

# Chapter one

## Phrase Structure in Principles and Parameters Syntax

---

### The issues

This chapter explains and analyzes the concepts which we examine in the three alternative P & P approaches to PS. As these concepts are to orient and guide us through the subsequent chapters, it is crucial that we have a tolerably clear understanding of them at the outset. However, it is not necessary to try for absolute clarity (whatever that might be) – largely because it is not possible. It is not possible, of course, because ultimately theoretical concepts such as the ones we explore in this chapter can only be truly well understood within a particular theoretical setting. One of the signal benefits of working at the end of the 20th century is that we no longer labor under the delusion that in empirical inquiry we must “define our terms”, preferably at the outset. We *discover* what our terms mean in a constant, reciprocal process of refining and bootstrapping as we engage in both theoretical and analytic work.<sup>1</sup> So, we shall rough out the contours of our investigations here, and in the later chapters we will see just how they are sharpened and articulated in different settings.

We have already listed some concepts and questions explored in this chapter.<sup>2</sup> I repeat that listing now.

- 1 The relation of argument structure to syntactic structure:
  - What is the position of Subjects in PS?
  - What are the statuses of Adjuncts (that is, non-argument modifiers of a Head) and adjunctions (that is, the results of non-substitution movement)?
- 2 Headedness & endocentricity:
  - Are all constructions endocentric?
  - Is there a unique Head?
- 3 Functional vs. Lexical Categories:
  - What is the theoretical justification for the “generalization of X-Bar” to Functional Categories?
- 4 X-Bar Theory and its status:
  - If X-Bar rules should not be explicitly stated, how are their effects to be derived?
  - Is all branching binary?

- 5 The respective statuses of precedence and dominance:  
Are there two formal primitives for PS or just one?

Though there is inevitably a large degree of interconnectedness among them, hence some degree of arbitrariness in doing so, I shall nonetheless now take them up in the order just presented.

## 1.0 Structuralization and argument alignment

---

The basic issue here is: how does lexical information get “structuralized”?<sup>3</sup> The guiding idea here is that there must be some systematic relation between individual lexical requirements (i.e., argument structure) and syntactic structure – lexical items cannot appear in arbitrary places in a well-formed sentence – and that characterizing this relation is to some large degree what the theory of PS is about and for. In particular, then, there must be (a set of) “canonical structuralizations” which are the systematically encoded/decodable syntactic representations of lexical information. The reasoning behind this view is pretty familiar and clear. We are interested in what one knows when one knows a language. Part of that knowledge is knowledge of lexical items. Part of what it means to know a lexical item is to know the particular restrictions that item carries with it in terms of what it can or must cooccur with. This sort of lexical knowledge is not, strictly, grammatical knowledge, and has to be rendered in a format appropriate to (knowledge of) the grammar, given that grammar is the means by which lexical information is made recoverable for wider cognition (i.e., thought). In the P & P tradition, this systematic grammatical format is PS. The issue then is sharpened: what are the forms of information in the lexical entry, how much of this information is structuralized, and according to what principle(s)? Borer (1994) calls this “the lexical-entry-driven-approach” and credits Ann Farmer’s 1980 MIT dissertation (cf. Farmer 1984) with its first explicit statement.<sup>4</sup> Stowell (1992: 11–13) refers to this central issue as “argument alignment”; Speas (1990) provides discussion within “standard” P & P; Pesetsky (1995) offers a detailed nonstandard approach.<sup>5</sup> We shall only indicate briefly some main lines of inquiry and dispute.

Chomsky (1981) suggested that syntactic subcategorization information is structuralized as D-structure using defined predicates such as “government” to regulate the mapping. Grimshaw (1979, 1981) argued that as well some sort of semantic selection was required in the lexicon and in structuralization. Chomsky (1986b), drawing on David Pesetsky’s unpublished 1982 MIT dissertation, argued that only the semantic selection (s-selection) was needed, and that syntactic selection (c-selection) was redundant. Rothstein (1991a, 1991b) countered that not all c-selection is in fact eliminable.

Chomsky (1981, 1986b) also suggests that D-structure is “a pure representation of theta structure” (1986b: 100). This idea is taken up and its implications extensively traced out by Speas (1990), Lebeaux (1988, 1990), and Chametzky (1996).<sup>6</sup> An influential proposal concerning argument alignment was put

forward by Baker (1988) in his Uniformity of Thematic Assignment Hypothesis (UTAH), given in (1); Bouchard (1991) proposed a Relative Theta Assignment Hypothesis (RTAH), given in (2).<sup>7</sup>

- (1) UTAH: Identical thematic relationships between items are represented by identical structural relationship between those items at the level of D-structure. (Baker 1988: 46)
- (2) RTAH: The argument which is relatively highest in LCS [Lexical Conceptual Structure – RC] is linked to the position which is relatively highest in syntactic structure, and so forth, for the second highest, etc. (Bouchard 1991: 28–9)

We may note that RTAH differs from UTAH in that the former, but not the latter, seems to presuppose some sort of hierarchy of semantic/thematic arguments and that this hierarchy mediates structuralization. However, this is a misleading way to characterize Bouchard’s proposal. In fact, Bouchard is revising Jackendoff’s (1983) Conceptual Structure (CS), and the notion of “higher” in (2) is basically one of constituency both in LCS and syntax. Structuralization on this view means that “. . . all syntactic nodes correspond to CS constituents . . .” and “two nodes A and B in SS [syntactic structure – RC] . . . correspond to two nodes A’ and B’ in CS . . .” (Bouchard 1991: 24, 25). This is in contrast to “true” hierarchy positions with respect to argument alignment, which date back to Fillmore (1968) (see also Jackendoff (1972)), in which there is an ordered set of thematic/semantic roles, and these are structuralized in accordance with their positions in this order.

Stowell (1992: 12–13) points out that this approach is not usually one found in P & P work, as it is possible to map more directly from such a hierarchy to a “surface” string, without mediation by means of phrase structure. Thus, as he notes, such hierarchies are generally found in work that is not exclusively or at all PS-based, and this forms one area for disagreement among approaches to syntax. However, Speas (1990: 14–16) does adopt a version of the hierarchy approach, with the proviso that she is not committed to “. . . claiming that some primitive device like [the hierarchy she gives] exists as a part of the grammar. The thematic hierarchy *describes* the order of the arguments in the theta grid” (emphasis in original).<sup>8</sup>

More commonly accepted within P & P work is an asymmetry between an *external* and all *internal* arguments (Williams 1981).<sup>9</sup> The external argument is the “most prominent” argument of a lexical item, and is structuralized in a position outside a Maximal Projection headed by that item and containing any internal arguments. Which argument is the external argument is typically taken to be part of the lexical information specified in a lexical entry.<sup>10</sup> This distinction also leads us directly to the next subtopic, the position of Subjects in PS.

## 1.1 *Subjects in PS*

The intuition that Subjects are somehow different from any other arguments is an old one. The notion of “external argument” provides one way to begin to

theoretically reconstruct this intuition. It does not by itself, however, have necessary implications for a PS mediated interpretation of the intuition.<sup>11</sup> There are several issues in play with respect to Subjects and PS. Does the Subject appear within a Maximal Projection of the item of which it is an argument? Does it appear outside such a Maximal Projection? Are these mutually exclusive options? What principle(s) determine the answers to these questions? Are there principles which refer to Subjects as such?

Much, perhaps most, P & P work seems to accept the following idea: “. . . all of the arguments of [a] predicate are in some sense internal to a projection of the predicate.”<sup>12</sup> Speas calls this idea the Lexical Clause Hypothesis (LCH), and it is one of the main (though not the only) justifications for the currently widespread view that Subjects originate structurally within VP (or, more generally, any maximal category headed by a predicate).<sup>13</sup> The basic idea here is theoretical parsimony: if structure is a representation of lexical information, and if an argument-taker licenses structure within which its arguments are to be found, then the simplest proposal is that all of its arguments are to be so found. Accepting the LCH still leaves open the question of what category a Subject is sister to, however.

It might seem that this is not so. If Subjects are within maximal projections, then, presumably, they are not also sisters to them. This is the position of Speas (1990). Yet many proponents of versions of the LCH in fact have Subjects as both daughters and sisters to maximal projections.<sup>14</sup> There are both empirical and theoretical reasons for so doing. Speas herself (1990: 128–38) provides a range of empirical evidence that strongly suggests “that the underlying structure of English is hierarchically organized in such a way that the subject is structurally superior to the objects and is outside of a maximal constituent of VP” (Speas 1990: 138). Theoretical grounding is provided by the theory of Predication (Williams 1980; Rothstein 1985).

Predication Theory is a syntactic theory, and one that explicitly refers to Subjects. As developed by Rothstein (1985: 7) there are three central claims:

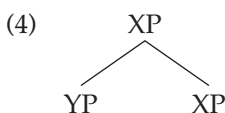
- (3) a. A predicate is an open one-place syntactic function requiring SATURATION.
- b. The syntactic unit which may be a predicate is a maximal projection.
- c. APs, VPs, and PPs must always be predicates; NP and S' may be predicates.

A Subject, on this view, just is that argument which saturates a Predicate. As Predicates must be maximal projections, it follows then that a Subject is outside of the maximal projection of which it is the Subject.<sup>15</sup> If you want both Predication Theory and the LCH, then Subjects are going to be both within and outside of the maximal projection of the verb.<sup>16</sup> But, as Speas (1990: 102) points out, such a proposal places the Subject “. . . in an underlying adjunction configuration [actually, an *adjunct* configuration–RC], which violates the strictest forms of X-bar theory. . . .” This means that in order to combine Predication Theory and the LCH, one needs an account of Adjunct structure. But one needs that anyway, and it is our next subtopic.<sup>17</sup>

1.2 *Adjunct(ion)s*

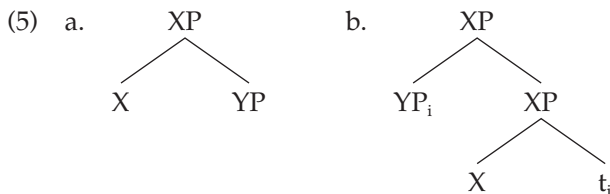
Here we double back to the notion of “canonical structuralization” in that these are exactly cases of **noncanonical** structures, as explained below. There are two distinct topics here which are often run together.<sup>18</sup> One is *adjunction*, a syntactic movement operation, one often taken to mediate, for example, the S-Structure-to-LF relationship. The other is *Adjuncts*, nonargument modifiers of a Head.<sup>19</sup>

It is not really so surprising that these two have been treated together. They share, so it seems, a structural configuration unique to them, exemplified in (4). As noted, this is a noncanonical structure. The crucial fact is that the mother and one of the daughters have the same label.<sup>20</sup>



I begin with adjunctions, then move on to Adjuncts.

