

## ON THE SIX-WAY WORD ORDER TYPOLOGY

MATTHEW S. DRYER

*State University of New York at Buffalo*

### ABSTRACT

A number of arguments are given against the traditional word order typology based on the six types SOV, SVO, VSO, VOS, OVS, and OSV, and in favour of an alternative typology based on two binary parameters OV vs. VO and SV vs. VS. The arguments given include ones based on various advantages of collapsing VSO and VOS into a single type, the infrequency of clauses containing a noun subject and noun object, the value of isolating the more predictive parameter of OV vs. VO, and the fact that the traditional typology ignores the position of intransitive subjects.

### 1. Introduction

There is a long tradition in linguistics that treats the basic word order of the subject, object, and verb as a fundamental typological parameter, that treats the question of whether a language is SOV, SVO, VSO, VOS, OVS, or OSV as one of the most important things to know about a language. The purpose of this paper is to present arguments in favour of an alternative to this six-way typology, one based on two separate 2-way typological parameters: OV vs. VO, and SV vs. VS. Together these two parameters define four types: VS&VO, SV&VO, SV&OV, and VS&OV. The first of these, VS&VO, which I will refer to as *verb-initial*, corresponds roughly to the two traditional types VSO and VOS; SV&VO corresponds roughly to SVO; SV&OV, which I will refer to as *verb-final*, corresponds to the two types SOV and OSV; and VS&OV corresponds to the rare type OVS.<sup>1</sup>

I will present eight arguments for the proposed typology. First, I will argue that it allows easy classification of languages which are indeterminate-

ly VSO/VOS. Second, I will argue that there is no evidence that the difference between VSO and VOS languages is predictive of anything: the properties that are typical of VSO languages are apparently also typical of VOS languages, and hence VSO and VOS are best treated as belonging to the same type. Third, I will argue that the difference between VSO and VOS order is a relatively unstable one, both orders being commonly found as basic orders within the same language family and within the same linguistic area. Fourth, I will argue that the traditional six-way typology is based on a clause type that occurs relatively infrequently, while the proposed typology is based on clause types that occur much more frequently. Fifth, I will argue that there are many languages which have word order sufficiently flexible that they are impossible to classify by the traditional typology but which are still classifiable by the proposed typology. Sixth, I will argue that there are other languages with word order even more flexible but which are still classifiable either for the order of subject and verb or for the order of object and verb. Seventh, I will argue that the proposed typology is superior because it isolates the order of the object and verb, the more fundamental typological parameter in terms of word order correlations. And eighth, I will argue that the traditional typology suffers because it overlooks the order of subject and verb in intransitive clauses, even though the order in such clauses is occasionally different from the order of subject and verb in transitive clauses, and even though intransitive clauses containing a noun subject are much more common than transitive clauses containing a noun subject.

## 2. The notion of *basic word order*

The term *basic word order* is used in various ways by different linguists, often without an apparent awareness that it is being applied to different notions.<sup>2</sup> The characterization of basic word order by Hawkins (1983: 13) is representative of criteria assumed by many linguists: he uses a set of different criteria, all of which tend to correlate with each other, though none of them are necessary properties. These include the most frequent order, the order that occurs in the broadest set of syntactic environments, and the order that is unmarked by a variety of other markedness criteria. But

apart from the description of particular languages, the notion of basic word order has played its most significant role in identifying the orders of various pairs or sets of elements that provide the empirical basis for the crosslinguistic generalizations originally discussed by Greenberg (1963) and pursued in works by various people (e.g. Lehmann 1973, Vennemann 1976, Hawkins 1983, Dryer 1992). Whatever the merits of the various criteria discussed by Hawkins, the fact remains that the empirical basis for these generalizations has largely been statements in grammatical descriptions which describe one order of some pair or set of elements as the normal order or the most common order. In other words, while criteria other than frequency may typically correlate with the most frequent word order, a large body of descriptive literature provides statements regarding the apparent relative frequency of different orders, but very little evidence regarding other criteria, excepting that for many languages, like English, one order is so clearly basic by all criteria that no questions arise. In short, we know that there are significant crosslinguistic generalizations based on a notion of basic order associated with most frequent word order, and the criteria for identifying an order as most frequent are relatively clearcut (though see below). For the purposes of this paper, therefore, I will simply define the *basic word order* of two or more elements as the most frequent order of those elements in the language. In defining the term *basic word order* in this way, I am not intending to deny that there may be other useful notions that one might apply the expression to, but the different notions should not be confused.<sup>3</sup> While some linguists assume a notion of basic word order that is not always the most frequent word order, this difference is simply a terminological one, with the term *basic word order* being used by different linguists as a label for related but distinct notions. Given the fact that most notions of basic word order at least correlate very strongly with the most frequent order, it seems likely that the conclusions of this paper are equally applicable to other notions of basic word order. But whether or not this is the case is difficult to determine, given the difficulties applying other criteria for identifying basic word order.

Despite the fact that I appeal to evidence based on frequency, I must concede from the start two fundamental problems with the notion of *most frequent order in a language*, one methodological, the other substantive. The methodological problem is that, despite the fact that it is often easy to identify a most frequent order in a given body of texts, questions can arise whether the differences in frequency between alternative orders are necessari-

ly general properties of texts in the language rather than accidental properties of a particular text or set of texts. What this means is that if we really want to determine whether a particular order is really most frequent in a language, we need to examine as wide a variety of texts representing different genres as possible. If a particular order is more common in most or all texts, then we can justifiably describe that order as most frequent. If no order is most frequent over most texts, however, or if the order varies from genre to genre or text to text, we should probably not describe any particular order as the basic order (in the sense of most frequent order) and we should say that the language is one that lacks a basic word order, as Mithun (1987) argues for a number of languages. In short, while it may be relatively easy to identify a most frequent order in a single text or in a small body of texts, it is necessary to examine a wide variety of texts before one can decide with confidence that a particular order is most frequent in the language *as a whole*. Much of the frequency data cited below falls short of this ideal, and thus the text counts cited should probably be considered in many cases as no more than pilot studies that at best suggest that a particular order is most frequent.

The substantive problem with frequency is that frequency is epiphenomenal relative to grammars of languages; typologies of languages are often assumed to be typologies of grammars of languages. Where one order is more frequent than another in a language, I assume it to be the case that the higher frequency ultimately reflects the discourse conditions under which the different orders are used and the order that is more frequent is more frequent only because the discourse conditions in which it is used tend to occur more frequently in normal discourse. Hence frequency is due to two factors, the linguistic factor of the rules or principles governing when particular orders are used, and the nonlinguistic factor of how often particular discourse conditions arise. The following example of word order in Papago should make this point clear. Payne (1987: 793–794) cites the text count given in Table 1.

*Table 1: Word Order in Papago*

SV	48 (23%)	OV	44 (29%)
VS	158 (77%)	VO	108 (71%)

Table 1 shows that both subjects and objects more commonly follow the verb in Papago. Payne argues, however, that word order in Papago is largely determined by the definiteness of noun phrases, that indefinite noun phrases generally precede the verb, while definite noun phrases generally follow. The greater frequency of VS and VO order is thus apparently largely a reflection of the fact that definite noun phrases occur more often than indefinite ones in typical discourse. The greater frequency of VS and VO orders is thus not a fact about the grammar of Papago, but reflects the interaction of the discourse principles governing word order and the frequency with which the conditions defined by these principles arise.

These considerations would seem to provide a reason to question the usefulness of a notion of basic word order based on frequency. If we accept Payne's account, the grammar of Papago apparently contains no statement that VS and VO are basic: rather, the sentence-level syntax of the language defines both orders as possible, and the discourse grammar of the language specifies the discourse conditions in which VS and VO are used and the discourse conditions in which SV and OV are used. If VS and VO are more common, then that is not a fact about the grammar of Papago, but a fact about Papago texts.

I argue at length in Dryer (1989a), however, that while the greater frequency of VS and VO order in Papago may not be facts about the grammar of Papago, it is necessary to appeal to these frequency facts to explain certain changes that have been occurring in the grammar of the language. Namely, Papago has apparently undergone a recent change in the discourse principles governing word order whose effect is that while the discourse principles now governing word order in the language result in more frequent VS and VO order, the discourse principles used to be ones that resulted in more frequent OV order (and perhaps SV order). An apparent side effect of this change in the discourse grammar of the language is that the language has been undergoing other changes in word order, with an increase in the use of prepositions and noun-genitive order. These changes

are apparently due to the increase in frequency of VS and VO order, and the factors underlying the crosslinguistic correlations between VO order and other word order characteristics (cf. Dryer 1992) have led to changes in the grammar of the language, changes that can only be explained in terms of the higher frequency of VS and VO, and not in terms of the discourse principles governing the choice of VS and VO vs. SV and OV. What this means is that even though the greater frequency of VS and VO order in Papago is epiphenomenal relative to the grammar of the language, they are features of language use that are essential to understanding the grammar of the language. In describing languages as, say, VS on the basis of frequency, I assume this to be a circumlocution for saying that the language is one whose discourse (or semantic) principles governing word order are such that VS generally ends up being more common in texts. I thus assume that these labels, while assigned on the basis of relative frequency, are ultimately a reflection of a deeper typology of underlying discourse principles governing word order.

