To appear in Sergio Scalise, Elisabetta Magni, Antonietta Bisetto (eds.),
Universals of Language Today, Berlin: Springer.

The Branching Direction Theory of Word Order Correlations Revisited[1]

Matthew S. Dryer
University at Buffalo

In Dryer (1992a), I presented evidence from a sample of 625 languages on which pairs of elements correlate with the order of object and verb, and which do not (Greenberg 1963). I argued that the word order correlations reflect a tendency for languages to be consistently left-branching or consistently right-branching, what I referred to as the Branching Direction Theory (the BDT), and proposed that this tendency reflects processing difficulties associated with mixing left- and right-branching. The predictions of the BDT depend heavily, however, on one's assumptions about constituent structure. A number of the correlations require assuming fairly hierarchical constituent structures, and are not predicted by the BDT if one assumes flatter constituent structures. In this paper, I discuss a number of these correlations, arguing that some of them can be explained by a combination of processing considerations and other principles, while a few remain unexplained under assumptions of flat constituent structures.

1. What correlates with the order of object and verb?

While my current database is considerably larger than it was at the time I wrote Dryer (1992a), now containing at least partial data for over 1500 languages, the additional languages do not change the evidence presented in Dryer (1992) as to which pairs of elements correlate with the order of object and verb. In this section, I summarize the conclusions regarding this.

In order to discuss the correlations, it is useful to have a way of referring to the various pairs of elements that correlate with the order of object and verb and the members of each pair. To do this, I will say that if a pair of elements X and Y is such that X tends to precede Y significantly more often in VO languages than in OV languages, then <X, Y> is a *correlation pair*, and X is a *verb patterner* and Y an *object patterner* with respect to this pair. In Dryer (1992a) I provided evidence that each of the pairs of elements in Table 1 is a correlation pair.

[1] Part of the research for this paper was made possible by Social Sciences and Humanities Research Council of Canada Grants 410-810949, 410-830354, and 410-850540, by National Science Foundation Research Grant BNS-9011190, and by support from the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. I am indebted to Lea Brown, Jack Hawkins, Bill Croft, an anonymous reviewer, and members of the audience at the conference in Bologna for comments on earlier drafts of this paper.

VERB PATTERNER	OBJECT PATTERNER	EXAMPLE
verb	object	*ate + the sandwich*
adposition	NP	*on + the table*
noun	relative clause	*movies + that we saw*
article	N'	*the + tall man*
copula verb	predicate	*is + a teacher*
'want'	VP	*wants + to see Mary*
tense-aspect auxiliary verb	VP	*has + eaten dinner*
negative auxiliary	VP	
complementizer	S	*that + John is sick*
question particle	S	
adverbial subordinator	S	*because + Bob has left*
plural word	N'	
noun	genitive	*father + of John*
adjective	standard of comparison	*taller + than Bob*
verb	PP	*slept + on the floor*
verb	manner adverb	*ran + slowly*

Table 1: Correlation Pairs

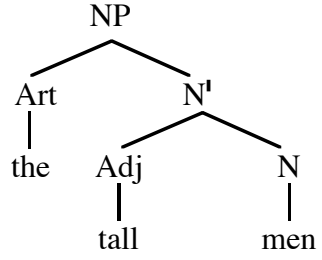
2. Explaining the word order correlations

In Dryer (1992a), I argued in detail that the word order correlations cannot be handled in terms of consistent ordering of heads and dependents. Such a theory would predict that adjectives, demonstratives, and numerals ought to be object patterners, but they are not; the order of these three elements with respect to the noun does not correlate with the order of object and verb (Dryer 2005a). Nor can the correlations be handled in terms of consistent ordering of heads and complements. Such a theory would fail to predict that the order of relative clause and noun and the order of adpositional phrase and verb do correlate with the order of object and verb.

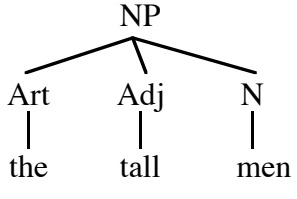
I proposed instead what I called the Branching Direction Theory (BDT), according to which verb patterners are nonphrasal while object patterners are phrasal, with the effect that languages tend towards being either consistently left-branching or consistently right-branching. In addition, I proposed that structures with consistent left-branching or consistent right-branching are easier to process than structures that involve a mixture of left- and right-branching.

A practical problem with the BDT is that it depends on one's assumptions regarding constituent structure. For example, in order to account for the fact that articles are verb patterners, the BDT requires that we assume a structure like that in (1a), where the article combines with a phrasal category like an N'. If, on the other hand, we assume a flatter constituent structure like that in (1b), then the BDT would fail to predict that articles are verb patterners, since they would not be combining with a phrasal category.

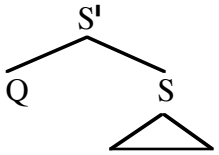
(1) a.

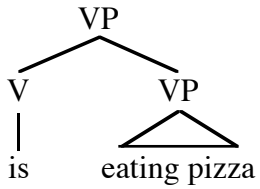


b.



Similarly, unless we assume that polar question particles combine with clauses, as in (2a), rather than simply being one constituent of the clause, the BDT fails to predict that polar question particles are verb patterners; and unless we assume that auxiliary verbs combine with VPs, as in (2b), the BDT fails to predict that auxiliary verbs are verb patterners.

(2) a. 

b. 

3. Flat constituent structure

Because the BDT depends on assumptions about constituent structure that not everyone believes, the BDT will be unconvincing to anyone whose assumptions about constituent structure are different from those required for the BDT to make the correct predictions, and more specifically, will be unconvincing to anyone who assumes that constituent structure is flatter than what is required for the BDT to work. Now perhaps this would not worry me particularly if the constituent structures required by the BDT were ones that I myself believed in. The problem, however, is that over the past fifteen years, my own views about constituent structure have changed, so that I now am one of those who believe that constituent structure is flatter than has often been assumed in generative grammar over the past thirty-five years. What this means is that given my own assumptions about constituent structure, the BDT fails to account for a number of the word order correlations.

The discussion here of why one might adopt flat constituent structures will not do the topic justice, but for reasons of space, my comments are necessarily brief. The idea that constituent structure is flatter than is often assumed in generative grammar is a view that is explicit or implicit in a number of approaches. For example, the idea that articles do not combine with N's is implicit in much work from the early days of generative grammar, where the trees assumed by generative linguists often involved a flatter structure along the lines of (1b). More recently, Culicover and Jackendoff (2005) argue for flatter constituent structures. Furthermore, flatter constituent structures are implicit in most work within Basic Linguistic Theory (Dixon 1997), the theoretical framework assumed in most descriptive grammars written within the last twenty years and implicit in much work in linguistic typology. Such descriptive grammars all assume that there are noun phrases and clauses, which means that they are implicitly assuming that these at least are constituents, but it is relatively rare that such descriptions will assume anything analogous to an N', and it is if anything even rarer for someone to describe a polar question particle as combining with a clause, rather than simply being one element in the clause. In discussing the structure of noun phrases, most such grammars treat articles simply as one type of modifier of the noun. Similarly, most such grammars treat question particles as simply one constituent within the main clause.

The fact that clauses and noun phrases seem to be universally recognized as constituents and have been since the time of traditional grammar reflects a property of the notion of constituent that often seems to be forgotten by linguists of very different persuasions. And that is that constituent structure is largely semantic. The reason that people have always recognized clauses and noun phrases as constituents is that they are clearly semantic units, the clause corresponding to a proposition, or a situation, or an event, or a state of affairs, the noun phrase corresponding to things. Students in introductory linguistics classes are generally better at identifying clauses and noun phrases as a constituents, not because they have natural talents as linguists, but because, in identifying constituents, they are primarily tapping in on their

knowledge of the meaning of the sentence and their awareness of which words go together to form semantic units.

