Grammar meets Discourse: the case of Yucatec Maya

Judith Tonhauser
Stanford University
juditht@stanford.edu

Main claim: The discourse context plays a role in the organization of the grammar of Yucatec Maya in a different way than has been described yet for another (Mayan) language.

Two domains of grammar:

D1. The mapping from lexical categories to syntactic functions is more flexible in Yucatec Maya than in more familiar languages like English. The discourse context determines whether a lexical predicate is realized as a syntactic argument or a syntactic predicate. In particular, the focus of a proposition is always realized as the main syntactic predicate. (cf., Beck (1997) for Salishan)

D2. Argument linking (i.e., the linking of overt arguments to semantic roles) does not make reference to case-marking or configurationally definable argument positions, but relies on the discourse function of event participants.

Overview of the talk:

§1 Yucatec Mayan syntax, and the role of the discourse context its organization (D1).
(based on several months of fieldwork and a corpus study; Tonhauser (2003a,b,c, ms.))

⇒⇒ The syntactic analysis I propose does not make reference to case-marking or configurationally definable argument positions.

§2 Argument linking in discourse (D2).

§3 Comparison with a standard, configurational analysis of Mayan languages (Aissen (1992, 1996, 1999)).

§4 Conclusions: the way in which the discourse context affects the organization of the grammar of Yucatec Maya has interesting implications for syntactic typology and theories of argument linking.

---

1The results reported in this paper are based on fieldwork on Yucatec Maya in Tulum, Quintana Roo, during the summers of 2002 and 2003. I would like to thank my language consultants, in particular Evaristo Dzul Caamal, for their time and patience, and the whole Dzul Caamal family for their friendship. I am grateful to the Stanford Center for Latin American Studies for financial support. For helpful comments and discussions I thank Judith Aissen, David Beaver, John Beavers, Maria Bittner, Lev Blumenfeld, Jürgen Bohnemeyer, Ashwini Deo, Veronica Gerassimova, Florian Jaeger, Paul Kiparsky, Lynn Nichols, Andrew Koontz-Garboden, Beth Levin, Ivan Sag, Peter Sells and Tom Wasow, as well as the participants of SULA II, CLS 39, the Stanford Semantics Workshop (all 2003), the Berkeley Syntax/Semantics Circle, the Berkeley fieldwork class, the Stanford Semantics Fest, and the Stanford QP Fest (all 2004). The usual disclaimers apply.

2Yucatec Maya is a Mayan language spoken by about 800 000 people on the Yucatan peninsula of Mexico, and bordering parts of Belize and Guatemala. I use the following glosses: AF=agent focus voice; A1-3=set A cross-reference marker; B1-3=set B cross-reference marker; CMP=completive status; D1-3=deictic markers; DEF=definite marker; IMPF=imperfective aspect; INC=incompletive status; PERF=perfective aspect; PREP=preposition; SBJ=subjunctive status; TA=transitive active voice; TERM=terminative marker.
1 The Syntax of Yucatec Maya

The relationship between lexical categories and syntactic functions in English:

<table>
<thead>
<tr>
<th>language</th>
<th>lexical category</th>
<th>prototypical semantic denotation</th>
<th>prototyp. syntactic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>nouns</td>
<td>individuals</td>
<td>argument</td>
</tr>
<tr>
<td>(e.g., Croft (1990))</td>
<td>adjectives</td>
<td>properties</td>
<td>adjunct</td>
</tr>
<tr>
<td></td>
<td>verbs</td>
<td>events</td>
<td>syntactic predicate</td>
</tr>
</tbody>
</table>

Compared to English, and other more familiar Indo-European languages, the relationship in Yucatec Maya between lexical categories and syntactic functions is more flexible: whether a member of a lexical category realizes a syntactic argument or a syntactic predicate depends on the discourse context.

§1.1: Lexical categories

§1.2: The syntax of Yucatec Maya

§1.3: The focus is realized as the main syntactic predicate, irrespective of the lexical category.