In assigning languages to a particular category on the basis of frequency, I will treat a particular order as basic in a language if it is at least twice as frequent in texts as the order or orders it contrasts with. I will also assume that where two orders differ in frequency by a ratio of much less than two-to-one, neither order should be considered the basic order. Although it is obviously somewhat arbitrary to divide the line at a ratio of two-to-one, some empirical evidence is presented in Dryer (1989a) that it is meaningful to treat an order as basic if it is more common by a ratio of only two-to-one: in languages in which one order of object and verb outnumbers the other by only two-to-one, one tends to find other word order characteristics that are associated with the more common order. There is less evidence of this for languages in which one order is more frequent by a ratio of less than two-to-one.

### 3. Collapsing VSO and VOS

#### 3.1 *Argument 1: Languages which are indeterminately VSO/VOS*

The first three arguments that I will present for the typology being defended here are based on the fact that this typology differs from the traditional six-way typology in collapsing the two types VSO and VOS, by treating them both as instances of the type VS&VO. These arguments (and

others discussed below) argue that the difference between the two types VSO and VOS is a minor one, particularly when compared with the difference between these types and the SVO or SOV types.<sup>4</sup>

While the type VS&VO corresponds roughly to the two types VSO and VOS, this correspondence is not exact. For one thing, there are a number of languages which are clearly VS&VO but which are difficult to classify by the traditional six-way typology because, while noun arguments of the verb typically follow the verb, it is not clear that the language is specifically VSO or VOS. For example, in Fijian (Dixon 1988: 242), the two normal orders are VSO and VOS. But both are sufficiently common that there appears to be little basis for treating one of them as the basic order. What this means is that Fijian cannot be assigned to one of the six traditional word order types; either it remains unclassified, or it must be described using the hybrid type VSO/VOS. On the typology being proposed here, however, Fijian is clearly an instance of a VS&VO language. Hence the typology being proposed here allows for the straightforward classification of languages which are not easy to classify on the traditional typology. The situation in Tagalog is somewhat similar to that in Fijian, but worse because of the difficulties deciding how (if at all) one is to apply the notion of subject to Philippine languages (cf. Schachter 1976, 1977; Mulder & Schwartz 1981). However, since all of the relevant candidates for subject follow the verb, Tagalog can be classified as VS&VO in the sense that both arguments of a transitive verb typically follow the verb. Other languages that are like Fijian and Tagalog in being indeterminately VSO/VOS but clearly VS&VO include Rukai (Li 1973), an Austronesian language of Taiwan, and Cariri (Adam 1897), a language of eastern Brazil (sometimes classified as Equatorial).

### 3.2 *Argument 2: The similarity of VSO and VOS languages*

A typological distinction is more significant if it provides the basis for predicting other typological characteristics. Thus the significance that has been associated with the contrast between OV and VO languages derives from the extent to which there are other characteristics that can be predicted, given the order of object and verb. But there are no known differences between VSO and VOS languages, apart from the order of subject and object. Keenan (1978) documented a set of properties of what he called "subject-final" languages, but since the existence of OVS languages had not been documented at that time, the properties he listed were actually proper-

ties of VOS languages. Furthermore, as noted by Hawkins (1983: 31), there is no evidence that any of the properties he listed are any more typical of VOS languages than they are of VSO languages.

Since Greenberg (1963), various word order properties have been claimed to correlate with VSO order. Table 2 provides data showing that the same characteristics are associated with VOS languages. The data in this table comes from a large database containing the word order properties of 772 languages (cf. Dryer 1988, 1989b, 1991, 1992). The languages are divided up into *genera*, language groups with a time depth roughly comparable to the subfamilies of Indo-European. The numbers in Table 2 represent the numbers of genera containing languages of the sort indicated. Counting genera rather than languages avoids counting two closely related languages with the same characteristics twice.<sup>5</sup>

The VSO columns in Table 2 show that VSO languages generally possess the characteristics listed on the lefthand side of the table: in each case the number of genera containing languages with the given characteristic is considerably larger than the number of genera containing languages with the opposite characteristic. In each case the VOS columns illustrate a similar tendency for VOS languages. Hence these properties can simply be treated as properties of languages which are VS or VO.<sup>6</sup>

### 3.3 *Argument 3: The instability of VSO vs. VOS*

The claim here that the difference between VSO and VOS languages is a relatively unimportant difference is further supported by considerable evidence that languages can easily change from one order to the other. There are a number of language families or subfamilies that contain both VSO and VOS languages. Within the Oceanic branch of Austronesian, there are many VSO languages (e.g. Hawaiian, Easter Island, Tahitian, Yapese, Niuean), but Kiribatese (also known as Gilbertese) is VOS. Both orders are also represented among more western Austronesian languages: Toba Batak and Malagasy are VOS, while a number of Philippine languages have been described as VSO (e.g. Agta, Pangasinan, Manobo, Mamanwa, Hiligaynon, and Tboli).<sup>7</sup> Among the Arawakan languages of South and Central America, Island Carib and Goajiro are VSO, while Baure is VOS. Among Mayan languages, some are VSO (e.g. Jacalteco, Mam, Ixil) while others are VOS (e.g. Tzotzil, K'ekchi).<sup>8</sup> Similarly among Aztec languages, both VOS (Pipil) and VSO (Huasteca Nahuatl) are represented. And among Oto-Manguean languages,



Table 2: Word Order Characteristics Typical of VSO an VOS languages

	VSO		VOS	
	With Property	With Opposite Property	With Property	With Opposite Property
Prep	20	4	12	0
NGen	24	3	11	3
NRel	22	0	6	0
ArtN	14	3	7	0
NumN	19	5	8	1
V-PP	20	0	8	0
NegV	23	1	10	0
AuxV	16	6	5	1
Initial Q	12	7	3	0
Initial <i>wh</i>	19	5	6	2

Key: Prep: prepositional vs. postpositional; NGen: noun-genitive vs. genitive-noun; NRel: noun followed by relative clause vs. noun preceded by relative clause; ArtN: article-noun vs. noun-article; NumN: numeral-noun vs. noun-numeral; V-PP: PP (adpositional phrase) following verb vs. PP preceding verb; NegV: negative word preceding verb vs. negative word following verb vs. PP preceding verb; NegV: negative word preceding verb vs. negative word following verb; AuxV: auxiliary verb or nonverbal tense-aspect particle preceding main verb vs. auxiliary verb or tense-aspect particle following main verb; Initial Q: sentence initial question particle in yes/no questions vs. sentence-final question particle; Initial *wh*: *wh*-expression obligatorily in sentence-initial position vs. *wh*-expression not obligatorily in sentence-initial position (typically *in situ*). Numbers represent numbers of genera containing languages of the given sort. Discrepancies in totals reflect the fact that many languages are not coded for all characteristics, for one of three reasons: (1) insufficient relevant information in my source; (2) both orders occur, with neither order basic (e.g., both orders of noun and genitive are common); (3) characteristic not applicable to the language (e.g., language lacks articles).

Otomi and Lealao Chinantec are VOS, while Mixtecan languages, Zapotecan languages, Tlapaneca, and some other Chinantecan languages (e.g., Comaltepec and Palantla Chinantec) are all VSO.

The last three genetic groups mentioned (Mayan, Aztecan, and Oto-Manguéan) all fall within the Meso-American area, which Campbell, Kaufman & Smith-Stark (1986) show exhibits a large number of areal

features, among them a strong tendency towards VO order. Their data further show that verb-initial order is especially common in this area, more common than SVO. But among the verb-initial languages in this area, we find both VSO and VOS sprinkled amongst each other, with both orders often found within the same family. And in addition to the examples of these orders cited above, Copainalá Zoque and Xinca are VOS. Similarly, within the Pacific Northwest, in which the languages are overwhelmingly verb-initial, both VSO and VOS are found: while Lower Chinook and Kutenai are VOS, various other groups in the area contain languages exhibiting VSO order: Salish (Squamish, Kalispel, Bella Coola), Wakashan (Kwakiutl), Chimakuan (Quileute), Tsimshian, and Sahaptin-Nez Perce (Northern Sahaptin). In fact the only area of the world in which verb-initial order is widespread and in which all the verb-initial languages in my database are VSO is north and north-east Africa. But the general pattern of finding VSO and VOS together suggests that it is relatively easy for languages to change between these two orders or that contact with a language which is VSO or VOS can influence change in a second language toward the other of the two types. These facts suggest that the difference between VSO and VOS is a minor one.