The strongest claim one might make, in fact, is that constituent structure is entirely semantic, that constituent structure trees simply represent one aspect of the meaning of the sentence, of which words go together to form semantic units. I think that in its strongest form, such a claim will not work. For example, the two sentences in (3) have the same (surface) constituent structure, but arguably different semantic structures.

(3) a. John is likely to win the election.
b. John is eager to win the election.

There is a clear sense in which the syntactic structure of these two sentences is closer to the semantic structure of (3b), and (3a) in some sense has a syntactic constituent structure that is somewhat different from its semantic structure. Similarly, the two structures in (4) are two ways of expressing the same meaning across different languages, but have different constituent structures.

(4) a. John caused [Bill to fall].
b. John [caused-fall] Bill.

The different syntactic structures here do not involve a difference in what is a semantic unit.

While examples like those in (3) and (4) argue against a strong claim that syntactic structure is purely semantic, this does not alter the fact that these examples are exceptional. To a large extent constituent structure is semantic, much like the fact that membership in word classes is usually largely predictable from the meaning of words, but not entirely. The extent to which constituent structure is semantic has also been obscured by the sorts of constituent structures that have been popular in generative grammar over the past thirty-five years, first in adopting more hierarchical constituent structures which claim that a lot of sequences of words are constituents that were not viewed as such before, with the added effect that many of the new constituents are less clearly semantic units, and second by assuming constituent structures that are more abstract relative to surface structure. If one assumes flatter constituent structures, then the constituents in such flat constituent structures are more likely to be semantic units.

But the notion of semantic units or semantic constituents is also of specific relevance to the BDT. Underlying that theory is the idea that the word order correlations reflect parsing or processing difficulties associated with certain sorts of syntactic structures. But ultimately, the reason that people parse sentences is to understand the meaning of the sentence. There is often a tendency to view parsing as a process of assigning syntactic structures to sentences, but the final result is the hearer's assigning a meaning to the sentence. Parsing is ultimately a matter of determining the semantic units or semantic constituents of a sentence. In fact, one possible view of parsing is that that is all it is: a process of determining what the semantic units of a sentence are, of determining which words go together semantically. Thus whether or not syntactic constituent structure is flat or hierarchical may not really matter, if what really matters is the semantic units. In other words, one might claim that syntactic structure is not flat, but that when people parse sentences, they only try to identify syntactic constituents that are also semantic units. And while I will formulate the rest of this paper in terms of flat syntactic structures, I could equally well have formulated it in terms of flat semantic structures, while remaining neutral on the question of whether syntactic structures are also flat.

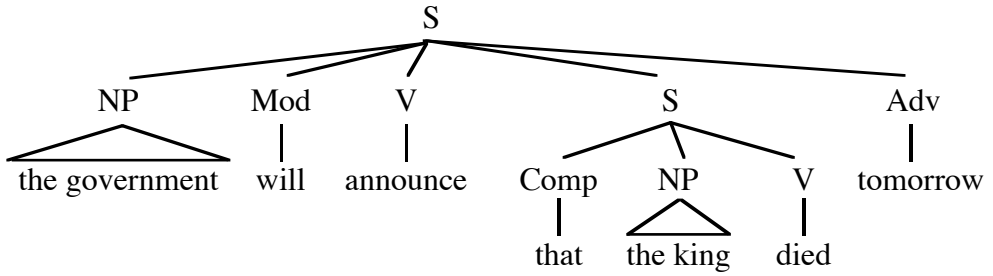
To summarize what I have said so far, parsing sentences correctly means assigning sentences the right meaning. Structures that are difficult to parse are ones that present difficulty for hearers to assign the correct meaning to. In other words, structures that are

difficult to parse are ones for which speakers have difficulty identifying the semantic constituents. Let me illustrate the point with a few examples. Consider the sentence in (5).

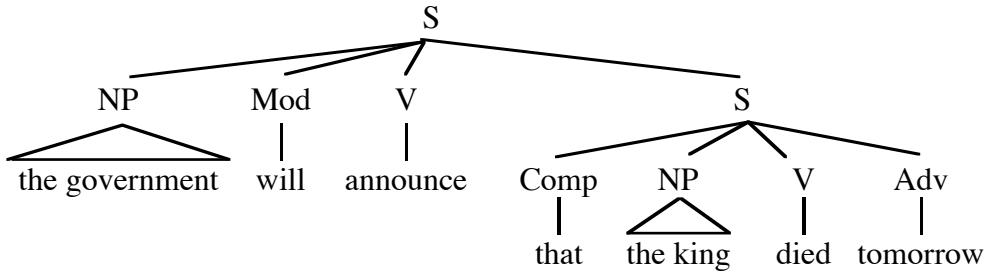
(5) The government will announce that the king died tomorrow.

From a purely syntactic point of view, this sentence has two possible structures, given in (6) and (7). Note that the trees I give are flatter than what is customary nowadays, though they are not radically different from the trees used by generative linguists in the 1960s, except my trees do not recognize a VP constituent (cf. Van Valin and LaPolla 1997: 217-218). But the flatness of these trees, while reflecting the structures I will assume below, is actually irrelevant to the point I am currently making.

(6)



(7)

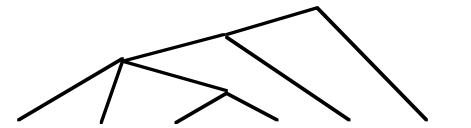


This sentence is interesting in that there is a tension between the automatic strategies of the human parser and what sort of meaning makes sense. The nature of the human parser is such that it tries to assign the sentence the structure in (7). This preference has been expressed in many different ways, one of them being the principle of Late Closure of Frazier (1978). But the meaning associated with (7) doesn't make sense: the past tense of the verb *died* is incompatible with the meaning of the adverb *tomorrow*. The human parser is such that people may never realize that the sentence has another meaning, that corresponding to (6).

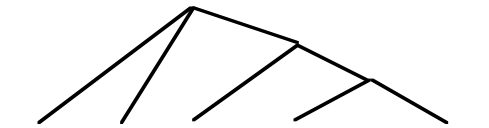
But the main point I want to make about this example is that if someone assigns the sentence the wrong constituent structure, that means that they have assigned it the wrong semantic structure: the syntactic difference between (6) and (7) is equally well a semantic difference, and the difference in the syntactic structures directly represents the difference in meaning: in (7), *the king died tomorrow* is a semantic unit, while in (6) it is not. In other words, the processing difficulty associated with (6) can be described as a difficulty assigning it the right syntactic structure, but that difficulty in assigning it the right syntactic structure is equivalent to the difficulty assigning the sentence the right meaning. If one doesn't assign the sentence the constituent structure of (6), one doesn't understand the meaning of the sentence.

Or consider the pair of constituent structures in (8). Both are possible syntactic structures, but again one of them, namely (8b), is semantically anomalous. If on hearing a sentence containing this phrase, someone assigns it the syntactic structure in (8b), or, equivalently, assigns it the semantic structure in (8b), that means that they have not understood the sentence.

(8) a.

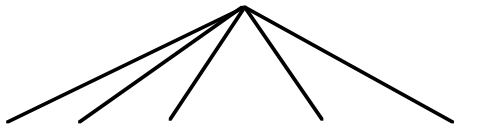


b.

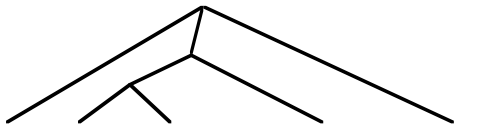


And if one were to assign the syntactic or semantic structure in (9a) to the phrase *a much more interesting idea*, that would mean that one had not understood the phrase, since understanding the phrase requires that one recognize that it has a semantic structure like that in (9b).

(9) a.

