On the nature of word order in Yucatec Maya[*]

Rodrigo Gutiérrez-Bravo
Jorge Monforte y Madera

1. Introduction

It is a standard assumption that in languages that have SVO as their
unmarked word order, this order is the result of a "strong" EPP requirement.
This requirement is understood in Minimalist analyses as the presence of an
EPP feature in I^0 or C^0, and in Optimality Theoretic analyses as resulting
from the ranking EPP » STAY. In this paper we argue that Yucatec Maya (a
Mayan language from the Yucatán Peninsula, Mexico: henceforth, *Yucatec*)
instantiates a different kind of SVO language. Yucatec displays a
discrepancy between clauses where there is a phonetically-overt object,
where the unmarked word order is SVO, and clauses with no object, which
are by and large VS. We propose that this phenomenon be referred to as
split word order and show that standard analyses of the SVO order cannot
account for it. We develop an alternative Optimality Theoretic analysis,
where we argue that Yucatec has the ranking STAY » EPP characteristic of
verb-initial languages. However, we propose that a higher ranked constraint,
which we name DISTINCT prevents the subject and the object to appear
simultaneously inside the VP. This constraint forces the subject DP to move
outside the VP, thus producing the SVO order. Since this constraint is
sensitive to whether both the subject and the object have phonetic content,
when there is no overt object the resulting order is VS. The paper as a whole
is structured as follows. In Section 2 we address the problem of the
unmarked word order of transitive clauses in Yucatec and provide evidence
that SVO is the unmarked word order, contrary to what is usually claimed in
the literature. In Section 3 we address the issue of the word order of clauses
without a direct object and conclude that it is mostly VS. We introduce the
term *split word order* to refer to this word order discrepancy. In Section 4
we develop an Optimality Theoretic analysis that accounts for this split, and
in Section 5 we present our conclusions.

2. Yucatec Maya as an SVO language
2.1 Background

Mayan Languages (like the majority of Mesoamerican languages) are mostly verb-initial. As such, Yucatec is generally classified as a VOS language, where the SVO order is taken to be the result of topicalization of the subject. Yucatec indeed displays VOS clauses in texts and spontaneous speech.[1]

(1) *Je' bin k-u   la'ach-ik-ø   u   jo'ol x-nuk*
     ASV CIT   HAB-ERG.3   scratch-IND-ABS.3sg   ERG.3   head FEM-great
     reyna-o'.
     queen-CL
     'And the great queen truly scratched her head...'   (Gigante-97)

   However, none of the works that have dealt in detail with word order in Yucatec (Durbin and Ojeda (1978), Hofling (1984), Briceño Chel (2002)) has actually concluded that the language's unmarked word order is VOS. Instead, what these works observe is the notorious frequency of the SVO order.

(2) *Le   ko'olel-o' t-u   ts'-aj-ø   u   ma'alob nook'.*
     DM   woman-CL CP-ERG.3   put-PRF-ABS.3sg ERG.3   good   clothes
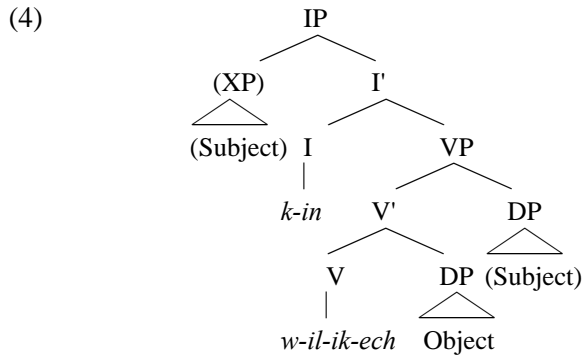   '...and the woman put on her good clothes...'   (Si'ipil-51)

   For instance, Durbin and Ojeda (1978) and Hofling (1984) both conclude, on the basis of qualitative criteria, that Yucatec is best understood as a language with two basic word orders: SVO y VOS. Hofling (1984) further acknowledges that if frequency alone is considered, SVO should be taken to be the basic word order. Similarly, based on frequency Briceño (2002) concludes that SVO is the unmarked word order of the language. A similar conclusion is arrived at in Gutiérrez-Bravo and Monforte (2008). As a first step in solving this puzzle, in 2.4 we briefly repeat the evidence which lead us to the conclusion in Gutiérrez-Bravo and Monforte (2008) that Yucatec is not a VOS language. In 2.4.4, we further present evidence from our own corpus of Yucatec oral narratives to support this conclusion.[2] Before presenting the relevant evidence, we lay out our assumptions about clause structure in Yucatec. We also briefly discuss the relation between topicalization, focus fronting and the SVO order in this language.

2.2 Transitive clauses in Yucatec

Transitive clauses in Yucatec consist minimally of the verb and a proclitic (glossed ERG) cross-referencing the subject of the verb.[3] Most of the time, the proclitic is preceded by an auxiliary particle or verb. The main verb in turn displays a series of suffixes (glossed ABS) that agree with the object.[4] This minimal structure is shown in (3).

(3) *K-in*   *w-il-ik-ech.*
     HAB-ERG.1s EP-see-IND-ABS.2s
     'I see you.'

    The subject and object DPs then appear to the left or right of this minimal structure, as in (1) and (2). Following Aissen (1996) and Aissen (1999a), we assume that the VP in Yucatec projects its specifier to the right. Hence we analyze the VOS order of (1) as a structure where both the subject and the object remain in their VP-internal positions, whereas for SVO we assume that the subject has been fronted from [Spec, V] to [Spec, I].

(4)

2.3 Preverbal subjects, topics and foci

As mentioned, in this paper we conclude that the unmarked word order of transitive clauses in Yucatec is SVO. Our definition of unmarked word order corresponds to the one developed in Costa (1998) and Gutiérrez-Bravo (2005). Specifically, following these works we assume that there is a set of purely syntactic constraints that determine the relative order of constituents in a language. These purely syntactic constrains in turn interact

constantly with pragmatic, semantic, intonational, and discourse constraints that have their own word order requirements. Unmarked word order is then defined as the order that results from the purely syntactic constraints in the language, whereas perturbations of this order resulting from the non-syntactic constraints are understood as derived orders.

Because of this interaction between syntactic and non-syntactic word order requirements, in order to determine the unmarked word order of Yucatec it is necessary to simultaneously take into account at least the word order perturbations that result from Information Structure considerations (topicalization and focusing). In terms of the problem we are addressing here, what is necessary is to shown that there are instances of the SVO order where the preverbal subject is not the result of focusing or topicalization. As in any other Mayan language (England (1991), Aissen (1992)), transitive subjects in Yucatec can appear in the preverbal position as the result of topicalization or focusing. Topicalized preverbal constituents (i.e. sentence topics), including the subject, typically take the topic clitic *=e'*, as in (5).

(5) *Juan=e'   t-u   ts'on-aj-ø   kéej.*
      Juan=TOP CP-ERG.3s shoot-PRF-ABS.3s deer
      '(As to) Juan, he shot a deer.'

Focused transitive subjects also appear to the left of the verb, and they further trigger the Agent Focus form of the verb (Bricker 1978), as happens in many other Mayan languages. In Yucatec, this form differs from the canonical structure of transitive clauses in (3) in that it lacks both the ergative proclitic and the auxiliary that appears to its left (Bricker (1978), Bohnemeyer (2002), Tonhauser (2003), Stiebels (2006)).[5] This is shown in (6).