1.1 Lexical categories in Yucatec Maya

<table>
<thead>
<tr>
<th>language</th>
<th>lexical category</th>
<th>prototypical semantic denotation</th>
<th>prototyp. syntactic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yucatec Maya</td>
<td>dynamic predicates</td>
<td>events</td>
<td>–</td>
</tr>
<tr>
<td>(Bohnemeyer (1998))</td>
<td>stative predicates</td>
<td>individuals (“nouns”)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>properties (“adjectives”, “quantifiers”)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>relations (“prepositions”)</td>
<td></td>
</tr>
</tbody>
</table>

Dynamic predicates (“verb”):

(1) T-inw il-ah-ech.
PERF-A1sg see-CMP-B2sg
‘I saw you.’

<table>
<thead>
<tr>
<th>transitive A set A cross-reference markers</th>
<th>transitive O set B cross-reference markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg</td>
<td>pl</td>
</tr>
<tr>
<td>1</td>
<td>in(w)</td>
</tr>
<tr>
<td>2</td>
<td>a(w)</td>
</tr>
<tr>
<td>3</td>
<td>u(y)</td>
</tr>
</tbody>
</table>

Stative predicates as main syntactic predicates:

(2) a. Péek’-ech.
daog-B2sg
‘You’re a dog.’

b. Chan áák, óotsil -ech!
little turtle poor -B2sg
‘Little turtle, you are a poor one.’ (Bohnemeyer (1998), p.182, E92c)

c. Ya’ab -o’on.
many -B1pl
‘We are many.’

3The main morphosyntactic difference between dynamic and stative predicates is that only the former occur with aspect/mood markers.
There are five of us.’ (Bohnemeyer (1998), p.182, E92d)

‘That’s how they ended their days.’ (lit: It finished for them like that.) (AM1:131)

Both dynamic and stative predicates can realize the main syntactic predicate of a clause.

3rd person set B cross-reference marker: phonologically empty.

<table>
<thead>
<tr>
<th>Number</th>
<th>sg cross-reference marker</th>
<th>pl cross-reference marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-en</td>
<td>-o’n</td>
</tr>
<tr>
<td>2</td>
<td>-ech</td>
<td>-e’x</td>
</tr>
<tr>
<td>3</td>
<td>-∅</td>
<td>-o’b</td>
</tr>
</tbody>
</table>

Stative predication with third person argument

1.2 Arguments versus adjuncts

Deictic-marked phrases are adjuncts: they can precede or follow the predicate (different discourse function).

Arguments (ummarked) must follow the predicate.

*Máax is a nominal predicate meaning ‘person’; it receives an interrogative interpretation when realized as the main syntactic predicate (cf. Tonhauser (2003c)).
Clauses headed by a transitive verb: deictic-marked phrases must follow unmarked phrases in postverbal position.

(8) a. K-uy il-ik-∅ xch’úup le winik -o’.  
IMPF-A3 see-INC-B3sg woman DEF man -D2  
‘The man is seeing a girl/girls.’

b. *K-uy il-ik-∅ le winik -o’ xch’úup.  
IMPF-A3 see-INC-B3sg DEF man -D2 woman

(9) The syntax of Yucatec Maya:

- predicate: verbal or non-verbal lexical predicate realizes the focus of the proposition
- argument: must follow the predicate
- deictic-marked phrase: adjunct to S

1.3 The focus of the proposition is realized as the main syntactic predicate

We’ve seen that both dynamic and stative predicates can realize the main syntactic predicate of a clause.

This section: the discourse context affects the mapping from lexical categories to syntactic categories. In particular, the focus of a proposition is realized as the main syntactic predicate (cf., Beck (1997) for two Salishan languages).

Assumption:

Following the literature on questions and focus (cf., e.g., Hamblin (1973), Karttunen (1977), Rooth (1992)), I assume that focus determines coherent question/answer pairs.

I.e., the question word (of a question) is in focus, as well as that part of the answer that corresponds to the question word.