Adjunctions actually relate the structure in (5a) to that in (5b), not that in (4).



The crucial properties of the adjunction rule are (1) it is a movement rule, (2) it creates a new node, (3) the new node is the after movement mother of the moved element and a second node, and (4) the new mother node bears the label of its nonmoved daughter (the host). Adjunction is taken to mediate both the S-Structure-LF mapping, as noted previously, and the D-Structure-S-Structure relationship. The issue for PS theory is how, if at all, the theory helps explain the existence of a rule with the four properties just listed. PS theory is only indirectly relevant to property (1). It is this property which marks adjunctions as noncanonical structures; that is, the attempt to develop an account of the nature of noncanonical structures within a PS-based approach to syntax is the essence of “movement”. So, PS theory may not shed any direct light on the movement, but the fact that the overall approach is strictly phrase structural throughout effectively requires movement if noncanonical structures are to be incorporated and, it is hoped, explicated. However, PS theory could or should shed light on each of the other three, as these involve elements (nodes) or relations (mother, labelling) ostensibly from PS theory.

For property (2), the issue is the status and licensing of nodes. Most P & P theorists take nodes as a basic building block of PS.<sup>21</sup> A theory that did not presumably could not have a rule with the second property; either there is no

adjunction rule in such a theory, or there is some variation on (2) that does not invoke “nodes”. Regardless of how this issue is resolved,<sup>22</sup> there is a deeper one: should a **rule** be such as to license additional structure? This engages two more fundamental issues. First, how is structure generally licensed – are rules involved or not? Second, what is the status of **additional** structure: is it necessary? if it is not necessary, how is it possible? why and how is it allowed? We should like our theory of structure to give us insight into the creation and licensing of all structure, canonical and noncanonical.

A relatively popular approach to adjunctions stems from work by May (1985, 1989; see also Chomsky 1986a).<sup>23</sup> May stipulates a rule with the four properties above and offers redefinitions of the basic PS predicates *dominates* and *category* to accommodate structures as in (5) above.<sup>24</sup>

In addition to issues about nodes (structure), there are issues about labels (categories).<sup>25</sup> What mechanisms and principles are involved? Given that there are both canonical and noncanonical structures, we need to consider how and why each are labelled as they are. Are the mechanisms and principles the same for both canonical and noncanonical structures? If not, how do they differ and how are they related? May, for example, assumes the same restrictions apply,<sup>26</sup> and that this apparently requires distinct mechanisms.<sup>27</sup> Speas (1990: 42–6) adumbrates an approach to labelling which “collapses the labeling function of X-bar theory with the implicit free generation of hierarchical structures” (Speas 1990: 43). Chametzky (1996: 17–19, 91–5) revises this position, uncollapsing labels and structures.<sup>28</sup> Chomsky (1995) takes a quite different approach to labels and labelling.<sup>29</sup>

Chomsky (1986a: 6) put forward a proposal for adjunctions, though at best for empirical reasons, it seems. His proposal is a two-part stipulation: (1) that adjunction to an argument is prohibited and (2) that only maximal projections may be hosts. This latter has been dropped in much subsequent work, including that by Chomsky, as Heads have been widely used as hosts. A number of theoretical issues are raised here. If the stipulations are useful, we should, of course, prefer to derive rather than demand their effects. And if they are not useful – as, for example, many find the restriction on hosts – then the theoretical investigation into the nature of possible adjunctions is entirely wide-open.<sup>30</sup>

We see with respect to adjunctions, then, that there has been widespread agreement – though little real theoretical work – both on the existence and nature of such a rule,<sup>31</sup> with the dissent of Chametzky (1994, 1996), and on the basic nature of nodes and labels, with the dissent of Chomsky (1995). Moreover, even putting aside the dissents, there are questions and issues that come up about adjunctions which PS theory could and should answer and address. We turn now to Adjuncts.

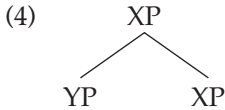
Adjuncts, recall, are nonargument modifiers of a Head. There are several diagnostics (NB, *diagnostics* are not a *definition*): (1) they are not mandatory<sup>32</sup> (2) they appear farther from the Head than do arguments (3) they iterate and (4) they do not appear in a specific order with respect to one another. Typical Adjuncts are, for example, locative and temporal PPs for V(P)s and relative clauses for N(P)s. Some examples illustrating the diagnostics are provided in (6).

- (6) a. Kim announced the winners.  
 b. \*Kim announced in the auditorium.  
 c. \*Kim announced at three o'clock.  
 d. Kim announced the winners in the auditorium.  
 e. Kim announced the winners at three o'clock.  
 f. \*Kim announced in the auditorium the winners.  
 g. \*Kim announced at three o'clock the winners.  
 h. Kim announced the winners in the auditorium at three o'clock.  
 i. Kim announced the winners at three o'clock in the auditorium.

Though the phenomena are distinct, similar questions and issues arise here as arose with adjunctions: why do we see the structure we do (that in (4))? How is this structure licensed: by rule? by general principle(s)? Why? Is there node creation? How is the labelling licensed? Is there any restriction on the label which a host may bear? Notice that if we can answer these questions satisfactorily with respect to Adjuncts, then we might actually have reason to doubt the existence of an adjunction rule with all the properties listed in the discussion above. In particular, as no one analyzes Adjuncts by means of a movement rule that creates and labels a new node, if we find we are able to license and label a new node without such a rule in this case, we would want to know why whatever was involved did not carry over to the situation of the new node and label in adjunctions. It could be, of course, that there would be good theoretical reasons ruling out such carry over; but then we would be faced with a situation in which these two phenomena would have entirely distinct routes to their convergence on a common structure found only in these cases. Not impossible, by any means, and reminiscent of the distinction in evolutionary biology between homologous and analogous traits; but still, in generative grammar at least, a highly disfavored theoretical state of affairs.

Within the approach to structuralization explored by Lebeaux (1988, 1990) and Speas (1990), which, as noted above, takes D-structures to be “pure representations of theta grid requirements” there is no obvious way to include Adjuncts – definitionally not required by theta considerations – at D-structure.<sup>33</sup> The approach thus offers a relatively transparent explication of the noncanonical status of Adjuncts, though no straightforward clue as to how to integrate them into structure.<sup>34</sup>

For adjunct(ion)s as in (4) (repeated here), then, the crucial point, to reiterate, is that the mother and one daughter bear the same label. We have the following issues and questions. Is (4) correct in that the host is always a maximal projection; why should this be either true or false? Why, in the case of adjunction, do we not see some other adjunction – daughter-adjunction, for example?<sup>35</sup> Why not a label distinct from either that of the host or that of the adjoiner? How are nodes licensed: is this separate from labels; is it done by rule or by principle? Given the noncanonical nature of such structures, might it be appropriate to invoke revisions of basic predicates and relations to license them? Why should there be noncanonical structures anyway? Should we expect the same licensing principles to apply to Adjuncts and to adjunctions?



In discussing Adjuncts, crucial mention (and use) has been made of the notion “Head”. We now turn to directly examining this concept.

## 2.0 Heads in PS

---

“Head” is the basic notion of DT; central to X-Bar theory has been the importation of this concept into PS-based syntax.<sup>36</sup> A Head is generally taken to be the lexical item which determines the syntactic nature of a larger unit (a phrase) of which the Head is a part. Typically, a Head is taken to determine the syntactic category of the larger unit, to be a mandatory constituent of that phrase, and to determine to some degree what else, if anything, may or must also occur in the phrase. Though there are questions and problems concerning the notion “Head” itself, we shall not address them.<sup>37</sup> Instead, we shall simply accept that “Head” is a real syntactic notion, with a role to play, as is now customary in PS-based P & P work. Our questions come down to the two already raised on p. 8. (1) are all constructions endocentric? and (2) is there a unique Head? These questions are related, since, if (1) is false, then (2) must be also, and if (2) is false, we might wonder about whether (1) is true. They are also related to the topic of the next section, Functional Categories, as these often are taken to expand the set of phrases to which (1) and (2) might or must pertain. “Head” is also crucial to the nature of X-Bar Theory.<sup>38</sup>

Endocentricity is largely assumed to hold for all phrases in current P & P work. Sometimes it is simply written into a set of rule schema (e.g., Chomsky 1986a); others, such as Speas (1990), Chomsky (1995), and Bouchard (1995) attempt an explication and analysis. PS-based work antecedent to or outside of P & P (e.g., Jackendoff 1977; Pullum 1991) does not always accept this view, and instead allows for exocentric analyses. The P & P analyst typically uses new categories (viz., Functional Categories) to cover those phenomena analyzed as exocentric by others. We shall presently have occasion to inquire into this move; for now, we simply note it.

Head uniqueness is also generally assumed in P & P work. Indeed, it seems to be taken as hardly worth discussion, more or less an inevitable result of other aspects of the theory (viz., endocentricity, the nature of “projection”). However, Bouchard (1995) and Williams (1994) disagree. Bouchard (1995: 75; 457, fn. 5) proposes a parameter of “endocentricity value” by which specific languages allow only one or more than one Head. Williams (1994: 11f.) argues specifically that coordinate structures and clauses are doubly headed. More generally, Williams (1994: 45–51) develops a notion of “relativized head” that goes along with, and is dependent on, the more usual notion, which he calls the “absolute head”. This notion allows a non (absolute) Head to have some Head-like properties (e.g., feature projection) in a given configuration.