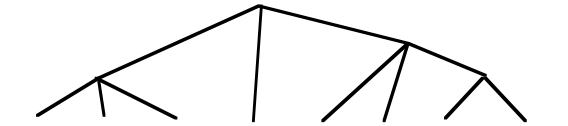


b.



And interpreting the sentence in (10) requires that one realize that *the old man* belongs to the clause with the verb *says*, but that *he* and *an angel* belong to the clause with the verb *saw*. If one doesn't realize this, one has not understood the sentence.

(10)



The general point is that when I talk about processing difficulties associated with certain constituent structures, I am talking about difficulties in communication, difficulties in assigning the correct meaning to a sentence.

This brings us back to the issue of flat constituent structure and the BDT. Structures that are difficult to parse are ones where hearers have difficulty recognizing which words go together semantically. But that means that the constituents that are crucial are those that are clearly semantic units, like clauses and noun phrases. Hence for the BDT to be a convincing explanation for the word order correlations, it must work for flat structures, structures that represent the semantic units. If the BDT depends on structures that are irrelevant to meaning, then it fails as an adequate account.

So which of the correlation pairs in Table 1 above involve pairs of elements where the object patterner is clearly a semantic unit? Admittedly, there is room for considerable disagreement as to what constitutes a semantic unit, so it may not be clear in all cases whether the object patterner is a semantic unit. However, the pairs in Table 2 are pairs for which I believe a good case can be made that the object patterners are semantic units, either because they involve a clause (including VPs without a subject, which can be analysed as clauses without a subject) or a noun phrase (possibly including an adposition).

VERB PATTERNER	OBJECT PATTERNER	EXAMPLE
verb	object	*ate + the sandwich*
copula verb	predicate	*is + a teacher*
'want'	VP	*wants + to see Mary*
noun	genitive	*father + of John*
noun	relative clause	*movies + that we saw*
adjective	standard of comparison	*taller + than Bob*
verb	PP	*slept + on the floor*

Table 2: Correlation pairs where the object patterner is a semantic unit

On the other hand, the pairs in Table 3 are ones for which it is less clear that the object patterners are semantic units.

VERB PATTERNER	OBJECT PATTERNER	EXAMPLE
article	N'	*the + tall man*
plural word	N'	
adposition	NP	*on + the table*
tense-aspect auxiliary verb	VP	*has + eaten dinner*
negative auxiliary	VP	
complementizer	S	*that + John is sick*
adverbial subordinator	S	*because + Bob has left*
question particle	S	

Table 3: Correlation pairs where is it less clear that the object patterner involves a semantic unit

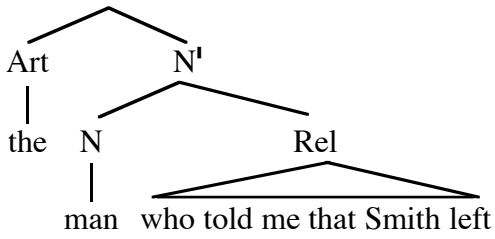
What I have given as object patterners in Table 3 are as I listed them in Table 1 above (and in Dryer 1992a). But under the assumption that these object patterners are not semantic units, i.e., that the verb patterners are sisters to the constituents of the object patterners, these correlations would need to be reformulated. In fact, if these verb patterners occur as the initial (or final) constituent amongst a number of sister constituents, then the very notions of object patterner and correlation pair become problematic. At most, we can say that the verb patterners tend to occur first within their mother constituent more often in VO languages than in OV languages. I will therefore formulate the discussion below in terms of the position of the verb patterners in Table 3.

The question therefore is whether we can find alternative explanations for why the verb patterners in Table 3 are verb patterners. For reasons of space, I will only be able to discuss three of these verb patterners, namely articles, complementizers, and auxiliary verbs.

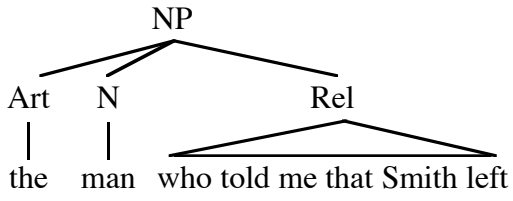
4. The position of articles

The first verb patterner I will discuss is articles. As mentioned above, for the BDT to work, one must assume a structure in which the article combines with an N', as in (11a) (or an analysis in which the article is a determiner and combines with an NP to form a DP); if, on the other hand, we assume a flatter structure, as in (11b), the article would not be combining with a semantic unit.

(11) a.



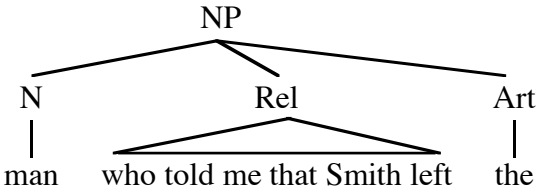
b.



In the particular example in (11b), one of the two constituents that the article is combining with is a relative clause, which is a phrasal category. However, it is not in general the case that articles combine with phrasal categories, since they may combine with just a noun, or just a noun plus one or more nonphrasal modifiers of the noun, like numerals or adjectives. In other words, if we assume a structure like that in (11b), articles no more combine with phrasal categories than adjectives or numerals do, but the latter modifiers are not verb patterners. In short, the BDT fails to account for the fact that articles are verb patterners while adjectives and numerals are not, if we assume flat structures like that in (11b).

So let us search for an alternative explanation for why articles might be verb patterners. In other words, what we need to do is explain why VO languages are more likely than OV languages to employ the word order in (11b) than the word order in (12), where the article occurs at the end of the noun phrase.

(12)



It turns out that there is a proposal in the literature that appears to explain this. Hawkins (1994, 2004) proposes a principle like that in (13), which I will refer to as Hawkins' Principle. Hawkins' theory is more complex than I will portray it in this paper, and my discussion will simplify it in some ways, though I believe that this simplification does not affect the argument.

(13) Hawkins' Principle

Structures with shorter constituent recognition domains (CRDs) are easier to process.
(Hawkins 1994, 2004)

The notion of *constituent recognition domain* (CRD) is central to Hawkins' theory. Again simplifying things somewhat, the CRD of a constituent is the stretch of words in that constituent starting with the first word that allows the parser to construct the first daughter of that constituent and ending with the first word that allows the parser to construct the last daughter of that constituent, where a word allows the parser to construct a phrase if the parser can infer the category of the daughter phrase from that word. Hawkins calls words that allow the parser to construct phrases *mother-node constructing categories* (MNCCs). The clearest instances of MNCCs are heads. For example, if the parser encounters a noun, it can normally infer that the noun is part of a noun phrase, and the noun therefore allows the parser to construct a noun phrase. In addition to heads, various other words also serve as MNCCs. For example, articles are MNCCs for noun phrases; i.e. if the parser encounters an article, it can thereby infer that there is a noun phrase. If the first daughter of a constituent consists of one

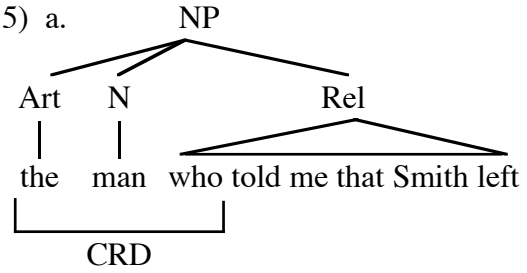
word, the CRD will start with that word. Similarly, if the last daughter of a constituent consists of one word, the CRD will end with that word.

Hawkins' theory is similar in spirit to the BDT in that it predicts that certain sorts of structures are less common in language because they are more difficult to process. Strictly speaking, the BDT is just an hypothesis about what provides a general categorization of verb patterners and object patterners; the proposal that the tendency to avoid mixing left and right branching is motivated by processing difficulties associated with such structures is actually a separate hypothesis. Hawkins' Principle furthermore makes similar predictions to the BDT. It predicts, for example, that structures that involve a mixture of left and right branching will be less common in language because they in general have longer CRDs (see, for example, Hawkins 1994: 96). On the other hand, it is a broader theory in a number of ways, for example in that it predicts lower frequency of structures that are independent of the word order correlations. It predicts, for example, that (14a) is preferred over (14b) because the CRD for the VP (or the clause if one assumes a flat structure) is shorter in (14a).