(10) Q: Who did John see?  
A: John saw MARY. (coherent answer)  
A’: #JOHN saw Mary. (incoherent answer)

In Yucatec Maya, where the focus of a proposition is realized as the main syntactic predicate, this means that both the question word (of the question) as well as the focus of the answer is realized as the main syntactic predicate.

The question word is in focus and therefore realized as the main syntactic predicate:

(11) a. Máax -ech?  
who -B2sg  
‘Who are you?’ =(5a)

b. Máax -∅ Juan?  
who -B3sg Juan  
‘Who is Juan?’ =(5b)
c. \( \text{Máax} \, \text{-l} \, \text{t-aw} \, \text{il-ah-} \, \text{t-aw} \, \text{-o’} \) \( \) (adjunct)
who -B3sg DEF PERF-A2 see-CMP-B3sg -D2
‘Who is the one you saw?’ \( \)= (6)

d. \( \text{Máax} \, \text{-l} \, \text{t-aw} \, \text{il-ah-} \) \( \) (argument)
who -B3sg PERF-A2 see-CMP-B3sg
‘Who did you see?’

Question/answer pairs: (12b) and (13b) realize the same lexical material, but differ with respect to which of the lexical material is realized as the main syntactic predicate, as well as with respect to the question to which they are a coherent answer.

(12) a. \( \text{Ba’ax} \, \text{-l} \, \text{k-a} \, \text{meet-ik} \) \( \)?
what -B3sg IMPF-A2 do-INC-B3sg
‘What are you doing?’

b. \( \text{K-in} \, \text{jan-tik} \, \text{bu’ul} \) \( \) denotation of verb in focus
IMPF-A1sg eat-INC-B3sg beans \( \Rightarrow \) verb is main syntactic predicate
‘I am eating beans.’

(13) a. \( \text{Ba’ax} \, \text{-l} \, \text{k-a} \, \text{jan-tik} \) \( \)?
what -B3sg IMPF-A2 eat-INC-B3sg
‘What are you eating?’

b. \( \text{Bu’ul} \, \text{-l} \, \text{k-in} \, \text{jan-tik} \) \( \) denotation of noun in focus
beans -B3sg IMPF-A1sg eat-INC-B3sg \( \Rightarrow \) noun is main syntactic predicate
‘Beans is what I’m eating.’

1.4 Intermediate summary

- Yucatec Maya has two main lexical categories, namely dynamic predicates (“verbs”) and stative predicates. The mapping from lexical to syntactic category is flexible, i.e., members of both classes can be realized as syntactic arguments and syntactic predicates.

<table>
<thead>
<tr>
<th>lexical category</th>
<th>syntactic category</th>
</tr>
</thead>
<tbody>
<tr>
<td>dynamic predicate</td>
<td>e.g., (12b)</td>
</tr>
<tr>
<td>stative predicate</td>
<td>e.g., (11)</td>
</tr>
</tbody>
</table>

- The mapping from lexical to syntactic category is sensitive to the discourse context: the focus of a proposition is realized as the main syntactic predicate.

- Based on a study of a Yucatec Mayan corpus, I propose in Tonhauser (ms.) that the position of all lexical expressions of a clause is determined by the discourse function of the denotation of the expression.
(9) The syntax of Yucatec Maya:

![Syntax Diagram](diagram.png)

**Advantage:** the analysis accounts for the data.

**Disadvantage:** highly non-standard syntax; no reference to configurational argument positions (and note that Yucatec Maya doesn’t have case-marking on nominal arguments).

**Challenge:** How is argument linking determined? I.e., how are lexical arguments linked to the appropriate semantic roles if arguments are not case-marked and the syntax does not provide for argument positions where semantic roles are assigned?

2 Argument Linking and Voice in Yucatec Maya

Yucatec Maya is a head-marking language: the cross-reference markers identify the person features of the A-, O-, and S-arguments of syntactic predicates.

Argument linking in clauses with one or two local cross-reference markers: determined by the cross-referencing system.

(14) **T-inw** il-ah-ech.
    PERF-A1sg see-CMP-B2sg
    ‘I saw you.’

