OBJECT—VERB ORDER AND ADJECTIVE—NOUN ORDER: DISPELLING A MYTH

Matthew S. DRYER*

University of Alberta, Edmonton, Alberta, Canada

It is widely believed (Greenberg (1963), Lehmann (1973), Vennemann (1974)) that there is a relationship between the order of object and verb and the order of adjective and noun, that OV languages tend to be AdjN, while VO languages tend to be NAdj. Evidence against this is presented here from a sample of 316 languages.

1. The evidence

Consider first the following data for OV languages:[1]

	# Lgs	# Families	# Families (where dominant)
OV&AdjN	64	22	11
OV&NAdj	94	21	13
			(Neither 6)

Three types of data are given in the above table. The first column lists the total number of languages from my sample of each of the respective types. These figures show NAdj order to be somewhat more frequent among OV languages, although both orders are well-represented. There are severe methodological problems, however, with using such numbers. It can be shown that using raw language numbers can provide a distorted picture since

[1] The individual languages are listed by family in Appendix A.

a small number of language families that are well-represented in a sample can 'swamp' the rest of the sample. In such cases, the fact that languages in a sample with one property tend to have some second property may reflect accidental properties of one or two language families rather than a significant property of language in general.

One way of avoiding this problem is to construct a sample containing only one language from each family, as in Perkins (1980) and Bybee (1985). An alternative approach is to contruct a sample containing many languages per family, but then count families rather than languages. This second approach is employed here. The second column in the table above lists the number of families containing languages in my sample with the given order. Here we see a roughly equal number of the two kinds of languages. A somewhat more representative count is given in the third column of the table. The numbers here denote the number of families in which the given order of noun and adjective is dominant among the languages in the family, where one order is considered dominant if it is found in more than two thirds of the relevant languages in the family. In determining this figure, I calculate a *weighted proportion* of the languages in each family, assigning a weight to each language so that each parallel genetic grouping is counted equally.[2] This weighting procedure is described in detail in Appendix B. The figures in the third column of the above table point to the same conclusion as those in the first column: although NAdj order is slightly more common than AdjN order, the two orders are roughly equally common among OV languages. In 11 families the OV languages are primarily AdjN, while in 13 families the OV languages are primarily NAdj. In 6 families, neither order is dominant; in these families, both AdjN order and NAdj order make up at least $\frac{1}{3}$ of the OV languages in each family.

Although the figures given above show that both orders of adjective and noun are common among OV languages, they fail to reveal a strong areal pattern governing the distribution of the two language types. This becomes clear when we examine the particular families in which each order is dominant:[3]

---

[2] In calculating the weighted proportion, I assume the genetic classifications given in Voegelin and Voegelin (1977).

[3] Some of the families assumed in this paper are highly controversial, if not doubtful. I follow the classification of Voegelin and Voegelin (1977), in effect making them adjudicators in these disputes. But because of the controversial nature of many of these groups, the precise numbers should not be taken too seriously. I believe that we can only draw reliable inferences when the data exhibits very strong statistical tendencies, and in such instances, the conclusions are

Among OV languages:

Families where AdjN dominant: Afro-Asiatic, Austroasiatic, Burushaski, Caucasian, Chukchi-Kamchatkan, Dravidian, Indo-European, Nivkh (Gilyak), Ural-Altaic, Yeniseian, Yukaghir.

Families where NAdj dominant: Australian, Austronesian, Basque, Eskimo-Aleut, Ge-Pano-Carib, Indo-Pacific, Macro-Algonkian, Macro-Siouan, Niger-Kordofanian, Nilo-Saharan, Sino-Tibetan, Sumerian, Yukian.

Families where neither dominant: Andean-Equatorial, Aztec-Tanoan, Hokan, Macro-Chibchan, Na-Dene, Penutian.

This data is summarized in the following table, distinguishing those families spoken in Eurasia from those spoken outside Eurasia.[4]

Number of families by area

	Eurasia	Non-Eurasia
AdjN dominant among OV lgs	10	1
NAdj dominant among OV lgs	3	11
Neither order dominant	0	6

This table shows that in Eurasia, the OV languages are predominantly AdjN. The three families that do not conform to this are Sino-Tibetan along with the two isolates Sumerian and Basque, the latter on the periphery of the area. Conversely, outside Eurasia, the OV languages are overwhelmingly NAdj. In only one out of 18 language families outside Eurasia are the OV languages predominantly AdjN. This one family is Afro-Asiatic, the OV languages in question being spoken in the area around Ethiopia and Somalia, just outside Eurasia. Although I have noted the dangers of using raw languages numbers, we find similar results hold when we count languages:

---

sufficiently robust that they do not appear to depend on any questionable genetic assumptions. Furthermore, the data in this paper reveals strong areal patterns, which for sampling purposes are just as significant as genetic considerations. See Appendix B for a discussion of some of these issues.

[4] By Eurasia, I mean continental Eurasia; I thus exclude Austronesian and Indo-Pacific languages spoken in Indonesia, the Philippines, and New Guinea. I treat Afro-Asiatic as non-Eurasian, mainly because the OV Afro-Asiatic languages are spoken in Africa. Treating Afro-Asiatic as Eurasian would have sharpened the areal pattern, since it is the only non-Eurasian family in my sample in which the OV languages are predominantly AdjN.

Number of languages by area

	Eurasia	Non-Eurasia
OV&AdjN	36	28
OV&NAdj	11	83

This table illustrates that among the languages in my sample, 77% of the OV languages in Eurasia employ AdjN order, while only 25% of the OV languages outside Eurasia do so. In short, the previously believed tendency for OV languages to be AdjN is simply an Asian areal phenomenon, and the area exhibiting the phenomenon is an extension of one that Masica (1976) has shown exhibits other examples of areal phenomena. Furthermore, the fact that OV languages outside this area do not exhibit a tendency to be AdjN that the use of AdjN order throughout the OV languages of Eurasia cannot have been triggered independently by OV order, but must itself have spread as a result of contact, or must be an instance of a more general typological parameter which has spread through contact.[5]

It must be emphasized, however, that none of the evidence cited so far demonstrates that there is not a tendency for OV languages to be AdjN, at least in terms of the claims of Greenberg (1963). Greenberg suggested that the order of adjective and noun is governed not only by a *harmony* principle whereby the order of object and verb tends to be harmonic with the order of adjective and noun, but also by a *dominance* principle whereby there is a general preference for the adjective to follow the noun. Thus, in his 30-language sample, there is a roughly equal number of OV&AdjN and OV&NAdj languages:

Data from Greeberg's 30-language sample

	AdjN	NAdj
OV	6	5
SVO	5	8
V-initial	0	6

Greenberg suggested that the high number of OV&NAdj languages reflects the conflict between the harmony principle and the general dominance of NAdj

---

[5] It is possible that the tendency for the OV languages in Eurasia to be AdjN is partly due to remote genetic relationships. There is evidence (Joseph Greenberg, p.c.) that is highly suggestive of a remote genetic relationship between many of the families of northern Eurasia, in particular Indo-European, Altaic, Uralic, Yukaghir, Nivkh (Gilyak), and Chukchee-Kamchatkan. However, given the extent to which less remote groupings exhibit word order variation, any retention of word order characteristics after such a length of time would have to have been reinforced by contact.

order. Thus, despite the data given for OV languages, there might still be a harmony principle linking the order of object and verb to the order of adjective and noun, if, as seems to be the case in Greenberg's 30-language sample (at least for V-initial languages), VO languages exhibit a significantly stronger preference for NAdj order. Hence, to show that there is no such harmony principle, one must show that VO languages do not exhibit any such stronger preference for NAdj order.

Following Greenberg's practice, I will examine SVO languages and V-initial languages separately. Consider first SVO languages:

	#Lgs	#Families	#Families (where dominant)
SVO&AdjN	23	8	5
SVO&NAdj	67	13	11
			(Neither 1)

We find here at best a weak preference for NAdj order among the SVO languages in my sample. The particular families are as follows:

Among ~~V-initial~~ SVO languages:

Families where AdjN dominant: Austr~~al~~ alian, Indo-European, Macro-Algonkian, Penutian, Ural-Altaic.
Families where NAdj dominant: Afro-Asiatic, Andean-Equatorial, Austro-Asiatic, Austronesian, Aztec-Tanoan, Ge-Pano-Carib, Indo-Pacific, Khoisan, Niger-Kordofanian, Nilo-Saharan, Oto-Manguean.
Families where neither order dominant: Sino-Tibetan.