We should also take note of the question of “empty Heads”. The issue here concerns when or if there is projection from a Head. Given the “lexical-entry-driven” approach, this comes down to what must there be in a lexical entry for a Head to project? This can have some important implications, particularly for the theory of movement.<sup>39</sup> We shall briefly mention two proposals.<sup>40</sup>

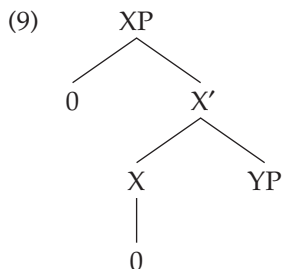
Speas (1994: 186, her (11)) proposes “a principle of Economy” on representations (7). As she notes, this can only do any work if its notion *content* is itself given content. She therefore suggests (8) (Speas 1994: 187, her (13)):

- (7) Project XP only if XP has content.
- (8) A node X has content if and only if X dominates a distinct phonological matrix or a distinct semantic matrix.

Speas includes both phonological and semantic matrices because “all structures must be interpreted at both of the interface levels, PF and LF . . .” (1994: 187). She applies (8) to a structure as in (9) (1994: 187, her (12)) in the following way (emphasis in original):

If XP in [(9)] dominates no phonological material except that which is in the complement YP, then XP dominates no distinct phonological matrix. Similarly, if XP dominates no semantic material except that which is in the complement YP, then XP dominates no distinct semantic matrix. (1994: 187)

The implications for movement are explicitly drawn by Speas (1994: 187): “Thus, radically empty projections with the sole purpose of serving as landing sites for movement are disallowed.”



Bouchard (1995: 25) develops an approach to syntax which arrives at basically the same conclusion that “there cannot be semantically unlicensed ‘open’ positions, which only appear as landing sites for syntactic operations.” He argues (1995: 93) for a particularly strong version of the “lexical-entry-driven” approach: “[t]he null assumption . . . is that the semantic formatives in the lexical representation of an item are all identified in some way in the sentence where the item occurs.” He enforces this assumption, which drives his entire theoretical undertaking, with (10), his Principle of Full Identification (1995: 93–4, his (31)).

- (10) *Principle of Full Identification*: Every syntactic formative of a sentence must have a corresponding element in the semantic representation. Every formative of a semantic representation must be identified by a morphosyntactic element in the sentence with which that representation is associated.

The point in bringing up this issue is the following. If one adopts both some version of the “lexical-entry-driven” view of PS and some kind of “accountability” requirement (Bouchard 1995: 95)<sup>41</sup> on what can and must appear in a syntactic structure, then the justification of “empty Heads” can become a problem. It may be possible to dissolve the problem (Chomsky 1995), but it should not be assumed that analyses developed within other assumptions can be simply carried over without problem. Very often, I think, analytic work proceeds without this sort of theoretical reconnoitering, sometimes, at least, to bad effect.

Even this brief roughing out of the Head territory has proven impossible to do without reference to the topic of the next section, Functional Categories. So we move directly to that discussion.

### **3.0 Functional vs. Lexical Categories**

---

The founding notion of the P & P approach to PS is that structure is not arbitrary, but is rather a means of representing information carried by lexical items in their lexical entries, where this information is both the category of the lexical item itself and restrictions on its cooccurrence with other items or phrases.<sup>42</sup> Understandably enough, this project began with those items which traditionally were thought to carry such lexical information: Nouns, Verbs, Adjectives, some Prepositions – the categories traditionally dubbed “content words” or, more theoretically, “Open Class items” or “Lexical Categories”. Opposed to these are the “function words”, or “Closed Class items” or “Functional Categories”: Determiner, Complementizer, INFL, perhaps Conjunction.<sup>43</sup> A striking feature of much P & P work of the last decade has been the extension of the structuralization relation from Lexical Categories to Functional Categories, along with a concomitant proliferation (or, perhaps, discovery) of further Functional Categories.<sup>44</sup>

For the PS theorist, the basic question is what theoretical justification is there for this development. There may be a temptation to point to various analytic successes that this development has afforded, but, while analytic success is, of course, the ultimate arena of evaluation, it is not the only such arena. And, from a theoretical point of view, we need not be overly impressed, perhaps, by the analytic successes, as it should be expected that expansion of the theoretical apparatus leads to increased analytic coverage. More interesting is to achieve greater analytic coverage without expanding the theoretical apparatus. Nonetheless, there is no inherent virtue in waiting for theoretical or conceptual clearance for expansion before proceeding with analytic work, and, indeed,

such pusillanimity may create a pointless obstacle to inquiry.<sup>45</sup> But as we are here engaged in theoretical analysis and inquiry, we do ask for something other than analytic justification. And, indeed, the discussion in this subsection will ultimately be more evaluative than has so far been the case, as we shall see.

### 3.1 Chomsky (1986a) and the “generalization” hypothesis

Though early in P & P work (e.g., Chomsky (1981), Stowell (1981)) it is suggested that the syntactic constituents then labelled *S* and *S'*, and which were only problematically analyzed endocentrically,<sup>46</sup> might better be seen as headed by INFL or COMP, these constituents were nonetheless not afforded “full projection” status until Chomsky (1986a). Chomsky there asks whether the system regulating Lexical Categories “extend[s] to the nonlexical categories as well.” He immediately answers, “[e]vidently, the optimal hypothesis is that it does” (1986a: 3). This is the extent of the theoretical argument and justification. The argument is familiar and not unconvincing: if there is a set of clear cases (in this instance, the Lexical Categories) for which we have a theoretical proposal, then the null hypothesis is that this carries over to less clear or central cases (Functional Categories) until and unless there develops strong evidence to the contrary.<sup>47</sup>

So, starting with Chomsky (1986a), we get IP and I (= INFL), CP and C (= complementizer), and then the further “generalization” to DP and D (= determiner) (this is developed especially in Steven Abney’s unpublished 1987 MIT dissertation), the “splitting” of IP into AGRP and AGR (= agreement) and TP and T (= tense) (beginning with Pollock (1989)), along with subsequent or contemporaneous suggestions for various other Functional Heads and Categories (e.g., NEGP and NEG for negation, ASPP and ASP for aspect). Spencer (1992: 313) calls the general idea the Full Functional Projection Hypothesis (FFPH).<sup>48</sup>

#### (6) Full Functional Projection Hypothesis

Any morphosyntactic formative which corresponds to a Functional Category in a given language is syntactically the head of maximal projection.

Despite its apparent status as a null hypothesis, there is nonetheless a basic theoretical question we might ask of the FFPH. If the guiding idea in P & P approaches is that PS is indeed “lexical-entry-driven”, why should Functional Categories be involved? Really, this is just questioning that FFPH is the null hypothesis. Sharpening the question a bit: given that *everyone* seems to agree that there is some real difference between the two sorts of categories, is it truly so obvious and straightforward to “generalize” as Chomsky (1986a) suggests?

There are a number of possible positions that some sort of negative answer to this question allows. Most radical is simple denial of the generalization,

limiting projection to Lexical Categories. No one in P & P seems to take this position explicitly, so we leave it (for now).<sup>49</sup> Instead, analysts suggest ways of maintaining both the distinction and the generalization. We shall look at Fukui and Speas (1986) (FS, hereafter) and Fukui (1995), Speas (1990) and unpublished works by Abney (1987), Grimshaw (1991), and Lebeaux (1988). We should bear in mind, however, that there is a tension between the “distinction” and the “generalization”: the greater the one is, the less motivation there is for the other, and any particular attempted reconciliation may strike us as unstable.

### 3.2 *Fukui and Speas (1986), Fukui (1995), and Speas (1990): Functional Heads*

FS (1986: 5–6) and Fukui (1995: 14) list the following four properties as ways in which Functional Categories contrast with Lexical Categories.

- (11) (i) Functional Heads have one and only one (i.e., non-iterable) Specifier, while the Specifiers of Lexical Heads may be iterable ones.
- (ii) The Specifiers of Functional Heads are often (in our model, always) moved from within their complement.
- (iii) All Functional Heads can have Specifier positions; it is not at all clear that all Lexical Heads have Specifier positions.
- (iv) Languages which lack Functional Heads also lack Specifier positions.

Fukui (1995: 12, 13) also incorporates the observations that Functional Categories, unlike Lexical Categories, do not “have theta-grids or ‘Lexical Conceptual Structures’” and, citing Abney (see below), that Functional Categories “are closed-class items, that they lack the sort of semantic value associated with Lexical categories, and that they always select a unique complement.”

Having characterized a distinction between Functional and Lexical Categories, FS (1986: 13–16) and Fukui (1995: 27–30) distinguish between two kinds of Functional Categories. They do this on the basis of *Functional Features* (*F-Features*). The idea here is to generalize the commonly accepted difference between tensed versus infinitival AGR/INFL (or whatever) in that the former, but not the latter, assigns nominative Case. The generalization is the suggestion that every Functional Category includes some such distinction, with F-Features being the cover term for the class of assigned properties. Besides nominative Case, genitive Case, assigned by ‘s and +WH, assigned by a WH-COMP are given as F-Features. A further new term Kase is also introduced, where this is defined as in (12) (Fukui 1995: 27, his (29)). Functional Categories have a specifier position only when Kase is assigned to that position.

- (12) Kase = Case U F-Features

The paradigm in (13) is derived on these assumptions (FS 1986: 14; Fukui 1995: 28, his (30)).

(13)

|                   | C    | I         | DET |
|-------------------|------|-----------|-----|
| Kase assigner     | +WH  | Tense/AGR | 's  |
| non-Kase assigner | that | to        | the |

These authors go on to work out various consequences and extensions of these ideas, but these do not directly concern us. The point, rather, is to get a sense of how they flesh out the bare suggestion of generalizing “projection from the lexicon” from Lexical Categories to Functional Categories.

