4

WORD ORDER

As has already been indicated in passing, word order typology has played a major role in the recent development of language typology. In large measure, this is because the current interest in language typology using data from a wide range of languages has taken its impetus from Greenberg's seminal article on word order typology: this article not only talked about doing this kind of language universals and typology research, but actually set about doing it. Although Greenberg himself is very cautious about the reliability of his results ('the tentative nature of the conclusions set forth here should be evident to the reader' is how his article starts), this caution has not been shared by all of those who have further developed his ideas, with the result that, as we shall see, generalizations have been claimed that go far beyond anything warranted by the data to hand, and attempts have been made to make word order the basic parameter in a holistic typology. In the present chapter, we will examine Greenberg's original work, then the attempts to generalize beyond his results, and finally some of the more recent critiques of such generalization. Although on occasion critical remarks will be directed at Greenberg's original contribution, it should be borne in mind that these are criticisms that can be made with hindsight, and in no way detract from the pioneering insights provided by Greenberg.

Although we retain the term word order typology, which has become established for referring to this area of typology, it should be noted that, strictly speaking, we are concerned not so much with the order of words as with the order of constituents, i.e. it would be more correct to speak of constituent order typology (cf. Greenberg's term 'the order of meaningful elements'). On the one hand, in saying, for instance, that a given language has subject – verb – object basic word order, it is irrelevant whether the constituents referred to consist of one or more words, so that this characterization applies equally to *John hit Mary* and to *the rogue elephant with the missing*

tusk attacked the hunter who had just noticed that his rifle was unloaded. Secondly, in addition to being concerned with the order of constituents that contain one or more words, we are also, in principle, interested in the order of morphemes less than a word, for instance in the relative order of affixes (prefixes, suffixes, infixes) and stems.

4.1 WORD ORDER PARAMETERS

This section examines the various major word order parameters that have been used in recent typological literature, in particular the order of the major constituents of the clause (subject, object, verb) and of the noun phrase, although other constructions are introduced where relevant. In typologizing a language on each of these parameters, we are concerned with the basic word order of the language in question. Although, in many instances, the assignment of a given basic word order to a language is unproblematical, there are also numerous instances where the assignment is more complex or even, perhaps, impossible. We will discuss examples of this as they arise.

The order of constituents of the clause is one of the most important word order typological parameters, indeed, as we will see in section 4.2, some linguists have made it into the major typological parameter. In its original form, this parameter characterizes the relative order of subject, verb, and object, giving rise to six logically possible types, namely SOV, SVO, VSO, VOS, OVS, OSV. As has already been noted in passing, in chapter 1, the distribution of these types across the languages of the world is heavily skewed in favour of the first three, more especially the first two, but we can now cite solidly attested examples of each of the first five basic word orders, and it is probably only a matter of time before reliable attestations of OSV languages become available:

Hasan öküz-ü aldı. Hasan ox ACCUSATIVE bought 'Hasan bought the ox.'	(Turkish: SOV)	(1)
The farmer killed the duckling.	(English: SVO)	(2)
Lladdodd y ddraig y dyn. killed the dragon the man 'The dragon killed the man.'	(Welsh: VSO)	(3)
Nahita ny mpianatra ny vehivavy. saw the student the woman 'The woman saw the student.'	(Malagasy: VOS)	(4)

Toto yahosiye kamara. (Hixkaryana: OVS) (5)
man it-grabbed-him jaguar
'The jaguar grabbed the man.'

Although, in the languages illustrated above, there is general agreement as to the basic word order, there are many languages where the situation is less clear-cut, and perhaps even languages where we are forced to say that, in terms of subject, object, and verb, there is no basic word order: such languages would then be irrelevant to word order typology on this parameter, reducing its range, but not its over-all validity. First, the parameter is only applicable to languages in which the grammatical relations of subject and object(s) exist, and, as we will see in more detail in chapter 5, there are many languages where the criteria identifying subjects seem to split across two noun phrases, thus making it difficult or impossible to specify the linear order of subject with respect to other constituents. Secondly, the parameter is only applicable to languages in which there is a basic word order determined, at least in part, by grammatical relations relative to the verb, and there are some languages where this seems not to be the case. For instance, in Dyirbal and Walbiri, all permutations of major constituents give rise to grammatical sentences, and if there is any preference for one word order over another, it is so slight as to be almost imperceptible. It should be noted that the problem with these languages is inability to determine a basic word order for the language as a whole. It is not just the case that certain limited constructions have a word order differing from that found elsewhere. If this were all that was involved, then we could agree to disregard such limited constructions in favour of the major sentence type in the language. Thus, when we classify English as being basically SVO, we abstract away from the fact that in special questions the word order of the *wh-* element is determined not by its grammatical relation, but rather by a general rule that places such elements sentence-initially, thus giving rise to such OSV orders as *who(m) did John see?* Even in many languages that are often described as having free word order, there is some good indication that one of the orders is more basic than the others. In Russian, for instance, any permutation of S, O, and V will give a grammatical sentence, but the order SVO is much more frequent than all of the other orders put together, and is moreover the preferred interpretation for sentences with the sequence NP – V – NP when the morphology, exceptionally, does not indicate which noun phrase is subject and which one is direct object (as in sentence (106) of chapter 3).