Here we do not find any strong areal pattern, although in all four families of Africa, the SVO languages tend to be NAdj. This latter fact is an instance of an apparent pan-African tendency to place modifiers after the noun regardless of the order of verb and object.

The data for V-initial languages is as follows:

	#Lgs	#Families	#Families (where dominant)
V-initial&AdjN	15	9	5
V-initial&NAdj	24	9	4
			(Neither 4)

This data is quite at odds with Greenberg's Universal 17: "With overwhelmingly greater than chance frequency, languages with dominant order VSO have the adjective after the noun." Greenberg's universal refers specifically to VSO languages, but there is no evidence that VSO languages behave differently from other V-initial languages, either VOS languages or V-initial languages which are neither clearly VSO nor clearly VOS (cf. Keenan (1978), Hawkins (1983)). Greenberg's universal is apparently based on the fact that in his 30-language sample, all 6 V-initial languages are NAdj, and on the fact that a preference for NAdj order is reflected among the majority of the V-initial languages in his appendix. But although the number of NAdj languages in the above table is slightly higher than the number of AdjN languages, this disappears when we count families instead of languages: in fact, we find slightly more families preferring AdjN order, by 5 to 4. And there are four additional families in which neither order is dominant. Hence only 4 out of 13 families with V-initial languages conform to Greenberg's Universal 17. The individual families in question are as follows:

Among V-initial languages:

Families where AdjN dominant: Chimakuan, Hokan, Penutian, Salish, Wakashan.

Families where NAdj dominant: Afro-Asiatic, Andean-Equatorial, Aztec-Tanoan, Nilo-Saharan.

Families where neither dominant: Australian, Austronesian, Indo-European, Oto-Manguean.

Here again, a clear areal pattern emerges. Four of the five families in which the V-initial languages are predominantly AdjN are spoken in the northwest of North America: Chimakuan, Penutian, Salish, and Wakashan.[6] The number of families cited above is thus rather misleading, and given the greater areal diversity of the four families in which the V-initial languages are predominantly NAdj, one might argue that V-initial languages do exhibit a weak preference for NAdj order. But there is no evidence that this preference is any stronger among V-initial languages than it is among SVO languages or OV languages. In fact, outside of Eurasia, OV languages seem to exhibit, if anything, a stronger preference for NAdj order than either SVO or V-initial

---

[6] Penutian languages are spoken outside this area, but it is the V-initial languages from this area in my sample that make Penutian predominantly AdjN among the V-initial languages.

languages. In short, there is no evidence of any relationship between the order of verb and object and the order of adjective and noun.

In the remainder of this paper, I will discuss a number of independent points related to the above empirical results, in terms of both their theoretical implications and areas for future research.

2. Theoretical implications

A widely accepted view of the fundamental principle underlying the Greenbergian word order correlations is that they reflect a tendency towards consistent ordering of head and dependent.[7] But this theory is based, at least in part, on the assumption that there is a correlation between the order of object and verb and the order of adjective and noun. The fact that there is no such correlation thus presents a problem for that theory.

In Dryer (in press), I propose an alternative theory for what underlies the Greenbergian word order correlations:

Branching Direction Theory

Languages tend toward consistent left-branching or consistent right-branching. I.e. languages tend toward one of two ideals, one in which branching categories *precede* nonbranching categories, the other in which branching categories *follow* nonbranching categories.

In the context of modifiers of nouns, this theory only requires that we should find correlations between the order of object and verb and the order of modifier and noun for modifiers which involve branching categories. This accounts for the fact that we do not find a correlation with the order of adjective and noun since adjectives are single words,[8] but predicts that we

---

[7] This theory is implicit in Hawkins' (1983) Cross-Category Harmony Principle, and is explicitly defended by Hawkins (1984). See section 6 below for a discussion of Hawkins' Cross-category Harmony Principle in relation to the claims of this paper.

[8] This ignores adjective phrases that consist of more than one word, as in the English examples in (i) to (iii).

(i)   a *very tall* man
(ii) a *tall and thin* man
(iii) a man *taller than Bill*

Adjective phrases like those in (i) are very limited in their branching in that only one or two words can be added (cf. *much more interesting*). There is no limit in principle to forming adjective phrases

should with the order of genitive and noun and that of relative clause and noun, since genitives and relative clauses are both branching categories, genitives being NP's or PP's, relative clauses being clauses. Thus in VO languages, in which the branching object follows the nonbranching verb, the Branching Direction Theory predicts that the branching categories genitive and relative clause will tend to follow the nonbranching noun. A language possessing each of these word order characteristics will be consistently right-branching with respect to these constructions:[9]
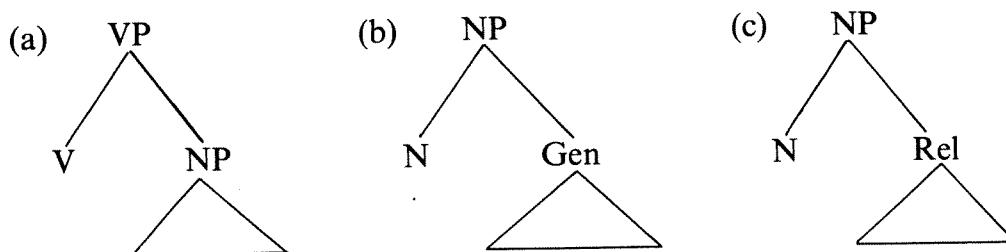
And such a language will still be consistently right-branching whether it is AdjN or NAdj, since neither the noun nor the adjective are branching categories.

The predictions that the Branching Direction Theory makes for genitives and relative clauses are borne out. Table 1 shows that OV languages exhibit a strong tendency to be GenN, and that V-initial languages exhibit a strong tendency to be NGen, while both orders are common in SVO languages.[10]

---

by conjoining adjectives, as in (ii), but the branching is still limited in the sense that adjective phrases of this sort can only contain adjectives (or adjective phrases of the sort in (i) and (ii)). Adjective phrases like those in (iii) are not limited in this way, since they can contain any other category; they feed into the rest of the grammar, so to speak. In making the Branching Direction Theory precise, I would restrict the notion *branching category* to fully recursive branching categories like the adjective phrase in (iii). But adjective phrases of this sort do not appear to be very common cross-linguistically as modifiers of nouns. Thus in most languages, it appears to be the case that adjective phrases modifying nouns are not fully recursive branching categories. Further research is required to determine whether this is really true.

[9] The Branching Direction Theory would also seem to imply that VO languages should be VS. But I assume, along with Hawkins (1983) and Keenan (1979), that there is a separate dominance principle, independent of the harmony principle that underlies the Greenbergian word order correlations, whereby subjects tend to occur in clause-initial position. Such a principle is described by Keenan as 'Subjects Front'. Functional considerations underlying such a principle are discussed in Tomlin (1986).

[10] The figures for 'Families where dominant' represent the number of families where the given order of genitive and noun is dominant among those languages in the language with the given order of clausal constituents.

Table 1

	# Lgs	# Families	# Families (where dominant)
OV&GenN	139	24	21
OV&NGen	11	7	2
			(Neither 2)
SVO&GenN	33	12	7
SVO&NGen	60	9	5
			(Neither 2)
V-initial&GenN	5	4	1
V-initial&NGen	44	13	12
			(Neither 1)

The order of relative clause and noun shows a similar pattern, although less clearly (table 2).

Table 2

	# Lgs	# Families	# Families (where dominant)
OV&RelN	29	10	8
OV&NRel	32	15	13
			(Neither 1)
SVO&RelN	3	1	0
SVO&NRel	61	12	11
			(Neither 1)
V-initial&RelN	0	0	0
V-initial&NRel	28	8	8

We see here a slight preference for NRel order among OV languages, but this seems to be due to the effects of a strong dominance principle whereby NRel is preferred over RelN.[11] This view is supported by the extreme rarity of VO languages that are RelN. We can represent this situation schematically as in figure 1.