(15) **T-inw** il-ah-ð Juan.
    PERF-A1sg see-CMP-B3sg Juan
    ‘I saw Juan.’

**Contextual resolution:** In clauses with a third person cross-reference marker but no lexical material, the referent(s) of third person cross-reference markers need to be contextually resolved (cf., Tonhauser (ms.)):

(16) **T-inw** il-ah-ð.
    PERF-A1sg see-CMP-B3sg
    ‘I saw him/her/it.’
Clauses with two third person arguments: argument linking is not determined by cross-referencing system.

(17) a. **Máax -∅** t-uy il -ah -∅ Maria? [transitive active voice]

Who PERF-A3sg see.TA -CMP -B3sg Maria

‘Whom did Maria see?’

b. **Máax -∅** il -eh -∅ Maria? [agent focus voice]

who see.AF -SBJ -B3sg Maria

‘Who saw Maria?’

**Argument linking in Yucatec Maya relies on syntax and voice:**

- Syntax: encodes the discourse function of the denotation of a lexical element (§1).
- Topicality hierarchy: topic > non-salient > focus, new ‘>’: higher topicality
- Voice: sensitive to the relative topicality of the event participants (cf. Cooreman (1993)).

<table>
<thead>
<tr>
<th>Voice Type</th>
<th>Ag/Th</th>
<th>Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive</td>
<td>ag &gt;</td>
<td>pt/th</td>
</tr>
<tr>
<td>active</td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent focus</td>
<td>ag &lt;</td>
<td>pt/th</td>
</tr>
<tr>
<td>intransitive</td>
<td>ag &lt;&lt;</td>
<td>pt/th</td>
</tr>
<tr>
<td>passive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>antipassive</td>
<td>ag &gt;&gt;</td>
<td>pt/th</td>
</tr>
</tbody>
</table>

Table 1: Relative topicality of the four main voices of Yucatec Maya

The discourse status of the event participant is reflected in the position and morphological marking of the lexical material that denotes the event participant.

- (17a) and (17b): máax ‘who’, the main syntactic predicate, is in focus; Maria is an unmarked argument and therefore non-salient.

  ⇒ on the topicality hierarchy: Maria > máax

The transitive predicate il ‘see’ is realized in the active voice in (17a) and the agent focus voice in (17b): both voices are transitive voices in Yucatec Maya (cf., Tonhauser (2003b)), but they differ with respect to the relative topicality requirement on the two arguments.

- (17a): transitive active voice: ag > pt/th
- (17b): agent focus voice: ag < pt/th

Given these two sets of constraints, argument linking is determined for the two examples in (17):

- (17a): ag > pt/th and Maria > máax

  ⇒ Maria is linked to the agent role, and máax to the patient role

- (17b): ag < pt/th and Maria > máax

  ⇒ Maria is linked to the patient role, and máax to the agent role

⇒ The cross-referencing system determines argument linking in most cases. In transitive clauses with two 3rd person arguments, argument linking relies on voice and word order (encodes discourse functions).
Summary

The analysis I propose for Yucatec Mayan syntax and argument linking is highly non-standard but it brings together two long-standing observations about Mayan languages:

1. Word order is highly sensitive to the discourse context. (cf., e.g., Brody (1984))
2. Mayan languages have voices that are sensitive to the discourse function of event participants. (cf., e.g., Ayres (1983), Zavala (1997))

3 Comparison with a standard, configurational analysis

\[
\text{EP} \quad \text{CP} \\
\text{topic} \quad \text{wh-words} \quad \text{IP} \\
\text{focus} \quad \text{VP} \\
\text{V} \quad \text{subject} \\
\text{verb} \quad \text{object}
\]

- configurational postverbal argument positions
- preverbal \( \overline{A} \)-positions for topics, content question words, and focused arguments

Advantages:
- argument linking can proceed as known from more familiar languages like English
- syntax can be easily compared to more familiar languages

3.1 The assumption of a basic word order

The existence of a basic word order with configurational argument positions is a central assumption of the configurational analysis of Mayan languages: arguments are base-generated in postverbal positions where they are assigned a semantic role by the verb.

