Specifiers as Secondary Heads

Ronnie Cann
University of Edinburgh

1 X-bar Theory

One of the important tasks of a syntactic theory is to provide an account of the different dependency relations that hold between elements in a phrase.[1] In Dependency Grammar (following various traditional grammatical traditions (cf. Lyons 1968), there are two basic types: complement and adjunct (or modifier). The first defines an obligatory relation between a functor (or head) and its argument and the second an optional modification of a category (or head). In Categorial Grammar, these two are the only types of relation allowed by the basic theory[2], as shown in (1) where X and Y are any category labels.

(1) Categorial Grammar:[3]
(i) $X/Y + Y \Rightarrow X$ *Complement*
(ii) $X/X + X \Rightarrow X$ *Adjunct*

These two dependencies are easily reconstructed in phrase structure terms with complements introduced as (phrasal) sisters to a lexical category, both dominated by a phrasal category (2.i), and adjuncts introduced as sisters to a phrasal category, dominated by an instance of the latter (2.ii).

(2) Phrase Structure Grammar
(i) $XP \rightarrow X + YP$ *Complement*
(ii) $XP \rightarrow XP + YP$ *Adjunct*

In standard X-bar theory, a third type of dependency was introduced into phrase structure formalisms, the so-called specifier. Although interpreted in works such as Jackendoff (1977) as a linear relation to a head (precedence), the specifier relation is more commonly defined in structural terms as sister to a non-minimal head and daughter of a maximal projection (3.i). Within the standard version of X-bar theory (see, for example, Radford 1988), the complement relation (3.ii) is analysed in the same way as in the PSG rule in ((2).i), except that the dominating node is defined as a category that is neither maximal (phrasal XP) nor minimal (lexical X), i.e. X'.

(3) X-bar Theory:
(i) $XP \rightarrow X' + Y$ *Specifier*
(ii) $X' \rightarrow X + YP$ *Complement*

The introduction of the specifier relation enabled another traditional functional notion, that of the subject, to be reconstructed within Phrase Structure Grammar in a generalised (cross categorial) way. Although the latter appears to have been the principal reason for its introduction in Chomsky (1970), the relation has in fact been used in various ways since that time to encompass minor grammatical formatives, operators, escape hatches for movement, and so on.

In X bar theory, the adjunct relation has been somewhat problematic, but it has generally been assumed that a recursive schema, like that in ((2).ii) is necessary to analyse adjuncts like attributive adjectives and so on. Thus, the theory is extended to incorporate the schema in (4), where X^n ranges over levels of a category X and permissible values of n are defined for some variant of the theory.[4] This schema in effect adopts into X-bar syntax Harris' repeatable substitution equations (Harris 1951).[5]

---

1. I am grateful to all those who commented on the earlier version of this paper delivered at the Specifiers Conference at York University; to Mary E. Tait with whom some of the ideas presented here were worked out; to a number of postgraduate students, particularly Diane Nelson and Martha Robinson; to Annabel Cormack for e-mail comments on one of the drafts; and to two anonymous referees. The usual caveats apply.

2. Modifications on the basic type of functional application are found that yield different structural relations, see Steedman (1988), etc.

3. Here and below, order of elements within rule schemata is immaterial to the discussion and so '+' should be interpreted as unordered concatenation.

(4) $X^n \rightarrow X^n + YP$ *Adjunct*

These three dependency relations, complement, adjunct and specifier, have been very fruitful in providing the theoretical vocabulary for discussing the syntactic behaviour of different types of expression. However, in the drive for explanatory adequacy, there remains a strong tendency to reduce the number of structural interpretations of these three relations. This is evident in the recent work of Kayne (1994) and Chomsky (1995a,b), (basically following e.g. Henk Verkuyl (1981), Frits Stuurman (1985) and others in the 1980s). In Kayne, there are two levels of projection, maximal (XP) and minimal (X) with the two basic relations of CG, complement and adjunct. Specifiers are subsumed under a restricted concept of adjunct, as in (5).

(5) Kayne (1994):

(i) $XP \rightarrow X + YP$ *Complement*

ii $XP \rightarrow YP + XP$ *Adjunct/Specifier*

Multiple adjunction is prohibited in this theory through the operation of the *Linear Correspondence Axiom* which is taken to derive word order from hierarchical relations (asymmetric c-command). A consequence of this (controversial) move is that all movement must take place to, and through, a unique adjunct/specifier position, leading to a proliferation of (often null) functional heads to provide the necessary positions.

In Chomsky (1995a,b), all bar levels are eschewed and the specifier relation is defined by the same structural relation to the head as a complement, as in ((6).i) where *x* and *y* are labels derived from lexical expressions (this is discussed below), and *y* is the non-head variously interpreted as specifier, complement or adjunct. Pure adjuncts, on the other hand, are restricted to lexical adjunction.[6] Since multiple specifiers are not disallowed in principle (as in Kayne), phrasal adjuncts are defined by the same basic relation to the head as complements and specifiers.

(6) Chomsky (1995):

(i) $x \rightarrow x + y$[7] *Complement/Specifier/Adjunct*

ii $\langle x,x \rangle \rightarrow x + y$ *Adjunct (lexical)*

Although the basic structural definition of complementation, specification and (phrasal) adjunction is the same, the distinction between complements and specifiers/ adjuncts is defined in terms of locality to a head. A complement is the most local (least embedded) expression to a head and a specifier is anything else in the structural domain of the head. If the notions, specifier and complement, continue to have content

4. In the ‘pre-minimalist’ version of transformational grammar, n was restricted to 2 or 0, thus allowing only XP (maximal) or X (lexical) adjunction. In Chomsky’s current theory, however, n is restricted to 0 only (Chomsky 1995b). It should be noted that Chomsky does not discuss adjunction that does not result from movement (like attributive modification) and specifically denies that ‘there is any good phrase structure theory of’ such matters (Chomsky op. cit ch. 4:fn 22). It, therefore, remains to be seen whether the extreme restriction to lexical adjunction is tenable.

5. See Stuurman (1985:16-26) for a discussion of the resemblances between Harris’ and Chomsky’s proposals.

6. The notation, $\langle x,x \rangle$ in (6.ii) indicates a two segment category which is not a term of a phrase marker and so not visible at LF.

7. Chomsky uses a set theoretic notation for the structures induced by the tree-building operation ‘Merge’ so that the rule schemata in (6.i and ii) are represented as i. and ii., respectively. For the purposes of this paper, there are no consequences of the differences in notation.

i $\{x,\{x,y\}\}$

ii $\{\langle x,x \rangle,\{x,y\}\}$

(as they do in Chomsky (1995b), since only specifiers can check features), then these assumptions simply reconstruct X-bar theory. Although it does so without the postulation of extra features such as bar-levels, it still ascribes different properties to different levels constructed by the same operation, Merge. The checking operation must, therefore, be able look inside a tree, i.e. beyond the local domain, in order to ascertain whether a term, *x*, is minimal or not (since *x* as a term and *x* as a label are non-distinct). While this may not seem to be much to propose in order to maintain a purely minimal definition of structure building, it is not clear why the distance between a head and its complement/specifier should have such significance nor why only non-local expressions (specifiers) enter into checking relations. Further problems, of course, arise with the consideration of adjuncts, all specifiers according to Chomsky, but which differ in their syntactic properties. Despite the interesting attempts of Chomsky and Kayne to reduce the number of distinct X-bar relations, the differences between complements, specifiers and adjuncts remain significant and still need to be stated independently. There is as yet no reason to suppose that the differences do not reside in differences in their structural realisation.

2 Specifiers and complements

It is well known that constructions containing a specifier typically have a distribution that differs from either of its subconstituents on their own and that properties of the specifier may be selected or otherwise determined by a higher governing head (cf. e.g. Chomsky 1986). In Cann (1993), it is suggested that this observation be adopted into the grammar by allowing two expressions, a specifier and its head, to combine to form a constituent that has more grammatical properties than either expression on its own. This is done by requiring that the category of a local tree consisting of a specifier and a head be determined by the unification of the categories of its immediate subconstituents. The statement in (7) is interpreted there as forming part of the phrase structure component and has the effect of making specifiers secondary heads.

(7) If α is a specifier of β, then the category of the minimal tree, T containing α and β, is given by the unification of the categories of α and β.

