-*, see (2)) is class 4 (subject marker *i-*).

(12) a. Mu-dzi u-ku-kula.
*3-village SM3-*PRES*-grow*
'The village is growing'
b. Mi-dzi i-ku-kula.
*4-village SM4-*PRES*-grow*
'The villages are growing'

Here we can talk of a lexically specified gender combining with a number to give class 3 or 4. A coordination of two class 4 nouns is unsurprisingly class 4.

(13) mi-peni ndi mi-phika i-ku-sowa.
4-knife and 4-pot *SM4*-PRES-*missing*
'Knives and pots are missing'

However, a coordination of two class 3 nouns controls class 8 agreement (subject marker *zi-*), as shown in (14).

(14) m-peni ndi m-phika zi-ku-sowa.
3-knife and 3-pot *SM8*-PRES-*missing*
'A/The knife and pot are missing'

In (14), class 8 appears to be introduced by the computation. We may still be able to avoid the conclusion that the computation introduces features, if the gender feature of singulars is not available to a probe outside the coordinate structure, and agreement is in number only, class 8 being a kind of default plural for nonhumans. Going further into the complexities of feature resolution would take us too far afield. See Dalrymple and Kaplan (2000) for discussion.

3 Categories and distinctive features

3.1 Lexical categories and subcategories

No grammatical description can make do without identifying distributional categories of words and morphemes such as noun, verb, conjunction, and so on. Chomsky (1965) suggests doing this with distinctive features such as [+N[oun]] and [+V[erb]], and later work attempted to connect such features with formal content, for example the [+V] categories V and A are predicates, and the [–N] categories V and P assign case (Chomsky 1981; see Hale and Keyser 1999 and Baker 2003 for alternative proposals concerning the properties distinguishing the basic lexical categories).

For finer discriminations, for example subcategories of verb such as transitive and intransitive, Chomsky (1965) proposes 'strict subcategorization' features, which are category features like [+N] but specified as restrictions on complementation. On that proposal, subcategorization is a kind of second order property in the sense of Adger and Svenonius (2011). A transitive verb has both a [+V] feature and a [+N] feature, but the [+N] feature on a transitive verb has the second-order property of being a subcategorization

feature, a need for a complement bearing that feature.

Chomsky (1965) also tentatively proposes a distinction between the syntactic features invoked in subcategorization and those involved in semantic 'selection' (e.g., the verb *admire* requires an animate subject). Pesetsky (1982) suggests the terms C-selection (C for category) and S-selection (S for semantic).

In these terms, (15a), where an obligatory noun phrase object is omitted, would be a violation of the subcategorization or C-selection properties of *frighten*, and (15b), where an inanimate object is found in a position reserved for animate objects, would be a violation of that verb's selection, or S-selection properties.

(15) a. *Sincerity frightened.
b. #The boy frightened sincerity.

Subcategorization allows the larger category of verb to be subclassified into smaller subcategories on the basis of the categories of complements, and similarly for any other complement-taking category. For example, Emonds (1985) proposed unifying the category P with the category of subordinating conjunction. In traditional grammar, a word like *before* is a preposition when it takes a noun phrase complement (*before lunch*) and a subordinating conjunction when it takes a phrasal complement (*before we ate*). Emonds suggested (building on observations by Klima (1965)) that such cases could be unified with each other, and with unambiguous prepositions such as *from* and unambiguous subordinating conjunctions such as *although*, under a more general category P, using subcategorization to distinguish the different cases. Verbal particles such as *up* and *down* can then be treated as intransitive members of category P.

