

39

Frequency and pragmatically unmarked word order *

Matthew S. Dryer
SUNY at Buffalo

1. Introduction

Discussions of word order in languages with flexible word order in which different word orders are grammatical often describe one of the orders as the (pragmatically) *unmarked* or neutral word order, while other grammatical orders are all described as being *marked* in some way. In most languages in which one order has been so characterized, the order described as unmarked is also the order which occurs most frequently in spoken or written texts. It is widely assumed, in fact, that this is a necessary characteristic of unmarked word order, that it is part of what it means to be unmarked that the unmarked word order be most frequent. For example, Greenberg (1966:67) claims explicitly that the unmarked order in a language is “necessarily the most frequent”. There are instances, however, in which this assumption has been questioned, in which descriptions of word order in particular languages have claimed that a particular order is unmarked or neutral, even though that order is not significantly more frequent than other orders, and may in fact be less frequent than at least some other orders. I will discuss four instances from the literature in which claims of this sort have been made. The purpose of this paper is to explore the question of whether such claims make sense, whether there is a useful notion of pragmatically unmarked word order that is not necessarily the most frequent order. I will propose that a particular word order can be described as pragmatically unmarked if it is the *default* word order, if

there are concise ways to characterize the situations in which other word orders are used, with the pragmatically unmarked word order being most easily characterized as the order that is used elsewhere, once the situations in which the pragmatically marked orders are used have been characterized.

I discuss a number of at least hypothetical situations where it would make sense to say that one order can reasonably be described as pragmatically unmarked, even though not more frequent than other orders. I will also show that there are possible situations where two orders might differ in relative frequency, but not differ in pragmatic markedness. But I will argue that in none of the four actual cases where a claim has been made that the pragmatically unmarked word order is not significantly more frequent than other orders is it clear that the order in question is pragmatically unmarked in the way proposed here. And while it is possible for there to be mismatches between frequency and markedness, it is often not clear when people describe a word order as pragmatically unmarked that their characterization is based on anything more than frequency. Finally, I will propose that while frequency is epiphenomenal relative to the grammars of individual languages, it plays a much larger role than pragmatic markedness in explaining why languages are the way they are.

Three of the four actual examples I discuss here involve the word order of clauses containing a lexical subject and a lexical object, but no significance should be associated with this. I assume that questions of which order is unmarked can be asked of any elements in language whose order is determined by discourse factors. In fact, a number of the cases I discuss here involve alternations other than word order alternations. I assume that these questions regarding the relationship of pragmatic markedness to frequency apply to any alternation that is governed by discourse factors. As various people have noted (cf. Du Bois 1987), clauses containing both a lexical subject and a lexical object tend to be relatively infrequent in texts, and I assume that one could also argue that such clauses are pragmatically marked, relative at least to clauses containing a single lexical argument (or no lexical argument). While I believe that typological discussion often assigns such clauses a special significance they do not deserve, it still makes sense to ask what order, if any, among such clauses is pragmatically unmarked in a language, just as it makes sense to ask of a language what the unmarked order is in noun phrases containing a demonstrative, a numeral, an adjective, and a noun, even though such noun phrases will be very infrequent in any language.

2. Case 1: Tojolabal

In order to ground the discussion, I will start off by presenting one of the four cases I will discuss in which a particular order is claimed to be pragmatically unmarked, though not most frequent. This case is a description of clause order in Tojolabal, a Mayan language of Guatemala, by Brody (1984). Brody argues that while VOS order in Tojolabal is the least marked word order, it is infrequent, particularly by comparison with SVO order, which she claims is a marked word order. But I do not find her arguments for this conclusion convincing. Her only argument that SVO order is pragmatically marked is that it involves what she calls “topicalization” of the subject. But she does not explain what discourse property is associated with what she calls “topicalization”, except to say that it gives “prominence” to the subject. Since the terms “topic” and “prominence” are used by different linguists in a wide variety of ways, some of them even contradictory, this tells us very little about exactly what discourse factors are associated with SVO order in Tojolabal. For example, on one common use of the term “topic”, subjects are typically topics. I assume that most linguists would not want to describe utterances in which the subject is topic in that sense as pragmatically marked.

The following quotation from Brody (1984) provides further insight into the nature of her claim that VOS is unmarked, even though less frequent than SVO:

The VOS sentence communicates very little; in fact, it communicates nothing beyond propositional content. It reveals nothing about the speaker’s prejudices or presuppositions, nothing about the direction the discourse is moving toward or coming from, and nothing about the relative importance of the participants. In every significant sense, the VOS sentence in Tojolabal is an unimportant sentence type in discourse, albeit the basic sentence by most of the criteria examined above. It is simply not a very useful kind of sentence for furthering the flow of information. It is thus not surprising that an uninformative sentence type should occur with low frequency in discourse. The SVO sentence, on the other hand, is more frequent in discourse because it is less neutral, both semantically and from a discourse perspective, and it is also more informative, giving prominence to the subject NP. (Brody 1984:726)

I have great difficulty making sense of Brody’s claims. Cleft sentences in English are less neutral than basic sentences and are more informative in the sense that they give prominence to the clefted element. The logic of Brody’s argument would seem to imply that cleft sentences in English should therefore

be more common than non-cleft sentences. Similarly, transitive clauses in which both arguments are lexical are presumably in some sense “more informative” than clauses in which one of the arguments is pronominal. But then by Brody’s logic, one might argue that transitive clauses in which both arguments are lexical ought to be the most frequent transitive clause type. But as she herself notes, such clauses are infrequent in Tojolabal, much as they are in other languages. In fact, following suggestions by Chafe (1987), we might say that such clauses are infrequent precisely because they are *too* informative, and the information expressed by such clauses is often broken down into two separate clauses. Word orders that are described as pragmatically marked often involve some sort of unexpectedness, some information that involves a change in the direction of the flow of information, either because some information is counter to expectations or because a new participant is introduced to the discourse. Word orders that are described as pragmatically unmarked, on the other hand, are often used in clauses which continue the existing flow. Brody’s claim that pragmatically unmarked word order should be expected to be infrequent because it “reveals nothing ... about the direction the discourse is moving toward or coming from” ignores the fact that the default type of clause need not reveal anything of that sort, precisely because the discourse is proceeding in an expected fashion. But regardless of these problems, it is difficult to evaluate Brody’s claim that SVO sentences in Tojolabal are more informative than VOS sentences in the absence of any indication as to the nature of the prominence SVO order gives to subjects.

One is tempted to respond to Brody’s claim with the following counterargument. If we accept her claim that SVO order is most frequent in Tojolabal, then this fact alone tells us something about the discourse factors associated with SVO order. Namely, whatever discourse properties are associated with what she calls topics, these properties are apparently ones that subjects more often have, at least in clauses containing both a lexical subject and a lexical object. That in itself eliminates various possible interpretations of what she means by “topic”, such as that associated with so-called topicalization in English, since that type of situation occurs infrequently. But it also implies that being topic is something that is normal for subjects in Tojolabal, and thus we might say that the unmarked situation in Tojolabal is for subjects to be topics. And if that is the case, we might say that SVO order is the unmarked word order, not a marked one. Conversely, although we do not know what properties

are associated with subjects in VOS sentences in Tojolabal, whatever properties they are are ones that subjects more commonly lack, so we might say that VOS is a marked word order in the language.