If one puts this into the terms of Chomsky's 'Bare Phrase Structure' (Chomsky 1995a), we replace the definition of specifier in (6.i) with (8), (where the relation $\cup$ is to be made more precise below):

(8) $x \cup y \rightarrow x + y$ *Specifier*

Chomsky observes that the construction of the label of a mother node from the union of the labels of both its daughters is logically possible, but dismisses the idea out of hand:

> 'The label [of a phrase marker] must be constructed from the two constituents α and β. Suppose these are lexical items, each a set of features. Then the simplest assumption would be that the label is [one of]:
> ithe intersection of α and β
> iithe union of α and β
> iiione or the other of α, β
> The options [i] and [ii] are immediately excluded: the intersection of α, β will generally be irrelevant to output conditions, often null; and the union will not only be irrelevant but contradictory if α, β differ in value for some feature, the normal case.' Chomsky (1995b): 397

Chomsky does not attempt to show that the 'normal case' is, in fact, for α and β to differ in the value for some feature and thus that their union is contradictory.[8] Presumably, he would base his claim on his hypothesis that, since α and β are labels projected from lexical items, they contain phonological features whose union is likely to lead to contradiction. However, if we restrict our attention to syntactic features (I will return below to the question of phonology), then it is not true that the union of two labels will generally lead to contradiction. If, as sometimes suggested, functional categories are not inherently verbal or nominal, then incoherence will not normally arise where one of the unified labels is a functional expression.[9] For example, if the formal features of *the student* and *Agr* are as in (9a) and (9b) then the union of the two items is the non-contradictory set in (9c) (assuming for the moment that $\cup$ is to be interpreted simply as set union of formal, syntactic, features).

(9a) *the student*: {Det:def,3,sg}
(9b) *Agr*: {3,sg}
(9c) *the* $\cup$ *Agr*: {Det:def,3,sg}

Indeed, the compatibility of formal features is a sine qua non for Specifier-Head agreement (and consequent feature checking) and must hold not only of noun phrases in the specifier position of Agr (assuming the independence of this node, pace Chomsky 1995b), but also in Operator positions in order to allow the satisfaction of the Neg and WH criteria (see Rizzi 1990b and Haegemann and Zanuttini 1991).

In terms of tree building operations, then, I am suggesting a new operation for syntactically induced dependencies like those generally ascribed to specifiers. This does not involve simple substitution of a tree for a (c-selected) node, like Chomsky's Merge, but the combination of two trees (simple or complex) and the creation of a new categorial label for the mother, determined by the union of the root labels of the two trees combined by the process. This process of union may also be seen as a validation of the syntax-lexicon interface: checking in Chomsky's terms. If contentives come out of the lexicon with their morpho-syntactic features fully specified (perhaps as part of the numeration process, as suggested in Chomsky 1995b:225 ff.), and if functional categories are essentially syntactic constructs, then matching of features becomes essential to maintain congruence between the two domains. The union process can thus be viewed as ensuring compatibility between lexically and syntactically derived feature specifications.

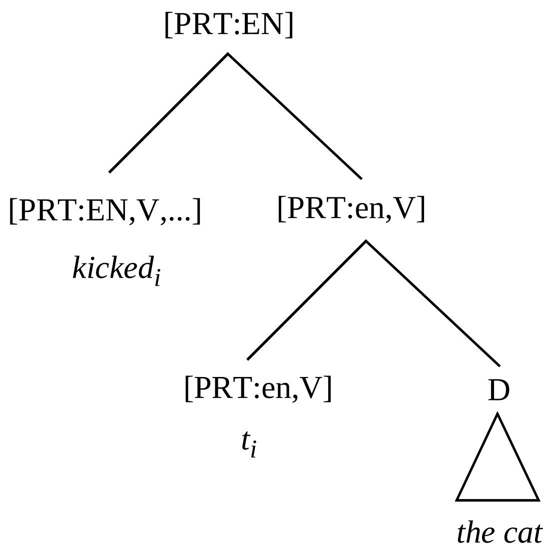
Technically, this can be quite easily achieved if we treat $\cup$, not as set union, but as category unification as commonly construed (see Pollard and Sag (1994:19) and much other literature). Assume that certain features in lexical entries, those that have syntactic significance like Tense, AGR, etc., have values that are variables which range over a restricted subset (possibly unary) of the values associated with the feature. Feature checking can then be viewed as instantiation of this variable with a particular value as determined by a particular tree configuration. Such a checking procedure will be necessary under the natural assumption that feature variables are not interpretable at LF.

Let us consider first feature checking through head movement, taking as example,

8. Notice that Chomsky is assuming that it is not the simple union of feature sets that is ill-formed, but that the resulting set contains instances of the same feature with different values: an incoherent category, rather than an incoherent set. See Gazdar et al. (1985) for discussion of such matters.
9. Even if all labels carry N and V features, however, it will often arise that the union of two labels will not be incoherent.

the movement of a verb to a participle head position. Following Cann and Tait (1995), participles in English may be analysed as realisations of a (language specific) functional category which we may label PRT. This has two variants which can be encoded by letting PRT take one of two values, EN (for the passive/perfect participle) and ING (for the present participle).[10] A participle verb form, like *kicked*, may then be analysed as including as part of its label, the feature PRT with a variable value, which we may represent as *en*, that ranges over only the single passive/perfect value of PRT, EN. This variable value of PRT on the verb must be instantiated during the derivation which can be achieved directly through head movement if this operation does not create a lexical adjunction structure like that in ((6)ii), but unifies the features of the target node with that of the moved element. Thus, abstracting away from irrelevant details, the movement of *kicked* to the participle head involves the creation of structures like that in (10) which shows the way the variable value *en* associated with the label of the verb is identified with the given value EN of the functional head. In this way, morphological features are straightforwardly checked against syntactic ones.

(10)



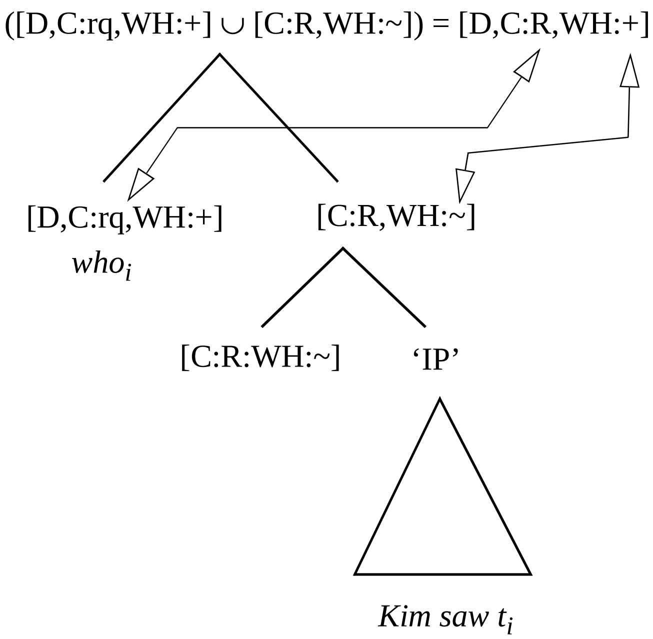
Exactly the same process of feature instantiation can be used to account for feature checking in typical specifier constructions, reinterpreted here as doubly headed structures formed through the unification (as opposed to union) of the labels of the two heads.[11] The creation of tree structures involving the process envisaged in (8) will thus instantiate variable feature values. Consider WH-movement, for example. To ensure that WH elements move to the complementizer position within the Minimalist framework, it is necessary to assume that a WH phrase contains a C feature that must be checked (pre- or post- Spell Out depending on strength). Let us assume (following suggestions in Gazdar et al. (1985) and Pollard and Sag (1994), etc.) that C is a feature that may take number of values, depending on whether the clause is a question (Q), relative (R), etc. As part of the label of the WH expression, the feature C takes a varia-

10. See the next section for further discussion of this.

11. Note that the process must be unification and not union here since [PRT:en] and [PRT:EN] are different formal objects whose union would be {[PRT:EN],[PRT:en}} but whose unification is [PRT:EN].

ble value, *rq*, ranging over (at least) the two values, Q and R,[12] which will need to be instantiated during the course of a derivation to ensure convergence. The WH expression moves and combines with the complementizer clause that contains it. The latter may itself contain a variable feature value for WH ranging over + and - (to distinguish constituent questions from root questions and WH relatives from those with just a complementizer)[13]. As part of the combination of the two trees the formal features of the labels at their roots are unified, instantiating the variable feature values with fully specified ones, as shown in (11), where the arrows relate the instantiated feature values with the appropriate variables (irrelevant details are omitted or simplified, so whatever is under C is just labelled 'IP' for convenience).[14]

(11)



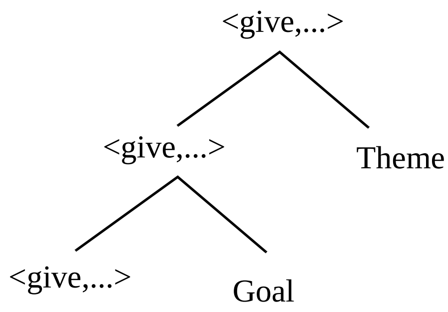
Thus, feature checking can be specified as a single process: the instantiation of variable feature values through unification. There are, however, two operations that involve it: substitution head movement, interpreted as unification of one lexical node with another[15]; and phrasal movement, interpreted as the 'specifier' tree-building operation in (8), for the validation of syntactic dependencies. In both cases, however, the moved element becomes a head, singulary in the case of head movement, and secondary in phrasal movement.