A further problem in assigning basic word order is where the language has a split, i.e. different basic word orders in different constructions. In some instances, this does not lead to undue difficulty in assigning basic

word order, where one of the word orders is clearly much more restricted than the other. Thus, the presence of special questions in English where the object precedes the subject does not seriously jeopardize the claim that English is a SVO language, and one can establish a general principle that word order of statements is more basic than that of questions (the more marked sentence type). In many languages, the order of pronouns is different from that of other noun phrases, so that in French, for instance, clitic object pronouns precede the verb, whereas other objects follow:

Le garçon a vu la jeune fille. (6)
'The boy has seen the girl.'

Le garçon l'a vue. (7)
'The boy has seen her.'

However, it is known that unstressed constituents, such as clitic pronouns, are often, cross-linguistically, subject to special positioning rules only loosely, if at all, relating to their grammatical relation, so sentences with pronouns can be discounted in favour of those with full noun phrases.

There are, however, examples of splits where no such ready solution is forthcoming. A classic example is from German, which has the word order SVO in main clauses but SOV in subordinate clauses:

Der Mann (NOMINATIVE) *sah den Jungen* (ACCUSATIVE). (8)
'The man saw the boy.'

Ich weiß, daß der Mann den Jungen sah. (9)
'I know that the man saw the boy.'

Controversy continues to rage over which, if any, of these word orders should be considered basic, with a certain tendency for analysts to split on ideological grounds: surface structure typologists tend to opt for main clause order as less marked, while transformational-generative grammarians tend to opt for subordinate clause word order as more basic. Since the parameter does not specify what kind of object is most relevant, a similar problem arises in languages like Kpelle, where the direct object precedes the verb but other objects follow:

È seŋ-kâu tèe kâloŋ-pɔ́. (10)
he money sent chief to
'He sent the money to the chief.'

Turning now to word order within the noun phrase, we may start with the relative order of adjective (A) and noun (N). Here, as with most of the following parameters, there are only two possibilities for basic order (if there is a basic order), namely AN and NA. The former is illustrated, for instance, by English *the green table* or Turkish *büyük şehir* 'large city'; NA order is illustrated by French *le tapis vert* 'the green carpet' or Welsh *llyfr bach* 'a little book'. The examples given here illustrate the basic, by far the most usual, order of adjective and noun in these languages, although in both French and Welsh it is possible for at least some adjectives to precede their noun, and in both languages there is a set of adjectives that usually precede, as in French *le petit prince* 'the little prince', Welsh *yr hen wlad* 'the old country'. It seems to be generally true that languages with the basic word order NA are more tolerant of exceptions of this kind than are languages with the basic word order AN (Greenberg's universal number 19): English examples like *court martial, envoy plenipotentiary*, are marginal, and often not felt synchronically to be sequences of noun and adjective.

Related to adjective–noun order, at least conceptually, is the order of head noun (N) and relative clause (Rel) in the relative clause construction. Again, there are two possible orders, either the head precedes the relative clause as in English, or the relative clause precedes the head as in Turkish:

adam-ın	*kadın*	*-a*	*ver*	*-diğ-i*	*patates*	(11)
man GENITIVE	woman DATIVE		give		his potato	

'the potato that the man gave to the woman'

For further discussion of relative clauses, including this Turkish example, reference should be made to chapter 7, where we will see that there is a further, third, possible order relation between head and relative clause, with the head internal to the relative clause. Although adjectives and relative clauses are similar conceptually, and indeed hard to separate from one another in some languages (e.g. Malay), in many languages they differ in word order: English is AN but NRel, for instance. In English, moreover, many heavy adjectival phrases have the same order as relative clauses, as in *people fluent in three languages*. This suggests that in characterizing languages as AN or NA, preference should be given to the order of simple adjectives than to that of more complex adjectival phrases.