(6) *Juan il-ik-ech.*
      Juan see-IND-ABS.2sg
      'JUAN sees you.'   (Tonhauser 2003: 211)

Although these two processes have been widely reported in the literature on Yucatec, it has mostly gone unobserved that preverbal transitive subjects in this language can also appear without the topic clitic or the agent focus form. This is illustrated with the preverbal subject *leti'* 'he' in (7) and *máak* 'person' (i.e. 'someone') in (8).

(7) *Leti'   t-u   ordenar-t-aj-ø   ka'a túul   saapo*
    3SG   CP-ERG.3   order-TRNS-PRF-ABS.3sg two   NUMC toad
    u   bi-s-ej-ø.
    ERG.3 go-CAUS-IRR-ABS.3sg
    'He commanded two toads to take it away (a box).'   (Sapo-1)

(8) *Pero wa máak   u   y-ojel   u   ts'ak-ø=e',*
    but if   person   ERG.3 EP-know   ERG.3 cure-ABS.3sg=TOP
    pues séeb u   ts'a'ak-al.
    then   fast   ERG.3 cure+PASS-IND
    'But if someone knows how to cure it, then it is cured fast.'
                          (Sonámbulo-26)

As such, Yucatec displays *three* different kinds of transitive preverbal subjects. This is sketched in Table 1, where the numbers refer to the example sentences that illustrate each of the preverbal subject types.

Table 1. Preverbal transitive subjects

		Morphosyntactic cue
Topicalized subject	(5)	*=e'*
Focused subject	(6)	Agent Focus Form
"Neutral" subject	(7) and (8)	None

In what follows we present evidence that the subjects labeled as "neutral" above are pragmatically and syntactically unmarked subjects. This points to the conclusion that the unmarked word order of Yucatec is SVO.

2.4 Evidence for SVO as unmarked
2.4.1.Out-of-the-blue questions ("what happened?")

It has been observed that answers to this kind of interrogative are only felicitous when they have the language's unmarked word order (Contreras (1976), Vallduví (1992), Lambrecht (1994); for Spanish see also Gutiérrez-Bravo (2005), and Gutiérrez-Bravo (2008). In order to test for this felicity contrast, different pairs of sentences (one SVO, the other one VOS) were read out loud to five native speakers (in random order with respect to

5

one another) so that they would identify which one was the most natural answer to the question in (9).

(9) *Ba'ax k-u   y-úuch-ul?*
     what   HAB-ERG.3   EP-happens-IND.
     'What's happening?'

(10) *Le   koolnáal-o' t-u   jats'-aj-ø   le   máak-o'.*
      DM   peasant-CL   CP-ERG.3 beat-PRF-ABS.3sg DM   person-CL
      'The peasant beat the man.'

(11) *T-u   jats'-aj-ø   le   máak   le   koolnáal-o'.*[6]
      ASP-ERG.3 beat-PRF-ABS.3sg DM   person   DM   peasant-CL
      Idem.

   We present the results of this test in Table 2, where it can be seen that speakers consulted expressed their preference for SVO as the answer to (9) 80% of the time. We take this to be a first piece of evidence that the unmarked word order of transitive clauses in Yucatec is SVO.

Table 2. Preference to 'Out of the blue' questions

order		percentage
SVO	8/10	80%
VOS	2/10	20%

2.4.2 Definiteness effects

Many languages display definiteness effects in their derived (i.e. marked) word orders, but not in their unmarked word order (Gutiérrez-Bravo 2008). In other words, in constructions displaying different kinds of perturbations from the unmarked word order (topicalization, object shift, existential constructions, etc.) it is not rare to find that the relevant perturbation is only tolerated if the argument displaced from its canonical position is either obligatorily definite or obligatorily indefinite. In contrast, to the best of our knowledge no such effects are observed in a language's unmarked word

6

order. Hence, the presence of definiteness effects can be taken to be an indication that the order in which it is observed is not the language's unmarked word order. For Mayan languages in particular, this can be illustrated with Tz'utujil, a Mayan language from Guatemala. Tz'utujil, like any other Mayan language, allows SVO constructions like (12).

(12) TZ'UTUJIL
    Ja Taa' ma t-uu-ya' r-paq r-xaayii.l   SVO
    the señor NEG ASP-A3SG-give A3-money A3SG-wife
    'The Señor doesn't give Money to his wife.'   (Dayley 1985: 314)

However, as noted in Aissen (1999a), it is fairly clear that SVO is a derived order in Tz'utujil and not its unmarked word order. This is because the preverbal subject needs to be definite in these cases, as shown by the contrast between (13) and (14). This definiteness restriction on preverbal subjects is an indication that the preverbal subject must be a sentence topic, and consequently that the SVO order of (12) is derived by subject topicalization.

(13) *X-in-ruu-ti'*   *jun kaab'.*   VS
    CMPL-B1SG-A3SG-eat   a wasp
    'A wasp bit me.'

(14) **Jun kaab' x-in-ruu-ti'.*   SV
    a wasp CMPL-B1SG-A3SG-eat   (Aissen 1999a: 172)

In contrast, this definiteness effect is not observed in Yucatec SVO clauses. Indefinite DPs and bare NPs in the preverbal position are accepted by speakers without hesitation even in the absence of any context, as shown by elicitation of clauses like (15). SVO clauses with indefinite DPs and bare subject NPs are also readily found in texts, as in (8) and (16). In (16), from a monolingual speaker, the preverbal NP is actually the first instantiation made in the narrative of the referent in question.

(15) *Jun túul máak t-u kíin-s-aj-ø le*
    one NUMC person CP-ERG.3 die-CAUS-PRF-ABS.3sg DM
    koolnáal-o'.
    peasant-CL
    'A person killed the peasant.'

(16) *Jun   túul   máak   te'   Peto ku   y-a'al-ik-ø*
    one   NUMC person   LOC Peto HAB-ERG.3   EP-say-IND-ABS.3sg
    teen=e'…
    1.SG=TOP
    'A man from Peto used to tell me…'   (Don Pablo-170)

To corroborate this observation further, we applied an *elicitation-by-translation* test, the results of which indeed show that the contrast *indefinite-verb-definite* of (15) does not affect the SVO interpretation of this kind of clause. The *elicitation-by-translation* test was carried out as follows: two different clauses like (15) were read to five Yucatec speakers (ten trials in total). The speakers were asked to provide a translation in Spanish for each sentence (see also Skopeteas and Verhoeven 2005). One of the trails was discarded because the subject provided an answer unrelated to the current experimental conditions (hence Table 3 below shows the results of only nine trials).

From the translation it is possible to determine if these clauses are being interpreted as SVO (*a person killed the peasant*) or OVS (*the peasant killed a person*). Crucially, in this test an OVS interpretation indicates that definiteness, and not word order, is the factor that determines which of the two DPs is interpreted as the subject/agent. Following our assumptions about the relation between definiteness effects and word order, this would point to the conclusion that the order *argument-verb-argument* is derived (i.e. a marked word order). As shown in Table 3, this was not what was observed. The preferred interpretation of this kind of clause was still SVO. This is further evidence that SVO is not a derived order.

Table 3. Preferred interpretation of indefinite-verb-definite clauses.

Interpretation		Percentage
SVO	8/9	88.88%
OVS[7]	1/9	11.11%

We further applied this same test with clauses like (17), in which both argument DPs appear after the verb (we henceforth refer to this order as *verb-argument-argument* order). In these clauses the nominal expression to the immediate right of the verb (the canonical object position) was definite,

whereas the second DP was indefinite.[8] Again, each speaker was presented with two different clauses, giving ten trials in total.

