

**SYNTACTIC CATEGORIES IN THE
CORRESPONDENCE ARCHITECTURE**

Jean-Philippe Marcotte
University of Minnesota

Proceedings of the LFG14 Conference

Miriam Butt and Tracy Holloway King (Editors)

2014

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

Existing approaches to the notion of syntactic category in Lexical-Functional Grammar are either formally explicit but theoretically inadequate (Kaplan, 1987), or detailed but ill-integrated in the correspondence architecture (Bresnan, 2001; Toivonen, 2003). This paper develops a third approach, excising syntactic categories from the c-structure, modeling them as sets of privative features, and situating them in a corresponding x-structure. This allows for the elimination of X' levels as theoretical primitives, while maintaining straightforward definitions of the notion of syntactic projection, of non-projecting words (Toivonen, 2003), and of endocentric structure–function mappings (Bresnan, 2001). An application of the system in the domain of paradigmatic morphology (Stump, 2001) is also suggested.

1 What's in a syntactic category

The internal structure of syntactic categories in Lexical-Functional Grammar is a topic which has received some attention in the literature. We find hints about the nature of this internal structure in Kaplan (1987: 351), writing about levels of representation:

There's the constituent phrase structure, which varies across languages, where you have traditional surface structure [...] and parts of speech labeling categories, *perhaps a feature system on those categories* (although in the case of LFG *if there is one it's a very weak one*). [emphasis added — JPM]

Kaplan does not exploit the possibility he alludes to in the quote above: in §2, following a short presentation of his formal model of LFG, I review his expositionally expedient adoption of atomic categorial symbols, and conclude that it is inadequate because it lacks the properties needed to define the notion of syntactic projection as currently understood. In this paper, after reviewing previous LFG models of syntactic categories, I take up Kaplan's idea by formulating a weak feature system for them, and exploring its consequences.

There are at least two distinct LFG-specific kinds of attempts at giving syntactic categories an internal structure: complex categories (Butt et al., 1999; Crouch et al., 2008) and the X' theory of Bresnan (2001) and Toivonen (2003). In §3, I argue that complex categories are more a solution to an engineering problem than a theoretically interesting model. In §4, I point out that the X'-theoretic categories defined by Bresnan and Toivonen are both somewhat baroque and ill-integrated in the correspondence architecture.

In §5 I introduce the level of x-structure, from which X'-type relations can be derived, and infuse it with three privative categorial features. These features serve to define lexical and functional syntactic categories, and restate Bresnan and Toivonen's X' theory:

c-structure rules, category types, combinatorial constraints on these categories, and universal endocentric structure–function mapping principles. However some problems remain, in particular an inability to distinguish between the functional categories I and C.

I demonstrate in §6 that a tweak of the formal properties of x-structure, with a slightly different assortment of categorial features, allows this deficiency to be remedied, with the ability to specify distinctions between inflectional categories as a side-effect; I offer speculation that this is a beneficial outcome.

2 The correspondence architecture of Lexical-Functional Grammar

This section recapitulates some foundational design principles of Lexical-Functional Grammar, setting up an apparatus for subsequent formal gymnastics.

In his exposition of the formal underpinnings of the LFG architecture, Kaplan (1987) proposes to model the grammatical mapping between sound and meaning as a function Γ from a form to a meaning:

$$(1) \quad \text{form} \bullet \xrightarrow{\Gamma} \bullet \text{ meaning}$$

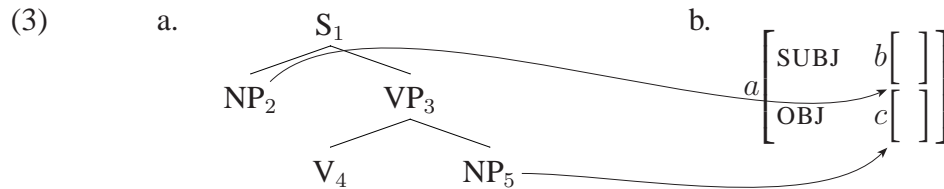
The mapping is obviously complex, and stating it explicitly requires making generalizations of various types, which are best modeled in structural levels with congenial formal properties. Formally, we can assume that Γ is the composition of functions which state correspondences between intermediate structural levels, for example:

$$(2) \quad \begin{array}{c} \Gamma = \psi \circ \phi \circ \pi \\ \text{form} \bullet \xrightarrow{\pi} \bullet \xrightarrow{\phi} \bullet \xrightarrow{\psi} \bullet \text{ meaning} \end{array}$$

The precise assortment of correspondence functions and the structural levels they mediate is to be determined based on careful linguistic argumentation over relevant generalization types. As such we need a c-structure tree for modeling generalizations about constituency, linear order, and syntactic category; we need an f-structure for modeling generalizations about grammatical function, agreement, long-distance dependencies, binding, control, raising, etc.; and we need correspondence functions to serve as interfaces between these structural levels. Essentially, we factor generalizations out of Γ and allocate them to structural levels according to their formal type and relationship to other generalizations.

2.1 Structural description

Trees and attribute–value matrices are merely visually perspicuous ways of displaying consistent STRUCTURAL DESCRIPTIONS. Thus the c- and f-structures in (3) are perspicuous visualizations the structural descriptions in (4).



In (4) \mathcal{M} is a function from nodes to (their mother) nodes; \mathcal{L} is a function from nodes to their category labels;¹ $<$ is a precedence relation between nodes with the same mother.

(4) a. $\mathcal{M}(n_2) = n_1$ $\mathcal{L}(n_1) = S$ c. $(a \text{ SUBJ}) = b$
 $\mathcal{M}(n_3) = n_1$ $\mathcal{L}(n_2) = NP$ $(a \text{ OBJ}) = c$
 $n_2 < n_3$ $\mathcal{L}(n_3) = VP$

 b. $\mathcal{M}(n_4) = n_3$ $\mathcal{L}(n_3) = VP$
 $\mathcal{M}(n_5) = n_3$ $\mathcal{L}(n_4) = V$
 $n_4 < n_5$ $\mathcal{L}(n_5) = NP$

Notice that (4a) and (4b) are respectively $S \rightarrow NP VP$ and $VP \rightarrow V NP$. Consequently we can use standard phrase structure grammar notation to abbreviate structural descriptions.