Completing our list of constituents of the noun phrase is the relative order of possessive (genitive) (G) and head noun (N), again giving two possible orders, GN and NG. The former is illustrated by Turkish *kadın-ın çavuğ-u* 'the woman's chicken', literally 'woman-GENITIVE chicken-her'; the latter is illustrated by French *la plume de ma tante* 'the pen of my aunt' or Welsh *het y dyn* 'the man's hat', literally 'hat the man'. Although we

have not always illustrated problems caused by conflicting word orders within the noun phrase, we may do so here in discussing the characterization of English, which has two possessive constructions, the prenominal Saxon genitive, e.g. *the man's hat*, and the postnominal Norman genitive, e.g. *the roof of the house*. Although the Norman genitive is, textually, the more frequent of the two, and has become more frequent over the historical development of English, it is far from clear, for the modern language, whether one can specify that one of these two constructions is the basic order of head noun and possessive in English.

The last among the major word order parameters to be examined here is whether a language has prepositions (Pr), such as English *in the house* or Welsh *yn y tŷ* (same meaning), or postpositions (Po), such as Turkish *adam için* 'for the man'. The terminology of traditional grammar, though providing the two terms preposition and postposition, does not provide a single term to cover both of these, irrespective of order, and recent typological work has filled this gap by coining the term adposition. If we abbreviate this to Ap, then we can say that English has the order ApN (=PrN), while Turkish has the order NAp (=NPo). Most languages clearly have either prepositions or postpositions, though there may be occasional exceptions (thus Persian is basically prepositional, but has one postposition *-rā* for direct objects); however, there are also languages which are more mixed, such as Estonian, for which it is difficult to say, other than on the basis of slight statistical preponderance, whether the language is prepositional or postpositional. Most Australian languages have neither prepositions nor postpositions. Languages like Estonian and the Australian languages can thus be judged irrelevant, rather than counterexamples, to generalizations about prepositional versus postpositional languages.

Other parameters discussed, though less centrally, by Greenberg and figuring in some of his universals are the following. First, whether auxiliary verbs typically precede the main verb (as in English *will go*) or follow (as in Japanese *aisite iru* 'loves'). Secondly, whether in comparative constructions, the standard of comparison precedes the comparative (as in Turkish *Ankara'dan daha büyük* 'bigger than Ankara', literally 'Ankara-from more big'), or follows it (as in English *bigger than Ankara*); Finnish has both constructions here, the standard following when introduced by the conjunction *kuin* 'than' (e.g. *vanhempi kuin Helsinki* 'older than Helsinki'), but preceding when the standard is in the partitive case (e.g. *Helsinki-ä vanhempi*). Finally, we may distinguish between languages which are overwhelmingly suffixing as opposed to those which are overwhelmingly prefixing; while there are few good examples of the latter type, and probably none where a large number of prefixes can be added to

a given stem, there are some languages with long sequences of suffixes but virtually no prefixes, such as Turkish *bil-mi-yor-um* 'I do not know', literally 'know-NEGATIVE-PROGRESSIVE-1SINGULAR'.

4.2 CORRELATIONS AMONG WORD ORDER PARAMETERS

Most of the parameters listed in section 4.1 are logically independent of one another, for instance in that there is no a priori expectation that the presence of SOV basic word order in a language should correlate more or less well with the presence of AN rather than NA word order. Even in those instances where one might expect, a priori, there to be some correlations, as between AN order and RelN order (these are different kinds of attributive constructions), there are sufficient languages that do not have this correlation – such as English, with AN but NRel – to demonstrate that the correlation is far from necessary. Despite this, it turns out to be the case that there are many statistically significant correlations that can be drawn among these various parameters, and it is one of Greenberg's more specific merits, in addition to initiating general interest in this approach to language typology, to have established so many of these correlations. In section 4.2.1, we shall discuss some of Greenberg's correlations in more detail.