However, the identification of a basic word order for Mayan languages is problematic because the criteria that are applied do not converge on the same word order (cf., e.g., Durbin and Ojeda (1978:69) for Yucatec Maya, Brody (1984) for Tojolabal, Hofling (1984) for the Yucatecan branch, Broadwell (2000) for Kaqchikel).

<table>
<thead>
<tr>
<th>criterion</th>
<th>word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>“least pragmatically marked”</td>
<td>VOA 2</td>
</tr>
<tr>
<td></td>
<td>VAO 2</td>
</tr>
<tr>
<td>“most frequent”</td>
<td>AVO 6</td>
</tr>
<tr>
<td></td>
<td>OVA 2</td>
</tr>
<tr>
<td></td>
<td>12 (29.3%)</td>
</tr>
<tr>
<td>one overt argument</td>
<td>27</td>
</tr>
<tr>
<td>zero overt arguments</td>
<td>29 (70.7%)</td>
</tr>
<tr>
<td>total</td>
<td>41</td>
</tr>
</tbody>
</table>

Table 2: Word order of transitive sentences of corpus in Tonhauser (ms.)

VOA word order is not pragmatically unmarked: it requires that the O-argument is “discourse-new” and the A-argument is “non-salient”.

The analysis of Yucatec Mayan syntax I propose in (9) does not assume such a basic word order.
Further restriction

The VOA word order is ungrammatical when the O-argument is deictic-marked and the A-argument is unmarked. The configurational analysis needs to assume an additional constraint to rule out (8b).

(8) a. K-uy il-ik-∅ xch’úup le winik -o’.
   IMPF-A3 see-INC-B3sg woman DEF man -D2
   ‘The man is seeing a girl/girls.’

b. *K-uy il-ik-∅ le winik -o’ xch’úup.
   IMPF-A3 see-INC-B3sg DEF man -D2 woman

My analysis straight-forwardly accounts for the data.

3.2 Content questions and focus construction

3.2.1 Co-occurrence restriction:

A single clause cannot realized a focused element and a content question, even in the appropriate context.

(19) *Ba’ax Maria jant-eh?
   what Maria eat.AF-B3sg
   (intended: What did MARIA eat?)

The configurational analysis needs to make further assumptions to explain why it is impossible to move both the subject and the object to the preverbal positions for wh-words and focused elements.

The answer according to the analysis in (9): focus constructions and content questions realize the focus and question word as the main syntactic predicate. Since there can be only one main syntactic predicate per clause, examples like (19) are ungrammatical.5

3.2.2 Focus and question word are predicates

According to the analysis presented in §1, the focus and question word are the main syntactic predicate of focus constructions and content questions, respectively.

Under the configurational analysis, the verb is the main predicate of focus constructions and content questions; the focus and the question word are arguments of the verb.

3.2.3 Focusing of referential expressions

In Yucatec Maya, it is possible to focus a proper name as in (20a), but not a definite noun phrase as in (20b).

(20) a. Juan -∅ il-∅-ech!
    Juan -B3sg see.AF-SBJ-B2sg
    ‘It was Juan who saw you!’

---

5In Tonhauser (2003c), I demonstrate, based on the assumption that content questions and focus constructions instantiate the structure in (9), that the two constructions differ with respect to the lexical semantics of the nominal that instantiates the main syntactic predicate: content questions realize one of the five ‘general’ nominals as the main syntactic predicate (i.e., mťan ‘person’, ba’ax ‘thing’, tu’ax ‘place’, buka’an ‘quantity’, or bix ‘manner’) while focus constructions realize a non-general nominal as the main syntactic predicate.
The configurational analysis needs to explain why certain referential expressions can be $\bar{A}$-moved to the focus position and other referential expressions can’t, and can only be $\bar{A}$-moved to the topic position.