Speas (1990: 110–16) revises somewhat the proposals made in FS. Speas (1990: 112) points out that FS made “an implicit assumption that the theory of phrase structure is relevant in restricting the occupants of the specifier position for functional categories, and that bar level is a primitive which may vary both cross-linguistically and from category to category.” Both of these ideas are rejected in Speas (1990), though the elaboration of Kase Features is retained and is used to account for the facts which the rejected assumptions accounted for in FS.

Speas also adumbrates an account of grids and Lexical Conceptual Structures (LCS) for Functional Heads, “grids which specify what sort of complement they take, and specify the Kase features that they have” (Speas 1990: 114). The examples of functional LCS are those for DET and INFL, which are analyzed as “theta binders” rather than “theta assigners”, with concomitant differences in LCS form:

The difference between lexical and functional heads is that the position in the Kase grid of a functional head is never linked to a variable in LCS. The reason is that there are no referential variables in the LCS of a functional head. A functional head is in an informal sense semantically parasitic on a predication, and so although it has Kase features to assign, it has no relevant variable in its LCS to which these might be linked. (Speas 1990: 116)

We see, then, that Speas (1990) disagrees with Fukui (1995) with respect to whether Functional Heads have LCSs, but agrees with the notion, which Fukui gets from Abney, “that they lack the sort of semantic value associated with Lexical categories”. Speas also carries over the restriction on number of specifiers of a Functional Head (see (11i) above), though now rejecting the PS account given by FS. Again, our point is not what can be done with the generalization once it is effected, but rather how and to what extent, if at all, the distinction between Lexical and Functional Categories is maintained given this generalization. We turn now to the work of Abney (1987), already alluded to by Fukui (1995).

### 3.3 Abney (1987): Functional Elements

Fukui (1995) takes over some ideas from Abney for characterizing Functional Categories.<sup>50</sup> Abney (1987: 64) also notes that the Functional versus Lexical distinction seems to play a role outside of grammar proper, in that children acquire the former later than the latter, and that in some aphasias there can be loss of ability to process Functional, but not Lexical, Categories. Abney (1987: 64–5) lists the following five properties, stressing that “none of the following properties are critical for classification as a functional element. . . .” (emphasis in original)

- (14) (i) Functional Elements constitute closed lexical classes.  
 (ii) Functional Elements are generally phonologically and morphologically dependent. They are generally stressless, often clitics or affixes, and sometimes even phonologically null.  
 (iii) Functional Elements permit only one complement, which is in general not an argument. The arguments are CP, PP, and . . . DP. Functional Elements select IP, VP, NP.  
 (iv) Functional Elements are usually inseparable from their complement.  
 (v) Functional Elements lack . . . “descriptive content”. Their semantic contribution is second-order, regulating or contributing to the interpretation of their complement. They mark grammatical or relational features, rather than picking out a class of objects.

Though they differ in their respective statuses (e.g., (iii) is evidently theory internal in a way that (i) and (ii) are not), none of these by now seem especially surprising or novel. Abney takes (v) to be the central characteristic, and it is crucial for motivating his distinction between *c-projection* and *s-projection* (Abney 1987: 57–8). C-projection (= “category projection”) is just the usual notion of syntactic projection (viz., V to VP, I to IP, etc.). S-projection (= “semantic projection”) “is the path of nodes along which its descriptive content is ‘passed along’.” Thus, Functional Categories are part of the s-projections of Lexical Heads, as Functional Elements have no “descriptive content” of their own, and serve to “pass along” the content of the Lexical Heads. Abney gives the following definition (1987: 57, his (47)):

- (15) b is an s-projection of a iff  
 a.  $b = a$ , or  
 b. b is a c-projection of an s-projection of a, or  
 c. b f-selects an s-projection of a

“F-selection” is the name Abney (1987: 56) gives to “the syntactic relation between a functional element and its complement.” He gives the example in (16) (1987: 58, his (48)), where the circled nodes are, respectively, the c-projection and s-projection sets of the lower V. We note that CP is the maximal s-projection of V, I, and C.



|      |   |         |      |      |
|------|---|---------|------|------|
| (17) | V | [+V,-N] | {F0} | (L0) |
|      | I | [+V,-N] | {F1} | (L0) |
|      | C | [+V,-N] | {F2} | (L0) |
|      | N | [-V,+N] | {F0} | (L0) |
|      | D | [-V,+N] | {F1} | (L0) |
|      | P | [-V,+N] | {F2} | (L0) |

Grimshaw (1991: 3, 4, 9, her (3), (4), and (12)) also provides the following definitions for *perfect Head/projection*, *extended Head/projection*, and the *Generalized Theta Criterion* (GTC). Perfect projection is just ordinary projection, as the F-feature is effectively irrelevant. Extended projection allows and constrains projection in terms of the F-feature value.

- (18) x is the *perfect head* of y, and y is a *perfect projection* of x iff
- (a) y dominates x
  - (b) y and x share all categorial features
  - (c) all nodes intervening between x and y share all categorial features
  - (d) the F value of y is the same as the F value of x
- (19) x is the *extended head* of y, and y is an *extended projection* of x iff
- (a) y dominates x
  - (b) y and x share all categorial features
  - (c) all nodes intervening between x and y share all categorial features
  - (d) If x and y are not in the same perfect projection, the F value of y is higher than the F value of x
- (20) *Generalized Theta Criterion*
- Every maximal projection must either
    - a. receive a role or
    - b. be part of an extended projection that receives a role.

As noted above, the goal of the analysis is to explain (im)possibilities in the combinatorics of Heads and their complements. Grimshaw (1991: 8 her (9)–(11)) provides the following three classes of cases. In (21) we find combinations which form extended projections as defined in (19) – an F-head with a complement that it is both categorially identical to and higher valued on the F-feature value than. In (22) we have combinations of an L-head with a complement. In (23) are combinations of an F-head with a complement with which it cannot form an extended projection. The combinations in (21) and (22) are common ones which must be licensed, while those in (23) do not occur, and thus must not be licensed.<sup>52</sup>

- (21) C-IP, P-DP  
 I-VP, D-NP  
 C-VP, P-NP
- (22) V-PP, V-DP, V-NP, C-CP, V-IP, C-VP  
 N-PP, N-DP, N-NP, N-CP, N-IP, N-VP



- (23) I-NP, I-DP, I-PP, I-CP  
 D-VP, D-IP, D-CP, D-PP  
 C-NP, C-DP, C-NP, C-VP  
 P-VP, P-IP, P-CP, P-NP  
 C-CP, P-PP, I-IP, D-DP

The two clauses of the GTC (20), along with the assumption that only the complements and specifiers of L-heads can be theta-marked, will distinguish the cases. The maximal projection complements in (22) fall under (20a), theta-marked complements of L-heads, while those in (21) fall (potentially) under (20b), as they form extended projections with F-heads which may be theta-marked. But those in (23) fall under neither, as they are complements of F-heads, but not ones with which they form extended projections. The combinations in (23) are impossible in principle: the complement can neither be theta-marked by the F-head nor form an extended projection with the F-head, and so cannot be assigned a role, as the GTC requires. As Grimshaw observes (1991: 9) this provides an account of a property we have seen already (emphasis in the original): “each f-head . . . occurs only with a very limited set of complements, quite typically only with one. This will follow if *f-heads take only complements that they form extended projections with*. . . .” Given the definitions and category analysis, L-heads cannot form extended projections with their complements: L-heads and their projections are {F0}, and so the projection cannot have a higher F-feature value than a complement, as required by (19d) for an extended projection.

The *perfect projection* versus *extended projection* distinction is clearly very similar to Abney’s *c-projection* versus *s-projection* distinction.<sup>53</sup> Indeed, it is not unreasonable, I think, to see Grimshaw’s work as something of a culmination of the entire line of Functional-Head Theory inquiry. So, before moving on to Lebeaux (1988), which, as we shall see, is importantly different from these other works, it is appropriate to look somewhat more closely at this line, and its zenith in Grimshaw (1991).

As we have noted at the outset of this section, everyone agrees that Functional and Lexical Categories are different. And, again as pointed out earlier, the issue in Functional-Head Theory is to investigate these differences within a framework that generalizes PS theory from Lexical to Functional Heads. Grimshaw develops Functional-Head Theory most directly and explicitly, and reaches the statement in (24) (1991: 41 her Principle 1.)

- (24) Only Lexical Heads Select, Syntactically and Semantically.

As a corollary of this, Grimshaw states that “[f]unctional heads have no selectional powers at all.” She argues for this view as follows (1991: 39–40). The usual view in Functional-Head Theory is to suppose that the generalization to Functional Categories is essentially a generalization of the *selection* relation from Lexical Heads to Functional Heads. But this position then takes the combinatoric (im)possibility facts illustrated in (21)–(23) above as arbitrary facts. And it ignores the salient dissimilarities between the two cases; Grimshaw (1991: 40) cites the familiar facts that Functional Heads take only one category

as a complement and that, as a result, there is little or no lexical variation within Functional Categories with respect to complement categories. Further, it ignores what Grimshaw (1991: 40) calls the “stability” of the relations between Functional Heads and their complements “both within a language and cross-linguistically.” Grimshaw cites as an example the fact that “C is always on top of I rather than vice versa. . . .” Again, this is unexpected if the mechanisms in play are “isolated stipulations about what goes with what.” She concludes (1991: 40) “that there is nothing substantive to the claim that the relationship between an F-head and its complement is one of selection.” Making the claim “does not bring to bear a set of principled restrictions; it merely allows the description of any observed combination.”

Grimshaw’s alternative is that not selection, but rather projection, determines the nature of complements to Functional Heads. Her notion of extended projection (along with the GTC and her analysis of categories) completely determines the range of (im)possible complements to Functional Heads. And, correspondingly, because Lexical Heads do not form (extended) projections with their complements, projection, whether extended or perfect, plays no role in determining the complements of Lexical Heads (1991: 41). “The character of the relationship between a functional head and its complement,” Grimshaw writes (1991: 40), “is quite dissimilar from that between a lexical head and its complement.”