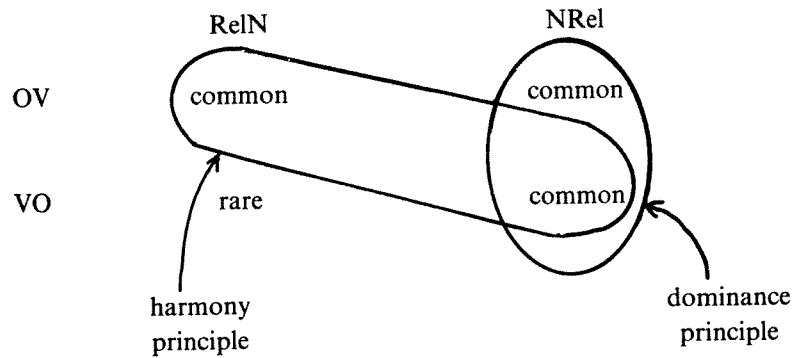[11] See note at the end of this section.

Fig. 1.

The rarity of VO&RelN languages can be attributed, on this approach, to the fact that they conform to neither the dominance principle nor the harmony principle. The three types which are more common all conform to at least one of the two principles. In short, once the dominance principle that involves the preference for NRel order is controlled for, the data given here provides strong evidence that there is a harmony principle whereby OV languages are RelN and VO languages are NRel, or equivalently, that the order of object and verb correlates with the order of relative clause and noun.

It should be emphasized that these harmony and dominance principles are intended as *descriptive* principles, motivated by the relative frequency of the different language types, and are not intended to be themselves explanatory. I assume, however, that they are amenable ultimately to deeper explanations. The claim that there is a dominance principle favouring NRel order is thus simply a claim that NRel languages are significantly more common than RelN languages. The claim that there is a harmony principle preferring OV&RelN and VO&NRel (or equivalently that there is a correlation between the order of object and verb and that of relative clause and noun) is simply a claim that the proportion of OV languages that are RelN is significantly higher than the proportion of VO languages that are RelN.[12]

---

[12] John Hawkins (p.c.) questions the appropriateness of my use of the term *correlation* in instances such as these in which only one of the four possible types is rare, in which the facts can be described naturally a 'unilateral' implicational universal, but not a 'bilateral' one, suggesting that the term should be restricted to instances where two of the four possible types are common, and two rare, instances in which the facts can be described by a bilateral implicational universal. My use of the term is based on my assumption that even in instances in which three of the four types are common, the situation is best analysed in terms of the interaction of a harmony principle (which is like a bilateral implicational universal) with some other principle, like a dominance principle. Hence the high frequency of one of the types is to be understood as the result of some

The evidence above shows that although there is no correlation between the order of object and verb and the order of adjective and noun, there is a correlation between the order of object and verb and the order of modifier and noun for modifiers which involve branching categories. This difference between adjectives and modifiers that involve branching categories thus provides evidence for the Branching Direction Theory over the view that languages tend towards consistent ordering of head and dependent.

Note. It is worth noting that the split between the two types of OV languages, those which are RelN and those which are NRel, also follows an areal pattern similar to that exhibited by the order of noun and adjective:

Number of families by area

	Eurasia	Non-Eurasia
RelN dominant among OV lgs	6	2
NRel dominant among OV lgs	2	11
Neither order dominant	0	1

The OV languages of Eurasia are generally RelN, while those outside Eurasia are generally NRel. Interestingly, the families that did not fit the areal pattern with adjectives are not always the same as those that do not fit the pattern with relative clauses: two of the three OV&NAdj families of Eurasia, Sino-Tibetan and Basque, are RelN. Despite this areal pattern, we are still justified in claiming that there is a harmony principle whereby OV languages tend to be RelN, VO languages NRel, because VO&RelN languages are so rare. The three VO&RelN languages listed are all so-called 'dialects' of Chinese.

---

other principle which 'overrides' the harmony principle (and correlation) in certain instances. At any rate, I define my use of the term *correlate* as follows. Given two pairs of elements $\{X_1, X_2\}$ and $\{Y_1, Y_2\}$, I will say that the order of X_1 and X_2 correlates with that of Y_1 and Y_2 if the proportion of languages in which Y_1 precedes Y_2 is significantly greater among languages in which X_1 precedes X_2 than it is among languages in which X_2 precedes X_1. By this definition, the order of relative clause and noun does correlate with the order of verb and object. This use of the term is consistent with its normal use. For example, we can say that there is a correlation between smoking and lung cancer if the proportion of people who get lung cancer is significantly greater among smokers than it is among nonsmokers, even if only a minority of smokers get lung cancer.

3. The possibility that some languages are head-ordering languages

Despite the apparent lack of correlation between the order of object and verb and the order of adjective and noun, there exists a possibility that there may be a subtype of language in which the order of object and verb and the order of adjective and noun are related. Consider, for example, the order of adjective and noun among Afro-Asiatic languages. The OV languages in this family are predominantly AdjN (the exceptions in my sample being restricted to a subset of the Cushitic languages), while the VO languages in my sample are all NAdj. Or consider the fact that the Celtic languages are V-initial and NAdj while the majority of other Indo-European languages are AdjN and not V-initial. Or consider the fact that the Mon-Khmer languages, which are consistently VO, are also consistently NAdj, while the majority of languages of Asia are OV and AdjN. If there is no relationship between the order of object and verb and the order of adjective and noun, then these facts are simply coincidental. An alternative possibility, however, is that there exists a subset of the languages of the world in which the word order rules *are* sensitive to head and dependent, languages we might call *head-ordering* languages. Exactly what might be the defining characteristic of this class of languages is not entirely clezar, but there might be some fundamental typological parameter that is shared by the languages of Eurasia due either to extended contact or to extremely remote genetic relationships. If such a typological property exists, and if it is possessed by the Afro-Asiatic languages, the Celtic languages, and the Mon-Khmer languages, then the properties of these languages that were noted above would not be coincidental, despite the evidence presented in section 1, showing that there is no overall relationship between the order of object and verb and the order of adjective and noun.

It turns out, in fact, that in other areas of the world, one often finds the opposite pattern from that found in Eurasia. Among the languages of Australia, for example, the OV languages in my sample are overwhelmingly NAdj (15 out of 16), while the VO languages are overwhelmingly AdjN (6 out of 7). A similar pattern is found in the languages of North America north of the US-Mexican border (but exluding languages in the Aztec-Tanoan and Hokan families, these families being ones that straddle the border). Of the languages in my sample from this area, 15 of the 17 OV languages are NAdj, while all 10 of the VO languages are AdjN. Thus both Australia and North America exhibit the opposite relationship between the order of adjective and noun and the order of object and verb from that previously believed to be universal. Although these facts may be coincidental, it is also possible that

they reflect some fundamental difference between the languages noted and those of Eurasia, and that the languages of Australia and North America possess some property that leads them to exhibit the opposite correlation from that found in the languages of Eurasia.[13]

4. Types of 'adjectives'

In collecting data for this paper, I have followed the implicit practice of Greenberg (1963) in employing essentially semantic criteria in identifying adjectives. Namely, a word is identified as an adjective if it expresses the kind of meaning associated with descriptive adjectives in English. It is clear, however, that in many languages, the words in question do not constitute a word class in the way they do in English. In many languages, the words in question are simply verbs; in other languages they are nouns. The use of the term *adjective* in word order studies is thus potentially misleading, and the expression *property term* might more accurately convey the fact that it is a semantic category rather than a grammatical category. Not only are there languages in which there is no grammatical class identifiable as adjectives, but there are also languages (described in great detail by Dixon (1977)) in which there is an adjective class, but it is a closed class of items, often with a relatively small number of members. In many of these languages, some property terms are adjectives while others are verbs.

It is possible that the way in which a language categorizes property terms, whether by an open adjective class, by a closed adjective class, by verbs, or by nouns, may provide a basis for more accurately predicting the order of 'adjective' and noun, given other word order characteristics. It is possible, for example, that in languages in which adjectives form a small closed class of items, the order of adjective and noun may exhibit an inverse correlation with the order of object and verb. Conversely it seems plausible that in languages in which property terms are verbs, the order of property term and noun will correlate with the order of object and verb, since when such property terms

---

[13] Munro (1985) suggests that some languages that are OV and NAdj may be 'predicate-final' languages. Extending this idea, it may be that in addition to a class of head-ordering languages, there is also a class of 'predicate-ordering' languages, of which there would be two types, 'predicate-initial' and 'predicate-final'. The verb-initial AdjN languages of northwest North America are possible candidates for the status of 'predicate-initial' languages. A variant of this idea is that some languages are 'function-argument' languages, with consistent ordering of funtion and argument.

modify nouns, they may be formally relative clauses, and, as shown above, the order of relative clause and noun does correlate with the order of object and verb. However, Van Valin and Garcia (1985) suggest exactly the opposite, that OV languages in which adjectives are verbs tend to be NAdj. Clearly, this area requires further research.