4.2.1 GREENBERG'S CORRELATIONS

Since Greenberg's proposed universals are gathered as an appendix to the work cited in the notes and references to this chapter, we will not simply repeat this list here, but rather state and comment upon some of the more salient of his results. The universals listed by Greenberg contain both absolute universals and tendencies, both non-implicational and implicational universals (though there are in fact more implicational than non-implicational universals – whence our characterization of them as correlations). Throughout, Greenberg's statements are very careful and cautious, based meticulously on his sample of languages and other languages from which he had relevant data. For instance, in the first universal, 'in declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object', the statement is as a (strong) tendency, rather than as an absolute, because Greenberg was aware of claims that certain languages have word orders violating this – although the data actually available to Greenberg were not always reliable on this score, we have cited examples above of reliably attested languages with VOS and OVS basic word order.

Another instance of Greenberg's care, especially in contrast to much later work, can be seen in the fact that he consistently avoids generalizing unilateral implications to bilateral implications, where the material does not justify doing so. Thus despite universal 27: 'if a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional', there is no corresponding universal that would say 'if a language is postpositional, then it is suffixing; if a language is prepositional, it is prefixing' – and this is clearly justifiable, since there are many languages like Huichol that have postpositions but also widespread prefixing, and like Persian that have prepositions but also widespread suffixing.

Thirdly, Greenberg does not take any one single parameter as being the basic determiner of word order typology, and again this caution is amply justified by the nature of the data. Thus word order in the clause is a good predictor of adposition order, at least for VSO languages (exclusively prepositional, by universal 3) and for SOV languages (overwhelmingly postpositional, by universal 4). However, it turns out that it is the order of adposition and noun that provides the best predictor for that of genitives, as per universal 2: 'in languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes'. One, and only one, of the three major basic clause order types gives a good prediction for adjective order: by universal 17, 'with overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun'.

Fourthly, many of the correlations are stated, where required by the data, not as holistic correlations across all parameters or as simple correlations involving only two of the parameters, but as complex correlations involving conditions among several parameters, as in universal 5, which correlates certain instances of clause order, genitive order, and adjective order: 'if a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun'. Perhaps the most extreme example of such a complex condition is universal 24: 'if the relative expression precedes the noun either as the only construction or as an alternative construction, either the language is postpositional or the adjective precedes the noun or both'.

In keeping with the general principles of the Dobbs Ferry Conference on Language Universals, at which Greenberg's paper was first presented, the emphasis at this stage of research was on establishing a wide range of language universals on a reliable cross-linguistic basis, with little or no attempt to find explanations, or farther-reaching generalizations, underlying these universals. Some of the individual universals proposed by Greenberg do have plausible, reasonably clear, explanations, although it was not the task of the original paper to explore this avenue. In section 1.3.3 we

suggested that the tendency for subjects to precede objects (universal 1) may be explainable in terms of the correlation between subject and agent, the correlation between object and patient, and the tendency for agents to be more salient perceptually than patients. Likewise, an explanation can readily be found for universal 15: 'in expressions of volition and purpose, a subordinate verbal form always follows the main verb except in those languages in which the nominal object always precedes the verb'. There are many instances where language has a tendency to mirror temporal order of events by linear order (e.g. in coordinate constructions like *John arrived and sat down*); a wish necessarily precedes its realization, a statement of purpose necessarily precedes its realization, therefore one would expect, other things being equal, that the main clause verb, expressing the wish or intention, would precede the subordinate verb, expressing the (potential) result thereof. The rider 'other things being equal' is necessary to account for the exception noted by Greenberg: if a language otherwise has a strict requirement of sentence-final main clause verb, then this can override the universal correlation of linguistic and temporal order.

Perhaps one of the reasons why many linguists following on from Greenberg's results have been less careful in their statements about correlations is that a large number of Greenberg's universals, however valid they may be as statements of limitations on cross-language variation, do not seem to lend themselves to any ready explanation or generalization. If one looks at universal 2, discussed above, for instance, then it is quite unclear why precisely the order of adpositions should play such a significant role in determining the order of a genitive relative to its head noun: there is no obvious conceptual link between adpositions and genitives, and one would hardly suspect a priori that adpositions would be the central parameter in a holistic typology of word order. Even a generalization of the kind 'adpositions and genitives tend to be placed on opposite sides of the noun' is little more than a formal restatement, with no indication of why this formal generalization should hold. Universal 24 above, in which prehead relative clause position predicts either postpositions or postnominal adjectives (or both) is perhaps the clearest example of a universal which is, in this sense, unintuitive.