I think there can be no denying but that Grimshaw is correct in her theoretical arguments and conclusions. Using *selection* as the mechanism to account for the relation between Functional Heads and their complements is unprincipled, un insightful, and, ultimately, no more than word play. In this, she, and now we, disagree with the Functional-Head Theory tradition; so much the worse for the tradition. However, we might now want to inquire a bit more closely into Grimshaw’s own approach. For if the tradition’s view of things is theoretically hopeless, it would be nice to know that the alternative bears up under some scrutiny, as otherwise a considerable problem arises. I happen to think Grimshaw’s approach does not bear up very well at all, and that there is, indeed, a considerable problem. I move first to Grimshaw, then the problem, and finally to Lebeaux (1988), who suggests a way out, I believe.

The {F} feature is crucial to Grimshaw’s project.<sup>54</sup> There are two crucial aspects of the {F} feature. First, it is not categorial. Second, it is not binary. We take up each in turn. Because it is not categorial, specification for {F} does not distinguish e.g., DP from NP. In order to help understand what {F} is, Grimshaw invokes “Bar-level specification”, represented as the feature (L), as also neither categorial nor binary. But it is an interesting fact about the most careful contemporary analyses of X-Bar theory – viz., Speas (1990) and Kornai and Pullum (1990) – that they do away with “Bar-level” as a primitive of the theory.<sup>55</sup> While this is not itself sufficient to give up on {F}, it is not particularly encouraging, either. We need to look a bit more deeply. We need to ask what it can mean, syntactically, to say, as Grimshaw (1991: 3) does, that “a category label . . . is analyzed as a triple” only one part of which is categorial. We need to ask how it is that information in a category label that is not categorial is available in and for syntax. Notice that if we did still believe in the need for “bar-levels”, there would be a disanalogy here between (L) and {F}. It is certain

that no one would want to maintain the position that, say, N, N' and NP are all *identical* in category. It was precisely the point of X-Bar projection that these are both not identical syntactic categories and also closely related syntactic categories.<sup>56</sup> This is of no small importance, as to claim that syntactic entities are of identical syntactic category is to be committed to their having identical syntactic behavior. At least, it is to be so committed within a PS-based approach to syntax. There just is no other sort of information in such an approach other than the syntactic category labels, based in word classes, and the constituent structure positions of such category labels. Grimshaw (1991: 2) says, “[t]he {F} value of a node is, in this theory, not part of its categorial analysis.” But this is to say that it is not part of syntax “in this theory”; either the feature makes a categorial contribution, distinguishing subcategories, or it is syntactically invisible and inert. Grimshaw herself makes the salient point, though she does not draw our conclusion:

The categorial theory which is the basis for extended projection makes explicit the hypothesis that a functional category is a relational entity. It is a functional category by virtue of its relationship to a lexical category. Thus DP is not just a functional category it is the functional category for N, as IP is for V. (Grimshaw 1991: 3, underlining in original)

The point to stress is that PS-based syntax can refer to a “relational entity” only if it is definable in PS terms. This is why, for example, we have the famous structural/categorial definitions of Subject and Direct Object as, respectively, NP immediately dominated by S and NP immediately dominated by VP.<sup>57</sup> This is not to say that Grimshaw’s observation might not be correct. It is to say, rather, that how to encode this observation in the vocabulary of PS-based syntax is none too obvious.<sup>58</sup> One may attempt to do as Grimshaw has done, but this changes the theory in obscure ways, allowing noncategorial information to play a categorial role.

It might be objected that I have placed undue weight on a name. That is, I have read too much into Grimshaw’s use of “categorial” for the [+/-N +/-V] features. If some other term had been used – say, “lexical” – no one would be so tempted to make the argument I have. What we have is not, on this view, some kind of major theoretical incoherence, but rather, at worst, a relatively minor terminological equivocation. This response is serious, and requires an answer. My answer depends on examining Grimshaw’s practice in two areas. One is what she says about the category status of Functional versus Lexical Categories. The other is how the {F} feature is actually put to use – which leads us back to our second aspect of the {F} feature, its nonbinarity. I note that this answer may not totally convince a hard-liner in that Grimshaw could be wrong or confused in her practice and this would not necessarily undermine the theoretical concepts: *abusus non tollit usum*. However, there would be considerable *prima facie* evidence in favor of my position, and a substantial burden of argument would be placed on my hypothetical interlocutor.

Grimshaw says explicitly (1991: 2) “D and N are of the same syntactic category, once we have abstracted away from the lexical/functional distinction.” This may still seem equivocal. Again perhaps equivocally, she writes (1991: 24)

“the neutralizability of N and V holds equally for D and I, P and C, since these are categorially the same as N and V.” However, Grimshaw also says (1991: 2), “[s]ince I and D are of the same category as their lexical counterparts, V and N, they must be distinguished by another property. . . .” And, later, she writes of C, I and V (1991: 39) that “[t]hese three are of the same syntactic category, and have F values which permit them to form an extended projection.” These seem to me about as unequivocal as could be hoped for. I think it is fair to say that for Grimshaw, at least, the terminological choice is a motivated, not an arbitrary, one: the “categorical” features alone really do determine an item’s syntactic category.

Let us now consider the use of {F}. How does Grimshaw account for the differences in distribution between, e.g., NP and DP? It is here that we confront the second crucial aspect of {F}, its nonbinary nature. For, of course, Grimshaw must use {F} to distinguish DP and NP, as they are categorially identical. It is in conditions (18d) and (19d) of her definitions for perfect and extended projections, repeated here, that the crucial distinctions are made. Given these, the analyses of DP as {F1} and NP as {F0}, and the GTC, also repeated here, DP and NP will not be licensed in the same environments. Thus, as argued above, noncategorical information not reducible to some combination of the PS conceptions of category and structure is used to account for syntactic distribution, a theoretical development entirely new in – because it goes against the basic theoretical commitments of – PS-based theory. Notice, too, how much depends on a notational convention, viz., that {F} should have values for which “higher than” can be defined. This is not a necessary consequence of a nonbinarity; imagine that the three values in question were notated by #, @, and \$, to pick three symbols pretty much at random from the keyboard in front of me. We could then stipulate an ordering among them, of course, but we would do so only in order to be able to state the definitions as in (18) and (19). I think we are entitled to serious suspicion when apparently deep linguistic properties hinge on the vagaries of notation. As Gauss said, “*non notationes, sed notiones.*”

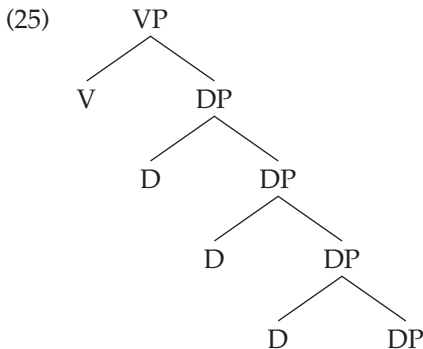
- (18) *x* is the *perfect head* of *y*, and *y* is a *perfect projection* of *x* iff
- (a) *y* dominates *x*
  - (b) *y* and *x* share all categorial features
  - (c) all nodes intervening between *x* and *y* share all categorial features
  - (d) the F value of *y* is the same as the F value of *x*
- (19) *x* is the *extended head* of *y*, and *y* is an *extended projection* of *x* iff
- (a) *y* dominates *x*
  - (b) *y* and *x* share all categorial features
  - (c) all nodes intervening between *x* and *y* share all categorial features
  - (d) If *x* and *y* are not in the same perfect projection, the F value of *y* is higher than the F value of *x*
- (20) *Generalized Theta Criterion*
- Every maximal projection must either
    - a. receive a role or
    - b. be part of an extended projection that receives a role.

And, in fact, it is not even clear that the apparatus Grimshaw sets up does the work she claims it does – or, at least, not in the way she claims. Consider again the set of “impossible combinations” (23), repeated here. In particular, consider the last row: C-CP, P-PP, I-IP, D-DP.

- (23) I-NP, I-DP, I-PP, I-CP  
 D-VP, D-IP, D-CP, D-PP  
 C-NP, C-DP, C-NP, C-VP  
 P-VP, P-IP, P-CP, P-NP  
 C-CP, P-PP, I-IP, D-DP

Grimshaw’s claim and point, recall, is that “[i]t is only combinations of L-heads and complements that are not governed by projection” (1991: 40). Or, put the other way round, the combinatoric (im)possibilities of F-heads are governed by projection. Grimshaw points out (1991: 6) that such configurations as in the last line of (23) are ruled out by (19d). True, but they are all allowed in by (18d). While they cannot form extended projections, they are (perfectly) in accord with the conditions on perfect projections. Notice, further, that what would rule out such combinations in traditional Functional-Head Theory, viz., a stipulation on the “selectional” properties of the Functional Head in question – is not possible for Grimshaw, given that she has shown that selection plays no role with respect to Functional Heads and their complements.

Now, as it happens, these perfect projections do run afoul of (20), the GTC. They do so because in a structure such as (25), in which all the DPs form a perfect projection, none of the lower ones are in accord with either clause of the GTC. They are not assigned a role *simpliciter*, clause (20a), because, unlike the topmost DP, they are not complements of a Lexical Head. And because they form a perfect projection, not an extended projection, they cannot fall under clause (20b). But now, contrary to Grimshaw’s stated claim, it is not just the theory of projection which accounts for the combinatoric (im)possibilities of Functional Heads and their complements. While evidently not in itself any sort of theoretical catastrophe, it does indicate a theoretical weakness in exactly the place where the claim is for most theoretical strength.



To sum up: I have pointed to two sorts of problems with the {F} feature. One is that it seems to advance a merely notational approach to a presumably deep linguistic issue. Second, the major theoretical advantage it claims to afford, *viz.*, directly and entirely accounting for the combinatorics of Functional Heads and their complements, is not exactly true (though, perhaps, not exactly false, either). I now claim that such technical problems are exactly what we should expect to find in an approach that is in fact conceptually unsound, as I have already argued Grimshaw's is. Foundational problems ought to – presumably must – give rise to technical problems at some point. If there were not the sort of unsoundness I have demonstrated, then we could count the technical problems merely as defects in the particular analysis or analyst. Now, however, there would be a terribly heavy burden of argument on one who would take such a position.