While I think that there is some merit to this argument, and that it is at least as strong as Brody’s argument for the opposite conclusion, it does beg the question being addressed in this paper, since it assumes that the most frequent order is necessarily pragmatically unmarked. The question I wish to address here is whether there exists a useful notion of pragmatically unmarked word order that is distinct from that of most frequent order. In order to do that, I turn briefly to the ways in which the terms “marked” and “unmarked” are employed in other domains, to determine whether those notions can be extended naturally into the area of discourse.

3. Markedness as a general notion

The terms “marked” and “unmarked” have been used by linguists in other domains in a wide variety of ways. It is possible to distinguish a broad and a narrow sense in which the terms are used. On its broad use (cf. Greenberg 1966; Croft 1990), markedness is an umbrella term that refers to a cluster of notions that correlate with each other or bear family resemblances to each other. On its narrow use (cf. Andersen 1988), the term is restricted to what are viewed as core instances of markedness and the other notions are simply viewed as related notions which may correlate with markedness, but which are not themselves instances of markedness. I assume that regardless of how one chooses to use the term, the broad and narrow notions must be kept distinct.

A further distinction can be made between crosslinguistic markedness and language-particular markedness. Most of the references to markedness in the typological literature involve a crosslinguistic notion of markedness that is related to naturalness. When it is claimed that [p] is typologically unmarked relative to [b], what is meant is that language-particular manifestations of the unmarked status of [p] relative to [b], such as its being the sound found when a phonemic contrast is neutralized, are to be explained in terms of the articulatory or perceptual properties of [p] compared to those of [b], the assumption being that the neutralization in the direction of [p] is a natural and explainable phenomenon. And in many cases, such as this one, the distinction between cross-linguistic markedness and language-particular markedness may not be

obvious. However, when Greenberg (1966:39) notes that Oneida is exceptional in treating feminine gender as unmarked relative to masculine gender, the distinction is somewhat clearer, since the criteria that identify feminine gender as unmarked in Oneida are clearly distinct from those that identify masculine gender as the unmarked gender crosslinguistically. The crucial evidence that masculine gender is the unmarked gender crosslinguistically derives from the fact that in most languages exhibiting markedness differences between masculine and feminine, the language-specific criteria point to the masculine being unmarked. Thus language-specific markedness involves a set of traditional criteria or tests, while crosslinguistic markedness simply involves the frequency over languages in which the criteria for language-specific markedness point to the same value as the unmarked value. Hence we can say, without contradiction, that feminine gender is unmarked in Oneida despite being universally marked, since the criteria for it being universally marked only require that it be marked in a significant majority of cases. The distinction is particularly clear in the domain of unmarked word order, where the notion is usually used in a language-particular sense. While one might say that SV order is crosslinguistically unmarked, because SV order is more common, such usage is clearly irrelevant to the question whether SV or VS order in a particular language is unmarked. Similarly, one can say that SO order is crosslinguistically unmarked, but this is clearly irrelevant to whether SO or OS (if either) is pragmatically unmarked in a given language. For these reasons, I will restrict attention here to markedness in a language-particular sense.

Among the ways the term “markedness” is used in a language-particular sense in other domains, we can distinguish at least four types of markedness: formal markedness, distributional markedness, semantic markedness, and frequency (cf. Lyons 1977; Croft 1990). *Formal* markedness, sometimes referred to as morphological, morphosyntactic, or structural markedness, refers to the presence vs. absence of morphemes, or, sometimes, to the relative phonological size of morphemes. Thus *lioness* is formally marked relative to *lion*, because *lioness* contains an additional morpheme, and a passive sentence like *Mary was kissed by John* is formally marked relative to its corresponding active sentence *John kissed Mary*, since it contains two more morphemes than the active form. Word order differences in language generally do not involve differences in formal markedness, but occasionally they do. For example, in Yagua, an Amazonian language discussed by Doris Payne (1990 *inter alia*), both VS and SV order occur, as illustrated in (1).

- (1) a. *sa-júúy Anita.*
3SG-fall Anita
'Anita falls.'
- b. *Anita júúy.*
Anita fall
'Anita falls.'

VS order is formally marked in Yagua, however, since a subject clitic occurs on the verb when that order is used, as in (1a), but is absent when SV order is used, as in (1b). As discussed below, however, Payne (1990) argues that the SV order of (1b), while formally unmarked, is both less frequent and pragmatically marked. Quite apart from questions of pragmatic markedness, this case is an interesting one, since it illustrates the more unusual situation in which the formally marked member of a pair is the most frequent.

Distinct from formal markedness is the notion of *distributional* markedness: one form is said to be distributionally marked if it occurs in a proper subset of the morphosyntactic contexts in which the other occurs. This notion is relevant in the area of word order since there exist instances in which one word order has a more restricted distribution than another. In Turkish, for example, both OV and VO order are found in main clauses, while only OV order is found in relative clauses (Utschig 1985: 161); since VO has a more restricted syntactic distribution, it is distributionally marked. As with formal markedness, we may expect a distributionally marked order to be pragmatically marked as well, but this is not necessary and the fact that an order is distributionally marked in a language provides no basis for saying that it is pragmatically marked.

Text frequency is treated by some linguists, like Greenberg (1966) and Croft (1990), as a criterion for markedness, but those who use the term in a narrower sense have pointed out how frequency can deviate from markedness (cf. Andersen 1988:30, Comrie 1976:116-117). In contrast, Greenberg (1966: 67) claims that of the various criteria for markedness, only frequency is relevant to the notion of unmarked versus marked word order. He suggests, in fact, that this might provide the basis of an argument that frequency is primary as a criterion for markedness. However, if one assumes a narrower notion of markedness that excludes text frequency as a criterion, and if one accepts Greenberg's conclusion that frequency is the only notion of markedness that is applicable to the notion of unmarked word order, then one could draw a very

different conclusion from Greenberg, namely that what people call marked or unmarked word order has nothing to do with markedness at all, since its only resemblance to true markedness is the family resemblance via text frequency. Or, as Andersen (1988:30) puts it, if “unmarked” means no more than “most frequent”, then the former term “might as well be dispensed with”.

4. Pragmatic markedness as additional meaning

There is reason to question Greenberg’s assumption that frequency is the only notion of markedness that applies to marked and unmarked word order. For example, one aspect of the markedness of *lioness* relative to *lion* is that it is semantically marked in that the meaning of *lioness* involves the meaning of *lion* plus an additional component of meaning. There are alternations in discourse that seem to involve a similar kind of opposition. For example, (2b) can be construed as having the meaning of (2a), plus the additional presupposition that John saw someone.

- (2) a. *John saw Mary.*
 b. *It was Mary that John saw.*

Furthermore, (2a) can be used with the same presupposition, with suitable stress, as in (3), just as *lion* can be used to refer to a female lion, so the set of contexts in which (2b) is appropriate is a proper subset of the set of contexts in which (2a) is appropriate.

- (3) *John saw MARY.*

In general, we can say that a construction is pragmatically marked relative to another if the range of contexts in which it is appropriate is a proper subset of the set of contexts in which the unmarked construction is used.¹

It is not clear, however, how often discourse-governed alternations can be characterized in this way, how often a difference can be adequately characterized in terms of an added component of meaning in one alternant that is absent in the other. For example, while there have been attempts to describe the difference between active and passive sentences in English in this way, these attempts are unconvincing. Battistella (1990:108) attempts to characterize the active-passive contrast in this fashion, claiming that “passivization in effect topicalizes the promoted object, making it the information focus of the sen-

“tence”. But while I think there are a variety of problems with this characterization of passive sentences, the central point is that even if it were meaningful and true, we could equally well say, stripping away the derivational terminology, that “active sentences in effect treat the ‘notional subject’ as topic, making it the information focus of the sentence”. On such a view, active sentences would involve the addition of a component of meaning absent in the passive just as much as passive sentences involve the addition of a component of meaning absent in the active, so the argument provides no basis for identifying a sense in which passive sentences in English are pragmatically marked relative to active sentences.

