KAQCHIKEL SVO: V2 IN A V1 LANGUAGE*

LAUREN EBY CLEMENS

Harvard University

1 Introduction

As a whole, Mayan languages are predicate-initial, i.e. VOS (England 1991, Aissen 1992). A small subset of Mayan languages from the Q’anjob’alan and Mamean branches, including, Chuj of San Sebastián Coatán, Eastern Jakalteko, Mam and Q’anjob’al, are strictly VSO (Mayers 1966, England 1983, Mateo Toledo 2008). Nonetheless, many Mayan languages allow subject-initial structures and a few Mayan languages even prefer subject-initial word order in transitives, while maintaining verb-initial word order in nonverbal predicates and other intransitives (England 1991, Edmonson 1988, Quizar 1994, Gutiérrez-Bravo and Monforte 2008). It is somewhat surprising then, that at least the present variety of Patzún Kaqchikel is predominately SV(O) and that this characterization holds for all types of intransitives. The data presented in this paper are representative of both elicited and narrated material and come from multiple speakers in Patzún and one speaker living in the United States at the time of data collection.

For matrix declarative clauses in this variety of Kaqchikel, subjects can always be clause-initial, but in order for them to surface to the right of the predicate, some other A'-element (adverb, locative, focused phrase, etc.) must surface clause-initially. The word order in embedded clauses is not subject to the same restrictions that we see in matrix clauses. In embedded clauses, verbs may be the first element. In other words, embedded clauses illustrate the verb- or predicate-initial tendencies of Mayan, while in matrix clauses some other requirement renders this characteristic opaque. Borrowing a common solution from the literature

---

* My field work in Patzún, Guatemala was supported by a summer research grant from the Harvard GSAS Graduate Student Council. I am especially thankful to Maria Polinsky, Amy Rose Deal and Judith Aissen for their guidance in the different stages of this project. Thank you also to my classmates in Ling117r (Harvard University 2011), 24.942 (MIT 2011), 24.896 (MIT 2012) and participants at the 86th meeting of the LSA and FAMLi 2 for valuable comments. While writing this paper I have benefited from conversations with many people including Sam Alxatib, Jessica Coon, Michael Erlewine, Laura Grestenberger, Robert Henderson, Michael Kenstowicz, Yusuke Imanishi, Junya Nomura, Pedro Mateo Pedro, Omer Preminger, Craig Sailor and Kirill Shklovsky. I am grateful to Ana Lopez, Jairo Lopez, Victoriana Sipac, Dominga Sipac, and Isabel Sipac for their hospitality, patience and Kaqchikel judgments. All errors in this paper are my own.
of V2 languages, I propose that the relevant clausal requirement is that the specifier of matrix declarative CPs is obligatorily filled. I argue that SVO in Patzún Kaqchikel is an artifact of this requirement.

The paper is organized as follows: In Section 2, I introduce basic Kaqchikel morphosyntax and constituent order. Section 3 begins by comparing Patzún Kaqchikel to V2 languages. As in V2 languages, I argue that the specifier of matrix declarative CP is obligatorily filled. Unlike V2 languages, other constituents can surface between the initial element and the verb, so long as some element is in CP. In this section I also present evidence that the CP requirement is syntactic and exists independently of prosody and information structure. In Section 4, I formalize the analysis and consider an alternative where non-focused preverbal elements are located in IP instead of CP. I draw from Kaqchikel’s particular brand of syntactic ergativity to settle the issue. In Section 5, I discuss relevant variation within Kaqchikel and the intrinsic tension between SVO and the agent focus construction. Section 6 concludes.

2 Kaqchikel Basics

2.1 Morphosyntax

As with other Mayan languages, Kaqchikel morphology is ergative-absolutive. Case is not marked overtly on DPs, but the data in (1) illustrate that the subject of an intransitive clause and the subject of a transitive clause condition the appearance of a different set of agreement markers on the verb.\(^1\)

\begin{align*}
\text{(1a)} & \quad \text{Ri a Juan i ri xta Kotz’ij x-e-ok.} \\
& \quad \text{DEF CL Juan COOR DEF CL Kotz’ij COM-3PL.ABS-enter} \\
& \quad \text{‘Juan and Kotz’ij entered.’} \\

\text{(b)} & \quad \text{Ri a Juan i ri xta Kotz’ij x-at-ki-tz’et.} \\
& \quad \text{DEF CL Juan COOR DEF CL Kotz’ij COM-2SG.ABS-3PL.ERG-see} \\
& \quad \text{‘Juan and Kotz’ij saw you.’}
\end{align*}

The agreement marker associated with intransitive verbs is the same as that which indicates the objects of transitive verbs. Compare (1b) to (2):

\begin{align*}
\text{(2)} & \quad \text{X-at-ok.} \\
& \quad \text{COM-2SG.ABS-enter} \\
& \quad \text{‘You entered.’}
\end{align*}

---

\(^{1}\) Abbreviations in glosses are 1, 2, and 3 for first, second and third person singular; ABS absolutive marker; AF agent focus; CL classifier; COM completive aspect; COMP complementizer; COOR coordinating conjunction; DEF definite article; ERG ergative marker; EXT existential; FOC focus; INCOM incompletive aspect; INDEF indefinite article; PL plural, POSS possessive; and SG singular. I attempt to conform to the orthographic convention of the Academia de las lenguas Mayas de Guatemala with help from (Macario Cutzal and Cali 1998). The standard orthography does not entirely represent this variety of Kaqchikel, which does not maintain the tense/lax contrast for vowels at all places of articulation and is pronounced with a good deal of syncope in functional morphology.
The table in (3) provides the complete paradigm of predicate-argument agreement, for which allomorphy is phonologically conditioned.

(3) Agreement Markers

<table>
<thead>
<tr>
<th>Person</th>
<th>Number</th>
<th>Ergative</th>
<th>Absolutive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SG</td>
<td>in-/w-q(a)-</td>
<td>i(n)-oj-</td>
</tr>
<tr>
<td>2</td>
<td>SG</td>
<td>a(w)-i(w)-</td>
<td>a(t)-ix-</td>
</tr>
<tr>
<td>3</td>
<td>SG</td>
<td>r(u)-k(i)-</td>
<td>Ø e-</td>
</tr>
</tbody>
</table>

Aspect is also marked with a verbal prefix: y-/ni- is used for incompletive aspect and x- is used for completive. These morphemes are considered aspectual and not temporal, because both can be used to describe past and present events.

2.2 Constituent Order

Although Mayan languages are by and large verb-initial, SVO is not uncommon in the family. Yet, Kaqchikel stands out from other Mayan languages with SVO in the prevalence of subject-initial word order across predicate types. The standard analysis of SVO in Mayan comes from Aissen (1992), who proposes that VOS is base-generated and SVO is derived for Tzotzil, Jakaltek and Tz’utujil. According to her analysis, the subject enters the derivation in a rightward oriented specifier, resulting in the base-generation of VOS word order.

(4) Right-branchingSpecifier Account of VOS (X = V, A or N)

```
XP
  X'
    Subject
X''
    Object
```

This account is the prevailing approach to Kaqchikel phrase structure within the extended projection of the verb or equivalent in the case of non-verbal predicates (pace Kim 2011).

Aissen (1992) argues that when arguments appear in a pre-verbal position, they are associated with a left-branching functional projection higher than the extended projection of the verb. Concretely, SVO arises when preverbal subjects are in A'-positions that are associated with either topic or focus. Focused constituents reside in IP, while topics reside in the specifier of CP or are in a projection adjoined to CP (what she calls internal and external topics). The languages Aissen considers in the paper exhibit preverbal elements in the position associated with focus and one or both positions associated with topic. The non-focused preverbal elements I address in this paper are structurally analogous to the (CP) internal topics of Tz’utujil, a closely related language to Kaqchikel.
The analysis presented in this paper is structurally congruent with Aissen’s, but the Kaqchikel facts reported in this paper dictate a divergence from Aissen’s account in a couple of ways. First, I argue that that preverbal arguments located in CP are not topics in a discourse-relevant sense, but are fulfilling a structural requirement of the clause. Second, the structural requirement that preverbal elements in CP fulfill can also be satisfied by non-DPs (see Section 3.2). Finally, I argue that the specifier of the matrix declarative CP in Patzún Kaqchikel must be filled by some constituent. In effect, matrix declarative clauses may surface with any word order other than verb-initial. All known exceptions to this generalization are discussed in this section.

In the few Mayan languages for which SVO is reported to be the preferred word order in transitive clauses, predicate-initial word order is either preferred or obligatory with non-verbal predicates and other intransitives (see Edmonson 1988 for Huastec, Gutiérrez-Bravo and Monforte 2008 for Yucatec, Kim 2011 for Kaqchikel and Quizar 1994 for Ch’orti’). In contrast, Patzún Kaqchikel speakers report a preference for subject-initial word order in transitives and intransitives; and I have encountered no counterexamples in narratives.

In this variety of Kaqchikel, nonverbal predicates have subject-initial word order, regardless of the weight or complexity of the subject, as illustrated in (6). The V1 renditions of the data in (6) are polar questions.

(6a) Rije’ e-tijoxel-a’.
   3PL 3PL.ABS-student-PL
   ‘They are students.’

(b) Ri tz’i’ nüm.
   DEF dog big
   ‘The dog is big.’

(c) Ri ru-tz’i’ ru-chaq’ a Juan jeb’elix.
   DEF 3SG.Poss-dog 3SG.Poss-brother CL Juan beautiful
   ‘Juan’s brother’s dog is beautiful.’

In general, V1 matrix clauses produced with falling intonation, characteristic of statements, are reported to be poorly formed questions, which should have rising intonation. For example, when (7) is produced with rising intonation, it is a polar question, incompatible with declarative interpretation. When (7) is produced with falling intonation, it is ungrammatical.
Kaqchikel SVO: V2 in a V1 Language

(7) X-e-r-ch’áj ri laq ri xta Kotz’ij.
COM-3PL.ABS-3SG.ERG-wash DEF dishes DEF CL. Kotz’ij
‘Kotz’ij washed the dishes?’
*‘Kotz’ij washed the dishes.’

Kaqchikel is robustly pro-drop, and when an argument is identifiable by verbal morphology only, verb-initial matrix clauses make fine declaratives.

(8) Ø X-e-r-ch’áj ri laq.
pro COM-3PL.ABS-3SG.ERG-wash DEF dishes
‘(S)he washed the dishes.’

Thus, when (8) is produced with rising intonation, it is a polar question (and presumably pro is located after the object); when it is produced with falling intonation, it is a declarative clause (and presumably pro is located before the verb). This variety of Kaqchikel does not have the matrix polar question particle la found in other varieties (see Section 5 for a discussion of la). Examples (6-8) were elicited via translation and grammaticality judgments, but it is worth repeating that in roughly 40 minutes of narration (all from one speaker), I do not have a single token of a verb-initial declarative that is not an instance of pro-drop.

In the context of elicitation, some speakers report that some verb-initial intransitive matrix clauses make better statements than others, although both are less preferred than their S-V counterparts. Dayley (1985) reports that discourse sensitive VS-SV alternations arise in Tz’utujil; however, in the case of Patzún Kaqchikel, information structure is unlikely to explain why some verb-initial intransitive matrix clauses make better statements than others. The judgments reported in (9-10) were elicited in a single context. If discourse considerations were a factor, we would expect a uniform degree of acceptability across VS examples in a single context and no contrasting degrees of acceptability. Instead, the SV-VS contrast, to the extent that one exists, appears to be an issue of verb class. The VS declaratives in (9) are marked ‘?’ to indicate that they are not as well formed as their corresponding polar questions, but somewhat better than the declaratives in (10).

(9a) X-Ø-jiq’ ri xta Maria / rija’.
COM-3SG.ABS-drown DEF CL Maria / 3SG
‘Maria / she drowned?’
*‘Maria / she drowned.’

(b) X-Ø-jison ri xta Maria / rija’.
COM-3SG.ABS-hiss DEF CL Maria / 3SG
‘Maria / she hissed?’
*‘Maria / she hissed.’

2 The judgments and generalizations pertaining to (9-10) conflict with those found in Kim (2011), where it is reported that Patzún Kaqchikel freely allows VS and SV alternations in intransitives.
The contrast in grammaticality between the VS order in examples like (9) on the one hand and those like (10) on the other may fall along the line of unaccusativity. However, without an independent test for unaccusativity, this hypothesis is tentative. Importantly, there is a clear preference for SV as opposed to VSO order even in examples like (9). See also Gutiérrez-Bravo (2007) for a discussion of word order according to verb class.

In contrast, embedded clauses have VOS and SVO variants that are reported to be equally acceptable and indistinguishable in meaning. An SVO and VOS version of an embedded intransitive and transitive verb are provided in (11-12) below.³

³ Kim (2011) derives VOS surface order via predicate fronting in the spirit of Coon (2010). I do not attempt to provide an analysis of word order variability in polar questions or embedded clauses in this paper.
3 A V2 Doppelgänger

This section is based on an analogy between Patzún Kaqchikel and the Germanic V2 languages. I propose that this variety of Kaqchikel requires that a constituent originate in or move to the specifier CP in matrix declarative clauses. This explains why a subject may surface post-verbally exactly when one of the following three conditions obtains: 1) the object undergoes focus movement; 2) an adverb surfaces clause-initially; or 3) an oblique argument surfaces clause-initially. In a sense, this is a V2 conspiracy, because ultimately we will see that this language is not V2 in the descriptive sense that the verb must be the second element of a clause (it can be the third element, fourth, etc.), but in the technical sense that in certain cases there is a position in CP that must be filled for structural reasons. The practical result of the CP requirement is that SV(O) is the most common word order of this variety of Kaqchikel.

3.1 Focus Movement

The first type of object focus we will consider is when the object is fronted and preceded by the particle *ja*. Subjects may surface after the verb when the object undergoes focus movement, as shown in (13a):

(13)  Oₑ,SV
Ja ri achi’a’ x-e-r-tz’êt ri a Juan.
FOC DEF man-PL COM-3PL.ABS-3SG.ERG-see DEF CL Juan
‘It was the men who Juan saw.’

Although (13) is translated as a cleft in English, I assume that this structure is monoclausal. Canonical SVO order is unavailable when the object is focused (14), which illustrates that the focused object must move into a higher (focus) position.

---

4 As one reviewer points out, it is also possible to treat these structures as headless relatives. This type of biclausal analysis would consist of positing a null pronominal in the embedded clause that controls agreement and coindexes with the focused constituent. The focused constituent would be base-generated in the matrix clause and take a headless relative as its subject. However, data with multiple preverbal elements make this sort of analysis less viable. For example, it would be difficult to rule out examples like (13”), because SV is freely allowed in embedded contexts. A version of (13) that also has two preverbal arguments is allowed (13”), but here, the unfocused argument precedes the focused argument. The ungrammaticality of (13”) is not predicted by a biclausal analysis and the grammaticality of (13”) would be more difficult to account for, if not impossible.

(13”)  Oₐ,SV
*Ja ri achi’a’ ri a Juan x-e-r-tz’êt.
FOC DEF man-PL DEF CL Juan COM-3PL.ABS-3SG.ERG-see
Intended: ‘It was the men who Juan saw.’

(13”)  Oₐ,SV
Ri a Juan ja ri achi’a’ x-e-r-tz’êt.
DEF CL Juan FOC DEF man-PL COM-3PL.ABS-3SG.ERG-see
Intended: ‘It was the men who Juan saw.’
Another type of object focus occurs when the object is preceded by *xaxe* ‘only.’ This type of movement has similar properties to what has been shown for *ja*. The data in (15) illustrate that the focused object must be located in a preverbal position.

Finally, objects may surface before the verb, allowing the subject to follow the verb, when the object is separated from the rest of the clause with an intonational break (indicated by a hash in (16a)). If and only if the extra auditory cue is present, is it possible to get the OVS interpretation. Furthermore, if the extra auditory cue is present, only the OVS interpretation is possible. Therefore, neither (16a) nor (16b) is ambiguous, and in (15a), the object contrasts with some other possible object, i.e. ‘It was Juan not Kotz’ij who Maria saw.’

I assume that the examples in (16) are structurally analogous to other focus constructions in Kaqchikel, for example (13) and (15a), and that in (16), focus is marked with focus prosody instead of focus morphology.

### 3.2 Initial Adverbs and Obliques

When a matrix clause begins with an adverbal phrase, V(OS) word order becomes grammatical. Removing the initial adverb and leaving the rest of the clause intact, however, is only possible under a polar question interpretation. The data in (17) show a manner adverb, a temporal adverb, and a heavy adverbial phrase with predicates of various types illustrating this point.
Kaqchikel SVO: V2 in a V1 Language

(17a)#(Ütz) x-Ø-u-ch’aj ru-tzyāq ri xta Kotz’ij.
good COM-3SG.ABS-3SG.ERG-wash 3SG.POSS-clothes DEF CL Kotz’ij
‘Kotz’ij washed her clothes well.’

(b) #(Jantäq) n-Ø-b’e pa jay ri syan.
sometimes INCOM-3SG.ABS-go PREP house DEF cat
‘Sometimes the cat goes into the house.’

(c) #(Pa oxi’ juna’ pa nab’ey lunes chin saq’ij)
PREP three year PREP first Monday COMP dry.season
n-Ø-b’e pa tinamït ri achin.
INCOM-3SG.ABS-go PREP town DEF man
‘In three years, on the first Monday of the dry season, the man will go to town.’

In much the same way, an oblique argument that fronts above the verb licenses a post-verbal subject, as illustrated by the data in (18). Here too, if the oblique arguments were removed from these examples, they would become polar questions:

(18a)#(K’in jun kuchara) x-Ø-u-taj ri sopa ri xta Maria.
PREP INDEF spoon COM-3SG.ABS-3SG.ERG-eat DEF soup DEFCL Maria
‘Maria ate soup with a spoon.’

(b) #(Pa ruwi ri qejoj) x-Ø-tzopin ri tz’i’.
PREP top DEF sticks COM-3SG.ABS-jump DEF dog
‘The dog jumped over the sticks.’

The data presented so far can be summarized as follows: in embedded clauses, subjects may surface to the left or the right of the predicate (SVO, VOS); in matrix clauses, subjects can always be clause-initial, but in order to surface to the right of the predicate, there must be an A’-element (focused object, adverb or oblique) in initial position. Many types of constituents can satisfy this requirement. So far I have given examples of subjects, adverbs, focused objects, and indirect objects and other obliques.5

Together these data suggest that matrix clauses in Kaqchikel have a position in CP that must be filled; and in subject-initial clauses, it is the subject that meets the CP requirement. One final addition to the list of elements that can satisfy the CP requirement is the yes/no particles ja and ne. The data in (19) illustrate a context where this might occur. They are organized in a readable fashion, but the first example of ‘yes/no particle-V-O-S’ was spontaneously produced by a consultant, i.e. it was not elicited as a grammaticality judgment.

---

5 Notably absent from this list is non-focused objects. In elicitation context, non-focused objects in preverbal position are ungrammatical (see examples 16a and b). A reviewer informs me that for multiple related languages, OVS without object focus is rejected in elicitation contexts but is nonetheless attested in speech. This word order does not appear in the narratives I’ve considered, but I will refrain for the time being from making any claims on the basis of an apparent lack of OVS without object focus.
(19a) Question: Xutäj oxi’ saq’molo’ ri xta Kotz’ij? / ‘Did Kotz’ij eat three eggs?’

Answers: Ja  x-Ø-u-täj      oxi’ saq’molo’ ri xta Kotz’ij.
yes  COM-3SG.ABS-3SG.ERG-eat three egg  DEF CL  Kotz’ij
‘Yes, Kotz’ij ate three eggs.’

Ja  ri xta Kotz’ij x-Ø-u-täj      oxi’ saq’molo’.
yes  DEF CL  Kotz’ij  COM-3SG.ABS-3SG.ERG-eat three egg
‘Yes, Kotz’ij ate three eggs.’

(b) Question: Xutäj oxi’ saq’ul ri xta Kotz’ij? / ‘Did Kotz’ij eat three bananas?’

Answers: Ne  x-Ø-u-täj      oxi’ saq’molo’ ri xta Kotz’ij.
no  COM-3SG.ABS-3SG.ERG-eat three egg  DEF CL  Kotz’ij
‘No, Kotz’ij ate three eggs.’

Ne  ri xta Kotz’ij x-Ø-u-täj      oxi’ saq’molo’.
no  DEF CL  Kotz’ij  COM-3SG.ABS-3SG.ERG-eat three egg
‘No, Kotz’ij ate three eggs.’

For good measure, if the yes/no particle is omitted and agreement/disagreement is instead gestured, the verb-initial responses are ungrammatical, as is expected.

3.3 Possible Combinations of Preverbal Elements

Section 3.1 included examples where focused constituents, which must occur to the left of the verb, allow subjects to occur to the right of the verb. Focused objects and preverbal subjects are not in complementary distribution, and both can occur before the verb. When a clause contains two preverbal elements and one of them is focused, however, the focused constituent must follow the other preverbal element. (20) provides an example.

(20a) Subject » FocusJa

Ri  a  Juan  ja  ri  achi’-a’  x-e-r-tz’ët.
DEF CL Juan FOC  DEF  man-PL  COM-3PL.ABS-3SG.ERG-see
‘It was the men who Juan saw.’

(b)  *FocusJa » Subject

*Ja  ri  achi’-a’  ri  a  Juan  x-e-r-tz’ët.
FOC  DEF  man-PL  DEF  CL  Juan  COM-3PL.ABS-3SG.ERG-see
Intended: ‘It was the men who Juan saw.’

Quantified elements like the examples with xaxe ‘only’ in (12) behave like ja-focused elements (20) with regard to their position relative to preverbal subjects, as shown in (21).
(21a) Subject » Focus\textsubscript{XAXE} \\
Ri a Juan xaxe ri achi’-a’ x-e-r-tz’ët. \\
DEF CL Juan only DEF man-PL COM-3PL.ABS-3SG.ERG-see \\
‘Juan saw only the men.’

(b) *Focus\textsubscript{XAXE} » Subject \\
*Xaxe ri achi’-a’ ri a Juan x-e-r-tz’ët. \\
only DEF man-PL DEF CL Juan COM-3PL.ABS-3SG.ERG-see \\
Intended: ‘Juan saw only the men.’

The strict word order requirement illustrated by (20) and (21) suggests that preverbal subjects and focused constituents are associated with different structural positions. As already noted, many Mayan languages have a designated preverbal topic position and a designated preverbal focus position. In Mayan languages that have both positions, topic precedes focus (Aissen 1992, Norman 1977). In the next section, I will present arguments that preverbal arguments located in matrix declarative CPs are fulfilling a structural requirement of the clause and are not topics in the discourse-relevant sense. Yet, the general pattern of preverbal constituent order in Patzún Kaqchikel is consistent with the rest of the family, in so much as focused constituents must be left-adjacent to the verb and right-adjacent to any non-focused preverbal elements.\textsuperscript{6}

3.4 On the Nature of the CP Requirement

I have generalized across a set of data that some element must appear before the verb in matrix declarative clauses in Patzún Kaqchikel. This restriction is at least in part syntactic: the element that appears before the verb must be a syntactic constituent. In theory, it is also possible that this restriction is the result of an interaction between syntax and phonology or syntax and information structure. Would it be more accurate to say that the relevant preverbal element be a syntactic constituent or a prosodic word? Can these preverbal elements be any type of syntactic constituent or are they actually topics? I will discuss these possibilities in turn.

Researchers have turned to prosodic considerations to explain anomalous behavior in otherwise well-behaved V2 languages. For example, Rice and Svenonius (1998) show that in some dialects of Norwegian, the verb must be the second prosodic word, not syntactic element, and thereby integrate what appear to be V3 constructions into the standard V2 system. The question then becomes: to what extent can this analysis be applied to V2 constructions in other languages? For this variety of Kaqchikel, this approach does not make it off the ground. Kaqchikel, like many Mayan languages, is robustly pro-drop (England 1991). Verbs can be the first prosodic word in a statement so long as they are the second syntactic element. Example (8) is repeated as (22) below.

\begin{verbatim}
(22) Ø X-e-r-ch’äj ri läq. \\
pro COM-3PL.ABS-3SG.ERG-wash DEF dishes \\
‘(S)he washed the dishes.’
\end{verbatim}

\textsuperscript{6} An exception to this rule is that some adverbs can optionally surface between the focused constituent and the verb.
The second possibility, that Patzún Kaqchikel has a designated topic position that must be filled, seems equally unlikely. While there is a fair amount of disagreement in the literature with regard to how to diagnose a topic (see Lambrecht 1994 for an overview), topics are commonly associated with old information or what a sentence is about (Halliday 1967, Putnam 1975). In the context of discourse, topics usually contain information that has been previously introduced. This does not appear to be the case for this language. In (23), _ri kaq’iq_ ‘the wind’ and _jun ta’ala_ ‘the boy’ are both introduced for the first time in the narrative, and both are in the preverbal position:

(23a) 
\[
\text{Ri kaq’iq’ x-Ø-u-tzaq-kën ru-pawi’}
\]
\[
\text{DEF wind COM-3SG.ABS-3SG.ERG-drop-LOC 3.SG.POSS-hat}
\]
\[
\text{‘The wind blew off his hat.’}
\]

(b) 
\[
\text{Jun ta’ala’ x-Ø-b’e naqaj ri chakäch…}
\]
\[
\text{INDEF boy COM-3SG.ABS-go near DEF basket}
\]
\[
\text{‘A boy went near the basket…’}
\]

The data in (23) show that a discourse element may make its first appearance in the preverbal position, which makes it an unlikely that the preverbal position is a designated topic position.

Next, many people working on information structure have proposed that topics must contain a constituent that is familiar to both the speaker and listener (Givón 1990, Lambrecht 1994). The fact that definite and indefinites can appear in a preverbal position is further evidence that the relevant preverbal elements are not topics. In addition to (19b), two elicited examples with preverbal indefinites are given (24).

(24a) 
\[
\text{Jun tz’i’ x-Ø-r-aq’ij ri xta Kotz’ij.}
\]
\[
\text{INDEF dog COM-3SG.ABS-3SG.ERG-bite DEF CL Kotz’ij}
\]
\[
\text{‘A dog bit Kotz’ij.’}
\]

(b) 
\[
\text{Jun tz’i’ nu-Ø-ba’on.}
\]
\[
\text{INDEF dog INCOM-3SG.ABS-bark}
\]
\[
\text{‘A dog is barking.’}
\]

The examples in (24) were elicited in a context where neither the speaker nor the listener could identify the dog that did the barking or the biting. Since the indefinite article _jun_ can also be the numeral ‘one,’ the examples in (24) are also felicitous in a context where the culprit is one particular dog in a group of dogs familiar to both speaker and listener. I avoided the alternative numeric interpretation by controlling the elicitation context and verifying that the intended interpretation was possible.

For a Mayan language to freely allow indefinite subjects in preverbal position is somewhat surprising, and there is a general tendency in the family to avoid indefinite and inanimate subjects (England 1991 and sources therein). Dayley (1985) reports that most subjects are

---

7 An example of an inanimate subject is given in (i):

(i) 
\[
\text{Ri aq’om x-Ø-r-aq’omaj ri a Juan.}
\]
\[
\text{DEF medicine COM-3SG.ABS-3SG.ERG-cure DEF CL Juan}
\]
\[
\text{‘The medicine cured Juan.’}
\]
marked with the definite article or are unmarked in Tz’utujil, and indefinite subjects are strictly disallowed in preverbal position in transitive clauses when the object is definite. As shown by (24a), this variety of Kaqchikel allows preverbal indefinite subjects where Tz’utujil does not (see also Aissen 1999). Kim (2011) also reports that Patzún Kaqchikel does not restrict the type of subject that can surface in preverbal position.

In sum, data from the previous section demonstrate that a syntactic constituent must occupy CP for matrix declarative clauses to be well-formed. Data from this section demonstrate that this requirement is neither prosodic nor pragmatic in nature. Across Mayan, it is commonly understood that preverbal positions are discourse sensitive, as arguments in preverbal position can be shown to be associated with topic or focus (Aissen 1992, England 1991, Norman 1977, and others). For Patzún Kaqchikel, this does not appear to be true: there are no restrictions on the type of subjects that can appear in preverbal position (see also Broadwell 2000 and Kim 2011), and there are no restriction on the discourse context in which preverbal subjects can appear.

4 Proposal

This section develops a proposal that seeks to capture the generalization that the subject of a matrix declarative clause can surface after the verb only when some A’-element surfaces before the verb. Specifically, matrix declarative $C^0$ has a generalized EPP feature that does not appear to be constrained by intervention. Recall that Kaqchikel is rather indiscriminate about which constituent appears before the verb so long as some constituent does.

An alternative analysis would be that preverbal subjects in this variety of Kaqchikel are located in the specifier of IP, unless a null expletive resides there, in which case the subject can remain low (see Goodall 2001 for Spanish). Other preverbal elements would be in higher projections such as CP or independent focus projections. This is a more traditional view of EPP in that it targets subjects. This analysis could account for the SVO/VOS alternations in this paper, but it is unsatisfactory because it fails to capture the generalization of the data as a whole: only when some other element surfaces before the verb can the subject surface below the verb. In Section 4.4, I rule out a generalized EPP feature on $I^0$ with data from Kaqchikel’s ‘Agent Focus’ (AF) construction (see e.g. Stiebels 2006; Coon et al. 2011).

4.1 Kaqchikel V2

The proposal that $C^0$ has a generalized EPP feature captures the noteworthy similarities between structures that surface with preverbal subjects on one hand and structures with other preverbal elements and postverbal subjects on the other. The higher functional projections of an SVO clause are schematized below, without making any claims about the derivation of $vP$, i.e. whether it is ultimately verb- or subject-initial.
(25) SVO

\[
\text{CP} \\
\text{Subject} \rightarrow C' \rightarrow C_0^{[+EPP]} \rightarrow \text{IP} \rightarrow \text{IP}^0 \rightarrow \text{vP} \\
\text{Verb Object}
\]

(25) shows that SVO word order is derived via A'-movement of the subject from a vP-internal base-generated position to the specifier of CP.

Preverbal subjects do not appear to be generated in CP. One familiar diagnostic for CP-generated subjects (or topics) relies on scope ambiguities. If the subject is base-generated in CP, an object should not be able to scope over the subject. The subject cannot reconstruct to a lower position in the clause, because it never resided in a lower position in the clause. Even in embedded clauses, Kaqchikel does not demonstrate the scope profile one would expect if preverbal subjects were base-generated.

(26) X-Ø-in-waxaj chin chekonojel ri tijoxel-a’
    COM-3SG.ABS-3PL.ERG-hear COMP every DEF student-PL
    x-Ø-ki-sik’ij jun sikiwuj.
    COM-3SG.ABS-3PL.ERG-read INDEF book
‘I heard that every student read a book.’
\(\forall \geq \exists\)
\(\exists \geq \forall\)

An O_FOC VS clause is schematized below in (27). Note that the focused object moves first to the specifier of IP, before moving into the specifier of CP to satisfy EPP. I suspect that the first movement to IP is feature driven, since the class of arguments located in IP is uniformly scope-taking. Examples of ja FOC and xaxe ‘only’ focus have been provided, other items include the existential operators k’o EXT and majun ‘no one’ and DP negation.

(27) O_FOC VS

\[
\text{CP} \\
\text{Focused Object} \rightarrow C' \rightarrow C_0^{[+EPP]} \rightarrow \text{IP} \rightarrow \text{IP}^0 \rightarrow \text{vP} \\
\text{Verb Subject}
\]
Finally, (28) schematizes the later part of the derivation of XVOS, where X refers to the set of non-arguments that can surface before the verb. Recall that this set includes adverbs, oblique arguments and even yes/no particles

(28) XVOS

\[
\text{XP} \quad \{\text{oblique adverb yes/no}\} \quad \text{C'} \quad \text{IP} \quad l^0 \quad \text{vP}\]

I assume that the yes/no particles are base-generated in the specifier of CP and that some adverbial phrases may move to CP from vP and IP and others may be base-generated in CP.

Moving to obliques, I adopt Henderson’s (2008) movement analysis for a class of obliques that optionally surface before the verb. When members of this class of obliques appear in preverbal position, the clitic wi surfaces at the extraction site (extraction out of some other types of adjuncts triggers wi as well). Example (29) is from Henderson (2008).

(29a) X-Ø-in-lōq’ pa k’yab’āl.
    COM-3SG.ABS-3PL.ERG-buy PREP market
    ‘I bought it in the market.’

(b) Pa k’yab’āl x-Ø-in-lōq’ *(wi).
    PREP market COM-3SG.ABS-3PL.ERG-buy wi
    ‘I bought it in the market.’

It is possible that other preverbal obliques are base-generated in CP, if wi is a consistent diagnostic for movement. See Henderson (2008) for a detailed description of the types of adjuncts that trigger wi when extracted and those that do not.

4.2 Other Word Orders

There are two data points that suggest that EPP on C does not necessarily attract the closest available constituent. The first was example (29b) where the oblique argument moves to CP from its base-generated position lower in the clause, as evidenced by the clitic wi that surfaces at the extraction site. If the mechanism resulting in V2 effects were sensitive to intervention, the subject should always move to CP, as it is the closest available constituent. A second reason to assume that locality is not relevant to the movement at hand is the fact that two constituents can surface before the verb if the lower one is focused. This sort of example is schematized in (30).
The fact that the CP requirement can be satisfied by any number of constituents and is unconstrained by intervention is atypical of an EPP feature. If one particular lexical head entered the derivation with a feature that checked the same feature on C₀, such as a topic feature, locality would not be the same concern. Yet, many subjects in simple SVO clauses do not have the characteristics of topics (see Section 3.4). A combined approach could say that matrix declarative CPs attract topics, and if no lexical head enters the derivation with a topic feature, the closest available constituent, i.e. the subject, moves as a last resort. Investigating the pragmatics of XVOS sentences would be the first step in testing this alternative hypothesis.

The data at my disposal does not support an analysis that posits that the highest preverbal element bears with a topic feature that checks a topic feature on C₀; instead, I propose a generalized EPP feature on C₀. Either way the derivation of polar questions and embedded clauses remains an unsolved problem. To my knowledge, all of the word order possibilities of matrix declarative clauses in addition to the VOS order are options in both polar questions and embedded clauses. I leave the derivation of VOS and the word order variability of polar questions and embedded clauses for future research.

### 4.4 SVO and Agent Focus

The question of whether subjects are base-generated in the preverbal position or move there is an important question, because Kaqchikel is not only morphologically ergative, but it is also syntactically ergative, i.e. the language generally restricts A-bar movement of the ergative argument. When the ergative argument undergoes relativization, wh-movement, or focus-movement, AF morphology stands in for regular transitive agreement morphology. Therefore, if preverbal subjects move to CP, we expect AF morphology.

The particular realization of AF constructions differs somewhat from language to language. In Kaqchikel, when an ergative subject is extracted from a root transitive, -ō appears suffixed to the verb; when an ergative subject is extracted from a derived transitive, the AF suffix is -n. Some exceptions to this generalization will be discussed below. The second morphosyntactic peculiarity of AF constructions pertains to agreement. Despite the fact that the verb takes two arguments, it only agrees with one. The agreement is always in the form of absolutive, but whether the absolutive agreement targets the extracted subject or the object depends on an

---

8 See Kazenin 1994 and Manning 1996 for a broader use of that term.
agreement hierarchy described for Kaqchikel by Ajsivinac and Henderson (2011) as Local > Plural > 3rd person (see also Dayley 1985 for Tz’utujil, Mondloch 1981 and Smith-Stark 1978 for K’ichee’). An example of a simple transitive compared to two AF versions of that transitive is given in (31).

(31a) Simple Transitive

\[
\text{Rix} \quad x-e-i-tz’et \quad rje’. \\
2\text{PL} \quad \text{COM-3PL.ABS-2PL.ERG-see} \quad 3\text{PL} \\
\text{‘Y’all saw them.’}
\]

(b) Agreement with 2\textsuperscript{nd} person subject

\[
\text{Ja} \quad ix \quad x-ix-tz’et-ö \quad rje’. \\
\text{FOC} \quad 2\text{PL} \quad \text{COM-2PL.ABS-see-AF} \quad 3\text{PL} \\
\text{‘It was y’all who saw them.’}
\]

(c) Agreement with 3\textsuperscript{rd} person object

\[
\text{Achike} \quad x-e-tz’et-ö \quad rje’? \\
\text{who} \quad \text{COM-3PL.ABS-see-AF} \quad 3\text{PL} \\
\text{‘Who saw them?’}
\]

A restatement of the problem is that preverbal subjects of transitive clauses are base-generated in their argument position and A’-extracted into a CP surface position (see previous section), so we expect AF morphology in all subject-initial clauses, but none of the examples of preverbal subjects in previous sections have AF morphology. Why is there no AF morphology in simple SVO clauses?

The solution to the movement/base-generation puzzle for preverbal subjects is rooted in the observation that the correlation between AF constructions and A’-extraction of transitive subjects does not hold in all cases in this variety of Kaqchikel. Instead, we find a number of AF blocking effects, i.e. instances where we would expect to find AF but do not (see Erlewine, this volume). The data in (32-33) provide one such example. In (32-33), both the subject and object of a transitive clause undergo focus movement. Judgments were elicited with a visual aid: a grid that contained images of different characters with different food items. In this case, all of the characters were eating multiple food items with the exception of Maria who only ate soup. In (32), the extracted subject requires AF morphology (compare 32a and b). In (33), the extracted subject does not allow AF morphology (compare 33a and b).

(32a) AF Required

\[
\text{Xaxe ri sopa ja ri xta Maria x-Ø-tj-ö k’en jun kuchara.} \\
\text{Only DEF soup FOC DEF CL Maria COM-3.SG.ABS-eat-AF PREP INDEF spoon} \\
\text{‘It was Maria who only ate soup with a spoon.’}
\]

(b) *Xaxe ri sopa ja ri xta Maria x-u-tej k’en jun kuchara. \\
\text{Only DEF soup FOC DEF CL Maria COM-3.SG.ERG-eat PREP INDEF spoon} \\
\text{Intended: ‘It was Maria who only ate soup with a spoon.’}
(33a) AF Blocking

*Ja ri xta Maria xaxe ri sopa x-Ø-tj-ö k’en jun kuchara.
FOC DEF CL Maria only DEF soup COM-3.SG.ABS-eat-AF PREP INDEF spoon

Intended: ‘It was Maria who only ate soup with a spoon.’

(b) Ja ri xta Maria xaxe ri sopa x-u-tej k’en jun kuchara.
FOC DEF CL Maria only DEF soup COM-3.SG.ERG-eat PREP INDEF spoon

‘It was Maria who only ate soup with a spoon.’

When the extracted subject is adjacent to the verb, AF morphology is obligatory (compare the grammaticality of (32a) to the ungrammaticality of (33b)). When the extracted object intervenes between the extracted subject and the verb, however, AF morphology is banned (32b and 33a). In other words, when transitive subjects are extracted, but are not adjacent to the verb, standard transitive morphology is the only option.

These data demonstrate that there is not a one-to-one correlation between AF morphology and the extraction of transitive subjects in this variety of Kaqchikel. The lack of AF morphology in SVO structures is not as condemning to a A'-movement analysis of preverbal subjects as it might first appear. Furthermore, they are helpful in determining whether the EPP is associated with CP or IP.

We know that it is descriptively inadequate to say that AF arises whenever an ergative argument is extracted. We do not know whether AF morphology arises when the subject moves to a position that is either linearly adjacent to the verb or to a phrasal projection that immediately dominates the verbal complex. The data in (32-33) do not distinguish between these hypotheses, but (34) does. Elements assumed to be located in I0, such as the progressive marker tajin, intervene between the focused subject and the verb in AF constructions.9 Therefore, linear adjacency does not predict when AF is required.

(34a) Ja ri xta Maria tajin n-Ø-tj-ö ri sopa.
FOC DEF CL Maria PROG INCOM-3.SG.ABS-eat-AF DEF soup

‘It is Maria who is eating the soup.’

(b) *Ja ri xta Maria tajin n-u-tej ri sopa.
FOC DEF CL Maria PROG INCOM-3.SG.ERG-eat DEF soup

Intended: ‘It is Maria who is eating the soup.’

I follow Aissen’s (1992) analysis of related languages that the specifier of IP is the location of focused constituents, but add that if there are two focused constituents, one is in IP and the other is in CP. This is the only possibility short of an iterated IP. Focused constituents cannot be lower than IP, because they do not appear in non-finite clauses. While I will not attempt to provide an analysis of AF in Kaqchikel, I propose that the mechanism that triggers AF morphology is sensitive to ergative arguments that pass through IP and not to ergative arguments that avoid passing through IP (pace Erlewine this volume).

Before beginning this section I mentioned an alternative to a analysis with EPP on C0, which is EPP on I0. If word order in SVO clauses in Kaqchikel were the result of EPP on I0, AF morphology would occur with all ergative arguments. On the other hand, if word order in SVO

---

9 Thanks to Robert Henderson for pointing this out to me.
clauses were the result of EPP on C⁰, only those preverbal subjects that move through IP should trigger AF morphology.

5 Variation and Change

In this section, I discuss variation in Patzún Kaqchikel, how the loss of the matrix polar question particle *la* may have caused this variety to become more stringently SVO than either Classical Kaqchikel or other modern varieties, the tension between SVO and AF, and how Kaqchikel-Spanish bilingualism may have also contributed to this change.

5.1 Variation in Patzún Kaqchikel

Kaqchikel-internal variation is not well understood. For instance, the variety of the Kaqchikel in this paper comes from Patzún, but there exist at least two Patzún-area varieties as evidenced by the observation that the data in Broadwell (2000), Broadwell and Duncan (2002), and Henderson (2009) and the data presented in this paper differ from one another in systematic ways. Two of the ways Patzún varieties differ from one another are particularly relevant to this paper and are likely related: 1) the way in which polar questions are formed; and 2) whether verb-initial matrix clauses are well-formed.¹⁰ Some Kaqchikel varieties have a polar question particle *la* that surfaces clause-initially in matrix clauses, which the Kaqchikel variety discussed in this paper has lost. In the Kaqchikel varieties that have *la*, the following pattern arises in polar matrix questions:

(35) Polar Questions in Kaqchikel varieties with *la*

(a) \(\text{la } x\text-a\text{-b’e } iwir?\)
\(\text{Q COM-2SG.ABS-go yesterday} \)
‘Did you go yesterday?’

(b) \(\text{la } iwir x\text-a\text{-b’e?}\)
\(\text{Q yesterday COM-2SG.ABS-go } \)
‘Was it yesterday that you went?’

(c) \(*iwir la x\text-a\text{-b’e?}\)
\(\text{yesterday Q COM-2SG.ABS-go } \)
‘Was it yesterday that you went?’

In the varieties without *la* the data in (35) would be restructured, rephrased with rising intonation, and reinterpreted as follows:

---

¹⁰ Thanks to Robert Henderson (p.c.) for bringing these differences to my attention, and for helpful discussion about this section generally.
(36) Polar Questions in Kaqchikel varieties without la

(a) X-a-b’e iwir?
    COM-2SG.ABS-go yesterday
    ‘Did you go yesterday?’

(b) Iwir x-a-b’e
    yesterday COM-2SG.ABS-go
    ‘Did you go yesterday? ’ (with rising intonation)
    ‘You went yesterday.’ (with falling intonation)

Losing la means losing the morphosyntactic realization of polar questions. Subsequently, it is possible that prosody replaced la as the morpheme went from optional to absent. This could serve as an explanation for why the V2 pattern emerges in varieties without la. Along with question intonation, obligatory fronting of some constituent serves to disambiguate polar questions from statements. Finally, even in varieties of Kaqchikel without la, the polar question particle for embedded questions, wi, is still present. If EPP features on matrix C⁰ arose from the need to disambiguate questions from statements, then the retention of wi could explain the divergent behavior of embedded clauses.

5.2 SVO and Agent Focus Revisited

Recall that there are certain environments where the subject of a transitive clause is extracted, but where transitive agreement is grammatical while the expected AF morphology is not. I explained this phenomenon earlier by positing that in order for an ergative argument to trigger AF, it must be located in IP. While AF is a prominent feature of Mayan languages, there are some varieties, primarily in the Ch’olan branch, that have lost AF entirely. To my knowledge, Kaqchikel is unique among the syntactically ergative Mayan languages with respect to AF blocking effects. My primary consultant has also reported on multiple occasions that standard transitive agreement in canonical AF environments does not sound absolutely terrible or that she can imagine hearing someone else use transitive agreement where she would not. Based on her intuitions, it seems altogether possible that AF in Kaqchikel is in an intermediate stage of decline.

Recall also that Kaqchikel (at least this variety) is unique among Mayan languages in its steadfast avoidance of verb-initial matrix clauses. While it would be imprudent to suggest that the rise of SVO causes the decline of AF or vice versa, it is hard to imagine how a language could maintain both an AF/transitive agreement contrast as well as resolute SVO word-order. If consistency is a consideration in syntactic change, a language would have one of two options: either lose AF entirely, so the situation is avoided where some but not all extracted subjects trigger AF; or apply AF everywhere, resulting in the invariable marking of subjects (abstracting away from person hierarchy effects) and ultimately leading to a loss of the ergative system (recall that Kaqchikel does not have case marking, only agreements markers on the verb).

The steady rise of bilingual Kaqchikel-Spanish and even Spanish dominant Kaqchikel speakers is an obvious if not verifiable explanation for why modern Kaqchikel, or even just the Patzún Kaqchikel of my consultants, has come to favor SVO word-order when sixteenth century Kaqchikel certainly did not (England 1991; Maxwell and Hill 2006). In fact, many researchers
maintain that SVO in Spanish is derived via the A'-movement of subjects into CP from an underlying verb-initial structure (see for example Bordelois 1974, Contreras 1991, and Alexiadou and Anagnostopoulou 1998). While my analysis of Kaqchikel differs from these analyses of Spanish in substantial ways, it should go without saying that bilingual Kaqchikel-Spanish speakers could have adopted elements of the surface structure of Spanish while not adopting the derivation of Spanish.

6 Conclusion

This paper has provided a uniform account for the following three properties of Kaqchikel word order: the usual order of SVO; the possibility of VOS in embedded contexts; and the possibility of VOS in matrix contexts when an A’ element surfaces in initial position. I have argued that an EPP feature on matrix C0 underlies the generalization that some constituent must surface before the verb. I explain the fact that standard SVO sentences in Kaqchikel do not trigger AF morphology with independent evidence suggesting that IP is the projection associated with AF and that an extracted ergative argument must travel through IP in order for agent focus morphology to appear.

References

Coon, Jessica. 2010. VOS as predicate fronting in Chol. Lingua 120.


Gutiérrez Bravo, Rodrigo. 2007. ‘Prominence scales and unmarked word order in Spanish’ *Natural Language and Linguistic Theory* 25.


Lauren Eby Clemens
Harvard University
Boylston Hall, 3rd Floor
Cambridge, MA 02138 USA
lreby@fas.harvard.edu