The relevant difference between the two types of expressions is that proper names can be realized as syntactic predicates (cf. (21a)), and can therefore realize the main syntactic predicate in focus constructions, while definite noun phrases cannot realize the main syntactic predicate (cf., (21b)), and therefore need to be focused with a complex focus construction as in (22).

3.3 Agent focus voice

The “agent focus” voice is restricted to content questions, focus constructions and relative clauses.

Content questions:

(17) a. Máax $\emptyset$ t- uy il $\emptyset$ Maria? [TRANSITIVE ACTIVE] who PERF A3sg see.TA -CMP B3sg Maria ‘Whom did Maria see?’

b. Máax $\emptyset$ il $\emptyset$ Maria? [AGENT FOCUS] who see.AF SBJ B3sg Maria ‘Who saw Maria?’

According to the configurational analysis (cf., e.g., Klaiman (1991), Hale (1998), Aissen (1999)), the “agent focus” voice marks the extraction of an A-argument to an $\bar{A}$-position.

However, according to the criteria given in Aissen (1992), topics in Yucatec Maya are extracted to SpecCP, an $\bar{A}$-position. But the “agent focus” voice is ungrammatical with topics:

(23) *Le winik-o’ il-$\emptyset$-ech! 
DEF man D2 see.AF SBJ B2sg (intended: It was the MAN who saw you.)

Thus, not all extractions of A-arguments to $\bar{A}$-positions are marked with the “agent focus” voice. The configurational analysis needs to make further assumptions, e.g., Klaiman (1991): the “agent focus” voice only marks extraction to particular $\bar{A}$-positions.

Under my analysis (cf., Tonhauser (2003b, ms.)), the function of the “agent focus” voice is best understood...
as complementing the transitive active voice, which is restricted to realizing transitive A arguments with the “topic” discourse function. While there exist contexts in which a non-topical transitive A argument can simply be demoted (passive voice), there exist constructions in which a transitive A argument is not topical but nevertheless too important to be demoted, e.g., in content question, focus constructions and relative clauses, i.e., constructions where the transitive A is in focus. There exist languages where focused A-arguments are simply demoted, too, (e.g., Sesotho, Demuth (1995)) but Yucatec Maya has an additional transitive voice, the “agent focus” voice, to realized focused A-arguments without demoting them.

Thus, the agent focus voice occurs whenever a transitive A is in focus, i.e., realized as/by a syntactic predicate. The constructions in which this is the case are focus constructions, content questions, relative clauses, and nominalizations.

4 Conclusions, Implications and Future Work

- Although the analysis of Yucatec Mayan syntax and argument linking that I propose is rather non-standard, I believe that it compares favorably to standard, configurational analyses.

- The grammar of Yucatec Maya is highly sensitive to the discourse context:
  - The mapping from lexical to syntactic categories is flexible, and determined by the discourse context: the focus of a proposition is realized as the main syntactic predicate.
  - Argument linking cannot rely on configurational argument positions or case-marking, but proceeds via the discourse status of event participants.

- Argument linking in Yucatec Maya relies on word order and voice, just like English. The crucial difference seems to be that word order encodes discourse functions in Yucatec Maya, and grammatical functions in English. I therefore believe that the two systems of argument linking are more similar than it might seem at first glance.

  Yucatec Maya and the Salishan languages are surely not the only languages without case-marking on nominals and a word order that is highly dependent on discourse. Are there other languages for which it has been proposed that argument linking cannot rely on case-marking or configurational argument positions?

- Regarding syntactic typology, a finding of this work is that Yucatec Maya is not a discourse configurational language like Russian (King (1995)), Hungarian (Kiss (1995)) and other Mayan languages (Aissen (1992)) because it cannot be said to realize arguments with certain discourse functions in particular positions which are definable relative to the verb.

  Rather, Yucatec Maya looks like the Salishan languages discussed in Beck (1997). It’s not clear how these languages fit into current syntactic typologies concerned with the effect of discourse on syntax because the effect of the discourse context on the grammar of these languages seems to go deeper than the level of the clause (word order).
References