Returning to Chomsky's headed schema in (6.i), repeated in (12a), we may now interpret the structures induced by this in a rather different way to produce a coherent

12. Some WH expressions, like *what* may be lexically specified as taking a more restricted value for C which ranges only over Q, thus disallowing it from appearing as a relative marker. There are, however, dialects of English that allow relative *what* so that this property does seem to be truly lexical and so open to dialectal variation.
13. This could be seen as an analogue of Cormack's double selection of C for its specifier and WH for C (see Cormack this volume). This is not a necessary assumption for my purposes.
14. In fact, the tree in (10) would never appear as such, since the process of unification instantiates all instances of the relevant variables throughout the tree.

difference between complements and other expressions. In Chomsky's schema, the head is a unique lexical element and so structures based on it may be considered to define the lexical domain of that head, i.e. its θ-marking and c-selectional domain (cf. Grimshaw 1990). Hence, a lexical entry like that in (12b) will induce a tree structure like that in (12c) which may be construed as the syntactic realisation of the lexical structure of the verb in that independent properties of the lexical item (like θ-roles) are mapped into nodes (merge sites) with the head determining part of the label of each lexically determined dependency. Functional categories just project their c-selected complements.[16]

(12a) $x \rightarrow x + y$ Lexical Projection
(12b) {{give,V,<Goal,Theme,Agent>}
(12c)



Where a tree is determined by lexical properties of a head, the tree building operation, Merge, may thus be construed as simple substitution of a non-head node by another tree, as in Chomsky (1995b), with concomitant unifying of node labels to guarantee that lexical properties are maintained (the remnant of the projection principle of earlier work).

3 Selecting specifiers: the passive/perfect participle

Specifiers may appear not just to check features, however. They may be introduced to satisfy lexical dependencies that would not otherwise be satisfied. In Cann and Tait (1995), the treatment of specifiers as resulting from the unification of two heads is used to provide a new account of the passive and perfect constructions in English. They take as their point of departure the hypothesis that, for grammatical elements, morphological identity entails syntactic (categorial) identity. For example, it is com-

---

15. I am here making no claims that all head movement is substitution. It may be that lexical adjunction remains a necessary operation within the grammar. For example, in the theory espoused below, the analysis of Subject-Auxiliary Inversion as involving I to C movement is not tenable if the movement is substitution, at least in WH-questions, since this would exclude the fronting of a plural WH-phrase with a singular auxiliary (e.g. *Which students does Mary like?*). It may be that I adjoins to C in these cases to allow number (and person) mismatches between the WH-specifier and auxiliary, or that the auxiliary moves to a position between I and C. However, whatever the correct account, the current theory predicts that the movement of auxiliary in these cases does not involve checking of a WH-feature with the WH-expression and that other factors must be involved to derive such movement. This is a matter for further research.
16. One could use Cormack (this volume)'s categorial notation to encode c-selectional properties of both contentive and functional expressions. This sort of lexical projection is also reminiscent of HPSG, see Pollard and Sag (1994). See also Tait (1991) for a discussion of lexically induced trees and Fukui (1986) for arguments that contentive categories realise arguments within X' domains.

mon in syntactic theory to treat the perfect and passive participles in English as the realisation of separate morphemes, despite the fact that there are no differences in the morphological forms of verbs, even down to suppletion, depending on whether they appear in passive or perfect constructions.[17] In other words, passive and perfect participles are treated as homonyms, morpho-phonologically identical realisations of different morphemes. However, there is no a priori reason to suppose that the use of the single form in both constructions in English is purely accidental nor that the labels traditionally given to these syntactic constructions by linguists are anything but a terminological convenience.[18] It is at least an interesting hypothesis that minimalist expectations apply also in the functional lexicon and that homonymy of grammatical elements is not tolerated.[19] Cann and Tait, therefore, analyse both the passive and the perfect participles as projections of a single category, which they label *en* after Chomsky (1957). It follows from this assumption that the differences in the distribution (and interpretation) of *en*-participles in English must derive from the expressions with which they combine.

The data in ((13a)-(13h)) illustrate some of the different contexts in which the perfect/passive participles can appear in English.

(13a) Harassed by the students, the lecturer finally left.
(13b) *Be harassed by the students, the lecturer finally left.
(13c) The lecturer was harassed by the students.
(13d) The students have harassed the lecturer.
(13e) *The lecturer was harassed the students.
(13f) *The students have harassed by the lecturer.
(13g) *The lecturer might harassed by the students.
(13h) *The students might harassed the lecturer.

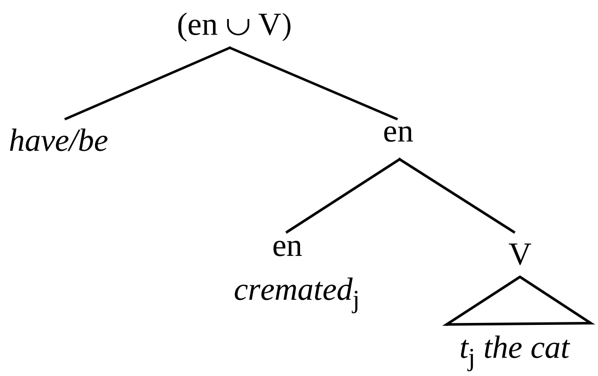
In (13a), the participle is in an adjectival context (cf. *Angry with the students, the lecturer finally left*), where no auxiliary can appear (13b); in (13c)-(13d), the participle appears in verbal contexts, passive in (13c) and in perfect (13d); (13e) and (13f) illustrate that the syntactic (case assigning) properties of the participle are dependent on which auxiliary appears; and (13g) and (13h) show that the auxiliary verbs are obligatory in situations where there is an element (like tense, modal or another instance of the participle) that requires a following verb.

Cann and Tait explain these differences as deriving from an analysis in which the auxiliaries appear as specifiers to syncategorematic *en* in the syntax where specifiers unify their properties with heads as suggested above. Revising the structures they suggest in the light of Chomsky (1995a), the basic structure of the perfect and passive constructions, differing only in which auxiliary is chosen, is as given in (14).[20]

---

17. Warner (1993) uses a feature +EN to identify the past (sic) and passive participles, but still differentiates them by the use of the feature ±PRD (predicative), the first being -PRD and the second +PRD.
18. Note that the claim made here is strictly with reference to (certain dialects of) English. No claim is being made that the passive and perfect morphemes are universally identical. Indeed, most languages that have both passive and perfect do differentiate them morphologically in which case the constructions will necessarily be distinct in analysis. Thus, in the discussion that follows, the specifics of the analysis pertain only to English, while only the general mechanisms are considered to be universal in any meaningful sense.
19. There is strong psycholinguistic evidence for this hypothesis with respect to functional elements. See Cann (1996) for discussion.

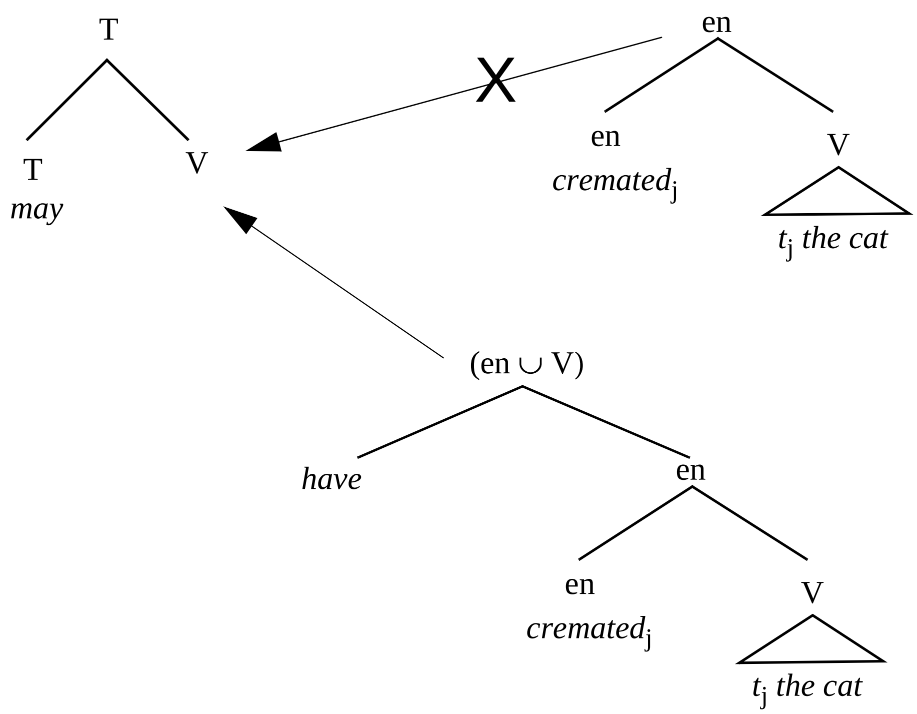
(14) [21]