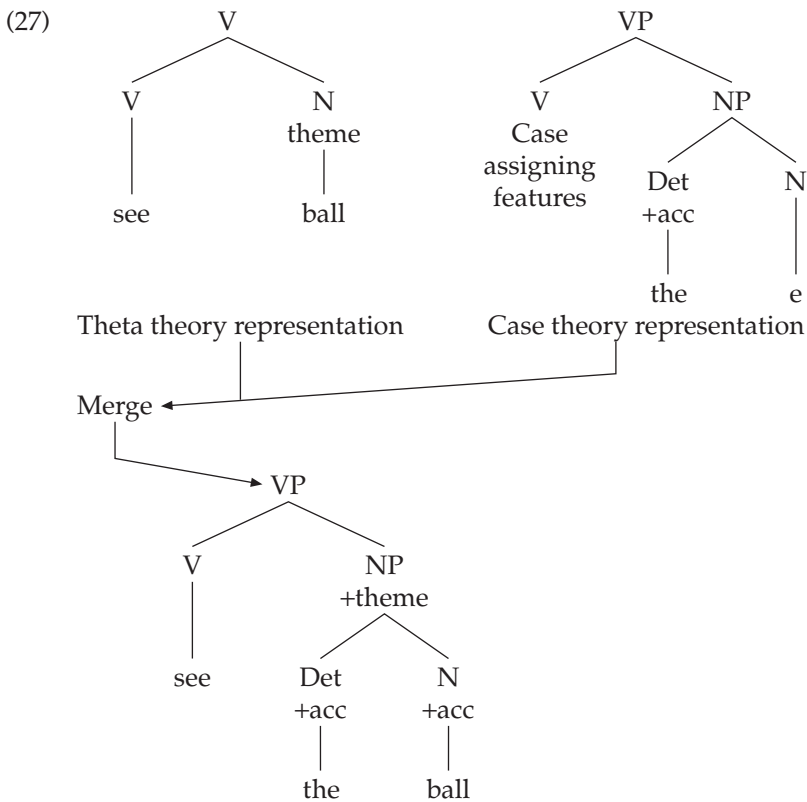
But now where are *we*? Grimshaw has shown us that the usual Functional-Head Theory is unsatisfactory, even empty, in crucial areas. Her own alternative is unsound conceptually and technically flawed. As these approaches – one based in selection, the other in projection – would seem to partition the ways in which to carry out the generalization of structuralization from Lexical to Functional Categories, the prospects for a theoretically contentful “generalization” do not look promising. This should not really be so surprising, given Grimshaw's argument that selection has essentially nothing to say about the role of Functional Heads in syntax and her observations concerning the “stability” of the relations Functional Heads do have in syntax. The central issue with respect to structuralization and PS, recall, is precisely to tame the (selectional) idiosyncrasies of individual lexical items. But if Functional Heads have no such idiosyncrasies, then there just is nothing for the generalization to do. Hence the theoretical emptiness of the traditional selectional approach and the conceptual unsoundness of Grimshaw's projectionist alternative. I think the work of Lebeaux (1988), to which we now turn, can offer us a way out of this impasse, though the way out does, in fact, require giving up rather a lot of what is generally taken as established by the “generalization” within P & P approaches to PS. What we will not have to give up, however, is *some* way of integrating Functional Categories into a PS-based approach to syntax. Instead of a spurious “generalization”, Lebeaux's work offers us an alternative that respects, indeed takes as fundamental, the distinction between Lexical and Functional Categories, and constructs a theoretical architecture for structuralization that is explicitly based in and builds on this distinction.

### 3.5 Lebeaux (1988): Closed Class items and a theoretical opening

Lebeaux asks us to take language acquisition facts into serious account when hypothesizing a syntactic architecture. In particular, Lebeaux (1988: 65, his (37)) suggests that syntactic theorizing be constrained by what he calls the *General Congruence Principle* (GCP), given in (26). Crucial from our perspective is the well-known acquisitional stage of “telegraphic speech” (see Lebeaux 1988: chapter 1 and 225f.) in which Closed Class items are absent.<sup>59</sup>

- (26) General Congruence Principle:  
Levels of grammatical representation correspond to (the output of) acquisitional stages.

Lebeaux argues for a radical reorganization of the syntax.<sup>60</sup> He suggests separate representations for Theta and Case relations, where these comprise, respectively, Open Class elements and Closed Class elements. These two representations undergo an operation called Merge, as in (27) (Lebeaux 1988: 243, his (43)). Each representation is “a pure representation of the particular vocabulary it invokes (Case theory vs. theta theory), and indeed, the crucial categories that are mentioned in each theory (determiner vs. nominal head) are distinct as well” (Lebeaux 1988: 242).



Lebeaux characterizes the two representations as follows. The Theta representation has as its vocabulary “theta roles . . . and category labels of the 0 bar level.” And it includes Open Class items instantiating the terminal nodes (Lebeaux 1988: 243), as these are the relata in Theta relations. The Case representation “factors out the closed class aspect of the V-NP representation in a principled way. What it contains is the following: (1) a subtree in the phrasal syntax (it projects up to at least V’), (2) where the case assigning features of the

verb are present, but not the verb, and (3) in which Case has been assigned to the determiner" (Lebeaux 1988: 244).<sup>61</sup>

The Merge operation is also characterized by Lebeaux (1988: 244):

First, it inserts two lexical items into the slots provided by the case frame: the head verb and the theta governed noun. Second, it percolates the theta relation already assigned to the noun to the NP node (theme, in this case). Third, it copies the Case that was originally associated with the determiner position onto the head noun. This means that *ball*, as well as *the*, is marked for (abstract) accusative case.

Finally, he notes that "the fixed character of the closed class elements is modelled by having such elements be the frame into which the theta representation is projected" (Lebeaux 1988: 245).<sup>62</sup> As he puts it at the start (1988: 1) "it is the need for CC [closed class–RC] elements to be satisfied which gives rise to phrase marker composition from more primitive units." This is a crucial observation from our perspective. We need to know what this "satisfaction" for Closed Class items is supposed to come down to.

In fact, Lebeaux has put things somewhat badly here. It is not, in fact, that Closed Class items must be satisfied; it is, rather, that Open Class items must be syntactically licensed. We require some further analysis of the general situation. Any item that receives a semantic interpretation must, presumably, be both semantically and syntactically licensed. Those are the respective functions of the theories of Theta and Case; the former is semantic licensing, the latter is syntactic. Roles and Cases must be assigned and borne, or else some element will be unlicensed. Now, Lebeaux's insight is to see that the theories of Case and Theta are rather more separate than has usually been supposed; and this is to say that syntactic and semantic licensing are more separate than generally supposed. The two kinds of licensing correspond to the two kinds of elements, Closed and Open Class. Open Class elements license semantically, Closed Class elements license syntactically, giving rise to two distinct representations. In order for all meaningful elements to be both syntactically and semantically licensed, the two representations must merge and the distinct licensings must "spread" within the newly composed structure, as Lebeaux (1988: 240f.) proposes. So, it is not quite right to say, as I did above, that what is required is for Open Class elements to be syntactically licensed. While true, this misses the point, which may have been what Lebeaux was getting at, that Closed Class elements are also in need of further licensing. If NPs require Case, and Case is assigned to D Heads, and Ns bear Theta roles, and Ds are Theta binders (as Speas 1990 suggests), then in order for both the Open and the Closed Class item to be both syntactically and semantically licensed, the merger must bring them together.

But there remains a difference between the two sorts of elements and their representations. Closed Class items, to repeat, create the "frame into which the theta representation is projected" (Lebeaux 1988: 245). Theta structures have just enough of the vocabulary of syntax to undergo merger. And theta structures are themselves usable, and used – by adults as well as children – without merger; that, after all, is what telegraphic speech is.



We can now see just what is wrong with the Functional-Head Theory/“generalization” tradition. It takes Functional Elements to be part of the problem, when in fact they are a part of the solution. The problem, recall, is the systematic structuralization of the idiosyncratic (selectional) properties of Lexical (Open Class) elements. Having made headway on this issue, the tradition extends – generalizes – its findings to Functional elements, despite the fact that there is no corresponding problem. This has allowed for an ever increasing inventory of Functional Heads and their projections, with apparently no theory guiding or constraining these “discoveries”.

Lebeaux, on the other hand, argues that the Functional versus Lexical elements distinction is a fundamental one, not just in the lexicon, but in the architecture of the syntax as well. Functional elements and Lexical elements give rise to distinct representations incarnating distinct licensing conditions, which representations ultimately are merged into a single object. Notice that, in a sense, this simply extends the common “lexical-entry driven” approach to structuralization while respecting the different kinds of elements in the lexicon. Moreover, and this is crucial, Lebeaux’s approach suggests that there are principles and constraints on the inventory of Closed Class elements and their structuralizations. The idea here – not one that Lebeaux himself advances – is the following. Functional elements and their structuralizations provide the syntactic licensing which Lexical elements require – recall the example of the Case marking on D spreading to the N(P) which requires Case for syntactic licensing. This means that analysts should posit Functional elements and their structuralizations only if it can be plausibly argued that there is an independent syntactic licensing relation which can both legitimate that Functional element and spread to a Lexical element that requires such licensing.<sup>63</sup>

We find ourselves, then, not with a new set of answers and analyses, but rather with a new way of looking for answers and analyses. There remain obscurities of both conception and detail in Lebeaux’s work. It is not at all clear which findings and analyses from the tradition would carry over, nor in what form. But the advantages are real enough, I think. First, different theoretical kinds (in the lexicon) are kept distinct, instead of being simply and unthinkingly assimilated. Second, this distinction forms the basis for a network of distinctions that are otherwise thought unrelated (e.g., Case vs. Theta representations in syntax, stages in acquisition), creating a theoretical architecture that is much more integrated and deductively complex than previously. Finally, the beginnings of a contentful theory of Functional Heads and their structuralizations is adumbrated, so that Chomsky’s (1986a: 3) original question of whether the system for structuralizing lexical categories “extend[s] to the nonlexical categories as well. . . .” no longer takes the null hypothesis “yes” as its answer.