5. Other word order parameters

I have shown in section 1 that there is no correlation between the order of adjective and noun and the order of object and verb. The question arises, however, whether there might be correlations between the order of adjective and noun and other word order parameters.

Hawkins (1983), for example, argues that whether a language is prepositional or postpositional is in general a better predictor of other word order characteristics than is the order of object and verb. Although he does not make predictions for pairwise correlations of the sort discussed here, it is worth examining the relationship between adposition type and the order of adjective and noun. The data for this from my sample is given in table 3.

Table 3

	#Lgs	#Families	#Families (where dominant)
Po&AdjN	65	21	12
Po&NAdj	77	17	13
			(Neither 4)
Pr&AdjN	29	10	4
Pr&NAdj	76	14	9
			(Neither 4)

This data is similar to that given in section 1 for the order of object and verb, but less unequivocal. One might argue that the data exhibits a weak correlation between adposition type and the order of adjective and noun, since NAdj order seems to be somewhat more common than AdjN order among prepositional languages. However, again, one must consider the areal pattern, which resembles that discussed above for the relationship between the order of object and verb and the order of adjective and noun.

Among Postpositional languages:

Families where AdjN dominant: Afro-Asiatic, Austroasiatic, Burushaski, Caucasian, Chimakuan, Chukchi-Kamchatkan, Dravidian, Hokan, Indo-European, Nivkh (Gilyak), Ural-Altaic, Yeniseian.

Families where NAdj dominant: Australian, Austronesian, Basque, Ge-Pano-Carib, Indo-Pacific, Khoisan, Macro-Algonkian, Macro-Chibchan, Macro-Siouan, Niger-Kordofanian, Nilo-Saharan, Sino-Tibetan, Sumerian.

Families where neither dominant: Andean-Equatorial, Aztec-Tanoan, Na-Dene, Penutian.

Among Prepositional languages:

Families where AdjN dominant: Macro-Algonkian, Penutian, Salish, Wakashan.

Families where NAdj dominant: Afro-Asiatic, Andean-Equatorial, Austroasiatic, Austronesian, Aztec-Tanoan, Ge-Pano-Carib, Indo-Pacific, Niger-Kordofanian, Nilo-Saharan.

Families where neither dominant: Australian, Indo-European, Oto-Manguean, Sino-Tibetan.

9 of the 12 families in which AdjN order is dominant among the postpositional languages are spoken in Eurasia. Hence postpositional languages outside Eurasia tend, if anything, to be NAdj. One might argue that Prep&AdjN order is sufficiently rarer than Post&AdjN order to provide a basis for claiming that adposition type and order of adjective and noun are not independent. 3 of the 4 families in which AdjN order is dominant among the prepositional languages are spoken in northwest North America, and the fourth family is Macro-Algonkian, based solely on Yurok, spoken on the periphery of the northwest, in northern California. On the other hand, Penutian, one of the families in this category, is a controversial group, and two of the 'Penutian' languages in question are spoken outside that immediate area: Yokuts in central California and Totonac in Mexico. In addition, there are four other families in which neither order of noun and adjective is dominant among the prepositional languages,[14] so that in only 9 out of 17, or barely half, of the families with prepositions, is NAdj order dominant. In short, the data is not unequivocal; there is an apparent trend towards a weak correlation between adposition type and order of adjective and noun, but the

---

[14] One of these families is Indo-European, English being one of the Prep and AdjN languages.

tend is sufficiently weak as to be within the realm of accident. The evidence from adposition type thus provides no clear basis either for or against the theory that the Greenbergian word order correlations reflect a tendency towards consistent ordering of head and dependent.

When we turn to the relationship between the order of genitive and noun and that of adjective and noun, the evidence is again equivocal (table 4).

Table 4

	#Lgs	#Families	#Families (where dominant)
GenN&AdjN	77	22	11
GenN&NAdj	91	18	9
			(Neither 8)
NGen&AdjN	19	10	4
NGen&NAdj	83	13	10
			(Neither 3)

The pattern here is fairly similar to that with adposition type, although here the correlation is a little stronger. Nevertheless, even if it is significant, it does not provide evidence for any cross-categorial tendency towards consistent ordering of head and dependent, since genitives and adjectives are both dependents of the noun. Thus it is possible that *within* individual categories like the noun phrase, there is a tendency towards consistent ordering of dependent and head noun, but that there is no tendency toward consistent ordering of dependent and head *across* categories.

When we turn to the order of relative clause and noun, we find that it does clearly correlate with the order of adjective and noun (table 5).

Table 5

	#Lgs	#Families	#Families (where dominant)
RelN&AdjN	28	11	11
RelN&NAdj	4	3	1
NRel&AdjN	27	9	4
NRel&NAdj	92	20	15

The numbers are skewed by the general preference for NRel order, but there is a clear tendency for RelN languages to be AdjN and for NRel languages to

be NAdj. This correlation is due partly to the fact that in many of these languages, adjectives are really verbs and hence formally relative clauses when they modify nouns. In fact, in the absence of data on whether adjectives are verbs in the languages in question, the correlation might be due *entirely* to this factor. But even if it is not (and I suspect that it is not), the correlation would be attributable to the close functional similarity of adjectives and relative clauses, which both, loosely speaking, restrict the reference of the noun to a subset of the set defined by the noun.[15]

The evidence cited so far does not provide any (unequivocal) evidence that the order of adjective and noun correlates with any other pairs of dependent and head except for pairs in which the head is the noun. The order of adjective and noun does, however, correlate with the order of manner adverb and verb (table 6).

Table 6

	# Lgs	# Families	# Families (where dominant)
AdvV&AdjN	35	14	8
AdvV&NAdj	37	13	7
			(Neither 5)
VAdv&AdjN	8	5	1
VAdv&NAdj	55	12	10
			(Neither 2)

Although both orders of adjective and noun are common among AdvV languages, VAdv languages are generally NAdj. This pattern can be described by a statistical implicational universal 'If VAdv then NAdj'. The situation is thus somewhat analogous to the relationship between the order of object and verb and the order of relative clause and noun discussed in section 2 above. Namely, we might characterize the situation in terms of the interaction of dominance and harmony principles, as in figure 2.

---

[15] There is a small number of descriptive adjectives in English for which this generalization does not hold, such as the adjective *fake*: the set of *fake guns* is not a subset of the set of *guns*.
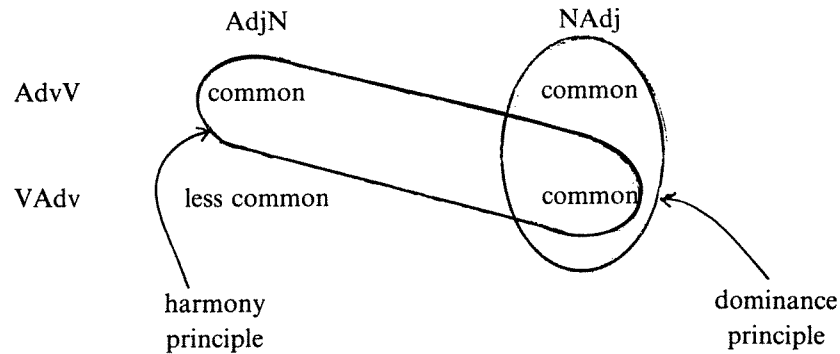
Fig. 2.

The dominance principle here would be the weak preference for NAdj order over AdjN order, a preference we saw exhibited in OV, SVO, and V-initial languages in section 1. This preference may, however, be too weak to account for the rarity of VAdv&NAdj languages. Nevertheless the harmony principle is more clearly motivated since the proportion of AdvV languages that are AdjN is significantly greater than the proportion of VAdv languages that are AdjN. This harmony principle is clearly cross-categorial, since it involves a correlation between order in the VP and order in the NP.