The grammaticality of participial expressions without an auxiliary as in (13a) shows that the auxiliary does not appear in (14) to satisfy any of the properties of the participial head: it is neither selected by *en* nor checks any of its features. It must, therefore, by economy principles, appear in order to satisfy the requirements of some other element. Apart from feature checking, the other principal mechanism that drives minimalist derivations is selection, determined by the merging of one tree with a node lexically projected from the label of some head, as noted above. For selection to have significance, any property of the merge site must be a property of the root node of the tree to be merged. In other words, if a head selects a verb, then the root node of any tree that satisfies this selectional property must be verbal.[22] Since the *en* morpheme (PRT:EN) is syncategorematic, it has no verbal features (or only incompletely specified features, see note 20.) and so cannot satisfy the selectional properties of a tense or modal node. This automatically accounts for the ungrammaticality of (13g)-(13h) where no auxiliary appears in such contexts. However, the appearance of an auxiliary to give the structure in (14) means that the root now has a verbal specification (through the unification of the properties of the auxiliary with those of the participial head) which can locally satisfy the selection property of tense/modal, as shown in (15).

20. Where (en ∪ V) may be construed as [PRT:EN,+V,-N]. Note that if *en* has any major features, then this may only be [+V] with N undefined to allow for the participle to appear in both verbal and adjectival contexts.
21. In this and the following examples, I have not shown the original position of subjects in order not to commit to, or deny, the VP Internal Subject Hypothesis.
22. Thus, the relation between the categories of the root and merge sites is one of extension as defined in Gazdar et al. (1985:39).

(15)



The grammaticality of (13a) thus follows from the fact that the participle is in a position where a verb is selected, while the ungrammaticality of (13b) results from the lack of any licenser for the auxiliary.

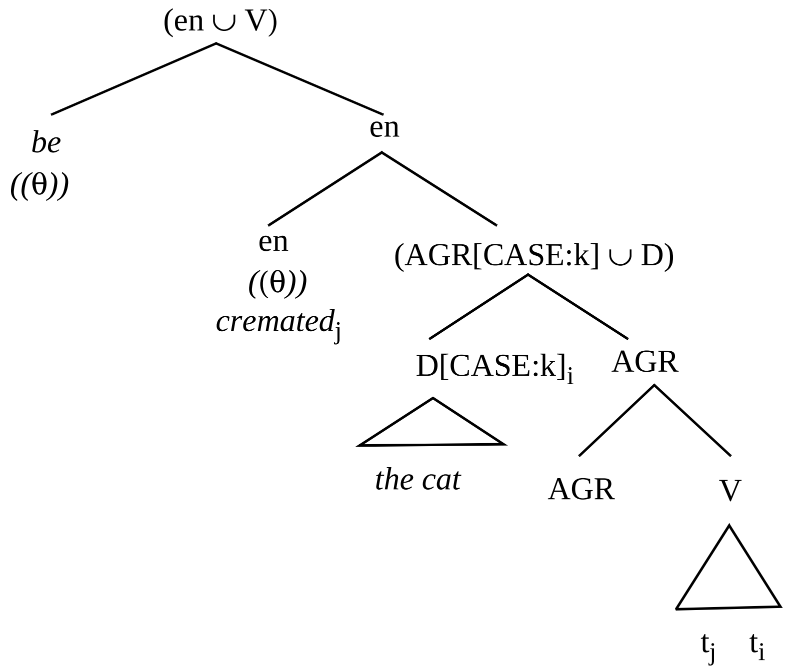
The different case-assigning properties of the two constructions (illustrated in (13c)-(13f)) are fairly straightforwardly derived under the current analysis. Burzio's generalisation states that a verb assigns object case if, and only if, it assigns an external theta-role (Burzio 1986). Within the current framework, we may interpret this statement as providing a means of licensing an expressed case value on some functional head, which I shall here take to be AGR.[23] Thus, I assume that certain functional heads will normally be generated with certain features having a variable value which can only be instantiated as a non-variable value when something else licenses that value. While we may assume that person and number are inherent features of AGR (or D), and that such features have values instantiated as part of the numeration, the case value is dependent on the position of the node in the tree. Nominative (subject) AGR gets its case value by virtue of its selection of Tense as a complement. Accusative (object) AGR, however, requires something else to determine its case value which, by Burzio's generalisation, is the ability of an immediately dominating element to assign an external theta-role.

Hence, as the functional head, *en,* does not itself assign an external theta-role,[24] it does not license an accusative case value on AGR in its complement position. When V moves to *en*, its external theta-role is absorbed (see Haegemann 1991 and references

23. Unlike Chomsky (1995b), it is necessary in the current framework for case to be checked by an independent functional head. This is because a DP specifier of a VP is impossible, since the unification of these two categories gives rise to an incoherent category (one that is both nominal and verbal).
24. This stance differs from that taken in Cann and Tait (1995).

cited therein, and, with very different assumptions leading to the same effect, Cann 1995) and again an accusative case value fails to be licensed on AGR. This situation remains constant when the specifier is *be*, because this expression, too, lacks an external theta-role. We may assume, therefore, that where an *en* participial phrase appears without an accompanying auxiliary verb or with *be* in its specifier position, the case value on AGR remains as a variable (signalled below as k, a variable value ranging over any grammatical case).[25] For the variable case value on the object DP to be instantiated, therefore, it must move to some appropriate position which contains a proper non-variable value for the feature CASE, as illustrated in (16).

(16)

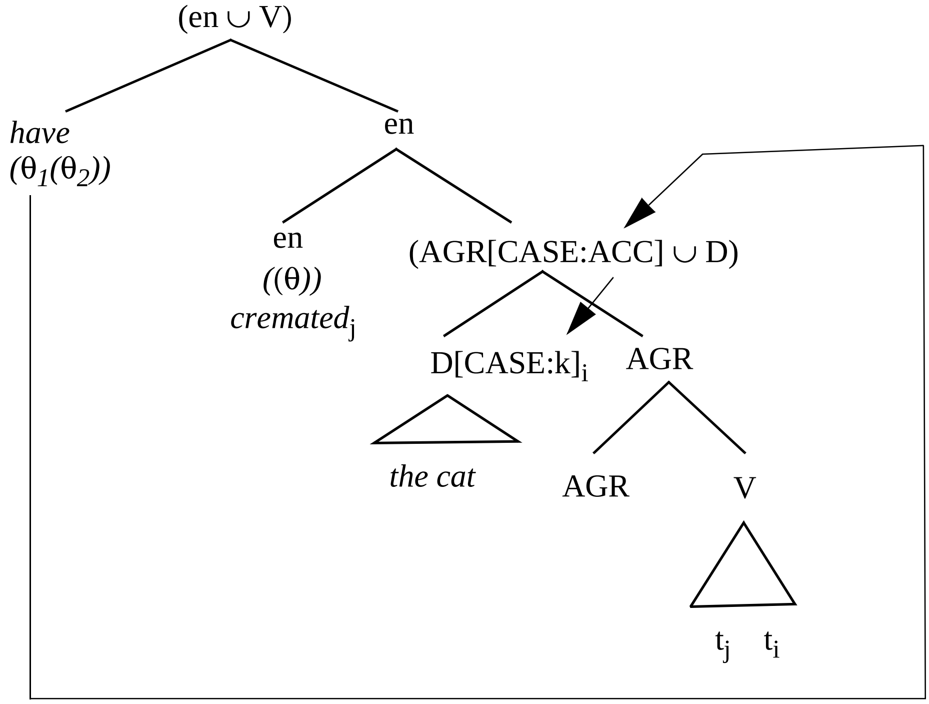


On the other hand, *have* does assign an external theta-role (on the assumption that perfective *have* is the same as main verb *have*, see section 4). Although *have* itself does not immediately dominate the AGR projection, its properties are shared with the node that does (*en*) through the unification operation. Hence, an accusative case value is licensed on AGR. Since the object noun phrase can instantiate (check) its case variable in the specifier position of AGR, it will move no further, as shown in (17).[26]

25. It is possible, of course, that it is the AGR projection itself that is licensed by the external theta-role, in which case it will simply not appear in the stated structures. The effect will be the same.

26. The structure in (17) embodies a number of further assumptions that I do not have space to go into here. The principal one of these is that head movement takes place into the nearest available compatible position, as in Rizzi (1990a). V cannot move into AGR in (17) since its feature specification would not unify with that of the DP is the specifier position. Another aspect of this analysis that is not discussed here is the way the external theta-role of the verb is identified with the subject. This matter is discussed in Cann (1995).