An f-structure is a recursive function from attributes to values, where values can themselves be such functions. In (4c), the structural description of (3b), a is an f-structure (that of the sentence), as are b and c (the f-structures of the SUBJ and OBJ, formally the values of those attributes).

2.2 Structural correspondence

The correspondence between the c- and f-structure is specified in terms of the immediate dominance (motherhood) relation native to c-descriptions.

Let the symbol $*$ stand as a variable for a node, and let ϕ be a correspondence function from nodes into f-structures. Then $\phi(*)$ is the f-structure of $*$; $\mathcal{M}(*)$ is $*$'s mother, and $\phi(\mathcal{M}(*))$ is $*$'s mother's f-structure.

¹Kaplan (1987) uses λ to notate this function; I have changed the notation to eliminate the potential for confusing it with the λ correspondence function from the a-structure to the f-structure, widely-accepted since Butt et al. (1997).

For notational perspicuity, Kaplan (1987) defines \downarrow to stand for $\phi(*)$, and \uparrow to stand for $\phi(\mathcal{M}(*))$. A grammar can then be formulated as follows:

$$(5) \quad \begin{array}{l} \text{a. } S \rightarrow \quad \text{NP} \quad \text{VP} \\ \quad (\uparrow \text{ SUBJ}) = \downarrow \quad \uparrow = \downarrow \\ \text{b. } VP \rightarrow \quad V \quad \text{NP} \\ \quad \quad \quad \uparrow = \downarrow \quad (\uparrow \text{ OBJ}) = \downarrow \end{array}$$

The functional schema annotations in (5a) read as: S's f-structure is (i) a function from the attribute SUBJ to the f-structure of NP, (ii) equal to the f-structure of VP.

2.3 Co-description

Now let $\phi(n_i) = f_i$ for every node n_i . Then (5) co-specifies the c- and f-descriptions in (6), which are equivalent to (4) given the following substitutions: a for $f_1 = f_3 = f_4$, b for f_2 , and c for f_5 .

$$(6) \quad \begin{array}{l} \text{a. } \mathcal{M}(n_2) = n_1 \quad \mathcal{L}(n_1) = S \quad (f_1 \text{ SUBJ}) = f_2 \\ \quad \mathcal{M}(n_3) = n_1 \quad \mathcal{L}(n_2) = \text{NP} \quad f_1 = f_3 \\ \quad n_2 < n_3 \quad \mathcal{L}(n_3) = \text{VP} \\ \text{b. } \mathcal{M}(n_4) = n_3 \quad \mathcal{L}(n_3) = \text{VP} \quad f_3 = f_4 \\ \quad \mathcal{M}(n_5) = n_3 \quad \mathcal{L}(n_4) = V \quad (f_3 \text{ OBJ}) = f_5 \\ \quad n_4 < n_5 \quad \mathcal{L}(n_5) = \text{NP} \end{array}$$

A partial lexical entry for the verb *yawns* might be something like:

$$(7) \quad \begin{array}{l} \text{yawns} \quad \mathcal{L}(\mathcal{M}(*)) = V \quad \text{'My mother's category label is V'} \\ \quad (\uparrow \text{ TENSE}) = \text{present} \quad \text{'My mother's f-structure's is present tense'} \\ \quad (\uparrow \text{ SUBJ PERS}) = 3 \quad \text{'My mother's subject's f-structure is 3rd person'} \\ \quad (\uparrow \text{ SUBJ NUM}) = \text{sg} \quad \text{'My mother's subject's f-structure is singular'} \end{array}$$

Both this lexical entry and the c-structure rules in (5) co-describe the c-structure and the f-structure: statements about the c-structure and the f-structure appear in the context of each other.

Every rule or lexical entry is a set of statements about one or more structural levels; the grammar is the disjunction of all such sets; a sentence is grammatical only if this grammar is true of every single one of its parts.

2.4 Atomic syntactic categories

The relative simplicity of this formal model is appealing. However, the atomic approach to syntactic categories in Kaplan (1987), illustrated in (4) and (6), is obsolete and inadequate.

Taking a specific instance: the categories of n_3 and its daughter n_4 are respectively VP and V. These are displayed with a shared character ‘V’, and are implied to be a phrase or not by the presence or absence of a character ‘P’. But these typographical conventions are just that, and in no way represent a formal assertion of relation (same category) and differentiation (distinct levels) within a syntactic PROJECTION.

To be exact: looking back at the tree in (3) and its structural description in either (4) or (6), there is no sense in which V_4 is formally represented as the categorial head of VP_3 , or as projecting VP_3 . Nor does the specific formal model in Kaplan (1987) contain an implicit theory precluding a tree in which VP is the daughter and categorial head of V, or N the daughter and categorial head of VP.

In this treatment, no mechanism is specified for the sharing of categorial information between mother and daughter nodes, and consequently the notion of syntactic projection is left completely undefined.

3 Complex categories

Another take on syntactic categories in LFG is the concept of COMPLEX CATEGORY, as documented for XLE in Crouch et al. (2008). Complex categories are intended not as a theoretically interesting formal device, but rather as an efficiency-maximizing engineering solution to the problem of near-duplicate c-structure rules in industrial grammars. Because they do this by allowing a degree of information-sharing between mother and daughter nodes, and thus appear to have properties required to model the notion of projection, it is worth considering here whether this is in fact the case.

Complex categories are justified for ParGram grammars in Butt et al. (1999: 192) with the following examples and text, in which the phrase structure rule for $NP[_type]$ is intended to generalize over standard, interrogative, and relative noun phrase subtypes (respectively $NP[std]$, $NP[int]$, $NP[rel]$):

- (8) $NP[_type] \rightarrow \{ (D[_type]: _type = std)$
 $\quad \quad \quad | D[_type]: \{ _type = int \mid _type = rel \} \}$
 $\quad \quad \quad NPap.$
- (9) a. $NP[std] \rightarrow (D[std]) NPap.$
b. $NP[int] \rightarrow D[int] NPap.$
- ⋮

The advantages of such parameterization over rules via the use of complex categories is that again large parts of rules can be shared across types of constructions that differ systematically in one respect, but which work in essentially the same way in other respects.

