Studies in the Linguistic Sciences
Volume 15, Number 2, Fall 1985

THE CASE AGAINST PLAIN VANILLA SYNTAX*

Arnold M. Zwicky

The adequacy of phrase structure theories of syntax has been in question for thirty years; despite the empirical attractions of such theories and their conceptual simplicity, it has seemed to many that these theories must be abandoned completely in the face of a long catalogue of defects. This paper first gives a careful account of the assumptions lying behind an utterly simple, or *plain vanilla*, (phrase structure) syntax, and then lists the presumed deficiencies of such a theory. The point of these exercises is to demonstrate what a theory of syntax must be able to do, and to show just how plain vanilla syntax falls short. More adequate theories of syntax can then be seen as particular packages of amendments to the basic model.

1. The issue. Transformational syntax arose in the context of American structuralism. Its immediate antecedents were *phrase structure* (or *constituent structure*) approaches to syntax, exemplified by three papers collected in Joos 1966 (Harris 1946, Bloch 1946, and Wells 1947). Chomsky (1957) observed that such approaches associated syntactic theory with discovery procedures - he argued that this association placed unreasonable demands on syntactic theory - and related such approaches to fully formalized models for the description of syntax. On the second score, the topic of this paper, Chomsky's (1956) formalizations (while making a major contribution to a new branch of discrete mathematics[1]) gave rise to two fundamental questions in syntax: Do these formalizations adequately capture the intentions of syntacticians like Harris, Bloch, and Wells? In any event, are the formal systems adequate for the description of the syntax of all languages?

Chomsky (1957) and Postal (1964) argued that the syntactic proposals of the American structuralists could all be interpreted as versions of *context-free grammar* (CFG), and they maintained that CFG was provably inadequate as a theory of syntax. On the textual issue, the question of whether CFG fully reconstructs the proposals of syntacticians like Bloch and Harris, I shall have little to say here; undoubtedly the core of each of these proposals is CFG, but virtually every one has features (often not formalized) beyond those of unadorned CFG, so that a case can be made on either side.

On the issue of adequacy, I shall have a great deal to say. The fact that early champions of transformational syntax cited multiple inadequacies of CFG and drew much of their evidence from familiar languages suggested quite strongly that no language is a context-free language (that no language has a context-free syntax) and provided considerable support for abandoning CFG in favor of some less simple and more powerful theory, such as transformational syntax. Such exceedingly powerful theories are attractive largely because they appear to provide solutions to *all* of the presumed inadequacies of CFG at once.

I outline (in section 2) a particularly stark and unadorned version of syntactic theory, according to which every language has a context-free syntax. Like the approaches of the American structuralists, this theory (which I dub *plain vanilla* (PV) syntax) takes as the central observation about sentences the fact that they are organized hierarchically. And in common with most varieties of American structuralist and transformationalist syntax, it identifies syntax as one autonomous component among a number of components in a full grammar of a language.

In taking hierarchical organization as the fundamental syntactic property of sentences, PV syntax contrasts with theories that focus instead on the head/modifier relation (as in dependency grammar, as summarized by Matthews 1981: 78-84), on grammatical relations like subject and direct object (as in the relational grammar of Perlmutter and Postal - see their papers in Perlmutter 1983), on the semantic and pragmatic functions of syntactic units (as in the *role and reference grammar* of Foley and Van Valin 1984), or on gap-filler and anaphoric relations (on the former as in the base generated syntax of Brame 1978, on both as in the government and binding syntax of Chomsky 1982). In insisting on hierarchical organization as the one central syntactic property, PV syntax also contrasts with theories in which two or more properties play a focal role: with, for instance, the daughter dependency grammar of Hudson (1976), the corepresentational grammar of Kac (1978), and the lexical functional grammar of Bresnan and Kaplan (see Bresnan 1982).

Section 3 summarizes the case against PV syntax. Much of this is familiar ground. Every textbook of transformational syntax provides some 'arguments for transformations' (which is to say, arguments that a much simpler phrase structure theory is inadequate), though different texts emphasize different arguments. And theoretical works have regularly cited one or another of the faults of PV syntax as support for other models, though again there has been no synthesis or summary of the evidence.[2] My intention here is not to dash cold water on CFG as a framework for syntax - in fact, I am recommending that phrase structure approaches as in Gazdar (1982) be vigorously explored - but rather to make it as clear as possible what conditions might be required of a syntactic theory that is to be fully adequate. Section 4 contains a few reflections on the lessons of section 3.

2. Plain vanilla syntax in summary. Three aspects of PV (or, indeed, any) syntax must be specified: first, the nature of the objects described by the syntactic rules (A and B below); second, the nature of the rules themselves (C); and third, the place of the syntax in a full grammar (E).

The first aspect I treat in two parts, both framed in tree terms: one (A) having to do with syntactic structure in a narrow sense, that is, with the hierarchical and sequential organization of words and phrases within sentences (in tree terms, with relations between nodes in trees); the other (B) having to do with the set of syntactic categories (in tree terms, with the stock of labels on nodes in trees) and its relationship to syntactic structure. In addition, the program of constituent structure grammar, both in its original structuralist versions and in later generativist developments, has had a distinctly formalist cast. Part D below gives explicit voice to some of the formalist assumptions of the program.