(17) *Tun   jats'-ik-ø   le   máak   jun   túul*
     DUR+ERG.3 beat-IND-ABS.3sg DM person one NUMC
     koolnáal-o'.
     peasant-CL

In principle, clauses like (17) should have two possible interpretations, 'A peasant beat the man' (i.e. VOS) or 'The man beat a peasant' (i.e. VSO). The first interpretation corresponds to what has been traditionally claimed to be the canonical word order of Yucatec. The second interpretation instead indicates that a definite post-verbal DP is preferably interpreted as an agent even if it does not occupy the canonical position of the subject/agent. Now, in sharp contrast with what is observed in Table 3, in *verb-argument-argument* clauses definiteness has a robust effect, to the extent that it thoroughly overrides a potential VOS interpretation. This is shown in Table 4.[9]

Table 4. Interpretation of *Verb-Definite-Indefinite* clauses

order		percentage
VOS	0/10	0%
VSO	10/10	100%

Under our assumption that definiteness only affects derived orders, this result points to the conclusion that verb-initial orders in YM transitive clauses are derived and not unmarked.

2.4.3 Neutralization of the preverbal position

We also carried out a test that can be described as "neutralization of the preverbal position". This was also an e*licitation by translation* test in which we "neutralized" the preverbal position by having an interrogative or relative operator occupy it, as in (18).

(18) [interrogative or relative pronoun] *verb-argument-argument*

We constructed a number of clauses with this structure (an example is presented in (19)), which were then read to native speakers. The native speakers were asked to provide a translation in Spanish of each example.

(19) *Tu'ux   t-u   jats'-aj-ø   le máak le   koolnáal-o'?*
     where CP-ERG.3 beat-PRF-ABS.3sg DM person DM peasant-CL

In principle, each clause like (19) could have two possible interpretations; '*Where did the peasant beat the man?*' (i.e. a VOS interpretation where the second DP is taken to be the agent) or '*Where did the man beat the peasant?*' (i.e. a VSO interpretation). This test was applied to corroborate one of two possible hypotheses.

(20) *Hypothesis A.* Yucatec is an SVO language. Under this hypothesis, the interpretation of which of the two postverbal DPs is the subject/agent should be affected by the presence of a third element in the preverbal position. This is because the preverbal position is the canonical subject position, but it has now been occupied by a different constituent. Yucatec is a strict head-marking language, so now there is no cue left to determine which of the two postverbal DPs in (19) is the subject. As a result, we expect the interpretation of these constructions *not* to be systematically VOS.

(21) *Hypothesis B.* Yucatec is a VOS language. In this case, the presence of a third element in the preverbal position should not affect the interpretation of clauses like (19). This is because neither the canonical subject position nor the canonical object position (both postverbal) are being affected by the presence of this third element. Interpretation is expected to be systematically VOS.

This test was applied to six speakers of Yucatec. Five different clauses like (19) were presented to each speaker, adding up to thirty trials in total. Out of these, four trails were discarded because the subjects provided answers unrelated to the current experimental conditions. Our results are presented in Table 5, where it can be seen that VOS, while being the preferred interpretation, was still clearly not the systematic interpretation of these clauses. We take this to be support for Hypothesis A and consequently

as further evidence that the unmarked word order of Yucatec transitive clauses is SVO and not VOS.

Table 5. XP-Verb-Argument-Argument clauses

Interpretation		Percentage
VOS	16/26	61.53%
VSO	10/26	38.46%

2.4.4 Text count and comparison with Tzeltal

Finally, we present evidence from a frequency count of the occurrences of SVO and VOS in four oral texts from Yucatec speakers.[10] Like all text counts of this kind, ours is also subject to the limitation that the frequency of occurrence in texts of transitive clauses with two full DP arguments is fairly low. Still, we believe that our results are significant, especially when they are subject to cross-linguistic comparison, as we do below.

The texts in our sample were segmented by clause (not sentence). After segmentation, we counted the total number of occurrences of SVO and VOS declarative clauses.[11] We present our results in Table 6, where it can be seen that the frequency of occurrence of SVO is three times higher than the frequency of VOS.

Table 6. Text Count in Yucatec oral texts 1

Order	No. of occurrences	% (*N=372*)
SVO	9	2.41
VOS	3	0.80

However, our result of the total number of SVO clauses in Table 6 includes cases where the subject has been clearly topicalized (see Table 1). In order to distinguish these cases from cases where SVO could not be attributed to topicalization, the number of occurrences of each was calculated separately. For this purpose, our criteria for identifying sentence topics were: (a) the fronted subject DP showed the topic clitic =*e'* or, (b) the subject DP was fronted outside the clause it originated in (i.e. either long topicalization had taken place, or the topic appeared to the left of an

interrogative pronoun or a complementizer). When clauses with topicalized subjects are considered as a separate category, the results are clearly more complex. In fact, as shown in Table 7, they actually reflect the original observations of Durbin and Ojeda (1978) and Hofling (1984) in that Yucatec appears to have two unmarked word orders.

Table 7. Text Count in Yucatec oral texts 2

Order	No. of occurrences	% (*N=372*)
S$_{=TOP}$VO	6	1.61
SVO	3	0.80
VOS	3	0.80

In order to shed light into this situation, we compared our results with two word order text counts from Tzeltal Mayan, a Mayan language for which there is very strong evidence that its unmarked word order is VOS (Keenan (1978), Robinson (2002)). The results of these text counts are presented below, where it can be observed that the comparative frequencies of occurrence of SVO and VOS in Yucatec and Tzeltal are considerably different. Specifically, even allowing for subject-initial constructions as the result of subject topicalization (or focusing), VOS in Tzeltal has a higher frequency than SVO, contrary to what is observed in Yucatec. We conclude from this comparison that Yucatec is not showing the behavior characteristic of true VOS languages.

Table 8. Tzeltal text count 1: Keenan (1978: 280)

Order	No. of occurrences	% (*N=288*)[12]
SVO	2	0.7
VOS	16	5.55

Table 9. Tzeltal text count 2: after Robinson (2002: 60-61)

Order	No. of occurrences	% (*N=495*)
SVO	32	6.46
VOS	67	13.53

Summing up, by themselves the results of our text count are not enough to conclude that YM is an SVO language. But when we compare them with the behavior observed in a VOS language like Tzeltal, then our results point to the conclusion that the unmarked Word order of YM is not VOS. This, added to the evidence presented in earlier subsections, leads us to conclude that the unmarked word order of Yucatec transitive clauses is SVO.

3. Split word order

However, in sharp contrast with the results in Tables 6 and 7, Skopeteas and Verhoeven (2005) had already noted that VS is much more frequent than SV in intransitive clauses. We in fact observed from the analysis of our own corpus that this was true not only of intransitive clauses, but also of transitive clauses where the object was *pro*-dropped. In order to bring intransitive and object-drop clauses together for comparative purposes we henceforth refer to these clauses as "subject-only" clauses. As mentioned, we took clauses where there is no overt object to belong to this category. However, we do not include in this definition clauses that display a sentential complement (CP, IP) instead of a DP object. Neither did we consider clauses where the subject DP occupies the postverbal position when an interrogative or relative pronoun had been fronted, since Yucatec displays obligatory inversion in these cases. Examples of "subject-only" clauses are presented below.