I have taken rather more space, and been rather more argumentative, in this subsection than in previous ones. I think it has been warranted, in that the issue is an important one, and the discussion would not find a natural home in the later chapters. The positions I have argued for might be controversial; in any event, they are certainly minority views. In the next subsection, I return to the less tendentious mode.

## 4.0 X-Bar Theory

---

We consider two issues here. The first is the status of rules (or rule schema) in the grammar. The second is branching: is it always and only binary?

### 4.1 Rules

Stowell (1981) initiated the program of “X-Bar reduction” that reached its zenith in Speas (1990). A crucial enabling observation for the program is the following. Once an independent lexicon was introduced by Chomsky (1965), the grammar contained a massive redundancy, in that essentially the same information was specified in both the phrase structure rules and subcategorization frames. Eliminating the redundancy could evidently proceed in one of two ways: get rid of the phrase structure rules, or get rid of the subcategorization frames. The Stowellian project pursues the former line.<sup>64</sup> Despite some criticism, notably by Pullum (1985) and Kornai and Pullum (1990), this is the standard view in P & P work.<sup>65</sup> Speas (1990) is the fullest working out of this line of inquiry, and contains much valuable review and criticism of other work. Despite this widespread agreement, there is still occasional equivocation about the status of rule schema, if not rules themselves. It is sometimes the case that rule schema such as those in (28) are given and it is claimed that “categories are projections . . . in terms of . . .” these schema (Chomsky 1986a: 2–3). However, the status of these schema is generally not explicitly discussed, and concomitantly it is left obscure just how and why structuralization is “in terms of” them.<sup>66</sup> If there is a commitment to elimination of phrase structure rules, we should be told explicitly what other commitments afford this elimination; a desire for reduction in redundancy does not, by itself, count.

- (28) a.  $X' = X X''^*$   
 b.  $X'' = X''^* X'$

### 4.2 Binary branching

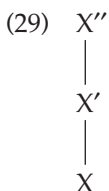
There are two issues here, though typically only one gets discussed. One, the usual topic, is whether there is more than binary branching in a phrase marker (i.e., can any mother have more than two daughters?). The other is whether there is less than binary branching (i.e., can any mother have fewer than two daughters?). Most discussion of the first question traces itself back to Kayne (1981). Kayne there proposed an “unambiguous path” condition on various syntactic relations, and for this condition to hold, branching would have to be, by and large, at most binary.<sup>67</sup> Unlike many who purport to follow him in this line, Kayne does not stipulate that branching is at most binary. Instead, binary branching is required if there is to be an unambiguous path mediating some other syntactic relation, so it is an effect, rather than a cause.

There is relatively little actual attempt to motivate at most binary branching in the literature. Such argument as there is generally amounts to observing that

such a restriction would restrict options analytically (or acquisitionally), hence is desirable, and therefore it will be assumed. Often virtually no consideration is given to the fact that there are empirical reasons to suppose branching may be more than binary,<sup>68</sup> though Williams (1994: 30–1) notices that such facts result in no more than binary branching having effects on the “locality” requirement on the “argument-of” relation, viz., that if a predicate has more than one argument, then not all the arguments can be sisters of that predicate. Williams seems to prefer giving up no more than binary branching to weakening the locality condition, as he finds the latter independently justified.

Kayne (1994) advocates no more than binary branching, once again, it seems, requiring this for satisfaction of some other syntactic relation (here his central Linear Correspondence Axiom rather than his earlier unambiguous path concept) rather than stipulating it outright for its own sake. Chomsky (1995: chapter 4) appears to take the position that no more than binary branching is required by “virtual conceptual necessity” (1995: 249). Chametzky (1996: 35–6, 112) allows that no more than binary branching might be justified as an empirical generalization and that it provides help in the analysis of Adjuncts. This issue will loom large in our Conclusion.

We turn now to the other issue, whether branching must be at least binary; that is, is there nonbranching domination? Often those who proscribe more than binary branching seem to allow for less than binary branching. The matter itself is examined in detail in Richardson and Chametzky (1985: 337–40) and resumed in Chametzky (1996: 27–8). They argue that no well-behaved PS theory ought to have such a relation.<sup>69</sup> There are two parts to their argument, a conceptual and an analytic. The conceptual portion is simply to observe that constituency is a part-whole relation, and to claim that a whole with one part is in the same relation to that part as a whole with two (or more) parts is to its parts is to make a nonobvious, quite plausibly spurious, claim. That the whole in the latter case is distinct from the parts seems clear, perhaps necessary; that this is so in the former case is not at all clear, though possible. The analytic portion is an examination of the actual range of cases of nonbranching domination in the literature. There are some four types examined; two have some antecedent plausibility.<sup>70</sup> One is the relation between lexical items and the “zero level” categories in a phrase marker. This, however, is not the part-whole relation of constituency, so the use of dominance is inappropriate.<sup>71</sup> The final case is that as in (29), nonbranching within a single Head’s projection. The alternative suggested is multiple labelling of a single node; both formal and substantive reasons are advanced for supposing this is possible, even desirable.<sup>72</sup> The general conclusion is the following. If nonbranching domination is conceptually unsound, then there ought to be no clear and compelling instances of it – and there are not.



We move now to our final question, the statuses of precedence and dominance.

## 5.0 How to order a phrase marker

---

The issue here is whether phrase markers are specified as formal objects with two ordering relations or only one. No one suggests that the *dominance* ordering relation does not formally specify a phrase marker (though there is some disagreement with respect to whether this relation is reflexive or not<sup>73</sup>), so the question comes down to whether a *precedence* ordering relation also does. The issue really only arises once PSRs are given up; PSRs specify both immediate dominance relations among mother and daughters and linear precedence relations among daughters.<sup>74</sup> Within the P & P tradition, it is typically assumed that these two dimensions are separable and separate.<sup>75</sup> Following the lead of Stowell (1981), it is also typically assumed that precedence relations are the result of the interactions of various substantive subparts of the theory (see Speas 1990: 19–24 for discussion); this provides one type of “precedence rejection”. A different type is argued for by Chametzky (1995, 1996: 6–14) where a novel argument against precedence as a basic ordering relation specifying phrase markers as formal objects is given.<sup>76</sup> Notice that the first, more common, type of “rejection of precedence” does not require this second type, while this second type still allows for the first. Under the first type, precedence relations are derived in substantive syntax, rather than specified in some sort of rule. But this is nonetheless consistent with a formal requirement that a precedence ordering relation be part of the specification of phrase markers. Under the second sort, there can be no such formal requirement, but it can still be the case that substantive syntactic considerations result in a precedence ordering.

Williams (1994: 178–98) provides a spirited, and unusual, defense of precedence as “a fundamental syntactic notion” (178). Most of the argument is with Larson’s (1988) development of “shells” and use of verb movement to solve problems in binding theory, given the assumption that binding relations are mediated entirely by hierarchical (command) relations, and not at all by precedence. Williams’s brief is that over a range of data types and analyses, Larson’s approach is empirically worse than alternatives that do use precedence. Williams also examines two other areas, the Nested Dependency Condition and Weak Crossover, with similar results. However, it is not actually clear even if Williams is correct that therefore either of the “rejectionist” stances outlined above need be given up. If independent substantive syntactic considerations result in a precedence ordering, then it may well be open to the “rejectionist” to refer to this ordering in accounting for the empirical phenomena Williams discusses. What such a “rejection of precedence” would actually then amount to might be less than entirely clear; indeed, it might be empty or pointless.

Kayne (1994) makes much of precedence facts. Indeed, that entire book is an attempt to make something theoretically deep and revealing out of precedence facts. Interestingly, however, I do not think that Kayne’s theory requires taking

precedence to be a formal ordering relation that is part of the specification of phrase markers. Chomsky (1995: 335–40) discusses Kayne (1994) and explicitly relegates precedence relations to the nonsyntactic portion of the grammar (viz., the “phonological component”).

This concludes our preliminary investigations. We turn now to the particular cases, the works through which we inquire into the fate of our concepts and questions, and of phrase structure more generally. Examination of the PS concepts and questions with which we are now familiar will structure much of our inquiry.

## Notes

- 1 See Chametzky (1996: xvii–xviii) on *theoretical* versus *analytic* work.
- 2 See p. 2.
- 3 This is often called “projection” in the literature. I use the less elegant, but more transparent, “structuralization” in Chametzky (1996).
- 4 Borer rejects the lexical-entry-driven-approach and adumbrates an alternative driven by “the part of the subpredicate dealing with the verb in combination with aspectual projections” (1994: 45).
- 5 There is an extensive literature on the form and content of lexical information and its relation to syntactic structure. Besides Speas and Pesetsky, within P & P beginnings can be made with: Grimshaw (1990) and Levin and Hovav (1995).
- 6 We shall examine details in Chapter Two.
- 7 Compare these with Perlmutter and Postal’s (1984) Universal Alignment Hypothesis (UAH) in Relational Grammar (RG). Rosen (1984) is an important discussion of argument alignment couched within RG assumptions. Chapters 1 and 2 of Pesetsky (1995) extensively review problems and prospects for what he calls U(T)AH, combining both UAH and UTAH. It should be noted that Bouchard (1995) elaborates an alternative general approach to these issues (and to syntax, semantics, and their relation) that he argues makes such statements as U(T)AH and RTAH – and explicit statements concerning PS – unnecessary.
- 8 Williams (1994: 159–61) has insightful remarks on argument ordering, including what he calls *t-normal order*: “ordering of elements according to theta roles . . . simply one of the factors that determines order, and a weak factor at that” (1994: 160).
- 9 This distinction remains absolutely central to the rather different theory developed in Williams (1994).
- 10 Speas (1990) accepts that there may be a “most prominent” argument lexically, but does not accept that it is structuralized outside a Maximal Projection of the lexical item of which it is an argument.
- 11 See note 15 below.
- 12 See, e.g., Speas (1990: 17–18, 102, and the works cited in her footnote 12, p. 25).
- 13 Whether Subjects which so originate also move to another position is a separate question, one that will not much concern us.
- 14 Chametzky (1996) is one example.
- 15 Well, almost. You also need a structural relation to mediate the Subject–Predicate relation; typically, some command relation (Barker and Pullum 1990; Chametzky 1996: ch. 2) is invoked, where the Subject must X-command the Predicate. See Chametzky (1987: 58–65) for some discussion of alternatives suggested by different versions of Predication Theory.