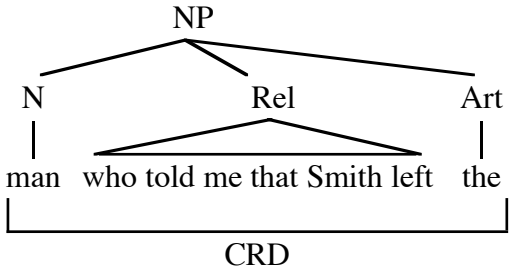
(14) a. He brought over the books that I asked for.
b. He brought the books that I asked for over.

But most important for this paper is the fact that Hawkins' Principle appears to provide at least a partial explanation for some of the correlations that the BDT fails to account for under assumptions of flat structures. For example, consider the structure in (11b) and (12) above, repeated here as (15a) and (15b).

(15) a.



b.



Consider first (15a). Since the article consists of a single word, the CRD of the noun phrase begins with the word *the*. If we assume that the relative pronoun *who* serves as a MNCC for the relative clause, then the relative pronoun will be the MNCC for the last daughter of the noun phrase, so that the CRD will end with the relative pronoun. Hence, the CRD for (15a) will be *the man who*. On the other hand, the CRD for (15b) will consist of the entire noun phrase *the man who told me that Smith left*, since the first and last daughters consist of single words. Hence Hawkins' principle correctly predicts a preference for the word order in (15a) over that in (15b).

There is one problem, however, with this explanation in terms of Hawkins' Principle. Namely, it implies a preference for Art-N-Rel order over N-Rel-Art order regardless of whether the language is VO or OV, since the argument makes no reference to the order of object and verb. But what we are trying to explain is why ArtN order is more common in VO languages while NArt order is more common in OV languages. The point is a significant one since the two orders of relative clause and noun are about equally common among OV languages (Dryer 1992a, 2005b). The appeal to Hawkins' Principle would seem to predict that that OV languages which are NRel should show the same preference for ArtN order as that found in VO languages. However, it turns out that this is not the case. To the contrary, OV&NRel languages exhibit the same (weak) preference for NArt order as that found in

OV&RelN languages. An example from an OV&NRel language that employs the N-Rel-Art word order of (12) is shown in (16).

(16) Lepcha (Tibeto-Burman; Mainwaring 1876: 43)

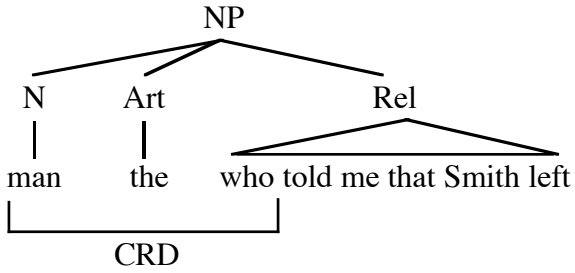
máro	[to	nun	zuk]	re
person	who	ERG	do	DEF

'the person who did it'

Hawkins' Principle does correctly predict NArt order in a language that is OV&RelN, since placing the article at the beginning of the noun phrase would mean that the CRD for the noun phrase would start at the beginning of the noun phrase, so that the CRD would be the entire noun phrase, analogous to (12) (but with the mirror image word order), while placing the article at the end of the noun phrase would mean that the CRD for the noun phrase would not start until the noun was encountered, after the relative clause.

One possible explanation for the high occurrence of N-Rel-Art order in OV languages is that the article helps processing in that it signals the end of the relative clause. This is more useful in OV languages than in VO languages because an object noun phrase containing a relative clause will end inside the main clause in an OV language but at the end of the main clause in a VO language. Signaling the end of a subordinate clause within a main clause is more useful for processing than signaling the end of a subordinate clause that is also the end of a sentence.[2] The above discussion considers two possible positions for an article in an NRel language, one at the beginning of the noun phrase, with Art-N-Rel order, the other at the end of the noun phrase, with N-Rel-Art order. But if we assume flat structures, then there is a third logical possibility for the placement of the article; namely it might occur immediately after the noun, as in (17).

(17)



This word order is found in Koyraboro Senni, as in (18), where there are actually two morphemes that could be treated as definite articles, the definite clitic *=oo* and the anaphoric demonstrative *din*.

(18)

hondu	beer=oo	din	[kaŋ	n=ga	diy-aa]
dune	big=DEF.SG	ANAPH	REL	2SG.SUBJ=IMPF	see-3SG.OBJ

'the great dune that you see' (Heath 1999: 244)

But this order is less common than NRelArt order. In (19) are listed languages of the two types NRelArt and NArtRel, first for VO languages and then for OV languages. This shows NRelArt outnumbering NArtRel by 13 to 6 among VO languages and by 11 to 3 among OV languages.

[2] This idea was suggested to me by Lea Brown.

(19) VO&NRelArt: 13 languages: Busa, Mupun, Maybrat, Kera, Gimira, Sobei, Sudest, Bali-Vitu, Sisiqa, Longgu, Fongbe, Bagirmi, Sundanese

VO&NArtRel: 6 languages: Koyra Chiini, Tetun, Kaulong, Linda, Nadrogā, Paamese

OV&NRelArt: 11 languages: Tshangla, Kanuri, Tubu, Maba, Seri, Busa, Arrernte (Mparntwe), Runga (Maban), Lepcha, Takia, Dogon

OV&NArtRel: 3 languages: Koyraboro Senni, Kairiru, Ute

We must look for some factor other than Hawkins' Principle to predict that the NArtRel word order in (17) tends to be avoided, since the length of the CRD is the same as ArtNRel in (15a) and shorter than the length of the CRD with NRelArt order in (15b).

The relative infrequency of structures like that in (17) can be seen as reflecting a generalization that is independent of word order correlations. Namely, crosslinguistically, articles tend to occur on the periphery of noun phrases, either as the first word, or as the last word. Hence, if we can explain why this is the case, we would have an explanation for why VO languages prefer the word order in (15a) over those in both (15b) and (17), and hence why ArtN order is preferred to NArt order in VO languages.

Now some linguists have an easy answer to why articles tend to occur on the periphery of noun phrases. Namely, those linguists who believe in more hierarchical structures than those I have been assuming claim that articles combine with N'. If articles combine with N's (or combine with NPs to form DPs), then structures like (17) are impossible, or at the very least, they involve a discontinuous N', and would tend to be avoided for that reason. However, since the goal of this paper is to discuss how to explain the word order correlations if we assume flat structures, that line of explanation is not open to me.

I would suggest as an alternative that the position of articles on the periphery of noun phrases is simply one manifestation of the more general principles governing the order of noun modifiers, that leads to adjectives tending to be closer to nouns than numerals, demonstratives, and articles, explaining why (20b) and (20c) are ungrammatical in English and to certain preferred orders among different sorts of adjectives, explaining why (20a) is preferred over (20d) .

(20) a. the three large black dogs
b. *three the large black dogs
c. *three large black the dogs
d. ??the three black large dogs

While I know of no completely satisfactory account of these preferences (but see Bache 1978, Posner 1986, Sproat and Shih 1988), they seem to involve some principle according to which words that denote more inherent properties of the referent of the noun phrase tend to occur closer to the noun. Thus in (20a), the adjectives denote properties that the dogs have had for a long time, while the numeral *three* denotes a property that has been true only as long as the dogs are together, while the pragmatics of the definite article do not denote a property of the dogs at all, but simply their status in the discourse. I suspect it is some principle like this that explains why articles tend to occur on the periphery of noun phrases.