One complication in this discussion is, I believe, unavoidable. In order to represent what is intended in pretheoretical observations about constituent structure, it is necessary to distinguish two slightly different notions of 'constituent structure' - *preterminal structure*, which is purely syntactic (and so contains no information about particular lexical items) and *terminal structure*, which is preterminal structure as augmented by pointers to specific lexical items. The sentences Birds eat and Birds drink have 'the same constituent structure' in the sense that their preterminal structures are identical; but they are not 'the same constituent structures' (they would not count as identical for the purposes of ellipsis, for example), because their terminal structures are different. I discuss only preterminal structures in B, reserving terminal structures for the presentation of assumptions about the lexicon in E1.

A. Relations between nodes in trees

A1. The theory of syntax refers to exactly two primitive relations between nodes in a tree: immediate dominance, the relation of mother to daughter; and immediate precedence, the relation of left sister to right sister.

A2. The Single Mother Condition: A node has no more than one mother.

A3. The No-Tangling Condition: If nodes M and N are sisters, then none of the descendent nodes of M is a sister of any of the descendent nodes of N.

A4. The Acyclic Condition: No node is the mother of one of its ancestor nodes.

A5. The Unique Root Condition: There is only one node without a mother.

A6. The Total Ordering Condition: All sister nodes are linearly ordered with respect to one another.

B. Node labels

B1. The number of node labels is finite.

B2. Every node has a label.

B3. No node has more than one label.

B4. The set of labels can be exhaustively partitioned into two (nonempty and disjoint) subsets, of *phrasal* and *lexical* labels.

B5. The set of phrasal labels has a designated member, the 'initial symbol' S.

B6. Labels are unanalyzable wholes.

C. Syntactic rules

C1. A syntax for a language is a finite set of rules.

C2. Taken together, these rules define a set of preterminal structures.

C2a. Preterminal structures are objects of the sort specified in A and B above. Accordingly:

C2a1. A rule may concern the branching of a mother node with a specified label M into daughter nodes with specified labels $D_1,\ldots,D_n$.

C2a2. A rule may concern the linear ordering of a set of daughter nodes with specified labels.

C2b. Preterminal structures constitute the only type of syntactic representation provided in a grammar.

C3. Rules are positive only; a rule licenses a particular branching or linear ordering. There are no negative (syntactic) conditions, or (syntactic) filters.

C4. A rule specifies the linear ordering of a set of daughter nodes by means of the operation of concatenation.

C5. Each rule both licenses some branching and specifies a linear ordering of the daughter nodes. Accordingly, each rule can be expressed in the form $D_1 \ldots D_n = M$.

C6. Rules receive the generative interpretation (as opposed to the parsing interpretation or the checking interpretation; see the Introduction to Dowty, Karttunen, and Zwicky (1985)). That is:

C6a. A rule of the above form is to be understood as licensing the replacement, within some partial tree, of a node with label M by a subtree having a mother node labeled M and n daughter nodes labeled, from first to last, D_1 through D_n. In more familiar phrasing, the rule expands the symbol of the mother into the symbols of the daughters, in order: $M \dashrightarrow D_1 \ldots D_n$.

C6b. A tree represents a preterminal structure described by the syntax if, and only if, it satisfies three conditions: (i) it has a root node labeled S; (ii) it can be obtained by successive expansions, each

licensed by a rule of the syntax; (iii) no further rule of the syntax is applicable to it. (Condition (iii) ensures that no phrasal category symbol is left dangling because an applicable rule has not been used.)

C7. There is a rule with mother label *M* if and only if *M* is a phrasal label. In more familiar terms:

C7a. No syntactic rule expands a lexical category symbol.

C7b. Every phrasal category symbol is expanded by some rule. This condition ensures that no phrasal category symbol is left dangling because there is no rule to rewrite it.

C8. The number of daughters in a rule must be at least 1. This condition ensures that there are no dangling phrasal category symbols stipulated by the grammar, which is to say that there are no 'zero expansions' of such symbols.

C9. Rules are not subject to conditions referring to context:

C9a. Rules are not subject to conditions referring to horizontal context; their applicability cannot be conditional on the nature of the mother's sister nodes. That is, they are not 'context-sensitive'.

C9b. Rules are not subject to conditions referring to vertical context; their applicability cannot be conditional on the nature of nodes ancestral to the mother or descendent from a daughter.

C10. There are no constraints on the application of rules.

C11. There are no generalizations over sets of rules.

C11a. There are no existential generalizations, or *rule schemata*.

C11b. There are no implicational generalizations, or *metarules*.

D. Substantive content not assumed

D1. The stock of node labels from which those in any particular grammar are drawn is not assumed to be finite; new labels can be devised as needed in any particular grammar.

D2. The stock of node labels is not assumed to be universally specified.

D3. Node labels are not assumed to have any intrinsic content, any 'meaning' outside the formal system of grammar.

D4. The list of possible rules is not assumed to be finite; new rules can be formulated as needed in any particular grammar.

D5. The list of possible rules is not assumed to be subject to any universal constraints on form; in principle, any configuration of symbols meeting the conditions in C is a rule.

D6. The list of possible rules is not assumed to be subject to any universal constraints on substance, having to do with the 'meaning' of rules outside the formal system of grammar.

E. Components of grammar

E1. Lexicon

E1a. The lexicon is a finite collection of *lexical entries*.

E1b. Each lexical entry is a quintuple of representations, together with an index serving as a name for the lexical entry.

E1c. The five types of representations are: syntactic, morphological, phonological, semantic, pragmatic/stylistic. Each is germane to the operation of a component of grammar other than the lexicon itself.

E1d. These representations are structured constructs of properties.

E1e. The lexicon articulates with syntax via, and only via, a finite set of rules deriving terminal structures from preterminal structures.

E1f. Each such rule is of the form L ---> I,

E1f1. where L is a lexical label,

E1f2. and where I is the index of a single lexical item,

E1f3. whose syntactic properties do not make it distinct from L.

E1g. There are no useless rules of this sort:

E1g1. There is at least one such rule for each lexical label.

E1g2. There is at least one such rule for each index.

E1h. There are no generalizations (*lexical redundancy rules*) relating properties of lexical items to one another.

E2. Morphology

E2a. The morphological representations of lexical items are labeled tree structures, formally of the same type as those presented in A and B above, with the following exceptions (cf. B4 and B5):

E2a1. The set of morphological labels is disjoint from the set of phrasal labels of syntax.

E2a2. The set of morphological labels can be exhaustively partitioned into two (nonempty and disjoint) subsets, of *constituent* and *morpheme* labels.

E2a3. The lexical labels of syntax are all constituent labels of morphology.

E2a4. Every lexical label of syntax is an initial symbol of morphology, and no other labels are.

E2b. There is a finite set of *word formation rules*, whose function as a set is to describe the morphological representations in a potential lexicon.

E2c. These rules are formally entirely parallel to the phrase structure rules of syntax as presented in C above, with constituent labels and morpheme labels playing the roles in morphology that phrasal labels and lexical labels, respectively, play in syntax.

E2d. The provisos of D above apply to the labels and rules of morphology.

E3. Semantics

E3a. Semantics provides semantic interpretations for terminal structures.

E3b. The interpretation for a node with a lexical label, dominating a node labeled with an index, is the semantic representation in the lexical entry having that index.

E3c. Semantics is 'rule-to-rule': For each syntactic rule with mother label M and daughter labels $D_1 \ldots D_n$, there is a principle by which the semantic interpretations of the D_i can be regularly related to the semantic interpretation of M. By this means, interpretations are assigned to phrasal nodes in a terminal structure.

E3d. Other principles express equivalences and implicational relationships between semantic interpretations.

E3e. Others place conditions of well-formedness on semantic interpretations.

E3f. However, semantic principles are *positive* only; there are no explicitly negative semantic principles, or *semantic filters*.

E4. Pragmatics/Stylistics

E4a. Stylistic, registral, discourse-structural, etc. information may be associated with the constructs provided in terminal structures, that is, with particular lexical items, with lexical categories, with phrasal categories, and with constituent structure configurations.

E4b. Such information is associated with syntactic phenomena only in this fashion.

E5. Prosody/phonology

E5a. Terminal structures provide the domains within which rules of sentence prosody apply, and supply all the nonphonological information available for conditioning or constraining such rules.

E5b. The terminal structures of syntax and those of morphology, taken together, provide the domains within which phonological rules apply, and supply all the nonphonological information available for conditioning or constraining such rules.

E6. Syntax is autonomous

Syntactic rules are not subject to any conditions or constraints referring to

E6a. the morphological composition of words,

E6b. the prosodic features of phrases,

E6c. the phonological features of words or phrases,

E6d. the semantics of the construct or of the constituents involved, or

E6e. the pragmatic/stylistic effects of this construct or these constituents.

3. A checklist of presumed inadequacies of PV syntax. I now turn to a compact presentation of all types of defects that I know of that might be laid at the door of PV syntax.

A. Inabilities in principle

A1. The basic black, or total, inadequacy. The *stringset* associated with any particular PV syntax - the set of strings of words appearing at the bottom of terminal structures - is context-free (CF). There are, it has been claimed, natural languages that do not have CF stringsets; if so, PV syntax is fundamentally inadequate. Pullum and Gazdar (1982) disputed all such claims in the literature, but Pullum (1984a, b) concedes that there might indeed be natural languages with non-CF stringsets.

There are several possible complexities here. The constructions at issue might be very rare across languages (like the particular sort of object incorporation found in Mohawk) or they might be of very marginal status in their languages (like the respectively construction in English). Faced with valid but rare phenomena, one might claim, unsatisfactorily, that natural languages were 'mostly' CF. Faced with examples of marginal status, one might claim either that (a) the troublesome constructions do not belong to the 'core syntax' of the language in question (a risky proposal, in the absence of supportable guidelines for distinguishing the center of language from its periphery), or that (b) the conditions on the troublesome constructions are to be described not by syntax, but rather by principles of semantics, or by extragrammatical principles (Pullum and Gazdar dispose of several apparently damaging cases in this fashion). It is also possible that PV syntax can be amended so that the resultant theory countenances only a small extension of the set of CF languages; this is the tack taken in such works as Joshi (1984) and Pollard (1984).

A2. Sublethal but very unpleasant. Here I refer to cases in which, it is claimed, a CF syntax can describe the stringset of a language...but there is no CF syntax that correctly describes the *constituent structures* of the language. If this is so, then the information supplied by any CF syntax for the language is deficient; cf. section D below. Worse, the structures that are supplied will be inappropriate for the functioning of other components of grammar (semantics, lexicon, and phonology, in particular); cf. section E below.

B. Gigantism

Beyond the presumed inadequacies-in-principle of the preceding section, the standard criticisms of PV syntax amount to saying that it is clunky, inelegant, unilluminating. There are four prongs to this criticism, taken up in this section and the three following it:

--Gigantism;

--Irretrievable loss of generalizations;

--Missing information;

--Insufficient support of other components.

The criticism of gigantism turns on the fact that a moderately adequate PV syntax for a natural language has an astounding number of necessarily distinct syntactic categories, and an even more astounding number of phrase structure rules, that is to say, a quite astounding diversity of (surface) constituent structures.

To see how the number of categories and rules expands as a PV syntax for a language moves towards adequacy, one has only to consider the problem of describing English NPs. To account for the mutual restrictions between various determiner types (including the absent determiner in museums and rice) and singular count, plural count, or (singular) mass

nouns, the place of the single category Det must be filled by several (six, in fact) categories, and the place of the single category N must be filled by at least three categories, NSg, NPl, and NM. Moreover, the category Nom (functioning in the rules NP ---> Det Nom, Nom ---> AP N) must be divided in the same way as N, into NomSg, NomPl, and NomM. And of course the rules referring to NP, Nom, N, and Det must be replaced by rules referring to NPSg, NomM, NPl, and the like.

In this replacement categories like Nom simply disappear. Labels like 'NomPl' are indivisible entities. The categories replacing Nom could just as well have been 307, 578, and 722; names like 'NomPl' are chosen for their mnemonic value, but 'NomPl' is just as distinct from 'NomSg' as '307' is from '578', or 'Zebra' from '#@$2'. The categories Det, Nom, and N have been *atomized* out of existence, in favor of a set of new categories.

Our original rules have also been atomized. Where there were once the two rules NP ---> Det Nom and Nom ---> AP N, there are now many.

Clearly, the atomization of categories and rules will only increase when we consider further facts about the internal composition of NPs in English, not to mention further facts about the distribution of noun phrases in sentences (obviously, NPSg and NPPl will have to be distinguished, if subject-verb agreement is to be described, and that is only the beginning of our troubles).

The problem with atomization is not just that the whole syntactic apparatus becomes unwieldy. The real problem is that it is mind-boggling to suppose that the description has any psychological reality whatsoever - that is, that speakers of English, or any other language, somehow assign a constituent to one out of an astronomical (or larger) number of categories, or that these speakers have provided any sort of indvidual internal representations for a set of rules whose number may well exceed the number of neurons in the brain. (Note the diminishing returns of CF-ness here: CF syntaxes are vastly more learnable, in the technical sense, than more powerful syntaxes, but as the number of categories and rules explodes towards astonomical size, or even beyond, it becomes less and less likely that human beings could experience the necessary inputs in the time available to them, no matter how quick they might be at inducing CF syntaxes.)

Moreover, as we shall see in the next sections, as the number of categories and rules increases, the amount of information directly represented in constituent structures decreases; as a result, it is more and more difficult to see the set of rules as a repository of generalizations about the syntax of the language; and the syntactic description is less and less satisfactory in its interfaces with other components of the grammar. The four species of clunkiness shamble along hand in hand.

C. Loss of Generalization

C1. Gigantism and lost generalizations. The loss-of-generalization criticism is the other side of the gigantism coin. If you've got a quarter of a million rules, it's very hard to find any generalizations in them.

Much worse, when categories and rules are atomized there is no direct way to refer to categories or construction types; the only way to manage such reference is by lists, and nothing in the syntax will associate one such list with another. To see the sort of problem that arises, consider the rules needed to achieve the effect of subject-verb agreement in English declarative sentences, then to describe yes-no questions, including their agreement possibilities. Each declarative rule has an interrogative counterpart, but this generalization will nowhere be made (indeed, the system would be simpler if some or all of the counterpart rules were absent), and there will be no entity that we could identify as 'declarative sentence' or 'yes-no question', just as there will be no N entity or S entity.

C2. Missing information and lost generalizations. The loss-of-generalization criticism also lies behind all the criticisms having to do with information that is not represented in constituent structures: if the information isn't there, there are no generalizations about it. Insofar as there are syntactic generalizations to be made in terms of Subject, Direct Object, and Indirect Object relations to verbs...or in terms of coconstituency between nonadjacent elements...or in terms of transformational relatedness...or whatever from the list following in section D: If constituent structures do not contain this information, the generalizations cannot be stated.

Even when the information is available, if it is not *directly* represented, then the appropriate generalizations cannot be stated. They will instead be dissociated into a swarm of instances. If the pairs of constituents that stand in a determining-determined relationship with respect to agreement have to be listed, for instance, then any generalization dissolves. If lexical and phrasal categories are atomized, then generalizations explode into lists of subcategories.

C3. Types of missing generalizations. There are two different sorts of missing generalizations. The first is an *existential* generalization over rules, a statement that anything of a specified form is a rule of the language. The second is an *implicational* generalization over rules, a statement that predicts the existence of one set of rules on the basis of the existence of another. Stating either sort of generalization will require that we treat category names as complex and divisible entities.

D. Missing information

Claims that particular types of information are missing - that is, not represented - in constituent structures vary in their import. The less potent variety involves types of information that could be reconstructed from what is present in constituent structures, but are not

directly represented there. The more potent variety involves types of information that are, it seems, in principle not reconstructible from the information in constituent structures. The potency of any individual case is, alas, not always easy to determine. The claim that some sort of information is 'missing' presumes, of course, that it *ought* to be present in, or at least reconstructible from, an adequate syntactic description. This is a claim about the goals of syntactic theory. One response to a missing-information criticism of PV syntax is simply to maintain that there is no reason to suppose that a syntactic description must register head/modifier relations, systematic paraphrase relationships, anaphoric linkages, or whatever the type of information at issue is. So far as I can see, there are two variants of this response: (a) the respondent proposes that the information in question plays no role whatsoever in grammatical description, that it figures not at all in generalizations about linguistic form; (b) the respondent proposes that the information is indeed relevant to the workings of a whole grammar, but maintains that the information plays no role in syntax, only in some other component or components.

There are then four possible responses to the observation that PV syntax lacks some type of information:

--Totally Irrelevant: The information is not linguistically relevant.

--Syntactically Irrelevant: The information is linguistically, but not syntactically, relevant.

--Syntactically Relevant But Obscure: The information is syntactically relevant; it is not directly represented in constituent structures but can be derived from them.

--Syntactically Relevant But Absent: The information is syntactically relevant, but cannot be reconstructed from constituent structures. The inadequacies of PV syntax that seem to involve syntactically relevant, but obscure or absent, information, are the darlings of the literature on transformational grammar. All of the standard 'arguments for deep structures' and 'arguments for transformations' concern missing information. I now enumerate these.

D1. The difference between the 'slot' filled by a constituent (its 'external syntax', which concerns its cooccurrence and alternation with other constituents and its ordering within larger constructs) and the 'fillers' of the constituent (its 'internal syntax') is not generally represented. (Note that in some instances this difference can be coded in constituent structure by assigning a constituent to two different categories, as when penguins can't fly in I know penguins can't fly is assigned to S, on the basis of its internal syntax, and also to NP, on the basis of its external syntax.)

D2. Specific types of relations between coconstituents are not (directly) represented: what PV syntax tells us about two constituents is only that they are sisters; PV syntax does not represent any dependency of

one on the other, much less any specific type of dependency.

D2a. Head/modifier relations are not represented; despite the mnemonic similarity in labels, nothing says that N is the head of NP, V of VP, etc.

D2b. Consequently, the distinction between endocentric and exocentric constructions is not represented.

D2c. Relations between syntactic operators and their operands are not systematically represented; in particular, grammatical relations between verbs and their various NP arguments are not (fully) represented - neither the fact that some NP is related to one V and not to another, nor the particular relation it bears (Subject, Direct Object, etc.).

D2c1. In particular, the fact that one NP can bear several different relations with respect to different Vs (as in Geoff intended to be hard to persuade) is not represented.

D2c2. Similarly, the fact that, for different purposes, a single NP can bear different relations with respect to a single V (as penguins does with respect to are in There are penguins on the porch) is not represented.

D2d. Relations of selection among coconstituents are not represented.

D2e. Relations of subcategorization among coconstituents are not represented.

D3. The topic constituent of a sentence is not distinguished, nor its focus constituent(s).

D4. Dependencies between noncoconstituents are not represented. There are at least two important subtypes (beyond the dependencies already mentioned, all of which can involve noncoconstituents as well as coconstituents).

D4a. Redundant marks.

D4a1. The relationship between the determining element and the determined element in grammatical agreement is not represented.

D4a2. The relationship between an element in one part of a structure and its repetition in another part is not represented.

D4a3. The ('agreement') relationship between an anaphor and its antecedent is not represented.

D4a4. The relationship between a governing and a governed element is not represented.

D4a5. The fact that some other marks are predictable from structural configurations (as is the Poss attached to an English NP that is a Det) is not represented.

D4b. Effective coconstituency of elements that are not actually coconstituents. These elements might happen to be adjacent to one another (as the man and who are in the man who discovered Poughkeepsie), but usually they are *discontinuous*. At least three dimensions of syntactic discontinuities seem to be important.

D4b1. Whether the discontinuities are short-range (whether the intervening material is specifiable as a finite sequence of constituent types), or whether they are long-range (whether the intervening material is a stretch that cannot be specified by a finite sequence of constituent types - that is, whether the discontinuity covers an 'essential variable').

D4b2. Whether the discontinuities are nested or crossing.

D4b3. Whether the discontinuity involves a single pair of elements, a multiple but bounded number of pairs, or an unbounded number of pairs

D5. Zeros are not represented. There are at least four subtypes. The names below are chosen with some desperation, since different writers use entirely different terms and symbols, or use the same terms and symbols in novel ways.

D5a. Gaps: a gap is a missing constituent corresponding to a *filler* constituent occurring elsewhere in the sentence; the sentence is interpreted as if the filler occurred in place of the gap. (Example: Who do you think ∅ murdered Cock Robin?)

D5b. Zero anaphors: a zero anaphor is a missing constituent which must be understood as coreferential with some constituent occurring elsewhere in the sentence (or possibly beyond). (Example: Helga knew ∅ to go.)

D5c. Elliptic zeros: an elliptic zero is a missing constituent which is neither a gap nor a zero anaphor; elliptic zeros are interpreted entirely from context. (Example: the subject and object ellipses in ∅ Finally finished ∅ last week.)

D5d. Absent words: an absent word stands in alternation (contrast or variation) with one or more present words. (Examples: the absent indefinite determiner in penguins, in contrast with the definite determiner the in the penguins; the absent complementizer in I know penguins can't fly, in variation with the complementizer that.)

D6. In order to state a variety of conditions on the occurrence of constituents in larger constructs ('horizontal' conditions, relating constituents within a construct; and 'vertical' conditions, relating a string of constituents to the context in which their construct occurs),

both lexical and phrasal categories must be atomized, broken into a number of distinct categories. Thus, V is atomized into Vint, Vtr, Vdtr, etc. when we state subcategorization restrictions; and S is atomized into Sroot and Semb when we distinguish between constructions that occur only in root sentences and those that occur only in embedded sentences. At least two sorts of information are consequently not represented in constituent structures specified by an adequate CF syntax:

D6a. The general lexical categories N, V, A, etc., and the general phrasal categories NP, VP, S, etc. are not represented. Recall the discussion of missing generalizations above.

D6b. Nothing corresponding to a general 'construction type' (e.g., topicalized sentence, regardless of which NP is topicalized, whether the sentence is embedded or not, whether it is active or passive, etc.) is represented.

D7. Harris' *transformational relatedness* is not represented.

D7a. No account is given of distinct, but systematically related, structures that stand in a paraphrase relationship (e.g., the English active and passive constructions).

D7b. Including the special case of alternative word orders, both of coconstituents and of noncoconstituents.

D7c. Nor is any account given of systematic relationships between structures where there is a corresponding relationship between the semantic/pragmatic content of these structures (e.g., the English declarative-interrogative distinction in sentence types).

D8. No differentiation is made between identical constituent structures with different relationships among their parts. There are two standard subtypes.

D8a. *Faux amis*: The structures are identical, but different classes of lexical items occur in them (e.g., the celebrated <u>easy/eager to please</u> cases).

D8b. Ambiguous sentences whose ambiguity can be traced neither to a lexical ambiguity nor to an ambiguity in constituent structure (the classical 'transformational' ambiguities): The structures are identical down to lexical items.

D9. The relationship between interruptive material and its surroundings is not adequately represented.

D9a. The structural (and phonological) separation of parenthetical items (including, at least, whole-sentence parentheses, adverbial parentheses, appositives, and vocatives) from surrounding material is not represented.

D9b. Nor is the similar separation of expletives and (filled or unfilled) hesitations from surrounding material.

E. Other Components

The problems in articulating a CF syntax with other components of a grammar are apparently manifold. One sort of difficulty is loss of generalization, in particular loss of generalizations in other components directly following on loss of generalizations in syntax: If syntax doesn't make the generalizations or supply the pieces of information, then other components don't have them to refer to.

A distinct difficulty is that even well-supported systems of constituent structures (whether describable by a CFG or not) do not always provide the hierarchical structuring needed to state generalizations in other components. I begin with three cases of this sort, then turn to two situations where generalizations appear to cut across components. In the final set of cases, syntax hands on a loss of generalization to another component.

E1. Hierarchical structuring in syntax and in other components.

E1a. Phonology: Although in general the domains for phonological rules applying within, at the boundaries of, or between, phrases are those of syntactic constituent structure, there are not infrequent cases where phonological phrase domains differ from syntactic phrase divisions.

E1b. Semantics: Although in general principles of semantic interpretation can be stated as operating on the interpretations of coconstituents to yield an interpretation for the construct, there are not infrequent cases where this simple scheme of semantic composition seems to fail.

E1c. Lexicon: Although lexical items are normally single words, *idioms* are lexical items comprising a number of words, usually making a syntactic phrase. However, some idioms (like take umbrage at) involve a sequence of words that does not constitute a syntactic phrase.

E2. Generalizations across components.

E2a. Lexicon: The generalization about the lexicon appears to be that a lexical item can be either a word or a syntactic phrase.

E2b. Morphology: Although in general words can be treated as unanalyzed units in the statement of phrase structure rules, there are not infrequent cases where a morphological construction alternates with a syntactic construction, and others where a morphological construction is an obligatory concomitant of a syntactic construction, so that it appears that principles of alternation and cooccurrence must embrace both syntactic and morphological structure.

E3. Loss of generalization inherited by other components from syntax.

E3a. Lexicon/Morphology:

E3a1. Since lexical categories are atomized in PV syntax descriptions, different forms in the same inflectional paradigm belong to different categories, and there is no automatic relationship between them in the lexicon. Horse and horses end up belonging to different categories, and the existence of one has nothing to do with the existence of the other.

E3a2. Since PV syntax does not permit the expression of metarule relationships, it does not permit expression of the generalization that the lexical subcategories relevant in one set of rules can be identical to the lexical subcategories relevant in another. This relationship holds whenever rules introducing a lexical category are implicationally related to other rules introducing that lexical category.

E3b. Semantics: Since PV syntax does not permit the expression of metarule relationships, it does not permit expression of the generalization that when set A of rules is derivable by metarule from set B of rules, then the principles of semantic interpretation for A are also derivable from the principles of semantic interpretation for B.

E4. Loss of generalization in extragrammatical domains. Since 'construction types' are not represented in a PV syntax, there is nothing to associate stylistic, registral, discourse-structural, or other pragmatic markers with.

There are two distinct defects here. First, the atomization of syntactic categories makes it impossible to refer directly even to such abstract patterns as A + N (cited by Hudson 1980: 23 as a simple example of a syntactic construction that might figure as a 'linguistic item' referred to by a sociolinguistic statement). Instead, only a giant list of patterns is available.

Second, even if there were no atomization, we would still have no way of distinguishing faux amis or disentangling ambiguities that are neither lexical nor phrase-structural. An abstract pattern like V + AccNP + InfVP lumps together believe him to be a spy and persuade him to be a spy, despite the fact that these VPs differ in their stylistic and pragmatic properties as well as in their semantics. In transformational syntax the construction types are distinguished by virtue of the application of different transformations in their derivation, but the notion of construction type can be reconstructed in some non-transformational frameworks (as in GPSG, where a complete phrase structure rule includes a semantic interpretation principle as well as a list of mother and daughter categories) and is taken as a theoretical primitive in others (as in the Grammatical Construction Theory described by Lakoff 1984). But PV syntax has no place for such a notion.

F. Meta-Arguments

F1. One meta-argument against the adequacy of PV syntax notes that the typical example illustrating one presumed deficiency almost always illustrates several other presumed deficiencies; the phenomena are closely, complexly, linked with with one another.

The other side of this argument is that a good solution to one of the inadequacies solves many of the others as well - a fact that is usually viewed as an argument in favor of a syntax with a flavor other than plain vanilla.

But this particular sword cuts both ways. If the inadequacies of PV syntax are linked to one another, then supposing that an amendment to PV syntax solves one of the inadequacies (while preserving the mathematical character of the resultant system as CF) means that it will solve at least some others as well.

F2. A final meta-argument against PV syntax comes from *bootstrapping* - a situation in which assuming a transformational treatment of one phenomenon means that a transformational treatment of other phenomena is forced; Bach (1974: sec. 7.74) refers to these as 'arguments from other rules'. Again, the sword cuts both ways: if we don't have to assume the transformational analysis in one case, we are not obliged to assume such a treatment in the other.

4. Observations. How are we to respond to this catalogue of defects? In part, we have to question whether the defects are as claimed. Where they are, the response depends on the nature of the defect. Recall the discussion of responses to 'missing information' criticisms in part D of section 3 above. The defect might be linguistically irrelevant or syntactically irrelevant, in which case no alteration of PV syntax is called for.

Clearly, PV syntax must be amended in various ways to take account of the defects that are both genuine and relevant - but some of these amendments are more serious than others. A large part of the catalogue, in particular all the criticisms that follow from the atomization of categories, can be handled by giving up assumption B6 of section 2, 'Labels are unanalyzable wholes'. If assumption B1, 'The number of node labels is finite', is maintained, abandoning B6 is *innocuous*, in the sense that neither the describable stringsets nor the describable sets of constituent structures would be affected by such a change. On the other hand, if B6 is abandoned as well, say by permitting arbitrary integers to serve as indices within node labels, then the amendment is *substantial*, because the syntactic theory that results has all the power of unconstrained transformational syntax (despite looking like a simple variant of phrase structure syntax).[3]

Any one or more of the assumptions listed in section 2 might be abandoned or altered in an attempt to meet the case against PV syntax. In fact, for nearly every assumption - whether it concerns relations between nodes in trees, node labels, syntactic rules, the substantive content of

syntax, or the interfacing of syntax with other components of a grammar - there is someone, somewhere or other in the literature on syntax, who has proposed replacing that assumption. Some of these proposals have been made repeatedly; for instance, node labels are analyzed into features in almost all current theories, against assumption B6, and it is now standard to posit some principles reorganizing the terminal structures of syntax so as to provide the domains within which rules of sentence prosody and phonology apply, against the assumptions of E5. Others have been put forward only occasionally; most syntacticians assume the Single Mother Condition (A2), the No-Tangling Condition (A3), the Acyclic Condition (A4), the Unique Root Condition (A5), and the Total Ordering Condition (A6), though each of these has been challenged at least once.

Each particular package of amendments makes an alternative theory of syntax. Indeed, it is useful to view existing syntactic theories as such packages of amendments to the basic model of PV syntax. Generalized phrase structure grammar, for example, drops B6 (thus permitting the analysis of node labels), C5 (thus separating the rules that specify linear ordering from those licensing branchings), C8 (thus allowing zero expansions for phrasal category symbols), C11 (thus admitting generalizations over sets of rules), and E1h (thus allowing for lexical redundancy rules), replaces C6 by an assumption that rules receive the checking interpretation, and maintains the rule-to-rule semantics of assumption E3c while allowing for the possibility that distinct rules (with distinct semantic parts) can have identical syntactic parts.

In any individual instance it can be very difficult to decide just how substantial a package of amendments is. The amendments of GPSG are all innocuous, but assessing the effects of much greater deviations from PV syntax is no easy task. It can also be very difficult to choose a package that is worth pursuing, since the number of plausible combinations of amendments is enormous, greater even than the figure of thirty million that McCawley (1982) gives as an estimate of the number of viable combinations of positions on the generative semantics vs. interpretive semantics question.