Some of Greenberg's empirically ascertained universals do have plausible explanations, and these are the greatest factual merits of his list of universals: progress has been made both empirically and explanatorily. With the less intuitively plausible universals, however, one senses a certain tension between, on the one hand, empirical validity without a coherent conceptual system, and, on the other, plausible coherent conceptual systems which, however, lack empirical validity. This tension will play a major role in the following two sections.

4.2.2 GENERALIZATIONS OF GREENBERG'S RESULTS

In the appendix to his article on word order typology, Greenberg lists 24 logically possible types of language, based on the combinations of the four parameters VSO/SVO/SOV, Pr/Po, NG/GN, NA/AN; of these 24, 15 are actually attested in his sample or in other languages used by him in this piece of work. However, it is noticeable that the distribution of languages among these fifteen attested types is far from even. In fact, four types each contain far more languages than does any of the other eleven, as follows:

(a) VSO/Pr/NG/NA
(b) SVO/Pr/NG/NA
(c) SOV/Po/GN/AN
(d) SOV/Po/GN/NA

} (12)

On the basis of this observation, one might think that in order to establish universal tendencies, rather than absolute universals, of word order typology, it would be possible to work with just these types, neglecting the relatively few languages that fall into the other eleven attested types. If one makes this assumption, then a number of other generalizations seem to emerge from the four types listed above. First, except for the position of the subject in clause order, types (a) and (b) are identical. If one were to omit the subject from consideration, then types (a) and (b) could be combined into a single VO type; types (c) and (d) would then both be characterized as OV. Secondly, on most parameters, types (a) and (b) are precisely the inverse of types (c) and (d): the former are VO, Pr, NG, and NA; the latter are OV, Po, GN, and either AN or NA, the only embarrassment to this generalization being the widespread occurrence of NA basic order in OV languages. However, since we are working with tendencies, we might be prepared to overlook this complication, and work with only two major types in terms of word order, (e) and (f):

(e) VO, Pr, NG, NA
(f) OV, Po, GN, AN

} (12)

Some further support for this view might seem to come from the fact that VOS languages, not included in the original list, tend strongly to adhere to type (e). Data on OVS languages, and even more so OSV languages, which should behave like type (f), are less readily available – indeed, detailed infor-

mation is really available only for the one language Hixkaryana, which does more or less adhere to type (f), except that, to the extent that Hixkaryana can be said to have adjectives (and relative clauses), these follow the head noun.

Moreover, some of the other parameters discussed tend to correlate with this distinction into two types: type (f) tends also to have prenominal relative clauses, a strong tendency towards suffixing, auxiliary verbs after the main verb, and the standard of comparison before the comparative; while type (e) tends to have postnominal relative clauses, some tendency towards prefixing, auxiliary verbs before the main verb, and the standard of comparison after the comparative.

The kinds of generalization of Greenberg's results noted above are associated with two linguists in particular, Lehmann and Vennemann, and we will examine their contributions in turn. Lehmann argues, first, that the order of subject is irrelevant from a general typological viewpoint, so that we may indeed work with two major types of language, OV and VO. Two comments are in order here. First, as we shall see in more detail in section 4.2.3, while the collapsing of VSO and VOS into a single word order type seems reasonably justified – on other parameters, these two kinds of languages generally behave alike – the inclusion of SVO languages within this same type is questionable. In particular, while the existence of verb-initial word order or of SOV word order seems to correlate highly with various other typological parameters of word order, the existence of SVO word order does not seem to correlate particularly well with any other parameter. Knowing that a language is VSO or VOS, we can predict its values for other word order parameters; knowing that a language is SOV, we can with considerable reliability predict its other word order parameter values; knowing that a language is SVO, we can predict virtually nothing else. Secondly, there is potential terminological confusion in the use of the terms OV language and VO language, and it is essential to be aware of the particular use that each author makes. On the one hand, one could use these terms strictly to refer to the relative basic order of verb and object. On the other hand – and this is Lehmann's usage – one could use VO language to mean a language that has all or most of the word order properties of type (e) above, and OV language to mean a language that has all or most of the word properties of type (f) above. An actual example will make this clear. In Persian, the basic word order in the clause is SOV, so by the first use of the term OV language it would be an OV language. However, Persian has prepositions, postnominal genitives, postnominal adjectives, and postnominal relative clauses, so that under the second usage it would be a VO language, even though it does not actually have VO basic word order. To avoid the confusion, it is preferable, following Vennemann (see below) to refer to type (e) as operand–operator (or head–adjunct), and type (f) as operator–operand (or adjunct–head).