I should note that even if one accepts the idea that some principle like this governs the order of noun modifiers, one might claim that this principle simply reflects a deeper fact that the semantic structure of noun phrases with multiple modifiers is such that in a noun phrase like (20a) with an article, a numeral, and an adjective, the adjective plus noun forms a semantic unit to the exclusion of the article and numeral and that the numeral plus adjective plus noun forms a semantic unit to the exclusion of the article. According to this approach, the noun

phrase *the three large dogs* has the semantic structure *[the [three [large dogs]]]*. The fact that languages avoid the word order in *the large three dogs* would be explained on this approach by saying that *three dogs* does not form a semantic unit to the exclusion of *large*. A proponent of this idea could argue that this explains why modifiers denoting inherent properties tend to occur closer to the noun, rather than the other way round.

While this approach may have some merit, it denies the premise of this paper, that the semantic structure of noun phrases is flat. In other words, I am trying to explain the word order correlations if we assume flat structures. If this premise is false, if articles do in fact combine with semantic units, then the BDT will account for the fact that articles are verb patterners. In other words, the premise may be false, but then the fact that the order of article and noun correlates with the order of verb and object would be correctly predicted by the BDT.

But there is another reason why it is not crucial whether the semantic structure of noun phrases is more hierarchical than I am assuming in this paper. Namely, what I am considering a semantic unit is a set of words which the hearer must recognize as a semantic unit in order to understand the noun phrase. But even if noun phrases are viewed as having an hierarchical semantic structure, it is not clear that the hearer must recognize the various units in this hierarchical structure as semantic units in order to understand the noun phrase. In other words, even if a noun phrase is viewed as having a nested hierarchical semantic structure, the semantic units in this hierarchical structure are not fundamental semantic units in the way that noun phrases and clauses are. The claim of this paper is that it is only these fundamental semantic units that must be recognized as semantic units if a sentence is to be understood.

Before leaving discussion of articles, there is one additional correlation that is worth discussing. Namely, not only do VO and OV languages differ in the order of article and noun, but they also differ in how frequently they employ articles. The data in (21) gives the number of languages in my database that have articles and the number that do not have articles for both OV and VO languages.

(21)

	OV	VO
Has articles	183	312
Does not have articles	101	57

What (21) shows is that languages with articles outnumber those without by less than 2 to 1 among OV languages, but by over 5 to 1 among VO languages, so that VO languages are more likely to have articles than OV languages.[3]

Interestingly, Hawkins' theory provides a possible explanation for this difference in the frequency of articles. Namely, as discussed above, both nouns and articles serve as MNCCs for noun phrases. Now an important difference between parsing right-branching languages and parsing left-branching languages that is predicted by Hawkins' theory is that in parsing right-

[3] There is one feature of the numbers in (21) that is highly misleading and purely an artifact of the way in which the data was collected. Namely, the numbers in (21) suggest that the majority of languages of the world have articles. But this is probably not the case. Rather the higher numbers for languages with articles in (21) simply reflects the fact that it is easier to infer from a grammatical description of a language that it has articles than it is to infer that it does not have articles (since the article might be optional). Namely, if I found evidence in a description that a language has articles, then I coded it as having articles. But a description had to be fairly detailed and fairly clearly written for me to conclude that the language lacks articles. If a description was brief, or if it was less clearly written, or if for some reason I had only limited time to examine it, I would not code the language as lacking articles in the absence of finding evidence of articles. It is only for this reason that the numbers for languages with articles in (21) are greater than those without. It is my educated guess that languages with articles amount to at most half the languages of the world, and probably somewhat less than half.

branching languages, the CRD of a constituent will tend to be shorter the more quickly the last daughter of that constituent is recognized. It is thus an advantage in a right-branching language to have more than one MNCC for a constituent, since the more MNCCs there are, the more quickly the constituent will be recognized, and the shorter the CRD of the mother of that constituent will be. Conversely, in a left branching language, Hawkins' theory predicts that the CRD of a constituent will be shorter, the later the first daughter of that constituent is constructed. In other words, while someone processing a right-branching language wants to recognize constituents as quickly as possible, someone processing a left-branching language wants to recognize constituents as late as possible. But having multiple MNCCs is thus a disadvantage in a left-branching language, because having multiple MNCCs means that the constituent with multiple MNCCs will be recognized earlier. But since a language with articles will have two MNCCs for noun phrases while a language without articles will have only one, this means that Hawkins' theory predicts that right-branching languages are more likely to have articles than left-branching languages, and since VO languages are generally right-branching while OV languages are generally left-branching, this predicts that VO languages are more likely to have articles than OV languages.

5. The position of complementizers

The second pair of elements that I will discuss is the order of complementizer and clause. Some languages have clause-initial complementizers, like English, as in (22), while other languages have clause-final complementizers, like Japanese, as in (23).

(22) John knows [***that*** we have left].

(23) OV&FinalComp: Japanese
John wa [nihongo ga muzukasii ***to***] it-ta
TOPIC [Japanese SUBJ difficult ***COMP***] say-PAST
'John says [that Japanese is difficult].'

Since I did not present data on the order of complementizer and clause in Dryer (1992a), I present in (24) data for this, using the format in that paper, where the numbers denote numbers of genera rather than numbers of languages, although I give the total number of languages of each sort in the rightmost column.

(24)	Africa	Eurasia	SEAsia&Oc	Aus-NewGui	NAmer	SAmer	Total	#Lgs
OV&FinalComp	2	[5]	[3]	1	[2]	[1]	14	27
OV&InitComp	[6]	4	1	[3]	0	0	14	22
VO&FinalComp	0	0	0	0	0	0	0	0
VO&InitComp	[23]	[9]	[13]	[4]	[10]	[4]	63	140

What (24) shows is that the two positions of complementizers are about equally common in OV languages, but that all of the VO languages for which I have data on the position of complementizers place the complementizer at the beginning of the clause. The data in (24) shows 63 genera containing languages with clause-initial complementizers; the total number of languages is 140.

Note that although the two positions of complementizers are about equally common in OV languages, it is still the case that the order of complementizer and clause correlates with the order of object and verb and that complementizers are verb patterners since OV languages

use clause-final complementizers significantly more often than VO languages. The asymmetry in (24) requires some further explanation, but it is an instance of a broader generalization, and that is that while structures of the form phrase+word (e.g., NP + Postposition) are as common as structures of the form word+phrase (e.g. Preposition + NP), structures of the form clause+word are much less common crosslinguistically than structures of the form word+clause (Dryer 1992b). Another instance of this generalization is that prenominal relative clauses are much less common crosslinguistically than postnominal relative clauses.

The example in (23) above from Japanese illustrates an OV languages where the complementizer is clause-final; the examples in (25) from Supyire illustrate an example of an OV language in which the complementizer is clause-initial; (25a) illustrates the OV word order, while (25b) illustrates the clause-initial complementizer.

(25) OV&InitComp: Supyire (Gur, Niger-Congo; Mali; Carlson 1994: 339, 423)

a. u ná naŋiyááyi kàni ŋya
3SG REM.PAST wild.animals.DEF only see
'he saw only the wild animals'

b. mìi a lì ŋyɛ̀ [na u a kàrè]
1SG PERF it see [COMP 3SG PERF go]
'I saw [that he had gone]'

But there is another important difference between Japanese and Supyire. Not only do these two languages differ in the position of complementizers, but they also differ in the position of the complement clause: in Japanese, the complement clause precedes the main verb, while in Supyire the complement clause follows the main verb. And this difference in the position of the complement clause among OV languages correlates very strongly with the position of complementizers, in that we have the two generalizations stated in (26).

(26) a. OV languages in which the complement clause precedes the verb normally have clause-final complementizers rather then clause-initial complementizers.

b. OV languages in which the complement clause follows the verb normally have clause-initial complementizers rather then clause-final complementizers.