(22) *T-u   tuukl-aj-ø   bey-o'   le   xi'ipal-e'.*
    CP-ERG.3   think-PRF-ABS.3s   thus-CL   DM   boy-CL
    'The youngster reasoned it thus'.   (Gigante-14)

R. Gutiérrez-Bravo and J. Monforte y Madera

(23) *Chan   k   janal   to'on*.
DIM   ERG.1.pl eat   we
'We are going to eat.'   (Don Pablo-675)

(24) *Lekan lúub-uk-ø   le   ka'anal ja'…*
when   fall-IRR-ABS.3s DM   tall   water
'When the rain falls...'   (Cultivo-21)

The verb in Yucatec shows different aspectual and mood morphology depending on whether the verb is transitive or intransitive, so there is no doubt that (22) is a transitive object-drop clause, whereas (22) and (24) are intransitive clauses. The results of our text count of "subject-only" clauses are presented below:

Table 10. Text count in Yucatec oral texts 3: "subject-only" clauses

Order	No. of occurrences	% (*N=372*)
SV	17	4.57
VS	58	15.59

Hence, on the one hand we have considerable evidence that the unmarked word order of transitive clauses in Yucatec is SVO. However, in the absence of an (overt) object, it is quite clear that the most frequent order is VS. We propose that this discrepancy between transitive declaratives and "subject-only" declaratives be referred as **split word order**. Split word order should be kept distinct from the phenomenon observed in Spanish, German, Japanese, etc. whereby different kinds of clauses display different word orders depending on the semantic role of the arguments of the verb (Gutiérrez-Bravo 2007). In split word order the split/discrepancy in word order is instead (mostly) dependent on whether one or two arguments of the verb are overtly expressed.

The split word order observed in Yucatec represents a considerable analytical challenge. If the SVO order is analyzed as the result of a strong EPP requirement in a functional projection above VP, as is typically the case in SVO languages, then the analysis incorrectly predicts that the strong EPP requirement will produce an SV order in "subject-only" clauses. On the other hand, an analysis that claims that Yucatec is a verb-initial language where SVO is the result of subject topicalization (as in the analysis of Greek

and Spanish in Alexiadou and Anagnostopoulou 1998), is incompatible with the evidence presented so far that SVO is the unmarked word order of Yucatec transitive clauses. In the following section we argue that Optimality Theory, where constraints are ranked and violable, is a framework that is ideally suited to provide an analysis of this word order split.

4. An Optimality-theoretic analysis
4.1 Deriving SVO and VS

The classic analysis of SVO languages in the OT literature is developed in Samek-Lodovici (1996), Grimshaw (1997), Costa (1998) and Costa (2001). In this analysis, the SVO order results from a constraint that requires [Spec, I] to be filled (henceforth EPP, following the definition in Gutiérrez-Bravo 2007) outranking the STAY constraint (Grimshaw 1997), which penalizes the presence of traces in the syntactic representation.
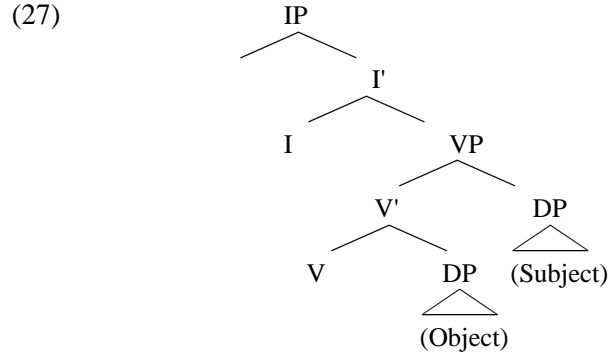
(25) EPP
      The specifier of the highest inflectional projection must be filled.

(26) STAY
      No traces

This analysis assumes that the subject DP is generated in [Spec, V], and this is the position from where it moves directly to [Spec, I]. As stated before, following standard analyses of Mayan languages like Aissen (1996) and Aissen (1999a), we assume that [Spec, V] is projected to the right in Yucatec. These assumptions are schematized in (27). Lastly, recall that I^0 in Yucatec transitive clauses is always filled in transitive clauses by an ergative proclitic and (in most cases) an auxiliary.

(27)

```
                        IP
                       /  \
                        I'
                       /  \
                      I    VP
                          /  \
                         V'    DP
                        /  \
                       V    DP   (Subject)
                           /\
                          /__\
                        (Object)
```

Here it is important to note that Yucatec also shows VSO clauses, although much less frequently than either SVO or VOS clauses. Our data shows that in VSO clauses the object is discourse-old. For instance, in (28) the object *maaya* 'Mayan' had been previously introduced in the discourse two clauses earlier.

(28) *K'abéet   a   kan-ik-ø   teech   maaya.*
      necessary ERG.2 learn-IND-ABS.3sg   2SG   Mayan
      'You need to learn Mayan.'   (Aurelia-3)

As such, we take VSO clauses be to cases where the object corresponds to the *tail* in the sense of Vallduví (1992). Hence, we assume the object is right-dislocated in these cases. With respect to the structure in (27), this means that the object is right-adjoined to IP (or perhaps VP). The result is a different structure altogether from that of VOS clauses, and so it is not targeted by the constraints that will be introduced in what follows.

Now, returning to the issue of split word order, if we analyze the SVO order of Yucatec transitive clauses as being indeed the result of the ranking EPP » STAY, we get an incorrect result for "subject-only" clauses, as illustrated in (29). This is because under the ranking EPP » STAY, the winning candidate for subject-only clauses is the one displaying an SV order.

(29) Subject only clauses = (22)

INPUT: < think (x,y), x=the boy, y=*pro* >	EPP	STAY
✘ a. [IP [DP le xi'ipale']i tu [VP tuklaj [DP *pro*] beyo' *h*i]] **SV**: *the boy thought.it thus*		*
☞ b. [IP tu [VP tuklaj [DP *pro*] beyo' [DP le xi'ipale']]] **VS**: *thought.it thus the boy*	*!	

Now observe that the opposite ranking does not give us the correct result either. Following Costa (1998) and Costa (2001), we can take the unmarked word order of verb initial languages to be the result of the ranking STAY » EPP. As the reader can verify for himself, this ranking will give the correct result for subject-only clauses, but now we get an incorrect result for transitive clauses. This is because the STAY » EPP ranking incorrectly selects the verb-initial VOS candidate as the winner. This is shown in (30).

(30) Transitive clauses = (2)

INPUT: < put (x,y), x=the woman, y=her good clothes>	STAY	EPP
✘ a. [IP tu [VP ts'aj [DP u ma'alob nook'] [DP le ko'olelo']] **VOS**: *put.on her good clothes the woman*		*
☞ b. [IP [DP le ko'olelo']i tu [VP ts'aj [DP u ma'alob nook'] *h*i] **SVO:** *the woman put.on her good clothes*	*!	

Hence, by itself the interaction between STAY and EPP is insufficient to account for the word order split observed in Yucatec. We now propose that Yucatec indeed has the ranking STAY » EPP characteristic of verb-initial languages, but that STAY is in turn outranked by a well-formedness constraint that forces the subject to move to [Spec, I].

Working on word order phenomena in Greek and Italian, Roussou and Tsimpli (2006) propose that two DPs cannot appear in the same syntactic domain if they share the same features (see also Mohanan (1994), Belletti (2004)). In their proposal, the syntactic domains in consideration are the lexical, inflectional and complementizer layers of the clause. They use this proposal to explain the absence of the VSO order in Italian. As illustrated in (31), from Roussou and Tsimpli (2006: 318), in this order both the subject and object DPs are found inside the lexical domain (which we label as *VP* below). Since in Italian these DP are identical in all of their features with overt morphological realization, this configuration is ruled out. In contrast,

17

VSO is freely allowed in Greek because overt case morphology in this language makes the subject and object DPs formally distinct.