Lehmann also proposes a formal explanation, or rather generalization, of the observed correlations. He argues that V and O are primary concomitants of each other, and that modifiers are placed on the opposite side of a constituent from its primary concomitant. Thus, in a VO language, the primary concomitant of V is the postverbal O, so modifiers of V (in particular, auxiliary verbs) go to the left of V (AuxV); likewise, V is the primary concomitant of O, so modifiers of O (in particular, adjectives, relative clauses, and possessives) go to the opposite side from V, namely to the right. Conversely, in an OV language: the primary concomitant of V is the O to the left, so other modifiers follow the V (e.g. VAux); the primary concomitant of O is V, to the right, so other modifiers of O go to the left, i.e. adjectives, relative clauses, and possessors precede the object noun.

Apart from problems stemming from generalizing Greenberg's universals, to which we return in section 4.2.3, there are two other specific problems in this explanation. First, the explanation for order within the noun phrase applies strictly only to object noun phrases, and does not generalize directly to subjects or noun phrases in adverbials. One could presumably argue that the order is generalized from objects to other noun phrases, but if this were so one might expect to find languages where the order of constituents within the noun phrase was different for objects and other noun phrases, and such instances are either non-existent or rare. Secondly, the explanation, as is clear from Lehmann's exemplification, makes no distinction between modifiers which are expressed as separate words and those which are expressed as affixes. With regard to modifiers of verbs, this creates few problems, as there is a high correlation between having the auxiliary after the verb and having suffixes, and between having the auxiliary before the verb and having prefixes. With noun modifiers, however, to the extent that there is any correlation it is the reverse: certainly, across a wide range of operator-operand languages, e.g. Turkic languages, most Uralic languages, Quechua, Armenian, possessors precede their head noun, but possessive affixes are suffixed (see further section 10.3.2).

The explanation proposed by Vennemann for the correlations represented diagrammatically by types (e) and (f) above does not suffer from these disadvantages (in part in that it does not consider relations below the word level), although it, too, remains a formal explanation, without any further consideration of the question why this particular explanation should hold. Vennemann argues that in each of the construction types under consideration, i.e. the relation between verb and object (but not subject), between noun and adjective, etc., one of the constituents is an operator (corresponding to the traditional structuralist syntactic term adjunct), and the other the operand (corresponding to the traditional term head), the assignment being as in the following table:

OPERATOR	OPERAND
Object	Verb
Adjective	Noun
Genitive	Noun
Relative clause	Noun
Noun phrase	Adposition
Standard of comparison	Comparative adjective

The assignment of operator (adjunct) and operand (head) status is in most instances uncontroversial, though some linguists have been less comfortable with declaring the head of an adpositional phrase to be the adposition, rather than the noun (phrase). However, this assignment can be justified, for many languages, by the usual structuralist syntactic test of substitution: in English, for instance, the prepositional phrase of *John is in the house* can be substituted by *in* but not by *the house*, cf. *John is in*, but not, as a similar construction, **John is the house*, and the traditional term prepositional phrase attests to the view that the preposition is the head (just as the noun is head of a noun phrase, the verb of a verb phrase). For present purposes, at any rate, we may assume this assignment of operator and operand to individual constructions, bearing in mind that these assignments have been made by other linguists working independently of the particular correlations that Vennemann wishes to establish.

It is then clear what the general principle underlying types (e) and (f) above is: in languages of type (e), the operator is placed consistently after the operand, whence our suggestion above, following Vennemann, that they be called operand–operator languages; in languages of type (f), the operator consistently precedes the operand, thus giving rise to the operator–operand type. For languages which are typologically consistent in this regard, we need only specify whether they are operator–operand or operand–operator, and this one specification will then predict each of the individual word order parameter values. For languages which are inconsistent, i.e. which do not follow type (e) or (f), the language may be describable as being in general operator–operand or operand–operator, in terms of predominance of one or other ordering among the parameters, but even so special mention will have to be made of those parameters on which the language is exceptional. In the case of Persian, for instance, we would say that Persian is an operand–operator language (prepositions, postnominal adjectives, relative clauses, and possessors), but exceptionally it has OV word order. Thus the deviation of a language from one of the consistent types can be measured in terms of the number of special statements that need to be made about its word order.