Even clearer evidence for a correlation between the order of adverb and verb and the order of adjective and noun is found if we restrict our attention to VO languages (table 7).

Table 7

	#Lgs	#Families	#Families (where dominant)
VO&AdvV&AdjN	7	6	6
VO&AdvV&NAdj	0	0	0
VO&VAdv&AdjN	8	5	1
VO&VAdv&NAdj	46	11	8
			(Neither 3)

We have evidence here for an exceptionless implicational universal

If VO and AdvV then AdjN.[16]

---

[16]   Since performing the computations for this paper, I have discovered that Jicaltepec Mixtec, a language not in the sample used for this paper, is an apparent exception to this universal.

This universal can be explained as resulting from the interaction of two harmony principles, as in figure 3.

AdjN   NAdj

VO&AdvV   (not rare)   unattested

VO&VAdv   (not rare   very common)

            harmony   harmony
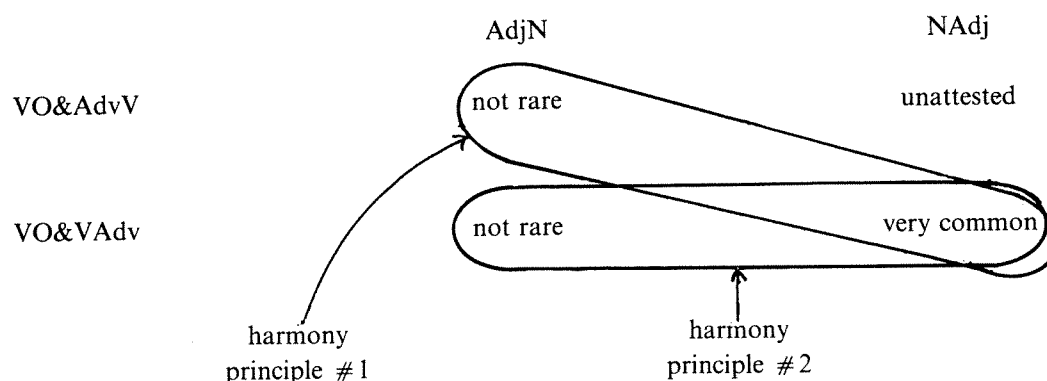   principle #1   principle #2

Fig. 3.

Harmony principle #1 is one whereby the order of adverb and verb tends to mirror the order of adjective and noun. Harmony principle #2 is a non-crosscategorial one whereby dependents of a given head tend to occur on the same side of that head. The most common language type is the one which satisfies both harmony principles. The unattested type satisfies neither principle.[17]

But why should the order of adjective and noun correlate with the order of adverb and verb but not with the order of object and verb? A plausible answer is that adverbs *modify* verbs just as adjectives modify nouns, but that objects do *not* modify verbs, at least in the traditional sense of the term *modify*. Hawkins (1983, 1984) uses the term *modifier* in a more general sense, equivalent to my use of the term *dependent*. But the facts cited in this paper suggest that the narrower, traditional notion of modifier is useful. That notion seems to be definable in terms of the interaction of head-dependent structure and function-argument structure. Namely, X modifies Y if X is a dependent of Y and X is a function that takes Y as its argument, or, more generally, if X is a function that takes Z as its argument where Z contains Y. (This more general version is necessary for cases like *John kissed Mary gingerly* where *gingerly* is a dependent of *kissed* but arguably takes *kissed Mary* as its argument rather than just *kissed*.) On the most natural approaches, a manner adverb is semantically a function taking a verb (or verb phrase) as argument, whereas in the verb–object combination, the verb is the function and the object is the argument. Thus, even though there may be no

---

[17] See note at the end of this section.

evidence for a cross-categorial tendency towards consistent ordering of head and *dependent*, there may be a tendency towards consistent ordering of head and *modifier*.

Similar evidence is available for a correlation between the order of *qualifier* and adjective and the order of adjective and noun. By a qualifier I mean a function word (characterized sometimes in generative work as a degree word and in traditional grammar – misleadingly – as a type of adverb) that qualifies or intensifies an adjective. Among the qualifiers in English are *very, more,* and *somewhat*. The following data show that although both orders of noun and adjective are common in QualAdj languages, AdjQual languages exhibit a clear tendency to be NAdj (table 8).

Table 8

	# Lgs	# Families	# Families (where dominant)
QualAdj&AdjN	41	14	9
QualÁdj&NAdj	22	22	7
			(Neither 3)
AdjQual&AdjN	4	3	0
AdjQual&NAdj	45	14	12
			(Neither 2)

As with the evidence from the order of manner adverb and verb, this data supports the idea of a cross-categorial tendency for consistent ordering of modifier and head (but not necessarily dependent and head).

Note. What remains unexplained on the present account is why a pattern like that in table 7 is not found with OV languages. The data we find is given in table 9.

Table 9

	# Lgs	# Families	# Families (where dominant)
OV&AdvV&AdjN	27	11	6
OV&AdvV&NAdj	32	12	8
			(Neither 4)
OV&VAdv&AdjN	0	0	0
OV&VAdv&NAdj	7	3	3

The pattern we found with OV languages might lead us to expect OV&-AdvV&AdjN to be by far the most common. But it is not; in fact OV&-AdvV&NAdj is slightly more common. On the other hand, the absence of

OV&VAdv&AdjN languages is parallel to the absence of VO&AdvV&NAdj languages, and we can generalize the implicational universal to say

If the adverb occurs on the opposite side of the verb from the object, then the order of adjective and noun will mirror the order of adverb and verb.

A possible explanation for the contrast with VO languages is that the dominance principle preferring NAdj order is augmenting the effects of the harmony principles in the case of VO languages, contributing to the preference for VO&VAdv&NAdj over VO&VAdv&AdjN but cancelling, in the case of OV languages, the effect of the harmony principle whereby the order of verb and adverb mirrors that of adjective and noun (see figure 4).
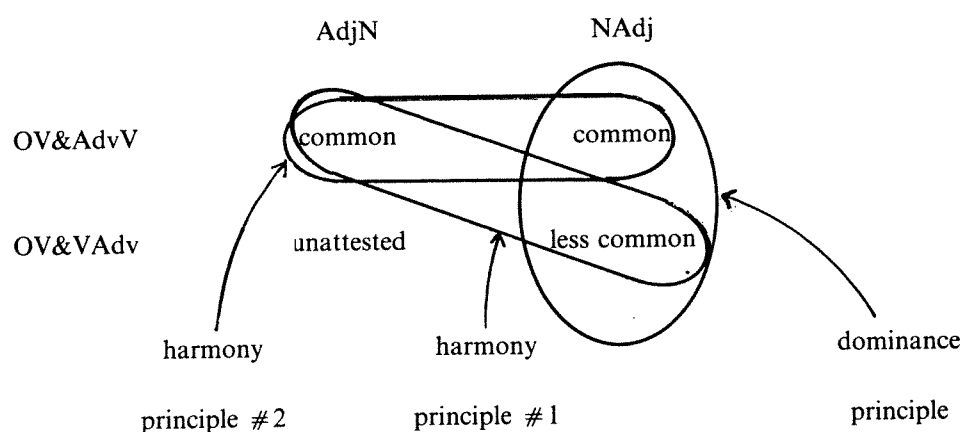

Fig. 4.

Each of the attested possibilities satisfies two of the three principles, whereas the unattested type satisfies none. The two most common types would seem to be the most common because harmony principle #2 is a more powerful factor.

6. Hawkins' principle of cross-category harmony

The claims of this paper are at apparent odds with the claims and evidence of Hawkins (1983). His principle of Cross-Category Harmony claims, loosely speaking, that the most common language types should be those in which the relative number of dependents preceding and following one head should be the same as the relative number of dependents preceding and following other

heads, and that languages should become increasingly uncommon as we deviate from these ideals. This predicts, for example, that SOV languages should tend to be AdjN, and V-initial languages NAdj. But, as shown in section 1, this prediction is not borne out. The question thus arises: Why does Hawkins' data apparently support his claims while mine seems not to?