(31) ITALIAN

    Ha   riparato$_i$ [$_{VP}$ [$_{DP}$ *Gianni*] h_i [$_{DP}$ *il   mio   computer*]].
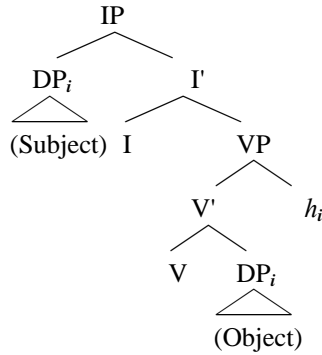   has   repaired   Gianni   the mine computer

We adopt Roussou and Tsimpli's proposal in its essentials, with the following modifications. First, it seems to us that the definition of domain in their proposal can be simplified to specific maximal projections (VP, IP, CP) if unexploded versions of these projections are adopted (as we do in the OT analysis that follows). More importantly, we further propose that the crucial features that trigger the effect in (31) are category features, ([D], [N], [V], etc.), not general feature identity.[13] Finally, following recent work where it has been argued that constituents must have phonetic content in order to meet certain well-formedness conditions (Holmberg (2000), Landau (2007)), we propose that the effect in (31) requires that the XPs that share the same category feature have both phonetic content (see also Belletti 2004). Taking these considerations into account, we arrive at the following constraint, which we name DISTINCT.

(32) DISTINCT
      Two phonetically overt phrases *XP* and *YP* with the same category
      feature [Z] should not be dominated by the same set of maximal
      projections.

Crucially, any structure where the subject is located in [Spec, I] satisfies DISTINCT. This is because both argument DPs are dominated by IP, but only the object DP is dominated by VP. In contrast, when the subject DP remains in its VP-internal position (thus deriving a verb-initial order, as in the VOS schema in (27)), DISTINCT is violated.

(33)

```
                    IP
                   /  \
                DPᵢ    I'
               /\     /  \
         (Subject)  I    VP
                        /  \
                      V'    hᵢ
                     /  \
                    V   DPᵢ
                        /\
                    (Object)
```

Having introduced the DISTINCT constraint it is now possible to account for the word order split observed in Yucatec with the ranking DISTINCT » STAY » EPP. The analysis of "subject-only" clauses under this ranking is presented in (34).

(34) "Subject-only" clauses = (22)

INPUT: < think (x,y), x=the boy, y=*pro* >	DIST	STAY	EPP
☞ a. [IP tu [VP tuklaj *pro* beyo' le xi'ipale']] **VS:** *thought.it thus the boy*			*
b. [IP le xi'ipale'ᵢ tu [VP tuklaj *pro* beyo' **hᵢ**]] **SV:** *the boy thought.it thus*		*!	

First observe that, even though the clause under consideration is transitive, both candidates satisfy DISTINCT. This is because the direct object is realized as a *pro*, which is not phonetically overt. Since both candidates satisfy the highest-ranked constraint, movement of the subject to [Spec, I] does not improve the structure in any respect. Hence it gratuitously violates STAY, and this makes SV candidate (34) suboptimal when compared to the VS candidate (34). In this way, the VS order of "subject-only" clauses is derived.

However, a different situation results when both argument DPs are overt. In this case there is no default compliance of DISTINCT, and now this constraint plays a crucial role in selecting the winning candidate, as shown in tableau (35).

(35) Clauses with overt subject and object = (2)

INPUT: < put (x,y), x=the woman, y=her good clothes>	DIST	STAY	EPP
a. [IP tu [VP ts'aj [u ma'alob nook'] [le ko'olelo']]] **VOS:** *put.on her good clothes the woman*	*!		*
☞ b. [IP le ko'olelo'i tu [VP ts'aj u ma'alob nook' *h*i] **SVO:** *the woman put.on her good clothes*		*	

In this case, the verb initial candidate (35) fatally violates DISTINCT because both DPs are dominated by the same maximal projections (VP and IP). In contrast, the candidate where the subject surfaces in [Spec, I] satisfies DISTINCT, and so the SVO structure emerges as the winner. In this way, our analysis accounts for the word order split observed in Yucatec.

Observe that our analysis further makes the prediction that when the complement of the verb is not a DP (but instead a PP or a CP), the unmarked word order should be verb-initial.[14] Testing this prediction fully requires a corpus much larger than the one used for this paper. However, the data we do have does confirm this prediction. All else being equal, the subject is postverbal when the complement of the verb is not a DP, as shown in (36) and (37).

(36) *Ken   k'uch-uk-ø*   [DP *jun p'éel lu'umkab*] *bey* [PP *te'*
    when   arrive-IRR-ABS.3sg one NUMC mortal   thus   LOC
    t-u   kaajtal] *bey-o'...*
    DUR-ERG.3 dwell   thus-CL
    'When a mortal arrives to where he (the giant) dwells...' (Gigante-32)

(37) *K-u   taal   teech* [DP *k'ank'ubul*][PP *ti'   a   nal*]
    HAB-ERG.3 come 2SG   plague   PREP ERG.2 corn
    tun   jóok'-ol   u   y-i'ij-o'ob-o'.
    DUR+ERG.3 exit-IND   ERG.3 EP-corn.bud-PLUR-CL
    'Plague comes (on you) to your corn as its tips are coming out.'
                                          (Cultivo-75)

More importantly, the two instances in the corpus where the complement is not a DP and the subject is instead preverbal are clearly cases of topicalization of the subject DP. In (38), the preverbal subject bears the topic clitic *=e'*, and in (39) it appears to the left of the complementizer *ka'*, which indicates that it has been displaced to a topic position (Aissen 1992).

(38) [_{DP} *Lu'umkab-e'*] *chéen   mix   u   y-ojel*   [_{CP} *bix*
       mortal-TOP   only   neither   ERG.3   EP-know   how
    ko'ol-ik-ø-i'].
    clear+PASS-FOC-*ABS.3sg*-LOC
    'The mortal didn't even know how (the jungle) had been cleared.'
                                    (Gigante-83)

(39) [_{DP} *Le chan   lu'umkab Isaac*]-*o',   ka'   j-k'uch-ø-i'*
       DM small   mortal   Isaac-CL   COMP CP-arrive-ABS.3sg-LOC
    ka'   t-u   y-e'es   u   baj [_{PP} *ti'   le   rey-o'*].
    COMP CP-ERG.3   EP-show ERG.3 RFLX   PREP DM   king-CL
    'The little mortal Isaac, when he arrived there, he introduced himself
    to the king.'   (Gigante-35)

   It is now worth addressing two theoretical issues that make our optimality theoretic analysis preferable to possible alternatives that do not make use of OT. First, observe that DISTINCT only states that two XPs with the same category feature should not be dominated by the same set of maximal projections. As such, DISTINCT does not specifically derive movement of the subject DP to [Spec, I]. DISTINCT would equally be satisfied if instead the object was moved to [Spec, I], or if the subject moved to *any* position outside the VP. Our OT analysis accounts for the fact that neither option is observed.

   With respect to the possibility of the object being the DP displaced to [Spec, I], we adopt the OT analysis of unmarked word order developed in Gutiérrez-Bravo (2007). This analysis claims that the highest inflectional specifier of the clause is sensitive to the semantic role of the XP that occupies it. Specifically, structures where this specifier position is occupied by an *agent* are less marked (all else being equal) than clauses where a *patient* or *theme* occupies this position. As such, a structure where the object is displaced to [Spec, I] to comply with DISTINCT will always be more marked than a structure where the subject-*agent* is displaced to this position. As we leave for the reader to verify, an OVS candidate that satisfies DISTINCT will consequently always be eliminated because of its more severe violation of the relevant markedness constraints proposed in Gutiérrez-Bravo (2007).