Formally, both subcategorization and selection can be treated in terms of an asymmetric feature-checking relationship: the verb *frighten* has an uninterpretable probe for a feature which is interpretable on its complement, whether the feature is syntactic or semantic. Work since the 1970's has tended to view semantic selection as non-syntactic (hence the cross-hatch in (15b), rather than a star; see e.g., McCawley 1968, Jackendoff 1972).[10]

[10]McCawley pointed out that unlike subcategorization, where the features involved were independently motivated, semantic selection appears to require a large number of additional features. Jackendoff pointed out that semantic selection would have to be nonlocal to account for contrasts like that in *I ate something that was the result of what Bill ac-*

However, an exact line between subcategorization and selection is difficult to draw. It has been suggested that subcategorization be treated together with selection, as a semantic matter outside of grammar proper. See Grimshaw (1979), Pesetsky (1982), and Borer (2005) for arguments that category selection is epiphenomenal, at least in some cases, and Emonds (1991; 2000) and Wurmbrand (2012) for a defense of syntactic subcategorization.

3.2 Functional categories

Traditional grammar makes use of some very general categories such as 'adverb,' and 'particle,' and features have been proposed to account for distributional distinctions among them. In English and many other languages, a class of auxiliaries can be distinguished from lexical verbs, and distinctions can be made among the different kinds of auxiliaries. Inspired by the theory of distinctive features, some work attempts to capture multiple oppositions with a small number of features cutting across categories.

In Abney's (1987) proposal, each category can be analyzed as a combination of of two features: a part-of-speech value like N or V, and a functional feature F, specified as minus for lexical categories and positive for functional categories. Grimshaw (1991) accommodated more elaborated extended projections by proposing that the F feature was multivalued. In the manuscript circulated in 1991, she suggested that there was one lexical level, 0, and two functional levels, 1 and 2, so that for example D would have the part-of-speech category of a noun and the F value 1, while T was a verb with F value 1. Above D and T were the level 2 functional categories P and C (cf. Emonds 1985 on the unification of P and C, as mentioned just above). Adger (2003) and Grimshaw (2005) adapt Grimshaw's 1991 proposal to higher numbers of categories.

The theory of Ritter and Wiltschko (2009) and Wiltschko (2014) proposes four levels and identifies a characteristic functional interpretation for each level. The lowest level is CLASSIFICATION, the next level is POINT OF VIEW, and these are dominated by ANCHORING, which is in turn dominated by LINKING. For a similar proposal which attempts to bridge cartography and minimalism, see Ramchand and Svenonius (2014). More distantly similar suggestions can also be found in the framework of Functional Grammar (Dik 1989, Rijkhoff 2002).

---

knowledged to be a {new baking process/#syntactic transformation}.

The overall logic of a hierarchy of functional categories works the same way regardless of how many there are. Cinque (1999), on the basis of a cross-linguistic survey, proposes a fine-grained 'cartography' for clausal structure in which functional categories such as auxiliaries are differentiated according to their canonical order with respect to other elements. He treats free function words such as modals in English on a par with bound morphemes such as modal suffixes in Turkish, and assigns a distinct category label to each class with a distinct distribution. The result is an inventory of about three dozen hierarchically arranged categories in the clause, including things like past tense, epistemic modality, and progressive aspect. A similar approach is taken to the noun phrase (Cinque 1994).

Some features are associated directly with semantic content, specifying logical notions like universal quantification, set-theoretical notions like containment, or mereological notions like part. Hale (1986) proposed that an opposition between central coincidence and non-central coincidence lies at the heart of several paradigms in Warlpiri, including locative enclitics on nouns, directional enclitics on verbs, the system of finite complementizers, and the aspectual system. He posited a a feature ±central-coincidence which is found in several distinct categories. Variants of Hale's coincidence feature have been extended to additional domains and languages by Demirdache and Uribe-Etxebarria (2000), Hale and Keyser (2002), Mateu (2002), Cowper (2005), Ritter and Wiltschko (2005), and others.[11]

In the same vein, Harbour (2007; 2011) has argued that the same features are used in Kiowa to differentiate gender in one domain and number in another; the features have contextually determined interpretations, so that for example a [−augmented] feature in the Class domain signals membership in a class of referents which are not composed of subparts with the same properties, namely animates (e.g., a part of a deer is not a deer), while a [−augmented] feature in the Number domain indicates numerical nondivisibility, namely nonplurals, which lose their numerical properties if divided (a part of a dual is not a dual).