We can now profitably contrast this position, or more specifically Vennemann's, with Greenberg's work discussed in section 4.2.1. Vennemann presents us with a schema that is conceptually very simple and very elegant; however, in order to establish this schema, certain liberties have to be taken with the data, as we will see in more detail in section 4.2.3. Greenberg's approach, on the other hand, is truer to the data, but ends up rather with a series of specific universals that do not fit together as a coherent conceptual whole.

4.2.3 CRITIQUE OF THE GENERALIZATIONS

The generalizations by Lehmann and Vennemann discussed in section 4.2.2 can clearly only stand, if they stand at all, as tendencies, since there are numerous counterexamples to them as absolute universals. Being satisfied with universal tendencies is sometimes necessary, forced upon one by the data, but it also brings with it the danger that one will cease to look further for absolute, or stronger, universals. One can see this even by a comparison of Greenberg's work with that of Lehmann or Vennemann. Greenberg did succeed in establishing some absolute universals, for instance the claim that VSO languages invariably have prepositions; however, within Vennemann's operator–operand schema, this exceptionless generalization is stated no differently from one with very low validity, for instance the correlation between SOV word order and adjective–noun order, since almost as many SOV languages have the adjective after the noun as have it before.

One way of comparing Vennemann's schema more directly with Greenberg's universals is to reformulate Vennemann's as a network of implicational universals, thus making them more directly comparable in form to Greenberg's. When reformulated in this way, it becomes clear that Vennemann's universal principles of word order are all expressed as bilateral implications: thus if it is the case that OV order predicts postpositions, then it is equally the case that postpositions predict OV order. Nearly all of Greenberg's universals, however, are unidirectional. Thus Greenberg is saying that some word order parameter values are good predictors for some other parameter values, but that this cannot be generalized to all parameters, some of which do not show any good correlation with anything else. Adjective order can serve as a good example here: from knowing that adjectives usually precede or follow the noun in a given language, one can tell virtually nothing about other word order parameter values. More generally, the over-generalization of Greenberg's universals with which we were concerned in section 4.2.2 fails to make any distinction as to the

reliability of individual implications, effectively treating them all as equivalent. Applied to Lehmann's generalization, the criticism would be rather different, because Lehmann does claim that one word order parameter is more important than any other, namely the relative order of O and V. But again, the earlier, more detailed work of Greenberg should have demonstrated that while clause word order is sometimes a good predictor of other word order parameter values, it is not always so (almost as many OV languages have postnominal as have prenominal adjectives), and there are many instances where other parameters have better predictive value (e.g. adpositions are good predictors of genitive order).

In reaction to the simplifying schema discussed in section 4.2.2, Hawkins has suggested that, even basing ourselves essentially on Greenberg's original data, it is possible to come up with a set of universals which are exceptionless, and which moreover are significant in that they tie together the various logically independent parameters used in word order typology. This can therefore be regarded as in some sense a compromise between Greenberg's position and that of Lehmann and Vennemann: further generalizations beyond those claimed by Greenberg are posited, but the claims are stronger than Lehmann's or Vennemann's in that they are said to have no counterexamples. On Vennemann's schema, for instance, over half the world's languages turn out to be exceptions, although, admittedly, some of them deviate only minimally from the norms of operator–operand or operand–operator languages, so that some kind of norm does still exist.

Hawkins's universals, like Greenberg's, are unilateral implications. However, they differ from most (not all) of Greenberg's universals by looking not just at correlations between two word order parameters, but rather at more complex implicational relationships using three or more parameters. The first proposed universals are reproduced as (13) and (14) below:

$$\text{SOV} \rightarrow (\text{AN} \rightarrow \text{GN}) \tag{13}$$

$$\text{VSO} \rightarrow (\text{NA} \rightarrow \text{NG}) \tag{14}$$

Let us make explicit the claims contained in these universals. First, it is claimed that clause order is a good predictor of certain other word order parameters, but only if the basic word order is SOV or VSO, i.e. SVO word order is not a good predictor, at least not in this case. (Word orders other than SOV, SVO, and VSO are not considered.) Once we know that a language has one of these two word orders, then we can make a further prediction, but this further prediction is itself in the form of an implication: if a language has SOV word order, then if it also has adjectives before the noun, it will necessarily also have genitives before the noun;

likewise, if a language has VSO word order, then if it also has adjectives after the noun, it will necessarily also have genitives after the noun. The two universals are clearly related formally to one another. It is possible to take these implicational universals and set out all the logical possibilities, then seeing which of these possibilities are in fact disallowed, although we will not carry out this task here. Suffice it to say that the excluded types are, by (13), SOV languages with AN and NG, and, by (14), VSO languages with NA and GN.