5. Pragmatic markedness in terms of discourse factors

Many recent approaches to discourse-governed alternations describe the alternation, not in terms of a difference in meaning, but rather in terms of the discourse factors that govern or determine the choice of one construction over another. I assume that one component of a speaker’s knowledge of a language is the discourse grammar of their language, the set of rules or principles that defines the association between particular constructions and the discourse factors that determine the use of those constructions. Thus, where languages allow flexibility of word order, I assume that the discourse grammars of those languages define exactly the circumstances in which the different orders are used. And I assume that questions about pragmatic markedness are questions about the particular rules or principles of discourse grammar, that to say that one word order is pragmatically unmarked in a language is to say something about the rules or principles of discourse grammar that govern word order in the language.

Similarly, I assume that the question of whether passive is pragmatically unmarked in English is a question about the nature of the rules or principles that govern the use of passive. Now there are theories of passive that characterize the choice between active and passive as involving some property or set of properties X such that the choice is governed by a principle like that in (4). Tomlin (this volume) assumes a theory of this form.²

- (4) a. Passive is used when the P (“patient” / “notional object” / “undergoer”) has property X;

- b. Active is used when the A ("agent" / "notional subject" / "actor") has property X.

But if the choice between active and passive is governed by a principle like that in (4), then the opposition would be an equipollent one, not a privative one, to use traditional markedness terminology from Trubetzkoy (1969), and there would be no sense in which passive is pragmatically marked. In other words, if we just examine the nature of the rule in (4), we find an essentially symmetric relationship between the specification of the conditions under which passive is used and the specification of the conditions under which active is used.

But there are also theories of passive that *do* treat the opposition of active and passive as a privative opposition, with the passive pragmatically marked. Both Tomlin (1983) and Thompson (1987) discuss a principle governing the use of passive that can be loosely paraphrased as in (5).³

- (5) Passive is used in English if and only if the P is more thematic than the A.

Exactly what "thematic" means is not crucial here; what is crucial is the logical form of the rule. Note that (5) is equivalent to the conjunction of the three propositions in (6).

- (6) a. Passive is used if P is more thematic than A.
 b. Active is used if A is more thematic than P.
 c. Active is used if A and P are equally thematic.

Implicit in (5), and made explicit in (6c), is the claim that the active is used when the A and P are equally thematic. This is an instance of neutralization, one of the classic criteria for markedness. If an analysis like that in (5)/(6) is right, then we *do* say that the active construction is pragmatically unmarked, in that when the factor determining the choice of active versus passive is neutralized, in this case the different thematicity of the A and the P, it is the active that is used. The contrast between (4) and (5)/(6) illustrates one possible way of distinguishing oppositions that do not involve pragmatic markedness and ones that do.

The two types of cases just discussed can also be used to illustrate the relationship between pragmatic markedness and frequency. The fact that passive clauses in English, especially ones with the A expressed, are far less frequent than active clauses provides no basis in itself for concluding that passive clauses are pragmatically marked and no evidence against an analysis

that takes the form of (4). For the greater frequency of active clauses might be entirely due to As having property X far more often than Ps do. The frequency of different constructions is a function of two things, the discourse factors underlying use of the construction, which are part of discourse grammar, and the frequency with which those factors tend to hold in typical human discourse, which is independent of discourse grammar. As Du Bois (1985:357) notes, "neither humanness nor agency are components of a notion of Topic Rather, only at the level of discourse tokens ... do the properties of humanness, agency, and topicality or thematicity tend to appear as recurrent clusters." It is possible, therefore, for an opposition to be equipollent, i.e. without a markedness relation holding between them, while one member of the opposition is more common than the other due to the nature of typical human discourse.

Another example illustrating this general point is provided by Payne's (1987) discussion of word order in O'odham (Papago). Payne shows that there is a strong association between definiteness and the position of nominals relative to the verb: some well-defined exceptions aside, indefinite nominals tend to precede the verb while definite nominals tend to follow. Thus, oversimplifying somewhat, we could describe O'odham word order as governed by the principle in (7).⁴

- (7) a. If the nominal is indefinite, then preverbal.
 b. If the nominal is definite, then postverbal.

The principle in (7) characterizes the opposition between preverbal and postverbal position in O'odham as an equipollent one, since the form of the rule provides no basis for treating one order or the other as unmarked. While Payne shows that postverbal position is more common in O'odham, this frequency difference does not in itself provide any reason to conclude that postverbal position is pragmatically unmarked. For the frequency difference can be explained entirely in terms of the fact that definite nominals tend to be more common than indefinites in typical human discourse. The difference in frequency does not reflect anything about the discourse grammar of O'odham.

These examples make clear that it is possible, at least in principle, for one word order to be more frequent than another without that order being pragmatically unmarked. Conversely, it is possible to describe at least hypothetical situations where one order is pragmatically unmarked in that it is the order that is used when the factors determining choice of order are neutralized, but where that order is not more frequent. Consider the hypothetical situation in (8).

- (8) Suppose there is some property X such that
- order AB is used when A has property X more than B;
 - order BA is used when B has property X more than A;
 - order AB is used when A and B do not differ in property X.

By the criteria discussed above, order AB is pragmatically unmarked relative to order BA since it is used in neutralized contexts. Suppose further that because of *extralinguistic* factors, the relative frequencies of the three situations in (8) are as given in (9).

- (9)
- | | | |
|----|-------------------------------------|-----|
| a. | A has property X more than B | 20% |
| b. | B has property X more than A | 60% |
| c. | A and B do not differ in property X | 20% |

Under such circumstances, the pragmatically marked order, namely BA, would occur more frequently than the pragmatically unmarked order AB, since BA order occurs with the single condition that arises 60% of the time. There are reasons to believe that such situations are atypical, but again the theoretical possibility of pragmatic markedness not matching up with frequency is illustrated. Because such situations are atypical, frequency may be a useful *diagnostic* for pragmatic markedness, even if ultimately it is not a defining characteristic.

Situations in which one word order is the order used when conditioning factors are neutralized is one type of situation in which a word order can be described as pragmatically unmarked. Such situations can be viewed as special instances of a more general type of situation in which one word order is the default order in the sense that the easiest way to characterize the contexts in which that order is used is to specify when other orders are used and then state that the order in question is used *elsewhere*. Such an approach is often implicit in descriptions of word order, in which linguists describe the contexts in which the marked word orders are used, with the understanding that the unmarked word order is used elsewhere. Similar comments apply to other discourse-governed alternations. For example, in her discussion of when passive is used in English, Thompson (1987) never directly addresses the question of when active is used, with the implicit understanding that active is used elsewhere. Similarly, Mithun (1990) describes a four-way contrast between four types of third person reference in Central Pomo. Her characterization of this contrast (p. 371) is summarized in (10).

- (10)
- If completely new, outside of speakers' consciousness, then use full lexical NP.
 - If already within focal consciousness, then no marker ["zero pronoun"] is used.
 - If referent is entity from whose point of view information is presented, then use empathetic *tʃi* pronoun.
 - Otherwise, use *muul* pronoun.