Sometimes morphological identity motivates the unification of features in distinct domains. For example in Kayardild, the same morpheme (*-kuru*,

[11]Category-neutral semantic properties have long been noted, for example Bach (1986) and Krifka (1987) on parallels between nominal reference and temporal constitution, but the references cited in the text explicitly motivate category-neutral syntactic features. Harbour (2014) formalizes the identity between number features and features used in aspectual systems.

with a range of allomorphs including *-(j)u* and *-(w)uru*) expresses potential mood, proprietive case (meaning 'having'), and a 'modal proprietive' which in conjunction with the potential mood expresses future tense (realized on VP-internal material). All three are seen in (16) (from Evans 1995:146; the suffixes glossed POT, PROP, and MPROP are all allomorphs of the same form).

(16) Ngada kurri-ju midijin-kuru-wuru dangka-wuru.
1SG.NOM *see*-POT *medicine*-PROP-MPROP *person*-MPROP
'I will see the doctor (medicine-having person)'

Of course, morphological identity does not entail syntactic featural unity; Aronoff (1994) proposes the notion MORPHOME for morphological forms with multiple syntactic functions. But if analyses like Hale's and Harbour's are correct (where the morphemes are distinct on the surface, but underlying syntactic features are shared), that suggests that learners are conservative with respect to positing new features, and liberal with respect to extending featural meanings to new domains (see §3.3). In that case a learner of Kayardild should be strongly motivated to posit a common underlying syntactic feature for the potential, the proprietive, and the modal proprietive.

3.3 The size of the feature inventory

Adult speakers of English make use of between twenty and fifty thousand words, and comprehend far more; already by the age of six, English-speaking children understand about 14,000 distinct words (Clark 1993). This means that English-speaking children must acquire an average of more than seven words per day between the ages of one and six (assuming that relatively few words are learned in the first year, when the child is still mastering the phonology). Many of those words, however, are featurally indistinct, from the perspective of syntax; the syntax of *stegosaurus* is indistinct from the syntax of *cat*, and so to acquire the word *stegosaurus* does not require any new syntactic features to be posited.

Suppose that the total number of syntactically distinct entries in the mental lexicon is on the order of one thousand. Each one could in principle correspond to a sui generis category feature. However, if the categories can be grouped into classes, then fewer features are necessary. Ten binary oppositions, freely combined, generate over one thousand distinct sets (2^{10}=1024), so an inventory of one thousand functional and lexical categories could be

fully differentiated by a set of just ten freely combining features.[12]

Whether the number of features employed in syntax is closer to ten or closer to one thousand depends on to what extent learners are conservative or liberal in positing new features to describe newly observed oppositions, as opposed to extending a previously posited feature to a new domain or category.[13]

4 Open issues

The feature system of HPSG is explicit and very rich. Minimalist feature systems are far more spare, even when made explicit as in Adger (2010) or Stabler (1997). A minimalist feature system consists of an inventory of 'first order' features (organized into a set of feature classes, shallow in the sense of §2.2), combined with a small inventory of second order properties which correspond to operations or constraints and relations. The number of second order properties found in natural languages may be very small; Chomsky (1977) pointed out the fundamental unity of unbounded dependencies, suggesting that such phenomena as wh-movement, topicalization, relative clause formation, and comparative deletion involve the same second order property.

A Minimalist Grammar feature calculus includes a small inventory of second order properties, for example if F^- is a category then F^+ is the second order feature selecting for that category. A verb selecting for a DP complement would have both V^- and D^+.[14]

[12]Compare Kayne (2005), where it is suggested that a language like English might have somewhat more than 100 distinct functional categories, each bearing a single syntactically relevant feature.