To explicate further: the upper rule expands $NP[_{type}]$ to a determiner $D[_{type}]$ followed by NP_{ap} , a noun phrase level within which adjective phrases attach. There is a disjunction over the determiner $D[_{type}]$ in the upper rule, within each disjunct of which the variable $_{type}$ is instantiated to one of the values *std*, *int*, and *rel*. This value is passed between $D[_{type}]$ and $NP[_{type}]$. Thus when $_{type}$ is instantiated to *std*, $D[_{type}]$ is instantiated to $D[std]$ (a node which specified as optional in the disjunction) and $NP[_{type}]$ to $NP[std]$; this fully-instantiated rule is shown in the quoted passage above for illustrative purposes, but does not need to be separately stated in the grammar since it is implied by the $NP[_{type}]$ rule. The complex category $NP[std]$ can now be called by some other rule in the grammar; when it is used it will expand via the $NP[_{type}]$ rule to and $D[std]$ followed by NP_{ap} .

Observe that the information-passing here involves not the passing of categorial information, but the passing of morphosyntactic information, which is likely to occur independently in the f-structure. This is intentional: computing over c-structures is more efficient than unifying f-structures, and complex categories are designed to shift the computational burden towards the former.² Nevertheless, as a formal device complex categories afford the possibility of specifying just the sort of mother-to-daughter information-passing which must be part of any model of the syntactic category and projection concepts. It is therefore worthwhile to contemplate briefly whether complex categories can be used to provide a satisfactory account.

In this spirit, I point out that the information-passing within the $NP[_{type}]$ rule above takes place between a mother node and its non-head daughter, and so represents an example of communication between different projections, not between levels of the same projection. It follows that complex categories impose no restrictions against categorial information-passing between projections.³ Nothing about this formal device prevents a restatement of the above as DP rules, in which the information-passing would be from a head D to a mother DP, preserving within a categorial projection the intuition of distinct determiners setting the type parameter of their mother node; however nothing about complex categories forces the latter formulation.

²I thank an anonymous reviewer for reminding me of this fact.

³In fact, it is possible to specify rules in which the mother is parameterized for two or more variables, whose instantiation can be distributed indiscriminately between the daughters. I myself have written such rules into the French ParGram grammar, at some point between 2000 and 2005.

Notice also that this device is not used above to specify different levels within a projection: instead within the noun phrase projection we have NP[_type] for broad noun phrase type and NP_{AP} for the daughter level. Shared categorial nomenclature is at the forbearance of the grammar writer, as is the ordering of the levels: the formal device imposes no substantive constraints.

But suppose that one did try to use complex categories to specify differences in level within a projection. One would now be faced with a fundamental problem: variables allow the passing of information that mothers and daughters share, not information which differentiates them. In short, the formal device of complex categories allows the passing of information to be shared between categorially distinct nodes, but not information that distinguishes otherwise categorially identical nodes.

To summarize: complex categories are not intended to model syntactic projection; at first glance they appear to have attractive information-passing properties that one could repurpose towards such modeling, but in fact they do not.

4 X' theory in Lexical-Functional Grammar

The X' theory of Bresnan (2001), revised and extended somewhat in Toivonen (2001, 2003), is the LFG literature's third type of syntactic category model. It uses the assortment of categorial features in Table 1.

	'predicative'	'transitive'
verbal	+	+
adjectival	+	-
adpositional	-	+
nominal	-	-

Table 1: Categorial features in the X' theory of Bresnan (2001: 120)

To these, Bresnan adds functionality features (F0–F2) and bar-level features (B0–B2).⁴ Syntactic categories can be exhaustively defined as in the following examples:

$$(8) \quad \begin{array}{ll} V : \langle [+Pr, +Tr], F0, B0 \rangle & IP : \langle [+Pr, +Tr], F1, B2 \rangle \\ VP : \langle [+Pr, +Tr], F0, B2 \rangle & C' : \langle [+Pr, +Tr], F2, B1 \rangle \end{array}$$

The nodes within a projection, like N–N'–NP, have the same features for category and functionality; nodes across projections may share category features, as for V–I–C and

⁴Bresnan does not use the prefix B for these; I add it for notational transparency.

D–N, but differ in functionality. The system owes much to Grimshaw (2000). Toivonen (2001, 2003) adds the notion of NON-PROJECTING WORD.

For Bresnan (2001: 120), “ ‘Predicative’ categories are those which cannot stand alone as arguments, but require an external subject of predication. [...] ‘Transitive’ categories are those which may take an object or direct complement function”.⁵ This establishes the idea of constraints on possible c- to f-structure mappings: for example, that a [–Tr] projection cannot accommodate a node annotated with (↑ OBJ) = ↓.

Following through on this idea, both Bresnan (2001) and Toivonen (2003) state several UNIVERSAL PRINCIPLES OF ENDOCENTRIC STRUCTURE–FUNCTION ASSOCIATION: constraints on which kinds of f-structure annotations a node can bear, given its own properties and those of its c-structure context.

The Bresnan–Toivonen feature system has the virtue of being explicit, but neither researcher formalizes it: it remains unclear where they intend these feature sets to dwell in the correspondence architecture.

A related wrinkle is that the universal phrase structure rules from Bresnan (2001: 99) seem to be the only mechanism for enforcing the immediate dominance of B2 over B1 over B0 categories:

- (9) a. $X' \rightarrow X, YP$ b. $XP \rightarrow YP, X'$

This dominance sequence is implied by the use of integers in the feature nomenclature, but does not constitute a formal requirement. A similar point holds for the nesting of projections which differ in their value for F: the nesting of VP inside IP inside CP is implied by their respective features F0, F1 and F2, but is not enforced formally.