For the first three reference types in (10), it is possible to characterize succinctly the conditions under which that reference type occurs. But Mithun notes (p. 371) that the pronouns that she refers to as the *muul* pronouns serve a constellation of functions and that they can be most easily characterized by contrasting their use with that of the other referential forms. In other words, they are the elsewhere case, and their use is most easily characterized by describing where each of the other three types is used. This constitutes a reason to describe them as the pragmatically unmarked referential form.

While it is probably the case that the default within a range of alternatives in opposition to each other is usually the most frequent alternative, this is not a necessity. For example, while Mithun does not discuss the relative frequency of the different referential forms in Central Pomo, there is no reason to believe that the *muul* pronouns, the pragmatically unmarked option, are more frequent than the alternative forms, and her characterization of the use of zero suggests that zero may even be more common. Similarly, within the context of word order, one can imagine hypothetical situations in which the default order occurs less frequently than other orders. Consider the hypothetical situation in (11).

- (11) BA if X and Y
 AB otherwise

Given a rule like that in (11), we can describe AB as the pragmatically unmarked order, since it is the elsewhere case. Now suppose that as a result of extralinguistic factors, X is true 80% of the time and Y is true 80% of the time. Then the expected frequency of the different possibilities (assuming X and Y independent) will be as given in (12).

- (12)
- | | | |
|-----|-----|-----|
| | Y | ~Y |
| X | 64% | 16% |
| ~ X | 16% | 4% |

But then, given the frequencies in (12), it follows that BA, the pragmatically marked order, will occur 64% of the time, while AB, the unmarked order, will occur only 36% of the time. In this situation, the marked order would be more frequent than the unmarked order.

While these examples point to the conclusion that one can identify a notion of pragmatic markedness that is independent of frequency, it is not clear how often linguists use the term in this way. In languages in which one word order has been described as unmarked, there is often not clear evidence that the given order is the default order, since linguists rarely provide even observationally adequate accounts of what determines word order alternations. Rather, claims that a particular order is unmarked generally seem to be based on nothing more than an impression that that order is most frequent, the assumption being made that higher frequency necessarily entails some property of the principles underlying the word order alternations. But I have argued here that while that assumption may generally be true, it is not necessarily the case. Thus, claims that a particular order is unmarked may often mean no more than that that order is most frequent.

One possible shortcoming of the approach to pragmatic markedness I have proposed here is that it appears to be dependent on a particular analysis. This problem is especially acute if two orders are in an essentially equipollent relationship, as in (7) above, or any case of the form in (13).

(13) AB if X; BA if Y

If X and Y exhaust the set of possible situations, (13) is equivalent to both (14) and (15).

(14) AB if X; BA elsewhere.

(15) BA if Y; AB elsewhere.

But if we take the default order to be the one that is specified as occurring elsewhere, then the situation in (13) to (15) presents a problem. If we assume (14), then order BA is pragmatically unmarked; if we assume (15), then AB is pragmatically unmarked; and if we assume (13), then neither is pragmatically unmarked. But since all three analyses describe the same set of facts, this is clearly an undesirable result.

What this example illustrates is that we need to revise the notion of pragmatically unmarked word order to be a default word order, not in the sense that one *can* provide an analysis under which the order is the elsewhere case,

but rather in the sense that one *cannot* easily specify when a construction is used without either listing a set of different conditions under which the order is used, or specifying the conditions in which *other* orders are used and then stating that the order in question is used elsewhere. Furthermore, we need to stipulate that a word order is pragmatically unmarked only if one *can* specify in relatively simple terms the situations in which all *other* orders are used, since one clear property that a pragmatically unmarked order must have is that it be in contrast to orders which are pragmatically marked.

The discussion so far has assumed that it makes sense to talk of the most frequent word order in a language. But in fact it is not entirely clear whether such a notion is always meaningful. Because the frequency of different linguistic constructions depends entirely on the frequency with which the different factors conditioning the choice arise, frequency will actually vary from discourse type to discourse type, from text to text, and from subtext to subtext. Where word order is sensitive to aspect, a word order that is more frequent in narrative may be less frequent than other discourse types. Where word order is sensitive to contrasting participants in a text, a text with two primary participants may exhibit rather different frequencies from one with a single primary participant. And clearly one order may be more frequent than another for similar reasons in one section of text but not in another section of the same text. Thus while the relative frequency of different orders may be well-defined for a given corpus of texts, it is less clear what it means as a feature of the language itself. In some languages, one order may be more frequent than another in most texts, and for such languages, it may make sense to say that that order is most frequent in the language. But for other languages, this may not be the case, and for such languages it presumably does not really make sense to say that one order is more frequent in the language.

6. Tojolabal again

Let me turn now to discussion of four cases from the literature where a word order has been claimed to be (pragmatically) unmarked though not significantly more frequent than some other order. The first case is the case of Tojolabal, discussed above. I argued above that the arguments by Brody (1984) for treating VOS as pragmatically unmarked are unconvincing since it is quite unclear what she means by "topic" and thus quite unclear what

determines the choice of SVO versus VOS in the language. If there is some argument that the easiest way to describe where VOS is used is by describing where SVO order is used and specifying that VOS order is used elsewhere, then we would have an argument that VOS is pragmatically unmarked. But Brody gives no argument for that conclusion. Without clear claims as to what determines word order in Tojolabal, supported by evidence from texts, this case remains unconvincing.

7. Case 2: Ojibwa

The second case involves Ojibwa, an Algonquian language. Tomlin and Rhodes (1979) claim that the unmarked word order in Ojibwa is VOS. They note that this is "obscured" by a number of principles. Rhodes (personal communication) informs me that SVO order appears to be at least as common as VOS in texts, and my own text counts of Ojibwa texts in Nichols (1988) revealed the frequencies given in Table 1.⁵ While the numbers here are too small to be taken too seriously, it is notable that only 5 out of 25 clauses with lexical subject and lexical object are VOS, the order claimed to be pragmatically unmarked.

The claim that VOS order is unmarked in Ojibwa is based on the fact that SVO order is claimed to occur in one of a set of marked situations, when the subject is indefinite, contrastive, or "the theme of a local section of text". VOS order is used when the subject lacks these properties. VOS order is used in situations more akin to ones where either the subject and the object are realized only by the pronominal affixes on the verb, ones where the subject or object is relatively predictable. The argument that VOS is unmarked runs as

Table 1. Ojibwa clauses with lexical subject and lexical object

SVO	12	(48%)
VOS	5	(20%)
VSO	4	(16%)
OVS	3	(12%)
SOV	1	(4%)
Total	25	

follows. It is assumed that the unmarked situation for an argument of the verb is one in which it is relatively predictable, and that the unmarked transitive clause is one in which both arguments are relatively predictable. In the majority of such clauses, one or both of the arguments is realized entirely by the verb morphology, and thus it does not contain two lexical arguments. Only in a minority of those situations in which both arguments are relatively predictable will both arguments be realized by full noun phrases, and in those cases VOS order will occur. But SVO order is as common as VOS because it is used in situations in which the subject is less predictable, and in those situations a full noun phrase is always used. Thus VOS is a subinstance of the more frequent situation in which both arguments are relatively predictable, but since it is used in only a subset of those situations, it is no more frequent than SVO.