    Secondly, we mentioned that in principle, the subject could move to any position outside the VP and still satisfy DISTINCT. However, this is clearly

not what is observed. For instance, although particles, adverbs and some adverbial XPs can appear between the main verb in V and the auxiliary in I^0 in Yucatec, insertion of the subject DP in this intermediate position is robustly ungrammatical.

(40) *[IP *tu*   [*le   ko'olelo'*]i [VP *ts'-aj-ø   u   ma'alob*
          CP-ERG.3   DM   woman-CL   put-PRF-ABS.3sg   ERG.3   good
          *nook'   h*i]].
          clothes

   Our OT analysis provides a straightforward account of this fact. Even though the EPP is at the bottom of the ranking in Yucatec, it is still an active constraint (as in any OT analysis). The structure in (40) satisfies DISTINCT, but it violates EPP because it leaves the highest inflectional specifier empty. Because of this, it is bound to lose against the S Aux VO candidate that satisfies both DISTINCT and EPP. This is shown in the tableau in (41), where the losing candidate (41a) corresponds to ungrammatical (40). Observe that there is no straightforward way to achieve this result in an alternative analysis that made use of DISTINCT, but where the EPP was not taken to be an active requirement in the language.

(41) Active EPP

INPUT: < put (x,y), x=the woman, y=her good clothes>	DIST	STAY	EPP
a. [IP tu [VP [le ko'olelo']i [VP ts'aj [u ma'alob AUX *the woman* *put.on her good* nook'] *h*i]]] *clothes*		*	*!

☞ b. [IP le ko'olelo'i tu   [VP ts'aj [u ma'alob nook']
      the woman   AUX   *put.on   her good clothes*
    *h*i]]
		*	

   Finally, DISTINCT is itself a violable constraint, and consequently we expect its effects not to be observed when it is dominated by another constraint. We suggest that this is the case of true VOS languages like Tzeltal. Specifically, VOS as the unmarked word order is derived with the ranking STAY » DISTINCT, EPP. This is shown for the example in (42) in the tableau in (43).

(42)   TZELTAL (Stross 1978:27; taken from Robinson 2002:61)
  La laj [_{VP} *s-ta-ø*   [_{NP} *alchaxiltik*]   [_{NP} *te   winik-e*]].
  CP HS   3ERG-find-3ABS   orange.orchard   ART man-CL
  '… the man found an orange orchard.'

(43) TZELTAL

INPUT: <find (x,y), x=the man, y=orange.orchard>	STAY	DIST	EPP
a. [_{IP} te winike_i la laj [_{VP} sta alchaxiltik **h_i**]] **SVO:** *the man found orange.orchard*	*!		
☞ b. [_{IP} la laj [_{FV} sta alchaxiltik te winike]] **VOS:** *found orange.orchard the man*		*	*

   In this way, our proposal is compatible with the analysis of verb-initial languages in Costa (1998) and Costa (2001), where it is proposed that what is characteristic of these languages is the very high ranking of STAY relative to the EPP.

4.2 Back to VOS clauses

It now only remains to be explained what derives VOS clauses like (1). In this final section, we briefly summarize the analysis of Yucatec VOS clauses developed in Gutiérrez-Bravo and Monforte (2008). Following the preceding sections of this paper, we conclude that VOS is a derived order, and specifically, we now propose that VOS is the order that corresponds to **thetic judgments**. In this we follow Kuroda (1972), Aissen (1999a), and Ladusaw (1994) in assuming a fundamental semantic difference between categorical and thetic judgments. Categorical judgments establish a predication relation between an argument (typically the subject) and the rest of the proposition. Categorical judgments are composed of two parts: a referent of which something is predicated (the logical subject), and the property that is ascribed to this referent. Following Aissen (1999a), we consider subject-initial constructions to prototypically instantiate this kind of judgment. Specifically, in an SVO clause like (2), repeated here as (44), the preverbal DP is not only the grammatical subject of the clause. It is also the logical subject, the entity of which the rest of the clause is predicated.

23

Hence subject-initial clauses can be taken as prototypical instances of categorical judgements.

(44) *Le ko'olel-o' t-u ts'-aj-ø u ma'alob nook'.*
DM woman-CL CP-ERG.3 put-PRF-ABS.3sg ERG.3 good clothes
'...and the woman put on her good clothes...' (Si'ipil-51)

In contrast with categorical judgments, thetic judgments present an event in a unitary way. The event is just presented as the simple perception of the situation, without highlighting any of the participants and without establishing the *logical subject-predicate* configuration of categorical judgments. Again, following Aissen (1999a), we take verb-initial constructions to be a typical instantiation of thetic judgments. Specifically, in a VOS clause like (1), repeated here as (45), there is no logical subject. In other words, in contrast with (44), there is no unique constituent of which the rest of the clause is predicated. What we mean by this is that there is no single constituent that the clause can be claimed to be about. Instead it appears that the event is merely being presented as the simple perception of the situation.

(45) *Je' bin k-u la'ach-ik-ø u jo'ol*
ASV CIT HAB-ERG.3 scratch-IND-ABS.3sg ERG.3 head
x-nuk reyna-o'.
FEM-great queen-CL
'And the great queen truly scratched her head...' (Gigante-97)

It does not seem to us that this particular state of affairs is specific to Yucatec. As discussed in Gutiérrez-Bravo and Monforte (2008), we believe that this same phenomenon is what is observed in a well-studied kind of word order alternation, namely, Transitive Expletive Constructions (TECs). TECs show a word order alternation that is very similar to the one observed in Yucatec (see Bobaljik and Jonas 1996, Alexiadou and Anagnostopoulou 1998). This is illustrated with examples from Icelandic in (46), from Alexiadou and Anagnostopoulou 1998:492).

(46) ICELANDIC

    a. *Einhverjir stúdentar lasu bókina.*   **SVO**
        some   students   read the book

    b. *Þađ   lasu* [*einhverjir stúdentar*] [*bókina*].   **Exp-VSO**
        EXPL read   some   students   the book
        'Some students read the book.'

Example (43b) is not strictly speaking a verb-initial construction. However, as argued by Alexiadou and Anagnostopoulou, this is merely a parametric difference, relative to the need to satisfy the EPP with a full XP. Greek, for instance, shows essentially the same SVO/VSO alternation, but no expletives. Thus it is feasible to equate the semantics of Icelandic TECs with Yucatec VOS clauses, given that the expletive adds no semantic content to the proposition. Now observe that, just as in the case of the Yucatec SVO/VOS alternation, the word order alternation in Icelandic TECs does not bring with it any truth-conditional difference.[15] As proposed in Gutiérrez-Bravo and Monforte (2008), TECs and VOS clauses in Yucatec appear to share a fundamental property. Whereas in the subject-initial constructions the preverbal subject functions as the target of predication and thus we have a categorical judgment, in the constructions where the subject appears in the post-verbal field there is no target of predication whatsoever. As mentioned previously with respect to Yucatec, TECs have no constituent that the clause can be understood to be about. Instead they simply present the content of the proposition in a unitary way. In this way, our claim that VOS in Yucatec is a marked order coexisting with unmarked SVO is really no different from the situation observed in Icelandic.

In order to integrate the distinction between categorical and thetic judgments into the OT analysis developed so far, we propose that the thetic nature of a proposition is established in the semantic information of the input (Gutiérrez-Bravo 2005). Hence, transitive inputs like (35) can further be specified with a [thetic] feature ([TH]):

(47) INPUT: < **[TH]** scratch (x,y), x=the great queen, y=her head >

The presence of the [thetic] feature in the input implies that the syntactic output should be interpreted as a thetic judgment and not as a *logical subject-predicate* configuration. Since subject-initial constructions are typically interpreted as *logical subject-predicate* configurations, we propose that the manifestation of the [TH] feature in output syntax is regulated by the following constraint.[16]

(48) THETIC-CONDITION (THET-CON)
    If a proposition bears the [thetic] feature in the input, then the
    highest inflectional head asymmetrically c-commands all the
    arguments of the lexical head of the clause.