#### **4. Argument 4: The infrequency of clauses with a noun subject and a noun object**

There are certain characteristics of a language which one can easily see from examining a few pages of text with interlinear glosses, while there are others which rarely manifest themselves in texts but which may only manifest themselves under elicitation. While characteristics of the latter sort are often important for theoretical purposes, there is a sense in which characteristics of the former sort are more fundamental characteristics of a language. By examining a few pages of text, perhaps even just a few sentences, one can usually roughly determine the morphological type of the language (whether it is analytic, synthetic, or polysynthetic), whether the language is a so-called pro-drop language, and perhaps even a rough idea of whether the language is head-marking or dependent-marking (Nichols 1986, 1992). One can also get an idea as to how flexible the word order is in the language, and if it is less flexible, whether the language is OV or VO and whether it is SV or VS. But it is generally rather difficult to assign the language to one of the six types in the traditional word order typology. The

reason for this is that this typology is based on clauses containing a noun subject and a noun object but such clauses do not occur very often in texts and may occur even less often in natural conversation. In text counts of Yagua, a language of Peru, Payne (1990: 222) reports that in a corpus of 1516 clauses, only 49 (or 3%), contained both a noun subject and a noun object. Payne (1987: 792–793) reports very similar results from examination of a corpus of 759 clauses of Papago text: only 22 (or 3%) contained both a noun subject and noun object. Weber (1989: 15–16) reports a somewhat higher frequency of such clauses in his Huallaga Quechua texts (99, or 8%, out of 1309), but Du Bois (1987: 818) reports only 5 such clauses (or 1%) in a corpus of 443 Sacapultec clauses.

On the other hand, the number of clauses containing a noun subject *or* a noun object in these same texts was much higher. 35% of the Yagua clauses examined by Payne, 55% of the Huallaga Quechua clauses examined by Weber, and 54% of the Sacapultec clauses reported by Du Bois contained at least one of the subject or object. It is because of this that one can easily determine from a few pages of text how flexible the word order of a language is and if it is less flexible, whether the language is clearly OV or VO and whether it is SV or VS. Consider the following example from my own informal counts of texts in Tacana (Ottaviano 1980), a language of Bolivia.<sup>9</sup> In only a few pages of text, I found only two examples of clauses containing a noun subject and a noun object, one of them SOV, the other OSV. Two examples, especially when they exhibit different orders, are clearly not sufficient for determining what order is most frequent. On the other hand, in the same text, there were 32 clauses containing a noun subject, of which 26 (or 81%) were SV and 6 were VS. These numbers give us grounds for classifying the language as SV rather than VS, since the number of SV clauses exceeds the number of VS clauses by well over two-to-one. Of course, with such a small sample there is the danger that the greater number of SV clauses is just an accidental property of this particular small text, but we can clearly have much greater confidence in the claim that the language is SV than we can about which of the six traditional types it is. The number of clauses in the text examined that contained a noun object is much smaller, only 14, but since 12 of these (or 83%) were OV and only 2 VO, we again can tentatively classify the language as basically OV. On the basis of the text examined, the language can be tentatively classified by the typology proposed here (as OV and SV) but cannot be classified by the traditional six-way typology. If one accepts the assumption that a characteris-

tic of a language is in one sense fundamental if and only if it is the kind of characteristic that one can determine from examination of a few pages of text, then the order of the object and verb and the order of the subject and verb are fundamental characteristics, but the classification of a language by the traditional six-way typology is not.

Various other people (cf. Mithun 1987: 285; Campbell, Bubenik, and Saxon 1988: 210, footnote 2; Dixon 1988: 242; Du Bois 1987: 818) have observed the infrequency of clauses containing a noun subject and a noun object either in general or in particular languages and some have pointed out the shortcoming this implies for the traditional six-way typology. The implication some people seem to have drawn is that this indicates a shortcoming of any typology based on the order of subject and object with respect to the verb. But such an inference is not warranted: the typology proposed in this paper is based on the order of subject and object with respect to the verb, but depends only on the much more frequent clause types containing a noun subject *or* a noun object.

Authors of grammars often ignore such clauses in addressing questions of word order. For example, Donaldson (1980), in discussing word order in Ngiyambaa, an Australian language, notes the infrequency of transitive clauses containing two noun arguments in her texts and the variation even among the few examples that are found, and comes to no firm conclusion. But examination of clauses in the same texts containing one noun argument show that such clauses overwhelmingly place such arguments before the verb, showing quite clearly that the language is an SV&OV language.

## **5. Languages which cannot be classified by the traditional typology**

### *5.1 Argument 5: Languages which can only be classified by the proposed typology*

I argued in the last section that it is not possible to classify Tacana by the traditional typology on the basis of a small sample of text. It is quite possible that Tacana could be classified by the traditional typology if sufficient texts were examined. It might turn out, for example, that 80% of clauses containing a noun subject and a noun object are SOV. But there are many other languages with order so flexible that no single type among the six orders emerges as the most frequent by such a large margin. For exam-

ple, in Dryer (1983), I reported the relative frequencies of different orders in a set of texts from Frachtenberg (1913) in Hanis Coos (an extinct language of the coast of Oregon, sometimes classified as Penutian) cited in Table 3.

*Table 3: Word Order in Hanis Coos (from Dryer 1983)*

SVO	6 (38%)	SV	30 (23%)	OV	17 (30%)
VOS	4 (25%)	VS	98 (77%)	VO	39 (70%)
VSO	3 (19%)				
OVS	3 (19%)				

The lefthand columns in Table 3 show the relative frequency among the different clauses containing a noun subject and a noun object. While SVO is the most common, we cannot tell from such small numbers whether this is a statistical accident or whether it would be maintained over a larger body of texts. But even if it would be, even if we found the same proportions in a body of texts one hundred times greater than the body of texts examined, I think we should be hesitant describing a language as SVO when that order occurs only 38% of the time, even if at 38% that order is more frequent than any other order. Not only do such clauses constitute a minority among clauses with a noun subject and a noun object, but SVO is less than twice as common as VOS. If we adopt the assumption proposed earlier in this paper that one of two orders should be considered basic relative to another order only if it occurs at least twice as frequently as the other, then we cannot say that SVO order in Hanis Coos is basic relative to VOS. As a result, it seems best not to assign Hanis Coos to any of the six traditional types. Nor can we say the language is SVO/VOS, since the difference between VOS compared to VSO and OVS is small and may be accidental. At most we could say the word order is quite flexible, though verb-final orders (SOV and OSV) are rare.<sup>10</sup>

While there is little that we can do about classifying Hanis Coos according to the traditional six-way typology, the middle and righthand columns of Table 3 provide considerable basis for classifying the language as VS and VO, since these two orders are both more than twice as common as SV and OV respectively. Not only was the traditional typology unable to classify the language, but it also obscured a basic regularity about Hanis Coos order: noun arguments of the verb, both subject and object, more

commonly follow the verb, a regularity that is immediately clear from casual examination of a few pages of text in the language. Hanis Coos is thus an example of a language which cannot be classified by the traditional six-way typology but can by the typology proposed here.<sup>11</sup>

Remarkably similar numbers are reported by Payne (1987: 793–794) for Papago, shown in Table 4.

*Table 4: Word Order in Papago (Payne 1987)*

SVO	8 (36%)	SV	48 (23%)	OV	44 (29%)
OVS	5 (23%)	VS	158 (77%)	VO	108 (71%)
VOS	5 (23%)				
VSO	3 (14%)				
SOV	1 (5%)				

As with Hanis Coos, SVO is most common, but does not occur with sufficient frequency to justify classifying the language as SVO. But again VS and VO occur with sufficient frequency to justify treating them as basic.

A question that might arise at this point is whether it is enough to say that Hanis Coos and Papago are VS and VO. Do we not also want to say something about what conditions the choice of VS and VO as opposed to the less frequent orders SV and OV? The answer of course is yes. In the case of Hanis Coos, I offered an account of this in Dryer (1983): definite nominals generally follow the verb, while indefinite nominals generally precede. Payne (1987) offers a very similar account for word order in Papago, this presumably being the explanation for why the relative frequencies for these two languages are so similar. And clearly what determines the choice among the different orders is more important than the relative frequency of the different orders. But, as discussed above, relative frequency is also important: languages in which one order of verb and object is twice as common as the other generally have word order characteristics that are typically associated with the more common order. As noted above, there is reason to believe that in the case of Papago, the VS and VO orders are relatively recent innovations, within the past few hundred years. Significantly, the language seems to be acquiring other word order characteristics typically associated with VS&VO order. As discussed in Dryer (1989a), these changes can apparently be explained only by reference to the greater frequency of VS and VO order,

not by any reference to the discourse factors that determine the choice of order.

A third language which is difficult to classify by the traditional typology is Yagua (an isolate spoken in Peru, sometimes classified as Macro-Carib). The figures in Table 5 are based on extrapolations from data in Payne (1990: 223–225).

*Table 5: Word Order in Yagua (Payne 1990)*

VSO	19 (42%)	SV	115 (31%)	OV	59 (24%)
SVO	15 (33%)	VS	257 (69%)	VO	186 (76%)
OVS	11 (24%)				

While VSO order is somewhat more common in Yagua than SVO was in either Hanis Coos or Papago, it still represents a minority type, found in only 42% of clauses with a noun subject and a noun object. Payne argues that VSO is the pragmatically unmarked word order, but it is not clear that her use of this expression is consistent with how others use the term (cf. Dryer, *in press*), and the assumption of this paper that an order must be at least twice as frequent as another order to be basic relative to that order leads us to conclude that VSO cannot be basic relative to either SVO or OVS. In this and many other cases, it is not clear that there is much to be gained by forcing the language into exactly one type within the traditional typology, rather than leaving it unclassified according to that typology. But again the data from clauses containing a noun subject *or* noun object provides us with a good basis for describing Yagua as VS and VO: these orders outnumber SV and OV by over two-to-one.