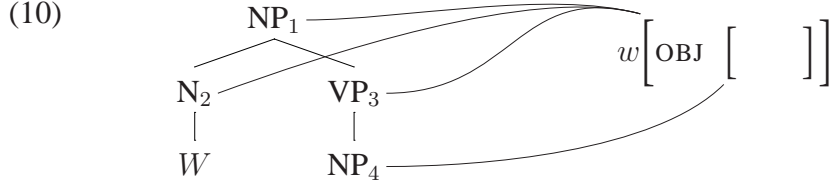
4.1 A note on mixed categories

In an analysis of Gĩkũyũ constituents which are categorially NPs but internally have some properties of VPs, Bresnan and Mugane (2006), rather than deploying the X' theory just presented, return to the atomic-category view of Kaplan (1987) discussed in §2 and criticized in §2.4.

Their treatment is as follows: for a lexeme W , take the f-structure w containing its PRED, obtain via ϕ^{-1} the functional domain of f (the set of c-structure nodes which corresponds to w via ϕ), and for each node in the functional domain obtain its atomic syntactic category via \mathcal{L} (see §2.1). The lexical category of W is a requirement for a maximal

⁵ With regards to the cross-classification of syntactic categories, these features correspond to those of Chomsky (1981: 48) as follows: ±Predicative corresponds directly to ±V (also called ‘predicative’ by Chomsky), and ±Transitive is ±N (‘substantive’) with the polarity reversed.

projection label (VP, NP, etc) be part of the set of syntactic categories thus achieved. A mixed-category verbal noun requires that both VP and NP be part of that category set: thus W in (10), while it occurs in N_2 , has a functional domain which also includes NP_1 and VP_3 , and is able to take both nominal and verbal dependents, the latter exemplified by the OBJ corresponding to NP_4 .

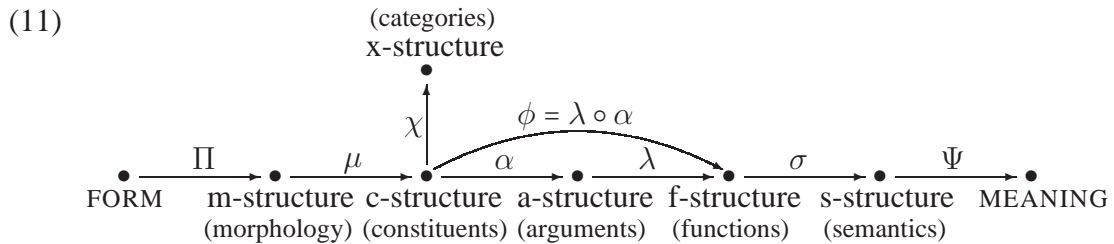


This account of mixed categories is far more formally explicit than the X' theory above. Precisely because the formal status of the category feature bundles in (8) is unclear, the possibility of its integration into this particular treatment of mixed categories remains unsettled. Presumably, a requirement for a functional domain to include a node with atomic category label VP would become a requirement for this functional domain to include a node with feature bundle $\langle [+Pr,+Tr],F0,B2 \rangle$. If this is the case, then an X' -theoretic account of this phenomenon follows straightforwardly.

5 Bare phrase structure for Lexical-Functional Grammar

Although some of the issues above can be resolved by formally integrating the Bresnan–Toivonen X' theory into the correspondence architecture without further modifications, what I propose in this section is a reformulation that will be reminiscent of the BARE PHRASE STRUCTURE of Chomsky (1995).

One version of the correspondence architecture is shown in (11). This version inspired by but not identical to that in Asudeh (2012) — in particular it is truncated at the extremities, with phonological and information-structural levels telescoped into Π and Ψ respectively, for horizontal spacing reasons. I have added a structural level of x-structure, in direct correspondence with the c-structure, as the locus of a formalization of syntactic category information.



I will provide more details on the correspondence function χ in §5.2; for the moment, imagine simply that it takes a c-structure node into its x-structure.

5.1 Projection

The native c-structural notion of immediate dominance (see §2.1) can serve to specify relationships between the x-structure of any node and that of its mother. Derived notions like (12a) and (12b) can now be defined as below, with the complementary notion in (12c):

- (12) a. PROJECTING NODE
 A node projects iff its x-structure is identical with its mother's x-structure:
 $Proj(*) \iff \chi(*) = \chi(\mathcal{M}(*))$
- b. MAXIMAL PROJECTION
 A node is a maximal projection iff it is not a projecting node:
 $Max(*) \iff \neg Proj(*)$
- c. TERMINAL
 A node is a terminal iff it has no mother:
 $Term(*) \iff \neg \exists n. \mathcal{M}(n) = *$

Note that (12c) is mutually exclusive with neither (12a) nor (12b). As such, it is technically possible to define the equivalent of a bar-level node as a projecting non-terminal; however, I leave this gap as a theoretical claim that the notion is not a universal.

A non-projecting word (Toivonen, 2001, 2003) can now be straightforwardly defined as a word whose lexical entry simultaneously contains both $Max(*)$ and $Term(*)$. In addition, the projection of any particular node n can be obtained via the inverse of χ , as the set of nodes $\chi^{-1}(\chi(n))$.

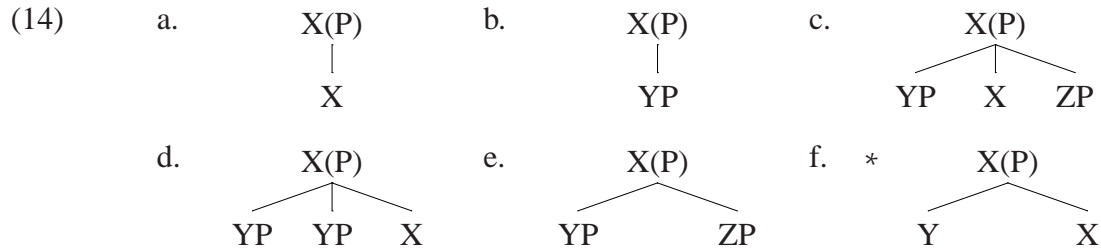
Finally, the two universal endocentric phrase structure rules in (9) can be replaced with the one in (13), which states that one (optional) projecting node can be shuffled⁶ among its potentially multiple (and optional) maximal projection sisters.

$$(13) \quad n_i \rightarrow \left(\begin{array}{c} n_j \\ Proj(*) \end{array} \right), \left(\begin{array}{c} n_{k\dots}^+ \\ Max(*) \end{array} \right)$$

Thus branching in this theory is potentially n-ary, with node ordering relegated to separate rule statements (in the style of Gazdar et al., 1985). Some sample configurations

⁶The comma in this rule is intended as the shuffle operator, as presented in Dalrymple (2001: 99), where its function is to shuffle the two lists that serve as its arguments.

are shown in (14), where X means a projecting node (not necessarily a terminal), XP means a maximal node, and X(P) means a node which can either be maximal or projecting. Note that (14a-e) are licensed by (13) but (14f), having two projecting daughters, is not:



Recall one of the problems with the Bresnan–Toivonen X' theory: it is the rules in (9) which enforce the dominance of phrasal nodes over heads, while the integer values of the features B0–B2 merely imply this dominance. In the current framework the problem does not arise: nodes are phrasal or terminal not by virtue of their feature specifications, but because of their position in the c-structure.