The specific data I have on this is given in (27).

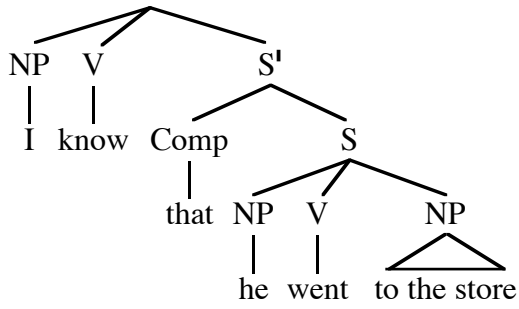
(27) *OV, preverbal complement clause, final complementizer:* 12 languages:
Dogon, Orkhon Turkic, Japanese, Ainu, Kannada, Hayu, Mao Naga, Angami, Lai Chin, Bawm, Amele, Slave
OV, preverbal complement clause, initial complementizer: 1 languages:
Harar Oromo
OV, postverbal complement clause, final complementizer: 1 language:
Khoekhoe (aka Nama)
OV, postverbal complement clause, initial complementizer: 18 languages:
Mauka, Supyire, Tunen, Latin, Hindi, Punjabi, Marathi, Wakhi, Pashto, Persian, Tajik, Wakhi, Turkish, Tsova-Tush, Gapapaiwa, Tawala, Sare, Djapu

Only two languages in (27), Harar Oromo and Khoekhoe, do not conform to (26), while the other 29 languages do conform. I will return to these generalizations shortly.

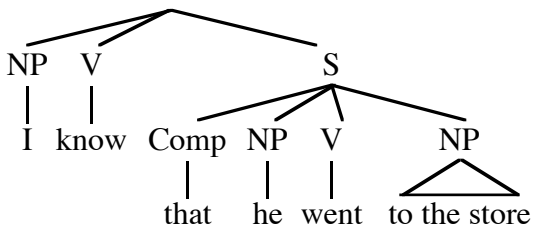
Let us turn to the issue of explaining why complementizers are verb patterners, occurring at the beginning of the clause significantly more often in VO languages than in OV languages. In Dryer (1992a), I explained this in terms of the BDT, assuming a structure like

that in (28a) in which the complementizer combines with an S to form an S'. On the other hand, if we assume a flatter structure like that in (28b) (which was the usual constituent structure assumed in generative grammar before around 1970), then the complementizer is no longer combining with a phrasal constituent.

(28) a.



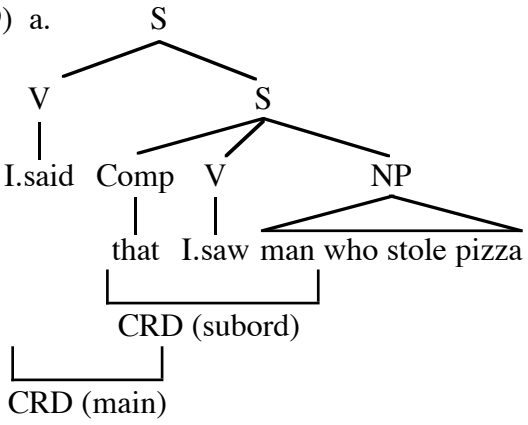
b.



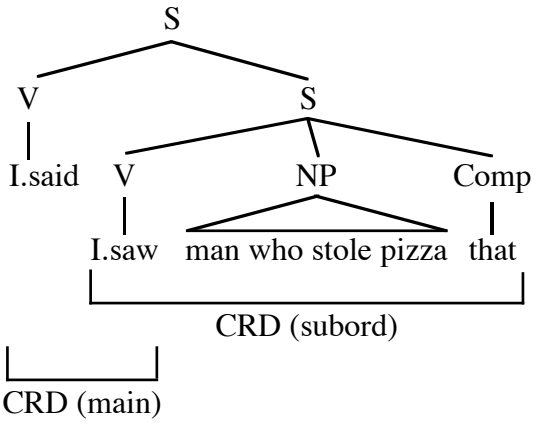
As a result, the BDT fails to predict that a VO language will place the complementizer at the beginning of the clause, rather than at the end.

Hawkins' Principle, however, does account for this. Compare the two structures in (29), where (29a) illustrates a clause-initial complementizer and (29b) a clause-final complementizer. The structures given represent the subject as being expressed in the verbal morphology, rather than by a separate pronoun, since this is the way the majority of languages express pronominal subjects, and this proves useful in comparing a number of different possible structures below, all intended as ways of expressing the meaning 'I said that I saw the man who stole the pizza'. I include the CRDs for both the complement clause and for the main clause, since considering both of these will be relevant below when I consider the possibilities in OV languages.

(29) a.



b.

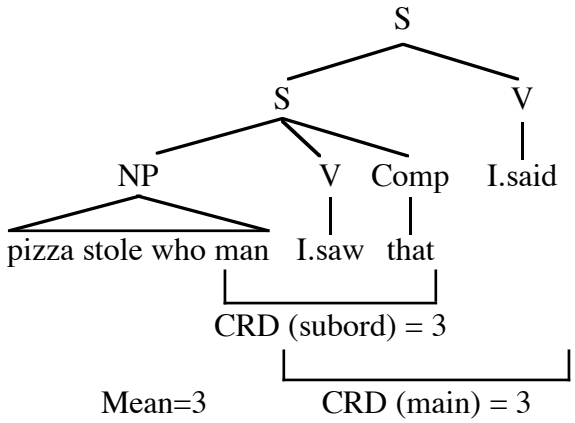


Hawkins' Principle does account for the preference for (29a) over (29b): although the CRDs for the main clause are the same in these two structures (both consisting of only two words), the CRD for the complement clause is shorter in (29a) than it is in (29b). The CRD for the subordinate clause in (29b) is the entire clause, while in (29a) it only includes the first word of the last daughter of the clause.

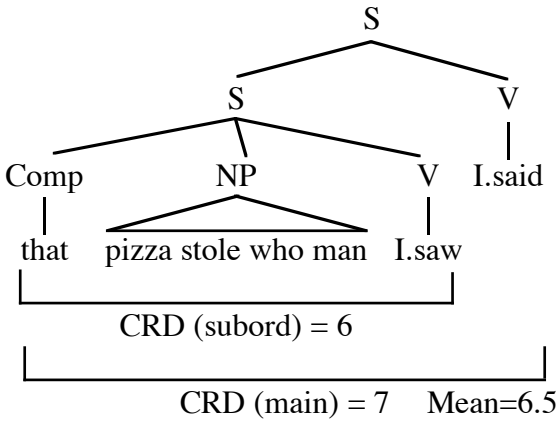
Now consider the situation in an OV language. There are in fact eight types of OV languages to consider, based on three binary variables, position of complementizer in clause, order of complement clause with respect to the verb, and the order of relative clause and noun. For all three of these variables, both orders are well attested among OV languages. These eight possibilities are given in (30) to (33). The CRDs for the subordinate clause and the matrix

clause, as well as the length in words of each CRD and the mean of these two values is given for each tree.