Various other languages are similar to the above three languages in exhibiting word order that is sufficiently flexible as to make them difficult to classify by the traditional six-way typology, but in still being easy to classify by the typology proposed here. The Quechuan languages are generally described as SOV, but frequency data from Weber (1989: 16) suggests that Huallaga Quechua is a borderline case: SOV order occurs in 48% of clauses with a noun subject and object, barely less than half such clauses, and less than twice the frequency of the next most frequent order, SVO at 28%. However, the language is more clearly SV (72%) and OV (66%). My own informal counts of Ojibwa texts in Nichols (1988) reveal SVO order making up less than half (48%) of clauses with a noun subject

and a noun object, but we can more easily identify the language as VS and VO: 72% of subjects and 84% of objects are postverbal.<sup>12</sup> Similarly, my counts of texts from Tonkawa (an isolate spoken in Texas), show SOV amounting to 58% of the clauses in question, but this is less than twice the frequency for SVO order, which amounts to 37%. But the language is clearly SV (96%) and OV (86%).<sup>13</sup> And my counts of texts from Patwin (a California Penutian language) reveal 45% SOV and 36% SVO, again close enough to prevent one being treated as dominant. But 84% are SV and 74% OV, again allowing us to classify the language as SV and OV.<sup>14</sup>

It should be stressed that there are still many languages whose word order is sufficiently flexible that they cannot be classified according to the typology proposed here. The data in Table 6 (from Dryer 1985: 8), from the Auk dialect of Tlingit (a Nadene language of southeast Alaska), provides little basis for classifying the language for order of subject and verb or of object and verb. While OV is more common than VO, the difference is less than two-to-one and hence, by the assumptions of this paper, neither order is basic.

*Table 6: Word Order in Auk Dialect of Tlingit (Dryer 1985)*

SV	156 (47%)	OV	143 (58%)
VS	177 (53%)	VO	102 (42%)

My own informal counts suggest Alawa and Ngalakan (both Australian languages), Hupa (an Athapaskan language of California), and Wintu (a California Penutian language), are also instances of languages of this type, where both orders of subject and verb and of object and verb are sufficiently common that no order can be isolated as basic.

## *5.2 Argument 6: Languages that are partially classifiable by the proposed typology*

While there are some languages that can be classified by both the OV: VO parameter and the SV: VS parameter and others that can be classified by neither, there are still other languages that can be classified for one of these parameters but not the other. Chacobo (Prost 1967) and Amahuacan (Rus-



sell 1975) are two Panoan languages which are fairly rigidly OV but which are indeterminately SV/VS. Southern Barasano (Jones and Jones 1991), a Tucanoan language of Colombia, and Apalai, a Carib language of Brazil, both appear to be similar.<sup>15</sup> Scancarelli (1987: 187-188) reports counts for Cherokee in which 79% of clauses with objects are OV, but in which SV and VS occur with equal frequency. My own counts of texts in Korana, a Khoisan language of South Africa, showed 89% OV but only 61% SV.<sup>16</sup> Again the language is clearly OV, but it seems best to leave it unclassified for order of subject and verb. There are also languages which are clearly VO, but in which both orders of subject and verb are sufficiently common that neither order appears basic. The dialect of Arawak described by Pet (1987) is an example of this. Okrand (1977: 335) describes Mutsun (a California Penutian language) as an instance of this. My own counts from of Madimadi (a Pama-Nyungan language) show 85% VO, but only 60% VS.<sup>17</sup> My own counts for Shasta (a Hokan language spoken in northern California) are based on small numbers but suggest that it is similar, with 90% VO, but almost equal numbers of SV and VS.<sup>18</sup> Masakin (a Kordofanian language of Sudan) employs VSO order in the past tense, SVO in the present (Tucker and Bryan 1966: 287). It is clearly VO, but again indeterminately SV/VS.

All of the languages mentioned in the preceding paragraph can be assigned either to VO or to OV, but the order of subject and verb is noticeably freer. There are also languages of the opposite sort, in which one order of subject and verb is more common but in which both orders of verb and object are common. According to Miller (1965: 175), subjects in Acoma (a Keresan language of New Mexico) precede the verb, but both orders of object and verb are found. Logbara (Crazzolaro 1960, Weber 1994), a Nilo-Saharan language spoken in Uganda, is clearly SV, but best classified as OV/VO since it employs OV in clauses in 'incompleted' aspect, VO in 'completed' aspect. Dinka (Nebel 1948: 9, 40, 82; Tucker and Bryan 1966: 437), a Nilotic language, employs SVO order in simple main clauses, but S-Aux-O-V order in main clauses containing an auxiliary, a pattern reminiscent of German and Dutch. These languages are predominantly SV, but seem best classified as OV/VO.<sup>19</sup> My own counts of texts in Southern Sierra Miwok (a California Penutian language) show 80% SV but equal numbers of OV and VO.<sup>20</sup> My own text counts for Wiyot (an Algic language, related to Algonquian languages) showed a similar pattern with 83% SV but only 59% OV.<sup>21</sup> The typology proposed here, based on two separate parameters, order of object and verb and order of subject and verb, allows us to more easily

isolate regularities of word order in languages with flexible word order. It is not clear, for most of the languages cited here, that they can be classified by the traditional typology. The proposed typology allows it to be clear exactly where one can assign a basic order.

## 6. **Argument 7: The importance of the OV: VO parameter**

While Greenberg categorized languages as SOV, SVO, or VSO (his sample did not contain languages of any of the other three types), the work of Lehmann (1973, 1978) and Vennemann (1974, 1976) has emphasized the importance of the simpler question of whether a language is OV or VO. Underlying this is the claim that among the three types Greenberg considered, SVO languages pattern much more like VSO languages than they do like SOV languages and that the order of subject and verb is much less important than the order of object and verb. By distinguishing these two parameters, as in the typology proposed here, we explicitly separate out the more important parameter from the less important one.

A number of linguists (Hawkins 1983; Comrie 1989: 96, 100–101; Mallinson and Blake 1981: 379) have criticized Lehmann and Vennemann for collapsing SVO, VSO, and VOS languages into a single category VO. They all argue that SVO is a mixed type, intermediate between verb-initial and verb-final. However, as argued in detail in Dryer (1991), this claim is not supported by the evidence from the large set of languages in my database: in most respects SVO languages differ little from verb-initial languages. It is shown in that paper, for example, that while verb-final languages are predominantly postpositional (96%), SVO languages are like verb-initial languages in being predominantly prepositional (86% for SVO, 91% for verb-initial). Similarly, while prenominal relative clauses are common in verb-final languages (43%), they are rare in both SVO languages (1%) and verb-initial languages (0%). For various other characteristics too, SVO languages pattern much like verb-initial languages. This demonstrates the validity of the claim that there is a fundamental distinction between VO and OV languages, a distinction which is isolated in the typology proposed here.

It is shown in Dryer (1991) that there are a few characteristics for which SVO languages do exhibit properties intermediate between those of verb-initial and verb-final languages. For example, verb-initial languages generally place the genitive after the noun and verb-final languages generally place the

genitive before the noun, but the two orders of genitive and noun are about equally common among SVO languages. This shows that for some characteristics, both the order of object and verb and the order of subject and verb are important. In other words, there are some characteristics for which both of the parameters in the typology proposed in this paper are relevant and other characteristics for which only the order of object and verb is relevant. The typology proposed in this paper makes this distinction clear by separating out these two parameters, and particularly in separating out the order of object and verb, which is the more predictive and hence more important parameter.

### **7. Argument 8: Differences between transitive and intransitive subjects**

The final argument against the traditional six-way typology is perhaps the strongest one. The traditional typology is based entirely on the word order of transitive clauses and ignores intransitive clauses. For many languages, this is fairly harmless, since the position of intransitive subjects relative to the verb in many languages is identical to the position of transitive subjects relative to the verb. But one does not need to look far to find languages in which this is not the case. VS order is quite common in a number of the languages of Europe, such as Spanish and Polish, sufficiently common that questions arise as to whether we are justified in describing these languages as SV. Moreover, the frequency of intransitive VS clauses in these languages is one of the main reasons for the high frequency of VS order. As discussed by Jacennik and Dryer (1992) for Polish and by Benvoglio and Weber (1986) for Spanish, the conditions favouring VS order include certain kinds of verbs that are generally intransitive. Jacennik and Dryer show that in the Polish corpus they examined, VS clauses were more often intransitive while SV clauses were more often transitive. On the other hand, although VS order is sufficiently common that it is questionable whether we are justified in describing these languages as SV, the order SVO is sufficiently common in clauses containing a noun subject and a noun object that we are justified in describing these languages as SVO. In the case of Polish, Siewierska (1993: 235) cites statistics on the relative frequency of the six orders of subject, object, and verb in a corpus containing 459 clauses containing an overt subject and object: SVO 72.5%, VOS 9.5%, OVS 7.4%, VSO 6.5%, SOV 2.4%, OSV 1.5%.<sup>22</sup> Since the frequency of SVO outnumbered all other orders combined by a ratio of over two-to-one, we are justified

by the criteria of this paper in describing Polish as SVO. But describing Polish as SVO is potentially misleading for two reasons. First, there is a tendency to infer from a description of Polish as SVO that it is SV; but, as I have argued, it is questionable whether we are justified in describing Polish as SV.<sup>23</sup> Second, while it is strictly correct that Polish is SVO, this fact is less significant than the fact that VS order is common, because VS clauses are more common than SVO clauses, since clauses with just a subject are more common than clauses with a noun subject and a noun object. Similar comments apply to Spanish: although VS is very common in intransitive clauses in Spanish, over 88% of clauses containing a subject and object in a sample of spoken Spanish discussed by Ocampo (1989) are SVO.<sup>24</sup>