5.2 Syntactic category features

Let χ now be a function from nodes to sets of the symbols Pr , Tr , f in any combinations. Let these symbols serve as privative features such that Pr is equivalent to a positive value for ‘predicative’ in Table 1 and its absence to a negative value for that feature, and let Tr do the same for ‘transitive’. Let the presence and absence of f in an x-structure signify the distinction between functional and lexical categories, respectively.

CATEGORY TYPE	
Lexical	Functional
$V:\{Pr, Tr\}$	$I:\{Pr, Tr, f\}$
$A:\{Pr\}$	$?:\{Pr, f\}$
$P:\{Tr\}$	$?:\{Tr, f\}$
$N:\{\}$	$D:\{f\}$

Table 2: Possible x-structure feature combinations

Table 2 shows the possible x-structure sets alongside the category labels they can be taken to define. Functional adjectival and adpositional feature sets are implied, although they define no standard category labels. Furthermore there is no way of distinguishing the category C from I; on this issue see §6.

The lexical entry in (7) is now as in (15), with the atomic category V replaced by the appropriate x-structure equation:

$$(15) \quad \text{yawns} \quad \chi(*) = \{Pr, Tr\} \quad (\downarrow \text{ SUBJ PERS}) = 3 \\ (\downarrow \text{ TENSE}) = \text{present} \quad (\downarrow \text{ SUBJ NUM}) = \text{sg}$$

This lexical entry also upends the arrow metavariable, to go along with the statement in (12c) that heads/terminals have phonological content: words *are* terminals instead of merely being *under* terminals.

5.3 Endocentric c- to f-structure mappings

The x-structure features asserted by (13) to be shared or not between nodes are supplied by heads. As in the Bresnan–Toivonen X' theory, all f-structure annotation in endocentric languages is handled by universal principles of endocentric structure–function association:

- (16) a. A projecting node shares the f-structure of its mother:

$$Proj(*) \implies \uparrow = \downarrow$$
- b. A SUBJ is a DP daughter of IP:

$$(\uparrow \text{ SUBJ}) = \downarrow \implies \begin{array}{ll} Max(*) & Max(\mathcal{M}(*)) \\ \chi(*) = \{f\} & \chi(\mathcal{M}(*)) = \{Pr, Tr, f\} \end{array}$$
- c. An OBJ is a DP with a V(P) or P(P) mother:

$$(\uparrow \text{ OBJ}) = \downarrow \implies \begin{array}{ll} Max(*) & \{f\} \notin \chi(\mathcal{M}(*)) \\ \chi(*) = \{f\} & \{Tr\} \subset \chi(\mathcal{M}(*)) \end{array}$$
- d. An OBL is a non-verbal/adjectival XP with a non-functional mother:

$$(\uparrow \text{ OBL}) = \downarrow \implies \begin{array}{ll} Max(*) & \{Pr\} \notin \chi(*) \quad \{f\} \notin \chi(\mathcal{M}(*)) \end{array}$$
- e. A POSS is a DP daughter of DP:

$$(\uparrow \text{ POSS}) = \downarrow \implies \begin{array}{ll} Max(*) & Max(\mathcal{M}(*)) \\ \chi(*) = \{f\} & \chi(\mathcal{M}(*)) = \{f\} \end{array}$$
- f. A node that shares its f-structure with a functional mother must be such that $\{f\}$ is the restriction of its mother's x-structure by its own x-structure:

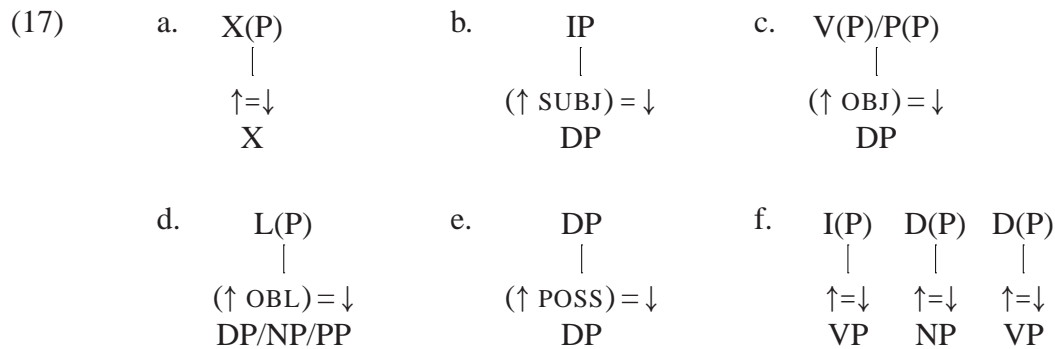
$$\begin{array}{l} \uparrow = \downarrow \\ \{f\} \subset \chi(\mathcal{M}(*)) \end{array} \implies \chi(\mathcal{M}(*)) - \chi(*) = \{f\}$$

These association principles are stricter than those of Bresnan (2001: 120), in which no categorial restrictions beyond lexicality or functionality are imposed on the nodes; indeed they may be too strict, if it is correct that that CPs can bear the SUBJ and OBJ functions (about which see respectively Bresnan, 1994 and Alsina et al., 2005). In this case, the constraints can be relaxed by using $\{f\} \subset \chi(*)$ to designate all functional categories and $\{f\} \notin \chi(*)$ to designate all lexical ones.