There are a number of problems with this argument. First, it is important to realize that markedness is a *relative* notion. A category A can be marked relative to B, even though A is a subinstance of a larger construction type C which is *unmarked* relative to B. Similarly a category A can be unmarked relative to B, even though both A and B are instances of a construction type that is *marked* relative to some other category C. One can plausibly argue that transitive clauses in which both the subject and object are lexically realized are pragmatically marked, relative to clauses in which either the subject or object is pronominal, realized either by an independent pronoun or by a pronominal affix on the verb. In any language, therefore, any SVO or VOS clause with lexical subject and object will be marked relative to clauses with a pronominal subject or object. The question of whether VOS or SVO is the unmarked order in a language is thus a question of which order (if any) is unmarked relative to other clauses with lexical subject and lexical object, even though the set of such clauses are all marked relative to other clauses. The fact that the discourse conditions under which VOS order occurs more closely resemble those in which the subject and object are pronominal thus constitutes no argument that VOS order is unmarked relative to SVO.

An analogy from phonology may be useful. Although the notion of markedness in phonology may differ in some ways from the notion of markedness in the context of discourse, the general point made in the preceding paragraph still holds. Namely, we may say that voiced sounds are unmarked overall relative to voiceless sounds, since, except for obstruents, there is an overall preference for voiced sounds. Among obstruents, however, voiceless

appears to be the unmarked value: in languages which lack voicing contrasts for obstruents, or for certain obstruents, the more common alternant is typically the voiceless one. The fact that voiced sounds are unmarked relative to voiceless sounds overall constitutes no argument against the claim that voiced obstruents may be marked relative to voiceless obstruents. Analogously, while transitive clauses with highly predictable subject and object might be unmarked overall, it might still be that among clauses with a lexical subject and a lexical object, the unmarked situation is one in which one of the arguments is less predictable.

To answer the question of whether postverbal position is pragmatically unmarked in Ojibwa, we must examine the form of the rule governing the distribution of the different orders. Tomlin and Rhodes' account basically is that given in (16).

- (16) a. An indefinite NP follows the verb if
- i. the verb is a quantifier verb, such as *baatiinak* 'be many' or *niizhig* 'be two';
 - ii. the NP is thematically irrelevant.
Otherwise, it precedes the verb.
- b. A definite NP precedes the verb if
- i. the NP is a quantifier;
 - ii. the NP is contrastive;
 - iii. the NP is the theme of the local section of text.
Otherwise, it follows the verb.

The notion of pragmatic markedness, interpreted as default choice, is relevant to the description of Ojibwa word order only in the sense that we can say that the pragmatically unmarked position for *indefinites* is preverbal while the pragmatically unmarked position for *definites* is postverbal. The reason that these are pragmatically unmarked is that the rule in (16) specifies these choices as the default or elsewhere choice. However, the form of the rule in (16) does not treat either preverbal position or postverbal position as pragmatically unmarked, since neither of these is defined as the default choice in general. In other words, the form of the rule in (16) treats the relation between preverbal position and postverbal position as equipollent. For this reason, there is no reason to say that postverbal position is pragmatically unmarked.

8. Case 3: Yagua

The third case I will discuss is Yagua, a language of eastern Peru. Payne (1990: 237) argues that VSO is the unmarked word order in Yagua. Her argument for this conclusion runs as follows. First, she argues that preverbal position for nominals is in general pragmatically marked in Yagua, regardless of their grammatical or semantic status. Second, she argues that although a variety of orders are common for clauses containing a lexical subject and a lexical object, VSO is the only order possible in which both follow the verb. Since postverbal position is in general the unmarked position for lexical noun phrases, VSO must be the unmarked order among clauses with a lexical subject and a lexical object.

VSO order is not, however, significantly more frequent than other orders. While it is true that in general, both subjects and objects more commonly follow the verb in Yagua, the data in Table 2 from Payne (1990) show that VSO order is only slightly more common than SVO in one set of texts and is found in only a minority of clauses containing a lexical subject and a lexical object.⁶

While VSO is more common than the other orders in this count, the difference is small and probably not statistically significant. Furthermore, VSO accounts for fewer than 40% of clauses containing a lexical subject and a lexical object. This would thus be a possible case of a language in which the pragmatically unmarked order is not significantly more frequent than other orders.

Table 2. Yagua clauses with lexical subject and lexical object

VSO	19	(39%)
SVO	15	(31%)
OVS	11	(22%)
O,SV	3	(6%)
S,OV	1	(2%)
Total	49	

Table 3. Types of pragmatically marked preverbal constituents in Yagua

Single focus contrast	38	(11%)	It was <i>Mary</i> that John saw.
Double focus contrast	43	(13%)	Her <i>husband</i> stayed at home to <i>babysit</i> , and <i>she</i> went to <i>work</i> .
Restatement	45	(13%)	They were eating apples; <i>apples they were eating</i> .
Added detail restatement	37	(11%)	They were eating apples; they were eating <i>green apples</i> .
(Wh-)question	81	(23%)	<i>Who</i> did Mary see?
Answer to (wh-)question	13	(4%)	John saw <i>Mary</i> .
Counter expectation	11	(3%)	He brought for us <i>soda pop</i> to drink.
Negation	4	(1%)	He doesn't want <i>manioc beer</i> .
Threats	6	(2%)	<i>The beetle</i> will plant you under the ground.
Unexplained	67	(19%)	

Payne's claim that preverbal position is pragmatically marked in Yagua is based on a detailed analysis of the types of clauses containing preverbal nominals. Her data for these types are given in Table 3.

I will discuss shortly Payne's arguments that these types are pragmatically marked. But it is worth noting first that her argument that VSO is unmarked assumes that the position of transitive subjects is governed by the same principles as those governing other nominals. However, transitive subjects precede the verb more often than intransitive subjects or objects, as shown by the frequency numbers given in Table 4, extrapolated from her data. Table 4 shows the relative frequency of the different orders of verb with respect to transitive subject (S_t), intransitive subject (S_i), and object.

Table 4 shows that while intransitive subjects and objects clearly follow the verb more often in Yagua, transitive subjects follow the verb only marginally more often.

There are three possible accounts of this difference between the figures for subjects of transitive clauses and those for intransitive subjects and objects. First, it might be random variation, without linguistic significance. However, the difference turns out to be statistically significant at the .001 level (chi-square). Second, it might be that, because of the nature of discourse, S_t 's more often fall into one of the pragmatically marked categories that Payne classifies preverbal elements in Yagua as falling into. It is already known (cf. Du Bois

Table 4. Order of nominal and verb in Yagua by type

VS_t	47 (55%)	VS_i	210 (74%)	VO	186 (74%)
S_tV	39 (45%)	S_iV	75 (26%)	OV	65 (26%)

1987) that transitive subjects generally differ from intransitive subjects and objects in being lexical far less often. This might be tied to a tendency to fall into one of Payne's pragmatically marked categories more often. A third possibility is that the discourse principles governing the position of transitive subjects are somewhat different from those governing other nominals. According to Payne (1990:219), 19% of the preverbal constituents do not fall into one of her nine pragmatically marked categories and are unexplained; perhaps a higher than random proportion of these are transitive subjects. If this third possibility is true, then Payne's case that VSO is pragmatically unmarked would be severely weakened. And in the absence of an explanation for the higher frequency of S_tV order, Payne's argument that VSO order is pragmatically unmarked is unconvincing.

