

NATURAL LANGUAGE SYNTAX: TRANSPARENCY AND BINARY BRANCHING

J. L. G. Escribano
Universidad de Oviedo

1. Introduction

This paper is an attempt to develop a parsimonious, psychologically realistic and computationally efficient model of syntactic-semantic 'tactics' in natural language from a 'generative' perspective. By 'generative', I mean 'explicit' and 'formal' in the sense in which such terms were originally used in Chomsky (1955). The overall inspiration, thus, is ultimately Chomskian, and the modular-interactive design of the grammar is directly inspired by Chomsky's current 'principles and parameters' model, but ours is mono-stratal, its default interpretation is strictly declarative (although a straightforward procedural interpretation is suggested), and a rather unorthodox approach has been taken in certain issues of principle and in many technical details. In fact, although several parts of the current 'principles and parameters' model (i.e., Case Theory, Theta Theory, Binding Theory and Move Alpha) are assumed in more or less the form they take in such publications as Chomsky 1981, 1986, 1988, 1989 and related research, the actual technical implementation chosen rather takes inspiration from Categorical Grammar, Functional Grammar, GPSG, HPSG and other current unification-based approaches.

2. Background Assumptions

The broad guiding principles underlying this research should sound familiar enough in TL, CL and AI circles and can be stated as follows:

1) At least in our current state of ignorance, such objectives as naturalness and explanatory adequacy call for 'tactic' principles of maximum simplicity and generality, since only such devices can at present be reasonably attributed to UG, if such an entity exists, or, for that matter, to its ultimately neural hardware implementation in the human brain.

2) Given present technology, computational efficiency in information-processing systems also implies the above strategic pre-requisite of simplicity and calls for a modular, interactive, approach.

3) Psycholinguistic poverty of the stimulus considerations, and the sheer diversity of existing, or, indeed, possible, languages reinforce the hypotheses in 1) and 2) above and further suggest a parameter-setting model of language acquisition and a modular parameterized syntax.

4) Given the richness and sophistication of the information activated in natural language processing, assumptions 1-3 above naturally lead to a strategy of minimizing complexity in the algorithmic-computational mechanisms of syntax-semantics and concentrating it where, as we shall argue, it naturally belongs (and where it admittedly cannot be avoided anyway!) i.e., at a central conceptual data-base interfacing language and other cognitive systems which, in its linguistic side, is a structured Lexicon.

According to the preceding line of reasoning, the model we shall sketch in this paper

will assume a rich data-base or repertory component and a maximally impoverished, in fact trivial, tactic one. The Lexicon, in other words, is quite complex, as is the associated pragmatic knowledge involved, but, in exchange, the syntactic and semantic rules that account for the compositionality and creativeness of NL are trivially simple, and it is my contention that precisely that distribution of labour is exactly correct in all the relevant theoretical-linguistic, psycholinguistic and computation-theoretic respects.

Needless to say, space limitations preclude a full discussion and justification of the empirical and theoretical underpinnings of the strategy we shall pursue here (cf. Escribano 1988b, 1989b, c, 1991b for analyses of English and Spanish syntax consistent with present assumptions), but a few summary remarks may be in order.

The central issue, of course, is whether our decision to concentrate complexity in the Lexicon is well founded, and as a matter of fact I believe that there are solid linguistic, psycholinguistic and computation-theoretic grounds for it. As to the first, for example, it is no accident that all linguists should so readily accept the existence in their grammars of a vast and highly complex lexical component: the world we want to talk about happens to have become quite complex and stipulative, or so we perceive it, and in order to make the fine-grained distinctions we need, we must depend on highly distinctive (= idiosyncratic) signs for them.

The lexical units of a language, as Saussure, Hjelmslev, Martinet and many others before and after them have emphasized, are essentially contingent, arbitrary, and historical products. No wonder, then, that lexical properties are largely language-specific and unpredictable by rules or principles such as those usually attributed to UG, but that implies that linguistic competence and our descriptions of it are bound to contain a vast deposit of lexical information.

If we now turn to the acquisition and processing side, the psycholinguistic and computation-theoretic evidence in support of this approach is no less compelling. Indeed, as to the former, there is little doubt that the acquisition of lexical competence, although far from a trivial accomplishment given the complexity of the knowledge eventually attained, depends on continued exposure to a wealth of linguistic information from various sources. It is a gradual, piecemeal, and relatively conscious, or even painful, process, in which ostension, massive contextual feedback, functional pressure and, in the later stages, straightforward instruction play a part to varying degrees. All that, along with obvious sources of motivation to learn on the part of the child, makes the long-term acquisition of lexical competence at least explicable in principle.

However, as Chomsky has repeatedly emphasized, those circumstances are totally absent in the acquisition of the tactic aspects of competence, and as a consequence the postulated architectural differences between the repertory and the algorithmic component seem strongly supported by extant psycholinguistic evidence. The tactic mechanisms of syntax and semantics are fixed once and for ever during a critical period, without conscious effort, and, apparently, under rather adverse conditions, since the relevant evidence is largely inaccessible to the child due to the fragmentary and corrupt nature of the data and to the intrinsic abstractness of the operations involved. That is the argument from poverty-of-the-stimulus in its classical form and has led Chomsky and his followers to attribute such devices directly to a genetically determined language-organ (LAD, as represented in UG).

It is fair to anticipate at this stage that one of the objectives of this paper is to explore a partial alternative to that conclusion, but, in the light of the complexities uncovered in NL in the last thirty years or so, it must be admitted that Chomsky's argument has its

force: the acquisition of syntactic competence is still largely a mystery to be explained. However, it must be pointed out that any simplification achieved in the algorithmic component reduces the force of the poverty-of-the-stimulus argument and proportionally weakens the case for a separate Language Faculty.

We shall initially assume with Chomsky that the linguistic mechanism is largely genetically determined and not learnt, but given the almost trivial generality and simplicity of the operations required in our account, it does not seem necessary to attribute them to a separate linguistic module. On the contrary, they rather look like primitive cognitive abilities such as are prominently encountered in many other cognitive domains. This is an unorthodox view in a GB framework but seems to me nothing but the logical dialectical conclusion of the restrictive research strategy introduced in the TG paradigm with the version known as EST around 1970, which is still under way in current GB: to the extent that consecutive versions of the theory have succeeded in discovering explanatory principles of growing simplicity and generality, the case for a neatly autonomous Language Faculty has become proportionally less compelling.

As to the computation-theoretic side, a similar argument can be constructed in support of the strategy complex lexicon cum trivial syntax. First, there is evidence that the algorithmic devices responsible for NLP are really quite efficient: we process language at great speed and with great accuracy. On the other hand, the currently known sources of efficiency in both natural and man-made information processing systems seem to be similar, and broadly depend on two general principles:

- 1) That memory (circuitry, computation space) is comparatively inexpensive and readily available in both computers and human brains. Thus, storage and retrieval of large amounts of pre-computed ready for use information is in general preferable to the *prima facie* more parsimonious and elegant expedient of storing only minimal units and assembling the rest on-line as required.

- 2) That, on the contrary, even though large data bases can be handled with relative ease by both computers and human brains, in both cases only extremely simple operations can be efficiently computed, and any increase in complexity in this respect tends to be very expensive in computational terms. As a consequence, efficiency in information processing systems is largely based on massive iteration of trivially simple tasks. Multiple, or even parallel, processing essentially require memory and computation space, which is cheap and readily available, as we said, but even a comparatively small upward shift in operational complexity is likely to require a totally different machine.