The joint effect of (16a) and (16f) ensure that, within an EXTENDED PROJECTION like I–V, the functional projection dominates the lexical projection (for a different version of this idea, see Grimshaw, 2000). Recall from the end of §4 that the Bresnan–Toivonen *X'* theory merely implies this dominance with integers as the values of its F0–2 feature.

The notion of extended projection can at this point be given an exact definition in terms of the inverse of ϕ , as implied by the definition of EXTENDED HEAD in Bresnan (2001: 132): $\phi^{-1}(\phi(*))$.⁷ This is exactly analogous to the definition of projection in §5.1.

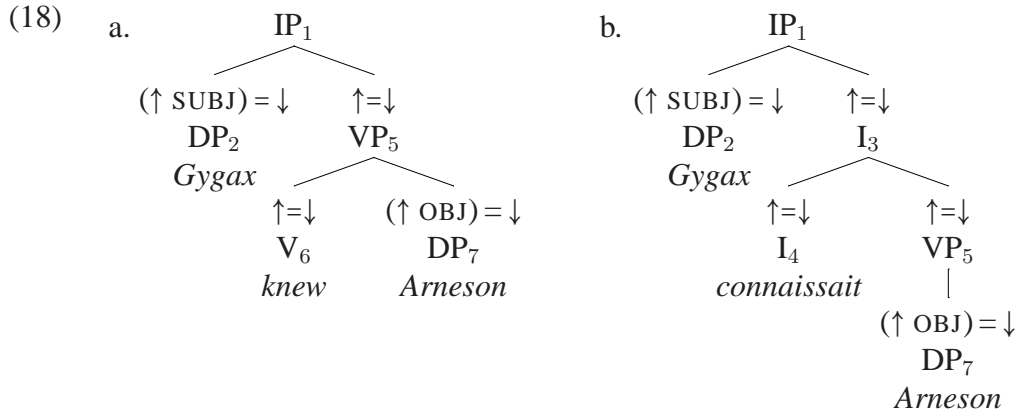
The respective configurations licensed by the principles in (16) are in (17), with category symbols as for (14); L in (17d) stands for any lexical category:⁸



Taken together, (13) and (16) minimally require the tree in (18a) for an English sentence with a finite verb form taking a SUBJ and an OBJ, and the tree in (18b) for its French translation, assuming an analysis of verb placement along the lines of Pollock (1989), implemented via head mobility within extended projections (Bresnan, 2001: chapter 7).

⁷With regards to this definition, it should be noted that Grimshaw (2000) considers P to be within the extended projection of D–N; for Bresnan (2001) this is not the case, as the c-structure complement NPs of prepositions are their f-structure OBJs. See §6.3 for more on Grimshaw’s idea.

⁸A reviewer points out that set theory’s Axiom of Extension — see Dalrymple (2001: 32) for an application to f-structure — prevents any configuration in which *a* is the mother of *b* and *c*, and all three are categorially identical. Nodes *b* and *c* cannot belong to separate functional domains: by Extension their x-structures, having the same membership, are identical, so by (16a) they share a functional structure. Among other things this makes Falk’s (1984) raising-verb analysis of English auxiliary verb sequences impossible; I do not take this to be an invalidating consequence, but rather a challenging prediction. Moreover, *b* and *c* cannot belong to the same functional domain, as they would be projecting sisters, which is impossible by (13). This would seem to prevent *any* categorially plausible analysis of English auxiliary sequences. The way out is to alter (13) to allow projecting sisters, and use other constraints to rein in the resulting overgeneration.



Applying the notational convention that $\chi(n_i) = x_i$ we find that the nodes in (18a,b) are in correspondence with the x-structures in (19a,b) respectively:

- (19) a. $x_1 = \{Pr, Tr, f\}$ $x_2 = \{f\}$ $x_5 = x_6 = \{Pr, Tr\}$ $x_7 = \{f\}$
 b. $x_1 = x_3 = x_4 = \{Pr, Tr, f\}$ $x_2 = \{f\}$ $x_5 = \{Pr, Tr\}$ $x_7 = \{f\}$

5.4 Summary

The x-structure model of syntactic categories altogether eliminates the Bresnan–Toivonen bar-level features, and reduces to privative features both the formerly binary \pm Predicative and \pm Transitive, and the formerly ternary F. It also reduces the number of universal phrase structure rules: Bresnan (2001) needs the two rules in (9); this system needs only (13).⁹

The added complication of a level of x-structure to house these privative features in fact fills a lacuna of the X' model, which does not formally accommodate information-passing between nodes, and leaves integer-valued features to imply dominance relations across projections, and among the bar levels of a projection. In the current model, information is explicit, dominance relations across projections are stipulated by a universal principle, and bar levels are derived from dominance relations.

However, in contrast with these improvements, the current model lacks sufficient resources to define a category C distinct from I. This issue is addressed in the next section.

Finally, in regards to mixed categories, it seems plausible to assume the following: the model of Bresnan and Mugane (2006) summarized in §4.1 can be revised such that

⁹The proper analysis of conjoined structures may require additional phrase structure rules. The topic of conjunction is not addressed in the X' theory of Bresnan (2001) or Toivonen (2001, 2003). Since I am concerned here with restating their treatments, I follow them in leaving the topic of conjunction aside.

the correspondence function χ replaces the labeling function \mathcal{L} , letting mixed category constraints be imposed on the set of x-structures thus obtained.

6 Syntactic categories and inflectional categories

In this section I propose a revision of the model just presented, which enables the statement of a C–I distinction, but otherwise downgrades the intention of hewing to the Bresnan–Toivonen line. A byproduct will be the possibility of using the syntactic category feature system to specify inflectional category distinctions.

The primary changes are as follows: a substitution of the feature assortment in Table 3 instead of those in Table 2, and the use of multisets instead of sets in the formalization of x-structure. Assume that the universal association principles of (16) are restated to accord with this new feature inventory.¹⁰

CATEGORY TYPE	
Lexical	Functional
A: $\{v, n\}$?: $\{v, n, f\}$
V: $\{v\}$	I: $\{v, f\}$
N: $\{n\}$	D: $\{n, f\}$
P: $\{\}$?: $\{f\}$

Table 3: Revised x-structure feature combinations

The feature assortment in Table 3 is a return to the X' features of Chomsky (1981), $\pm V$ 'predicative' and $\pm N$ 'substantive', in the sense that a positive value for those corresponds here to the presence of v and n in an x-structure, and a negative value corresponds to their absence.¹¹ The feature f is as before.

