

# Why Typology Doesn't Matter to Grammatical Theory\*

Frederick J. Newmeyer  
University of Washington

## 1. Introduction

Since the term 'typology' is not always used consistently, I should clarify at the very beginning how I plan to use it throughout this paper. Quite simply, it is the study of the distribution of grammatical elements in terms of their relative frequency (as in Table 1) and implicational relationships (as in Table 2):

Word order	Number	Languages
		%
SOV	180	45
SVO	168	42
VSO	37	9
VOS	12	3
OVS	5	1
OSV	0	0
Total	402	

Frequencies of basic constituency orders (Tomlin 1986: 22)

Table 1

	VSO	SVO	SOV
Prep	6	10	0
Postp	0	3	11

Correlations between word order and adposition order (Greenberg 1963)

Table 2

Given that definition, my title might seem nothing less than shocking. Indeed, it has implications 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 typological generalizations are (in the relevant sense) ‘real’, then they are in need of explanation.

What will I be arguing then? Quite simply that it is not the job of grammatical theory per se to explain typological generalizations. Take the correlation in Table 2, a correlation that Japanese upholds and Persian violates. I will argue that the theory of UG does not encode the fact that the grammar of Japanese is in some sense more common, consistent, natural, desirable, etc. than the grammar of Persian. As a corollary, grammars are not evaluated more highly if they are more common or more consistent. That is, I am arguing against the idea that there exist markedness or other types of implicational relationships among parameter settings by which typological generalizations might be derived. Likewise, there is no correlation between how ‘simple’ a grammar is and how common it is. To summarize in one pithy slogan:

(1) UG characterizes the notion ‘possible human language’, not the notion ‘probable human language’.

I will argue instead that many if not most typological generalizations fall out from a theory of language processing.

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:

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)

As the following more recent quote indicates, however, Chomsky now has taken a very different position on the intrinsic interest of typological generalizations:

There has also been very productive study of generalizations that are more directly observable: generalizations about the word orders we actually see, for example. The work of Joseph Greenberg has been particularly instructive and influential in this regard. *These universals are probably descriptive generalizations that should be derived from principles of UG.* (Chomsky 1998: 33; emphasis added)

It is the last sentence of the more recent Chomsky quote that I will be disputing here.

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. Section 4 demonstrates that this expectation has not been fulfilled. Section 5 argues that the correlations have an extragrammatical explanation, and §6 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 (1998a: 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. 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 in Table 2 above. 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. Such correlations, it was widely agreed, could not be due to chance.

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 3.

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 3

In the remainder of this paper, I will regard the generalizations expressed in Table 3 as facts in need of explanation and refer to them as 'the Greenbergian correlations'.

### 3. Typological generalizations and generative grammar

The central goal of generative grammar from its inception has been to characterize the notion 'possible human language'. The vocabulary of theoretical primitives, conventions for formulating rules, etc. of the theory are chosen 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:

$$(2) \quad W_1 - W_2 - W_3 - \dots - W_n \longrightarrow W_n - \dots - W_3 - W_2 - W_1$$

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. In phonology typology has acted as a guide to theory construction from early on. 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).

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’ (p. 22), is necessarily quite indirect.

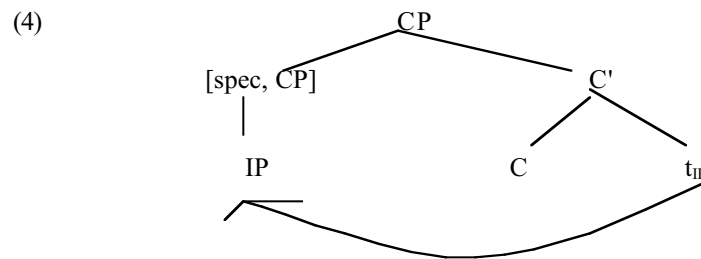
Nevertheless, there has been a guiding assumption that there is no significant gap between the notions ‘*typologically significant generalization*’ and

'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. As a case in point, consider the treatment of Chinese phrase structure in Huang (1994). Oversimplifying a bit, Chinese is consistently head-final, except for the rule expanding  $X'$  to  $X^0$ . 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:

- (3)
- |    |                  |                             |
|----|------------------|-----------------------------|
| 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.