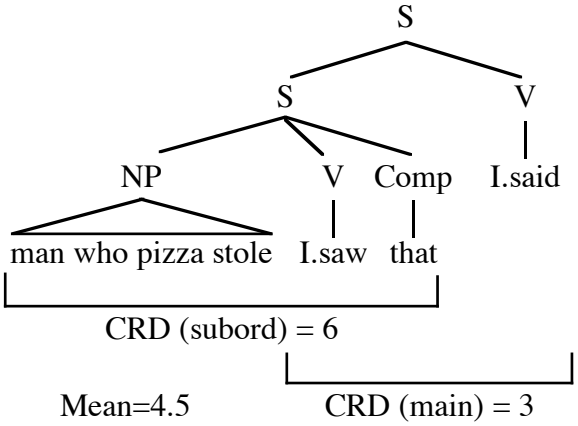
(30) a. CompClause-V, RelN, FinalComp



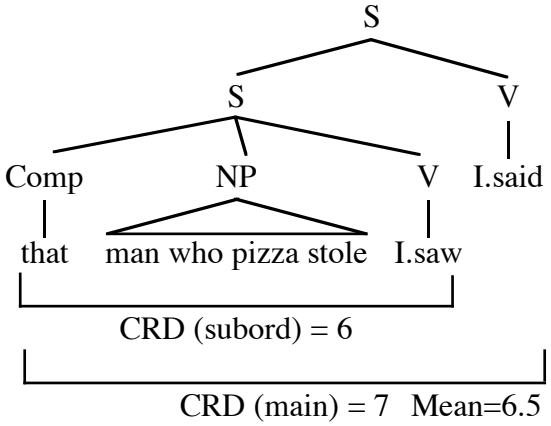
b. CompClause-V, RelN, InitialComp



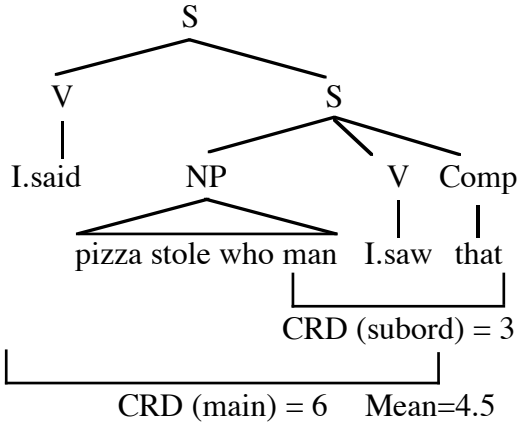
(31) a. CompClause-V, NRel, FinalComp



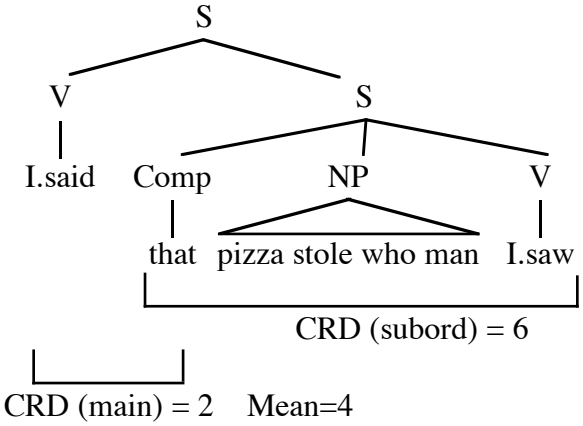
b. CompClause-V, NRel, InitialComp



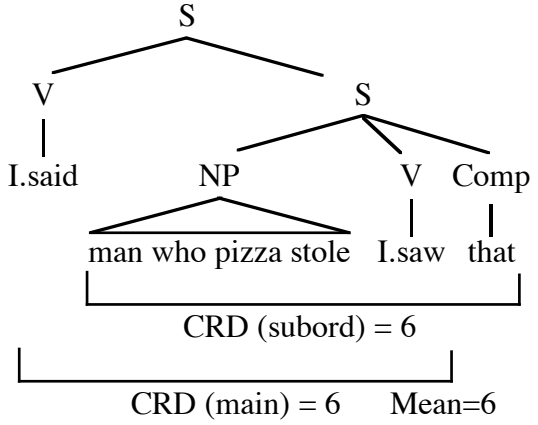
(32) a. V-CompClause, RelN, FinalComp



b. V-CompClause, RelN, InitialComp



(33) a. V-CompClause, NRel, FinalComp



b. CompClause-V, NRel, InitialComp

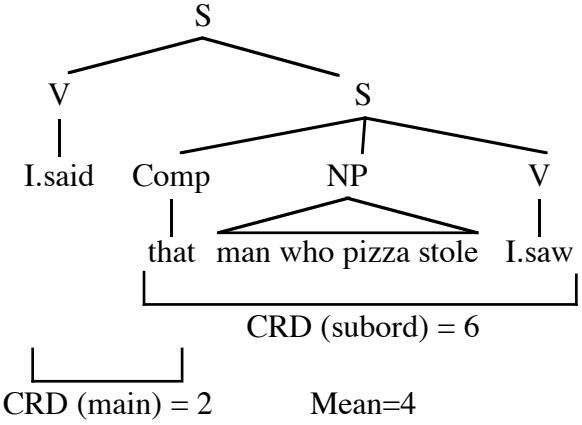


Table 4 summarizes the mean length of the CRDs for the eight structures in (30) to (33), and lists attested languages of each sort.[4]

		FinalComp		Initial Comp	
		Mean length of CRD	Attested languages	Mean length of CRD	Attested languages
ClauseV	RelN	3	8 languages: Japanese, Ainu, Orkhon Turkic, Kannada, Hayu, Mao Naga, Lai Chin, Bawm	6.5	Zero languages
	NRel	4.5	1 language: Amele	6.5	1 language: Harar Oromo
VClause	RelN	4.5	1 language: Khoekhoe (aka Nama)	4	3 languages: Marathi, Turkish, Tsova-Tush, Sare
	NRel	6	Zero languages	4	6 languages: Pashto, Persian, Tajik, Wakhi, Gapapaiwa, Tawala

Table 4: Position of complement clauses, complementizers, and relative clauses in OV languages

Table 4 shows that although the order of relative clause and noun has some effect, the basic generalization is that final complementizers lead to shorter CRDs if the complement clause precedes the verb while initial complementizers lead to shorter CRDs if the complement clause follows the verb. Hence Hawkins' Principle correctly predicts that VO languages should have

[4] Some of the languages listed in (27) above are not included in Table 4 because they employ either internally-headed or correlative relative clauses.

initial complementizers, since complement clauses in VO languages invariably follow the verb (Dryer 1980) and that OV languages will have final complementizers if the complement clause precedes the verb and initial complementizers if the complement clause follows the verb. In addition, the numbers of attested languages of each of the types in Table 4 matches the predictions of Hawkins' Principle: the three types whose mean CRD is 4 or less are exactly the types for which more than one language is attested.

The numbers in (24) above show another difference between OV and VO languages other than the position of complementizers. Namely, (24) includes data for 49 OV languages with complementizers but for 140 VO languages with complementizers. This difference reflects the fact that VO languages employ complementizers more often than OV languages do. This is analogous to the fact discussed in section 4 that VO languages have articles more often than OV languages do. And the explanation for the lower frequency of articles in OV languages in terms of Hawkins' Principle also applies to the lower frequency of complementizers in OV languages. Namely, complementizers, like verbs, are MNCCs for clauses. Having multiple MNCCs for a given category is an advantage in right-branching languages but a disadvantage in left-branching languages. This is reflected in the trees in (30) to (33) in that the shortest CRDs for preverbal complement clauses contained three words, while the shortest CRDs for postverbal complement clauses contain only two words. This is because with preverbal complement clauses, both the complementizer and the verb in the complement clause are part of the CRD for the matrix clause, since as soon as one of them is encountered, the complement clause can be constructed by the parser. With postverbal complement clauses, however, once the complementizer is encountered, the complement clause is constructed, so the verb in the complement clause will not be part of the CRD for the matrix clause.

As with articles, Hawkins' Principle only partially explains why complementizers are verb patterners. If we assume flat structures in which the complementizer is simply one of the constituents of the clause, what we need to explain is why complementizers do not occur inside clauses. If one considers what (29a) would look like with the complementizer immediately before the verb or immediately after the verb, one can see that the length of the CRD would be the same as in (29a). But languages with clause-internal complementizers are rare. As was the case with articles, complementizers normally occur at the periphery of clauses, either at the beginning or at the end. To fully explain why complementizers are verb patterners, we would need to explain this generalization.