Payne's argument that the types of clauses in Table 3 are pragmatically marked resembles Givón's (1979) attempts to define pragmatic markedness in terms of presuppositions. She characterizes unmarked assertions as "those assertions which largely rehearse what the speaker assumes are already-established links, or which attempt to make new links of a fairly incremental nature." She claims that with marked speech acts "the major portion of the predication ... is presupposed and is not asserted" and "the speaker assumes something about what the hearer holds to be true (or at least will accept without challenging), but takes pains to modify in some specific way what the hearer takes for granted". Payne also notes (personal communication) that her notion of pragmatic markedness can be characterized in terms of a notion of counterexpectation.

Payne's characterization of pragmatic markedness can be evaluated both from the perspective of its success in describing the particular Yagua clause types in Table 3 that are associated with preverbal constituents, and in terms of the question of whether it provides a general characterization of pragmatic markedness. With respect to the Yagua clause types in Table 3, her characterization does seem to fit a number of the types, like single focus contrast, added detail restatement, answer to a wh-question, counterexpectation, and

negated clauses. But it is less clear that it is successful in characterizing some of the other types, such as double focus contrast, where she herself admits (p. 202) that the situations are not necessarily taken as presupposed, and restatements, which she also admits fail to fit her characterization of pragmatic markedness. But these latter two types are in fact the two most frequent types with preverbal constituents, other than *wh*-questions, where the preverbal position of the interrogative expression may be grammaticized as it is in many other languages, like English. I therefore remain unconvinced that she has successfully justified her characterization of these clause types as pragmatically marked, even in her use of the term.

Payne's notion of pragmatic markedness differs from mine in that it is based on substantive pragmatic notions rather than the logical form of the underlying principles. In this respect, she is using the term "markedness" in a fashion that is different from its traditional use. Nevertheless, one might view the difference between Payne's use of the expression "pragmatic markedness" and mine as merely terminological. But the question remains whether her notion of pragmatic markedness succeeds in characterizing the range of phenomena in other languages that others might want to employ the term for. For example, if the choice between active and passive in English is governed by a principle like that in (6) above, according to which active is used in neutralized situations, then I assume we would want to say that passive is pragmatically marked relative to active. However it is quite unclear that passive would count as pragmatically marked under Payne's definition.⁷ Similarly, right dislocations in English, like (17), are presumably pragmatically marked, but again Payne's characterization would not treat them as such.

(17) *They're playing better now, the Red Wings.*

The same comment applies to inverted sentences in English like (18).

(18) *Beside a pond lived a little turtle.*

The sentence in (18) is the initial sentence from a children's story.⁸ But it does not seem to fit Payne's characterization of pragmatic markedness, since the speaker/writer presumably makes the fewest presuppositions at the start of a story. Yet the structure in (18) is most appropriate in precisely that kind of situation. It is therefore unclear that Payne's use of the term "pragmatic markedness" is consistent with how others use the term.

Payne's notion of pragmatic markedness seems, to some extent, to be language-specific. Even if there is some pragmatic characteristic shared by all

of the preverbal types in Table 3, it does not seem to be the case that word order in other languages is defined by exactly the same notion. It does not, for example, seem to fit marked word order in Slavic languages. I assume, however, that languages can vary widely in what discourse factors are associated with pragmatically marked word order. While one language may use a marked word order in certain situations, another language may use the unmarked word order in corresponding situations, and use a marked word order in situations in which the first language uses its unmarked word order. For these reasons, on my use of the expression "pragmatic markedness", any attempt to define pragmatic markedness in universal pragmatic terms cannot succeed. I have argued that what defines pragmatically unmarked word order is the general characteristic of being the default order, and the actual discourse conditions associated with the marked orders may vary considerably from language to language.

While I find Payne's arguments that postverbal position in Yagua is pragmatically unmarked unconvincing, both because it is not clear that all of the preverbal types satisfy her notion of pragmatic markedness and because it is not clear that her notion of pragmatic markedness is consistent with how others use the term, the question remains whether postverbal position in Yagua might be pragmatically unmarked in the sense of this paper, in being the default order. The fact that Payne lists the situations in which constituents precede the verb means one could reinterpret her analysis and simply say that postverbal position is pragmatically unmarked simply because it is the default order, the order used when none of the conditions in Table 3 holds. The primary problem facing this is the fact that 19% of the preverbal constituents in the texts she examined are unexplained by her criteria. In addition, she also notes the existence of a smaller number of postverbal constituents that fall into one of the types in Table 3 generally associated with preverbal position. Thus while her account goes a long way towards describing the factors conditioning Yagua word order (it accounts for 81% of preverbal constituents and over 98% of postverbal ones), the existence of unexplained instances means that it is not possible to identify a default order, since that concept requires that we have an observationally adequate description which accounts for all instances of the different orders. What this illustrates is that it is very difficult to provide convincing evidence that a given construction is pragmatically unmarked.

9. Case 4: Macushi

Abbott (1991:25) describes OVS as the least pragmatically marked order in Macushi, a Carib language of Guyana and adjacent areas of Brazil and Venezuela. She cites text counts, however, showing 39 instances of transitive clauses with S(O)V order and 32 instances with (O)VS order, where the "(O)" signifies that transitive clauses with or without an expressed object are included. Because these counts include transitive clauses in which the object is not expressed, they really signify the relative frequency of transitive subject with respect to verb and the markedness question is thus whether VS order is the pragmatically unmarked order for transitive subject with respect to verb. Unlike the other cases discussed here, pronominal subjects as well as lexical subjects are included. Abbott claims that VS order is pragmatically unmarked despite the fact that SV order is slightly more frequent, at least in this count.

Abbott's argument that SV order is pragmatically marked is that clauses employing this order either "highlight a change of topic" (33 cases) or involve fronting the subject "for other special discourse-pragmatic effects", while VS order is used when "the subject refers to a topic previously introduced and not especially highlighted" (p. 25). Her argument thus resembles Payne's argument for Yagua in that she assumes that whether an order is pragmatically marked or not depends on the intrinsic pragmatic properties of the clause, rather than the kind of notion I have proposed here, based on default order, which is independent of the actual discourse properties associated with the different orders. Her notion of pragmatic order also clearly resembles Payne's since a pragmatically marked order is assumed to be associated with a higher degree of unexpectedness than pragmatically unmarked clauses.

Although Abbott's argument that VS is pragmatically unmarked assumes a different notion of pragmatic markedness from the one assumed in this paper, the particular characterization she gives conforms precisely to the notion of pragmatic markedness discussed in Section 5. Namely, she provides two situations in which SV order is used (change of topic or special effect) and characterizes when VS order is used at least partly in terms of the *absence* of the conditions in which SV order is used (previously introduced topic and *not* especially highlighted). Characterizing the situation in which VS order is used partly in terms of a notion "not especially highlighted" essentially involves characterizing this use in terms of the absence of the conditions for VS order and is in effect a way of saying "elsewhere".

The Macushi case comes closer than any of the first three cases to satisfying the conditions required for the most frequent order not to be the pragmatically unmarked word order, but we would have to know that Abbott's characterization describes the Macushi facts accurately before we could conclude that it is a convincing case. One initial shortcoming is the vagueness of the characterizations "not especially highlighted" and "other special discourse-pragmatic effects"; testing claims of this sort is necessarily rather difficult. Furthermore, unlike Payne's analysis of Yagua word order, which was based on a detailed empirical study, the basis for Abbott's conclusions is unclear. Examination of the sample Macushi text she provides with her description of the language reveals cases that do not seem to conform to her description. In particular, at least one sentence, her (22) on p. 156, involves a preverbal subject that continues a subject from the preceding text and does not appear to be fronted for any obvious special pragmatic reason. Conversely, at least one sentence, her (45) on p. 159, contains a postverbal subject introducing a new topic, again contrary to her principle. But if her characterization has exceptions, then, as with Payne's, we have no way to evaluate which order is the default order. As a result, this case too is unconvincing.