A particularly extreme example of a language in which transitive and intransitive subjects exhibit very different word order properties is Salinan, an extinct Hokan language of California. As shown in Dryer (1989c), among clauses containing a noun subject and a noun object, SVO is clearly the dominant order, as illustrated by the lefthand columns in Table 7.<sup>25</sup>

Table 7: Word Order in Salinan (Dryer 1989c)

SVO	17 (81%)	S <sub>t</sub> V	24 (69%)	S <sub>i</sub> V	18 (11%)
VSO	2 (10%)	VS <sub>t</sub>	11 (31%)	VS <sub>i</sub>	147 (89%)
VOS	1 (5%)				
SOV	1 (5%)				

(S<sub>t</sub> = transitive object, S<sub>i</sub> = intransitive subject)

The lefthand columns of Table 7 provide ample justification for describing Salinan as SVO, since it occurs in the text examined more than four times as often as all other orders combined. The middle columns of Table 7 show that if we simply examine the relative order of transitive subjects with respect to the verb, SV order is more common than VS, by a ratio of just over two-to-one. The number of clauses included in these columns is higher than the number in the lefthand columns because these columns include transitive clauses that do not have a noun object, for example where the object is indicated by a pronominal affix on the verb. But when we examine the righthand columns, showing the relative order of *intransitive* subjects with respect to the verb, we find that there is an overwhelming tendency for intransitive subjects to follow the verb. Simply describing Salinan as SVO is

misleading, if no mention is made of the fact that intransitive clauses are overwhelmingly VS.

Various other languages have been described as exhibiting differences in the position of transitive as opposed to intransitive subjects. Stairs & de Hollenbach (1981: 335) describe Huave (an isolate spoken in southern Mexico, possibly related to Mayan and/or Mixe-Zoquean), as SVO, but state that the more common order in intransitive clauses is VS. Similar facts are reported for Huastec (a Mayan language, by Edmonson 1988: 565), Tepehua (a Totonacan language, by Watters 1988: 12), Michoacan Nahuatl (by Sischo 1977: 313), Paumar' (an Arauan language of Brazil, by Chapman and Derbyshire 1991: 163–166), and Muna (an Austronesian language of Sulawesi in Indonesia, by Van den Berg 1989: 150). Engel and Longacre (1963: 335–6) say that subjects more often follow intransitive predicates in Ostuacan Zoque but that subjects of transitive verbs precede or follow the verb with equal frequency. Merlan (1982: 26) describes OVS as the more common order in Mangarayi, a non-Pama Nyungan language of Australia, but she states that intransitive clauses are more often SV. And Pări (a Nilotic language) is similar to Mangarayi except that the OVS order in transitive clauses and SV order in intransitive clauses is fairly rigid (Andersen 1988). Describing Mangarayi and Pări as OVS is strictly speaking correct, but misleading, since intransitive clauses are more often SV. In other words, the word order in these languages is really following an ergative pattern, and might be better described as Abs-V-Erg. The fact that the word order in these languages patterns ergatively is quite plausibly one factor to be considered in explaining why they exhibit the highly unusual object-initial order. But the ergative pattern requires that we consider intransitive subjects as well as transitive ones.

Text count differences between transitive and intransitive subjects may justify treating one of these as having a normal order with respect to the verb but not the other. My own counts of Korana (Khoisan) reveal 86% SV among transitive clauses, but only 57% SV among intransitive clauses.<sup>26</sup> Conversely, the counts in Payne (1987) for Papago show 80% VS for intransitive clauses but only 59% VS for transitive clauses. The overall figure for Papago is 77% VS, but this is potentially misleading since the figure is much lower for transitive clauses. The Yagua data in Payne (1990) show 74% of intransitive clauses VS but only 55% of transitive clauses as VS. As noted above, Payne argues for VSO as the pragmatically unmarked word order of Yagua, but the fact that only 55% of transitive subjects follow

the verb raises questions about this claim.

By ignoring the position of intransitive subjects, the traditional typology is based entirely on one of the two types of subjects. But the situation is made even worse by the fact that intransitive noun subjects, the class of subjects that the traditional typology ignores, are much more common in texts than transitive noun subjects. The data for Salinan in Table 7 show that in the texts examined, intransitive subjects outnumber transitive subjects by almost five to one. Table 8 shows the relative frequency of transitive as opposed to intransitive subjects in text counts that have been made for a variety of languages.

*Table 8: Relative Frequency of Transitive and Intransitive Subjects*

	$S_t$	$S_i$	Ratio of $S_i$ to $S_t$	Source
Papago	37	169	4.6	Payne (1987)
Yagua	88	292	3.3	Payne (1990)
Huallaga Quechua	150	280	1.9	Weber (1989)
Sacapultec	11	126	11.5	Du Bois (1987)
Carib	64	131	2.0	Hoff (1978)
Tacana	8	24	3.0	Own counts from Ottaviano (1980)
Korana	7	35	5.0	Own counts from Meinhof (1930)
Tonkawa	15	40	2.7	Own counts from Hoijer (1972)
Patwin	19	38	2.0	Own counts from Whistler (1997)
Wiyot	6	29	4.8	Own counts from Reichard (1924)
Ojibwa	53	195	3.7	Own counts from Nichols (1988)
Salinan	35	165	4.7	Dryer (1989c)

The ratio of the frequency of intransitive noun subjects to that of transitive noun subjects varies from 1.9 to 11.5 in Table 8, with a mean of 4.1. This result is consistent with the general claim of Du Bois (1987) that subjects of transitive clauses are typically pronominal. Thus not only does the traditional typology ignore one class of subjects, but it ignores the class of subjects that is more common.

Two final examples from the literature will illustrate the extent to which intransitive subjects are overlooked in the literature. In his discussion of clause order in a description of Rumanian in the Croom-Helm series of

grammars, Mallinson (1986: 91) describes the language as SVO, but makes no mention of the fact that it is very common for subjects to follow the verb in the language. But in counts of Rumanian reported in Myhill (1986), VS order was found in 33% of clauses lacking an object. One would not have guessed this from Mallinson's description.<sup>27</sup> Second, in their discussion of the distribution of word order characteristics in the Meso-American area, Campbell, Kaufman & Smith-Stark (1986: 547) restrict attention to word order in transitive clauses and focus on the pattern of VO order throughout the area. But VS order in intransitive clauses also appears to be especially common throughout the area, not only among the verb-initial languages but also among the SVO languages. As noted above, Huave, Huastec, Tepehua, and Michoacan Nahuatl are all SVO in transitive clauses but VS in intransitive clauses. All four of these are spoken in Meso-America, but are in different genera (Huave, Mayan, Totonacan, and Aztecan respectively). Campbell et al list Tequistlatec as SVO, but Waterhouse (1962: 359) says that VS is the most common order when the verb is intransitive. By attending only to word order in transitive clauses, Campbell et al apparently missed further support for the areal patterns they were documenting: not only are VSO and VOS orders common in this area but VS order for intransitive clauses is even more common.

Thus the traditional typology not only ignores the order of intransitive subject and verb (despite the fact that these constitute the most common type of subject), but also leads linguists to forget intransitive subjects altogether, to think that in having described the order of clauses containing a noun subject and object they have thereby described the order of clauses containing a subject. But at most they have described the order of clauses containing a transitive subject.<sup>28</sup> Admittedly, these differences between transitive and intransitive subjects also present a problem for the typology being proposed here: if transitive and intransitive subjects pattern differently, that difference cannot be captured in a single statement regarding the order of subject and verb. However, in cases where such differences arise, our answer to the question as to what the order of subject and verb is should be that the situation is more complex, and that we need to break the question down into two subquestions, one regarding the order of position of intransitive subjects, the other regarding the position of transitive subjects. The traditional typology may yield an answer that is technically correct, like saying Salinan is SVO, but very misleading, since it leads people to make incorrect inferences about the position of intransitive subjects.