(17)



4 Double versus single headed structures: causative and perfect *have*

The analysis of auxiliaries as specifiers interpreted as secondary heads has further advantages in analysing other constructions involving the verb *have* in English. As is well-known, this verb appears in a wide range of constructions that, from a semantic point of view, can be divided up into contentive (main verb) uses (18a) and functional (grammatical) ones (18b).[27]

(18a) Contentive:	
(i) Possessive:	*Jo has three books*.
(ii) Relational:	*Jo had a sister once*.
(iii)Inalienable Possession:	*Jo has blue eyes*.
(iv)Experiencer:	*Jo had a headache*.
(v) Process:	*Jo had the new boy last night*.
(18b) Functional:	
(i) Perfect:	*Jo has gone home*.
(ii) Causative:	*Jo had the cat cremated*.
(iii)Modal:	*Jo has to go home*.

One of the interesting facts about this verb, however, is that not all semantically functional instances of *have* behave as auxiliaries and not all uses of the contentive behave as main verbs. Table 1: shows the properties of five of the different *have* constructions (two contentive and three functional) with respect to four normal tests for auxiliaryhood: Subject-Auxiliary Inversion, *n't* cliticisation, the disallowance of the 'dummy' verb *do* and the ability of the verb to cliticise to the preceding word. These are illustrated for the perfect in ((19a)-(19d)). [28]

27. For discussion of properties of *have* and different ways to analyse them, cf. Ritter and Rosen (1997), Cowper (1992), Kayne (1993), *inter alia multa*.

(19a) Has Jo cremated the cat?
(19b) Jo hasn't cremated the cat.
(19c) *Jo does have cremated the cat.
(19d) Jo's cremated the cat.

Table 1: Auxiliary properties of *have* constructions

	SAI	n't	~do	Clitic
Possessive	+	+	-	+
Process	-	-	-	-
Perfect	+	+	+	+
Modal	%	%	-	?-
Causative	-	-	-	-

The pattern in Table 1 shows the cline in auxiliaryhood shown in (20) from the main verb properties of the process and causative to the most auxiliary-like behaviour of the perfect. The modal construction appears in the middle and is most subject to idiolectal variation.

(20) Process/Causative > Modal > Possessive > Perfect

As with the perfect/passive participles, this variability in the properties of constructions based on a single morph might be analysed by postulating homonymy, i.e. distinct lexical entries for the different uses of the verb. However, such an approach fails to explain the apparent binary distinction between the constructions (ignoring the modal construction for convenience), i.e. an apparently straight Main Verb/Auxiliary split, and fails to account for why this distinction should cut across the contentive/ functional divide.

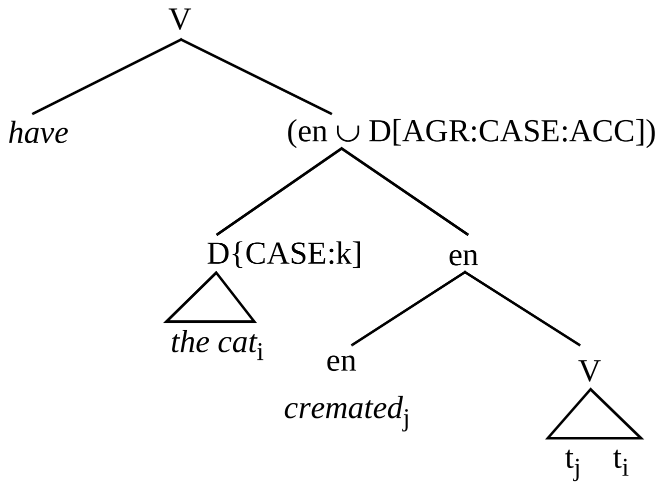
The observation of the syntactic differences in *have* constructions shows an interesting link with studies of grammaticalization processes, as described, for example, in Hopper and Traugott (1993). Grammaticalization is analysed by Hopper and Traugott as the development of a grammatical item ('bleached' of its contentive interpretation) from a single (contentive) expression via a polysemous stage where emerging grammatical and contentive uses are active side-by-side through a process of pragmatic enrichment. Given that the development of auxiliaries in this framework is analysed as the grammaticalization of main verbs, the variability shown above indicates that *have* is currently in the middle, polysemic, period in English. If this is the case, an explanation needs to be found for why the different polysemes should have different syntactic properties, since normally a polysemous item maintains its syntactic properties in its different interpretations.

If *have* is treated as a single polysemous expression and morphemes are treated monadically in the syntax, then the differences in the syntactic properties of the constructions must be explained in terms of the structures they induce. The structure for

28. The % for the modal use indicates variation in acceptability of such examples as (British English) *Has Lou to go home?*, *Lou hasn't to go home* versus the much more marginal *?Lou's to go home*. (All of these are grammatical in the author's idiolect.)

perfective *have* was given in (14) above, in which the auxiliary verb acts as a second head with the participle, unifying its properties with those of the participial head. In this construction, *have* exhibits strictly auxiliary behaviour, as indicated in Table 1. Causative *have*, on the other hand, exhibits only main verb characteristics: it does not allow SAI (**Had Jo the cat cremated*) or *n't* cliticisation (**Jo hadn't the cat cremated*); cannot cliticise to the subject (**Jo'd the cat cremated*); but does permit 'do-support' (*Jo did cremate the cat*). Under the assumption that main verbs, which all display the same behaviour, are analysed as singulary heads, *have* in its causative manifestation may also be analysed as an ordinary head taking a participial complement, as shown in (21).

(21)



This structure differs significantly from that in (14). In the first place, *have* is an independent element heading its own, uniquely determined, projection and thus it requires its c-selectional properties to be satisfied (taken here to be D, something that can be assigned an appropriate θ-role). The *en* complement alone cannot satisfy this requirement and so before Merge creates the structure in (21) a noun phrase must unify with it to yield a tree rooted in (D ∪ en) which is of the appropriate type (D) to satisfy the selectional property of *have*.[29] Since *have* does not unify its case-assigning properties with the participle in this construction, the direct object is not case-licensed in situ and so must move. The nearest position in which the case of the object can be checked is the specifier position of the participle, to which the object raises as observed.[30] [31]

---

29. The causative meaning is maintained if the participial form is *ing* as in *Jo had the cat howling*, *Jo had the cat eating smoked tofu*, which lends credence to the idea that the participle isn't the 'real' (selected) complement of the verb.

30. A full AGR projection is not shown in (21), but if it were there as in (i), the analysis would not be materially affected, since the object DP would still have to be fronted before the participle to have its case checked.

$[_V \text{ have } [_{(AGR \cup D)} [_D \text{ the cat}]_j \text{ AGR } [_{(en \cup D)} t_j' \, [_{en} \text{ cremated}] \, [_V \, t_i \, t_j \,]]]]$

31. Notice that the satisfaction of the selecting head, *have*, by a moved noun phrase requires that Merge has to apply to the tree $[_{en} [_{[en,V]}$ *cremated*$]\ [_{VP}$ t $[_{DP}$ *the cat*$]]]$ to raise the DP prior to the combination of the participle with the verb. Otherwise, selectional properties could be satisfied by covert (F-)movement after Spell Out. It must therefore be the case that Merge immediately satisfies lexical dependencies, as noted above. Since on this account movement must take place before the satisfaction of lexical selection, this analysis provides a further argument against the significance of d-structure as a level of representation.

No attempt can be made here to give a full account of the observed syntactic differences between *have* instantiated as an independent head, as in the causative, or as a second head, as in the perfect. However, the analysis sketched above suggests that overt movement of V to C in English - and perhaps elsewhere - involves the movement of auxiliary V through the *specifier* of T rather than through the head position.

In the doubly headed structures created by the unification operation, the status of both subconstituents remains maximal, under the assumption that movement applies only to maximal or minimal elements. The ‘specifier’ expression must be maximal, since it may be moved from such a position in Topicalisation and WH-constructions, etc., e.g. (22a). On the other hand, its sister constituent, the ‘head’ expression, is also maximal, since, for example the participle can be topicalised as in (22b). This follows from the fact that $x \cup y$ is not equivalent to either x or y, i.e. it is not strictly a projection of either one of them.