10. Some recent proposals by Givón

It is worth examining briefly recent proposals regarding markedness by Givón (1990). Givón discusses three notions of markedness, two of which are identical to two of the notions I have discussed here, namely formal markedness and frequency, and the third of which he calls cognitive complexity. His notion of cognitive complexity is similar to pragmatic markedness to the extent that both are cognitively-based markedness notions that are associated with, but distinct from, formal markedness and frequency. But otherwise, his notion is rather different from pragmatic markedness, at least as I use the term. As stated earlier, I assume that pragmatic markedness is a property of the rules or principles of discourse grammar, which determine the choice by speakers of a particular construction from a set of discourse-governed alternants. Since such rules or principles are relevant to explaining the choices that the speaker makes, pragmatic markedness is a speaker-based notion. But Givón's notion of cognitive complexity is purely a *hearer*-based notion. It involves the amount of effort or the number of cognitive operations that take place in the

mind of the hearer in processing. Givón argues, for example, that definite nominals have greater cognitive complexity than pronouns, since processing them involves a more complex mental search for the intended referent than is the case with pronouns.

Givón's notion of cognitive complexity differs from my understanding of pragmatic markedness in a number of other ways. First, many different factors contribute to cognitive complexity, among them, formal markedness and frequency. In some cases, formally marked structures will require more processing, simply because they involve more morphemes. In other cases, apparent differences in cognitive complexity may reflect no more than differences in frequency: to some extent, structures that occur quite frequently are presumably recognized and processed more quickly than structures that occur infrequently, simply because they are more familiar as a function of frequency.

Secondly, while the choices that speakers make are often governed by their model of the hearer, particularly in the case of reference, it is not clear to what extent hearers attend to all the choices that speakers make. This is particularly relevant in the case of flexible word order. In a language which permits all six orders of subject, object, and verb, I assume that there are *some* principles or determining factors in the mind of the speaker that completely determine the choice of order for these elements for every utterance that contains these elements. But there is no reason to believe that the hearer attributes some significance to every one of these choices. It is a matter of logical necessity that there be *some* cognitive events in the mind of the speaker underlying every choice of word order, since words cannot be unordered and the order results only from something in the mind of the speaker. But it is not a matter of logical necessity that hearers attend to every fact in the order of words in utterances they process — just as they presumably do not attend to all of the phonetic details of the words they hear — and there is very little empirical evidence bearing on just what hearers do attend to, and thus no evidence that hearers do attend to all details of word order. Hence, there may be word order alternations that are governed by something in the mind of the speaker, but which are ignored by hearers, so any cognitively based theory which refers exclusively to hearers is potentially inadequate.

A third feature of Givón's notion of cognitive complexity is that it is extremely context-sensitive. Since cognitive complexity is a function of an extremely broad array of factors, the cognitive complexity of a given utterance is a function of all of these factors. While we can talk of the cognitive

complexity of a given utterance in a given context, it is less clear how we can talk about the cognitive complexity of a given construction. For example, if passive sentences are cognitively less complex than active sentences in some contexts but more complex in other contexts, it is not clear how we compare the cognitive complexity of active and passive in general, or of one order compared to another, without appealing once again to frequency. Hence Givón's notion cannot provide the basis of a notion of unmarked word order distinct from frequency. Note, in contrast, that the notion of pragmatic markedness, as I am using the term, is not context-sensitive in the same way, since it is a property of discourse grammar, of the rules or principles that underly discourse-governed choices. But these rules or principles are not context-sensitive in the way cognitive complexity is. In so far as word order alternations are context-sensitive, that context sensitivity is built into the rules or principles. The rules or principles themselves do not vary with the context; rather the rules or principles define how the word order varies with the context. Hence it is possible to describe a construction as pragmatically marked or unmarked independent of the context of use. Whatever the merits of Givón's notion of cognitive complexity, it is at best a further notion of markedness, distinct from pragmatic markedness.

11. Conclusion

Let me finish with some general comments about the relative significance of frequency and pragmatic markedness. Pragmatic markedness, as I have used the term here, reflects the nature of the rules or principles underlying production, and thus is part of the grammar of particular languages. Frequency, on the other hand, is epiphenomenal relative to the cognitive structures or mechanisms underlying language production and understanding. It thus plays no role in the grammar of particular languages. On the other hand, while pragmatic markedness may play a role in the grammar of particular languages, it is not clear that it has any important role to play in explaining why languages are the way they are. Conversely, frequency, while it may be epiphenomenal relative to the grammar of particular languages, seems to play a major role in explaining why languages are the way they are. While a number of factors contribute towards differences in formal markedness, one of the primary factors appears to be what Haiman (1983) calls economic motivation, whereby the length of

linguistic structures will be inversely proportional to their frequency. Similarly, Du Bois (1987:851) argues that much of the nature of language can be explained in terms of a principle that “grammars code best what speakers do most”.

Such explanations assume that a notion of greater frequency is well-defined, and presumably the frequency differences in the relevant cases are sufficiently stable over texts that the greater frequency is well-defined. Furthermore, it is probably the case that the bulk of actual human discourse involves certain types of conversation and that while frequency differences may be great over different discourse types in a language, and even over different kinds of conversation, there may be certain core types of conversational discourse that occur with sufficiently high frequency relative to others that the frequency differences found in such discourse types are the ones that are crucial in driving the explanatory forces that are sensitive to frequency. It is precisely because of such considerations that a number of studies that examine the relative frequency of different constructions, with the goal of explaining why language is the way it is, base their frequency counts on conversational discourse (e.g. Thompson 1988; Fox 1987; Fox and Thompson 1990).

Within the context of word order, there are cases where word order changes seem explainable only in terms of frequency. I argue in Dryer (1989) that O’odham (Papago) has recently been undergoing a change from GenN order toward NGen order and a change from postpositions toward prepositions. This change is apparently the result of a recent change in the discourse factors governing word order variability in O’odham, with the side effect that VS and VO order have become more frequent than SV and OV. Payne (1987) argues, however, that O’odham is not grammatically VS and VO. Rather, as I discussed above, the greater frequency of VS and VO in O’odham is simply an epiphenomenal side effect of the fact that word order in O’odham is largely driven by definiteness, that VS and VO order are more frequent only because definite noun phrases happen to occur more frequently than indefinites. Thus while the greater frequency of VS and VO is epiphenomenal relative to the grammar of O’odham, it is necessary to appeal to the frequency facts in order to explain word order changes that have been occurring in the language. If all we knew was the grammar of the language, these changes would be a mystery.

