

# THE IRRELEVANCE OF TYPOLOGY FOR GRAMMATICAL THEORY

FREDERICK J. NEWMAYER

Many linguists believe that a parameter-setting model of grammar should capture typological generalizations. For example, a particular feature's cross-linguistic rarity might be 'registered' in a grammar that possesses that feature by means of a marked setting for the relevant parameter. I argue that such a view is in error. Grammars do not encode typological generalizations, either directly or indirectly. Put in a somewhat different way, 'universal grammar tells us what a possible language is, but not what a probable language is. The most robust typological generalizations –those arising from the seminal work of Joseph Greenberg– have an explanation based in language processing.

## 1. INTRODUCTION

I confess to have chosen a deliberately provocative title for this paper.<sup>1</sup> While it conveys a conclusion that I will stand by and defend –the conclusion that typology *is* irrelevant for grammatical theory– it may have connotations that I feel obligated to disassociate myself from immediately. The first is the possible implication that there is no need for grammatical theorists to undertake the intensive investigation of as many languages as possible. Indeed there is such a need, both for an appreciation of the range of processes that the languages of the world can manifest and for testing candidate universals that have been mooted on the examination of one or a small number of languages. After all, no investigation of a single language, no matter how thorough, could answer the question of whether overt *Wh*-Movement is subject to locality conditions if that language happened not to have overt *Wh*-Movement!

Second, I am not going to argue that typology lacks theoretical interest or importance. If a particular feature is manifest in 90% of languages in a properly constructed sample, then that is a fact in need of explanation. If feature A is correlated with feature B significantly greater than chance

would predict, then that too is a fact in need of explanation. But crucially, it does not follow that the explanations of such facts needs to reside internally to grammatical theory. That is, it might be wrong to derive the overwhelming preference for the feature that shows up in 90% of all languages from a principle, or set of interacting principles, within generative grammar. And similarly, the implicational relationship between A and B could fall out from the interaction of the grammatical module with others involved in the totality of language, rather than from the internal structure of grammatical theory itself.

And indeed, that is precisely what I will argue in this paper. It will be my conclusion that grammars do not encode typological generalizations, either directly or indirectly. Let us take, for example, some robust generalization, such as verb-final order tending to be associated with postpositions, rather than prepositions. I will argue that there is nothing in the theory of universal grammar (henceforth 'UG') in which this correlation is either stated directly or can be derived from its interacting principles. As a corollary to this claim, I will suggest that typological generalizations are not deducible from the inspection of the grammar of any individual language. Take Japanese, for example, which upholds the above mentioned correlation and German, which violates it. The grammar of neither language encodes, directly or indirectly, the information that the former language is typologically consistent and the latter inconsistent. Likewise, there is no grammatical provision of the information that the state of affairs represented by Japanese is relatively common cross-linguistically and that represented by German relatively rare.

It follows then that a grammar's fidelity to typological generalizations can play no part in its *evaluation*. Let us imagine two otherwise identical candidate grammars of German that differ only in how directly the typological mixedness of that language is expressed. All other things being equal, there is no reason, I will claim, to value the grammar with the more direct characterization of this state of affairs over the one that represents it less directly.

In pursuing such a line of argumentation, I will be going against quite the opposing trend in the community of generativist scholars. My sense is that typological generalizations have been increasingly regarded as relevant in the generativist community. The historical record, certainly, bears out such an idea. Most linguists would point to the publication of Joseph Greenberg's paper 'Some universals of language with special reference to the order of meaningful elements' (Greenberg, 1963) as marking the birth of modern typological studies. The first reference to this paper that I am aware of in the generative literature is a passage

from Chomsky's *Aspects of the theory of syntax* that can only be regarded as deprecatory.

Modern work has indeed shown a great diversity in the surface structure of languages. However, since the study of deep structure has not been its concern, it has not attempted to show a corresponding diversity of underlying structures, and, in fact, the evidence that has been accumulated in modern study of language does not appear to suggest anything of this sort. ... Insofar as attention is restricted to surface structures, the most that can be expected is the discovery of statistical tendencies, such as those presented by Greenberg (1963). (Chomsky, 1965: 118)

With the development of the principles-and-parameters approach of Chomsky (1981), however, the mainstream generative attitude to typological generalizations began to change. A central goal of syntactic theory now became to identify the various subsystems of grammar and to characterize the degree to which they may vary (be 'parameterized') from language to language. A consequence was thus to spur investigation of a wide variety of languages, particularly those with structures markedly different from some of the more familiar western ones. In this regard, it is instructive to observe Chomsky's changing attitude to Greenbergian typological work. In 1981, Chomsky offered what was perhaps his first favorable reference to this line of research:

Universals of the sort explored by Joseph Greenberg and others have obvious relevance to determining just which properties of the lexicon have to be learned in this manner in particular grammars –and to put it in other terms just how much has to be learned as grammar develops in the course of language acquisition. (Chomsky, 1981: 95)

By 1982 he was writing that 'Greenbergian universals ... are ultimately going to be very rich. ... They have all the difficulties that people know, they are "surfacy," they are statistical, and so on and so forth, but nevertheless they are very suggestive' (Chomsky, 1982: 111). And in 1986, they are 'important, ... yielding many generalizations that require explanation ...' (Chomsky, 1986: 21).

How, then, might typological generalizations be accounted for in the principles-and-parameters approach? Chomsky puts it succinctly:

Within the P&P approach the problems of typology and language variation arise in somewhat different form than before. Language differences and typology should be reducible to choice of values of parameters. A major

research problem is to determine just what these options are, and in what components of language they are to be found. (1995: 6)

My goal in this paper is not to argue against a parameter-setting model of grammar per se. Rather, I will question the presumed linkage between the structure of such a model and its ability to capture typological generalizations. I will claim that, as a point of fact, the model has failed to explain why such generalizations exist and, as a point of theory, it should not be expected to do so. In one pithy sentence, UG tells us what a *possible* human language is, but not what a *probable* human language is.

The paper is organized as follows. Section 2 outlines the 'Greenbergian correlations', the most robust cross-linguistic generalizations put forward in the literature. In §3, I outline why it has come to be expected among many generative linguists that UG might play a role in the explanation of these correlations and demonstrate that this expectation has not been fulfilled. Section 4 argues that the correlations have an extragrammatical explanation, and §5 is a brief conclusion.

## 2. THE GREENBERGIAN CORRELATIONS

The central presupposition underlying what follows is that there do indeed exist valid typological generalizations in need of explanation. Such is not self-evidently true. As I argue at length in Newmeyer (1998: ch. 6), it is by no means obvious that the cross-linguistic generalizations that can be gleaned from any sample, no matter how large, of presently-existing languages are robust enough to be regarded as brute facts in need of explanation. And worse, many such generalizations that have appeared (and are cited) in the literature are not even based on large samples. As a leading typologist has frankly admitted:

Yet the empirical basis for many typological claims is weak; generalizations are regularly made on untested convenience samples of *fewer than 50 languages (and usually considerably fewer)*. Even the most appealing work in syntactic typology, for example, Hopper and Thompson (1980) and Givón (1981), would be strengthened by a more systematic discussion of the database used in the research. This criticism by no means implies that the generalizations are false ones, or that no insight into the nature of language is to be drawn from the work; but it does mean that the reliability and the validity of their claims is compromised to some extent. (Tomlin, 1986: 17-18; emphasis added)

I do not believe, however, that there is any dispute that the most uncontroversially reliable typological generalizations are a subset of those that have grown out of the seminal Greenberg paper alluded to above. While the paper proposed several dozen typological universals, those that immediately attracted the greatest deal of attention and inaugurated the most extensive research program are the ones that correlate the basic order of subject, object, and verb with other grammatical features. Even though Greenberg worked with a convenience sample of only 30 languages, some of the correlations that he noted seemed too striking to be accidental. Consider, for example, the correlation between word order and adposition order. Greenberg's sample contained 6 languages with VSO order, all of which were prepositional; 13 SVO languages, which were overwhelmingly prepositional; and 11 SOV languages, all postpositional (see Table 1). Such correlations, it was widely agreed, could not be due to chance.

	VSO	SVO	SOV
Prep	6	10	0
Postp	0	3	11

*Correlations between word order and adposition order*

*(Greenberg, 1963)*

*Table 1*

The most exhaustive survey of typological correlations coming out of the Greenberg paper is Dryer (1992). Based on a study of 625 languages, Dryer found the statistically significant correlations of VO and OV order that are represented in Table 2.<sup>2</sup>

VO correlate	OV correlate
adposition - NP	NP - adposition
copula verb - predicate	predicate - copula verb
'want' - VP	VP - 'want'
tense/aspect auxiliary verb - VP	VP - tense/aspect auxiliary verb
negative auxiliary - VP	VP - negative auxiliary
complementizer - S	S - complementizer
question particle - S	S - question particle
adverbial subordinator - S	S - adverbial subordinator
article - N'	N' - article
plural word - N'	N' - plural word
noun - genitive	genitive - noun
noun - relative clause	relative clause - noun
adjective - standard of comparison	standard of comparison - adjective
verb - PP	PP - verb
verb - manner adverb	manner adverb - verb

*Correlation pairs reported in Dryer (1992)*  
*Table 2*

Interestingly, despite claims that had been advanced earlier that noun - adjective order is a VO correlate and adjective - noun order is an OV correlate (see Lehmann, 1973 and Vennemann, 1974), Dryer (1988) demonstrated that no such correlation exists.

In the remainder of this paper, I will regard the generalizations expressed in Table 2 as facts in need of explanation and refer to them as 'the Greenbergian correlations'.

### 3. TYPOLOGICAL GENERALIZATIONS AND GENERATIVE GRAMMAR

This section begins (§3.1) with a sketch of why it has come to be believed that typology is relevant to generative grammar. Sections §3.2 and §3.3 cast doubt on this belief by looking at properties of grammars and facts about first language acquisition respectively.

### 3.1. POSSIBLE AND PROBABLE LANGUAGES

The central goal of generative grammar from its inception has been to characterize the notion 'possible human language'. In an early formulation of this goal:

The theory thus constructed is a theory of linguistic universals... Specification of the form of grammars excludes certain infinite sets of sentences from consideration as possible natural languages. ... Procedures for evaluating grammars and determining structural descriptions impose strict conditions on the kinds of units that can be attributed to a natural language and the manner of their arrangement and interconnection. This general theory can therefore be regarded as a definition of the notion 'natural language'... (Chomsky, 1962: 536-537)

The vocabulary of theoretical primitives, conventions for formulating rules, etc. of the theory are therefore chosen, not on the basis of an appeal to 'simplicity' in the abstract, but rather with the view in mind of excluding from the very possibility of formulation any process outside of the definition of 'natural language'. For example, it would be just as simple, if not more so, for a language to form questions by regularly inverting the order of all the words in the corresponding declarative than by fronting some particular constituent of the declarative. UG, however, prohibits the former option by its failure to provide a mechanism for carrying out such an inversion operation. That is, the following rule type, while perhaps simple and elegant in the abstract, is not allowed by UG:

$$(1) W_1 - W_2 - W_3 - \dots - W_n \longrightarrow W_n - \dots - W_3 - W_2 - W_1$$

The question naturally arises, then, about the theoretical treatment of grammatical processes that are not fully excluded from UG, but rather are, in some pretheoretical sense 'unnatural', that is, unlikely to occur in the grammars of very many languages. The assumption has always been that there exists some theory-internal way of capturing their unnaturalness. A case of a grammatical process that straddles the border between literal impossibility and utter implausibility is discussed in Chomsky (1965: 42-45) and involves the English auxiliary. Aux may contain as its maximal expansion a tense morpheme, a Modal, a Perfect morpheme, and a Progressive morpheme. Among the ways that such a situation might logically manifest itself are the following. In (2a) the linear ordering

between elements is preserved, while in (2b) the relationship among them is 'cyclic':

(2) (a) Tense, Tense Modal, Tense *Perfect*, Tense *Progressive*, Tense Modal *Perfect*, Tense Modal *Progressive*, Tense *Perfect Progressive*, Tense Modal *Perfect Progressive*

(b) Tense Modal *Perfect Progressive*, Modal *Perfect Progressive* Tense, *Perfect Progressive* Tense Modal, *Progressive* Tense Modal *Perfect*, Tense *Perfect*, Modal *Progressive*

(2a) and (2b) contain an equivalent number of symbols and, from that point of view, are equally complex. However the linear regularity of (2a), not only represents the true situation for English, but is a commonplace one in the languages of the world. The cyclic regularity of (2b) is 'not characteristic of natural language' (Chomsky, 1965: 43). Hence the theoretical motivation for the parentheses notation, which allows the twenty symbols of (2a) to be collapsed to the four (to the right of the arrow) of (3):

(3) Aux → Tense (Modal) (*Perfect*) (*Progressive*)

No corresponding notational convention allows for the collapsing of (2b). But note that UG does not absolutely exclude the situation exemplified in (2b). That is, no constraint prohibits the following set of rules from being part of the grammar of English, or their analogs from being part of the grammar of some other language:

(4) Tense Modal *Perfect Progressive*  
 Modal *Perfect Progressive* Tense  
 Aux → *Perfect Progressive* Tense Modal  
*Progressive* Tense Modal *Perfect*  
 Tense *Perfect*  
 Modal *Progressive*

However, such a grammar must 'pay' for its typological unnaturalness by requiring a complicated set of rules, uncollapsible by any abbreviatory convention provided by UG.

As we have seen, then, typological considerations have acted as a guide to theory construction since the earliest days of transformational generative grammar. Nowhere is this so true as in phonology. In the



earlier chapters of *The Sound Pattern of English* (Chomsky and Halle 1968), the naturalness of a phonological rule was considered essentially as the inverse of the number of distinctive feature specifications needed to formulate it. That is, the design of UG provided an evaluation metric such that natural processes (say, those embodying natural classes of elements) were 'easier' to state, and hence valued more highly, than unnatural ones. The problem, addressed in chapter 9 of that book, was that feature counting alone did not suffice to distinguish typologically natural processes from typologically unnatural ones. For example, all other things being equal, no more feature specifications are required for a language to unround all rounded back vowels than to unround all rounded front vowels. Yet, the former process is extremely rare cross-linguistically, while the latter relatively common. Hence Chomsky and Halle introduced a set of *marking conventions* into the theory, which tied naturalness to evaluation. The natural unrounding process would be cost free in terms of the metric, while the unnatural one would be counted. These conventions were further developed in Kean (1975).

As early as the mid 1960s, the above adaptation of Praguean markedness theory was being applied to syntactic analysis as well as phonological. For example, Bach proposed to handle the Greenbergian word order correlations by means of marking conventions internal to the grammar. He wrote that since OV languages like Japanese typically 'have preposed desentential nominal modifiers ... we do not have to state the rule shifting 'REL' to a position before the noun for Japanese separately but can state in our general theory that this rule is predictable from the basic order of Japanese sentences' (1965: 10-11). In other words, the fewer marked processes a language turns out to have, the fewer descriptive statements are necessary in the formulation of its grammar. Bach's criteria for determining whether a process is marked or not appears to have been determined largely, if not wholly, on typological distribution: more common crosslinguistically was considered to be less marked grammatically.

Analogously, Hale (1976) called attention to the typologically rare phenomenon of defining the notion 'subject' in terms of the language's case system (found in Warlpiri and other 'nonconfigurational' languages) instead of configurationally. He suggested that marking conventions single out this phenomenon as highly marked, presumably by requiring more language-particular descriptive machinery than is necessary to characterize configurationally-assigned grammatical relations.

Now, as any generative theoretician would freely acknowledge, typological distribution cannot serve *in and of itself* as a factor determining

the principles of UG and the relative markedness of rules and principles provided by UG. Typological generalizations belong to the domain of E-language, that is, aspects of language 'understood independently of the properties of the mind/brain' (Chomsky, 1986: 20). Our minds/brains, after all, have no clue as to the typological status of any aspect of any element of our mental grammars. The relationship between typological generalizations and I-language, 'some element of the mind of the person who knows the language' (Chomsky, 1986: 22), is necessarily quite indirect.

Nevertheless, there has been a guiding assumption that there is no significant gap between the notions '*typologically* significant generalization' and '*linguistically* significant generalization'. That is, generative grammarians have generally taken it for granted that if investigation of the grammatical properties of a reasonably large set of languages leads to the discovery of a pervasive and profound structural pattern in those languages, then there is probably something mentally 'preferable' about that pattern, and this mental preference should be reflected by UG being organized to 'favor' that pattern.

A case in point is the X-bar schema, proposed in Chomsky (1970) and given its greatest development in Jackendoff (1977). It soon became apparent that the generalizations expressed by X-bar theory (which was initially formulated on the basis of English data) were borne out typologically: languages tend to favor a consistent ordering of their heads and complements. Consider Table 2 above. Many of the correlations pairs are head-complement correlations of one sort or another, e.g. those between adpositions and their objects, complementizers and their associated sentences, verbs and their PP complements, and so on. Hence 'inconsistent grammars' that violate X-bar generalizations require special 'marked' formulations.

The treatment of Chinese phrase structure in Huang (1994) provides an illustrative example. Oversimplifying a bit, Chinese is consistently head-final, except for the rule expanding X' to X<sup>0</sup>. If the head is verbal (i.e. a verb or a preposition), then the head precedes the complement. Huang captured this situation by a phrase-structure schema that complicates the X-bar schema somewhat:

- (5) (a)  $XP \rightarrow YP X'$   
 (b)  $X' \rightarrow YP X'$   
 (c)  $X' \rightarrow (c')X^0 YP$  iff  $X = [+v]$   
 (c'')  $YP X^0$  otherwise

So, deviation from typological naturalness is reflected by a more complex grammar.

A parallel treatment of regularities and irregularities of head-complement order has been provided within the framework of interacting parameterized UG principles. Travis (1989: 271) calls attention to eight possible orderings of the verb, direct object NP, complement PP ('PP<sub>1</sub>'), and adjunct PP ('PP<sub>2</sub>')

- (6) WORD ORDERS:
- (a) PP<sub>2</sub> PP<sub>1</sub> NP V
  - (b) PP<sub>2</sub> PP<sub>1</sub> V NP
  - (c) PP<sub>2</sub> NP V PP<sub>1</sub>
  - (d) PP<sub>2</sub> V NP PP<sub>1</sub>
  - (e) PP<sub>1</sub> NP V PP<sub>2</sub>
  - (f) PP<sub>1</sub> V NP PP<sub>2</sub>
  - (g) NP V PP<sub>1</sub> PP<sub>2</sub>
  - (h) V NP PP<sub>1</sub> PP<sub>2</sub>

Travis proposed three separate parameters to allow for the possibilities in (6a-h), which she designated 'headedness', 'direction of theta-role assignment', and 'direction of case assignment' (see also Koopman, 1984 for a very similar proposal). If these three parameters were independent, then all eight orderings would be predicted to exist, by virtue of the combinations of settings illustrated in Table 3.<sup>3</sup>

	HEADEDNESS	THETA	CASE	LANGUAGE
a.	final	left	left	Japanese
b.	final	left	right	Chinese (future)
c.	final	right	left	*
d.	final	right	right	Chinese (present)
e.	initial	left	left	Kpelle (past)
f.	initial	left	right	*
g.	initial	right	left	Kpelle (present)
h.	initial	right	right	English

*Combinations of the headedness, direction of theta-role assignment, and direction of case assignment parameters (Travis, 1989)*

*Table 3*

However, no language manifesting (c) and (f) appears to exist, nor is there evidence that such a language ever existed. Travis therefore proposed implicational relations among these 3 parameters whose effect is not only to predict (c) and (f) impossible, but to characterize the unmarked 'expected' situations to be (a) and (h), where the three parameters conspire to keep *all* complements on the same side of the head.<sup>4</sup>

Many linguists have felt there to be a degree of circularity in the claim that some feature of grammar (a violation of X-bar, a special parameter setting) is more 'marked' than another. The problem is that markedness is concluded from cross-linguistic rarity, but then cross-linguistic rarity is explained in terms of markedness. With this problem in mind, David Lightfoot has suggested that claims of markedness require independent motivation:

For specific proposals concerning marked values to entail testable claims, these claims will have to hold in an 'external' domain, a domain other than that of the distribution of morphemes or grammatical well-formedness. Claims to explanatory adequacy will have to be grounded in such domains. Natural candidates for such a domain wherein markedness proposals make empirically testable claims are language change and acquisition. (Lightfoot, 1979: 76-77)

What is the empirically testable claim about language acquisition that follows from a markedness proposal? The null hypothesis is that '[t]he "unmarked case" can be understood as the child's initial hypothesis about language (in advance of any data) ...' (Williams, 1981: 8). In terms of grammatical development, '[w]e would expect the order of appearance of structures in language acquisition to reflect the structure of markedness in some respects...' (Chomsky, 1981: 9).

If the order of acquisition is a function of the markedness of the construct being acquired and claims of markedness are based on part on cross-linguistic frequency, then we would naturally expect that early-acquired constructs would be cross-linguistically frequent. And indeed, two prominent specialists in the field of language acquisition have drawn just such a conclusion:

[I]n determining which notions are encoded in a language's morphology, the child is faced with a formidable search problem ... [B]y imposing a weighting on the child's hypotheses, one could account for the large disparities in the prevalence of various grammatical encodings in the world's languages, and in the speed of acquisition of various encodings by children. (Pinker, 1984: 168-171)

One intriguing possibility is that the relative accessibility for children of alternative schemes for partitioning meaning in a given conceptual domain is correlated with the *frequency with which these schemes are instantiated in the languages of the world*. ... It is plausible that relative frequency is correlated with 'ease' or 'naturalness' for the human mind. (Bowerman, 1985: 1306).

So, we have arrived at the following hypotheses linking typological generalizations to aspects of I-language:

- (7) (a) Cross-linguistically frequent properties of language are reflected by correspondingly simple (unmarked) properties of grammars.
- (b) Cross-linguistically frequent properties of language are acquired early by the child.
- (c) Cross-linguistically frequent properties of language are diachronically stable.

If (7a-c) were correct, then typology would indeed be relevant to grammatical theory in two complementary ways. First, we could appeal to grammatical theory to explain the typological distribution of any particular feature of language. Second, the typological distribution of a feature of language would serve as a reliable heuristic for the correct grammatical analysis of that feature. However, as we will see in the following sections, (7a-c) are not correct.<sup>5</sup>

### 3.2. THE FAILURE OF UG TO CAST LIGHT ON TYPOLOGICAL PATTERNING

In this section I will question the assumption driving the marriage of grammatical theory and language typology, namely that optimal grammars necessarily reveal profound cross-linguistic patterns of the distribution of grammatical elements. I will take several much investigated grammatical phenomena and demonstrate that popular analyses of these phenomena fail to shed any light on typology. Section 3.2.1 argues that one robust set of typological correlations does not follow from anything intrinsic to theory; in §3.2.2 evidence is presented that the Greenbergian correlations hold better at surface, rather than at deep, levels of grammar; section 3.2.3 demonstrates that in an important case where generative work *has* led to the formulation of a profound typological generalization, that generalization is incorrect; and §3.2.4 shows that for one well-studied phenomenon, the maximally 'simple' analysis makes the wrong typological predictions.

### 3.2.1. ROBUST TYPOLOGICAL GENERALIZATIONS DO NOT FALL OUT FROM UG PRINCIPLES: THE CASE OF WORD ORDER AND WH-MOVEMENT

It has long been known that verb-final languages are much less likely to exhibit *Wh*-Movement than VO languages, but much more likely to have sentence-final question particles. Table 4 from Dryer (1991: 455-466) provides the data supporting such an idea:<sup>6</sup>

	V-final	SVO	V-initial
<i>Wh-in situ</i>	71	42	16
Final Q particles	73	30	13

*Proportion of languages with Wh-in situ and final question particles, by word order type (Dryer, 1991)*

*Table 4*

The root of the typological correlation between verb finality, lack of *Wh*-Movement, and final question particles has been on the generative research agenda for almost three decades. Baker (1970) proposed a universal rule of Question Movement, in which a *wh*-type element moves to the left to replace an abstract question morpheme 'Q'. Bresnan (1970) identified 'Q' with the category 'COMP' (i.e. 'Complementizer') and suggested that only languages with clause-initial COMP permit a COMP-substitution transformation. Thus her proposal went beyond Baker's in addressing the question of why long-distance movements are *in general* excluded in OV languages. And Bach (1970b), working with a slightly different set of assumptions from Bresnan, derived the prediction that OV languages have no *Wh*-Movement (though they may have other movements) from the hypotheses that the element moved by this rule is attracted to a governing verb and that UG allows only leftward movement rules to be unbounded.

More recently, Fukui (1986) has attributed the impossibility of overt *Wh*-Movement to the lack of a Specifier for COMP, thereby denying the *wh*-element a landing site and Kim (1990) has argued that the analogs of *wh*-elements in movementless languages are actually quantifiers and therefore undergo Quantifier Raising at LF, rather than overt *Wh*-Movement. And still more recently, Cheng (1991/1997) has addressed the correlation between the presence of final question particles and the

impossibility of syntactic *Wh*-Movement. Based on the economy principles provided by the Minimalist Program, she proposes a theory of 'clausal typing', whereby a language must choose one of these two methods of 'typing' questions.

The problem with all of these accounts, as far as explaining how the typological generalizations in question are concerned, is that they are highly stipulative. Why should unbounded movements be only to the left? Why should a set of languages that are typologically similar in respects that do not involve COMP lack a specifier for that category? And Cheng in particular fails to address the third element of the typological correlation: the fact that final Q-particle/*wh* in-situ languages are overwhelmingly OV and movement languages are overwhelmingly VO. Furthermore, she doesn't explain (or even mention) the fact that languages that have *Wh*-Movement for questions almost always have it for other processes (relatives, clefts, and so on), a fact that would seem not to follow from her clausal typing hypothesis. The accounts just mentioned thus take the correlations essentially as primitives of theory, rather than as consequences that fall out from independently-motivated principles.

The Minimalist Program (MP) has made the correlations even more difficult to explain nonstipulatively, as far as I can tell. Basic clause structure is assumed to be universal, with differences in surface order due to differences in the strength of particular features. Now, the problem is to explain why a weak *wh*-feature on C (preventing overt *Wh*-Movement) would correlate with whatever feature or combination of features are responsible for surface SOV order. None come to mind. The problem of the typological associates of *Wh*-Movement is particularly difficult to explain vis-à-vis surface VSO languages. As Table 4 shows, verb-initial languages are far more likely to have *Wh*-Movement than SVO languages (not to mention verb-final languages). Why should this be? Since Emonds (1980), the predominant position has been that such languages 'start out' as verb-medial, but have a raising of the verb (for a recent account, see McCloskey 1996). Let us say, following the account presented in Marantz (1995: 372-373), that such movement is driven by strong V-features of T and/or AGR in the context of weak N-features for these functional heads. The question then is why this constellation of features would correlate *even more strongly* with strong *wh*-features on C (thereby guaranteeing overt *Wh*-Movement) than with the alternative feature strengths associated with T and AGR that 'preserve' SVO order. I cannot imagine how such a correlation might be derived, given any mechanisms accepted as intrinsic to generativist theory.

### 3.2.2. D-STRUCTURE IS NOT A GOOD PREDICTOR OF TYPOLOGICAL GENERALIZATIONS: THE HEAD PARAMETER AND RELATED ISSUES

As we saw in §3.1, the big typological success story associated with generative grammar is based on its providing an abstract level of grammatical structure at which X-bar principles or the parameters governing directionality of heads and complements are stated. Typological consistency, the story goes, is associated with simplicity at that deep level, typological inconsistency with complexity at that level. However, upon close examination, there turns out to be no story of success to tell at all. The properties of that level of grammar, however well motivated they may be for a particular language, bear but little on typological generalizations. To be specific, the Greenbergian correlations are more robust at surface levels than at deep levels of grammar.

Let us begin with German and Dutch. These languages are typologically peculiar in two different ways. First, while virtually all generativists agree that they are underlyingly head-final in VP (see Bach, 1962; Koster, 1975; Bennis and Hoekstra, 1984), they are uncontroversially head initial in other phrases. Second, a 'V2 rule' is responsible for VO order in main clause declaratives, while leaving intact OV order in embedded sentences. What this means is that in German and Dutch we find greater typological consistency at the surface, where VO order dominates by far in actual discourse (given the frequency of main clause declaratives), than at D-structure, where OV order clashes with post-head complements for N, P, and A.

An interesting case in point is Amharic. The surface order of this language is clearly SOV. In a classic paper, Bach (1970a) presented a series of arguments based on grammar-internal regularities that it is, in fact, a VO language in deep structure. An independently motivated rule of verb-shift was argued to create the verb-final surface order. And Bach did, in fact, point to a few correlates of VO order in Amharic. For example, it is prepositional, it exhibits 'gapping' behavior (Ross, 1970) more like a VO language than an OV language, and it shares with VO languages the property of placing honorifics before the name rather than after it. Bach might well have been correct that this language is underlyingly VO (though for a contrary opinion, see Hudson, 1972). However, its typological properties seem to be more in accord with the OV word order that predominates on the surface than with VO order. For example, as Bach himself noted, auxiliary verbs follow the main verb, rather than precede it –behavior that we would expect from an OV language. Bach went on to note other correlates with OV syntax, such as the genitive expression



always preceding the governing noun, the order in comparatives being standard-marker-adjective, and the relative construction usually preceding the modified noun. Even Amharic's prepositionality, a generally reliable marker of VO syntax, is not exceptionless. Bach pointed to a number of postposed elements denoting spatial relations ('inside', 'top', etc.). The other correlates of VO syntax cited by Bach are highly controversial (see Jackendoff, 1970 and Maling, 1972 on gapping) or are of unclear grammatical relevance (e.g. the positioning of honorifics).

There is another respect in which typological generalizations seem to be more robust on the surface than at a deep level. If we eliminate reorderings of elements whose principal function seems to be to place 'heavy' elements at the periphery of the clause, it is my impression that deeply inconsistent languages overwhelmingly allow variant surface order that fulfill the Greenbergian correlations, while deeply consistent languages much less frequently allow variant surface orders that violate them. For an example of the former case, consider Persian. That language is deeply inconsistent in the same sense that German is –of the four major phrasal categories, only VP is head-final. However, on the surface Persian allows a number of reorderings of S, V, and O, subject to purely grammatical conditions. For example, a direct object followed by the specificity marker *râ* can move freely within the verb phrase (for full discussion, see Karimi 1989). In other words, Persian *does* have head-initial VPs. Japanese illustrates the latter case. While that deeply consistent SOV language does indeed manifest surface orders of OSV and SVO, these orders occur, I believe, only as a result of 'scrambling', where it is not clear that we have an instantiation of Move-a. A number of linguists have put forward arguments, quite strong ones in my opinion, that the repositioning that we find in scrambling lacks many of the hallmarks of a transformational rule (see Lee, 1992; Bayer and Kornfilt, 1994; Kiss, 1994; Neeleman, 1994).

There are, of course, any number of languages for which controversy exists as to their underlying order or for which no intensive investigation of this feature has been undertaken. Many such languages allow a variety of surface orders of subject, object, and verb, where the ordering among them is determined largely by discourse considerations. As it turns out, as far as the Greenbergian correlations are concerned, their underlying word order is irrelevant. In an important study, Matthew Dryer (1989) has shown that languages with discourse-governed word order often exhibit the word order characteristics associated with *the most frequent surface word order* in the language. Table 5 illustrates Dryer's findings for 10 such languages.<sup>7</sup>

Frequency of $QV$		OV characteristics
Ute	72%	GenN, Po
Tlingit	67%	GenN, Po, RelN
Huallaga Quechua	69%	GenN, Po, RelN
Trumai	65%	GenN, Po
Koryak	66%	GenN, Po
Tacana	86%	GenN, Po, Clause-final subordinator
Takelma	85%	GenN, Po
Hupa	53%	GenN, Po
Cherokee	(more frequent)	GenN, Po
Korana	89%	GenN, Po, Clause-final Q, etc.

*Frequency of OV order and OV characteristics (Dryer, 1989)*  
*Table 5*

In other words, the D-structure order of elements in such languages, whatever it might be, seems irrelevant to determining their typological properties. And finally, there exists a set of languages for which strong arguments exist that there *is no* underlying order of subject, object, and verb (see, for example, Kiss, 1987 on Hungarian and Hale, 1992 on Warlpiri). It goes without saying that in such languages relations among D-structure elements could not bear on the Greenbergian correlations.

The above discussion has presupposed an approach to syntax containing a level of D-structure over which grammatical generalizations can be formulated. The recent trend in principles-and-parameters work toward 'minimalist' models lacking such a level fails as well to provide a nonstipulative theory-internal explanation of the Greenbergian correlations. The MP, which provides no 'basic order' among grammatical elements or would have all languages being underlyingly SVO (Kayne, 1994), must capture cross-categorial generalizations (and exceptions to these generalizations) by means of relations holding among feature strengths. So the correlations would presumably be captured in terms of the strength of the features that check object case. Under one realization of this possibility, if the case features of N, V, A, and P are weak, we would get head-complement order; if strong, then complement-head order. Marked inconsistency might be derivable by allowing the features

associated with the functional projections of these categories to differ (e.g. a strong feature for N, but a weak one for V).

There are two problems with such an approach for our concerns, one identical to those faced by models containing a level of D-structure and one unique to the structure of minimalism. As far as the former is concerned, if any argument for a D-structure order of elements in GB carries over to an argument for a derivationally-prior order in the MP, as I assume that it does, then the MP fails as well to capture the generalization that surface order, rather than deep order, is the best predictor of the Greenbergian correlations. But another problem arises in the MP as a result of its inability to distinguish base orders of grammatical elements from transformationally-derived orders. Consider a language which manifests all the Greenbergian correlations with OV order and to which a principled GB account would, indeed, assign a SOV D-structure order. Let's say that this language allows SVO order as a marked variant under extremely restrictive grammatical conditions. In GB the marked order would be transformationally derived and hence theoretically distinguishable from the basic SOV order. But there is no mechanism internal to the MP (novel stipulations aside) that would distinguish the feature-driven SOV order from the equally feature-driven SVO order. Hence the MP would fail to capture the 'essential SOV-ness' of this language.

### 3.2.3. THE FAILURE OF GENERATIVE THEORY-DERIVED TYPOLOGICAL PREDICTIONS: THE CASE OF THE NULL SUBJECT PARAMETER

Grammarians have long been aware of a typological difference between, say, English and French on the one hand and Spanish and Italian on the other. The latter languages, but not the former, allow the omission of thematic subjects. Hence:

- (8) (a) (\*I) arrived yesterday. (English)  
(b) (Yo) llegué ayer. (Spanish)

Perlmutter (1971) suggested that if a language allows the omission of thematic subjects then it will also allow null nonthematic subjects (as, for example, the subject of weather verbs) and also the extraction of a subject from an embedded clause headed by an overt complementizer (later called 'that-trace filter violations'). Rizzi (1982) added another typological correlate, namely, the possibility of free subject inversion in simple sentences. Other correlates have been proposed as well, such as long

*Wh*-Movement of subjects and empty resumptive pronouns in embedded clauses (Chomsky, 1981).

A number of proposals have been put forward over the years to capture some or all of the above-mentioned putative correlations. A guiding assumption in generative research is that a single setting of the 'Null Subject Parameter' suffices to derive them. It is not my intention to summarize the various proposals here (for an overview of the issues involved, see the papers collected in Jaeggli and Safir, 1989). However, the predictions of two well-developed ones will be subject to scrutiny: Rizzi (1982) and Safir (1985). Both proposals make claims about four of the typological features that have been claimed to be associated with this parameter, namely, null thematic subjects in tensed clauses, null nonthematic (expletive) subjects, subject inversion, and *that*-trace violations. As is shown in the 'LANGUAGE TYPES' box of Table 6, there are 16 ways that these 4 properties might in principle be distributed in a particular language.

Neither Rizzi nor Safir predict that these four features be rigidly correlated. Rizzi, for example, predicts a language type without null thematic subjects but the presence of the other three features. Hence, he predicts the existence of types (a, i, p). Safir, who proposes a somewhat more permissive theory with three interacting parameters, also predicts (a, i, p), but, in addition, types (d, m). These predictions are indicated in the left column of the 'PREDICTIONS AND ATTESTATIONS' box of Table 6 ('LR' refers to Rizzi's predictions and 'KS' to Safir's).

Rizzi's and Safir's predictions were put to the test by Gilligan (1987), who worked with a 100 language sample, which he attempted to correct for areal and genetic bias. Of these 100, there were 10 with data for all four properties. To these he added 19 more languages, for which data on these properties was available in the literature. As the right column of the 'PREDICTIONS AND ATTESTATIONS' box of Table 6 illustrates, neither Rizzi's nor Safir's predictions appear to be borne out. All three of the language types that Rizzi predicts to exist are attested, as are four of the five Safir predicts (and the missing one could easily be a function of the small sample). But disconcertingly, five types that neither theory predicts are attested.

	LANGUAGE TYPES				PREDICTIONS AND ATTESTATIONS	
	null thematic subjects	null nonthem. subjects	subject inversion	<i>that</i> trace filter violations	predicted	number in sample
a.	+	+	+	+	LR, KS	7
b.	+	+	+	-	no	4
c.	+	+	-	+	no	6
d.	+	+	-	-	KS	2
e.	+	-	+	+	no	0
f.	+	-	+	-	no	0
g.	+	-	-	+	no	0
h.	+	-	-	-	no	0
i.	-	+	+	+	LR, KS	1
j.	-	+	-	+	no	2
k.	-	+	-	-	no	1
l.	-	+	+	-	no	0
m.	-	-	+	+	KS	0
n.	-	-	-	+	no	3
o.	-	-	+	-	no	0
p.	-	-	-	-	LR, KS	3

*The null-subject parameter and its predicted correlations (Gilligan, 1987)*  
Table 6

Now, one must be clear that these results in and of themselves do not necessarily refute either Rizzi or Safir. It is possible that the cases of nonpredicted subject inversion, for example, are the result of something other than the null subject parameter. As Gilligan himself points out: 'Perhaps the Rizzi hypothesis is correct but its effects are obscured in [Brazilian Portuguese and Mandarin –two languages with a cluster of properties predicted not to exist] because of some as yet unanalyzed aspect of these languages' (1987: 90). Or, perhaps we have another example of a sampling problem, which, if corrected, would bear out Rizzi or Safir. Or, again, perhaps the null subject parameter itself is an epiphenomenon, whose effects are to be attributed to other parameters. Nevertheless, the fact that even the most extensively investigated generative parameter appears to lack typological support makes one wonder what the status would be of the myriad of others, were they put to a similar test.

### 3.2.4. SIMPLER GRAMMARS ARE NOT NECESSARILY MORE COMMON GRAMMARS: THE CASE OF PREPOSITION STRANDING<sup>8</sup>

As noted above in §3.1, there is no theory-independent way of characterizing one proposed grammar of a language as being 'simpler' than another. However, as we have seen, there has long been the assumption that we *can* compare two grammars (or at least corresponding subparts of two grammars) in terms of simplicity, so long as both are formulated within the same set of theoretical assumptions. The more complex grammar will have an extra rule of some sort, the same number of rules, but with more of them 'marked', and so on. And by hypothesis, the more complex grammar will represent a cross-linguistically rarer state of affairs.

For one reasonably well-studied phenomenon, this prediction is false. The simpler grammar is far rarer cross-linguistically than the more complex one. The phenomenon is 'preposition-stranding', illustrated in (9a-b) for English. In (9a) *Wh*-Movement has extracted and fronted the object of *to*, leaving the bare preposition behind. In (9b) NP-movement has taken *Mary*, the underlying object of the preposition *to*, and moved it into subject position, stranding the preposition:

- (9) (a) Who did you talk to?  
(b) Mary was spoken to.

Stranding is extremely rare cross-linguistically. In fact, it is attested only in the Germanic family (though not in German itself) and in some varieties of French. Surely, then, if a typologically rare state of affairs were to be represented by a more complex grammar, we would expect a grammar with stranding to be vastly more complicated in relevant respects than one without. Such is not the case, however. In GB terms, grammars without stranding can be captured by generalization (10a), those with stranding by (10b):

- (10) (a) NON-STRANDING LANGUAGES: The lexical categories N, V, and A are proper governors. The lexical category P is not a proper governor.  
(b) STRANDING LANGUAGES: All four lexical categories are proper governors.

When P is not a proper governor, extraction of its object is impossible, since the resultant trace would be ungoverned. A properly governing

preposition, however, allows extraction and may therefore occur 'stranded' on the surface.

It is difficult to imagine how a grammar incorporating (10a) could be regarded as simpler than one incorporating (10b). Aside from the pure (and nonexplanatory!) stipulation that it is the unmarked state of affairs in UG for P not to properly govern, there is no natural reason why P should be exceptional in this respect. Like other lexical categories, it assigns theta-roles, Case, and along with N, V, and A, it can be characterized by the distinctive features  $\pm N$ ,  $\pm V$ .

To be sure, there is no dearth of analyses of stranding that do complicate the grammars of languages that have it. For example, in one popular approach (Hornstein and Weinberg, 1981), P is never a proper governor. In languages that allow stranding, prepositions have the ability to overcome this defect by undergoing 'reanalysis' with an adjacent verb, thereby creating a complex verb that can properly govern the trace of movement, as shown in (11a-b):

- (11) (a) You talked  $P_p$ [to who] > You  $v$ [talked to] who > Who<sub>i</sub> did you  
 $v$ [talk to] e<sub>i</sub>?  
 (b) e was spoken  $P_p$ [to Mary] > e was  $v$ [spoken to] Mary > Mary<sub>i</sub> was  
 $v$ [spoken to] e<sub>i</sub>

The reanalysis approach to preposition stranding is riddled with problems, however. A number of tests show that, in general, the reanalyzed material does not behave as a single lexical item. For example, reanalysis would have to be assumed to create utterly implausible lexical items, such as *walk across Europe in* and *pay twice for*, as in (12a-b):

- (12) (a) Which shoes did you [walk across Europe in]? (Jones, 1987)  
 (b) Which of the two knives did you [pay twice for]? (Inada, 1981)

Furthermore, as noted in Koster (1986), Gapping does not treat the verb-preposition complex as a verb (13a-b), nor does Heavy NP Shift (14a-b). Even more problematically, reanalysis demands the possibility of Extraposition out of a lexical item, as in (15) (Levine, 1984), and, as pointed out by Hornstein and Weinberg (1981), in the very article in which reanalysis was first proposed, it demands mutually incompatible analyses, as in (16a-b), where Wh-Movement and Passive have applied in the same sentence:

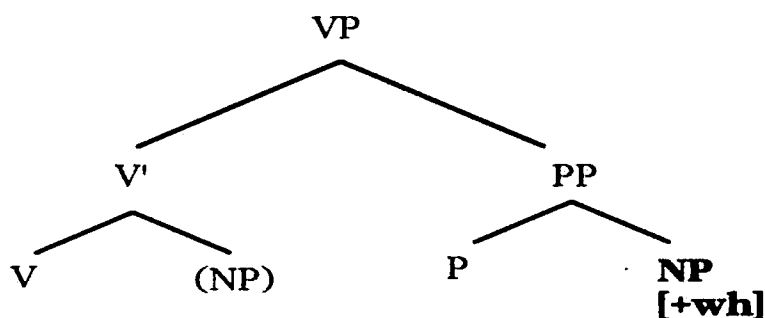
- (13) (a) \*John looked at Mary and Bill \_\_\_ Sue.  
 (b) John looked at Mary and Bill \_\_\_ at Sue.
- (14) (a) John looked at [the woman he loved] very often.  
 (b) John looked very often [at the woman he loved]  
 (c) \*John looked at very often [the woman he loved].
- (15) What did you [talk to that guy \_\_\_ about] \_\_\_ who was here yesterday?
- (16) (a) Which problems has Harry been [[talked to] e about] e?  
 (b) Who would you like to be [[sung to] e by] e?

Let us therefore abandon a reanalysis approach to stranding and adopt in its place the proposal first put forward, I believe, in Jones (1987) that P is a proper governor in English and other stranding languages. If such is correct, it is predicted that within V', V and P need not be adjacent. As the sentences of (17) illustrate, this is indeed the case:

- (17) (a) Who did you give all those books about golf to?  
 (b) Which burner did you leave the pot on?

The most interesting prediction of this analysis is that stranding should be possible with the extraction of NP from PP adjuncts to VP, i.e. in situations like (18):

(18)



Extraction of the bold-faced *wh*-phrase leads to the crossing of only one barrier, the PP itself. As predicted, then, sentences like (19a-d) are grammatical:



- (19) (a) Which shoes did you walk across Europe in?  
(b) Which ball park did Ruth hit the most home runs in?  
(c) Which knife shall we use to cut the turkey with?  
(d) Which red-headed man is Mary standing beside?

Now, it is a curious fact that many previous analyses of stranding have deemed analogous sentences ungrammatical. For example Hornstein and Weinberg (1981) point to the famous ambiguity of (20a), which they contrast to the seeming nonambiguity of (20b):

- (20) (a) John decided on the boat.  
(b) What did John decide on?