Hawkins bases his cross-category harmony calculations on data from the appendix in Greenberg (1963). Greenberg's appendix contains 28 Type 23 languages (SOV&Po&GenN&AdjN), the 'ideal' for SOV languages, according to Hawkins' Cross-Category Harmony Principle. The next most common type in Greenberg's appendix is Type 24 (SOV&Po&GenN&NAdj), of which there are 24 languages.[18] Although this difference is small, Hawkins' principle does correctly predict that Type 23 languages are more common.

Nevertheless, I have already mentioned the dangers in counting languages in a sample containing more than one language per family. In fact, it strikes me as conceptually strange that Hawkins' Cross-Category Principle is formulated in terms of numbers of *languages*, since actual language numbers are also the effect of various non-linguistic factors. Clearly, there is a *ceteris paribus* clause implicit in his principle, but one way to control for certain non-linguistic factors is to count families (or use a sample containing only one language per family).

It is therefore useful to test Hawkins' principle against my sample, but to do so counting *families* rather than languages. Consider the case of the order of genitive and adjective with respect to noun in SOV languages. Hawkins' principle (revised to refer to families rather than languages) predicts that the number of families should decrease as the number of cross-category harmony deviations increases:[19]
The relative size of these figures is exactly as Hawkins' principle predicts. But the lower frequency of the third type also follows from the Branching

[18]   Not surprisingly, when we examine the areal distribution of these languages, we find that they exhibit roughly the same pattern described in section 1. Namely, of the 28 Type 23 languages, 18 are spoken in Eurasia, and of the 10 languages spoken outside Eurasia, 5 are Afro-Asiatic. Conversely, of the 24 Type 24 languages, only 6 are spoken in Eurasia. Another interesting property of the languages in Greenberg's appendix is that when one ignores adposition type and order of noun and genitive, it turns out that Greenberg's appendix actually contains more SOV&NAdj than SOV&AdjN languages. Namely, of the 64 SOV languages, 35 are NAdj while only 29 are AdjN.

[19]   The type SOV&NGen&AdjN is excluded by Hawkins by an implicational universal equivalent to
      If SOV and NGen then NAdj.
In fact, my sample contains one apparent instance of such a language, Djapu.

	# Families
SOV&GenN&AdjN	20
SOV&GenN&NAdj	17
SOV&NGen&NAdj	7

Direction Theory discussed in section 2, since that theory predicts that SOV languages should be GenN. The Branching Direction Theory does not predict any difference between the first two types, and this is essentially true: although 20 is larger than 17, the difference is well within the realm of chance. Furthermore, 10 of the 20 SOV&GenN&AdjN families are spoken in Eurasia, whereas only 2 of the 17 SOV&GenN&NAdj families are. Hence the greater number of families of the former sort is apparently attributable to Eurasia.

I will not examine the analogous cases of SVO and V-initial languages here, but similar remarks apply. The only instances in which differences predicted by the Cross-Category Harmony Principle are significant are precisely those predicted by the Branching Direction Theory.

It is useful to consider the predictions of the Cross-Category Harmony Principle as applied to the order of adjective and demonstrative with respect to the noun, since neither the adjective nor the demonstrative are branching categories. In this instance, the Branching Direction Theory does not predict the differences in numbers of languages that the Cross-Category Principle predicts, and we would expect, modulo possible other principles and random variation, roughly equal numbers of the different language types. And this is essentially what we find in table 10.

Table 10

	SOV	SVO	V-initial
DemN&AdjN	*17*	7	7
DemN&NAdj	15	7	6
NDem&AdjN	1	1	2
NDem&NAdj	13	8	7

The numbers of families cited in this section is simply the number of families containing at least one language of the type in question, the statistic given in the middle column of various other tables in this paper, rather than the number of families in which that order is dominant. The latter statistic is only well-defined for comparison of two language types, rather than three (as in the present instance) and it is not clear how best to extend the notion to such cases.

The three types whose figures are underlined in the table are those that Hawkins' Cross-Category Harmony Principle predicts should be most frequent. But the only significant pattern is the consistent rarity of NDem&AdjN languages, something that Hawkins (1983) observes by his universal

If NDem then NAdj.

Although my sample contains exceptions, this universal is nevertheless a significant statistical implicational universal. But within each column, the other three figures do not differ significantly from each other. In short, when we restrict attention to nonbranching modifiers, like the adjective and demonstrative, we do not find any significant decrease in language frequency as we decrease in cross-category harmony. In short, differences in language frequency predicted by Hawkins' principle seem only to be appreciable when branching categories are involved.

7. Conclusion

The important results of this paper are empirical. I have shown that there is no evidence for any universal relationship between the order of adjective and noun and the order of object and verb. I have also shown that the order of adjective and noun among OV languages exhibits a strong areal pattern, the OV languages of Eurasia tending towards AdjN order, those outside Eurasia tending toward NAdj order. Both results may bolster the scepticism of those who question the possibility of making inferences from numbers of languages to properties of language. But although the alleged relationship between the order of adjective and noun and that of object and verb has been shown to be a myth, the same cannot be said for other word order characteristics; the figures cited for the order of genitive and noun and those for the order of relative clause and noun show strong tendencies that must be attributed to general properties of language. I have argued that these facts support a principle of consistent direction of branching rather than one of consistent ordering of head and dependent.

Appendix A: Listing of languages

The following is a listing of the languages in my sample for each of the two orders of noun and adjective and the three clause order types, OV, SVO, and V-initial:[20]

OV&AdjN
  AFRO-ASIATIC: Kemant, Ometo, Amharic, Gourague.
  ANDEAN-EQUATORIAL: Quechua, Jibaro.
  AUSTRALIAN: Djapu.
  AUSTROASIATIC: Mundari, Kurku.
  AZTEC-TANOAN: Shoshoni (Big Smokey Valley), Comanche, Yaqui, Cahuilla, Hopi.
  BURUSHASKI: Burushaski.
  CAUCASIAN: Georgian.
  CHUKCHI-KAMCHATKAN: Koryak.
  DRAVIDIAN: Gondi, Kolami, Koya, Kuvi, Telugu, Tulu, Kannada.
  GE-PANO-CARIB: Carib.
  HOKAN: Chimariko, Washo.
  INDO-EUROPEAN: Dumaki, Shina, Hindi, Punjabi, Bishnupriya Manipuri, Maithili, Nepali, Pali, Wakhi, Pashto.
  INDO-PACIFIC: Kewa, Agarabi, Awa, Gadsup, Usarufa, Hua, Yagaria, Yareba, Kiwai.
  MACRO-CHIBCHAN: Cayapa.
  NA-DENE: Tlingit.
  NILO-SAHARAN: Lendu (North).
  NIVKH: Nivkh (Gilyak).
  PENUTIAN: Klamath.
  SINO-TIBETAN: Lahu, Dafla, Gurung.
  URAL-ALTAIC: Buriat, Dagur, Khalkha, Karakalpak, Azerbaijani, Turkish, Chuvash, Korean.
  YENISEIAN: Ket.
  YUKAGHIR: Yukaghir.

OV&NAdj
  AFRO-ASIATIC: Somali, Geleba, Iraqw.

---

[20] For reasons of space I am unable to list my sources in the bibliography. such information can be obtained from the author.

ANDEAN-EQUATORIAL: Piro-Arawakan, Iranxe.

AUSTRALIAN: Maranungku, Malakmalak, Dalabon, Wageman, Ngandi, Mangarayi, Alyawarra, Diyari, Anguthimri (Mpakwithi), Uradhi, Guugu Yimidhirr, Gugada, Western Desert, Djingili.

AUSTRONESIAN: Manam, Wedau, Iduna, Balawaia.

AZTEC-TANOAN: Kiowa, Ute.

BASQUE: Basque.

ESKIMO-ALEUT: Eskimo.

GE-PANO-CARIB: Apinaye, Kraho, Cashibo (Ra-Txa Hu-Ni-Ku-I), Chacobo, Eseejja.

HOKAN: Diegueno, Walapai.

INDO-EUROPEAN: Persian.

INDO-PACIFIC: Waskia, Golin, Salt-Yui, Kobon, Sentani, Guhu-Samane, Daga, Kunimaipa, Barai, Omie, Koita, Koiali (Mountain), Kapau, Siroi, Yeletnye.

MACRO-ALGONKIAN: Chitimacha, Tonkawa, Tunica.

MACRO-CHIBCHAN: Miskito, Guaymi.

MACRO-SIOUAN: Wichita, Biloxi, Hidatsa, Dhegiha, Lakota, Yuchi.

NA-DENE: Navajo, Slavey, Haida.

NIGER-KORDOFANIAN: Rashad, Tenyer, Seme, Bambara, Mandinka, Mandinka (Gambian), Vai, Mende.

NILO-SAHARAN: Berta, Mamvu, Dongolese Nubian, Nyimang, Tama, Barya, Kunama, Maba, Tubu, Fur, Songhai.

PENUTIAN: Takelma.

SINO-TIBETAN: Garo, Kokborok, Rong, Chingpaw, Burmese, Ao, Ladakhi, Sherpa.

SUMERIAN: Sumerian.

YUKIAN: Wappo.

SVO&AdjN

AUSTRALIAN: Gunbalang, Jiwadja, Maung, Yukulta, Tiwi.

AUSTRONESIAN: Palauan.

INDO-EUROPEAN: Lithuanian, Danish, Swedish, English, Russian, Polish.

MACRO-ALGONKIAN: Blackfoot, Yurok.

NIGER-KORDOFANIAN: Gbaya Kaka, Gbeya Bossangoa, Sango, Nzakara.

PENUTIAN: Totonac.

SINO-TIBETAN: Cantonese, Mandarin.

URAL-ALTAIC: Hungarian, Finnish.

SVO&NAdj

    AFRO-ASIATIC: Hausa, Margi, Tera, Arabic (Colloquial Egyptian).

    ANDEAN-EQUATORIAL: Cocama, Guarani, Siriono.

    AUSTROASIATIC: Temiar, Chrau, Sre, Stieng, Brao, Palaung, Vietnamese, Cambodian, Khasi.

    AUSTRONESIAN: Indonesian, Mor, Nissan, Tigak, Nguna, Futuna-Aniwa, Rotuman, Dehu, Mokilese, Ponapean, Kusaiean, Kaliai-Kove, Patep, Halia, Sakao, Lenakel.

    AZTEC-TANOAN: Nahuatl (Michoacan), Nahuatl (Tetelcingo).

    GE-PANO-CARIB: Mataco.

    INDO-EUROPEAN: Albanian.

    INDO-PACIFIC: Arapesh.

    KHOISAN: Xu.

    NIGER-KORDOFANIAN: Katla, Katcha, Mba, Engenni, Ewe, Grebo, Fanti, Twi, Nkonya, Igbo, Lelemi, Yoruba, Swahili, Ewondo, Mwera, Luvale, Noni, Duka, Fulani.

    NILO-SAHARAN: Bari, Sara-Ngambay, Shatt, Temein, Bor, Acooli, Koma.

    OTO-MANGUEAN: Pame.

    SINO-TIBETAN: Nung, Thai.

V-initial&AdjN

    AUSTRALIAN: Wembawemba.

    AUSTRONESIAN: Chamorro, Agta (Central), Manobo (Western Bukidnon).

    CHIMAKUAN: Quileute.

    HOKAN: Chontal (Huamelultec Oaxaca).

    INDO-EUROPEAN: Nuri.

    OTO-MANGUEAN: Otomi.

    PENUTIAN: Chinook (Lower), Nez Perce, Northern Sahaptin, Tsimshian (Coast).

    SALISH: Squamish.

    WAKASHAN: Kwakiutl.

V-initial&NAdj

    AFRO-ASIATIC: Hebrew (Biblical), Arabic (Modern Literary).

    ANDEAN-EQUATORIAL: Goajiro, Island Carib, Baure, Guajajara.

    AUSTRALIAN: Garawa.

    AUSTRONESIAN: Tboli, Tahitian, Hawaiian, Samoan, Niuean, Tongan, Gilbertese, Yapese.

AZTEC-TANOAN: Nahuatl (Huasteca).
INDO-EUROPEAN: Welsh.
NILO-SANHAEAN: Didinga, Tepeth, Pokot, Maasai.
OTO-MANGUEAN: Chatino, Zapotec (Mitla).
PENUTIAN: Jacaltec.

Appendix B: The sampling procedure

The data cited in this paper is based on a sample of approximately 425 languages for which various word order characteristics have been collected from primary sources. The 316 languages that form the basis of this study are those for which I have been able to determine a basic order of noun and adjective. For some languages, this was not possible, either because there was insufficient information in my sources, or because both orders of noun and adjective are found and there is no apparent basis for choosing one order as basic. The larger sample of 425 languages is a convenience sample: I have attempted to include as wide a variety of languages as possible, and there are only two reasons that I have employed for excluding languages:

(1) I could not find any source for the language;
(2) within certain well-documented groups, I selected a subset of languages on the basis of personal convenience.

As discussed by Perkins (1980), there are severe methodological problems with the use of samples containing related languages. Namely, it becomes impossible to determine, when one language type occurs more frequently in one's sample, the likelihood that this is due to linguistic factors. The methodology of Perkins and of Bybee (1985) is to construct a sample in which each language is independent. The methodology employed here is similar in spirit: the sample contains languages that are not independent but controls for this by counting families rather than languages, more specifically by counting families in which the given order is dominant.

In calculating which order is *dominant* (if any) among the languages in a family, I calculate a *weighted proportion* of the languages in each family so that at each level in the genetic classification, parallel genetic groupings are counted equally. I illustrate this procedure here by discussing in detail my claim that in the Afro-Asiatic family, the OV languages are predominantly AdjN.

My sample contains 13 Afro-Asiatic languages. Of these, 6 are VO and all 6 are NADj. Of the 7 OV languages, 4 are AdjN while 3 are NAdj. Since only 4 of the 7 OV languages are AdjN, my claim that AdjN order is dominant among the OV languages might seem surprising. I classify an order as dominant if the weighted proportion of languages exhibiting that order is greater than $\frac{2}{3}$. It turns out that although the 4 OV&AdjN languages in Afro-Asiatic constitute an *absolute* proportion of only 0.57 (or $\frac{4}{7}$) of the family, their weighted proportion is 0.78, which is greater than $\frac{2}{3}$.

This figure of 0.78 is calculated as follows. The OV languages in Afro-Asiatic fall into 3 subfamilies, Cushitic, Omotic, and Semitic. (My sample also contains 3 Chadic languages, but these are all VO.) Each of these three subfamilies is assigned a weight of 0.33 (or $\frac{1}{3}$), the total weight for each family being 1. My sample contains one Omotic language. Ometo, which is OV. Since this is the only language from my sample from this subfamily, Ometo is assigned the same weight as the subfamily, namely 0.33. My sample contains 5 semitic languages, but only 2 of these, Amharic and Gourague, are OV; they are each assigned a weight of 0.17, one half of the weight of their subfamily Semitic, which is 0.33. The remaining four OV languages are Cushitic and they fall into three further subgroups, each of which is assigned a weight of 0.11, one third of the weight assigned to Cushitic. Kemant, the sole language in my sample from Central Cushitic, is thus assigned this weight of 0.11, as is Iraqw, the sole language from Southern Cushitic. The remaining two languages, Somali and Geleba, are each assigned a weight of 0.055, one half of 0.11, the weight assigned to the subgroup they belong to, Lowland East Cushitic. The following summarizes this, listing as well the order of noun and adjective.

AFRO-ASIATIC (weight = 1) (OV languages only)
  OMOTIC (0.33)
    Ometo (0.33) AdjN
  SEMITIC (0.33)
    Amharic (0.17) AdjN
    Gourague (0.17) AdjN
  CUSHITIC (0.33)
    CENTRAL (0.11)
      Kemant (0.11) AdjN
    SOUTHERN (0.11)
      Iraqw (0.11) NAdj
    LOWLAND EAST (0.11)
      Somali (0.055) NAdj
      Geleba (0.055) NAdj

We can now calculate the weighted proportions for AdjN and NAdj languages. The sum of the weighted proportions for the four OV&AdjN languages is 0.78. Although 3 of the 7 languages are NAdj, their weighted proportion is only 0.22. This is because they occur in two subgroups of one subfamily, while the rest of that subfamily, plus the other two subfamilies, are AdjN. There thus seems to be a clear sense in which the classification of the Afro-Asiatic family as one in which the OV languages are predominantly AdjN is intuitively justified.