The appeal to grammar-complicating extra statements to capture typological rarity is particularly developed in work based on Kayne's 'antisymmetry hypothesis' (Kayne 1994). Kayne develops a very restrictive theory of Universal Grammar allowing movement only to the left. The book contains discussion of a number of typological generalizations that appear to follow from the hypothesis. For example, in general COMP-final languages do not allow *wh*-Movement. Kayne provides an explanation: final complementizers arise from movement of IP into [spec, CP], thereby denying *Wh*-Movement a landing site:<sup>1</sup>



Now, some languages, like Vata, do have final COMP with *Wh*-Movement to the left. Kayne sketches some remarks on how such languages will have more complex grammars in this regard.

Cinque (1996) sets out to explain a large set of typological generalizations in Kayne's framework. For example, Kayne predicts that no language will have both N-Dem and Num-N. But some languages, including Berber, Hebrew, Welsh, and Zapotec, do have this correlation. Cinque posits an extra movement of demonstratives for these languages.

Kayne and Cinque are hazy on the formal mechanism for evaluating grammars so that more typologically grammars manifest themselves as more complex. In general, one has posited implicational relations among parameter settings for this purpose. For example, 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>')

(5) 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 (5a-h), which she designated 'headedness', 'direction of theta-role assignment', and 'direction of case assignment'. 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 4.

	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 4

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.

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:

[In] 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:

- (6) 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 (6a-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 section, (6a), at least, is not correct (for reasons of space there will be no discussion of language acquisition or language change).

#### **4. 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.

Let us begin by looking more deeply at the Kayne and Cinque attempt to capture typological generalizations. If the generalizations that they set out to explain were exceptionless, then we would have no problem saying that their accounts were successful. But they are not exceptionless. For example, as noted, some languages, despite the null hypothesis provided by Kayne's theory, do allow nouns to precede demonstratives and number words to precede nouns. As noted, this typologically dispreferred ordering requires that demonstratives undergo an extra movement. Now, that may very well be the correct analysis. But Cinque provides no explanation for why only 10% of the world's languages (let us say) have this extra movement. Nor does it follow — as far as we are told — from any other facts about the languages that have it. It could just as easily have been the case that 80% of the world's languages have the extra movement, thereby vitiating the typological generalization entirely.

Now Kayne and Cinque might reply that the more 'extra movements', the greater the degree of typological rarity. So the extra movement of demonstratives might be appealed to to explain why the dispreferred correlation

is so rare. But such a tack would go against Kayne's own assumptions. Kayne has all languages start out with specifier-head-complement order. Complement-head order derives from the movement of the complement to the left over the head. If so, then SOV languages have more complicated grammars than SVO languages. But typologists agree that complement-head order is more frequent in the world's languages than head-complement order. So for Kayne, extra movement correlates with a *more* common grammar.

The more deeply one looks, the more problematic is the idea that there exists a simplicity metric such that the 'simpler' grammar is the more cross-linguistically frequent one. A word of caution is in order, however, There is no theory-independent way of characterizing one proposed grammar of a language as being 'simpler' than another. However 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 (7a-b) for English. In (7a) *Wh*-Movement has extracted and fronted the object of *to*, leaving the bare preposition behind. In (7b) NP-movement has taken *Mary*, the underlying object of the preposition *to*, and moved it into subject position, stranding the preposition:

- (7) 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 (8a), those with stranding by (8b):

- (8) 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 (8a) could be regarded as simpler than one incorporating (8b). 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 (9a-b):

- (9) a. You talked<sub>pp</sub>[to who] > You<sub>v</sub>[talked to] who > Who<sub>i</sub> did you<sub>v</sub>[talk to] e<sub>i</sub>?  
b. e was spoken<sub>pp</sub>[to Mary] > e was<sub>v</sub>[spoken to] Mary > Mary<sub>i</sub> was<sub>v</sub>[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 (10a-b):

- (10) 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 (11a-b), nor does Heavy NP Shift (12a-b). Even more problematically, reanalysis demands the possibility of Extraposition out of a lexical item, as in (13) (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 (14a-b), where *Wh*-Movement and Passive have applied in the same sentence:

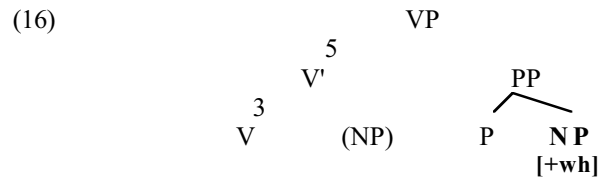
- (11) a. \*John looked at Mary and Bill \_\_\_ Sue.  
b. John looked at Mary and Bill \_\_\_ at Sue.  
(12) a. John looked at [the woman he loved] very often.  
b. John looked very often [at the woman he loved]

- c. \*John looked at very often [the woman he loved].
- (13) What did you [talk to that guy \_\_\_ about] \_\_\_ who was here yesterday?
- (14) 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 (15) illustrate, this is indeed the case:

- (15) 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 (16):



Extraction of the bold-faced *wh*-phrase leads to the crossing of only one barrier, the PP itself. As predicted, then, sentences like (17a-d) are grammatical:

- (17) 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?

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.

Kayne's ingenious explanation of why COMP-final languages tend to lack *Wh*-Movement is quite exceptional. In general, grammar-internal explanations of typological correlations have been pure stipulation. Let us consider another Greenbergian word order correlation. 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 5 from Dryer (1991: 455-466) provides the data supporting such an idea:<sup>2</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 5

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. But let us ask how this correlation might be handled within the Minimalist Program (MP). I can think of no nonstipulative means for doing so. 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 5 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.

In short, we have a robust typological generalization that seems not to follow from independently motivated principles of UG.

## 5. Explaining the Greenbergian correlations

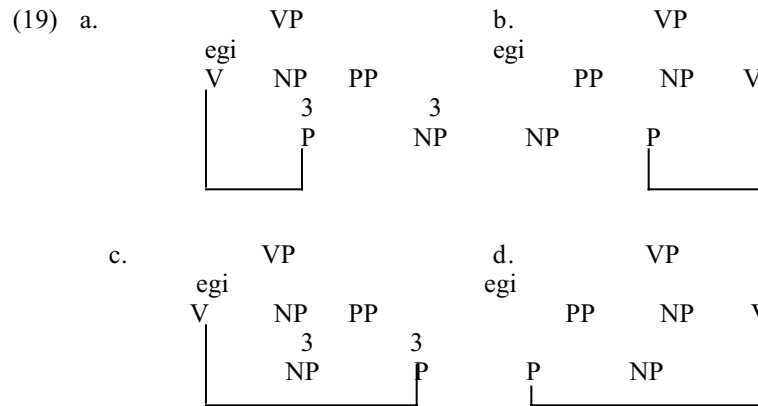
If grammatical theory per se cannot explain the Greenbergian correlations, then what can? The answer is a theory of language processing. That is, the theory of

grammar does not specify in any sense what correlates with what. But we still need to explain the fact that some grammars are more common than others and that some correlations are more common than others. The reason is, in a significant number of cases, that such grammars are easier to process. 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 (p. 77):

- (18) 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 (19a-d): VO and prepositional (19a); OV and prepositional (19b); VO and postpositional (19c); and OV and postpositional (19d):



Assuming that both NPs are two words long, in (19a) and (19b), the two typologically preferred structures, only 4 words have to be processed in order to identify the constituents of VP. But in (19c) and (19d), 6 must be processed. Furthermore, the longer the object of the prepositional phrase gets, the more

processing will be necessary for (19c) and (19d), while that for (19a) and (19b) 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 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.

It is generally assumed, I believe, that parsing-dictated orderings of elements are a feature of surface order, rather than deep order. (One thinks of parsing-dictated rightward movements of heavy objects and relative clauses.) Therefore, if typological generalizations have a parsing motivation, then we should expect more typological consistency at the surface level than at the deep level. Such an expectation, is in direct contradiction, of course, to the dominant idea that it is at an abstract level of structure at which X-bar principles or parameters of head directionality are stated. In fact, the parsing prediction is correct. 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.

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- $\alpha$ . 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).

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-categorical 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.

## 6. 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 does not matter to grammatical theory.

## Notes

\*Portions of this paper have appeared in Newmeyer (1998b) and are reprinted with permission.

<sup>1</sup> But see Bayer (1999) for a critique of Kayne's analysis.