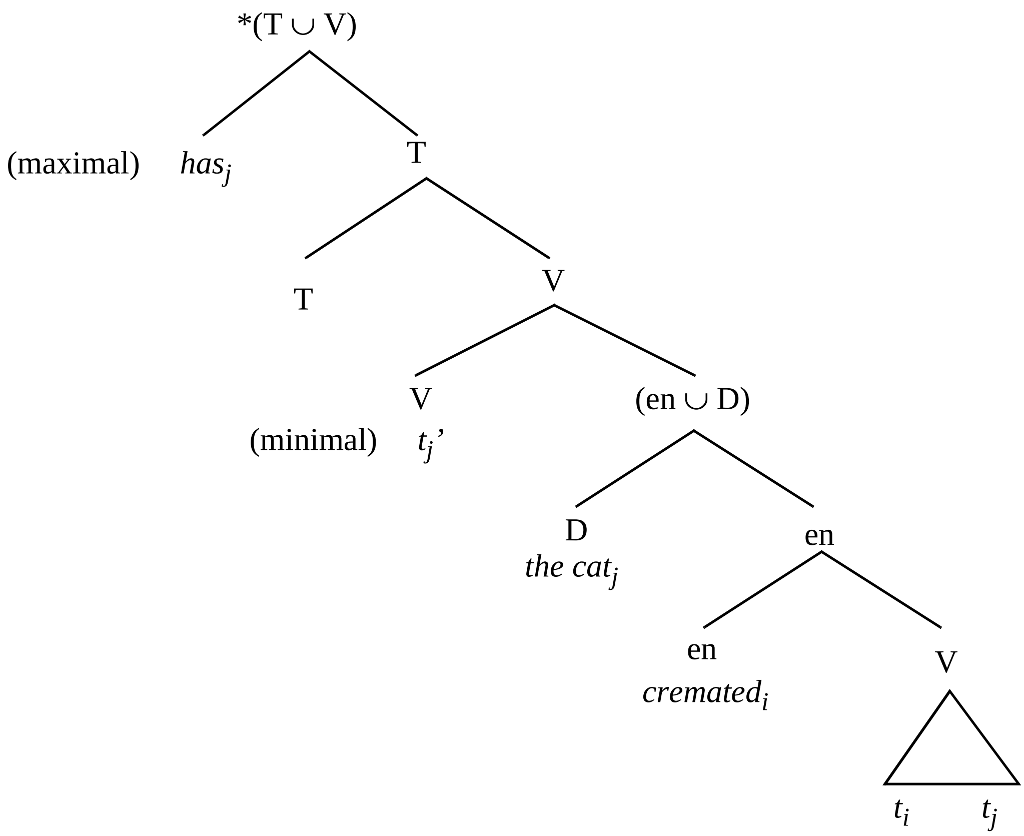
(22a) $[_{WH}$ which books$]_i$, did Jo think $[_{(WH \cup C)}$ t_j C [he lost t_j]]
(22b) $[_{en}$ cremated the cat$]_i$, Jo thought he $[_V$ had$]_j$ $[_{(V \cup en)}$ t_j t_i]]

The maximality of both constituents in the specifier construction accounts directly for why main verbs in English cannot appear in Comp (or preceding adjuncts, etc.). If the Uniform Chain Condition of Chomsky (1995b) is valid, then (23b) is correctly predicted to be ungrammatical, since the chain (t_j,t_j’) involves a maximal head and a minimal tail (as more clearly shown in the tree (11)).[32] (23a), on the other hand, involves a chain that is uniformly maximal in Chomsky’s terms and so the construction is grammatical. It could therefore be a matter of parametrisation whether verb movement occurs through the specifier or the head of the Tense projection, leading to differences in the ability of verbs to front.

(23a) $[_C$ $[_C$ $[_V$ has$]_j$ C] Jo $[_{(T \cup V)}$ $\mathbf{t_j}$ T $[_{(en \cup V)}$ $\mathbf{t_j'}$ cremated the cat]]
(23b) *$[_C$ $[_C$ $[_V$ has$]_j$ C] Jo $[_{(T \cup V)}$ $\mathbf{t_j}$ T $[_V$ $\mathbf{t_j'}$ the cat cremated]]

32. Covert F-movement of the tense feature of a main verb may then be analysed as movement to the head of T. It is possible that all covert movement should be analysed as F-movement to a head position. In the theory of this paper, since features are checked in the same way for both phrasal and head movement, no further operations need be defined and no extra structure need be created post Spell Out.

(23c)



Auxiliary cliticisation is also explicable in these terms. In this case, whatever syntactic position is correct for an auxiliary clitic, it is, under all accounts a maximal one, and so movement of the auxiliary from the secondary head position in T to this position does not violate uniformity, as illustrated in (24a). In the case of *have* in singulary head position, where the tail is minimal, the UCC predicts ungrammaticality in the same way as for (23b), as shown in (24b).

(24a) $[_{D}\ [_{D}\ \text{Jo}]\ \text{'d}_{j}]\ [_{(T \cup V)}\ t_{j}\ T\ [_{(en \cup V)}\ t_{j}\text{'}\ \text{cremated the cat}]]$
(24b) $*[_{D}\ [_{D}\ \text{Jo}]\ \text{'d}_{j}]\ [_{(T \cup V)}\ t_{j}\ T\ [_{V}\ t_{j}\text{'}\ [_{(en \cup D)}\ \text{the cat cremated}]]$

There are a number of ways in which examples with negative clitics can be analysed, depending on one’s general treatment of negation in the syntax. Because of the restriction of this process to finite contexts (**Kim wants to-n’t go*), it is clear that *n’t* must be associated with tense in some way. One possible analysis appears in (25a).[33] This adopts the analysis of Gazdar, Pullum and Sag (1982), where it is argued that verbs are specified lexically for the clitic, on the grounds of lexical variation (like *won’t* for **willn’t*, *ain’t* for %*amn’t*, etc.), and is consistent with the arguments in Zwicky and Pullum (1983) concerning the treatment of *n’t* as an inflection rather than a clitic. In the analysis here, the verb contains a variable NEG value that needs to be checked against the NEG head, before movement to the T projection to check tense. The resulting chain is uniformly maximal, unlike that formed by the movement in the causative example in (25b) where the tail of the chain is minimal and its head is maximal, as in the analyses presented above.

(25a) $[_D$ Jo] $[_{(T \cup V)}$ hadn’t_j T $[_{(NEG \cup V)}$ t_j NEG $[_{(en \cup V)}$ t_j’ cremated the cat]]]
(25b) *$[_D$ Jo] $[_{(T \cup V)}$ hadn’t_j T $[_{(NEG \cup V)}$ t_j NEG $[_V$ t_j’ $[_{(en \cup D)}$ the cat cremated]]]]

Whether or not the suggestions above for analysis of SAI, auxiliary or *n’t* cliticisation are the best that can be made for English is not important for the current argument, however. What is significant is that the two exemplar constructions, causative and perfect, involve two different structural realisations. This means that some syntactic account of their differing properties can be given without assuming that there are two (or more) different verbs, *have*, or two different morphemes, Passive and Perfect.

5 Grammaticalization

In addition to providing a structural difference between main verb and auxiliary-type uses of *have* in English, the hypothesis that certain verbs may appear as secondary heads (specifiers), provides a way to account for the diachronic process of grammaticalization. (26a) shows the hypothesised grammaticalization of verbs according to Hopper and Traugott (op cit). This is clearly a more articulated development than that implied by current formal ideas about structure which only provide two contrasts: between contentive and functional heads (26b). The current hypothesis, however, provides a third position in the development midway between the two heads, the specifier of a functional head, (26c).

(26a) Full Verb > Auxiliary > Clitic > Affix
(26b) C-Head > F-Head
(26c) C-Head > Specifier > F-Head

We have seen above how the development of auxiliaries as clitics is accounted for in the theory presented in this paper, but the syntactic development into an inflection/ functional head also follows naturally from the analysis of dependent auxiliaries as secondary heads. A specifier associated with a particular functional head (with a particular meaning) is dissociated from other instances of the original verb through phonological reduction, etc. At some point, the specifier merges completely with the functional head with which it combines to give a ‘composite’ category, maintaining a semantics based on the earlier specifier phase, but now ‘bleached’ of any information independent of the construction. An example that can be analysed in this way is the development of the Romance future from the Latin periphrasis *habere* plus infinitive construction. This is discussed in Hopper and Traugott (1993)[34] in terms of the reanalysis of an independent infinitival complement as part of the verb complex involving *habeo* ‘have’ with subsequent morphological changes, (27) (Hopper and Traugott 1993:44, example 13).