Thus, to the extent that the computer metaphor is relevant in the study of biological analogues responsible for natural language processing, we have good computation-theoretic reasons to concentrate complexity in the repertory component and keep the algorithmic one as simple as possible. Since, as we saw, the psycholinguistic evidence from acquisition and the sheer facts of linguistic history point in the same direction, I shall assume that strategy to be correct.

Consequently, it is my contention that, provided a suitably specified dictionary can be built, the syntactic and semantic rules of natural languages can be maximally reduced, to their virtual disappearance, and that such a conception of the algorithmic components yields better empirical predictions, is computationally more efficient and offers a more natural explanation of the facts of acquisition and use than current alternatives.

3. The 'tactic' component in current generative grammar

Although, following Heny 1979, Chomsky 1981 chapter 2, and Stowell 1981, rules like (1) no longer exist in GB grammars, and a great deal of lip-service has been paid to the theoretical convenience of dispensing with phrase structure rules altogether (cf. Speas 1990 for a recent statement to this effect), the fact remains that the usual organization of generative models of grammar still reflects the traditional division between the **syntax** and the **lexicon**.

Of course, there is no box labelled 'syntax' figuring as a module of GB, LFG or GPSG grammars, to mention only fully developed and clearly influential generative proposals currently under study, but context-free PS rules do appear under the guise of 'X-bar schemata', so-called 'node admissibility conditions' (GB), ID+LP rules (GPSG, HPSG), or even straightforward phrase structure rules of the conventional sort (LFG), and certain scholars have in fact cast doubt on the feasibility of Stowell's programme (cf. Pullum 1985 and Kornai & Pullum 1990).

$$(1) \quad V' \rightarrow V (N''') (P''') (V''')$$

Chomsky's and Stowell's main objection to such rules is that they are redundant, since the information they convey either must be stated in the Lexicon anyway or can be derived from general principles. Heny's criticism, thus, eventually led to the appearance of restrictive X-bar approaches, notably those in Stowell 1981, Stuurman 1985, Emonds 1985, Chomsky 1986, among others (cf. (2) below).

Of course, there is no consensus on such matters as a) the number of bar levels required, b) the maximum value of * in (2), (or the existence of n-ary branching, for that matter), c) the possibility of non-maximal non-heads, d) the inventory of lexical categories in the domain of X, or even e) the endocentricity of all phrases (Chomsky 1970, Emonds 1976 and Bresnan 1981 exclude the sentence), but we cannot offer a detailed account of such discrepancies here (cf. Escribano 1991a,b for discussion). Chomsky 1981, 1986, for example, has max=2, leaves * free (thus allowing for n-ary branching), and requires maximality of non-heads, and most GB linguists follow him. Stuurman 1985 has max = 1, though, and Jackendoff 1977, as is well known, raises it to 3, but more recently Speas 1990 (and Kornai & Pullum 1990) deny all relevance to this aspect (cf. Escribano 1991a,b, too, and *infra*, for a systematic approach in this spirit). As regards b), Kayne 1983 has argued for binary branching, and Speas 1990 has followed suit. On the other hand, Jackendoff 1977 and Emonds 1985 support thesis c), and there are wide discrepancies as to the domain of X (cf. Emonds 1987 for a standard account and see Abney 1987, Pollock 1989, Chomsky 1989, Ouhalla 1990 for recent additions in the area of so-called 'functional' categories and Escribano 1991a,b for relevant criticism of such approaches).

Stowell and Chomsky compress X-bar theory into a couple of schemata like (2) plus, perhaps, a general-purpose coordination schema like (3), where X is a variable ranging over the set [$\pm N$, $\pm V$] and '*' stands for 0 or more occurrences of the following category):

$$(2) \quad \begin{array}{l} \text{a. } X^{\max = n} \rightarrow * \text{Spec } X^{n-1} \\ \text{b. } X^1 \quad \quad \rightarrow X^0 * Y^{\max} \end{array}$$

- (2) $X^n \rightarrow *Y^{\max} X^n *Z^{\max}$
 (3) $X^* : X X X \dots$ (and/or) X

Stuurman 1985 was even more restrictive, since he operates within a two-tier X-bar frame with a single opposition between X' (= maximal projection) and X , a simplification that rests on a strategy of capturing distributional restrictions in terms of constraints on operator-variable linking. Stuurman's basic schema is (4), where ... stands for a sequence of zero or more maximal projections:

- (4) $X' \rightarrow \dots X/X' \dots$

Both approaches are on the right track, as widely recognized, but both are too permissive as to the strings allowed to accompany the head in such rules, and none of them develops the leading ideas to their only possible logical conclusion: the suppression of the X-bar component. This is also true of Speas 1990, who dismisses the rules, but preserves specific X-bar principles. Thus, the 'tactic' component of current GB models invariably includes first and foremost either a set of context-free schemata roughly of the form (2-3) (as in Chomsky 1986: 3) or an equivalent set of 'X-bar principles' or 'X-bar admissibility conditions' which can be spelt out as in (5):

- (5a). **Endocentricity**: all phrases are endocentric and are headed either by an X [Bar, 0] (i.e., a lexical category N, V, A, P , or a functional category like $Comp, Infl, Tns, Agr, Neg, Det$, etc.), or by a phrasal projection thereof (i.e., an X [Bar, n]) (Chomsky 1972, 1986, 1988, Stowell 1981, Abney 1987, Pollock 1989, Speas 1990, etc.).
 (5b). **Successivity**: X [Bar, n] dominates X [Bar, $n-1$] (Chomsky 1972, Jackendoff 1977, Stowell 1981, etc., cf. Kornai & Pullum 1990).
 (5c). **Maximality of non-heads**: complements, adjuncts and specifiers alike must be maximal projections (Jackendoff 1977, Stowell 1981, Chomsky 1986, etc.).
 (5d). **Uniformity of the value of X^{\max}** : $\max = n$ across categories (Chomsky 1972, 1986, Stowell 1981, etc.).
 (5e). **Peripherality of the head** (i.e., languages are either head-first or head-last; Chomsky 1972, 1981, Stowell 1981).
 (5f). **Parameterized alternating position of Specifiers and Arguments** with respect to the head (Chomsky 1972, Jackendoff 1977, and everybody else since).
 (5g). **Optionality of non-heads**: all nodes except H are structurally optional (Jackendoff 1977 and a standard view ever since).
 (5h). **Binary branching**: in certain versions (Kayne 1983, Speas 1990) only binary branching structures are licensed.

To either of those must be added substantial restrictions imposed by other 'modules', cf. (6), which, according to Kornai & Pullum 1990 and Speas 1990, in fact do most of the real work of constraining phrase structures:

- (6a). **Configurational definitions of grammatical functions** like Subject and Object (Chomsky 1972, 1981, 1986, Williams 1981, Speas 1990, etc.).
 (6b). **Case-Theory driven restrictions on linear order** of NP's, PP's, etc. (adjacency of the NP to its governor and case marker, cf. Stowell 1981).
 (6c). **Functionally inspired ordering restrictions**: the Clause-Last Principle (Stowell

1981, Radford 1988).

(6d). **Lexically or Theta-Theory-driven constraints on the projection process**, which determine that certain arguments be projected earlier than others according to a 'Thematic Hierarchy' and UTAH (= Uniformity of Theta Assignment Hypothesis): internal arguments like Theme are projected immediately on top of the lexical head, whereas the external argument Agent is attached directly under the maximal projection, etc.; cf. Jackendoff 1972, 1987, Williams 1981, Belletti & Rizzi 1987, Baker 1988, Speas 1990).