NOTES

*This paper began as a presentation, 'Arguing for Remote Structures', at the 1979 Conference on the Nature of Syntactic Representations at Brown University. The paper did not appear in the volume that resulted from this conference (Jacobson and Pullum 1982). My notes for it were written up into a partial draft in 1982 while I was a fellow at the Center for Advanced Study in the Behavioral Sciences; they were revised for courses on generalized phrase structure grammar at the Ohio State University in the autumns of 1983 and 1984; a nearly full draft was prepared while I was a visiting senior research associate at the Center for the Study of Language and Information, Stanford University, during the summer of 1984; and this version was completed during leave from the Ohio State University in the winter quarter of 1985. I am indebted to Pauline Jacobson, for her invitation to take part in the Brown conference; to the Spencer Foundation, for its financial support while I was at CASBS; to the System

Development Foundation, for its financial support while I was at CSLI; and to the Ohio State University, for its financial support while I was on leave in 1981-82 and in 1985. I have benefited from the comments and criticisms of many friends and colleagues, most especially David Dowty, Michael Geis, Frederick Newmeyer, and Geoffrey Pullum.

[1]Formal language theory, as summarized in works like Hopcroft and Ullman (1979).

[2]Serious efforts have been made, however, by Pullum and Gazdar (1982) and by Sampson (1983), and on the mathematical side by Perrault (1984).

[3]As my remarks in part A of section 3 suggest, there are also degrees of substance, corresponding to greater or lesser extensions of the set of CF languages.

REFERENCES

BACH, Emmon. 1974. Syntactic theory. New York: Holt, Rinehart and Winston.

BLOCH, Bernard. 1946. Studies in colloquial Japanese II: Syntax. Lg. 22.200-48. Reprinted in Joos 1966: 154-85.

BRAME, Michael K. 1978. Base generated syntax. Seattle: Noit Amrofer.

BRESNAN, Joan (ed.). 1982. The mental representation of grammatical relations. Cambridge MA: MIT Press.

CHOMSKY, Noam. 1956. Three models for the description of language. IRE Trans. on Information Theory 2.3.113-24.

_____. 1957. Syntactic structures. The Hague: Mouton.

_____. 1982. Lectures on government and binding: The Pisa lectures. 2nd ed. Dordrecht: Foris.

DOWTY, David R., Lauri Karttunen, and Arnold M. Zwicky (eds.). 1985. Natural language parsing: Psycholinguistic, computational, and theoretical perspectives. Cambridge: Cambridge Univ. Press.

FOLEY, William A. and Robert D. Van Valin. 1984. Functional syntax and universal grammar. Cambridge: Camridge Univ. Press.

GAZDAR, Gerald. 1982. Phrase structure grammar. In Jacobson and Pullum 1982: 131-86.

HARMAN, Gilbert H. 1963. Generative grammars without transformational rules: A defense of phrase structure. Lg. 39.597-616.

HARRIS, Zellig S. 1946. From morpheme to utterance. Lg. 22.161-83. Reprinted in Joos 1966: 142-53.

HOPCROFT, John E. and Jeffrey D. Ullman. 1979. Introduction to automata theory, languages, and computation. Reading MA: Addison-Wesley.

HUDSON, Richard A. 1976. Arguments for a non-transformational grammar. Chicago: Univ. of Chicago Press.

_____. 1980. Sociolinguistics. Cambridge: Cambridge Univ. Press.

JACOBSON, Pauline and Geoffrey K. Pullum (eds.). 1982. On the nature of syntactic representation. Dordrecht: Reidel.

JOOS, Martin (ed.). 1966. Readings in linguistics I. 4th ed. Chicago: Univ. of Chicago Press.

JOSHI, Aravind. 1985. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structure descriptions? In Dowty et al. 1985: 206-50.

KAC, Michael B. 1978. Corepresentation of grammatical structure. Minneapolis: Univ. of Minnesota Press.
KING, Margaret (ed.). 1983. Parsing natural language. New York: Academic Press.
LAKOFF, George. 1984. THERE-Constructions: A case study in grammatical construction theory and prototype theory. Berkeley Cognitive Science Report No. 18. Berkeley CA: Institute of Human Learning, Univ. of Calif.
MATTHEWS, Peter H. 1981. Syntax. Cambridge: Cambridge Univ. Press.
MCCAWLEY, James D. 1982. Thirty million theories of grammar. Chicago: Univ. of Chicago Press.
NEWMEYER, Frederick J. 1980. Linguistic theory in America. New York: Academic Press.
PERLMUTTER, David M. (ed.). 1983. Studies in relational grammar 1. Chicago: Univ. of Chicago Press.
PERRAULT, C. Raymond. 1984. On the mathematical properties of linguistic theories. Computational Linguistics 10.3-4.165-76.
POLLARD, Carl J. 1984. Generalized Phrase Structure Grammars, Head Grammars, and natural language. Ph.D. dissertation, Stanford Univ.
POSTAL, Paul M. 1964. Constituent structure: A study of contemporary models of syntactic description. Supp. IJAL 30.
PULLUM, Geoffrey K. 1984a. On two recent attempts to show that English is not a CFL. Computational Linguistics 10.3-4.182-6.
_____. 1984b. Syntactic and semantic parsability. Proceedings of Coling84, 112-22. Association for Computational Linguistics.
_____ and Gerald Gazdar. 1982. Natural languages and context-free languages. L&P 4.4.471-504.
SAMPSON, Geoffrey R. 1983. Context free parsing and the adequacy of context-free grammars. In King 1983: 151-70.
WELLS, Rulon S. 1947. Immediate constituents. Lg. 23.81-117. Reprinted in Joos 1966: 186-207.