

# On word order and information structure in Yaqui

*Lilián Guerrero*

*Valeria A. Belloro*

## 1. Introduction

This paper explores word order patterns in Yaqui (Uto-Aztecan) as they manifest in a corpus of oral narratives, and advances a preliminary account of the relation of word order and information structure in the language. Word order patterns can be conceived of as the arena where syntactic and pragmatic forces meet. The way in which languages resolve the potential tensions arising from this encounter serves as a typological criterion under which different language-types can be distinguished. According to this criterion, languages can be characterized in terms of whether the order of constituents in main clauses is primarily dependent on syntactic principles (e.g. grammatical relations) or on pragmatic ones (e.g. the (assumed) cognitive status of the referents involved) (Thompson 1978; Payne 1990, 1992). In those languages where word order is constrained by syntactic principles, the encoding of information structure is frequently carried out exclusively by prosodic means, leaving word order intact. These languages can be said to be syntactically rigid (Van Valin 1999). English is one example of this type. On the other hand, word order may be constrained by pragmatic principles. For instance, a language may ban the assignment of focus to preverbal subjects, as Italian (Bentley 2008), or reserve a specific syntactic position for particularly newsworthy information, as Cayuga (Mithun 1992). In Van Valin's typology, languages in which the potential domain for focus assignment excludes certain syntactic positions in the clause are considered pragmatically rigid. Rigidity and flexibility are conceived as tendencies along a continuum, and should not be understood as absolute opposites.

Syntactically rigid languages may be either flexible or rigid pragmatically. English and French represent examples of each type, respectively. In both languages word order is mainly determined by grammatical relations. However, in French, the syntactic rigidity of the SVO order meets a further pragmatic constrain which bans preverbal focus (with some exceptions that do not concern us here). Hence, to the question '*What happened?*' (which triggers a response where the full sentence is in focus) an English speaker may

answer with a canonical SVO structure; for instance ‘*My car broke down*’. A speaker of French, on the contrary, will have to produce an alternative in which both the syntactic constraint (SVO order) and the pragmatic constraint (no preverbal focus) are satisfied, as in ‘*J’ai ma voiture qui est en panne*’ (lit. ‘I have my car that is in breakdown’, Lambrecht 1994:15).

Pragmatically flexible languages may be syntactically rigid (like English), or syntactically flexible (like Russian). Russian is like English in that there is no constraint in the clause with respect to the domain for focus assignment. It is different from English, however, in that its word order is not determined by grammatical roles, so that it can be altered depending on the particular pragmatic status of the participants. Finally, pragmatically rigid languages may be syntactically rigid as well (as French) or syntactically flexible (like Italian). In Italian, unlike in French, the order of grammatical roles *can* be altered to accommodate the existing pragmatic constraints. Consequently, the Italian equivalent to the sentences above would be ‘*Mi si é rotta la macchina*’, a structure which avoids a preverbal focal subject with less dramatic effects on the syntax than the French alternative (Lambrecht 1994; Bentley 2008; Van Valin 1999, 2005).

English, French, Italian and Russian are examples of the four cross-linguistically attested results of the struggle between syntactic and pragmatic constraints over word order. Roughly, in the English type syntax overcomes pragmatics, and in the Italian type it is pragmatics which overcomes syntax. In languages like Russian there are no strong constraints of either type, and order depends on information structure (among other factors). Finally, in the French type the both syntax and pragmatics must be satisfied, potentially resulting in a complex structure. The analysis of the principles governing word order in different languages, thus, offers a privileged view of the interplay between syntax and pragmatics and, more generally, opens a venue for studying motivation in grammar.

With this as general background, the aim of this paper is to examine the interplay between word order and information structure in Yaqui.<sup>1</sup> So far, most of the studies on Yaqui have focused on the syntactic and/or semantic aspects of its grammar, and little has been said about word order and its possible correlation with the expression of information structure. In this regard, the goal of this paper is twofold: (i) to determine whether the basic word order assumed for the language is corroborated by the analysis of spontaneous corpora or whether it should be revised, and (ii) to examine to what extent word order variations are determined by discourse-pragmatic constraints.

Establishing if a language is syntactically flexible or not is, in our opinion, best determined by the analysis of texts. The more rigid the syntax of the language, the easier it should be to determine a “basic” word order based on frequency counts. On the other hand, one of the most straightforward ways to determine whether a language is pragmatically flexible or not is via direct elicitation, since this methodology makes it possible to create unambiguous contexts which trigger clearly distinct focus structures. Thus one may test, for instance, whether the same syntactic structure used to respond to ‘*What happened?*’ may also appear in relation to questions which assign different pragmatic roles to each the participants involved (suggesting pragmatic flexibility) or, instead, if a different constituent order or a different syntactic construction has to be employed (suggesting pragmatic rigidity).

In this paper we have taken both methodological approaches. First, we will present a brief sketch of Yaqui clause structure and its commonly assumed basic word order (Section 2), which will be evaluated with respect to the patterns found in a corpus of oral narratives (Section 3). The corpus data challenges the expected syntactic flexibility of the language with respect to the ordering of nominal arguments, and indicates that word order in Yaqui is strongly determined by grammatical relations (Section 3.1). It is found that, in fact, even so-called “second position” subject clitics frequently appear in the sentence-initial position expected for nominal subjects (Section 3.2). The facts discussed in this section show that the language should be categorized as syntactically rigid (in terms of the typology mentioned above) and, further, as strongly SOV. The most frequent marked construction involves the use of postnuclear subjects or objects. Finally, in Section 4 we draw on direct elicitation to provide a first description of the preferred means for encoding different focus structures in the language, based on Lambrecht’s (1994) characterization of focus-types. The evidence presented here indicates that Yaqui is pragmatically flexible and information structure is mainly encoded prosodically. However, marked syntax can be used to encode some discourse-pragmatic categories; we found evidence of postnuclear subjects and objects occurring in a right-dislocated position and functioning pragmatically as “antitopics” (Lambrecht 2001).

## **2. Basic clause structure**

Yaqui is a synthetic/agglutinative language, with a nominative-accusative case system. Lexical nominatives are morphologically unmarked, and accusatives are marked by the suffix *-ta* (1a). There is no dative case; rather, “in-

direct objects” are marked by directional postpositions (1b). Multiple object constructions (1c) are also very common (Guerrero and Van Valin 2004).<sup>2</sup>

- (1) a. *Amani wa'a rancheo-Ø wakas-ta o kaba'i-ta ta'aru-k.*  
 there DEM farmer-NOM cow-ACC or horse-ACC loose-PFV  
 ‘This farmer lost a cow or a horse over there.’ (Jiak nokpo’s 5: 10)
- b. *U o'ou-Ø jamut-ta-u nooka-k.*  
 DET man-NOM woman-ACC-DIR talk-PFV  
 ‘The man talked to the woman.’
- c. *Aurelia-Ø Ivan-ta mo'obei-ta jinu-ria-k.*  
 Aurelia-NOM Ivan-ACC hat-ACC buy-APPL-PFV  
 ‘Aurelia bought Ivan a hat.’

The examples above also illustrate the canonical arrangement of constituents in the Yaqui sentence, which justifies its categorization as a verb-final language. As most of the languages of the same type, Yaqui employs postpositions (2a) and verbal suffixes (2b), and within a genitive phrase the genitive precedes the head noun (2c).

- (2) a. *Bw'awa-ta betchi'ibo.*  
 soup-ACC for  
 ‘For the soup.’
- b. *Ili uusi-Ø bwaan-taite-k.*  
 little child-NOM cry-begin-PFV  
 ‘The little child started crying.’
- c. *Joan-ta juubi = ne bicha-k.*  
 Joan-ACC wife =1SG.NOM see-PFV  
 ‘I saw John’s wife.’

The examples in (1), (2b) and (2c) also show an arrangement where subjects precede objects, although it is commonly argued that the ordering of subjects and objects with respect to each other is relatively free (Escalante 1990; Rude 1996; Dedrick and Casad 1999: 39; Félix 2000; Guerrero 2006). Both SOV and OSV patterns are judged equally grammatical, the preference of one over the other presumably dependent on pragmatic factors. This word order flexibility is illustrated in (3), from Escalante (1985: 36).

- (3) a. *Juan-Ø Maria-ta bicha-k.*  
 Juan-NOM María-ACC see-PFV  
 ‘John saw Mary.’

- b. *Maria-ta Juan-Ø bicha-k.*  
 María-ACC Juan-NOM see-PFV  
 ‘John saw Mary.’

There are also constraints with respect to the ordering of pronouns, depending on their type. Traditionally, Yaqui pronouns have been divided into independent (“full”) and dependent (“reduced”) forms (Langacker 1977: 124-6 for Uto-Aztecan). “Full” pronouns are expected to behave as lexical elements in terms of their distribution. On the other hand, “reduced” nominative pronouns have been traditionally analyzed as occupying a fixed position in the clause (namely, as “second position” clitics). “Reduced” accusatives (available only for third person) cliticize to the verb. A third set of “reduced” forms occur as objects of postpositions. The relevant paradigms are presented in Table 1.

Table 1. Yaqui pronominal system

	Nominative	Accusative	Object of postposition
1 Sg	<i>inepo =ne</i>	<i>Nee</i>	<i>ne-</i>
2 Sg	<i>empó =’e</i>	<i>enchi</i>	<i>e-</i>
3 Sg	<i>aapo =Ø</i>	<i>apo’ik a=</i>	<i>a-</i>
1 Pl	<i>itepo =te</i>	<i>Itom</i>	<i>ito-</i>
2 Pl	<i>eme’e =’em</i>	<i>enchim</i>	<i>emo-</i>
3 Pl	<i>bempo =mme</i>	<i>apo’im am=</i>	<i>ame-</i>

The possible relation between alternative word orders and the expression of pragmatic categories has received, up to now, limited attention. Dedrick and Casad (1999: 43-45) comment on a structure where the first position in the sentence is occupied by a topicalized element. Characteristically, this constituent is separated from the rest of the clause by a so-called “introducer element” serving as discourse connector, such as *into ~ intok*<sup>3</sup> ‘and (thus)’ (4a) and *bea* ‘then’ (4b).

- (4) a. *Kauwis-Ø intok pocho’o-kun-bicha-u bwite-k.*  
 fox-NOM and woods-to-site-to run-PFV  
 ‘And the fox ran toward the woods.’ (Dedrick and Casad 1999: 43)
- b. *Name’e-Ø bea nee weeya-n reso-po.*  
 These-NOM then 1SG.ACC guide-PASTC pray-LOC  
 ‘And these (women) guided me in the prayers.’ (Maejto: 14)

Dedrick and Casad also mention a structure where the sentence-initial position is occupied by a focal element, like the pronominal *aapo* ‘he’ in (5). It is unclear, however, whether the authors believe that the positions occupied by the pronoun in (5) and the topicalized constituents in (4) should be structurally distinguished.<sup>4</sup>

- (5) *Aapo intok ejea-ta-betuk taawa-k.*  
 3SG.NOM and ironwood:tree-ACC-under stay-PFV  
 ‘And he stayed under the ironwood tree.’  
 (Dedrick and Casad 1999: 44)

To summarize, it is generally agreed upon that Yaqui is strongly verb final and that, typically, subjects precede objects, although SOV and OSV patterns are equally possible. With respect to pronominal arguments, it is commonly held that full forms may exhibit the same ordering as their corresponding nominal counterparts, whereas the order of reduced forms is fixed: accusatives attach to the verb and nominatives occur in second position, cliticizing to whatever element appears first. So far, two constructions where word order correlates with information structure have been described; one where the sentence-initial constituent is interpreted as topical, and another where it is interpreted as focal. For the most part, however, these claims are based exclusively on direct elicitation data. In this regard, the analysis of naturally-occurring texts appears a necessary step in order to examine whether the posited patterns are found in natural, connected discourse and, ideally, what are the functions served by ordering alternatives. With these goals in mind, the following section examines word order patterns in a corpus of oral narratives. The two sub-sections concern nominal arguments and full pronouns (3.1.) and the position of subject clitics (3.2.).

### 3. Word order in Yaqui texts

#### 3.1. Nominal arguments and strong pronouns

The following discussion on word order patterns is based on a corpus formed by the oral narratives listed in Table 2. The sample includes six folktale stories from *Jiak nokpo*’s collection (Silva *et al.* 1998), and three life stories (Silva 2004, n.d.; Félix n.d.), providing a total of nearly 1400 active clauses.<sup>5</sup>

Table 2. Yaqui oral texts

Text	Clauses
<i>Jiak nokpo's</i> (Silva <i>et al.</i> )	337
<i>Hilario</i> (Silva)	478
<i>Maejto</i> (Silva)	154
<i>CB</i> (Félix)	432
	1401

As mentioned above, it is traditionally assumed that the order of subject and object with respect to each other is relatively free when these arguments are realized by nouns. In order to examine whether this claim is consistent with natural discourse data, all transitive clauses with nominal Ss and Os were selected from the corpus; full pronouns and demonstrative pronouns (cf. (7)) were also included since they show a similar behavior within the clause. The relative frequency of each attested order is presented in Table 3.

Table 3. Word order in transitive clauses

	SOV	OSV	OVS	SVO
<i>Jiak nokpo's</i>	49	-	4	9
<i>Hilario</i>	21	-	5	3
<i>Maejto</i>	13	-	2	-
<i>CB</i>	18	1	1	4
	101 (78%)	1 (1%)	12 (9%)	16 (12%)
	102 (79 %)		28 (21 %)	

As expected, both arguments usually precede the verb (79%). Note, however, that there is only one instance of OSV. This suggests that this order, if not fully ungrammatical, is highly marked. The examples in (6) illustrate the most frequent SOV pattern, where S is a lexical noun (6a) or a full pronoun (6b); (7) presents the only attested OSV example.

- (6) a. *U kosineo-Ø itom tea-ka.*  
 DET cook-NOM 1PL.ACC find-PFV  
 'The cook found us.' (CB: 118)
- b. *Itepo into i-me'e yoeme-m aman bittua-kan.*  
 3PL.NOM and DEM-PL man-PL there send-PASTC  
 'We sent those men there.' (Hilario: 69)
- (7) *Bea ju-ka in jaboi-Ø ta-ta'a-k.*  
 then DET-ACC 1SG.GEN grandfather-NOM RED-know-PFV  
 'Then, my grandfather met that one.' (CB: 17)

The data in Table 3 also show that, after sentence initial, the next most frequent position for subjects is in fact postverbal (although no full pronouns nor demonstratives were found in this position). All postverbal subjects in the corpus are marked as definite by a determiner that is optional with preverbal subjects.<sup>6</sup> Some instances of OVS are presented in (8).

- (8) a. *Kia itom aman jika-u go'ota-∅ juna-me'e peron-im.*  
 only 1PL.ACC there up-DIR throw-PRES DET-PL pelón-PL  
 'They only throw us up (onto the train), those *pelones*.' (Hilario: 39)
- b. *Junaman itom to'o-bwite-k juna'a Mayor-∅*  
 there 1PL.ACC leave-run-PFV DEM Mayor-NOM  
 'He abandoned us there, the Mayor.' (Hilario: 185)

The third most common word order of nominal arguments found in the corpus is SVO. These postverbal objects exhibit the same morphosyntactic constraints as postverbal subjects: they exclude full pronouns and demonstratives, and the nominal form is obligatorily marked by a definite determiner. Further, postverbal objects must co-occur with a pronominal element occupying the canonical object position. The SVO pattern is illustrated with the examples in (9).

- (9) a. *junum bea inepo a<sub>i</sub> ansu-k u-ka eskuela-ta<sub>i</sub>.*  
 like.this and 1SG.NOM 3SG.ACC finish-PFV DET-ACC school-ACC  
 'I finished the school like this.' (CB: 346)
- b. *chukula itepo mura-m-met am<sub>i</sub> pu'akta-k*  
 later 1PL.NOM mule-PL-LOC 3PL.ACC load-PFV  
*ju-me'e koko-ka-me ito-betana<sub>i</sub>.*  
 DET-PL die.PL-PFV-REL 1PL-side  
 'Later on, we loaded our deaths in the mules.' (Hilario: 155)

With respect to intransitive clauses, subjects also occur most commonly in preverbal position (Table 4).

Table 4. Word order in intransitive clauses

	SV	VS
<i>Jiak nokpo's</i>	60	17
<i>Hilario</i>	28	7
<i>Maejto</i>	9	-
<i>CB</i>	59	3
	156 (85%)	27 (15%)

This preference is independent of the verb type, occurring for instance with both unergative (10a) and unaccusative (10b) verbs.

- (10) a. *Ju-me'e Caantora-m bwi-bwika-o.*  
 DET-PL singer-PL RED-sing-when  
 'When the singers sang.' (Maejto: 26).  
 b. *Junama in yoyo'owa-m jo'ak.*  
 there 1SG.GEN RED-parent-PL live.PFV  
 'My parents lived there.' (CB: 8)

Postverbal intransitive subjects obey the same constraints that apply to transitive clauses with respect to definiteness. An example of VS is presented in (11).

- (11) *Taawa-k ju'u boobok bo'o-jo-reo.*  
 stay-PFV DET toad road-make-AGT  
 'The wandering toad stayed.' (Jiak nokpo's 3: 29)

In sum, the frequencies from the narrative data question the previously assumed flexibility of preverbal nominal arguments and suggest that, instead, the order of nominal arguments is strongly determined by grammatical relations. Alternatives to the most frequent SOV order are relatively infrequent, and most commonly involve postverbal nominative and accusative phrases, which must be marked as definite (and in the case of accusatives, co-occur with a pronominal counterpart). The next section analyzes the ordering restriction of "reduced" forms, in particular so-called "second position subject clitics".

### 3.2. Subject clitics

We mentioned in Section 2 that it is commonly argued that Yaqui exhibits a widespread trait among Uto-Aztecan languages, namely the existence of so-called "second position clitics". The occurrence of these forms was first attested by an 18th century grammarian, who in describing the language "states as one of his most infallible rules that the pronominal subject [read clitic pronoun-Steele] *must* be the second word or element in the sentence..." (Steele 1977:553, citing from Mason 1923, emphasis in the original). This constraint is clearly observed in Tarahumara, a language closely related to Yaqui, belonging also to the Taracahita branch. In Tarahumara, second posi-

tion is defined in strictly linear terms, as illustrated in (12), where the subject clitic occurs after the first word of the initial constituent.

- (12) *Semati ne napaha rarimea aré.*  
 Nice 1SG shirt gonna:buy probably  
 ‘I am probably going to buy a nice shirt’ (Llaguno 1970: 29)

Although the pervasiveness of “second position clitics” in Yaqui is widely accepted, the literature is not clear on the criteria under which this second position should be determined. It has been claimed that a subject clitic occurs “following the first word” in the sentence (Dedrick and Casad 1999: 242), following the first “element” (Steele 1977:541), or “following the first constituent, whatever that may be” (Escalante 1990: 48). It is therefore ambiguous whether in Yaqui second position should be understood, as in Tarahumara, in strictly linear terms or in syntactic ones (i.e., after the first phrasal constituent). The examples provided in the grammars suggest that the second interpretation allows for a more consistent analysis (13).

- (13) a. *siika= ne*  
 go = 1SG.NOM  
 ‘I am going now.’ (Dedrick and Casad 1999: 242)  
 b. *Tutuka-biako-o = ne koche=su woo’o*  
 Yesterday time ago-when = 1SG.NOM sleep=EMP mosquito  
*nee ke’e-ka.*  
 1SG.ACC bite-PFV  
 ‘Last night while I was sleeping, a mosquito bite me.’  
 (Dedrick and Casad 1999: 48)

Be it as it may, the analysis of our corpus shows that either criterion (first word or first phrase) is far from uncovering an absolute pattern. In fact, subject clitics seem to exhibit a much greater flexibility in their ordering possibilities that was previously held. Table 5 illustrates the distribution of subject clitics in Yaqui transitive clauses (X stands for any non-subject phrasal constituent, either an object or an adjunct).

Table 5. Transitive clauses with “subject clitics”

	XS <sub>cl</sub> V	XXS <sub>cl</sub> V	S <sub>cl</sub> XV
<i>Jiak nokpo’s</i>	8	3	7
<i>Hilario</i>	22	12	19
<i>maejto</i>	21	5	13
<i>CB</i>	15	14	22
	66 (41%)	34 (21%)	61 (38%)

In terms of ordering possibilities, we did not find any instance where the reduced form occupied a “second position” cliticized to a verb (cf. (13a)). There were also no instances in the corpus where “second position” can be unambiguously defined, as in Tarahumara, in strictly linear terms. Whereas the first constituent in (14a) and (14b) can be analyzed either as a word or as a phrase, the last two examples are clear instances of phrasal constituents.

- (14) a. *Ka=ne ame beas kop-kopte-Ø*  
 NEG = 1SG.NOM 3PL.ACC about RED-forget-PRES  
 ‘I don’t forget them.’ (CB: 368)
- b. *Aman=te kari-ta am ya’a-ria-k*  
 there = 1PL.NOM house-ACC 3PL.ACC do-APPL-PFV  
 ‘There, we built them a house.’ (Hilario: 13)
- c. *Bea sejtul ta’a-po ket=ne a-u waate*  
 Then one day-LOC too=1SG.NOM 3SG-DIR remember  
 ‘I also remember one day.’ (CB: 108)
- d. *Tajkai huevona-m ke=ne bwa’e into jaiba-m*  
 tortilla huevona-PL no.yet=1SG.NOM eat and crab-PL  
 ‘I haven’t eaten yet ‘tortillas huevonas’ and crab.’ (Hilario: 260)

Even if “second position” is defined in terms of syntactic constituents, note that the assumed tendency is also belied by instances where reduced subject pronominals also appear following *more* than one phrasal element, as in (15).

- (15) a. *Amak goki-m-mea amak bisikleeta-t bea*  
 sometimes foot-PL-INSTR sometimes bike-LOC then  
 = *ne wee-ne*  
 1SG.NOM go-POT  
 ‘Sometimes I go by walking, sometimes by bike.’ (CB: 192)
- b. *Aman beseo-m=nee jajase-o.*<sup>7</sup>  
 there calf-PL=1SG:ACC track-when  
 ‘When I track the calf over there.’ (Hilario: 2)
- c. *Kaa bwe’ura Lio-nok-ta reso-ta jiba=ne*  
 NEG big God-word-ACC pray-ACC only=1SG.NOM  
*ta’a-n*  
 know-PASTC  
 ‘But I didn’t know all the “God words”, only the prayers.’ (Maejto: 5)
- d. *Junak bea batte cien metro-po bat=te kaate*  
 Later then almost 100 metro-LOC ahead=1PL.NOM go.PFV

‘And later, we were ahead by almost one hundred meters.’  
(Hilario: 138).

An anonymous reviewer suggested that some of these examples could be analyzed as small clauses functioning as matrix objects, e.g., *the calf over there* in (15b), in which case they could be consistent with the “second position” rule, if the small clause is considered to occupy the first position. Still, the relevance of “second position” with respect to this phenomenon is questioned by the fact that it can be defined with respect to a word, a phrase, or a small clause (under such analysis), and also by the instances where the pronominal follows more than one phrase.

Further, note that Table 5 shows that the “reduced” subjects appear sentence-initially almost as frequently as they occur in this putative “second position”. Some examples are presented in (16).

- (16) a. *Ne kaa into nappat a tekipanoa-Ø ian tajti.*  
1SG.NOM NEG and either 3SG.ACC work-PRES now until  
‘Up to now, I cannot work either.’ (Maejto: 98)
- b. *Ne in maala jariu-bae.*  
1SG.NOM 1SG.GEN mother look-DESID  
‘I want to look for my mother.’ (Hilario: 221)
- c. *Ne soda-ta je'e-ka merkao-wi.*  
1SG.NOM soda-ACC drink-PFV market-DIR  
‘I drank a soda in the market.’ (Hilario: 165)

In sum, the analysis of the corpus data presents some interesting puzzles for the common view of reduced nominative pronominals in Yaqui as “second position clitics”. That this view should be refined is particularly clear in light of examples such as those in (16), where there is no preceding material to which the pronoun may cliticize, and cases both as those in (15) and (16) where the pronoun appears out of its expected second position. These data suggest that more detailed studies of these forms should be pursued, which could further uncover the motivations behind their alternative orderings.

The findings about subject clitics presented in this section challenge current assumptions and justify the need for further studies related to the phenomena of cliticization in Yaqui, including the detailed analysis of the factors behind the occurrence of subject clitics in one position or the other and the presence of multiple pronouns. Leaving these issues for future endeavors, in the next section we go back to Van Valin’s (1999) typology on the interac-

tion between syntax and pragmatics in the determination of word order patterns. The analysis of corpus frequencies suggested that in Yaqui word order of major constituents is mainly determined by grammatical relations, justifying the categorization of the language as syntactically rigid. Now we examine how main information structure types are encoded in the clause, and how the language should be characterized pragmatically.

#### 4. Information structure in Yaqui

We saw that the corpus data strongly indicates that clausal constituents follow a basic SOV order. It is not surprising that this order is also the one used to encode predicate focus, since this structure also constitutes the unmarked focus type (Lambrecht 1994). A predicate focus structure in Yaqui is illustrated in (17). In this structure, the subject (typically a pronoun) is the topic, whereas the object and the verb constitute the focus (marked with small caps).<sup>8</sup>

- (17) a. *Jita-sa u jamut-Ø ya'a-Ø?*  
 What-Q DET woman-NOM do-PRES  
 'What is the woman doing?'  
 b. *U jamut-Ø /aapo KAFE-TA BWASA-Ø.*  
 DET woman-NOM 3SG.NOM coffee-ACC toast-PRES  
 'The woman / she is toasting the coffee.'

In Yaqui there is no constraint over focal preverbal subjects. Just as in English and other pragmatically flexible languages, the same basic word order is used in sentence-focus constructions, as illustrated in (18):

- (18) a. *Jita-sa weye-k?*  
 What-Q go.SG-PFV  
 'What happened?'  
 b. *U JAMUT-Ø TOTO'I-M JINU-K.*  
 DET woman-NOM hen-PL buy-PFV  
 'The woman bought the hens.'

The canonical SOV order is also used for narrow focus constructions, regardless of whether the focus targets the object (19) or the subject (20).

- (19) Q: -‘What did the woman buy?’  
 A: -U *jamut-∅ TOTO’I-M jinu-k.*  
 DET woman-NOM hen-PL buy-PFV  
 ‘The woman bought hens.’
- (20) Q: -‘Who bought the hens?’  
 A: -U *JAMUT-∅ toto’i-m jinu-k.*  
 DET woman-NOM hen-PL buy-PFV  
 ‘The woman bought the hens.’

On the other hand, our informants consistently considered OSV order ungrammatical, regardless of the focus structure assigned to it. This is illustrated in (21). Recall that we only found one such instance in the corpus (cf. Table 4). Their low frequency and the judgment of native speakers about their acceptability indicate that, if not fully ungrammatical for all speakers, OSV order is highly dispreferred.

- (21) Q: -‘Who bought the hens?’ / ‘What did the woman buy?’ / ‘What happened?’  
 A: \* *Toto’i-m u jamut-∅ jinu-k.*  
 hen-PL DET woman-NOM buy-PFV  
 ‘The woman bought the hens.’

The fact that the canonical SOV appears consistently in response to different focus structures indicates that, like English, Yaqui is pragmatically flexible, in as much as there does not seem to be any constraints in terms of focus assignment to any syntactic position within the clause. Speakers offer clauses with the canonical SOV order regardless of what constituent is in focus, and different interpretations are achieved by differences in the prosodic prominence associated with the different constituents.

We saw, however, that in connected discourse we found a marked pattern where a lexical argument, either S or O, occurred postverbally. As mentioned above, all postverbal arguments must be marked as definite and, in the case of accusatives, the canonical object position must be occupied by a pronominal form.

These postverbal constituents occupy a right-dislocated position and, consequently, cannot be focal. The examples in (22) illustrate the infelicity of

VS, OVS and SVO orders with a focal interpretation of the postverbal constituent (pragmatically unacceptable sentences are marked #).

(22) Q: 'Who danced?'/ 'What happened?'

- a. # *Ye'e-k*, U *QUETA-Ø*  
 dance-PFV DET Queta-NOM  
 'Queta danced.'

Q: Who bought the hens?

- b. # *U-me toto'i-m jinu-k*, U *JAMUT-Ø*  
 DET-PL hen-PL buy-PFV DET woman-NOM  
 'She bought the hens, the woman.'

Q: What did the woman buy?

- c. # *U jamut-Ø* *am<sub>i</sub> = jinu-k* *U-ME TOTO'I-M<sub>i</sub>*  
 DET woman-NOM 3PL.ACC = buy-PFV DET-PL hen-PL  
 'The woman bought them, the hens.'

The right-dislocated position in Yaqui is associated with the pragmatic function of "antitopic" (Lambrecht 2001). Note that, as is the case for most languages (cf. Lambrecht 2001 and references cited therein) the data suggests that Yaqui antitopics not only must be definite, and hence identifiable, but further, the source of the identifiability of the referent must be in the previous context. In all cases in the corpus the postverbal element refers to a "discourse-old" element (Prince 1992). We could not find any instances where the postposed element referred to an entity not previously introduced in the discourse context, even in the cases where these constituents denoted elements or characters which could be considered identifiable because of their widely known status in the community. The antecedent for the postverbal phrase is typically mentioned a few clauses prior in the context, suggesting that it is not maximally "active" (Chafe 1994) when the antitopic construction is used.

Below we illustrate some of the contexts where the examples of postnuclear arguments mentioned in Section 3 (repeated below) occur. The relevant sentences illustrate OVS (23), SVO (24) and VS (25) orders, respectively.

- (23) Context: *Abion ito-t siika, tettebote-sime. Bweta int-uchi notte-ka bea, ta'abwi=si itom tebo-tua-k. Metrayadora-e itom mumuki-su-k... Yokoria-po beja te aman bwij-wa-k. Nau konta-wa-k. Ju-me'e*

*Aguilenyo-m tea-me ama rejte into ju-me'e peron-im.*<sup>9</sup> *Junak bea te Pitaya-u te kom saja-k. Pitaya-u lula bea te weiya-wa. Junaman riel bo'o-u te toji-wa-ka, ama ja'abwa-wa-k, junaman bea te desarmar-oa-wa-k wiko'e-ka-me. Bweta juna-me'e into kaa parke-k. ¿Jita-e-sa nassua-ne kia bea wiko'e-ka? Ama rejte. Nakbea te vagone-u kima'a-wa-k.*

'And an airplane flew above us, greeting us. But then it returned and greeted us very differently. It shot us with a machine gun... The next day, we were captured. We were surrounded. The so-called *Aguileños* and **the Pelones** were there. Then, we went down to Las Pitayas; we were taken directly to *Las Pitayas*. And there, after we were taken to the railways, we were lined up, and those having rifles were disarmed. But those [rifles] didn't have bullets. How could we fight, just with the rifles? There we were. And later, we were put into a wagon.' (Hilario: 28-38)

*Kiaitom aman jika-u go'ota-Ø juna-me'e peron-im.*  
only 1PL.ACC there up-DIR throw-PRESDET-PL pelón-PL  
'They only throw us up (onto the train), those *pelones*.' (Hilario: 39)

- (24) Context: *Junak bea u jamut beeme inen au jüia-k*: "Nolia in o'ola ji'ibwa e; empo jumaa si tebaure kia ne kaa a'abo weye-n". *U yoeme into au nattemae-k*: "¿Jaisaakai?". "Bwe'ituk ne si obisaane-n". *U yoeme kia "mmm-po" taawa-k, kaachini into au jüia-k. Junuen luula yesjte-se-kan, aapo into bachi-ta chiu-taite-k.*

'Then, the young woman said: "C'mon, old man, eat; perhaps you are hungry and I wasn't coming [to bring you food]". And then the man asked: "Why?". "Because I was very busy". And the man just said "mmmhm", and said no more to her. At noon, he began to clean the corn.' (Jiak nokpo's 2:18-25)

*Jiba a<sub>i</sub> bitchu-k u-ka'a jamut-ta<sub>i</sub>.*  
always 3SG.ACC watch-PFV DET-ACC woman-ACC  
'He observed her all the time, the woman.' (Jiak nokpo's 2: 26)

- (25) Context: *Ju'u boobok juchi notte-k. Ju'u boobok bo'o-jo-reo-ta-u en jia-k: tu'i achai o'ola mabet-wa-ka, ta inim koobi'iku jiba yumjo'e-ne. Bwe tu'i chuba-la ne yumjoe-bae. Yooko ne ket jakko sim-bae, junuen jia-ka.*

'The toad came back. He told this to the wandering toad: "it is ok old man, you were accepted, but you must stay in this corner only". Ok then, I will rest for a little while. I will leave tomorrow, he said.'

(Jiak nokpo's 3: 24-28)

*Taawa-k ju'u boobok bo'o-jo-reo.*  
 stay-PFV DET toad road-make-AGT

'The wandering toad stayed.'

(Jiak nokpo's 3: 29)

With respect to the coding of arguments, the expected tendency was confirmed, where "reduced" pronominals encode continuous topics, full pronouns typically carry a contrastive meaning and nominals are the preferred means for introducing new participants. This is illustrated with the following passage from one of the folktales, where two turtles are discussing how to beat a coyote in a racing contest. Note the use of the full pronominals *empo* and *inepo* in (26b), (26c) and (26d) for contrasting the tasks assigned to speaker and addressee. In (27) the sequences of reduced forms after a full pronoun are used to encode maximally continuous topics.

- (26) a. *Ika'a = te ya'a-ne=ti bea a-u jiia:*  
 that.ACC = 1PL.NOM do-POT CLM then 3SG-DIR say  
 'We will do this, (the turtle) says to her.'
- b. *Empo bea bo'o jinko'ola naate'e-po yej-ne*  
 2SG.NOM then road competition begin-LOC stand-POT  
 'You will be where the racetrack begins'
- c. *Inepo into gojo'oria-po bo'o jinko'ola-ta*  
 1SG.NOM and hole-LOC road competition-ACC  
*lu'ute'e-po yej-ne.*  
 end-LOC stand-POT  
 'And I will stand in the hole where the racetrack ends.'
- d. *Into bea empo bwia-ta betuk emo*  
 and then 2SG.NOM land-ACC under REFL  
*bwiti-bae = ti jiu-ne.*  
 run.SG-DESID = CLM say-POT  
 'And then you will say (you) want to run underground.'
- e. *Apo'ik bwiiti-taite-k-o*  
 3SG.ACC run-start-PFV-when  
 'When he starts running.'
- f. *Empo junama'a ji'ibweji-taiti-ne,*  
 2SG.NOM there scratch-start-POT  
 'You will start scratching.'

- (27) a. *Inepo*      *1982-po = ne*      *muji-tua-wa-k*  
 1SG.NOM      1982-LOC = 1SG.NOM      consecrate-CAUSE-PASS-PFV  
 ‘I was consecrated in 1982.’
- b. *junu-ka’a*      *cargo-ta = ne*      *mak-wa-k.*  
 DET.ACC      cargo-ACC = 1SG.NOM      give-PASS-PFV  
 ‘I was given that position.’
- c. *Ne*      *kia*      *tukaria-t*      *naabujti*      *ino*      *majta-k.*  
 1SG.NOM      also      night-LOC      in.addition      1SG.REFL      learn-PFV  
 ‘In addition, I also prepared myself at nights.’ (Maejto: 103-105)

The evidence presented in this section indicates that Yaqui is pragmatically flexible and focus can be assigned to any position in the clause. However, this preliminary look at corpus data suggests that right-dislocated structures are used to encode relatively topical subjects and objects if they are not assumed to be maximally active in the mind of the interlocutor. Maximally active referent (i.e. continuous topics), on the other hand, are typically encoding with reduced pronominal forms.

## 5. Final comments

The aim of this paper has been to provide a first pass in the study of word order and information structure in the Yaqui language. It was shown that a basic SOV order can be determined with respect to frequencies in a corpus of oral narratives. The data called into question, however, the assumed flexibility of the language with respect to the ordering of nominal arguments, showing a strong tendency for placing subjects sentence-initially. We presented data which shows that Yaqui speakers make use of a right-dislocated position for placing antitopics. We did not find, however, clear arguments for distinguishing a left-dislocated structure used for topicalized participants. In terms of the typology mentioned in the introduction, we presented evidence indicating that the language is syntactically rigid but pragmatically flexible. There are however, many new questions waiting for investigation. We hope that this preliminary exploration can serve as a basis for future studies, and that through this we can begin to incorporate a discourse-pragmatic perspective into the study of Yaqui and other Uto-Aztecan languages.

## Notes

1. The Yaqui language is spoken mainly in Mexico, by more than 15,000 people living along the Yaqui River in the Central West part of the State of Sonora. Across the U.S.-Mexican border, in Pascua, Arizona, there are an estimated 1,000 speakers, who migrated there at the beginning of the 20th century. There are several grammatical studies on Yaqui; among the most significant are Crumrine (1961), Johnson (1962), Lindenfeld (1973), Escalante (1990), Jelinek and Escalante (2000), and Demers, Escalante and Jelinek's (1999) article on the principles that account for prosodic prominence in isolated words, all based on the Arizona dialect. Dedrick and Casad (1999), Félix (2000), Hernández (2002), Guerrero and Van Valin (2004), Silva (2004), Guerrero (2004a,b, 2006), Martínez (2007), as well as several articles in Estrada *et al.* (2008) and Estrada *et al.* (2007) are all on the Sonoran dialect.
2. Abbreviations; ACC: accusative, AGT: agentive, APPL: applicative, CLM: clause linkage marker, COM: commitative, DESID : desiderative, DET: determiner, DIR: directional, FUT: future; GEN: genitive, LOC: locative, NEG: negation, NOM: nominative, NMLZ: nominalizer, PASTC: past continuative, POT: potential, PPIO: participle, PFV: perfective, PASS: passive, PL: plural, PRES: present, PURP: purpose, Q: question particle, RED: reduplication, REFL: reflexive, SG: singular, TNS: unmarked tense.
3. The distribution and functions of *into(k)* within coordinated units have been explored by Martínez (2007). The author claims this particle cannot be considered a second position clitic (2007: 114) but nothing is said about its pragmatic functions. The extent to which *into(k)* and the other discourse connectors are used to set off the topicalized or focused material is still an open question.
4. In his study of Uto-Aztecan languages, Langacker (1977:25ff) recognizes as a common trait of this language family the displacement to sentence-initial position as a way to mark emphasis, although again this interpretation is linked to both topical and focal readings. We did not find clear evidence for these structures in our data and therefore they will not be discussed in this paper.
5. Non-predicative clauses and matrix clauses with speech act predicates were excluded from this analysis. We owe our gratitude to Crescencio Buitimea, Carlos Silva and Rolando Félix for letting us use their oral narratives.
6. In Yaqui, definiteness is associated with the presence of determiners. Determiners are marked by *-ta* when modifying an accusative NP, but with *-e* if the NP is marked by a postposition; if the N is plural, then the determiner is likewise plural. With preverbal subjects, the presence of the determiner is optional.
7. In adverbial and complement clauses, pronominal subjects are accusative and nominals are marked by *-ta* (Guerrero 2006). Except for this morphological coding, the distribution of major constituents in main and dependent clauses is the same.
8. The analysis in this section follows Lambrecht's (1994) characterization of focus types as adopted within the framework of Role and Reference Grammar.

For details, we refer the reader to Van Valin and LaPolla (1997:199-241) and Van Valin (2005:68-88). Some studies dealing with information structure within this model include Bentley (2008) for Italian and Sicilian; Shimojo (1995, in press) for Japanese; Belloro (2007, in press) for Spanish.

9. There is a prior mention of this participant, which was not included in the selected context for simplicity.

## References

- Belloro, Valeria  
 2007 The syntax-pragmatic interface. A study of Spanish clitic doubling. Ph.D. diss., University at Buffalo, The State University of New York.
- Belloro, Valeria  
 in press Spanish datives: remarks on the information-structure side of the story. In *Studies in Role and Reference Grammar*. Lilián Guerrero, Sergio Ibáñez, and Valeria A. Belloro (eds.). México: UNAM.
- Bentley, Delia  
 2008 The interplay of focus structure and syntax: evidence from two sister languages. In *Investigations of the Syntax-Semantics-Pragmatics Interface*. Robert D. Van Valin, Jr. (ed.). Amsterdam: John Benjamins.
- Barreras, Aguilar  
 1998 Orden de palabras básico en el guarijío de Sonora. *Memorias del IV Encuentro de Lingüística en el Noroeste*. A. Acosta, Z. Estrada Fernández, M. Figueroa, and G. López Cruz (eds). Universidad de Sonora, Hermosillo, 37-68.
- Buitimea, Cresencio  
 2007 *Preesio betana nottiwame*. Hermosillo: Universidad de Sonora.
- Cumrine, Lynne S.  
 1961 *The phonology of Arizona Yaqui, with texts*. Tucson: The University of Arizona Press.
- Chafe, Wallace  
 1994 *Discourse, Consciousness and Time*. Chicago: University of Chicago Press.
- Dedrick, John and Eugene H. Casad  
 1999 *Sonora Yaqui Language Structure*. Tucson: The University of Arizona Press.
- Demers, Richard, Fernando Escalante and Eloise Jelinek.  
 1999 Prominence in Yaqui words. *International Journal of American Linguistics* 65 (1):40-55.
- Escalante, Fernando  
 1990 Voice and Argument Structure in Yaqui. Ph.D. diss., The University of Arizona.

- Escalante, Fernando  
1985 A preliminary view of the structure of Yaqui. M.A thesis., The University of Arizona.
- Estrada, Zarina  
1985 El orden de palabras en pima bajo. Paper presented at the *Friends of Uto-Aztecan Working Conference*, Tucson, Az.
- Estrada, Zarina and Lilián Guerrero  
2007 Grammatical borrowing in Yaqui. In *Grammatical Borrowing in Cross-Linguistic Perspective*. Matras Yaron and Jeanette Sakel (eds.). Berlin: Mouton de Gruyter.
- Estrada, Zarina, Albert Álvarez, Lilián Guerrero and María Belem Carpio  
2007 *Mecanismos de voz y formación de palabra* [Mechanisms of voice and word formation]. México: Editorial Plaza y Valdes.
- Estrada, Zarina, Søren Wichmann, Claudine Chamoreau, and Albert Álvarez  
2008 *Studies in voice and transitivity (Estudios de voz y transitividad)*. Munich: Lincom Europa.
- Félix, Rolando  
2000 Las relaciones gramaticales en yaqui: un análisis en el marco de la Gramática del Rol y Referencia [Grammatical relations in Yaqui: a Role and Reference Grammar analysis]. M.A. thesis, Universidad de Sonora.
- Félix, Rolando  
n.d. Life story of CB (oral text). Manuscript.
- Guerrero, Lilián  
2004a The syntax-semantic interface in Yaqui complex sentences, a Role and Reference Grammar analysis. Ph. D. diss., University at Buffalo, The State University of New York.
- Guerrero, Lilián  
2004b Verbos de movimiento y posición en yaqui [Motion and position verbs in Yaqui]. In *Estudios en lenguas amerindias. Homenaje a Ken Hale*. Zarina Estrada, Ana Fernández, and Albert Álvarez (eds.), 199-228. Hermosillo: Universidad de Sonora.
- Guerrero, Lilián  
2006 *The Structure and Function on Yaqui Complementation*. (Studies in Native American Linguistics 54). Munich: Lincom.
- Guerrero, Lilián and Robert D. Van Valin, Jr.  
2004 Yaqui and the analysis of primary object languages. *International Journal of American Linguistics* 70: 290-319.
- Hernández Doode, Gabriela  
2002 Marcadores discursivos: una exploración en la narrativa yaqui [Discourse markers: A Yaqui narrative exploration]. M.A. Thesis., Universidad de Sonora.

- Jelinek, Eloise and Fernando Escalante  
2000 Unaccusative and unergative verbs in Yaqui. *Uto-Aztecan: structural, temporal and geographic perspectives*. Eugene H. Casad and Thomas L. Willet (eds). México: Universidad de Sonora.
- Johnson, Jean B.  
1962 *El idioma yaqui* [The Yaqui language]. México: Instituto Nacional de Antropología e Historia.
- Lambrecht, Knud  
1994 *Information Structure and Sentence Form. A Theory of Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lambrecht, Knud  
2001 Dislocation. In *Language Typology and Language Universals*. Martin Haspelmath (ed.), 1050-1078. Berlin: Mouton de Gruyter.
- Langacker, Ronald W.  
1977 *Studies in Uto-Aztecan Grammar. Vol. 1 An overview of Uto-Aztecan Grammar*. Texas: Summer Institute of Linguistics and University of Texas at Arlington.
- Lindenfeld, Jelinek  
1973 *Yaqui Syntax*. University of California Press.
- Llaguno, Jose  
1970 El tarahumara sin esfuerzo (in collaboration with D. Brambila and E. Uranga). Manuscript.
- Mason, J. Alden  
1923 *A preliminary sketch of the Yaqui language*. UCPAAE 20: 195-212.
- Martínez, Constantino  
2007 *Yaqui coordination*. (Studies in Native America Linguistics 59). Munich: Lincom.
- Mithun, Marianne  
1992 Is basic word order universal? In *Pragmatics of Word Order Flexibility*. Doris Payne (ed.), 15-61. Amsterdam and Philadelphia: John Benjamins.
- Payne, Doris  
1987 Information structuring in Papago Narrative Discourse. *Language* 63: 783-804.
- Payne, Doris  
1990 *The Pragmatics of Word Order: Typological Dimensions of Verb Initial Languages*. Berlin: Mouton de Gruyter.
- Payne, Doris  
1992 Introduction. In *Pragmatics of Word Order Flexibility*. Doris Payne (ed.), 1-13. Amsterdam and Philadelphia: John Benjamins.

- Prince, Ellen  
1992 The ZPG letter: Subjects, Definiteness, and Information Status. In *Discourse Description: Diverse Analysis of a Fund Raising Text*. Amsterdam and Philadelphia: John Benjamins.
- Rude, Noel  
1996 Objetos dobles y relaciones gramaticales: el caso del yaqui [Double objects and grammatical relations: the case of Yaqui]. In *Memorias del Tercer Encuentro de Lingüística en el Noroeste*. Vol. 2. Zarina Estrada Fernández, Max Figueroa Esteva, Gerardo López Cruz (eds.), 491-522. Hermosillo: Editorial Unison.
- Shimojo, Mitsuaki  
1995 Focus Structure and Morphosyntax in Japanese: *wa* and *ga*, and Word Order Flexibility. Ph. D. diss., University at Buffalo, The State University of New York.
- Shimojo, Mitsuaki  
in press Focus structure and beyond Discourse-pragmatics in Role and Reference Grammar. In *Studies in Role and Reference Grammar*. Lilián Guerrero, Sergio Ibáñez, and V. Belloro (eds.). México: UNAM.
- Silva Encinas, Carlos  
2004 Life Story of Hilario (oral text). In La secuencia temporal en el discurso narrativo Yaqui [The temporal sequence of narrative discourse in Yaqui] M.A. Thesis. Hermosillo: Universidad de Sonora.
- Silva Encinas, Carlos  
n.d. Life story of Maejto (oral text). Manuscript.
- Silva Encinas, Carlos, Pablo Álvarez Romero, and Cresencio Buitimea  
1998. *Jiák nokpo étejoim, Pláticas en Lengua Yaqui* [Conversations in Yaqui language]. Hermosillo: Universidad de Sonora.
- Steele, Susan  
1977 Clisis and diachrony. In *Mechanisms of syntactic change*. Charles N. Li (ed.), 539-579. Austin: The University of Texas Press.
- Thompson, Sandra  
1978 Modern English from a Typological Point of View: Some Implications of the Function of Word Order. *Linguistische Berichte* 54:19-35.
- Van Valin, Robert D.  
1999 A typology of the interaction of focus structure and syntax. In *Typology and Linguistic Theory: From Description to Explanation*. E. Raxilina and J. Testelec (eds.), 511-24. Moscow: Languages of Russian Culture.
- Van Valin, Robert D.  
2005 *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. and Randy J. LaPolla  
1997 *Syntax. Structure, Meaning and Function*. Cambridge: Cambridge University Press.