<sup>2</sup> 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.

## References

- Bach, Emmon. 1962. 'The order of elements in a transformational grammar of German', *Language*, **38**: 263-269.
- Bayer, Josef. 1999. 'Final complementizers in hybrid languages', *Journal of Linguistics*, **35**: 233-272.
- Bayer, Josef and Jaklin Kornfilt. 1994. 'Against scrambling as an instance of move- $\alpha$ '. *Studies on scrambling*, ed. by Norbert Corver and Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 17-60.

- Bennis, Hans and Teun Hoekstra. 1984. 'Gaps and parasitic gaps', *Linguistic Review*, **4**: 29-87.
- Bowerman, Melissa. 1985. 'What shapes children's grammars?' *The crosslinguistic study of language acquisition*, vol. 2. Theoretical issues ed. by Dan I. Slobin. Hillsdale, NJ: Erlbaum, pp. 1257-1319.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Studies in Generative Grammar 9. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 1998. 'Noam Chomsky's minimalist program and the philosophy of mind. An interview [with] Camilo J. Cela-Conde and Gisèle Marty', *Syntax*, **1**: 19-36.
- Chomsky, Noam and Morris Halle. 1968. *Sound pattern of English*. New York: Harper and Row.
- Cinque, Guglielmo. 1996. 'The 'antisymmetric' program: Theoretical and typological implications', *Journal of Linguistics*, **32**: 447-465.
- 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.
- Greenberg, Joseph H. 1963. 'Some universals of language with special reference to the order of meaningful elements'. *Universals of language*, ed. by Joseph Greenberg. Cambridge, MA: MIT Press, pp. 73-113.
- 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'. *Insights in Germanic linguistics I*, ed. by Irmengard Rauch and Gerald F. Carr. Berlin: Mouton de Gruyter, pp. 127-144.
- 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'. *Syntactic theory and first language acquisition: Cross-linguistic perspectives, vol. 1: Heads, projections, and learnability* ed. by Barbara Lust, Margarita Suárez, and John Whitman. Hillsdale, NJ: Erlbaum, pp. 15-35.
- Inada, T.. 1981. 'Problems of reanalysis and preposition stranding', *Studies in English Linguistics*, **9**: 120-131.
- Jones, Charles. 1987. 'P for proper governor'. *WCCFL* 6, 115-130.
- 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.
- Kac, Michael (1972). Action and result: Two aspects of predication in English. *Syntax and semantics*, vol. 1 ed. by John Kimball. New York: Seminar Press, pp. 117-124.

- Kean, Mary-Louise (1975). The theory of markedness in generative grammar. Unpublished Ph. D. dissertation, MIT.
- Kiss, Katalin É. 1994. 'Scrambling as the base-generation of random complement order'. *Studies on scrambling*, ed. by Norbert Corver and Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 221-256.
- 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.
- Lee, Young-Suk. 1992. 'Case and word order variations in nominal clauses', *Language Research*, **6**: 359-380.
- 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.
- Marantz, Alec. 1995. 'The minimalist program'. *Government binding theory and the minimalist program: Principles and parameters in syntactic theory*, ed. by Gert Webelhuth. Oxford: Blackwell, pp. 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'. *Studies on scrambling*, ed. by Norbert Corver and Henk van Riemsdijk. Berlin: Mouton de Gruyter, pp. 387-430.
- Newmeyer, Frederick J.. 1998a. *Language form and language function*. Cambridge, MA: MIT Press.'
- Newmeyer, Frederick J.. 1998b. 'The irrelevance of typology for linguistic theory', *Syntaxis*, **1**: 161-197.
- Pinker, Steven. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Tomlin, Russell S.. 1986. *Basic word order: Functional principles*. London: Croom Helm.
- Travis, Lisa. 1989. 'Parameters of phrase structure'. *Alternative conceptions of phrase structure*, ed. by Mark R. Baltin and Anthony S. Kroch. Chicago: University of Chicago Press, pp. 263-279.'
- Williams, Edwin S.. 1981. 'Language acquisition, markedness, and phrase structure'. *Language acquisition and linguistic theory*, ed. by Susan L. Tavakolian. Cambridge, MA: MIT Press, pp. 8-34.

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