Table 3 does not exhaust possible feature combinations in multiset x-structure, which allow elements to occur more than once: the x-structures $\{n\}$ and $\{n, n\}$ are distinct, as are $\{v, f\}$ and $\{v, f, f\}$. Iterations of f and iterations of v or n can be put to use as follows.

¹⁰The changes required are as follows: for statements of equality, the new feature sets can be substituted for the old with no further changes; for statements of set inclusion, v can be substituted for Pr without further changes, but where n is substituted for Tr a statement of inclusion must be converted to a statement of non-inclusion, and vice-versa.

¹¹See also footnote 5.

6.1 Iteration of f

The iteration of the feature f within x-structure is a syntactic device that serves to distinguish an unbounded number of functional projections for classifiers, determiners, case, negation, tense, complementizers, and so on, to the extent that these are found to be necessary. Because of (16f), these will be strictly hierarchically ordered by f cardinality, with one- f decrements: there will always be a projection $\{n, f\}$ between $\{n, f, f\}$ and $\{n\}$.

Is this too strict? I propose that it is not: it constitutes an argument for an alternative conception of syntactic categories, to be implemented not with $x = \{\dots\}$ equality statements but with $\{\dots\} \sqsubset x$ subset statements. A word lexically specified for $\{v\} \sqsubset x$ could in certain situations be additionally constrained by $\{f\} \sqsubset x$ and thus only be licensed in a $\{v, f\}$ projection. For instance, a French verb will minimally bear the constraint $\{v\} \sqsubset x$, but the further requirement $\{f\} \sqsubset x$ is imposed on a finite verb. Absent further constraints, a nonfinite French verb thus has x-structure $\{v\}$ (=V), and a finite verb has x-structure $\{v, f\}$ (=I); by (16f) and (16b) it follows that the latter but not the former will head the projection which licenses the SUBJ, which is the correct result, shown in (18b).

This reconception also has the benefit of not implying a supernumerary lexical inventory of category types, as found in the Cartography literature (Cinque, 2002): a projection is there when needed, with its f -ness jacked up contextually by additional constraints, based on a small number of lexically specified x-structure types.

6.2 Iteration of v or n

Iteration of f does some useful syntactic work; what about iterations of the features n and v ? I proposed that they are ignored in syntax, but serve within the morphology to distinguish inflectional classes of verbs, nouns, and adjectives — but not adpositions, since they have no features to iterate. So the x-structures $\{v\}$, $\{v, v\}$ and $\{v, v, v\}$ can serve to distinguish between Vs of three distinct conjugation classes; $\{v, f\}$, $\{v, v, f\}$ and $\{v, v, v, f\}$ will then be the x-structures of their respective IPs.

Suppose that a language does have these three conjugation classes, and furthermore that verbs of classes $\{v\}$ and $\{v, v\}$ truncate in some instances, while verbs of classes $\{v, v\}$ and $\{v, v, v\}$ epenthesize in some other instances. Then realization rules for conjugation can take the form: epenthesize only if $\{v, v\} \sqsubset x$, and truncate only if $\{v, v, v\} \not\sqsubset x$. A verbal form with either x-structure $\{v, v\}$ or $\{v, v, f\}$ will both epenthesize and truncate.

This is compatible with a Paradigm Function Morphology take on the morphological component of grammar: in Stump (2001),¹² inflectional class features are distinct from

¹²A finite-state implementation is given in Karttunen (2003), which seems to me to be lacking Panini's Principle, a central component of Stump (2001); Malouf (2005) supplies the necessary addendum.

morphosyntactic features, and constrain the application of realization rules at the same point as lexical categories do. In the revised x-structure approach, the features used to distinguish inflectional classes *are* the lexical category features, namely those whose iteration has no syntactic usage but is part of the model because of a need to distinguish C from I.

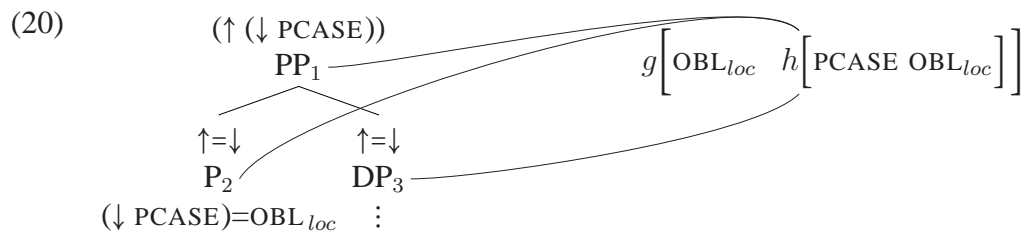
Of course, if the x-structure is going to have categorial import for both m- and f-structural concerns, then by the logic of the correspondence architecture it should be removed from its placement in (11) and interpolated between the m- and the c-structure. That would require more significant changes to the model in §5 than the tweaks I have applied in this section.

6.3 Minding the Ps

Lacking categorial features in the revised system of Figure 3, adpositions are distinct from the other categories in feature specification, and per §6.2 in their inflectional inertness.

Suppose now a variant of the system, which treats adpositions per Grimshaw (2000), as part of an extended projection P–D–N. In this system, an adposition will have the x-structure $\{n, f, f\}$, and will thus lack both featural distinctness (as n is shared by N and D) and inflectional inertness (since n can be iterated). There are other consequences.