[13]The principle of 'Feature economy' in Roberts and Roussou (2003) and Biberauer et al. (2014) is intended to characterize the conservatism of learners with respect to associating a lexical entry with an innate, UG-endowed feature. Giorgi and Pianesi (1997) and Hegarty (2005) are similar in this respect. But some similar principle could be stated to characterize the conservatism of learners with respect to positing a new feature, more in line with Chomsky's (2005) conception of a minimal UG.

[14]Stabler (1997) posits four second order properties: subcategorization (symbolized by "=," in Stabler 1997), the need to be licensed ("–"), the ability to license ("+"), and the property of triggering movement (signified by upper case), including both phrasal movement and incorporation, but with two variants of the latter, one for prefixation and one for suffixation. In that system, all second-order features can be said to be 'uninterpretable' as they are eliminated in the course of the derivation, but Stabler (2011) outlines a minimalist grammar with interpretable second-order features (which he calls 'persistent').

The issue of interpretability is linked to Chomsky's (2000, 2001) suggestion that operations and constraints are driven by the nature of the interfaces ("Bare Output Conditions"). An alternative (as argued by Rizzi 2010, Rizzi 2018 and Preminger 2014) is that that some second order properties invoke purely syntactic operations (they are 'syntax internal' features in the sense of Svenonius 2007). In Rizzi's terms, a feature is "an instruction for a certain syntactic action" (Rizzi 2010: 150).

The number of different second order properties depends on the extent to which phenomena such as selection, licensing, agreement, movement, and ellipsis can be unified or outsourced. For example, a feature can be copied into an adjunct under agreement (as when an adjoined adjective agrees in gender, number, and/or case with the noun phrase to which it is adoined, as in (1)), whereas a syntactic node cannot be moved into an adjunct. This could suggest that agreement and movement involve distinct second order properties.

On the other hand, Adger (2010) locates the difference not in the second order property, but in the classes of first order features. He uses a version of the second order property of valuation to analyze both movement and agreement. Movement involves an unvalued category feature (e.g., an unvalued D category feature on T attracts a DP, as in Chomsky (2001)), while agreement involves an unvalued feature from a distinct class of morphosyntactic features (e.g., unvalued person and number attributes on T are unified with person and number values in the search domain).

Greater restrictiveness in the theory of syntactic features can be achieved if some of these phenomena are outsourced, as advocated by Boeckx (2014). There are two directions for outsourcing. One is morphology: syntax can be rid of such morphosyntactic phenomena as case and agreement if they can be relegated to an active, computational morphological component (but at the cost of endowing the morphological module with syntax-like properties, see note 4).

The other direction in which grammatical phenomena can be outsourced from syntax is semantics. I have already mentioned (in §3.1). attempts to explain putative subcategorization in terms of semantic selection. In that case, no subcategorization features are needed in syntax. The issue is more broadly stated in terms of whether Merge is 'free' or 'triggered.' Triggered Merge means that there are features establishing which syntactic objects can be combined with which by Merge. Free Merge means that Merge can freely combine any syntactic objects; any impossible combinations must be ruled

out at the interface with semantic interpretation.

Another example of a phenomenon which has been variously treated as syntactic or semantic is polarity item licensing (Zeijlstra 2004). For example, phrases with *any* have a restricted distribution compared to phrases with other determiners, which can be described in terms of licensing: *any* requires an appropriate licensing context. Negation provides appropriate licensing, so that a phrase with *any* is licit in the scope of negation but not in the simple affirmative context in (17a).[15]

(17) a. *The police interviewed a witness who saw anything.
b. The police never interviewed a witness who saw anything.

This can be treated in terms of feature checking, for example *any* has a feature which needs to be checked by a feature borne by negation and similar operators. If polarity item licensing is feature checking, then feature checking can occur across island boundaries, since *never* and *anything* in (17b) are separated by a relative clause boundary; relative clauses are generally islands for extraction, as (18) shows.

(18) *What did the police (never) interview a witness who saw?

Another implication of polarity licensing is that features may have to be introduced by computations, rather than always being inherent to heads.

On the other hand, if licensing of polarity items like *any* is part of a semantic component, then it may not bear on syntax. In that case, the examples above do not show that feature-checking relations can cross island boundaries, nor that they require computation.

In general, the nature of the boundary between syntax and semantics is unclear, with many research questions remaining open. Clearly, conceptual content (the domain of the differences among *shoes* and *ships* and *cabbages*) is not based on features of the sort relevant to syntax. Syntactic features are plausibly minimally distinctive in the sense discussed in §1, whereas the conceptual content of lexical items is nothing of the sort. At the other end of the spectrum, there are syntactic features without semantic content. But in between the two extremes of semantically inert syntactic features and syntactically irrelevant conceptual content, there are a great many distinctions which can in principle either be analyzed as parts of a syntactic module or of

[15]I am setting aside the distribution of 'free-choice *any*,' which doesn't seem easily available in (17a).

a semantic module, with a relatively transparent mapping between the two. But the choice is not innocuous; managing these distinctions in the syntax requires a richer syntactic machinery, and relegating them to the semantics requires a semantic component which constrains as much as it interprets. The resolution of some of these issues would allow great progress in the theory of syntactic features.

References

Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*. Ph.D. thesis, MIT, Cambridge, Ma.

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

Adger, David. 2010. A minimalist theory of feature structure. In *Features: Perspectives on a Key Notion in Linguistics*, edited by Anna Kibort and Greville G. Corbett, pp. 185–218. Oxford University Press, Oxford.

Adger, David and Daniel Harbour. 2008. Why phi? In *Phi-Theory: Phi Features across Modules and Interfaces*, edited by Daniel Harbour, David Adger, and Susana Béjar, pp. 1–34. Oxford University Press, Oxford.

Adger, David and Peter Svenonius. 2011. Features in minimalist syntax. In *The Oxford Handbook of Linguistic Minimalism*, edited by Cedric Boeckx, pp. 27–51. Oxford University Press, Oxford.

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Linguistic Inquiry Monographs 22. MIT Press, Cambridge, Ma.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9 1: 5–16.

Baker, Mark C. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. No. 102 in Cambridge Studies in Linguistics. Cambridge University Press, Cambridge.

Baker, Mark C. 2008. *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.

Béjar, Susana and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40 1: 35–73.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45 2: 169–225.

Bjorkman, Bronwyn and Hedde Zeijlstra. to appear. Upward agree is superior. *Linguistic Inquiry* .

Blaho, Sylvia. 2008. *The Syntax of Phonology: A Radically Substance-Free Approach*. Ph.D. thesis, University of Tromsø.

Bobaljik, Jonathan David. 2002. A-chains at the PF-interface: Copies and covert movement. *Natural Language & Linguistic Theory* 20 2: 197–267.

Bobaljik, Jonathan David. 2008. Missing persons: A case study in morphological universals. *The Linguistic Review* 25 1-2: 203–230.

Boeckx, Cedric. 2014. *Elementary Syntactic Structures*. Cambridge University Press, Cambridge.

Booij, Geert. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In *Yearbook of Morphology 1995*, pp. 1–16. Springer.

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

Caha, Pavel. 2009. *The Nanosyntax of Case*. Ph.D. thesis, University of Tromsø.

Carstens, Vicki. 2016. Delayed valuation: A reanalysis of goal features, "upward" complementizer agreement, and the mechanics of case. *Syntax* 19 1: 1–42.

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

Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, edited by Peter Culicover, Thomas Wasow, and Adrian Akmajian, pp. 71–132. Academic Press, New York.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.

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. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1–22.

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.

Clark, Eve V. 1993. *The Lexicon in Acquisition*. Cambridge University Press, Cambridge.

Clements, George N and Elizabeth V Hume. 1995. The internal organization of speech sounds. In *The Handbook of Phonological Theory*, edited by John Goldsmith, pp. 245–306. Blackwell, Oxford.

Clements, Nick, Alexis Michaud, and Cédric Patin. 2010. Do we need tone features? In *Tones and Features: Phonetic and Phonological Perspectives*, edited by John A. Goldsmith, Elizabeth Hume, and W. Leo Wetzels, pp. 3–24. Walter de Gruyter, Berlin.

Collins, Chris and Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19 1: 43–78.

Corbett, Greville G. 1991. *Gender*. Cambridge University Press, Cambridge.

Corbett, Greville G. 2012. *Features*. Cambridge University Press, Cambridge.

Corbett, Greville G and Alfred D Mtenje. 1987. Gender agreement in Chichewa. *Studies in African Linguistics* 18 1: 1–38.

Cowper, Elizabeth. 2005. The geometry of interpretable features: Infl in English and Spanish. *Language* 81 1: 10–46.

Cowper, Elizabeth and Daniel Currie Hall. 2014. Reductio ad discrimen: Where features come from. *Nordlyd* 41 2: 145–164.

Dalrymple, Mary and Ronald M Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76 4: 759–798.

Demirdache, Hamida and Myriam Uribe-Etxebarria. 2000. The primitives of temporal relations. In *Step by Step. Essays on Minimalist Syntax in Honour of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, pp. 157–186. MIT Press, Cambridge, Ma.

Dik, Simon C. 1989. *The Theory of Functional Grammar*. Foris, Dordrecht.

Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories*. No. 19 in Studies in Generative Grammar. Foris, Dordrecht.

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

Emonds, Joseph E. 1991. Subcategorization and syntax-based theta-role assignment. *Natural Language and Linguistic Theory* 9 3: 369–429.

Enç, Mürvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18 4: 633–659.

Evans, Nicholas. 1995. *A Grammar of Kayardild: With Historical-Comparative Notes on Tangkic*. Mouton de Gruyter, Berlin.

Frampton, John and Sam Gutmann. 2006. How sentences grow in the mind: Efficient computation in minimalist syntax. In *Agreement Systems*, edited by Cedric Boeckx, pp. 121–157. John Benjamins, Amsterdam.

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

Giorgi, Alessandra and Fabio Pianesi. 1997. *Tense and Aspect: From Semantics to Morphosyntax*. Oxford University Press, New York.

Graf, Thomas. 2017. A computational guide to the dichotomy of features and constraints. *Glossa* 2 1: 1–36.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10: 279–326.

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

Grimshaw, Jane. 2005. Extended projection. In *Words and Structure*, edited by Jane Grimshaw, pp. 1–73. CSLI, Stanford, Ca. Revised version of 1991 ms.

Haegeman, Liliane and Terje Lohndal. 2010. Negative concord and (multiple) agree: A case study of West Flemish. *Linguistic Inquiry* 41 2: 181–211.

Hale, Ken. 1986. Notes on world view and semantic categories: Some Warlpiri examples. In *Features and Projections*, edited by Pieter Muysken and Henk van Riemsdijk, pp. 233–254. Foris, Dordrecht.

Hale, Ken and Jay Keyser. 1999. Bound features, Merge, and transitivity alternations. In *Papers from the UPenn/MIT Roundtable on the Lexicon*, edited by Liina Pylkkänen, Angeliek van Hout, and Heidi Harley, MITWPL, pp. 49–72. MITWPL, Cambridge, Ma.

Hale, Ken and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. No. 39 in Linguistic Inquiry Monographs. MIT Press, Cambridge, Ma.

Halle, Morris, Bert Vaux, and Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31 3: 387–444.

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

Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42 4: 561–594.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90 1: 185–229.

Harbour, Daniel. 2016. *Impossible Persons*. MIT Press, Cambridge, Ma.

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

Hegarty, Michael. 2005. *A feature-based Syntax of Functional Categories: The Structure, Acquisition and Specific Impairment of Functional Systems*, vol. 79. Walter de Gruyter, Berlin.

Heim, Irene. 2008. Features on bound pronouns. In *Phi-Theory: Phi Features across Modules and Interfaces*, edited by Daniel Harbour, David Adger, and Susana Béjar, pp. 35–56. Oxford University Press, Oxford.

Hiraiwa, Ken. 2005. *Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture*. Ph.D. thesis, Massachusetts Institute of Technology.

Hockett, Charles Francis. 1958. *A Course in Modern Linguistics*. Macmillan, New York,.

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

Hyman, Larry M. 2010. Do tones have features? In *Tones and Features: Phonetic and Phonological Perspectives*, edited by John A. Goldsmith, Elizabeth Hume, and W. Leo Wetzels, pp. 50–80. Walter de Gruyter, Berlin.

Iverson, Gregory K and Joseph C Salmons. 1995. Aspiration and laryngeal representation in germanic. *Phonology* 12 03: 369–396.

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

Jakobson, Roman. 1987. *Roman Jakobson: Selected Writings*. Mouton de Gruyter, Berlin. Edited by Stephen Rudy.

Jakobson, Roman. 1990 [1942]. The concept of phoneme. In *On Language: Roman Jakobson*, edited by Linda R. Waugh and Monique Monville-Burston, pp. 217–241. Harvard University Press, Cambridge, Ma. Originally presented in New York in 1942.

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

Jakobson, Roman and Linda Waugh. 1979. *The Sound Shape of Language*. Indiana University Press, Bloomington, In.

Kayne, Richard S. 2005. 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.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In *Linguistics in the Morning Calm*, edited by I. S. Yang, pp. 1–91. The Linguistic Society of Korea and Hanshin, Seoul.

Klima, Edward S. 1965. *Studies in Diachronic Syntax*. Ph.D. thesis, Harvard.

Koopman, Hilda. 2006. Agreement configurations: In defense of "spec head". In *Agreement Systems*, edited by Cedric Boeckx, pp. 159–199. John Benjamins, Amsterdam.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In *Proceedings of SALT VIII*, pp. 92–110. CLC Publications.

Krifka, Manfred. 1987. Nominal reference and temporal constitution: Towards a semantics of quantity. In *Proceedings of the 6th Amsterdam Colloquium*, edited by Jeroen J. Groenendijk, Martin Stokhof, and Frank Veltman, pp. 153–173. Institute of Linguistic, Logic and Information, University of Amsterdam, Amsterdam.

Kuhn, Jeremy. 2016. ASL loci: Variables or features? *Journal of Semantics* 33 3: 449–491.

Lombardi, Linda. 1996. Postlexical rules and the status of privative features. *Phonology* 13 1: 1–38.

Mateu, Jaume. 2002. *Argument Structure: Relational Construal at the Syntax Interface*. Ph.D. thesis, Universita autónoma de Barcelona, Barcelona.

McCawley, James D. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4: 243–269.

Mchombo, Sam A. 1998. Chichewa (bantu). In *The Handbook of Morphology*, edited by Andrew Spencer and Arnold M. Zwicky, pp. 500–520. Blackwell, Oxford.

Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25 2: 273–313.

Nevins, Andrew. 2011. Phonologically conditioned allomorph selection. In *The Blackwell Companion to Phonology*, edited by Colin Ewen, Elizabeth Hume, Marc van Oostendorp, and Keren Rice, pp. 2357–2382. Wiley-Blackwell, Oxford.

Nickel, Klaus Peter. 1994. *Samisk Grammatikk*. Davvi Girji, Karasjok, second edn.

Nielsen, Konrad. 1926. *Lærebok i Lappisk (Samisk), Bind I: Grammatikk*. Universitetsforlaget, Oslo.