Observe that this restriction is satisfied by a verb-initial clause like (1), since I^0 asymmetrically c-commands both of the arguments of the verb. In contrast, it is violated when the subject occupies the [Spec, I] position (as in SVO clauses) since in this case I^0 does not c-command the subject. Because of this THET-CON, necessarily conflicts with the requirements of DISTINCT, which forces movement of the subject to [Spec, I] when an object is also present.[17] We now account for the word order split observed in Yucatec through the interaction of these constraints.

When THET-CON outranks DISTINCT, if the input bears the [thetic] feature, the requirement that there should be no argument that is not c-commanded by I^0 has priority over the requirement that there shouldn't be two DPs dominated by the same set of maximal projections. In this case, in contrast with what is observed in (35), it is the verb initial candidate that emerges as the winner.

(49) Yucatec: thetic judgments/VOS clauses

INPUT: <[Tн] scratch (x,y), x=the great queen, y= her head>	THET-CON	DIST	STAY	EPP
a. [IP x-nuk reynao' k-u [VP la'achik *the great queen scratched* u jo'ol]] *her head*	*!		*	
☞ b. [IP __ k-u [VP la'achik u jo'ol *scratched her head* x-nuk reynao']] *the great queen*		*		*

In this way we account for the presence of VOS clauses in Yucatec even when the unmarked word order of its transitive clauses is SVO.[18]

5. Conclusions

In this paper, we have provided evidence that the unmarked word order of transitive clauses in Yucatec Maya is SVO, contrary to what is often assumed in the literature. We applied four tests to determine whether the unmarked word order of Yucatec is SVO or VOS (answers to *out-of-the blue* questions, definiteness effects in *argument-verb-argument* clauses, definiteness effects in *verb-argument-argument* clauses, and interpretation of *XP-verb-definite.NP-definite.NP* clauses). The results of none of these tests points to the conclusion that the unmarked word order of YM transitive clauses is VOS. In contrast, the first two provide strong evidence that the unmarked word order of Yucatec is SVO. We also carried out a text count to compare the relative frequency of occurrence of SVO and VOS clauses. Although our frequency results in text counts shed no light on whether the basic word order of Yucatec is SVO or VOS, a comparison of these results with Tzeltal indicates that Yucatec does not behave like a VOS language.

A frequency count of "subject-only" clauses, however, indicates that these clauses are by and large verb-initial. We have proposed the descriptive term **split word order** to refer to this discrepancy between the two different kinds of clauses considered. We then developed an Optimality Theoretic analysis to account for this word order discrepancy. We first showed that Yucatec is different from other SVO languages in that its unmarked word order cannot be accounted for as the result of the ranking EPP » STAY. Instead we proposed that the verb-initial nature of "subject-only" clauses in Yucatec results from the ranking STAY » EPP. The SVO order, we argued, is the result of an altogether different constraint outranking STAY in this ranking. We proposed that this constraint is the DISTINCT constraint, which states that two phonetically overt XPs with the same category feature [Z] should not be dominated by the same set of maximal projections. The resulting analysis, which crucially relies on the violable nature of the syntactic constraints involved, provides a straightforward account of the split word order observed in Yucatec.

Notes

[1] All examples are presented according to the orthographic conventions of the *Academia de la Lengua Maya de Yucatán* and so they do not necessarily reflect their phonetic form accurately. The abbreviations we use in the examples in this paper are the following:

ABS	absolutive	DUR	durative	IRR	irrealis
ART	definite article	EP	epenthesis	LOC	locative
ASV	assurative	ERG	ergative	NUMC	classifier
CAUS	causative	FEM	feminine (biological)	PASS	passive
CIT	reportative	FOC	focus	PRF	perfect
CL	clitic	HAB	habitual	TOP	topic
CP	completive	HS	hearsay	TRNS	transitive
DM	demonstrative	IND	indicative		

[2] In Gutiérrez-Bravo and Monforte (2008) we presented a comparison between the corpus counts made for Yucatec in Skopeteas and Verhoeven (2005) and the text count made for Tzeltal in Keenan (1978). There we argued that the results of this comparison were problematic for the analysis of VOS as unmarked in Yucatec. In section 2.4.4 of this paper we make use of the same argument, but we now resort to our own corpus count. We include an additional comparison with a second word order frequency count in Tzeltal, from Robinson (2002).

[3] In terms or traditional Mayan linguistics, these proclitics correspond to the *Series A* pronouns. Although glossed as ERG here, it should be noted that ergativity in Yucatec is split on the basis of aspect. Hence not every instantiation of these pronouns corresponds to an ergative function.

[4] These correspond to the *Series B* pronouns of traditional Mayan linguistics. Again, because of split ergativity, these suffixes do not always display an absolutive function.

[5] Specifically, the third person proclitic *u* and the completive aspect particle *t-* in (3), respectively. Agent focus phenomena in Mayan languages is most often classified depending on whether the agent focus form of the verb agrees with the subject or the object. In Yucatec, however, the verb agrees systematically with the object, and subject agreement is only observed in intransitive constructions in completive aspect or irrealis mood (which follow an ergative-absolutive pattern). Hence, perhaps the most accurate description of the agent focus construction in (6) is that it is a one-place predicate displaying transitive morphology, which further

always agrees with the object, and in which the proclitic that cross-references the subject is always absent. See Gutiérrez-Bravo and Monforte (2009).

[6] The clitic *–o'* attached to the right of the arguments of the verb in these examples is a deictic clitic that typically expresses the relative distance between the speaker and the referent of each DP; this particular clitic can be approximately translated into English as the demonstrative *that*. Yucatec does not allow more than one of these clitics to appear simultaneously in the postverbal field. Hence in (11) only the second argument DP bears this clitic.

[7] It is important to note that, under the correct pragmatic conditions, OVS is a perfectly acceptable (albeit marked) order in Yucatec. These pragmatic conditions, however, are not relevant for our study, which concentrates solely on the language's unmarked word order.

[8] Observe that in VOS analyses of Yucatec this second DP thus occupies the canonical subject position and hence is expected *not* to be subject to definiteness effects, contrary to what our results indicate.

[9] In this paper we do not attempt to provide an analysis of the definiteness effects shown in Table 4. This is because VSO is a marked order in Yucatec (see section 4.1) and the scope of this paper is restricted to the analysis of unmarked word order. However, it seems to us that a possible analysis of these definiteness effects can be developed in which definiteness is aligned with the subject grammatical relation, along the lines of the OT analysis in Aissen (1999b).

[10] All these texts correspond to the "eastern" variety of Yucatec, spoken in the state of Quintana Roo and the southern tip of the state of Yucatan.

[11] We restricted ourselves to declarative clauses since *wh*-interrogatives and relatives trigger obligatory subject inversion in Yucatec, hence forcing a VS order. We do not know if this variable was controlled in earlier texts counts of Yucatec or the in the Tzeltal text counts we discuss later.

[12] In Keenan's (1978) and Robinson's (2002) text counts, *N* corresponds to the number of sentences, not of clauses. In this respect their text counts are different from ours.