In their view, the adjunct (i.e. locative) reading is impossible in (20b). This follows, in their theory, from the restriction of reanalysis to subcategorized complements of V. But in fact, it is not hard to construct a sentence with just such a reading. Consider (21), imagining a situation in which John has been going from floor to floor in a department store specializing in vehicles of all sorts, trying to decide whether to buy a boat or a car. The sentence is impeccable:

- (21) Which floor did John decide on the boat on?

Clearly we would not want to say that (20b) is unambiguous.

One might object that if prepositions are proper governors, many sentences of dubious acceptability are predicted to be grammatical. For example, consider (22a-b):

- (22) (a) Who did you read a book about?  
(b) ?Who did you destroy a book about?

There have been a number of attempts to treat the deviance of (22b) in the grammar itself by devising grammatical operations to rule it out (Bach and Horn, 1976; Chomsky, 1977). That strikes me as entirely the wrong move. (22b) is in fact fully grammatical, as is suggested by the well-formedness of sentences like (23a-b):

- (23) (a) Which former party official did the Red Guard destroy more books about: Lin Piao or Liu Shao-Chi?  
(b) The party official that I would really like to destroy a book about is Chao En-Lai.

Destroying books is not a normal activity. If we create a discourse context in which we make it one –that is, if we make the extracted phrase the center of attention– extraction from the complement of *destroy* creates no problems.

To summarize, preposition stranding does not pay for its rarity by requiring complex rules for its formulation in grammars that license it. Even within the same general framework of theoretical assumptions, the more complex grammar is not necessarily the more cross-linguistically rare grammar.

### 3.3. TYPOLOGICAL GENERALIZATIONS AND LANGUAGE ACQUISITION

In this section I will question the existence of any significant links between typological generalizations and facts about the acquisition of language by children. The lack of such links provides one more argument that they are irrelevant to grammatical theory. I will argue that there has never really been any compelling evidence presented that more cross-linguistically frequent structural types are acquired earlier in acquisition than infrequent types.

A decade ago, a number of arguments based purely on learning-theoretic considerations were put forward that some element of grammar had to be chronologically present before some other. For example, the Subset Principle of Berwick (1985), departing from the observation that child learners are presented only with positive evidence, embodied the idea that children had to choose the narrowest possible language consistent with the evidence presented to them. As noted by Lightfoot (1991: 12) and others, obligatory (i.e. nonnull) subjects would therefore have to represent the unmarked setting of the null subject parameter and therefore children's grammars should manifest obligatory subjects before null subjects. Such a conclusion was immediately observed to clash with the typological generalization that the majority of the world's languages appear to be in the null subject class. But it also clashes with the fact that children acquiring null subject languages do not go through an initial obligatory subject stage (Valian, 1991). The problem is now believed to lie with the Subset Principle itself; it is inapplicable, it seems, given that there are (in principle) only a finite number of parameters to be set (Kapur, 1994). Other learnability-derived constraints on the order of acquisition have been challenged in like fashion, and are therefore ultimately irrelevant to our concerns.

There are, however, any number of disparities between the designation of a particular structural type as 'unmarked' and the idea that unmarked

phenomena are acquired early. Once again, preposition stranding is a case in point. Though a marked phenomenon if any is (for an explicit statement to this effect, see Riemsdijk, 1978), English-acquiring children produce sentences with stranding *before* those without stranding, even when the input in their (educated middle-class) surroundings contains sentences of both types (Karin Stromswold, personal communication). The recent trend in generative-driven studies of first language acquisition has been to present evidence for a very rich syntax at a very early age – regardless of the typological status of the property being acquired. As Nina Hyams puts it, ‘the parameters of Universal Grammar, which are tied to functional categories, are set quickly and without error’ (Hyams 1998). By way of example, French-speaking children appear to have mastered the verb-raising parameter from the earliest multi-word utterances (Pierce, 1992); there appears to be no period during which German-speaking children fail to set the V2 parameter (Poeppel and Wexler, 1993); the null subject parameter is set very early, regardless of whether the language is null subject or not (Valian, 1991); and children acquiring English, German, and French evidence strong knowledge of locality in *wh*-extraction domains at early ages (Roeper and De Villiers, 1994). Indeed, I believe it to be the case that a majority of language acquisition specialists now subscribe to the ‘Strong Continuity Hypothesis’, which holds that the parameters of UG are set early and are not subject to change in the course of development (see especially Lust, Suñer, and Whitman, 1994). If this hypothesis is correct, then there can be little connection between typology and acquisition.

#### 4. EXPLAINING THE GREENBERGIAN CORRELATIONS

If grammatical theory *per se* cannot explain the Greenbergian correlations, then what can? The answer is a theory of language processing.

Dryer (1992) points out that most grammar-internal explanations of the correlations have been based on the idea that there is pressure to maximize the parallelism between heads and dependents. Both X-bar theory and the head parameter, in somewhat different ways, illustrate such an approach, which he refers to as ‘head-dependent theory’. Dryer argues that head-dependent theory is incorrect. What is relevant is not what is a head and what is a dependent, but rather the *branching direction* of the elements involved. According to the branching-direction theory, languages tend toward one of two ideals: right-branching languages, in which phrasal categories follow nonphrasal categories, and left-branching languages, in which phrasal categories precede nonphrasal categories.

Now, of course, in many cases, head-dependent theory and branching-direction theory make the same prediction. Relative clauses follow their heads in VO languages. Head-dependent theory predicts this ordering because both objects and relatives are complements. Branching-direction theory makes the same prediction because both VO structures and noun-relative structures are right branching. But the two theories are not equivalent; where they make different predictions, it is the branching-direction theory that is correct.

Let us begin by examining auxiliary - verb and determiner - noun order. Traditionally, auxiliaries have been considered to be dependents of verbal heads –their specifiers in many early versions of generative grammar (see, for example Akmajian, Steele and Wasow, 1979). Precisely the same can be said for articles with respect to nouns (Jackendoff, 1977). To be specific, auxiliaries were considered part of the maximal projection of V, modifying their verbal head, while articles were part of the maximal projection of N, modifying their nominal head. Head-dependent theory, then, would predict that the order between verb and auxiliary in most languages should parallel that between noun and article and that both should parallel the order between verb and object. But typological research has concluded that such is not the case. VO languages show a near universal tendency for auxiliaries to precede verbs, not to follow them, while in OV languages auxiliaries follow verbs, rather than precede them. For determiners and nouns, the typological correlations are less robust, but there is a general tendency for determiners to precede nouns in VO languages and to follow them in OV languages.

Branching-direction theory, however, makes the correct predictions. Auxiliary-verb and article-noun structures are overwhelmingly right-branching in VO languages and left-branching in OV languages. Hence we derive the ordering correlations verb - object, auxiliary - verb, and determiner - noun; object - verb, verb - auxiliary, and noun - determiner.

It is quite interesting that in recent years many generativists have abandoned the idea that auxiliaries and determiners are specifiers of verbs and nouns respectively. Instead, it has become widely accepted that verbs (or, more properly verb phrases) are *complements* to auxiliary heads and that nouns are *complements* to determiners (see especially Pullum and Wilson, 1977 for the former, Abney, 1987 for the latter). In a sense then, we have seen head-dependent theory and branching-direction theory converging, in that they now make identical predictions with respect to auxiliary - verb ordering and determiner - noun ordering. The convergence is not complete, however. There are cases in which head-dependent theory fails, while branching-direction theory does not.

Consider the following three examples. First, many languages have a category 'Demonstrative' that is distinct from Determiner. In such languages, there is no robust correlation between the order of demonstrative and noun, even though in relevant respects demonstratives parallel determiners semantically. Second, tense-aspect *particles* do not show a consistent ordering with respect to the verb, even though semantically parallel *auxiliaries* do so. And, third, there is no robust correlation between the ordering of adjective and noun and that of dependents and heads.

Branching-direction theory explains why these correlations do not hold. In each of the three cases, a reasonable phrase-structure analysis involves the concatenation of two nonphrasal categories (e.g. A and N within NP). There is no dominant branching direction and hence no correlation with the order of verbs and adpositions, auxiliaries and verbs, and so on.

Dryer pointed to the roots of branching-direction theory in parsing ease. This idea has been developed by Hawkins (1994) in a comprehensive theory of the influence of processing considerations on grammar. The central parsing principle that Hawkins proposes is called 'Early Immediate Constituents' (EIC) and is stated as follows (1994: 77):

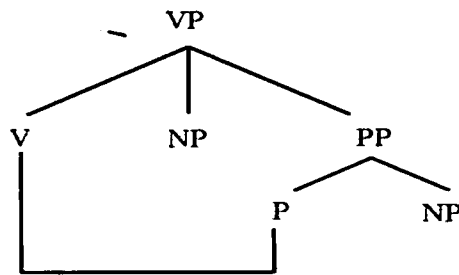
(24) Early Immediate Constituents (EIC)

The human parser prefers linear orders that maximize the IC-to non-IC ratios of constituent recognition domains (CRD).