Pesetsky, David. 1982. *Paths and Categories*. Ph.D. thesis, MIT, 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.

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

Preminger, Omer. 2014. *Agreement and its Failures*. MIT Press, Cambridge, Ma.

Ramchand, Gillian and Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46 B: 152–174.

Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax* 6 2: 156–182.

Rijkhoff, J. 2002. *The Noun Phrase*. Oxford University Press, Oxford.

Ritter, Elizabeth. 2014. Featuring animacy. In *Nordlyd, Tromsø Working Papers in Language & Linguistics: 41.1, Special Issue on Features*, edited by Martin Krämer, Sandra-Iulia Ronai, and Peter Svenonius, pp. 103–124. University of Tromsø – The Arctic University of Norway, Tromsø. Available at http://septentrio.uit.no/index.php/nordlyd.

Ritter, Elizabeth and Martina Wiltschko. 2005. Anchoring events to utterances without tense. In *The Proceedings of the 24th West Coast Conference on Formal Linguistics*, edited by John Alderete et al., pp. 343–351. Cascadilla, Somerville, Ca.

Ritter, Elizabeth and Martina Wiltschko. 2009. Varieties of INFL: Tense, location, and person. In *Alternatives to Cartography*, edited by Jeroen van Craenenbroeck Hans Broekhuis and Henk van Riemsdijk, pp. 153–202. Mouton de Gruyter, Berlin.

Rizzi, Luigi. 2010. On the elements of syntactic variation. *Studies in Linguistics* 4: 142–162.

Rizzi, Luigi. 2018. On the format and locus of parameters: The role of morphosyntactic features. *The Linguistic Review* 41 3-4: 159–190.

Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*, vol. 100. Cambridge University Press, Cambridge.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, MIT, Cambridge, Ma.

Schlenker, Philippe. 2016. Featural variables. *Natural Language & Linguistic Theory* 34 3: 1067–1088.

Shieber, Stuart M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. CSLI, Stanford, Ca.

Smith, Peter W. 2015. *Feature Mismatches: Consequences for Syntax, Morphology, and Semantics*. Ph.D. thesis, University of Connecticut.

Stabler, Edward. 1997. Derivational minimalism. In *Logical Aspects of Computational Linguistics*, edited by Christian Retoré, pp. 68–95. Springer, Dordrecht.

Stabler, Edward P. 2011. Computational perspectives on minimalism. In *Oxford Handbook of Linguistic Minimalism*, edited by Cedric Boeckx, pp. 617–642. Oxford University Press, Oxford.

Stabler, Edward P. 2013. Two models of minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5 3: 611–633.

Stowell, Tim. 2007. The syntactic expression of tense. *Lingua* 117 2: 437–463.

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

Toivonen, Ida. 2007. Verbal agreement in Inari Saami. In *Saami Linguistics*, edited by Ida Toivonen and Diane Nelson, pp. 137–166. John Benjamins, Amsterdam.

Wiltschko, Martina. 2014. *The Universal Structure of Categories: Towards a Formal Typology*. Cambridge University Press, Cambridge.

Wurmbrand, Susi. 2012. The syntax of valuation in auxiliary-participle constructions. *Coyote Papers: Working Papers in Linguistics, Linguistic Theory at the University of Arizona* .

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

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29: 491–539.

Zeijlstra, Hedde. 2014. On the uninterpretability of interpretable features. In *Minimalism and Beyond: Radicalizing the Interfaces*, edited by Peter Kosta, Steven L. Franks, Teodora Radeva-Bork, and Lilia Schürcks, vol. 11 of *Language Faculty and Beyond*, pp. 109–129. John Benjamins, Amsterdam.

Zwicky, Arnold M. 1986. Imposed versus inherent feature specifications, and other multiple feature markings. *The Indiana University Linguistics Club 20th Anniversary Volume* pp. 85–106.