(6e). **Binding Theory-driven restrictions** on the distribution of antecedent-anaphor and operator-variable chains. The strategy of eliminating X-bar rules of the conventional PS sort is quite pervasive, therefore, but the X-bar principles remain as a separate component, and thus the distinction between lexicon and syntax remains as strong as ever.

In the current standard T-model of GB grammar (Chomsky 1986, 1988, 1989, Speas 1990, Giorgi & Longobardi 1991, etc.), for example, a multistratal, dynamic model which implies sequentially ordered representations (thus, derivations), there is an initial level called D-Structure which is assumed to be a literal representation of underlying grammatical relations (Thematic Relations), and which results from the Lexicon, the Projection Principle **and** the X-bar principles. As adjuncts are not projected from the lexicon, and since D-S is taken to be a pure representation of thematic structure (cf. Chomsky 1981, Speas 1990: 1-2) we must count on an intermediate level between D-S and S-S at which adjuncts are added, but the rules or principles licensing adjunction structures, say (2') above, need not be different in kind from those in (2), so I will ignore that little ripple in what follows.

Contemporary generative grammars fall broadly into two classes: 'multistratal' and 'monostratal' (cf. Newmeyer 1983, Gazdar et al. 1985 for those terms). GB is a 'multistratal' and 'dynamic' (Stockwell's term, cf. Stockwell 1980) model. It depends on multiple representations which are not simultaneously available, since there are 'feeding relations' among them, i.e., S-Structure results from D-Structure only when Move Alpha has been allowed to apply, but until the basic configuration is constructed by the Lexicon, the Projection Principle and the X-bar principles, Move Alpha just cannot apply, since it presupposes nodes identified as X^0 and X^{\max} and spatial relations like C-command defined on such configurations. Thus, the separation between the Lexicon and the tactic component is quite neat in GB, and, indeed, the Lexicon is logically and procedurally prior to both the Projection Principle and the X-bar principles, since unless a lexical head has been 'activated', the Projection Principle has no information to project, the X-bar rules/principles have no nodes to configure, etc. Therefore, although Chomsky and others usually emphasize the neutrality of the grammar as to speaker / hearer orientation and its 'declarative' interpretation as just a 'static', modular set of well-formedness conditions, the most coherent and natural interpretation of a GB grammar is 'dynamic', and 'procedural', and has the Lexicon as the starting point of all derivations.

Mono-stratal theories, by definition, exclude derivations (and thus 'feeding' relations among representations), and therefore their most natural interpretation is in terms of a set of modules converging on a single 'surfacy' representation, but the dichotomy syntax/lexicon remains fully operative, and, correspondingly, a separate 'tactic component' is clearly detectable. In GPSG and HPSG, for example, there is a large 'box' containing the Lexicon, but then the structural configurations are determined by two sets of rules, context-free ID rules of the X-bar type, which express dominance,

endocentricity, and co-occurrence information related to subcategorization, and LP statements, which just determine linear order (cf. Gazdar et al 1985, Pollard & Sag 1987, and Escribano 1988a, 1989d, for discussion and criticism).

Thus, the differences with respect to GB are rather slight at best and may be summarized as follows: a) GPSG is a bit more conventional, and less reductive, than GB in that subcategorization information remains in the X-bar ID rules and is merely alluded to by special ad hoc features in the lexical entries, and b) linear order is directly and separately expressed by LP statements, instead of being indirectly imposed by the restrictions in (5e,f) and (6a-e) above. In all other respects, both the contents of the respective theories and the 'division of labour' among their components is very similar.

Now, one of the aims of this paper is to take issue with the opposition between a lexical component and a tactic one. Following the line of reasoning inaugurated by Heny and Stowell and recently pursued by Speas, on the one hand, and the criticism of the content of X-bar theory in Kornai & Pullum 1990 on the other, in what follows I will try to prove that no neat separation of the syntax and the lexicon is possible (since lexical information certainly includes not only the terms that must co-occur, but even the ways in which they must be projected), and, indeed, that under my interpretation of key notions like Government or Theta Theory, there is no job left for the X-bar component (in any of its guises) to perform. Our syntax becomes 'transparent' in the sense that it adds no information whatsoever (thus, it disappears from the grammar), and all configurational features defined in (5-6) above, including binary branching, follow from lexical information under a natural interpretation of indispensable notions like 'government' or 'theta-role'.

However, my argument rests on the hypothesis that X-bar rules are not only redundant, but inadequate, i.e., not the right theoretical device to express the empirical properties of the syntax and semantics of natural languages, as argued elsewhere (cf. Escribano 1989a, c), and, as regards the charge of redundancy, it focuses on a rather different kind of reasoning. In order to avoid redundancy and compare the two approaches in a more perspicuous way, I will start by briefly reviewing what those scholars have to say about the X-bar component.

First of all, it is clear that parsimony considerations such as Heny's or Chomsky's in criticizing rules like (1) above are not really applicable if the rules are as general as those in (2-4), in which only variables are mentioned. However, there are other empirical and conceptual considerations that force us to reject them.

An obvious drawback of rules like (2-4) is that, under what seems the most straightforward interpretation, they fit most naturally into a 'top-down' generation strategy which is rather alien to the spirit of current GB. The resulting grammar, on the other hand, overgenerates, and cannot be implemented without allowing for massive backtracking, the reason being that, as in all earlier versions of TGG, the starting point of the generative process is an axiomatic symbol (S, S', E, in earlier models, X" now). Starting from X^{\max} , then, rules like (2-4) would recursively generate a syntactic skeleton, and only when X nodes have been reached can actual lexical items be 'inserted' in the tree. Of course, nothing stops such rules from applying an infinite number of times and never reaching the point at which lexical pieces can be introduced, but, even if we disregard such extremes, there is no doubt that only a posteriori, when a given lexical head has been selected, can the grammar determine how many and which types of branches should have been generated in any particular derivation. In other words, the branches generated in excess must be pruned, and in modern GB versions of

the theory the pruning factor is Chomsky's Projection Principle (See Chomsky 1981). However, Chomsky's Projection Principle filters only the excess of arguments. The excess of adjuncts of various types must be corrected by some other pruning device, perhaps Emonds 1976 condition to the effect that a node remaining empty throughout the derivation must automatically be deleted.

Be that as it may, there is no doubt that a lot of backtracking will be involved in the process, and to that extent the model is bound to be computationally inefficient and, above all, unrealistic from the psycholinguistic point of view, since speakers are not likely to go about generating useless syntactic skeletons to leave them empty first and eventually discard them. A bottom-up generation strategy is rather more akin to the GB model and obviously preferable in this respect: if the generative process starts directly from lexical heads and we can develop an extended Projection Principle like Speas's 1990 'Project Alpha', only the strictly required configurations will be generated, in the first place, and none will have to be discarded. As a consequence, there will be no backtracking, the device will be vastly more efficient in computational terms, and the celerity with which natural language is processed will have become proportionally more understandable.

Of course, it is possible to interpret schemata like (2-4) in a non-dynamic fashion (as perhaps implied by Chomsky's symptomatic replacement of the earlier arrow '->' by a ':' sign in them), but the point is, even under a 'static', 'declarative' interpretation, what information do such schemata add to what we can deduce from general principles?

If we take a rule like (2), repeated below for convenience, all it says, under standard interpretive conventions, is that phrases of any category X^n contain or dominate other phrases of the same categorial type, except that their bar features may (2a-b) or may not (2') have a lower value. That may be one way of stating endocentricity, but notice that (2) does not directly tell us that the lower X is the head of the construction; that must be derived from other principles, be they the Projection Principle or Gazdar et al.'s Head Feature Convention.

On the other hand, rules (2a-b) do tell us that Spec precedes X and other elements constructed with X follow it, but they do so in a redundant and at the same time incomplete way. Since it is generally assumed that UG contains a parameter fixing the relative position of arguments and specifiers with respect to the head, that aspect of (2) is just a repetition of the value specification for that parameter in languages like English and should be deleted. Then, of course, (2) is not specific enough concerning the relative position of NP, PP and S' complements, and will overgenerate unless supplemented by extra principles to that effect, so the order of the symbols is not doing any real work in (2).

But notice that as soon as we remove the order of elements, the rules (2a-b) do not add anything to the Projection Principle: the fact that more 'expanded' versions of phrases dominate or contain less expanded ones follows from the idea of expansion and general logic considerations ($A+B$ includes A , but not vice versa), and the fact that the relevant items are a dominating maximal / non-maximal projection, and one or more daughters all but one of them maximal, follows from the Projection Principle and the concepts of Government and Theta Theory: any X^n must end up dominated by an X^{\max} , otherwise the PP would be violated; on the other hand, X^{\max} is a barrier to government, and that means that no non-maximal projection of X is licensed in construction with another node Y unless X is itself a governor and theta-assigner and Y its governee and theta-marked complement (and thus an X^{\max}). But if X is a governor and theta role

assigner, it should be impossible to construct it with any other node unless X is an unsaturated projection (thus: $X^{\text{non-max}}$).

Putting those conceptual pieces together, we expect combinations of the form $X^{\text{max}} + X^{\text{non-max}}$ to be in, and those of the form $*X^{\text{max}} + X^{\text{max}}$ or $X^{\text{non-max}} + X^{\text{non-max}}$ to be out, i.e., exactly the configurations defined in (2a-b). In fact, in a way (2a-b) is rather less informative than the Projection Principle, since it leaves the exact membership of $*Y^{\text{max}}$ unspecified, whereas the Projection Principle (together with the Lexicon it refers to) tells us what complements must accompany each head. In sum, as the schemata (2a-b) cannot possibly replace the Projection Principle, Government-Theory and Theta Theory, we would have to keep them all, but then (2a-b) would not be doing any work in the grammar.

Notice that if instead of using X-bar schemata like (2) we couch X-bar information in the form of a set of principles like (5) the content remains equally redundant, given the existence of the other modules. Endocentricity, Successivity, Maximality, Peripherality, and Optionality (5a,b,c,e,g above) all follow straightforwardly from the notion of dominance, the Projection Principle and Government Theory, i.e., Endocentricity follows from the Projection Principle under standard feature percolation conventions; Successivity follows from the notion of expansion as soon as optionality of non-heads is assumed; Maximality comes for free, too: only maximal projections are licensed in governed positions, as we have seen; Peripherality is a consequence of the directionality of government; lastly, Optionality, of course, is meant to be overridden by the Projection Principle.

The only exception might seem to be (5f), i.e., the relative position of specifiers and complements at opposite sides of the head, but notice that if Specs are functional heads, as forcefully argued by Abney 1987 and more or less generally established in current theorizing, then a single principle of Government accounts for both cases: governors precede/follow their governees. Uniformity, on the other hand, does not follow, but then, as pointed out by Kornai & Pullum 1990, this is one of the weakest claims of X-bar theory, as it is trivially possible to satisfy it by positing empty projections like Jackendoff's M, M', M'', M''', Art, Art', Art'', Art''', etc.

If anything, the only substantial information that remains so far unaccounted for is (5h), **binary branching**. As we said, this is one of the more controversial claims, not a generally accepted constraint. However, even **binary branching** follows from Theta Theory and the Theta Criterion under most natural assumptions, after all. If A and B are 'constructed' together, one is a governor and the other a governee, one is a theta role assigner and the other receives a theta role. Now, suppose we had a construction of the form A B C. One of them, say A, would have to be a predicate and theta role assigner, and the other two would have to be arguments taking theta roles from A. The possibility of there being two predicates, say A and B, and one argument, C, is ruled out by the Theta Criterion, since C would receive two Theta Roles, one from each of the predicates A and B. Let's now assume that unless predicates A and B govern and theta-mark C, they will not be 'constructed' with C in the relevant sense (i.e., we interpret 'construction' in a strict sense for which see Escribano 1991b chapter 3 and below). The impossibility of such an A B C construction follows without further stipulation.

But now notice that if we adopt that strict view of 'construction' much the same argument holds for the case in which, in a construction A B C, one, say A, is a predicate, and B, C are arguments theta-marked by A: B and C may each be independently constructed with the predicate A in the relevant sense, but B, an argu-

ment, may **not** be constructed with C, another argument, at all! An argument B may not govern or theta mark another argument C, nor vice versa. If B did theta-mark C, (or C, B) as A does, C would get two theta roles, and for all purposes we would be in the situation above, i.e., a violation of the Theta Criterion would follow. Thus, under a strict interpretation of 'construction' along such lines, no A B C construction will be possible in any natural language. In essence, this is the content of our Transparency hypothesis below.

Of course apart from redundancy considerations, there are also positive, but equally compelling, reasons to abandon the PS rules, X-bar rules or principles, ID+LP formats, and, in short, all purely 'structural' tactic principles. As I showed elsewhere (Escribano 1989c), the positive case against such rules, especially n-ary branching ones, primarily rests on the following facts:

1) That, under standard assumptions, the rigid canonical geometry X-bar rules impose on phrases clashes with the results of standard syntactic tests like pro-form substitutions, coordination, deletion, etc. The structures X-bar assigns to phrases precludes the recognition of many constituents activated by such processes as One-Pronominalization or Coordinate Deletion and force on us a choice between two equally uncomfortable theses, a) that One-pronominalization and the like are different from other anaphoric processes in that they do not invariably affect constants (i.e., categories of the X-bar system), or b) that phrases are simultaneously structured in many incompatible ways. The first option saves the X-bar geometry at the cost of relaxing one of the most important constraints in the theory. The second saves the constraint but leaves us with no methodological basis to do syntax. Of course, neither alternative is really viable.

2) That, under the reasonable assumption that semantic predicates and arguments should be syntactically represented as sisters, n-ary branching rules produce semantically uninterpretable structures. Complements to verbs and nouns, in particular, fall semantically into two types. If they are 'arguments', by definition, an argument is an argument of one and only one predicate, and therefore only a binary-branching sub-tree offers a semantically transparent syntactic structure. If they are predicates (different classes of adverbials in the environment of the verb, and adjectives, PP's or relative clauses, etc. as modifiers of the noun) they are predicated of the corresponding argument, i.e., a projection of their head, and, again, the only transparent configuration is a binary-branching one. (See Escribano 1989c, 1991 for sustained arguments in this spirit).

3) That, apart from such theory-bound considerations, in fact, in English, as in Spanish and other languages, an overwhelming variety of syntactic tests seem to support binary-branching as the only legitimate syntactic configuration, anyway.

There is, thus, substantial syntactic and semantic evidence that Jackendoff's, Chomsky's, or Stuurman's multiple-branching schemata above are both redundant and too permissive. If, as I contended in Escribano 1988b, 1989c, 1991b, natural languages are characteristically binary-branching, and syntactization means organization of the message in a binary-branching fashion according to pragmatically conditioned strategies of focus vs. presupposition, significant improvements in simplicity and transparency can be attained in both the syntax and the semantics.

Of course, that is an argument against multiple-branching X-bar schemata, not one against X-bar rules in general, but the previous objections hold against the existence of a separate X-bar component in general, and, to the extent that they are sound, the

conclusion is clear: it is the positive information embodied in the supplementary principles of (6) associated to the theories of Government, Theta Marking, Binding, etc., that is doing the real tactic work in the grammar, and not the presumed X-bar principles of (5).

Of course, (6) could turn out to be mere consequences of deeper logical, cognitive or information-processing principles. In particular, the prominent position reserved to the subject might be a consequence of the structure of predication at UG, the Thematic Hierarchy and UTAH surely follow from general cognitive constraints on prominence, as acknowledged by Jackendoff 1987, Baker 1988, Speas 1990, and others, the c-command requirement of operator-variable chains looks like another cognitive constraint, and the requirement that clauses appear in peripheral positions opposite the head looks like a perceptual constraint.

All that is plausible, but is a different issue, and does not impair the obvious conclusion that the X-bar module, as currently conceived, does no real work at all and should be dispensed with. In fact, the extreme ease with which it can be reduced to independently necessary and more general principles, without loss of empirical adequacy, offers additional support for the essential correctness of that approach. In the rest of this paper I shall explore such principles and sketch what seems to be a more promising conception of natural language tactics.

4. Transparent syntax

Following the line of reasoning above, the model of grammar we have in mind in this paper contains a fully specified Lexicon and assumes that 'linguistic tactics' is transparent, or indeed, mostly embodied into lexical structure. The crucial notions of Government and Theta Theory, in particular, are seen as aspects of lexical structure, so the model in fact trivializes the constructive side of traditional syntax and semantics. The Lexicon, of course, is independently required no matter the syntactic-semantic options taken, so we will not say anything else about it in this paper. Apart from the principles embodied in the Lexicon, the only substantive principles required are those governing the computation of long distance dependencies (operator-variable and antecedent-anaphor chains). Needless to say, the phonological side of the grammar is irrelevant to our purposes here and has been left out of consideration. Obviously, part of the phonological structure is also contained in the lexical entries. The rest is created by a late, surfacy phonological component.

In section 3 above, I suggested that, although the tactic aspect of lexical information may be logically distinguishable from other aspects (the number of arguments each predicate takes, its categorial features, etc.), in practice, all the 'tactics' required is directly implied in the lexical information contained in the lexicon; and vice versa: it would be impossible to express that lexical information without assuming the tactic principles of government and theta theory. Thus, if we mention them separately, this is merely an analytic and expository strategy.

Crucial to our approach is the **Transparency Principle** (TP), which, informally stated, says that syntax is just a way to secure an unambiguous and readily computable expression for semantically relevant relationships. This rather obvious fact implies that only interpretable semantic relations should find a place in syntactic representations, where 'interpretable' semantic relations include the **class-inclusion** relation (expressed in syntax by 'vertical' relations between nodes and their heads), and two 'horizontal'

relations, i.e., **conjunction**, formally expressed by syntactic coordination under sisterhood, and what we shall broadly call **modification**, formally expressed by two classes of government, G1 and G2.

G1 holds of unsaturated heads and their arguments and is directional and subject to parametric variation. G2 holds of adjuncts, and saturated projections (X^{\max}) and is not directional. Both are formally expressed under sisterhood, and both are interpretable according to predicate-argument schemata: complements of the head are obviously arguments in the first sense, and adjuncts are interpreted as predicates over their respective heads. It follows from TP, under this interpretation, that two constituents will be sisters only if they stand in a relation of conjunction or predicate-argument. Given the Theta Criterion, arguments cannot stand in any of the interpretable semantic relations to each other, as we saw above, so they just cannot be sisters; in fact they cannot even be arguments of the same head at all. This effectively imposes binary-branching throughout the grammar without stipulation. Given our notion of transparency, binary branching follows from Theta Theory and the Theta Criterion.

Of course, we must adopt some principle equivalent to Chomsky's Projection Principle (PP), because we must make sure that the inherent and contextual properties of lexical items and their projections are satisfied in well-formed representations at all levels. Let's assume, therefore, a **Generalized Matching Principle** (GMP) interpreted as implying at least the following 'sub-principles':

a) **Matching** of categorial (i.e., 'endocentricity') and other inherent features (roughly, the 'Head' features of GPSG and HPSG) of nodes and their phrasal expansions (i.e., Number, Gender, Person, Case, Tense, S(ematic)-Type, etc.). Informally, if in the case of the NP subject of the sentence 'the students went on strike' the mother node bears the features [Cat: N], [Number: Plur], [Gender: Masc], [Person: 3], [Case: Nom], [S-Type: Entity], [Theta Role: Ag], say, we want its head N to bear exactly the same [Attribute: value] specifications. Otherwise, the head node N will have been inserted in the wrong local configuration and its inherent and contextual properties will not have been observed, in violation of the GMP.

b) **Saturation**, i.e., the thematic roles projected by the predicate must be properly 'discharged', and they are so discharged when the associated variables are interpreted. This may be seen as a particular case of matching, i.e., an environment matches a predicate when it contains suitable referential XP's (typically NP's) capable of binding the predicate's variables and discharging its thematic roles, where 'suitable' NP's means NP's whose denotation is not incompatible with the semantic requirements culturally fixed for the corresponding theta roles (i.e., in our civilization, Agents must be animate, Places and Times must be inanimate and are restricted in other ways, etc.).

In order to implement the GMP we must count on a) a precise definition of the different kinds of linguistic information, and b) suitable feature instantiation and transmission mechanisms. Of course, detailed discussion of such matters is out of the question here, but a few suggestive remarks are surely in order.

In general, the approach we take to the representation of linguistic information is inspired by, and similar to, that taken in GPSG, HPSG, LFG and Categorical Grammar, but we generalize it so that it becomes **the only kind** of notation in use, i.e., we use [Attribute: Value] pairs to represent **all** linguistic information, without exception. Thus, syntactic category features are expressed by pairs like [Cat: N], [Cat: V], etc., and Gender, Number, Person, Tense, Case, for example, are expressed by comparable pairs [Gender: +/-Masc], [Number: +/-Plur], [Person: 1/2/3], [Tns: +/-Past], [Case:

Nom/Gen/Obj], etc., as in GPSG, but the Theta Roles associated with a lexical predicate are also expressed by pairs [Agent: x], [Theme: y], [Loc: z], etc., referential properties are expressed by pairs of the form [Ref: index], functional properties of nodes are noted by pairs of the form [F: x] (where F stands for one of the syntactic functions 'subject of', 'object of', 'head of', etc., and x denotes nodes). Even positional information like 'immediately precedes / follows X', 'immediately dominates Y' may be so expressed if necessary: [IP: x], [ID:y], etc. (Thus, trees, labelled bracketings, and other bidimensional formats we may use are mere representational devices, have no theoretical significance and can be dispensed with). Of course, not all that information need be directly associated with nodes. At least part of it may be inferred from general principles of the grammar by means of trivial inference rules.

The process by which lexical information gets projected is construed essentially as in Categorical Grammar. A lexical predicate like 'give' will have a lexical entry in which, apart from its categorial feature [Cat: V], there will be thematic attributes corresponding to all the arguments 'give' requires, with their corresponding free variables, i.e., (7):

- (7) GIVE [CAT: V]
 [LEVEL: -P]
 [FORM: Inf]
 [S-TYPE: Action-Pred]
 [REF : i]
 [TIME OF i : v]
 [THEME OF i : x]
 [GOAL OF i : y]
 [AGENT OF i : z].... etc.

(Where [LEVEL:] refers to the lexical/phrasal status of nodes, [FORM:] is like VFORM in GPSG, [S-TYPE:] expresses 'semantic type', and [REF:] establishes referential values for nodes of different types; cf. infra.)

According to the information expressed in the partial lexical entry in (7), the item 'give' will be legitimate if inserted under a categorial node which has non-incompatible values for those [A:v] pairs, i.e., it must be [CAT: V] (but not *[CAT: N]), and [LEVEL: -P] (i.e., not phrasal), and must show the same thematic attributes.

The phrasal node right on top of V, in its turn, will be [CAT: V] too, but the [LEVEL:] attribute will have a different value, +P(hrasal), instead of -P. Apart from that, of course, the value of the first of the thematic attributes, [THEME OF i:], will not be the variable that appears in the lexical entry and in the lexical category V, but a referential index shared by an appropriate NP governed by 'give'. If that index is '1', in the node right on top of V in the tree the pair [THEME OF i: x] will be replaced by [THEME OF i:1], i.e., the argument role THEME will have been discharged.

Now, for the resulting configuration to be well-formed, the attributes of the mother-node must 'match' the environment it dominates according to GMP, i.e., that node must dominate an XP, in this case an NP, with a [REF:] attribute valued '1' (i.e., [REF:1]) and a thematic attribute [THEME-OF:] valued 'i', where 'i' coincides with the value of the [REF:] attribute on the head (i.e., [THEME-OF: i]). In other words, either the NP is interpreted as THEME OF the event denoted by the projection of 'give', and as referred to the individual indexed '1', or the NP itself will be illegitimate in the suggested configuration, and the whole sub-tree will be ruled out. If no such NP is available in the

environment governed by V, or the only available NP is not interpretable as THEME OF an action predicate like 'give', or the referential index 'i' does not coincide, the phrasal node on top of V will not 'match' in the required sense, and the corresponding subtree will be excluded by GMP.

Similarly, the next ascending phrasal head node will have the same features as its daughter, except that another index, say '2', will figure instead of the variable 'y' as the value of the predicate [GOAL OF i:], provided it dominates an XP (say a PP, in this case) with an NP containing a [REF:] attribute valued '2', and a thematic attribute [GOAL OF: i], where i coincides with the [REF:] value on the head. (This NP, of course, will have to match the thematic structure of the preposition and its features will eventually appear on PP, too, thus becoming 'visible' for further computation, cf. infra).

We may assume that the same mechanism applies until all the thematic attributes associated with the head have been 'discharged' in the same way and no unbound variables remain in the head node. At that point, the G1-projection licensed by 'give' will be 'saturated'. A saturated G1-projection is a (minimum) 'maximal projection'. It is also the turn-off point for the directional relation of Government-1 (G1), and the turn-on point for Government-2 (G2), which, as we said above, takes care of the relations between non-arguments and heads.

Maximality must be conveniently marked in the grammar, for all sorts of reasons (among others, that only maximal projections refer, only X^{\max} may be affected by Move Alpha, etc.). Therefore, we must set up a special attribute [MAX: +/-] and link it to our previous attribute [LEVEL: +/-P] by a Feature Specification Default, i.e., (8):

(8) [MAX: +] > [LEVEL: +P].

The idea behind (8) is that the level of nodes may be lexical or phrasal, and phrasal nodes, in turn, may be maximal or non-maximal. This is an alternative to setting up Bar levels 0, 1, 2,...n, as in GPSG or GB, and obeys my conviction that only the two oppositions between Phrasal / Non-Phrasal and Maximal / Non-Maximal are relevant (cf. Escribano 1991, chapter 3 for discussion; Speas 1990 adopts the same point of view).

Of course, a saturated projection, by definition, cannot be further 'G1-constructed' with items interpreted as arguments, since by that time no unbound variables will remain in the frame, so excess of arguments is automatically ruled out. This captures one of the essential aspects of the Projection Principle.

However, nothing prevents a saturated projection from being G2-governed by another maximal projection (i.e., by an adjunct of a suitable type); so, we must define principles constraining G2-government and the well-formed configurations at this level.

The simplest hypothesis in this respect is to assume that adjuncts are themselves 'predicates' (we will call them 'P2-Predicates', to avoid confusion) G2-governing, and predicating over, a special kind of 'arguments' (again, say, 'A2-arguments', to avoid confusion with proper arguments), i.e., their respective heads. What kind of predicates they are depends on the [S(EMANTIC)-TYPE:] attribute on their lexical head, i.e., a prepositional phrase like 'in London' will be marked [S-TYPE: LOC(ATIVE)], a PP such as 'with a hammer' will be marked [S-TYPE: INS(TRUMENTAL)], an adverbial like 'yesterday' will bear the feature [S-TYPE: TIM(E)], etc.

Of course, the S-TYPE attribute is one of the 'Head' features and will have to match the S-TYPE on the lexical head (the prepositions 'in', 'with', the lexical item 'yesterday',

etc.). [S-TYPE:], however, must be 'visible' at the level of the corresponding (G1) maximal projection, since it is relevant for the purposes of defining its further possibilities of (G2) construction. Thus, whereas thematic attributes are saturated in the process of projection internal to the smallest X^{\max} , [S-TYPE:] attributes are inherent in each lexical head and necessarily remain visible at the top of its G1-projection. The next step is to devise principles establishing the domain of P2-predicates of the different S-TYPE's. Of course, a full taxonomy of semantic types is beyond the scope of this paper, but we may offer a rough suggestion as to how we would proceed.

Suppose we say that ENTITIES and EVENTS imply/are compatible with [S-TYPE: LOC] and [S-TYPE: TIME], that ACTIONS and PROCESSES in addition imply/are compatible with [S-TYPE: MAN(NER)], whereas only ACTIONS are compatible with [S-TYPE: INS(STRUMENT)], and that both ACTIONS and PROCESSES are types of EVENTS, etc., so that, for example, an action predicate like 'marry' bears not only its mandatory THETA ROLE attributes ([AGENT of i:] ([Theme of i:]), etc.) but, by implication, attributes like [TIME OF i: t] (where 'i' is the index associated with the maximal projection eventually licensed by 'marry', and 't' stands for a time variable).

Now, assume we set up a new type of attribute-value pair, say [SEMANTIC-ROLE, v], where the attribute stands for one of the values of S-TYPE interpreted as a relation to an entity, event, etc., i.e., [LOC OF:], [TIME OF:], [INS OF:], [MAN OF:], etc., and the values are entity-variables, action variables, event-variables, etc., depending on the domain of the predicate.

Now, further suppose that we encounter an X^{\max} like 'John and Mary married' (a sentence = V^{\max} , in this case) 'modified' by an adjunct like the PP 'in 1991'. According to our previous assumptions, the head of the VP will bear the attributes [S-TYPE: Action] and, by implication, [S-TYPE: Event] and [TIME of i: t], and so will its maximal projection (i.e., the VP itself), since these are Head features. In addition, the VP, being a maximal projection, is referential, so let's suppose we index events like 'John and Mary married' (or 'John and Mary's marriage', for that matter) by means of the same type of [REF: i] feature we used above, so that the VP 'John and Mary married' bears the attribute [REF: i] (where 'i' refers to an event and coincides with the value of 'i' in the expressions [TIME OF i:] and [REF:i]). The lexical head of the PP adjunct, P, in its turn, (and its maximal projection PP) will bear the attribute [S-TYPE: TIME], and PP, being a maximal projection, will also have a [REF:] attribute, say [REF: t]. According to our hypothesis, the S-ROLE of the PP adjunct will be [TIME OF: i], for 'i' an event variable.

Now, provided the value of 'i' in [REF: i] on the head and [TIME OF: i] on the adjunct, on the one hand, that of 't' in [TIME of i: t] on the head and [REF: t] on the adjunct on the other, match, the S-ROLE of the adjunct will be 'saturated' and the adjunct will be interpreted as 'time t of the event i', where 'i' refers to the event denoted by the VP. Correspondingly, the [TIME OF i : t] predicate on the VP head will be interpreted by the [REF: t] value on the adjunct, i.e., the expression 'in 1991' in our example.

In other words, the variable 't' in the [TIME of i: t] attribute implied by the event 'John and Mary married' will be bound by the value of 't' in the [REF: t] attribute of the adjunct ('1991'), and, correspondingly, the event variable 'i' of the attribute [TIME OF: i] on the adjunct will have been bound by the [REF: i] value associated with the VP head.

The direct consequence, of course, is that the adjunct cannot now be G2-constructed

with any other head, and the head VP cannot take any other temporal adjunct (since its [TIME of i: t] variable is already bound by the [REF: t] value on the PP). The description of the process is rather time-consuming and we will not insist, but suffice it to say that a parallel approach may be taken to all other adjuncts, quantifiers, etc.: to all non-arguments subject to G2, in short. (Of course, in order to develop this systematically, we need a full taxonomy of S-types which has not been elaborated yet, so this is merely suggestive of how the general approach sketched here might be fully developed in a way consistent with our assumptions so far).

On the whole, then, the well-formedness of representations depends on the compatibility of features at different nodes according to the Generalized Matching Principle. This means that a linguistic representation is ultimately (and exhaustively) reduced to a set of [A: v] pairs. Trees, labelled bracketings, and the like are just convenient bidimensional representations of such sets, but it must be kept in mind that it is the relations expressed by [A: v] pairs and the logical operations applying to them that are crucial to the approach.

Nevertheless, if we are going to use trees, we must count on certain conventions. Thus, in order to secure an accurate mapping and a systematic correspondence between abstract government relations and syntactic trees, we must adopt a (trivial) representational convention, the **Sisterhood Principle** (SP), according to which both classes of government (G1 and G2) translate into tree structure in terms of sisterhood. According to SP, then, any pair of constituents that should be interpreted as a predicate plus its argument, in both senses of such terms, must appear in a local syntactic tree in which such constituents are sisters. This is the binary branching condition of Kayne 1983 and Speas 1990, of course, but in our account nothing special need be said about it: it follows from Government theory and Theta Theory, as we saw above.

5. Conclusions and implications

In sum, granted the existence of a properly specified Lexicon, the 'tactic' component of a formal grammar roughly in the spirit of current GB, but with our additional constraint of Transparency, can be trivially reduced to just the Generalized Matching Principle as formulated above or slightly revised variants thereof, and such a trivial picture of what syntactic competence in natural language amounts to is exactly correct, we believe, in all the relevant respects:

a) At the descriptive level, such a grammar yields a highly restrictive binary-branching geometry that satisfies the usual syntactic constituency tests without introducing the rigidity attending more traditional PS or X-bar solutions. It, therefore, answers the objections in Escribano 1989c and above and allows for a maximally simple and transparent syntax-semantics mapping.

b) At the explanatory level, the great simplicity and naturalness of the matching principle invoked offers a credible account of the acquisition of syntactic competence and of the efficiency with which it is activated in natural language processing. Significantly, the nature of the principles invoked is such that it is not necessary to assume with Chomsky that they are manifestations of a separate and specific language faculty, as matching operations of that type can readily be attested in other cognitive systems; an obvious conceptual advantage.

c) Finally, at the performance/computational level, the model locates complexity where it is comparatively inexpensive in computational terms, the data-base component

and the associated Lexicon interfacing it, and simplifies the computational mechanism, any extra complexity of which is widely recognized as incomparably costlier.

The model sketched above, on the other hand, is inherently 'declarative', but it admits a straightforward interpretation in procedural terms, i.e.: 1) the generative process starts from so-called 'conceptual' structures in close interaction with the Lexicon (see Jackendoff 1983, 1987); 2) there follows a process of 'lexicalization' (basically matching of specifications and choice of perspective, but see Escribano 1991, Chapter 3 for some details and justification) that yields lexical choices for the semantic entities involved (individuals, classes, properties, relations, actions, processes, states,...etc.); finally, 3) the information implied in lexical items is 'projected', and nodes are amalgamated and computed in a bottom-up fashion according to the GMP, as described in section 4 above, until they eventually yield well-formed linguistic expressions.

One of the characteristic features of this model is the way it derives the binary branching (BB) property from general principles of Government and Theta Theory. So given the controversial nature of (BB) and the light it casts on the syntactization process, this issue surely merits a few final considerations.

Of course, the existence of **some** BB structures is not in question. After all, many predicates are clearly monadic and fit nicely into this pattern. What is at issue is whether **all** NL structures are BB structures. After all, it may be argued that predicates like 'sell', 'give', etc. are logically and conceptually polyadic, so why should n-ary branching structures be banned?

Indeed, 'give' is a polyadic predicate, (hence the multiple thematic attributes we associated with its lexical entry in the preceding discussion). However, NL is **not** Logic, nor Linguistic Structure can be identified with Conceptual Structure, and therefore a predicate may be logically polyadic and yet be forced into a BB structure in the process of 'syntactization', and this is exactly what we claim happens. Syntactization, in other words, implies organizing logical conceptual structure in a way compatible a) with UG principles, and b) with certain principles of discourse organization we briefly mention below. UG principles include GMP, Government Theory, Theta Theory and the Theta Criterion, and these, as it happens, exclude multiple branching. Mapping of possibly n-ary branching conceptual structure into a BB structure, in other words, is the essential import of syntactization. If this is correct, we surely are entitled to ask, is that just a biological accident, (= a contingent property of our linguistic mechanism), or is it functionally motivated?

In my view, it is functionally motivated, although the deep cause may still be in the way our brain processes information. Arguably, any sentence in discourse contains two parts, 'old', or 'presupposed', and 'new', or 'focused', information. Let's call them Focus and Presupposition for short. Now, in a predication like 'John sold his computer to Peter FOR A THOUSAND POUNDS', the constituent interpreted as Focus will normally be the last PP, but we may change the Focus while expressing the same propositional content by altering the order of constituents, i.e., we may say 'John sold his computer for a thousand pounds TO PETER'. Other possibilities include extra emphasis on certain constituents that remain outside the usual Focus positions. For example, we may say 'JOHN sold his computer to Peter for a thousand pounds', with focus on the subject, or emphasize the object HIS COMPUTER, etc. Of course, the distribution of each of those variants in discourse is rigidly constrained. (Notice, for example, that each of them is appropriate as an answer to only one of the corresponding WH Questions, is equivalent to a different cleft sentence, etc.).

Of course, a natural interpretation of the focusing pattern is precisely in terms of alternative binary branching configurations, most prominently so in the case of WH-Movement. Notice that movement into Comp, or, as we prefer to call it, movement into operator position (cf. Escribano 1988b, 1991b) is just a kind of restructuring of the sentence according to a binary branching pattern Operator-Domain, or equivalently, Focus-Presupposition, which is patently driven by discourse constraints on the organization of information.