The second set of universals is given by Hawkins in two forms. First, there is a weak form, which does have some counterexamples:

$$\text{Pr} \rightarrow (\text{NA} \rightarrow \text{NG}) \tag{15}$$

$$\text{Po} \rightarrow (\text{AN} \rightarrow \text{GN}) \tag{16}$$

Like Greenberg, these claim that the difference between prepositional and postpositional languages can be a significant predictor of other word order parameters, effectively (given the close similarity between the implicata of (13) and (16), and (14) and (15)) as good a predictor as clause word order. The excluded language types are (a) those with prepositions, postnominal adjectives, and prenominal genitives – in Greenberg's list, Arapesh is a counterexample; and (b) languages with postpositions, prenominal adjectives, and postnominal genitives – in Greenberg's sample, some Daghestan languages are exceptions. The number of exceptions, relative to the overall sample, is very small, so that one might let (15)–(16) stand as universal tendencies. However, Hawkins notes further that the exceptions are all SVO languages. As we have indicated several times, SVO is a much less good predictor of other word order parameters than either of SOVor VSO, and we can build this observation into the universals by requiring as implicans, in addition to the appropriate kind of adposition, that the languages by either SOV or VSO, thus giving Hawkins's final formulations:

$$\text{Pr} \,\&\, (\text{VSO} \vee \text{SOV}) \rightarrow (\text{NA} \rightarrow \text{NG}) \tag{17}$$

$$\text{Po} \,\&\, (\text{VSO} \vee \text{SOV}) \rightarrow (\text{AN} \rightarrow \text{GN}) \tag{18}$$

Although these universals may look quite complex in their final formulation, the preceding discussion should have made it clear how these formulations are built up from more basic observations.

Within Hawkins's over-all view of word order typology, then, the above implicational universals would stand as absolute universals. Skewings in the distribution of languages among the permitted types could be described, as for Greenberg, by means of universal tendencies (which Hawkins slightly reformulates as distributional universals). But crucially, in

opposition to Lehmann and Vennemann, a distinction will be made between those absolute universals that rigidly delimit possible variation across languages, and those that are only tendencies as seen in skewings in the cross-linguistic distribution of attested types.

4.3 THE VALUE OF WORD ORDER TYPOLOGY

As we have emphasized at several points in this chapter, one of the main roles of word order typology in the recent study of language universals and typology has been methodological–historical: the work originated by Greenberg demonstrated that it is possible to come up with significant cross-linguistic generalizations by looking at a wide range of languages and without necessarily carrying out abstract analyses of these languages; in addition, there were a number of more specific methodological lessons, such as improvements in techniques for language sampling (see section 1.1.2). However, the question does arise as to just how far-reaching word order typology is in terms of the over-all typology of a language. In Greenberg's original work, relatively few correlations between word order and other parameters were drawn. In Vennemann's work, essentially no further correlations are drawn, and as we have seen even the elegance of Vennemann's account of over-all word order typology is in certain respects questionable. Hawkins's work demonstrates that if word order typology is to be rigorous, then it must forsake the extreme elegance of Lehmann's or Vennemann's schemata. At present, the main proponent of word order typology as the basis of a holistic typology is Lehmann, but it has to be acknowledged that, in addition to qualms about the degree of generalization made in his account of word order itself, most of the detailed correlations between word order and other phenomena, including even phonology, remain in need of establishment on the basis of data from a wide range of languages.

NOTES AND REFERENCES

The seminal paper by Greenberg referred to throughout this chapter is Greenberg (1966b); his universals are listed in Appendix III of this article (pp. 110–13). For discussion of Malagasy as a VOS language, see Keenan (1976a); for Hixkaryana as an OVS language, see Derbyshire (1977) and, more generally, Derbyshire (1979). The Kpelle example is from Givón (1975b, 50). On the lack of adpositions in Australian languages, see Dixon (1980, 272).

The discussion in section 4.2.2 is based on Lehmann (1973) and Vennemann (1972). Section 4.2.3 relates primarily to Hawkins (1980); Hawkins's universals are set out on page 203 of this article.

Some of Lehmann's more wide-ranging typological correlations are included in Lehmann (1978a); for some criticism, see Smith (1980).

See also the notes and references for chapter 10 for work on word order change. Li (1975) is a collection of articles relevant to both chapters 4 and 10.