[13] Observe that we are not claiming that identical category features are the only features that trigger the effect in (31). What we propose is that identical category features are a *necessary* condition for this effect to occur in the first place. We will see that in Yucatec it is also a *sufficient* condition for this effect to occur. Our sense, however, is that in other languages (including other Mayan languages) this might not be by itself a sufficient condition. It is also fairly clear from our data that full NPs and pronouns in Yucatec do not belong to the same category for the purposes of the word order effects reported here. Further research is necessary to clarify this point.

[14] We are thankful to an anonymous reviewer for bringing this point to our attention.

29

[15] TECs are subject to the requirement that the object be indefinite, as it is well known. In this respect it seems to us to be very meaningful that various kinds of different word orders in Mayan languages are subject to definiteness restrictions (see England 1991 for a survey). We acknowledge the importance of studying this particular property in VOS clauses in Yucatec, but addressing this specific issue is beyond the scope of this paper.

[16] See Gutiérrez-Bravo (2005) for an earlier formulation.

[17] THET-CON also necessarily conflicts with the EPP constraint (Gutiérrez-Bravo 2005, Gutiérrez-Bravo and Monforte 2008), but this is not relevant in the case of Yucatec because of the low ranking of EPP in this language.

[18] A reviewer asks whether it would be possible instead to account for the SVO/VOS alternation with a stochastic OT analysis. At this point we cannot claim for certain whether such an analysis would be desirable or not. However, there is a reason why we have not attempted to develop this kind of analysis. In a stochastic analysis, the same kinds of clauses are predicted to freely alternate between the SVO and VOS orders. However, we have reason to believe that this is not what is observed in Yucatec. Our data indicate that there is a marked tendency for either the subject or the object to be a bare NP in VOS clauses, but not in SVO clauses. In principle, this could be the result of a semantic or syntactic restriction on VOS clauses that does not apply in SVO clauses, which would make the stochastic analysis undesirable. A full confirmation of this observation, however, requires a much more detailed corpus study than the one we have developed in this paper.

References

Aissen, Judith
   1992   Topic and Focus in Mayan. *Language* 68:43-80.
   1996   Pied-piping, abstract agreement and functional projections in Tzotztil. *Natural Language and Linguistic Theory* 14:447-491.
   1999a   External possessor and logical subject in Tz'utujil. In *External Possession*, Doris L. Payne and Immanuel Barshi (eds.), 167-193. Amsterdam/Philadelphia: John Benjamins.
   1999b   Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17:673-711.
Alexiadou, Artemis, and Elena Anagnostopoulou
   1998   Parametrizing AGR: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16:491-539.
Belletti, Adriana
   2004   Aspects of the Low IP Area. In *The Structure of CP and IP*, Luigi Rizzi (ed.), 16-51. New York: Oxford University Press.
Bobaljik, Jonathan, and Dianne Jonas

1996　Subject positions and the roles of TP. *Linguistic Inquiry* 27; 195-236.

Bohnemeyer, Jürgen
2002　*The Grammar of Time Reference in Yukatek Maya*. Munich: Lincom Europa.

Briceño, Fidencio
2002　Topicalización, enfoque, énfasis y adelantamiento en el maya yukateco. In *La organización social entre los mayas prehispánicos, coloniales y modernos*, V. Tiesler Blos, R. Cobos and M. Greene Robertson (eds.), 374-387. Mexico City/Mérida: INAH/UADY.

Bricker, Victoria
1978　*Wh*-questions, relativization and clefting in Yucatec Maya. In *Papers in Mayan Linguistics*, Laura Martin (ed.), 109-139. Columbia, Missouri: Lucas Brothers.

Contreras, Heles
1976　*A Theory of Word Order with Special Reference to Spanish*. Amsterdam: North Holland.

Costa, João
1998　Word order variation: A constraint-based approach, University of Leiden: Ph.D. Dissertation.

Costa, João
2001　The emergence of unmarked word order. In *Optimality-theoretic Syntax*, G. Legendre, J. Grimshaw and S. Vikner (eds.), 171-204. Cambridge, Mass.: MIT Press.

Dayley, Jon P.
1985　*Tzutujil Grammar* . Berkeley, CA: University of California Press.

Durbin, Marshall, and Fernando Ojeda
1978　Basic word order in Yucatec Maya. In *Papers in Mayan Linguistics*, Nora England (ed.), 69-77. Columbia: University of Missouri.

England, Nora
1991　Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57:446-486.

Grimshaw, Jane
1997　Projection, heads and optimality. *Linguistic Inquiry* 28:373-422.

Gutiérrez-Bravo, Rodrigo
2005　*Structural Markedness and Syntactic Structure*. New York: Routledge/Taylor and Francis.
2007　Prominence scales and unmarked word order in Spanish. *Natural Language and Linguistic Theory* 25:235-271.
2008　La identificación de los tópicos y los focos. *Nueva Revista de Filología Hispánica* 56: 363-401.

Gutiérrez-Bravo, Rodrigo, and Jorge Monforte

    2008   La alternancia sujeto-inicial/verbo-inicial y la Teoría de Optimidad. In *Teoría de Optimidad: Estudios de Sintaxis y Fonología*, R. Gutiérrez-Bravo and E. Herrera (eds.), 61-99. Mexico City: El Colegio de México.

    2009   Focus, agent focus and relative clauses in Yucatec Maya. In *New Perspectives in Mayan Linguistics*, H. Avelino, J. Coon and E. Norcliffe (eds.), 83-96. Cambridge, Mass.: MIT Working Papers in Linguistics.

Hofling, Charles

    1984   On proto-Yucatecan word order. *Journal of Mayan Linguistics* 2:35-64.

Holmberg, Anders

    2000   Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31:445-483.

Keenan, Edward

    1978   The syntax of subject-final languages. In *Syntactic Typology*, Winfred Lehmann (ed.), 267-327. Austin: University of Texas Press.

Kuroda, S. Yuki

    1972   The categorical and the thetic judgment. *Foundations of Language* 9:153-185.

Ladusaw, William

    1994   Thetic and categorical, state and individual, weak and strong. In *Proceedings of the Conference on Semantics and Linguistic Theory 4*, 220-229. Ithaca, New York: Cornell University.

Lambrecht, Knud

    1994   *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Landau, Idan

    2007   EPP Extensions. *Linguistic Inquiry* 38:485-524.

Mohanan, Tara

    1994   Case OCP: A constraint on word order in Hindi In *Theoretical Perspectives on Word Order in South Asian languages*, M. Butt, T.H. King and G. Ramchand (eds.), 185-216. Stanford: CSLI.

Robinson, Stuart

    2002   Constituent order in Tenejapa Tzeltal. *International Journal of American Linguistics* 68:51-80.

Roussou, Anna, and Ianthi-Maria Tsimpli

    2006   On Greek VSO again! *Journal of Linguistics* 42:317-354.

Samek-Lodovici, Vieri

1996   Constraints on Subjects: An Optimality Theoretic Analysis: Ph.D. Dissertation, Rutgers University.

Skopeteas, Stavros, and Elisabeth Verhoeven
2005   Postverbal argument order in Yucatec Maya. *Sprachtypologie und Universalienforschung* 58:347-373.

Stiebels, Barbara
2006   Agent focus in Mayan languages. *Natural Language and Linguistic Theory* 24:501-570.

Stross, Brian
1978   *Tzeltal Tales of Demons and Monsters.* Museum Brief 24. Columbia: Museum of Anthropology of Missouri.

Tonhauser, Judith
2003   F-constructions in Yucatec Maya. In *The Proceedings of SULA 2*, Jan Anderssen, Paula Menéndez-Benito and Adam Werle (eds.), 203-223. Amherst: GLSA, University of Massachusetts.

Vallduví, Enric
1992   *The Informational Component*. New York/London: Garland.