First, by (13a,f) an adposition with the x-structure $\{n, f, f\}$ will share a functional domain with its complement DP. This forces a turn away from a common analysis of adpositional phrases, in which the complement DP is an OBJ selected by the adposition. But it provides a theory-internal motivation for the (\uparrow (\downarrow PCASE)) analysis of obliques, going back to Kaplan and Bresnan (1982): in (20), the adposition lacks a PRED feature but has a PCASE attribute whose value sets the PP’s grammatical function.¹³



Second, the Grimshavian amendment leaves a gap in the category typology, as no categories obviously fit x-structures like $\{ \}$, or $\{f\}$, or $\{f, f\}$, etc. If such categories exist, then their f-structures are allowed to be OBJ-taking, per (16c). However, it may

¹³An alternative analysis path would be to generalize beyond case-marking the inside-out designators of Nordlinger (1998): a PP would be annotated (\uparrow GF)= \downarrow , and individual Ps would set GF with inside-out designators like ($\text{OBL}_{loc} \downarrow$).

be possible to motivate stipulations that x-structures cannot be empty, and that the feature f can only be present when n or v is. To enforce the latter restriction in grammars with a Paradigm Function Morphology, that framework's property cooccurrence restriction mechanism (Stump, 2001: 41) could be recruited.

7 Conclusion and prospect

The x-structure model of syntactic categories implements a feature system on part of speech labeling categories, although one perhaps not quite as weak as envisioned by Kaplan (1987: 351). Unlike Kaplan's atomic category approach, in a manner to which the complex categories of Crouch et al. (2008) are unsuited and for which they were not designed, the system as a whole elegantly models both the notions of projection and extended projection. The treatment I have proposed improves on the X' theory of Bresnan (2001) and Toivonen (2003): by allowing a reformulation of its insights with a reduced number of theoretical primitives, and by being formally integrated in the correspondence architecture.

The model can potentially be altered to simultaneously give an account of syntactic and inflectional categories, as discussed in §6, concretizing a possibility alluded to in Marcotte and Kent (2010).

That same paper assumed a realizational Paradigm Function Morphology in line with Sadler and Nordlinger (2004), but suggested eliminating a distinction of that treatment, in which morphosyntactic features are transduced into f-structure equations. Marcotte and Kent (2010) proposed letting the morphological system handle f-structure equations directly: the equations *are* morphosyntactic features.

I perceive the x-structure model presented here as part of a generalization of that idea, in which the morphology accesses structural description equations for all levels of structure: f-structure, x-structure, a-structure, s-structure, i-structure, etc. Such a morphology, if implemented as realizational in the PFM mould, would contain: a specification of possible structural description equations for each level of representation, a paradigm space defined by cooccurrence restrictions within equation sets, a list of lexemes each with one or more underlying stems, and rules associating a phonological form with every licensed combination of lexeme and equation set. This project I set aside for another day.

References

- Alsina, Alex, Mohanan, Tara and Mohanan, KP. 2005. How to Get Rid of the COMP. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG '05 Conference*, University of Bergen, Norway.

- Asudeh, Ash. 2012. *The Logic of Pronominal Resumption*. Oxford: Oxford University Press.
- Bresnan, Joan. 1994. Locative inversion and the architecture of universal grammar. *Language* 70(1), 2–131.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers.
- Bresnan, Joan W. and Mugane, John. 2006. Agentive nominalizations in Gĩkũyũ and the theory of mixed categories. In Miriam Butt, Mary Dalrymple and Tracy Holloway King (eds.), *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*, Stanford, CA: CSLI Publications.
- Butt, Miriam, Dalrymple, Mary and Frank, Anette. 1997. An architecture for linking theory in LFG. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG '97 Conference*, University of California, San Diego.
- Butt, Miriam, King, Tracy Holloway, Niño, María-Eugenia and Segond, Frédérique. 1999. *A grammar writer's cookbook*. Stanford, CA: CSLI Publications.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 1995. Bare phrase structure. In Hector Cambos and Paula Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, pages 51–109, Washington, DC: Georgetown University Press.
- Cinque, Guglielmo (ed.). 2002. *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Vol. 1.* Oxford: Oxford University Press.
- Crouch, Dick, Dalrymple, Mary, Kaplan, Ron, King, Tracy Holloway, Maxwell, John and Newman, Paula. 2008. XLE documentation, Palo Alto Research Center.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*, volume 34 of *Syntax and Semantics*. New York: Academic Press.
- Falk, Yehuda N. 1984. The English auxiliary system: A Lexical-Functional analysis. *Language* 60(3), 483–509.
- Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey and Sag, Ivan. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

- Grimshaw, Jane. 2000. Locality and extended projection. In Peter Coopmans, Martin Everaert and Jane Grimshaw (eds.), *Lexical Specification and Insertion*, pages 115–134, Amsterdam: John Benjamins.
- Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson and P. Bennett (eds.), *Linguistic Theory and Computer Applications*, pages 149–188, New York: Academic Press, also published in Dalrymple et al. (1995).
- Kaplan, Ronald M. and Bresnan, Joan. 1982. Lexical-Functional Grammar: a formal system for grammatical representation. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, pages 173–281, Cambridge, MA: The MIT Press.
- Karttunen, Lauri. 2003. Computing with Realizational Morphology. In Alexander Gelbukh (ed.), *Computational Linguistics and Intelligent Text Processing, Lecture Notes in Computer Science*, volume 2588, pages 205–216, Heidelberg: Springer Verlag.
- Malouf, Robert. 2005. Disjunctive Rule Ordering in Finite State Morphology, presented at *41st Meeting of the Chicago Linguistics Society*, <http://bulba.sdsu.edu/malouf/papers/cls05-slides.pdf>.
- Marcotte, Jean-Philippe and Kent, Kateryna. 2010. Russian verbal affixes in the projection architecture. In *Proceedings of the LFG 2010 Conference*, pages 353–373.
- Nordlinger, Rachel. 1998. *Constructive Case: Dependent Marking Nonconfigurationality in Australia*. Stanford, CA: CSLI Publications.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365–424.
- Sadler, Louisa and Nordlinger, Rachel. 2004. Relating morphology to syntax. In Louisa Sadler and Andrew Spencer (ed.), *Projecting Morphology*, pages 159–186, Stanford, CA: CSLI Publications.
- Stump, Gregory T. 2001. *Inflectional Morphology: a Theory of Paradigm Structure*. Cambridge, UK: Cambridge University Press.
- Toivonen, Ida. 2001. *The Phrase Structure of Non-projecting Words*. Ph. D.thesis, Stanford University.
- Toivonen, Ida. 2003. *Non-Projecting Words. A Case Study of Swedish Particles*. Dordrecht: Kluwer.