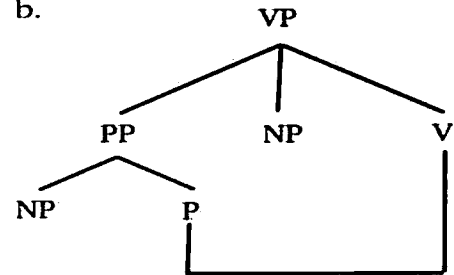
A 'constituent recognition domain' for a particular phrasal mother node M consists of the set of nodes that have to be parsed in order to recognize M and all of the ICs of M.

So consider how Hawkins derives the result that VO languages to be prepositional and OV languages to be postpositional. There are four logical possibilities, illustrated in (25a-d): VO and prepositional (25a); OV and postpositional (25b); VO and postpositional (25c); and OV and prepositional (25d):

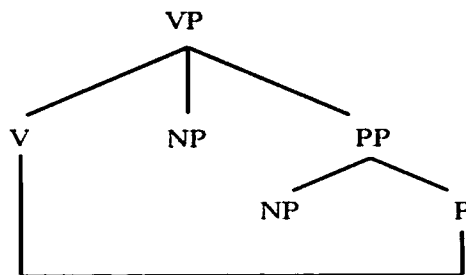
(25) a.



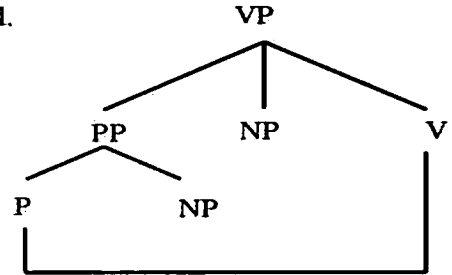
b.



c.



d.



Assuming that both NPs are two words long, in (25a) and (25b), the two typologically preferred structures, only 4 words have to be processed in order to identify the constituents of VP. But in (25c) and (25d), 6 must be processed. Furthermore, the longer the object of the prepositional phrase gets, the more processing will be necessary for (25c) and (25d), while that for (25a) and (25b) will remain the same. Analogous demonstrations can be made for other Greenbergian correlation pairs.

The correlation between verb-finality and lack of Wh-Movement also lends itself to a parsing explanation. Hawkins (1995) notes that heads, in general, are the best identifiers of their subcategorized arguments. If one hears the verb *give*, for example, one is primed to expect two associated internal arguments, one representing a recipient and the other an object undergoing transfer. On the other hand, a human NP might or might not be a recipient and an inanimate NP might or might not be an object undergoing transfer. Hence, if arguments precede their heads, as they do in SOV languages, extra cues are useful to identify their thematic status. Such can be accomplished by keeping them contiguous to the head (that is, by restricting their movement possibilities) and / or by endowing them with case marking that uniquely identifies their thematic role or helps to narrow down the possibilities.

In other words, the Greenbergian correlations are not at root facts provided by grammars. They are encoded in grammars only to the extent that to whatever degree the properties of grammars are a response to the pressures exerted by the mechanisms of language processing.

## 5. CONCLUSION

I have argued that typological generalizations are not encoded in grammars, either directly or indirectly. That is, there is no set of principles or parameters internal to a theory of UG from which cross-linguistic facts can be derived. It is not surprising, therefore, that attempts to provide UG-internal explanations for them have been failures. Nor is it surprising that there appears to be no correlation between the typological status of a grammatical feature and the order of appearance of that feature in child language. The task of explaining the most robust typological generalizations, *the Greenbergian correlations*, falls not to UG, but to the theory of language processing. In short, it is the task of grammatical theory to characterize the notion *possible human language*, but not the notion *probable human language*. In this sense, then, typology is indeed irrelevant to grammatical theory.

Frederick J. Newmeyer  
Department of Linguistics  
University of Washington  
Seattle, WA 98195-4340 USA  
fjn@u.washington.edu

## NOTES

1. A version of this paper was presented at the Workshop on Formalism, Functionalism, and Typology at the semi-annual meeting of the Linguistics Association of Great Britain, held in Lancaster in April 1998. I would like to thank the following people for their helpful comments and criticism: William Croft, Richard Hudson, Peter Sells, and Anna Siewierska.

2. Dryer conflated VSO and SVO languages into one category, given that their typological correlates are largely the same (see Dryer, 1991).

3. Travis speculates that Chinese will have parametric system (b) in the future and that Kpelle had system (f) in the past. I would guess that there are attested languages that manifest these systems, though I am not aware of them.

4. Huang argues that his account is superior to Travis's, since (c'') predicts correctly that noun complements in Chinese (unlike verb complements) will precede the head, while Travis's approach does not allow some heads to theta-mark to the left and others to the right within the same language.
5. To be accurate, we will see only that (7a-b) are not correct. For reasons of space, the relationship between typology and language change will not be discussed in this paper.
6. The figures in the 'Final Q particles' row give the proportion of final question particles out of the total number of final and initial particles. Languages with no question particles at all, or those whose particles occur nonperipherally, are not counted.
7. Dryer calls attention to a few languages in which his generalization does not appear to hold: Papago (higher frequency of VO, but typologically mixed), Yagua (VO more frequent, but GenN and postpositional), Hanis Coos (VO more frequent, but GenN more common than NGen), and Cree (VO more frequent, but postpositional). Dryer notes the (possibly) troublesome fact that in all of his instantiating languages, OV order is more common than VO and the languages exhibit OV characteristics.
8. The remarks in this section are developed in considerably more detail in Newmeyer (in press).
9. Dryer points out (p. 99) that in most principles-and-parameters accounts of inflectional elements, tense/aspect particles, as well as auxiliary verbs, are heads. Thus such accounts predict, incorrectly, that particle - verb order should parallel auxiliary - verb order.
10. For earlier parsing accounts of the Greenbergian correlations, see Kuno (1974) and Frazier (1985).

#### REFERENCES

- Abney, Steven P. (1987) *The English noun phrase in its sentential aspect*. Unpublished Ph. D. dissertation, MIT.
- Akmajian, Adrian, Susan Steele & Thomas Wasow (1979) "The category AUX in universal grammar", *Linguistic Inquiry* 10, 1-64.
- Bach, Emmon (1962) "The order of elements in a transformational grammar of German", *Language* 38, 263-269.
- Bach, Emmon (1965) "On some recurrent types of transformations". In Charles W. Kreidler (ed.) *Approaches to linguistic analysis, language and society, teaching language skills* (= Monograph Series on Languages and Linguistics No. 18; 16th Round Table Meeting), Washington: Georgetown University Press. 3-18.
- Bach, Emmon (1970a) "Is Amharic an SOV language?", *Journal of Ethiopian Studies* 8, 9-20.
- Bach, Emmon (1970b) "Questions", *Linguistic Inquiry* 2, 153-166.



- Bach, Emmon and George M. Horn (1976) "Remarks on 'Conditions on transformations'", *Linguistic Inquiry* 7, 265-299.
- Baker, C. L. (1970) "Notes on the description of English questions: the role of an abstract question morpheme", *Foundations of Language* 6, 197-219.
- Bayer, Josef and Jaklin Kornfilt (1994) "Against scrambling as an instance of move-a". In Norbert Corver and Henk van Riemsdijk (eds.) *Studies on scrambling*. Berlin: Mouton de Gruyter. 17-60.
- Bennis, Hans and Teun Hoekstra (1984) "Gaps and parasitic gaps", *Linguistic Review* 4, 29-87.
- Berwick, Robert C. (1985) *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.
- Bowerman, Melissa (1985) "What shapes children's grammars?" In Dan I. Slobin (ed.) *The crosslinguistic study of language acquisition*, vol. 2. *Theoretical issues*. Hillsdale, NJ: Erlbaum. 1257-1319.
- Bresnan, Joan W. (1970) "On complementizers: toward a syntactic theory of complement types", *Foundations of Language* 6, 297-321.
- Cheng, Lisa L. S. (1991/1997) *On the typology of wh-questions*. New York: Garland.
- Chomsky, Noam (1962) "Explanatory models in linguistics". In E. Nagel, P. Suppes, and A. Tarski (eds.) *Logic, methodology, and philosophy of science*. Stanford, CA: Stanford University Press. 528-550.
- Chomsky, Noam (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam (1970) "Remarks on nominalization". In Roderick Jacobs and Peter Rosenbaum (eds.) *Readings in English transformational grammar*. Waltham, MA: Ginn. 184-221.
- Chomsky, Noam (1977) "On *wh*-movement". In Peter Culicover, Thomas Wasow, and Adrian Akmajian (eds.) *Formal syntax*. New York: Academic Press. 71-132.
- Chomsky, Noam (1981) *Lectures on government and binding*. *Studies in Generative Grammar* 9. Dordrecht: Foris.
- Chomsky, Noam (1982) *The generative enterprise: a discussion with Riny Huybregts and Henk van Riemsdijk*. Dordrecht: Foris.
- Chomsky, Noam (1986) *Knowledge of language: its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam (1995) "Introduction". In Noam Chomsky (ed.) *The minimalist program*. Cambridge, MA: MIT Press. 1-11.
- Chomsky, Noam and Morris Halle (1968) *Sound pattern of English*. New York: Harper and Row.