Thus, if our hypothesis is correct, this is, at bottom, the explanation for the BB constraint in NL. It is one of the major theses of the present approach that, in principle, any complex string may be structured in n possible ways depending on the focus-presupposition relations prevailing in the context. What we seem to need, therefore, is a grammar that allows for phrases to be structured in mutually incompatible ways depending on the context of use, but, significantly, there is one property all such alternatives comply with, and that is, precisely, binary branching. BB, however, is the property that most prominently enters into conflict with the standard assumptions of X-bar syntax and the rigid geometry they generate. By contrast, the picture of syntax that emerges from our present assumptions is most flexible in this respect. Transparent Syntax imposes only very weak and unspecific constraints, essentially, apart from GMP, that the speaker should make a decision as to what in his message is being 'presupposed' and what, on the contrary, is being 'focused'. That is all syntactization amounts to, since the grammar allows multiple structures to be 'activated' in different contexts of use as long as they satisfy BB and the other general principles discussed. Yet, at the same time, natural language tactics as defined by Transparent Syntax is quite a straightjacket in a way, since only binary-branching structures are licensed, but that is what makes the system so efficient. If our hypothesis is correct, the tactic mould UG enforces is quite meagre: it boils down to the combined effect of Government, Theta Theory (thus, BB) and GMP. Of course, the apparently paradoxical combination of a most restrictive and computationally unambiguous architectural constraint like BB and a great organizational flexibility, possible only in the context of GMP, is the key to the system's immense power and efficiency.

BB, although a consequence of Government and Theta Theory, has substantial computational and psycholinguistic implications. In a BB grammar, all the creativeness and compositionality of NL results just from computing a depth-one relation on adjacent matching nodes, a function surely general and simple enough to be a likely candidate to be attributed to whatever general-purpose cognitive devices are available to the child in the initial phase of language acquisition.

Therefore, the present hypothesis not only affords a maximally transparent syntax-semantics mapping and a most efficient computational design. It also yields a promising basis for an explanation of the efficiency of natural language processing and the celerity with which linguistic competence develops. Correspondingly, the present view of NL tactics reduces the force of Chomsky's poverty of the stimulus argument and makes it somewhat less inescapable to postulate a separate Language Faculty as the only way to explain the acquisition of Language.

References

- Abney, S. 1987
The Noun Phrase in its Sentential Aspect. Ph. Diss., MIT
- Baker, M. 1988
Incorporation: A Theory of Grammatical Function Changing.
Chicago: The University of Chicago Press.
- Belletti, A. & L. Rizzi 1988
'Psych Verbs and Theta Theory'. *Natural Language and Linguistic Theory*, 6, pp. 291-352
- Chomsky, N. 1955
The Logical Structure of Linguistic Theory. New York: Plenum Press, 1975
- Chomsky, N. 1972
'Remarks on nominalization'. In *Studies on Semantics in Generative Grammar*. The Hague: Mouton Publishers, pp. 11-61.
- Chomsky, N. 1981
Lectures on Government and Binding. Dordrecht: Foris Publications.
- Chomsky, N. 1986a
Barriers. Cambridge, MA: MIT Press.
- Chomsky, N. 1986b
Knowledge of Language. New York: Praeger Elsevier.
- Chomsky, N. 1988
Language and Problems of Knowledge. Cambridge, MA: MIT Press.
- Chomsky, N. 1989
'Some notes on economy of derivation and representation'. *MIT Working Papers in Linguistics*, 10, pp. 43-74.
- Emonds, J. 1976
A Transformational Approach to English Syntax. New York: Academic Press.
- Emonds, J. 1985
A Unified Theory of Syntactic Categories. Dordrecht: Foris Publications.
- Emonds, J. 1987
'Parts of speech in generative grammar'. *Linguistic Analysis*, 17/1-2, pp.3-42.
- Escribano, J. L. G. 1988a
'Una crítica de la GPSG'. *Revista Española de Lingüística* 19/2, 1989, pp. 291-376.
- Escribano, J. L. G. 1988b
'Los efectos de inversión y la estructura subyacente de la oración en inglés'. *Revista Española de Lingüística*, 20/2, 1990, pp. 329-401.
- Escribano, J. L. G. 1989a
'Natural Language Syntax: a Restrictive Approach'. Unpublished Conference Paper presented to the First International Colloquium on Cognitive Science, San Sebastián, May 1989.
- Escribano, J. L. G. 1989b
'Sobre la estructura subyacente de la oración en español'. *Actas del V Congreso de Lenguajes Naturales y Lenguajes Formales*, pp. 553-574.
- Escribano, J. L. G. 1989c
'Sobre la naturaleza del componente táctico en las lenguas naturales: un argumento basado en el uso de la proforma *one* en inglés'. *Studia Patriciae Shaw Oblata*, Universidad de Oviedo, 1991, pp. 277-306.
- Escribano, J. L. G. 1989d
'GPSG.¿Una alternativa a la teoría GB? Estado de la Cuestión y Perspectivas'. *Actas del V Congreso de Lenguajes Naturales y Lenguajes Formales*, vol. 1, pp. 43-96.
- Escribano, J. L. G. 1991a
'En defensa de una sintaxis sin INFL: Observaciones críticas a ciertos análisis de la oración inglesa inspirados en la teoría de los principios y los parámetros'. En *Estudios de Lingüística Inglesa*, ed. M. T. Turell, AESLA, 1991 (in prep.).
- Escribano, J. L. G. 1991b
Una Teoría de la Oración. Oviedo: Servicio de Publicaciones de la Universidad.

- Gazdar, G. et al. 1985 *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.
- Giorgi, A. & G. Longobardi 1991 *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press.
- Heny, F. 1979 'Review of N. Chomsky, *Logical Structure of Linguistic Theory*. *Synthese*, 40, pp. 317-352.
- Jackendoff, R. 1972 *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jackendoff, R. 1977 *X-bar Syntax. A Study of Phrase Structure*. Cambridge: MIT Press.
- Jackendoff, R. 1983 *Semantics and Cognition*. Cambridge, MA: MIT Press
- Jackendoff, R. 1987 'The status of thematic relations in linguistic theory'. *Linguistic Inquiry*, 18/3, 369-411.
- Kayne, R. S. 1983 *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kornai, A. & G. K. Pullum 1990 'The X-bar theory of phrase structure'. *Language*, 66/1, pp. 24-50.
- Newmeyer, F. J., 1983 *Grammatical Theory*. Chicago: The University of Chicago Press
- Ouhalla, J. 1990 'Sentential negation, relativized minimality and the aspectual status of auxiliaries'. *The Linguistic Review*, 7/2, pp. 183-231.
- Pollard, C. & I. A. Sag 1987 *Information-Based Syntax and Semantics*. Stanford: CSLI
- Pollock, J. Y. 1989 'Verb movement, Universal Grammar and the Structure of IP'. *Linguistic Inquiry*, 20/3, pp. 365-424.
- Pullum, G. K. 1985 'Assuming some version of X-bar theory'. *CLS*, 21, pp. 323-353.
- Radford, A. 1988 *Transformational Grammar*. Cambridge: Cambridge University Press
- Ross, J. R. 1967 *Constraints on Variables in Syntax*. Ph. D. Diss., MIT.
- Speas, M. 1990 *Phrase Structure in Natural Language*. Dordrecht: Kluwer Academic Publishers.
- Stockwell, R. P. 1980 'Summation and assessment of theories'. In E. A. Moravcsik & J. R. Wirth, eds., *Syntax and Semantics*, 20, New York: Academic Press, pp. 353-381.
- Stowell, T. 1981 *Origins of Phrase Structure*. Ph.D. diss., MIT.
- Stuurman, F. 1985 *Phrase Structure Theory in Generative Grammar*. Dordrecht: Foris Publications.
- Williams, E. 1975 'Small Clauses in English'. In J. Kimball, ed. *Syntax and Semantics*, 4. New York: Academic Press, pp. 249-273.
- Williams, E. 1981 'Argument structure and morphology'. *The Linguistic Review*, 1, pp. 81-114.