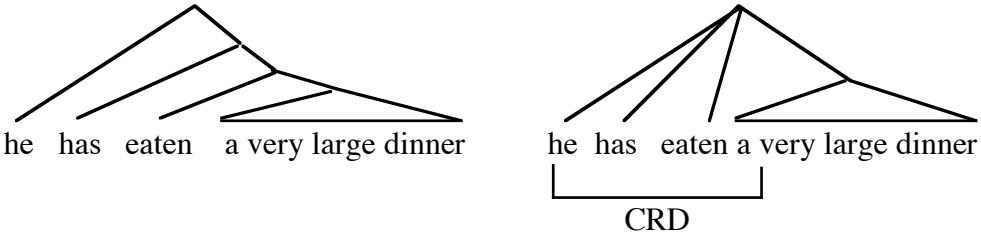
Once again, one obvious way to explain why complementizers occur at clause boundaries would be to say that they combine semantically and/or syntactically with clauses, as in (28a) above, where they combine with an S to form an S'. However, can we explain it without making such a claim? One possibility is that one of the functions of complementizers is to signal clause boundaries. Since identifying which words go together in the same clause is essential for understanding a sentence, having an overt signal of a clause boundary is advantageous for sentence processing. A complementizer inside a clause would clearly not be helpful in signaling a clause boundary.

The hypothesis that complementizers signal clause boundaries also provides an additional explanation for why languages would prefer (29a) over (29b): placing a complementizer at the end of a sentence is not going to be helpful in signaling a clause boundary within a sentence. This provides an additional explanation for why languages with preverbal complement clauses employ clause-final complementizers and why languages with postverbal complement clauses employ clause-initial complementizers.

6. The order of auxiliary verb and main verb

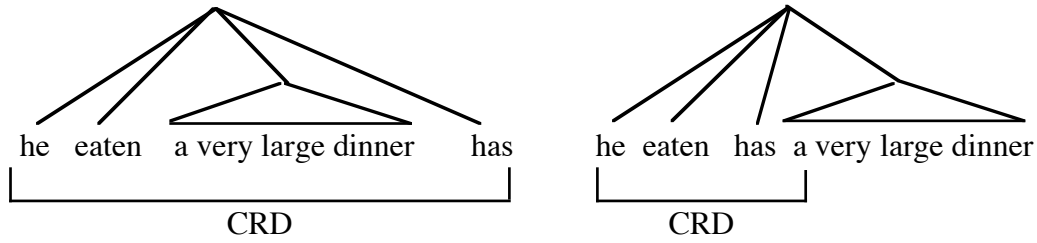
My discussion of this pair of elements will be fairly brief. The primary conclusion is that there is no obvious explanation for why auxiliary verbs tend to precede the main verb in VO languages but follow in OV languages. The BDT explains the correlation only if we assume a structure like that in (34a), where the auxiliary verb is combining with a verb phrase *eaten a very large dinner*. If we assume the flatter structure in (34b), the auxiliary verb is no longer combining with a phrasal category.

(34) a. b.



Hawkins' Principle does correctly predict that VO languages will prefer the structure in (34b) to that in (35a), with the auxiliary verb at the end of the clause, since in (35a), the CRD for the clause is longer, since it includes the entire clause. However, it does not explain why the word order in (35b) is avoided, since the length of the CRD in (35b) is the same as in (34b).

(35) a. b.



Nor do the sorts of factors that cause articles and complementizers to occur at the periphery of their mother constituents seem to be relevant here, since if we assume the flat structures in (34b) and (35b), the auxiliary verb is not in peripheral position in either structure, nor is it the case that auxiliary verbs tend to occur at the periphery of clauses, since in SVO languages they typically occur after the subject and immediately before the main verb.

My only suggestion is that explanations in terms of grammaticalization are especially plausible with auxiliary verbs since the processes of grammaticalization of main verbs to auxiliary verbs is one of the most frequent forms of grammaticalization. Under this approach, their position relative to the main verb reflects their position as main verb prior to the grammaticalization. Note that if this is the correct explanation, then the word order correlations are not a unified phenomenon, some being due to processing factors, some due to grammaticalization, and perhaps some due to other factors, and perhaps some due to a combination of factors. While a single explanation might seem a priori preferable, it is certainly undesirable if it is incorrect.

7. Conclusion

For reasons of space, I have not been able to discuss the other problematic pairs listed in Table 3 above, but let me make some very brief suggestions. If adpositions are simply constituents of noun phrases, then the problem of explaining their correlation is rather similar to the issues surrounding the order of article and noun, discussed in section 4, since they, like

articles, would serve as MNCCs for noun phrases. Grammaticalization also appears to be a factor in explaining the position of adpositions in many languages, so this is a case where two different sorts of factors may be working together. The situation with plural words may also be similar to that with articles, but this is less obvious, since unlike articles and adpositions, they frequently do not occur at the periphery of noun phrases. The explanation for the position of negative auxiliaries is presumably the same as that with tense-aspect auxiliaries. The situation with adverbial subordinators is similar to that with complementizers in many respects, and they plausibly serve as signals of clause boundaries (in addition to signalling a specific semantic relationship between the subordinate clause and the main clause). As for question particles, if we can explain why they tend to occur at the beginning or end of the sentence (perhaps in terms of the fact that they are modifying the sentence as a whole), then processing considerations would predict that in a verb-final language, they would tend to occur at the end of the sentence, since that would result in a shorter CRD for the sentence than if they occurred at the beginning of the sentence.

In conclusion, it appears that for some pairs of elements whose status as correlation pairs is not explained by the BDT under assumptions of flat constituent structure, the correlations can be explained, at least partly, by a combination of Hawkins' Principle with factors that lead certain sorts of words to occur on the periphery of their mother constituents. But there are other pairs, like auxiliary verb and main verb, that do not seem explainable in this way. Further examination of the other correlation pairs that become problematic for the BDT under assumptions of flat constituent structure is still needed.

References

Bache, Carl. 1978. *The order of premodifying adjectives in present-day English*. Odense: Odense University Press.

Carlson, Robert. 1994. *A grammar of Supyire*. Berlin: Mouton de Gruyter.

Culicover, Peter W. and Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press.

Dixon, Robert M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25: 123-195.

Dryer, Matthew S. 1992a. The Greenbergian word order correlations. *Language* 68: 81-138.

Dryer, Matthew S. 1992b. Adverbial subordinators and word order asymmetries. In *Performance principles of word order*, eds. John A. Hawkins and Anna Siewierska, 50-67. EUROTYP Working Papers, European Science Foundation.

Dryer, Matthew S. 2005a. Relationship between order of object and verb and order of adjective and noun. In *World atlas of language structures,* eds. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 394-397. Oxford University Press.

Dryer, Matthew S. 2005b. Relationship between order of object and verb and order of relative clause and noun. In *World atlas of language structures,* eds. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 390-393. Oxford University Press.

Frazier, Lyn. 1978. On comprehending sentences: syntactic parsing strategies. Univ. of Massachusetts Ph.D. dissertation.

Giridhar, P. P. 1980. *Angami grammar*. Manasagangotri, Mysore: Central Institute of Indian Languages.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. Joseph, Greenberg, 73-113. Cambridge, Mass: MIT Press.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Heath, Jeffrey. 1999. *A grammar of Koyraboro (Koroboro) Senni*. Köln: Rüdiger Köppe Verlag.

Posner, Roland. 1986. Iconicity in syntax: the natural order of attributes. In *Iconicity: Essays on the nature of culture: festschrift for Thomas A. Sebeok on his 65th birthday,* ed. P. Bouissac, M. Herzfeld and R. Posner, 305-337. Tübingen: Stauffenberg.

Sproat, Richard and Chillin Shih. 1988. Prenominal adjectival ordering in English and Mandarin. In *Proceedings of the North Eastern Linguistic Society* (NELS) 18: 465-488.