## 8. An alternative three-way typology

Some of the arguments in this paper against the traditional six-way typology are consistent with an alternative typology which has often been assumed in discussions of word order typology (e.g. Hawkins 1983; Nichols 1986, 1992), namely a three-way typology that distinguishes *verb-initial*, *verb-medial*, and *verb-final*. The typology proposed in this paper roughly corresponds to this typology in that the type SV&OV corresponds roughly to verb-final, the type VS&VO to verb-initial, and the type SV&VO to verb-medial. The most significant difference, however, is that what is a rare fourth type in the typology proposed here, namely VS&OV, corresponding to the traditional type OVS, is collapsed with SVO in the three-way typology into the type *verb-medial*. But collapsing SVO and OVS into a single type verb-medial is apparently completely unmotivated. In fact, from the limited evidence available for OVS languages, they do not exhibit properties associated with SVO languages: while SVO languages typically resemble VS&VO languages, as discussed above and in Dryer (1991), OVS languages often exhibit properties associated with SV&OV languages. For example, the clearest case of an OVS language, Hixkaryana (Derbyshire 1979), has postpositions and has inflected auxiliaries following the main verb, both properties that are relatively unusual for SVO languages. It is true that Hixkaryana also has some properties that are atypical of SV&OV languages: it normally places adpositional phrases and manner adverbs after the verb. The two other languages in my data base that could be classified as OVS are Mangarayi (Merlan 1982) and Pări (Andersen 1988), but as mentioned above, both of these languages are actually OVS for transitive clauses but SV for intransitive clauses, and thus are OVS only in a potentially misleading sense. Mangarayi exhibits a mixture of OV and VO characteristics: it is prepositional (a VO characteristic) but places adpositional phrases and manner adverbs before the verb and inflected auxiliaries after the main verb (all OV characteristics). Apart from OV order, Pări exhibits properties associated with VO languages, thus resembling other Nilotic languages, which are generally VO. While these languages provide little basis for saying what is typical of OVS languages, two of them (Hixkaryana and Mangarayi) are clearly unlike SVO languages and there is thus good reason to be suspicious of a verb-medial type that collapses OVS with SVO.

While the typology proposed here may appear to resemble the three-way typology apart from not collapsing SVO and OVS, it differs in another more



subtle but significant respect. Namely, unlike the typology proposed here, the three-way typology is like the six-way typology in being based entirely on the order of transitive clauses with noun subject and noun objects. But, as argued above, this is an infrequent clause type and any typology based on such clauses ignores the order in intransitive clauses. The typology proposed here, however, is based on both types of clauses. There are thus a number of languages which are SVO on the six-way typology, and thus verb-medial on the three-way typology, but which are *not* SV&VO on the typology proposed here: a number of languages discussed in the preceding section are examples of this, since although they are SVO, VS order is sufficiently common in intransitive clauses that they are not classified as SV. Similarly, both Pări and Mangarayi count as verb-medial languages on the three-way typology, but not as VS&OV languages on the typology proposed here, since intransitive subjects in these languages more often precede the verb; they are thus classified as SV/VS & OV. Salinan is an example of a language which can be accurately described as SVO while still being VS&VO on the typology proposed here. If we employ the term *verb-initial* as a label for VS&VO languages (rather than for languages which are VSO or VOS), then Salinan is an instance of a verb-initial language, a label which is justified given the fact that the verb normally precedes both subjects and objects. It differs from other verb-initial languages only in the relatively minor property that noun subjects in transitive clauses containing a noun object normally precede the verb. But this is no different from other instances in which verb-initial languages put the subject before the verb under certain special conditions. Clauses containing a noun subject and a noun object are fairly infrequent and the preverbal placement of subjects in clauses containing a noun object is just one kind of special condition. The problem with both the traditional six-way typology and the three-way typology is that they treat the order in such infrequent clauses as a fundamental typological parameter.

## 9. Conclusion

I have given a series of arguments in this paper against the traditional six-way typology and in favour of an alternative typology that classifies languages as OV or VO and as SV or VS. I have argued that this alternative typology allows for a superior classification in a number of ways. For one thing, the classification yields types that are more relevant to typological

predictions: it collapses VSO, VOS, and indeterminately VSO/VOS, all of which exhibit similar characteristics, into a single VS&VO type; and it isolates the OV: VO parameter, which is fundamental to predictions in word order typology. Second, it allows classification of many languages which are not classifiable by the traditional typology. Third, it does not depend on a clause type that is infrequent in actual language use, namely clauses with a nominal subject and a nominal object. And fourth, it includes consideration of the position of intransitive subjects, which are ignored entirely by the traditional typology.

The traditional typology treats the differences between SOV and SVO, between SVO and VSO, and between VSO and VOS on a par. But the first of these differences is a fundamental one, since they differ in the order of verb and object. The second of these differences is intermediate in importance; they are similar with respect to the important parameter of order of verb and object but they differ with respect to the lesser parameter of order of verb and subject. The third of these differences is the least important and it is ignored in the typology proposed here because it is not a difference that is predictive of anything else.

None of the arguments in this paper should be interpreted as arguing against the traditional taxonomy *per se*, but only as arguments against that taxonomy as reflecting a fundamental property of languages. Statements about the basic order of clauses containing a noun subject and a noun object are ultimately things that must be said about a language, just as we must say how comparisons are expressed in a language or what positions can be relativized. The fact that Salinan clauses with a noun subject and a noun object are normally SVO is a real property of the language. It is just not among the fundamental properties of the language. The fact that both subjects and objects in Salinan most commonly follow the verb, *is* a fundamental property, and it is this kind of fundamental property that the typology proposed in this paper focuses on.

*Author's address*

Matthew S. Dryer  
Dept. of Linguistics  
SUNY Buffalo  
NY 14260, USA  
Email: [lindryer@acsu.buffalo.edu](mailto:lindryer@acsu.buffalo.edu)

## ACKNOWLEDGMENTS

I am indebted to Bill Croft for useful comments on an earlier draft of this paper, and to two anonymous reviewers who provided detailed comments.

## NOTES

1. As discussed below, however, these correspondences between the two typologies are not as exact as might appear at first.
2. I assume here a use of the term *basic word order* which can be applied to the order of any pair or set of elements, and that one can ask of a language what the basic order is of, say, adjective and noun, or negative and verb. The term is sometimes used to refer specifically to the order of subject, object, and verb in transitive clauses. Note, for example, the title of Tomlin (1986) (*Basic Word Order: Functional Principles*), which deals exclusively with the order of subject, object, and verb.
3. See Dryer (1995) for discussion of how the pragmatically unmarked word order in a language need not be the most frequent order. However, most attempts at characterizing a notion of basic word order distinct from the most frequent order prove problematic on close examination. For example, England (1991) argues for either VSO or VOS order as basic in a number of Mayan languages, even though SVO is said to be most frequent. Among her arguments for treating SVO as nonbasic in a number of these languages is the fact that SVO involves “topicalization”. However, her discussion and that in a number of specific descriptions of Mayan languages suffer in a number of ways. First, in general, no even vaguely satisfying characterization of what is meant by “topic” is provided, which renders the claims obscure, given the wide range of uses of the term. Second, no or little evidence is given to support claims that the subject in SVO clauses is “topic” while the subject in VOS or VOS orders is not. And third, no or little argument is given for assuming that clauses in which the subject is “topic” are marked or nonneutral. If subjects in SVO clauses in these languages are topics and SVO order is most frequent, then whatever pragmatic property characterizes “topics” is one that subjects apparently generally have, which makes it unclear on what basis it is claimed that clauses in which the subject is “topic” are marked or nonneutral.
4. Various other people (e.g. Hawkins 1983) have also treated VSO and VOS as subtypes of a single type verb-initial. In section 8 below, I discuss an alternative three-way typology that distinguishes verb-initial, verb-medial, and verb-final languages.
5. See Dryer 1989b, 1992 for further discussion.
6. Greenberg’s Universal 17 claims that VSO languages tend to be NAdj. I have shown in Dryer (1988, 1992) that the order of adjective and noun does not correlate with the order

of object and verb and that the adjective follows the noun as often in verb-final languages as it does in verb-initial languages. Curiously, when we compare the VSO languages in my database with the VOS languages, we seem to find a difference as far as the order of adjective and noun is concerned: while NAdj outnumbers AdjN by 19 genera to 9 among VSO languages, AdjN outnumbers NAdj by 7 to 6 among VOS languages. I suspect that this difference is accidental and not indicative of anything linguistic. To some extent it reflects the fact that I have no VOS languages from Africa, where NAdj order is more common among the VSO languages by 8 genera to 2: if we leave out Africa, the preference for NAdj order among VSO languages drops to 11 vs. 7. On the other hand, we find a curious contrast in Meso-America: except for Jacaltepec, the VSO languages in my database from this area are NAdj (Huasteca Nahuatl, Cora, Jicaltepec Mixtec, Peoles Mixtec, Ocotepéc Mixtec, Comaltepec Chinantec, Palantla Chinantec, Quiotepec Chinantec, Zapotec, Chatino, and Tlapaneca), while, except for Lealao Chinantec, the VOS languages from this area are AdjN (Copainalá Zoque, Otomi, Pipil). In addition, Campbell, Bubenik, and Saxon (1988) mention two other VOS& AdjN languages of Meso-America not in my database that conform to this generalization: Xinca and Chontal of Tabasco (a Mayan language, not to be confused with Chontal of Oaxaca, a non-Mayan Tequistlatecan language). While I suspect that this pattern is accidental, it bears further investigation.

7. While there may be questions about the application of the term *subject* to some of these languages (cf. Schachter 1976, 1977; Mulder & Schwartz 1981), they are apparently VSO in the sense that the agent/experiencer precedes the other argument.
8. Much of the discussion of word order in Mayan languages (e.g., England 1991) assumes a notion of basic word order distinct from the one assumed here based on frequency and in some Mayan languages which are described as VSO or VOS, SVO order is described as being the most frequent order. In the languages cited here, however, SVO order appears to be less frequent than the verb-initial orders.
9. The accuracy and significance of the counts performed by myself for Tacana and a number of other languages cited below should be taken with a grain of salt, because of the small sample of texts examined and my very superficial familiarity with these languages. I use these examples to illustrate a general theoretical point. If the conclusions I draw about these particular languages are inaccurate, it is likely that similar conclusions could be accurately drawn for other languages.
0. Though not attested in the body of texts reported by Dryer (1983), which provide the data in Table 3, SOV order is found infrequently in other Hanis Coos texts.
  1. Hanis Coos is one of the languages that Mithun (1987) argues lacks a basic word order.
  2. The actual counts for Ojibwa are 12 SVO, 5 VOS, 4 VSO, 3 OVS, and 1 SOV; 72 SV and 186 VS; 25 OV and 130 VO. Proulx (1991: 200) cites somewhat similar figures for Ojibwa clauses with noun subject and noun object, based on texts in Bloomfield (1957), namely 17 SVO, 7 VSO, 2 OVS, and 1 VOS. Both my counts and Proulx's show SVO considerably more frequent, though the percentage SVO is higher in Proulx's counts (63%). He notes that these figures call into question claims that Ojibwa is basically VOS (cf. Tomlin and Rhodes 1979). VOS order is somewhat more common in my counts than in Proulx's, though clearly less frequent than SVO. It should be noted, however, that

Tomlin and Rhodes made no claim that VOS is most frequent, only that it was basic. Their notion of basic word order is different from the one assumed in this paper.

13. The actual counts for Tonkawa are 11 SOV, 7 SVO, and 1 OSV, with SV outnumbering VS 53 to 2 and OV outnumbering VO 24 to 4, based on texts in Hoijer (1972).
14. The actual counts for Patwin (based on a text in Whistler 1977) are 5 SOV, 4 SVO, 2 OVS; 48 SV, 9 VS; 32 OV, 11 VO; .
15. Both of these languages have been claimed to be OVS, but the evidence for this is not convincing. Jones and Jones (1991: 2) claim that Southern Barasano is OVS, but this claim is apparently based on the fact that independent subject pronouns generally follow the verb; nominal subjects are apparently common in both preverbal and postverbal position. Similarly Koehn and Koehn (1986: 34) claim that the basic order of Apalai is OVS, but in the text they examined, there were only four instances of clauses with a free subject and free object and of these three were OVS and one SOV, and they give no reason beyond this for classifying the language as OVS. It is certainly possible that further evidence might support their claim, but the available evidence is inconclusive. They also note that the texts they examined contained five transitive clauses without a free object and that all five of these were SV. For intransitive clauses, both orders of verb and subject are apparently common. Apalai provides a further example of the difficulties in attempting to classify a language according to the six-way typology due to its infrequency in texts. On the typology proposed here, the language is unambiguously OV and SV/VS.
16. The actual counts for Korana are 24 OV, 3 VO, 28 SV, and 18 VS, based on texts in Meinhof (1930).
17. The actual counts for Madimadi are 18 SV, 12 VS, 33 VO, and 6 OV, based on texts in Hercus (1986)
18. The actual counts for Shasta are 13 SV, 11 VS, 1 OV and 9 VO, based on a text in Silver and Wicks (1977).
19. Note that unlike Dutch and German, which are consistently SV and OV in subordinate clauses, Dinka is VSO for subordinate clauses not containing an auxiliary and Aux-S-O-V for subordinate clauses containing an auxiliary (Nebel 1948: 25, 75).
20. The actual counts for Sierra Miwok are 20 SV, 5 VS, 13 OV, and 12 VO, based on texts in Broadbent (1964).
21. The actual counts for Wiyot are 29 SV, 6 VS, 22 OV, and 15 VO, based on texts in Reichard (1924).
22. Siewierska's data apparently includes clauses with independent pronouns as subject or object. Based on data discussed by Jacennik and Dryer (1992), I suspect that the frequency of SVO would be even higher if one were to restrict attention to clauses containing a nominal subject and a nominal object.
23. In Siewierska's corpus, among all clauses, including both transitive and intransitive, SV outnumbers VS by 68% to 32%. By the two-to-one ratio criterion, this would mean that Polish is barely eligible to be classified as SV.
24. Ocampo's sample includes clauses with independent pronouns as subject or object but it

is clear from his data that clauses containing a noun subject and a noun object are also overwhelmingly SVO.

25. The figures in the middle column of Table 7 for transitive subjects include the clauses in the first column in which there is also a noun object.
26. The actual counts for Korana are 20 S<sub>i</sub>V, 15 VS<sub>i</sub>, 6 S<sub>i</sub>V, and 1 VS<sub>i</sub>.
27. Mallinson does note (p. 92) that the fronting of a constituent in a clause containing an object can result in the subject following the verb, but this leaves unaddressed the word order of clauses not containing a nominal object, either transitive or intransitive. He does refer to VS order elsewhere in his grammar, but given the questionnaire-answering format of his grammar, a reader seeking information about basic clause order is *unlikely* to notice these comments. He mentions in a section on emphasis (p. 183) that VS order is common. He also mentions in a section on topics (p. 190) that “verbs of coming into existence” normally precede their subject. Later in the same section, in a subsection on topicalizing adverbials (p. 192), he notes that when “a predicate adverbial is topicalized by fronting ..., the subject must be moved after the verb.” Similarly, he notes (p. 194) that “the usual effect of object fronting is for the subject to move into postverbal position.” Finally, he states, in a section on heavy shift processes (p. 195), that “in addition to unmarked SVO order, statements can also have V-S order ...” The placement of this discussion in sections discussing emphasis, topicalization, and heavy shift processes all conspire to suggest that VS order is a clearly marked order. Since Mallinson applies the term “unmarked” only to SVO clauses, the question of an unmarked or normal order for clauses not containing an object does not arise.
28. In fact, they may not even have done that. In some languages the word order of clauses containing a transitive subject but no free object pattern more like intransitive clauses than like transitive clauses containing a nominal object. In Salinan, for example (see Table 7), transitive subjects of clauses lacking a nominal object tend to follow the verb, like intransitive subjects, and unlike transitive subjects of clauses containing a nominal object. Myhill (1986: 340) reports similar facts for Rumanian.

## REFERENCES

- Adam, Lucien. 1897. *Materiaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille kariri*. Paris: J. Maisonneuve [Bibliothèque Linguistique Américaine]. [Reprinted 1968, Nendeln/Liechtenstein: Kraus].
- Andersen, Torben. 1988. “Ergativity in Pāri, a Nilotic OVS language. *Lingua* 75: 289–324”.
- Bentivoglio, Paola, and Elizabeth G. Weber. 1986. “A functional approach to subject word order in spoken Spanish”. In: Jaeggli, Osvaldo, and Silva-

- Corvalan, Carmen (eds.), *Studies in Romance linguistics*, 23–40. Dordrecht: Foris.
- Bloomfield, Leonard. 1956. *Eastern Ojibwa: grammatical sketch, texts, and word list*. Ann Arbor: University of Michigan Press.
- Broadbent, Sylvia M. 1964. *The Southern Sierra Miwok language*. Berkeley: University of California Press.
- Campbell, Lyle, Vit Bubenik, and Leslie Saxon. 1988. “Word order universals: refinements and clarifications”. *Canadian Journal of Linguistics* 33: 209–230.
- Campbell, Lyle, Terrence Kaufman, and Thomas C. Smith-Stark. 1986. “Meso-America as a linguistic area”. *Language* 62: 530–570.
- Chapman, Shirley, and Desmond C. Derbyshire. 1991. “Paumari”. In: Derbyshire, Desmond C. and Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages* 3, 161–352. Berlin: Mouton de Gruyter.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*. Chicago: University of Chicago Press. Second Edition.
- Crazzolara, J. P. 1960. *A study of the Logbara (Ma'di) language*. London: Oxford University Press.
- Derbyshire, Desmond. 1979. *Hixkaryana*. Amsterdam: North-Holland [Lingua Descriptive Studies 1].
- Dixon, Robert M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Donaldson, Tamsin. 1980. *Ngiyambaa: The language of the Wangaaybuwan*. Cambridge: University Press.
- Dryer, Matthew S. 1983. “Coos word order”. Paper delivered at the Western Conference on Language, Eugene, Oregon.
- Dryer, Matthew S. 1985. “Tlingit: an object-initial language?” *Canadian Journal of Linguistics* 30: 1–13.
- Dryer, Matthew S. 1988. “Object-verb order and adjective-noun order: dispelling a myth”. *Lingua* 74: 77–109.
- Dryer, Matthew S. 1989a. “Discourse-governed word order and word order typology”. *Belgian Journal of Linguistics* 4: 69–90.
- Dryer, Matthew S. 1989b. “Large linguistic areas and language sampling”. *Studies in Language* 13: 257–292.
- Dryer, Matthew S. 1989c. “Salinan word order”. In: Delancey, Scott (ed.), *Papers from the 1988 Hokan-Penutian Languages Workshop*, 40–49. Eugene: Univ. of Oregon [University of Oregon Papers in Linguistics, Publications of the Center for Amerindian Linguistics and Ethnography 1].

- Dryer, Matthew S. 1991. "SVO languages and the OV: VO typology". *Journal of Linguistics* 27: 443–482.
- Dryer, Matthew S. 1992. "The Greenbergian word order correlations". *Language* 68: 81–138.
- Dryer, Matthew S. 1995. "Frequency and pragmatically unmarked word order". In: Downing, Pamela, and Noonan, Mickey (eds.), *Word order in discourse*, 105–135. Amsterdam: John Benjamins.
- Du Bois, John W. 1987. "The discourse basis of ergativity". *Language* 63: 805–855.
- Edmonson, Barbara Wedemeyer. 1988. *A descriptive grammar of Huastec (Potosino dialect)*. Tulane University dissertation.
- Engel, Ralph, and Robert E. Longacre. 1963. "Syntactic matrices in Ostuacan Zoque". *International Journal of American Linguistics* 29: 331–334.
- England, Nora. 1991. "Changes in basic word order in Mayan languages". *International Journal of American Linguistics* 57: 446–486.
- Frachtenberg, Leo J. 1913. *Coos texts*. New York: Columbia University Press [Columbia University Contributions to Anthropology].
- Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". In: Greenberg, Joseph (ed.), *Universals of language*, 73–113. Cambridge, Mass: MIT Press.
- Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.
- Hercus, Luise. A. 1986. "Outline of the Madimadi language". In: *Victorian languages: a late survey*, Pacific Linguistics, Series B, No. 77, pp. 101–151. Revised edition.
- Hoff, B. J. 1978. "The relative order of the Carib finite verb and its nominal dependents." In: F. Jansen, ed., *Studies on fronting*, pp. 11–28. Lesse: The Peter de Redder Press.
- Hoijer, Harry. 1972. *Tonkawa texts*. Berkeley: University of California Press [University of California Publications in Linguistics 73].
- Jacennik, Barbara, and Matthew S. Dryer. 1992. "Verb-subject order in Polish". In: Payne, Doris L. (ed.), *Pragmatics of word order flexibility*, 209–241. Amsterdam: John Benjamins.
- Jones, Wendell, and Paula Jones. 1991. *Barasano syntax*. Summer Institute of Linguistics and The University of Texas at Arlington.
- Keenan, Edward. 1978. "The syntax of subject-final languages". In: Lehmann, Winfred (ed.), *Syntactic Typology*, 267–327. Austin: University of Texas Press.



- Koehn, Edward, and Sally Koehn. 1986. "Apalai". In: Derbyshire, Desmond C. and Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages 1*, 3–127. Berlin: Mouton de Gruyter.
- Lehmann, Winfred P. 1973. "A structural principle of language and its implications". *Language* 49: 42–66.
- Lehmann, Winfred P. 1978. "The great underlying ground-plans". In: Lehmann, Winfred P. (ed.), *Syntactic typology*, 3 — 55. Austin: University of Texas Press.
- Li, Paul. 1973. *Rukai structure*. Taipei: Institute of History and Philology Academia Sinica [Special Publication No. 64].
- Mallinson, Graham. 1986. *Rumanian*. London: Croom Helm.
- Mallinson, Graham; and Barry J. Blake. 1981. *Language typology*. Amsterdam: North-Holland.
- Meinhof, Carl. 1930. *Der Koranadialekt des Hottentottischen*. Berlin: Verlag von Dietrich Reimer (Ernst Vohsen) A. G [Reprinted 1969, Nendeln, Liechtenstein: Kraus].
- Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland [Lingua Descriptive Studies 4].
- Miller, Wick R. 1965. *Acoma grammar and texts*. Berkeley: University of California Press.
- Mithun, Marianne. 1987. "Is basic word order universal?" In: Tomlin, Russell (ed.), *Coherence and Grounding in Discourse*, 281–328. Amsterdam: John Benjamins.
- Mulder, Jean; and Arthur Schwartz. 1981. "On the subject of advancements in the Philippine languages". *Studies in Language* 5: 227–268.
- Myhill, John. 1986. "The two VS constructions in Rumanian." *Linguistics* 24: 331–350.
- Nebel, Arturo. 1948. *Dinka grammar (Rek-Malual dialect) with texts and vocabulary*. Verona: Missioni Africane.
- Nichols, Johanna. 1986. "Head marking and dependent marking grammar". *Language* 62: 56–119.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Nichols, John. 1988. *An Ojibwe text anthology*. London, Ontario: The Centre for Research and Teaching of Canadian Native Languages, University of Western Ontario.

- Ocampo, Francisco. 1989. *The Pragmatics of word order in spoken Rio-platense Spanish*. University of Southern California dissertation.
- Okrand, Marc. 1977. *Mutsun grammar*. University of California at Berkeley dissertation.
- Ottaviano, Ida de. 1980. *Textos Tacana*. Riberalta, Bolivia: Instituto Lingüístico de Verano.
- Payne, Doris. 1987. "Information structuring in Papago narrative discourse". *Language* 63: 783–804.
- Payne, Doris. 1990. *The Pragmatics of word order: typological dimensions of verb initial languages*. Berlin: Mouton de Gruyter.
- Pet, Willem A. P. 1987. *Lokono Dian, the Arawak language of Suriname: a sketch of its grammatical structure and lexicon*. Cornell University dissertation.
- Prost, Gilbert R. 1967. "Chacobo". In: Matteson, Esther (ed.), *Bolivian Indian Grammars 1*, 285–359. Norman, Oklahoma: Summer Institute of Linguistics.
- Proulx, Paul. 1991. Review of *Statement made by Indians: a bilingual petition of the Chippewas of Lake Superior, 1864*, John Nichols, (ed.) (1988). *Canadian Journal of Linguistics* 36: 198–201.
- Reichard, Gladys. 1924. *Wiyot grammar and texts*. Berkeley: University of California Press [University of California Publications in American Archaeology and Ethnology, Vol. 22, No. 1].
- Russell, Robert. 1975. *Una gramática transformacional del Amahuaca*. Yarinacocha, Peru: Summer Institute of Linguistics [Estudios Panos IV].
- Scancarelli, Janine. 1987. *Grammatical relations and verb agreement in Cherokee*. University of California at Los Angeles dissertation.
- Schachter, Paul. 1976. "The subject in Philippine languages: topic, actor, actor-topic, or none of the above". In: Li, Charles (ed.), *Subject and topic*, 491–518. New York: Academic Press.
- Schachter, Paul. 1977. "Reference-related and role-related properties of subjects". In: Cole, Peter and Sadock, Jerrold M. (eds.), *Syntax and Semantics 8: Grammatical Relations*, 279–306. New York: Academic Press.
- Siewierska, Anna. 1993. "Syntactic weight vs. information structure and word order variation in Polish". *Journal of Linguistics* 29: 233–265.

- Silver, Shirley, and Clara Wicks. 1977. "Coyote steals the fire (Shasta)". *International Journal of American Linguistics, Native American Texts Series 2*: 2: 121–131.
- Sischo, William R. 1977. "Michoacán Nahuatl". In: Langacker, Ronald W. (ed.), *Modern Aztec grammatical sketches, Studies in Uto-Aztecan Grammar 2*, 307–380. Summer Institute of Linguistics and The University of Texas at Arlington.
- Stairs, Emily F., and Elena E. de Hollenbach. 1981. "Gramatica Huave". In: Kreger, Glenn A.S., and Stairs, Emily F. (eds.), *Diccionario Huave de San Mateo del Mar*, 283–391. México, D.F.: Instituto Lingüístico de Verano.
- Tomlin, Russell. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.
- Tomlin, Russell, and Richard Rhodes. 1979. "An introduction to information distribution in Ojibwa". In: *Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society* 307–320. [Reprinted 1992 in Payne, Doris (ed.), *The pragmatics of word order flexibility*, Amsterdam: John Benjamins.]
- Tucker, A. N., and M. A. Bryan. 1966. *Linguistic analysis: the non-Bantu languages of north-eastern Africa*. London: Oxford University Press.
- Van den Berg, René. 1989. *A grammar of the Muna language*. Dordrecht: Foris.
- Vennemann, Theo. 1974. "Analogy in generative grammar: the origin of word order". In: *Proceedings of the Eleventh International Congress of Linguists (1972)*, 79–83. Bologna.
- Vennemann, Theo. 1976. "Categorial grammar and the order of meaningful elements". In: Juillard, A. (ed.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*, 615–634. Saratoga, California: Anma Libri.
- Waterhouse, Viola. 1962. "The grammatical structure of Oaxaca Chontal." *International Journal of American Linguistics* 28: 2, Part II.
- Watters, James K. 1988. *Topics in Tepehua grammar*. University of California at Berkeley dissertation.
- Weber, David J. 1989. *A grammar of Huallaga (Huánaco) Quechua*. Berkeley: University of California Press.
- Weber, Heather. 1994. *A syntactic and morphological sketch of Logbara*. State University of New York at Buffalo M.A. Thesis.
- Whistler, Kenneth W. 1977. "Deer and bear children (Patwin)". *International Journal of American Linguistics, Native American Text Series 2*: 2: 158–178.