---

33. Other analyses are possible. For example, (i) gives an analysis in which the negative clitic undergoes head movement from NEG to T with movement of the auxiliary to the secondary head (specifier) position of T. The analysis in (ii) treats the negative clitic as a secondary head of NEG which moves to become a secondary head of T. The auxiliary then moves to a second specifier position (yielding a construction with three heads). Only the analysis in (25a), however, provides a straightforward treatment of fronted negative auxiliaries, since only in this analysis do the auxiliary and the negative clitic form a constituent. (But see section 6 for a suggestion that would solve this in the phonology.)

i $[_D$ $[_D$ Jo]] $[_{(T \cup V)}$ had_j $[_T$ n’t_i] $[_{NEG}$ t_i $[_{(en \cup V)}$ t_j’ cremated the cat]]]

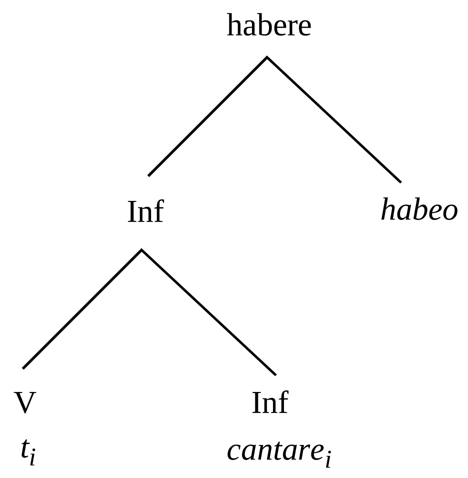
ii $[_D$ $[_D$ Jo]] $[_{((T \cup NEG) \cup V)}$ had_j $[_{(T \cup NEG)}$ n’t_i T $[_{NEG}$ t_i NEG $[_{(en \cup V)}$ t_j’ cremated the cat]]]]

34. And see Roberts (1993) for an analysis within the Principles and Parameters framework that differs substantially from that proposed below.

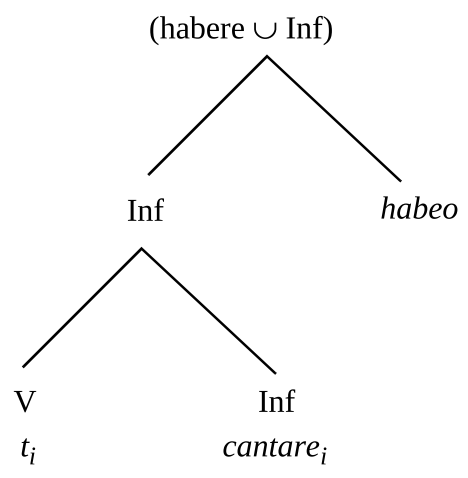
(27) Classical Latin [[cantare] habeo] >
Late Latin [cantare habeo] >
French: [chant -er-ai]

In the current theory, this reanalysis has a natural reconstruction in terms of the development noted in (26c). Simplifying and abstracting away from complex functional systems, the independent Latin verb *habeo* plus infinitival complement in (28a) is reinterpreted in later Latin as a doubly headed verb plus infinitive construction (28b). In the final stage, (28c), the properties of the verb and those of the infinitival projection have fully merged to give a single inflectional category (shown as *habere-Inf* below) interpreted as future tense.

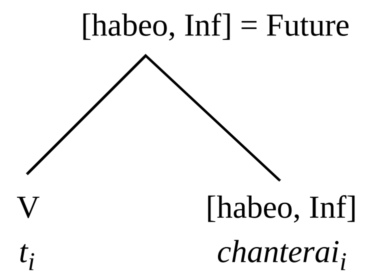
(28a) Classical Latin:



(28b) Late Latin:



(28c) French:



Analysing specifiers as secondary heads, therefore, not only provides a means of maintaining a minimal grammatical lexicon of one morph/one meaning, but provides a theoretical means of accounting straightforwardly for the process of grammaticalization.

6 Unifying labels

To end, I briefly return to the concept of unification as it applies to syntactic labels. As mentioned above, Chomsky rejects the idea that $x \cup y$ is coherent where x and y are lexical labels. In the discussion above, the operation was taken to be unification of formal syntactic features, thus excluding the semantic and phonological information that also comprise a lexical label. Treating $x \cup y$ in this way already puts some constraint on specifiers: only categorially compatible elements may appear in doubly headed (specifier-head) constructions. This is sufficient to rule out, for example, noun phrases unifying with verbal heads (all arguments must therefore be complements) or adjectives with nouns. Indeed, except for expressions of the same major category, the operation $x \cup y$ will tend to require one of x, y to be a functional category of a compatible sort.

Consideration of phonological information reinforces this restriction, given certain assumptions. Chomsky (1995b) takes the position that inflectional functional categories do not themselves contain phonological material (as was assumed in earlier work, e.g. Baker 1988), but that contentives appear in a numeration either fully morphologically determined (for number, tense, person, etc.) or with sufficient information that their morphological form can be determined by some morphological component. Assuming that phonological information is in the domain of the unification operation (as it should be if labels are determined by lexical expressions), and that the unification of the phonological matrices of two expressions is incoherent, then $x \cup y$ will be coherent if (and only if) one of x or y contains no phonological information, i.e. if one of the expressions is phonetically null. The constraint in English against doubly filled Comp positions follows automatically from this assumption.[35]

This will again tend to favour combinations where one of the combining constituents is an inflectional functional category (normally phonologically null). This would, however, rule out both of the alternative analyses of *n't* cliticisation in footnote 33., because a phonologically realised auxiliary appears with a phonologically realised

35. Where languages allow the equivalent of expressions like **the students who that I saw*, it further follows that the WH expression cannot combine with the complementizer.

negation element, a conclusion that would support the analysis in the text. However, it could be that certain elements phonologically specify a position that can be unified with the phonology of another element. For example, the negative clitic could have a phonological structure /ϕ+nt/, where ϕ is a phonological variable. This variable would then be instantiated as part of the unification with an auxiliary verb, i.e. /hæd/ ∪ /ϕ+nt/ gives /hædᵊnt/. This matter must be left open here, but whether or not phonological unification is allowed, it is likely to be very restricted (perhaps just to clitics) and is unlikely to be a property associated with contentives.

Finally, there is the status of the semantic features of a label to be considered with respect to the unification operation. The hypothesis that there is only one verb *have* in the causative and perfect constructions requires an account of the different interpretations of the two constructions, derived from some basic meaning of the item, possibly with pragmatic enrichment as suggested by Hopper and Traugott. The assumption that ∪ applies to the semantic structure of a label as well as its syntax and phonology provides the necessary means to achieve this. In other words, in x ∪ y the argument structures of two expressions merge into a single object with shared arguments and combined properties (cf. Grimshaw and Mester 1988). For the causative/perfect distinction, the difference in interpretation is thus due to the different way in which *have* combines with the participial phrase. In the causative, the participial (plus DP) phrase functions as the full argument of *have* whereas in the perfect, the argument structures of the verb and the participle merge so that the internal argument of *have* is not the whole participial phrase, but the verb itself. Cann (1995) shows that this distinction is sufficient to derive the basic semantic interpretation of the causative and the perfect constructions in English based on a single interpretation for both *have* and the participle morpheme, *en*. In this paper, a minimal semantics is assigned to *have* that captures the facts that: the event denoted by an expression involving the verb depends on its complement (*Jo has a headache* state, *Jo has a party* activity); the external argument of *have* has an underspecified theta-role which is dependent on the semantic content of the Theme (*Jo has a headache* Experiencer, *Jo had a baby* Agent, *Jo has a table* Possessor); and that the external argument bears a locative relation to the theme (cf. Kayne (1993), etc.). The participle head, *en*, is interpreted as denoting a state and having an internal event argument (associated with the verbal stem) that is contiguous with that state (and so may be pragmatically interpreted as cause of, or temporally prior to, the state).

In the causative construction, e.g. *Jo had the cat cremated*, the semantics of *have* and *en* are independent and the participial phrase is interpreted as the Theme of the verb. The event denoted by the *have* phrase and the thematic role of the subject is thus determined by the semantics of the participial phrase. The latter denotes a state resulting from the activity of someone's cremating the cat and gives rise to the interpretation of the whole sentence as involving some independent event that leads to a state of a cremated cat. The least marked relation between this event and the state is one of causation and we get the intended reading with an interpretation of the external argument of *have* as the Agent of the causing event.

In the perfect, however, the semantic structure of the participle head and the verb, *have*, are unified, so that the internal argument of *have* is identified as that of the participle, i.e. the event denoted by the verbal stem, and the state denoted by the participle head is identified as the eventuality denoted by the verb. This means, firstly, that there are just two eventualities in the perfect construction, not three as in the causative: the eventuality associated with the verb stem and the state that follows on from that, the classic structure of the perfect. Secondly, the role of the external argument of *have* is determined by the semantics of the verbal stem. Since the external theta-role has not been assigned directly (because of absorption by the participle head), it is a short step to identifying the unassigned role as being that of the external argument of the auxiliary. This identification is reinforced by the locative relation between the arguments of *have* which require the subject to be 'located at' (i.e. a participant in) the event denoted by the verbal stem.

If the reasoning outlined above is correct, then, semantically, the specifier relation provides a means of deriving pragmatically enriched interpretations that are compositionally determined by the semantics of two elements. It follows that doubly headed constructions will again tend to favour one of the constituents being a functional category and the other an expression whose interpretation is open to pragmatic enrichment, i.e. one that is starting to lose its independent contentive status.