- Dryer, Matthew S. (1988) "Object-verb order and adjective-noun order: Dispelling a myth", *Lingua* 74, 77-109.
- Dryer, Matthew S. (1989) "Discourse-governed word order and word order typology", *Belgian Journal of Linguistics* 4, 69-90.
- Dryer, Matthew S. (1991) "SVO languages and the OV:VO typology", *Journal of Linguistics* 27, 443-482.
- Dryer, Matthew S. (1992) "The Greenbergian word order correlations", *Language* 68, 81-138.
- Emonds, Joseph E. (1980) "Word order in generative grammar", *Journal of Linguistic Research* 1, 33-54.
- Frazier, Lyn (1985) "Syntactic complexity". In David Dowty, Lauri Karttunen, and Arnold M. Zwicky (eds.) *Natural language parsing: psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press. 129-189.
- Fukui, Naoki (1986) *A theory of categorial projection and its applications*. Unpublished Ph. D. dissertation, MIT.
- Gilligan, Gary M. (1987) *A cross-linguistic approach to the pro-drop parameter*. Unpublished Ph. D. dissertation, University of Southern California.
- Givón, Talmy (1981) "Typology and functional domains", *Studies in Language* 5, 163-193.
- Greenberg, Joseph H. (1963) "Some universals of language with special reference to the order of meaningful elements" In Joseph Greenberg (ed.) *Universals of language*. Cambridge, MA: MIT Press. 73-113.
- Hale, Kenneth (1976) "Linguistic autonomy and the linguistics of Carl Voegelin", *Anthropological Linguistics* 18, 120-128.
- Hale, Kenneth (1992) "Basic word order in two 'free word order' languages". In Doris Payne (ed.) *Pragmatics of word order flexibility*. Amsterdam: John Benjamins. 63-82.
- Hawkins, John A. (1994) *A performance theory of order and constituency*. *Cambridge Studies in Linguistics* 73. Cambridge: Cambridge University Press.
- Hawkins, John A. (1995) "Argument-predicate structure in grammar and performance: a comparison of English and German". In Irmengard Rauch and Gerald F. Carr (eds.) *Insights in Germanic linguistics I*. Berlin: Mouton de Gruyter. 127-144.
- Hopper, Paul J. & Sandra A. Thompson (1980) "Transitivity in grammar and discourse", *Language* 56, 251-299.
- Hornstein, Norbert and Amy Weinberg (1981) "Case theory and preposition stranding", *Linguistic Inquiry* 12, 55-92.

- Huang, C.-T. James (1994) "More on Chinese word order and parametric theory". In Barbara Lust, Margarita Suñer, and John Whitman (eds.) *Syntactic theory and first language acquisition: cross-linguistic perspectives*, vol. 1, *Heads, projections, and learnability*. Hillsdale, NJ: Erlbaum. 15-35.
- Hudson, Grover (1972) "Why Amharic is not a VSO language", *Studies in African Linguistics* 3, 127-165.
- Hyams, Nina M. (1998) "Underspecification and modularity in early syntax: a formalist perspective on language acquisition". In Michael Darnell, Edith Moravcsik, Frederick J. Newmeyer, Michael Noonan, and Kathleen Wheatley (eds.) *Functionalism and formalism in linguistics*, vol. 1, *General papers*. Amsterdam: John Benjamins. 385-414.
- Inada, T. (1981) "Problems of reanalysis and preposition stranding", *Studies in English Linguistics* 9, 120-131.
- Jackendoff, Ray (1970) "Gapping and related rules", *Linguistic Inquiry* 2, 21-35.
- Jackendoff, Ray (1977) *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
- Jaeggli, Osvaldo & Kenneth J. Safir, Eds. (1989) *The null subject parameter*. *Studies in Natural Language and Linguistic Theory* 15. Dordrecht: Reidel.
- Jones, Charles (1987) "P for proper governor", *WCCFL* 6, 115-130.
- Kapur, Shyam (1994) "Some applications of formal learning theory results to natural language acquisition". In Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt (eds.) *Syntactic theory and first language acquisition: cross-linguistic perspectives*. Hillsdale, NJ: Erlbaum. 491-508.
- Karimi, Simin (1989) *Aspects of Persian syntax, specificity, and the theory of grammar*. Unpublished Ph. D. dissertation, University of Washington.
- Kayne, Richard S. (1994) *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kean, Mary-Louise (1975) *The theory of markedness in generative grammar*. Unpublished Ph. D. dissertation, MIT.
- Kim, Soowon (1990) *Chain scope and quantificational structure*. Unpublished Ph. D. dissertation, Brandeis University.
- Kiss, Katalin É. (1987) *Configurationality in Hungarian*. *Studies in Natural Language and Linguistic Theory*. Dordrecht: Reidel.
- Kiss, Katalin É. (1994) "Scrambling as the base-generation of random complement order". In Norbert Corver and Henk van Riemsdijk *Studies on scrambling*. Berlin: Mouton de Gruyter. 221-256.

- Koopman, Hilda (1984) *The syntax of verbs: from verb movement rules in the Kru languages to universal grammar. Studies in Generative Grammar* 15. Dordrecht: Foris.
- Koster, Jan (1975) "Dutch as an SOV language", *Linguistic Analysis* 1, 111-136.
- Koster, Jan (1986) *Domains and dynasties: the radical autonomy of syntax. Studies in Generative Grammar* 30. Dordrecht: Foris.
- Kuno, Susumo (1974) "The position of relative clauses and conjunctions", *Linguistic Inquiry* 5, 117-136.
- Lee, Young-Suk (1992) "Case and word order variations in nominal clauses", *Language Research* 6, 359-380.
- Lehmann, Winfred P. (1973) "A structural principle of language and its implications", *Language* 49, 47-66.
- Levine, Robert D. (1984) "Against reanalysis rules", *Linguistic Analysis* 14, 3-30.
- Lightfoot, David W. (1979) *Principles of diachronic syntax. Cambridge Studies in Linguistics* 23. Cambridge: Cambridge University Press.
- Lightfoot, David W. (1991) *How to set parameters: arguments from language change. Cambridge, MA: MIT Press.*
- Lust, Barbara, Margarita Suñer, and John Whitman (eds.) (1994) *Syntactic theory and first language acquisition: cross-linguistic perspectives* 1, *Heads, projections, and learnability*. Hillsdale, NJ: Erlbaum.
- Maling, Joan (1972) "On 'Gapping and the order of constituents'", *Linguistic Inquiry* 3, 101-108.
- Marantz, Alec (1995) "The minimalist program". In Gert Webelhuth (ed.) *Government binding theory and the minimalist program: principles and parameters in syntactic theory*. Oxford: Blackwell. 349-381.
- McCloskey, James (1996) "On the scope of verb movement in Irish", *Natural Language and Linguistic Theory* 14, 47-104.
- Neeleman, Ad (1994) "Scrambling as a D-structure phenomenon". In Norbert Corver and Henk van Riemsdijk (eds.) *Studies on scrambling*. Berlin: Mouton de Gruyter. 387-430.
- Newmeyer, Frederick J. (1998) *Language form and language function*. Cambridge, MA: MIT Press.
- Newmeyer, Frederick J. (in press) "Preposition stranding: Parameteric variation and pragmatics", *Languages and Linguistics*.
- Perlmutter, David (1971) *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart, Winston.
- Pierce, Amy (1992) *Language acquisition and syntactic theory: A comparative analysis of French and English child grammars*. Dordrecht: Kluwer.

- Pinker, Steven (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Poeppel, David and Kenneth Wexler (1993) "The full competence hypothesis of clause structure in early German", *Language* 69, 1-33.
- Pullum, Geoffrey K. and Deirdre Wilson (1977) "Autonomous syntax and the analysis of auxiliaries", *Language* 53, 741-788.
- Riemsdijk, Henk van (1978) *A case study in syntactic markedness: the binding nature of prepositional phrases*. *Studies in Generative Grammar* 4. Dordrecht: Foris.
- Rizzi, Luigi (1982) *Issues in Italian syntax*. *Studies in Generative Grammar* 11. Dordrecht: Foris.
- Roeper, Thomas and Jill De Villiers (1994) "Lexical links in the *wh*-chain". In Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt (eds.) *Syntactic theory and first language acquisition: cross-linguistic perspectives*, vol. 2, *Binding, dependencies, and learnability*. Hillsdale, NJ: Erlbaum. 357-390.
- Ross, John R. (1970) "Gapping and the order of constituents". In Manfred Bierwisch and K. Heidolph (eds.) *Progress in linguistics*. Mouton: The Hague. 249-259.
- Safir, Kenneth J. (1985) *Syntactic chains*. Cambridge: Cambridge University Press.
- Tomlin, Russell S. (1986) *Basic word order: Functional principles*. London: Croom Helm.
- Travis, Lisa (1989) "Parameters of phrase structure". In Mark R. Baltin and Anthony S. Kroch (eds.) *Alternative conceptions of phrase structure*. Chicago: University of Chicago Press. 263-279.
- Valian, Virginia (1991) "Syntactic subjects in the early speech of Italian and American children", *Cognition* 40, 21-81.
- Vennemann, Theo (1974) "Theoretical word order studies: results and problems", *Papere zur Linguistik* 7, 5-25.
- Williams, Edwin S. (1981) "Language acquisition, markedness, and phrase structure". In Susan L. Tavakolian (ed.) *Language acquisition and linguistic theory*. Cambridge, MA: MIT Press. 8-34.