- 16 Chametzky (1996: 130–4) makes exactly this proposal.
- 17 Williams (1994: 141–51) develops an approach based in Predication Theory that neither obviously accepts nor denies the LCH.
- 18 Here and below I follow the discussion in Chametzky (1996: ch. 4).
- 19 Recall that “Adjunct” names a dependency relationship. It is not an essentially PS relationship.
- 20 A terminological note: I call the XP-daughter – the adjoined-to element – the “host”. I here ignore nonmaximal hosts, viz., “adjunction to a Head”. See below, however.
- 21 As we shall see in Chapter Four, however, Chomsky (1995: 244–5) does not.
- 22 We return to this issue in Chapters Four and Five.
- 23 It should be noted that May never discusses Adjuncts, hence never distinguishes them from adjunctions.
- 24 I have gone over this ground before, in Chametzky (1994, 1996: 89–106), and do so again in Chapter Two, section 2.3, so I will refrain from further comment here.
- 25 As pointed out in Chametzky (1996: 1) *structure* and *category* are the two basic **concepts** of the theory of syntax. *Immediate constituent node* and *word class label* are the **conceptions** given these concepts in virtually all P & P work. See, e.g., Rawls (1971: 5–6) or Dworkin (1977: 134–6) on the *concept* versus *conception* distinction.
- 26 Whether these are the result of similar, or any, principles, is not discussed, however.
- 27 Though this is somewhat unclear, as May does not explicitly discuss labelling and he may be assuming already labelled nodes as his primitives.
- 28 See Chapter Two.
- 29 See Chapter Four. Bouchard (1995: 85–6) adopts what was to become the position found in Chomsky (1995).
- 30 We note that Kayne (1994) relies on adjunctions to a very great degree; this is discussed in Chapter Three.
- 31 Bouchard (1995: 109–13) rejects all movement transformations, including adjunctions.
- 32 A more syntactic way to approach the same fact is to say Adjuncts do not subcategorize a Head.
- 33 Kornai and Pullum (1990: 44) in their formal (and somewhat hostile) analysis of “X-Bar grammars” notice this property as well.
- 34 We might also suggest that the approach appears to make the very existence of Adjuncts something of a surprise. That the existence of Adjuncts might pose a theoretical issue in need of explanation has not been explicitly raised before, to my knowledge. I think the theory in Chametzky (1996) can shed some light on this, as discussed in Chapter Two. We shall explore these problems and seeming conundra in some depth, particularly in section 2.3 of Chapter Two and section 3.3 of Chapter Four.
- 35 Or sister-adjunction – these are from the Standard Theory (see, for example, Bach 1974: 86–7).
- 36 See sections 2.1 and 3.0 of the Introduction.
- 37 See the references in note 7 of the Introduction. Croft (1996) argues for a “semantic definition of headhood” (1996: 69). I am not exactly sure what Croft takes the significance of his argument to be; we can note, however, that inability to *define* “Head” syntactically might be what one would expect were it a syntactic **primitive**. This is hardly a novel idea; it is basically the conclusion reached by Johnson (1977: 690) with respect to the failed attempt by Edward Keenan to define “Subject” in syntax, for example:

the best interpretation of K[Keenan]'s results is that they provide a *reductio ad absurdum* argument in favor of the RG view that grammatical relations such as SUBJ be taken as primitive, theoretical terms. As such, they are neither defined in terms of, nor directly connected to, observables or antecedently understood concepts. As uninterpreted, theoretical terms, however, they are indirectly connected to observables and/or antecedently understood concepts via the empirically testable predictions resulting from the statements in which they occur.

- 38 See section 4 below.
- 39 This is alluded to in Chametzky (1996: 182, fn. 16).
- 40 See also Emonds (1987) on his Empty Head Constraint.
- 41 As, for example, Bouchard's Principle of Full Interpretation (10), or Speas's statement, quoted above, concerning interpretation "at both interface levels."
- 42 In the terminology of the structuralist tradition, these might be seen as information concerning the "paradigmatic" and the "syntagmatic" dimensions.
- 43 Prepositions, famously, seem to cross-cut the classifications, being both "ClosedClass" and "Lexical". See, for example, the discussion in Emonds (1985: chs. 4, 6, 7) and Grimshaw (1991).
- 44 It should be noted that Jackendoff's (1977) "Uniform Three Levels Hypothesis" was not limited to Lexical Categories, but also included Article, Degree, and Quantifier Phrases.
- 45 See Kitcher (1983: 213–17, 229–41, 268–70) for insightful discussion of these issues in the context of "rigorization" in mathematics, with special reference to the difference in the development of calculus in Britain and on the European continent after its discovery by Newton and by Leibniz. A belief in the need to get a "legitimate inquiry license" may go along with the belief in the need to antecedently "define your terms" alluded to on p. 8.
- 46 See Chametzky (1987a) for discussion.
- 47 As Fukui and Speas (1986: 4) remark, *functional categories* is a much better term than *nonlexical categories* because on this view these categories "are projected from the lexicon and have independent lexical entries." Abney (1987: 54) makes a similar observation, using the term *functional element*.
- 48 Spencer, it should be noted, is critical of this position. Other critical discussion of "generalization" work can be found in Iatridou (1990), Ernst (1992), Janda and Kathman (1992), Joseph and Smirniotopoulos (1992), and Janda (1993).
- 49 Bouchard (1995: 255; 457, fn. 5; 479, fn.1) almost does, but in fact disallows only "contentless functional categories".
- 50 Abney uses the terms *Functional Element* and *Thematic Element* in his discussion.
- 51 Grimshaw (1991: 6) justifies the {F2} analysis of P and C as follows. "They do not act like the lexical categories, because they do not occur as complements of functional categories. They do not act like the functional categories because they do not take lexical complements. . . . P stands in the same relationship to DP and NP as C does to IP and VP."
- 52 There are some analytic details to consider here, as the generalization just stated is recognized by Grimshaw to be not obviously true. We shall not pursue the apparent counter examples or Grimshaw's attempts to explain or deny them.
- 53 Grimshaw (1991: 14) notes the similarity, and points out that the problems that require these moves are "more general – involving not just semantic relations but syntactic ones. . . . Hence, the solution must also be more general."

- 54 Though Grimshaw (1991: 8) writes, "... the theory of extended projection in no way rests on reference to the {F} feature; the same results will always be obtained even if the work of {F} is taken over by other parts of the theory." This is true enough, but, in the absence of anything remotely like a suggestion as to what such "other parts of the theory" might be, only abstractly and uninterestingly so.
- 55 As noted in Chametzky (1996: 20–1), these authors converge in some striking ways in their analyses of what is baby and what is bath water with respect to X-Bar theory, despite their rather different starting points and rather different goals.
- 56 One can understand the differences in bar-levels as differences in subcategories of the categories determined by the "categorical" features, as subcategory distinctions are needed anyway.
- 57 Obviously, whether or not these definitions are current or correct is not the point or issue. It is the necessity of some such definition that is crucial.
- 58 Recall here the discussion in section 2.1 of the Introduction with respect to the general issue of encoding dependency relations in PS syntax, of which this is evidently another example.
- 59 Lebeaux uses the terms *Open Class* and *Closed Class items* for what we have hitherto been referring to as Lexical and Functional elements. I shall follow his terminology in discussing his work.
- 60 Lebeaux argues not just from acquisitional facts and assumptions, but also from "pure" syntactic ones, which he does "not really differentiate between" (Lebeaux 1988: 6).
- 61 Lebeaux (1988: 16f) argues that the set of Case features on verbs is Closed Class, while the set of verbs is Open Class.
- 62 Lebeaux uses *Project-a* as another name for his Merge operation.
- 63 There must also be semantic licensing of the Functional element if it is to receive a semantic interpretation, of course.
- 64 Gazdar and Pullum (1981), within Generalized Phrase Structure Grammar, observe the same redundancy and pursue the alternative course, eliminating the subcategorization frames.
- 65 Speas (1990: 35–8, 56–60) discusses Pullum (1985); Chametzky (1996: 20–2) comments on this discussion and on Kornai and Pullum (1990).
- 66 See Chametzky (1996: 153–6) for discussion of this problem with respect to Chomsky (1993).
- 67 Chametzky (1996: 32–6) criticizes the "unambiguous path" concept.
- 68 See, e.g., Carrier and Randall (1992); Pesetsky (1995: chapter 7).
- 69 The theory of Lasnik and Kupin (1977), in fact, does not. Or, rather, in this framework nonbranching domination is a symmetric relation, so it effectively disappears.
- 70 The other two cases are exocentric labelling and using "functional" labels such as "subject" or "topic".
- 71 See Chametzky (1996: 5, and references cited there).
- 72 See Richardson and Chametzky (1985: 339–40); Chametzky (1996: 14–15, 27–8); Chapter Two in this volume, section 5.2.
- 73 Chomsky (1995: 338) takes dominance to be irreflexive. Chametzky (1996), following the normal practice in formal grammar theory, takes it to be reflexive. See Chapter Four and Chapter Two, respectively.
- 74 More accurately, PSRs are generally interpreted as specifying such relations among mothers and daughters; the rules can also be interpreted simply as string-to-string rewrite instructions, which can be further related to mother and daughter



relations. These matters, while of intrinsic interest, are orthogonal to our concerns (see McCawley 1968; Partee, ter Meulen, and Wall 1990).

75 See Gazdar and Pullum (1981) for elaboration of a rule formalism that separates these two within Generalized Phrase Structure Grammar: what they call ID/LP (immediate dominance/linear precedence) format.

76 We return to this in Chapter Two.