Summarising, the unification operation has the consequence that the combination of two expressions in a doubly headed construction is restricted as set out in (29).

(29) $x \cup y$ is valid iff.
(i) the formal features of x and y are non-contradictory;
ii either x or y is phonologically null;
iii and the argument structure and associated thematic structure of x and y can merge coherently.

Notice that (29) is not a stipulation, but follows from a number of straightforward assumptions. Furthermore, these conditions do not impose any total constraint on the inherent nature of *x* and *y*. Thus, while they tend to favour one of the combining constituents to be a functional category, they do not absolutely require it. This will allow some freedom (often missing from formal or semi-formal theories of syntax) for the development of grammatical expressions from contentive ones, as envisaged above.

Clearly, this analysis of specifiers as secondary heads and the conditions imposed on the relation by the unification operation has wide-ranging consequences that require further exploration. For example, the theory disallows movement from leaving a full copy of a moved constituent. Otherwise, expressions like *who$_i$ did the student say [$_{(C \cup WH)}$ t$_i$' that the lecturer was harassing t$_i$]* would be ungrammatical as the phonological information of the trace t_i' would be required to unify with that of the phonologically overt complementizer. Another important consequence is that the theory of word order to be found in Chomsky (1995a) (and Kayne 1994) is impossible under these assumptions. Since, as noted above, *x* and *y* in $x \cup y$ structures are both maximal and since $x \cup y$ is non-distinct from both *x* and *y*, notions of asymmetric c-command are problematic. Where only one of the expressions is minimal, then in the theory of Chomsky (1995a) there will be no problem: the minimal expression will asymmetrically c-command, and so precede, the minimal elements in the other (although this is a problem for Kayne). Where both *x* and *y* are non-minimal (or both minimal), c-command relations will not be asymmetric, however defined, and there will need to be something more required to define ordering relations.

Furthermore, once two constituents have been Merged through unification, it is impossible to tell the difference between them with regard to their status as heads (or specifiers). It is not clear that anything, in fact, follows from this. Since both expressions (where one does not have a trace embedded in the other) are 'maximal' from the point of view of further movement, the fact that one might be labelled as the head is not significant for this operation. From the semantic point of view, the unification of the semantics of the two elements again renders the headedness of the subexpressions of the construction irrelevant. However, if one wants to define a notion of primary headedness within ‘specifier-head’ constructions, then it should be possible to do so, as the two subexpressions will in general differ in one of two ways, as specified in (30a) and (30b). The first distinguishes ‘specifiers’ from co-structures into which they have a thematic dependency, while the second defines headedness in terms of selection (a typical pre-theoretical property ascribed to heads, cf. Zwicky 1985). Notice that on many occasions both properties will pick out a single expression (e.g. where X is C and Y is DP). On other occasions, as in the auxiliary construction in English, selection may be the principal property, as for example in the perfect where *have* is selected as the primary head in *have* ∪ *en*, again in conformity with pre-theoretical expectations).

(30a) In $\{x \cup y,\{x,y\}\}$, if X contains the trace of Y, then X is the primary head.
(30b) In $\{x \cup y,\{x,y\}\}$, if x is selected, then X is the primary head.

Whatever solutions are found for the apparent problems noted above, the introduction a second tree-building operation utilising unification provides an interesting theoretical model that appears to be able to analyse common constructions straightforwardly with the minimum of ancillary assumptions; provides a theoretical account of the process of grammaticalization; and allows selectional (and other) dependency relations to be satisfied strictly locally.

References

Baker, M C (1988) *Incorporation: A Theory of Grammatical Function Changing*. Chicago University Press, Chicago.

Burzio, L (1986) *Italian Syntax*. Dordrecht, Reidel.

Cann, R. (1993) ‘Patterns of headedness’. In G. Corbett, N. M. Fraser and S. McGlashan (eds.) *Heads in Grammatical Theory*. Cambridge, Cambridge University Press: 44-72.

Cann, R (1995) ‘Grammaticalization, compositionality and *have* in English’. Unpublished MS., University of Edinburgh.

Cann, R (1996) ‘Categories, labels and types’. *Edinburgh Occasional Papers in Linguistics* EOPL-96-3.

Cann, R and M E Tait (1995) ‘Raising morphology’. In C. S. Rhys, D. Adger and A. von Klopp (eds.) *Functional Categories, Argument Structure and Parametric Variation.* Edinburgh Working Papers in Cognitive Science vol. 9, Centre for Cognitive Science, University of Edinburgh: 1-23.

Chomsky, N. (1957) *Syntactic Structures*. The Hague, Mouton.

Chomsky, N (1970) ‘Remarks on nominalisation’. In R. A. Jacobs and P. S. Rosenbaum (eds.) *Readings in English Transformational Grammar*. Waltham, Mass., Ginn.: 184-221.

Chomsky, N (1986) *Barriers*; MIT Press, Cambridge, Massachusetts.

Chomsky, N (1995a) ‘Bare phrase structure theory’. In G Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program*. Oxford, Basil Blackwell: 385-439.

Chomsky, N (1995b) ‘Categories and Transformation’. In *The Minimalist Program*. Cambridge, Mass., MIT Press:219-314.

Cormack, A. (this volume) ‘Without specifiers’.

Cowper, E A (1993) ‘Infinitival complements of have’. *Canadian Journal of Linguis-*

tics 37: 115-135.
Fukui, N (1986) A theory of category projection and its applications. Unpublished doctoral dissertation, MIT.
Gazdar, G., G. K. Pullum and I. A. Sag (1982) ‘Auxiliaries and related phenomena in a restricted theory of grammar’; *Language* 58:591-638.
Gazdar, G., E. Klein, G. K. Pullum and I. A. Sag (1985) *Generalized Phrase Structure Grammar*; Oxford. Basil Blackwell.
Grimshaw, J (1990) *Argument Structure*. Cambridge, MA, MIT Press.
Grimshaw, J. and A. Mester (1988) ‘Light verbs and θ marking’. *Linguistic Inquiry* 19: 205-232.
Haegemann, L. (1991) *Generative Syntax*. Oxford, Basil Blackwell.
Haegemann, L. and Zanuttini, R. (1991) ‘Negative heads and the NEG criterion’. *The Linguistic Review* 8: 233-252.
Harris, Z. (1951) *Methods In Structural Linguistics*. Chicago, Illinois. Chicago University Press.
Hopper, P. J. and E. C. Traugott (1993) *Grammaticalization*. Cambridge, Cambridge University Press.
Jackendoff, R (1977) *X’ Theory*. Cambridge, Mass., MIT Press.
Kayne, R (1993) ‘Towards a modular theory of auxiliary selection’. *Studia Linguistica* 47: 3-31.
Kayne, R. (1994) *The Antisymmetry of Syntax*. Cambridge, Mass., MIT Press.
Lyons, J (1968) *An Introduction to Theoretical Linguistics*. Cambridge, Cambridge University Press.
Pollard, C and I A Sag (1994) *Head-Driven Phrase Structure Grammar*. Chicago, Chicago University Press.
Radford, A (1988) *Transformational grammar*. Cambridge, Cambridge University Press.
Roberts, I. (1993) ‘A formal account of grammaticalization in the history of Romance futures’. *Folia Linguistica Historica* 13: 219-258.
Ritter, E. and S. T. Rosen (1997) ‘The function of have’. *Lingua* 101: 295-321.
Rizzi, L. (1990a) *Relativized Minimality*. Cambridge, MA., MIT Press.
Rizzi, L. (1990b) ‘Speculations on verb second’. In M. Nespor et al. (eds.) *Grammar in Progress: a festschrift for Henk van Riemsdijk* Dordrecht, Foris Publications: 375-386.
Steedman, M (1988) ‘Combinators and Grammars’. In R. T. Oehrle, E. Bach and D. Wheeler (eds.) *Categorial Grammars and Natural Language Structures*, Dordrecht, Kluwer Academic Publishers: 417-442.
Stuurman, F. (1985) *Phrase Structure Theory in Generative Grammar*. Dordrecht, Foris Publications.
Tait, M. (1991) The syntactic representation of morphological categories. Unpubl. PhD. dissertation, University of Edinburgh.
Verkuyl, H (1981) ‘Numerals and quantifiers in X-bar syntax and their semantic interpretation’. In J. A. G. Groenendijk, T. M. V. Janssen, and M. B. J. Stokhof (eds.) *Formal Methods in the Study of Language*. Amsterdam, Mathematisch Zentrum: 567-599.
Warner, A. R. (1993) *English Auxiliaries: structure and history*. Cambridge, Cambridge University Press.
Zwicky, A (1985) ‘Heads’. *Journal of Linguistics* 21: 1-30.
Zwicky, A. and G. K. Pullum (1983) ‘Cliticization versus inflection: English *n’t*. *Lan-*

guage 59: 502-513.