Three shortcomings of the above weighting system should be noted. The first of these is that the calculation depends on a particular subclassification of the languages within a family. I have depended on the subclassifications of Voegelin and Voegelin (1977), but clearly many of their classifications are controversial, many more so than the above subclassification of Afro-Asiatic. But note that even if Omotic were classified within Cushitic, as has sometimes been done, the weighted proportions would change relatively little: even if it were treated as a fourth branch of Cushitic, the weighted proportion of AdjN languages would still be 0.75, which differs little from my figure of 0.78. It seems likely that alternative subclassifications would not in general affect the calculated proportions very much. At most, certain families that are classified as predominantly some order might shift into the 'Neither order dominant' category or vice versa, and thus the figures for 'Number of families where dominant' might shift a little accordingly.

A closely related and perhaps more severe shortcoming is that the system depends on assumptions as to which groupings are families. In some cases, this makes little difference. I treat Ural-Altaic as a group, but this matters little since both Uralic and Altaic are entirely AdjN, so distinguishing them would simply add one more family to the list of OV&AdjN families in Eurasia. I treat Caucasian as a group, but since my sample currently contains only one Caucasian language (Georgian), this doesn't matter. This failure to include more Caucasian languages is a problem; inspection of data listed in Hawkins (1983) reveals that the Kartvellian (Southern) and Northeast Caucasian languages are AdjN while the Northwest languages are NAdj. If we were to split up the Caucasian languages into the three more widely accepted families, then we would have one more OV&AdjN family in Eurasia, but also an additional OV&NAdj family in Eurasia, a more significant effect, given the relative rarity of this order in Eurasia.

A more interesting example of a questionable grouping is Penutian, a grouping that is not only highly controversial, but one that contains a large number of subgroups, seven of which are represented in my sample. Three of

these seven groups (Chinookan, Sahaptin-Nez Perce, and Tsimshian) are
V-initial and AdjN, but all three are spoken in northwest North America.
The remaining four groups each fall into different word order categories:
Totonac is SVO&AdjN, Mayan is V-initial&NAdj, Klamath is OV&AdjN
and Takelma is OV&NAdj. Suppose then that we were to treat these seven
groups as seven separate families. What effect would this have on the figures
given? Treating Klamath as an isolate would add a second family outside
Eurasia in which the OV languages were predominantly AdjN. But the areal
pattern is so strong that it would be little affected by this. And treating
Chinookan, Sahaptin-Nez Perce and Tsimshian as separate families would
only increase the number of V-initial&AdjN families, so the apparent evi-
dence against Greenberg's Universal 17 would if anything be further under-
mined. In short, the conclusions drawn are not particularly affected by the
treatment of Penutian as a family. It must also be emphasized that it is not
clear whether there is any more reason to reject the grouping Penutian than a
number of various other groupings that are also controversial. By assuming
the classification of Voegelin and Voegelin, I at least achieve a certain amount
of consistency, since they have applied the same criteria across the world to
the best of their ability. Furthermore, by using their classification, I at least
have an objective standard, something I would not have if I allowed myself to
be swayed by the biases of particular linguists who question Voegelin and
Voegelin's classification.

The third and to my mind most severe shortcoming of the system of
weighted proportions and counting families employed in this paper is that it
completely ignores the effects of contact or areal phenomena. The assumption
underlying the methodology of counting families is that one thereby obtains a
count of *independent* phenomena. But given the extent to which word order
phenomena are areal, such an assumption is simply false. The evidence for a
linguistic area stretching from Basque in Europe to Chukchee in eastern
Siberia means that no two languages in this area can be assumed to be
completely independent. One implication of this is that concerns about
genetic classification are rendered somewhat moot. Whether Chinookan,
Sahaptin-Nez Perce and Tsimshian should be grouped together into a single
genetic group is largely irrelevant in light of strong evidence that they do
belong to a single linguistic area (Northwest North America).[21] Similar
comments apply to Ural-Altaic and to Caucasian. If anything, for sampling

---

[21]   There is considerable evidence beyond that cited in this paper for a linguistic area in
Northwest North America. Cf. Thompson and Kinkade (to appear).

purposes, the conservative assumption is always to assume that any two languages are related, unless there is good reason to believe otherwise. Hence the inclusion of groups like Ural-Altaic and Penutian is a less severe problem than the failure to include broader areal groups like 'Northwest North America' and 'Central Asia'. Unfortunately, it is not clear at this time what the appropriate areal/genetic groups should be, especially given the continuum nature of areal phenomena. Hence the actual number of families cited in this paper should not be taken too seriously. Nevertheless, the basic empirical conclusions of this paper, that there is no relationship between the order of object and verb and the order of adjective and noun, and that the OV languages exhibit a large-scale areal split between Eurasia and the rest of the world, are sufficiently robust that they survive any reasonable objections to the particular classification assumed.

The problem of areal and contact phenomena has even more severe implications for the system of weighting languages according to the classification within families, because this system ignores the effects of contact between languages *within* the same family. When we consider the example of Afro-Asiatic discussed above, the question arises whether the fact that AdjN order is found in various branches of the family might be partly explained in terms of contact. In fact, the OV&AdjN order in the Semitic languages Amharic and Gourague is likely due to contact with Cushitic languages with the same order. But the system of weighting languages employed here actually assigns a higher weight to the OV&AdjN Semitic languages than to the OV&AdjN Cushitic languages (since all the OV Semitic languages in my sample are AdjN, while only some of the OV Cushitic languages are).

I have openly discussed the shortcomings of the sampling methodology employed in this paper. I would maintain that the methodology is nevertheless clearly superior to those generally employed in typological studies, with a few notable exceptions, like the methodology of Perkins (1980) and Bybee (1985). Although my attempts to control for the relatedness of languages (be it genetic or areal) are crude, I have a least attempted to address those problems in a serious fashion.

References

Bybee, Joan, 1985. Morphology: A study of the relation between meaning and form. Typological studies in language, Vol. 9. Amsterdam: John Benjamins.

Dixon, Robert M.W., 1977. Where have all the adjectives gone? Studies in Language 1, 19–80. Reprinted in: Where have all the adjectives gone? And other essays in semantics and syntax. New York: Mouton, 1982.

Dryer, Matthew S., in press. Universals of negative position. In: Edith Moravcsik, Jessica Wirth and Michael Hammond (eds.), Studies in syntactic typology. Amsterdam: John Benjamins.

Greenberg, Joseph H., 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: Joseph Greenberg (ed.), Universals of language. Cambridge, MA: MIT Press.

Hawkins, John A., 1983. Word order universals. New York: Academic Press.

Hawkins, John A., 1984. Modifier-head or function-argument relations in phrase structure? Lingua 63, 107–138.

Keenan, Edward, 1978. The syntax of subject-final languages. In: Winfred Lehmann (ed.), Syntactic typology. Austin, TX: University of Texas Press.

Keenan, Edward L., 1979. On surface form and logical form. In: Braj B. Kachru (ed.), Linguistics in the seventies: Directions and prospects. Studies in the Linguistic Sciences 8:2, Urbana, IL: Dept. of Linguistics, University of Illinois.

Lehmann, Winfred P., 1973. A structural principle of language and its implications. Language 49, 42–66.

Masica, Colin P., 1976. On defining a linguistic area. Chicago, IL: University of Chicago Press.

Munro, Pamela, 1985. A new harmony within Type 24. Paper presented at the 1985 Annual Meeting of the Linguistic Society of America, Seattle.

Perkins, Revere, 1980. The evolution of culture and grammar. SUNY Buffalo. (Ph.D. dissertation.)

Thompson, Laurence C. and M. Dale Kinkade. To appear. Linguistic relations and distributions. In: Handbook of American indians, Vol. VII, The Northwest coast.

Tomlin, Russell, 1986. Basic constituent orders: Functional principles. London: Croom Helm.

Van Valin, Robert D., Jr. and Cheryl R. Garcia, 1985. Adjective–noun order and relative clause–noun order in OV languages. Paper presented at the 1985 Annual Meeting of the Linguistic Society of America, Seattle.

Vennemann, Theo, 1974. Analogy in generative grammar: The origin of word order. In: Proceedings of the eleventh international congress of linguists (1972). Bologna.

Voegelin, C.F. and F.M. Voegelin. 1977. Classification and index of the world's languages. New York: Elsevier.