Notes

- * The research for this paper was supported by Social Sciences and Humanities Research Council of Canada Research Grants 410-810949, 410-830354, and 410-850540 and by National Science Foundation Research Grant BNS-9011190. I am indebted to Bill Croft, Doris Payne and Russ Tomlin for comments on earlier drafts of this paper.
- 1. Lambrecht (ms.) offers a definition of pragmatically unmarked word order that is similar: “Given a pair of truth-conditionally equivalent syntactic structures, one of its members may be said to be unmarked pragmatically if it can be used to express two discourse functions while the other member can express only one of them” (p. 10). The examples he discusses are all ones in which the set of contexts in which the marked construction is used is a proper subset of the set of contexts in which the unmarked construction is used. A possible problem with Lambrecht’s approach, however, is that it assumes that unmarked constructions serve multiple discourse functions. Consider the analogy of instances of semantic markedness. While we might want to say that *lion* can be used for either a male of the species or a female, we presumably would not want to say that *lion* has two meanings, one for male and one for female: rather the word is simply vague or unspecified for the sex of the animal. In some cases, like the one illustrated in (2a) and (3), we may want to say that a given construction has two uses, distinguished by intonation. In other cases, however, a pragmatically unmarked construction may be appropriate in a superset of the contexts in which the marked construction is appropriate, without there being two identifiable discourse functions associated with the set of contexts in which the unmarked construction is used.
- 2. I assume in (4) that property X is some property that exactly one of the A and the P can possess.
- 3. In Thompson’s analysis, the principle like (5) is only one of two principles governing use of passive.
- 4. The terms “definite” and “indefinite” are used here as the names for discourse-pragmatic categories rather than morphosyntactic categories. The terms “identifiable” and “non-identifiable” are alternative terms for these categories.
- 5. Ojibwa exhibits a contrast between two types of clauses, direct clauses and inverse clauses, and there are competing theories as to what is subject in inverse clauses. The traditional analysis assumes a notion of subject corresponding to notional subject. Rhodes (1976, 1990) provides an analysis of inverse clauses under which the notional object is the grammatical subject and the notional subject is the grammatical object. All 25 of the clauses with lexical subject and lexical object in the counts in Table 1 were direct rather than inverse clauses, so the problem of how to analyse inverse clauses does not arise. Proulx (1991) cites results of a count of a different set of Ojibwa texts that are similar to the ones in my count to the extent that SVO is most common: SVO – 17, VSO – 7, OVS – 2, and VOS – 1. The most notable difference is that only one VOS clause occurred in the texts he examined. He apparently assumes notions of subject and object corresponding to notional subject and notional object.
- 6. The cases of OSV and SOV may not belong here since the initial element is separated intonationally from the rest of the clause and may better be viewed as outside the clause,

somewhat analogous to left dislocation structures in English. The number of such clauses is relatively small, however, and their inclusion does not affect the logic of the arguments presented here.

7. Givón (1979:58-59) argues that passive sentences are more presuppositional than active sentences, but his sense of presupposition seems broader than Payne's and his arguments seem unconvincing. His arguments depend crucially on the properties of agentless passives. But since the crucial question is whether *agented* passives are pragmatically marked, the properties of agentless passives are irrelevant, despite the fact that they are more frequent than *agented* passives.
8. *Little Turtle's Big Adventure*, by David Harrison (New York: Random House, 1969).

References

- Abbott, Miriam. 1991. "Macushi." *Handbook of Amazonian Languages*, Vol. 3 ed. by Desmond Derbyshire and Geoffrey Pullum, 23-160. Berlin: Mouton de Gruyter.
- Andersen, Henning. 1988. "Markedness theory: The first 150 years." *Markedness in Synchrony and Diachrony* ed. by Olga Mišeska Tomić. Berlin: Mouton de Gruyter.
- Battistella, Edwin L. 1990. *Markedness: The evaluative superstructure of language*. Albany: State University of New York Press.
- Brody, Jill. 1984. "Some problems with the concept of basic word order." *Linguistics* 22.711-36.
- Chafe, Wallace. 1987. "Cognitive constraints on information flow." *Coherence and Grounding in Discourse* ed. by Russell Tomlin. Amsterdam/Philadelphia: John Benjamins.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Croft, William. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
- Dryer, Matthew S. 1989. "Discourse-governed word order and word order typology." *Belgian Journal of Linguistics* 4.69-90.
- Du Bois, John W. 1985. "Competing motivations." *Iconicity in Syntax* ed. by John Haiman, 343-65. Amsterdam: John Benjamins.
- Du Bois, John W. 1987. "The discourse basis of ergativity." *Language* 63.805-55.
- Fox, Barbara. 1987. "The noun phrase accessibility hierarchy revisited." *Language* 63.856-70.
- Fox, Barbara, and Sandra Thompson. 1990. "A discourse explanation of the grammar of relative clauses in English conversation." *Language* 66.297-316.
- Givón, T. 1979. *On Understanding Grammar*. New York: Academic Press.
- Givón, T. 1990. *Syntax: A Functional-Typological Introduction*, Vol. 2. Amsterdam/Philadelphia: John Benjamins.
- Greenberg, Joseph. 1966. *Language Universals (with special reference to feature hierarchies)*. The Hague: Mouton.
- Haiman, John. 1983. "Iconic and economic motivation." *Language* 59.781-819.
- Lambrecht, Knud. "Information structure and sentence form: the pragmatics of syntax in spoken French." Unpublished ms.

- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Mithun, Marianne. 1990. "Third person reference and the function of pronouns in Central Pomo natural speech." *International Journal of American Linguistics* 56.361-76.
- Nichols, John. 1988. *An Ojibwe Text Anthology*. London, Ontario: The Centre for Research and Teaching of Canadian Native Languages, University of Western Ontario.
- Payne, Doris L. 1987. "Information structuring in Papago narrative discourse." *Language* 63.783-804.
- Payne, Doris L. 1990. *The Pragmatics of Word Order: Typological dimensions of verb initial languages*. Berlin: Mouton de Gruyter.
- Payne, Doris L., and Thomas E. Payne. 1990. "Yagua." *Handbook of Amazonian Languages*, Vol 2 ed. by Desmond C. Derbyshire and Geoffrey K. Pullum, 249-474. Berlin: Mouton de Gruyter.
- 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.
- Rhodes, Richard A. 1976. "The Morphosyntax of the Central Ojibwa Verb." University of Michigan dissertation.
- Rhodes, Richard A. 1990. "Obviation, inversion and topic rank in Ojibwa." *Berkeley Linguistics Society 16, Special Session on General Topics in American Indian Linguistics*, 101-15.
- Thompson, Sandra A. 1987. "The passive in English: a discourse perspective." *In Honor of Ilse Lehiste* ed. by Robert Channon and Linda Schockey. Dordrecht: Foris.
- Thompson, Sandra A. 1988. "A discourse approach to the cross-linguistic category 'adjective'." *Explaining Language Universals* ed. by John A. Hawkins, 167-185. Oxford: Blackwell.
- Tomlin, Russell S. 1983. "On the interaction of subject, theme, and agent in English." *Journal of Pragmatics* 7.411-432. Reprinted 1985 in *Beyond the Sentence: Discourse and Sentential Form* ed. by J. Wirth Ann Arbor: Karoma.
- Tomlin, Russell S. This volume. "Focal attention, voice, and word order: an experimental, cross-linguistic study."
- Tomlin, Russell S., and Richard A. Rhodes. 1979. "An introduction to information distribution in Ojibwa." *Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society*, 307-320. Reprinted in, *The Pragmatics of Word Order Flexibility* ed. by Doris Payne. Amsterdam/Philadelphia: John Benjamins.
- Trubetzkoy, Nikolaj Sergeevitsch. 1969. *Principles of Phonology*. Translated by Christiane A. M. Baltaxe. Berkeley: University of California Press.
- Utschig, Anne M. 1985. "Main and Subordinate Clause Word Order." Ph.D. Dissertation, Univ. of Minnesota.

WORD ORDER IN DISCOURSE

Edited by

PAMELA DOWNING

MICHAEL NOONAN

University of Wisconsin-Milwaukee

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA