

# **Left peripheral arguments and discourse interface strategies in Yucatec Maya<sup>1</sup>**

Stavros Skopeteas and Elisabeth Verhoeven  
University of Potsdam, University of Bremen

## **Abstract**

Constituents in the left periphery are often assumed to bear information structural functions such as topic and focus. Yucatec Maya provides the empirical basis for a challenging case study in this respect, since it provides a distinction between a sentence-initial position that is characterized by a series of enclitics and is labeled ‘topic position’, and an immediately preverbal position that is labeled ‘focus position’. This paper addresses the issue where do the interpretational properties of the left peripheral constituents come from and considers two alternative hypotheses: (a) the left peripheral constituents occupy the Specifier positions of functional projections that bear information structural features such as ‘topic’ and ‘focus’ and (b) the syntactic positions in the left periphery are underspecified with respect to information structure. The data presented in this paper support the view of hypothesis (b) and show that the interpretational properties of the left peripheral positions can be accounted for through the interaction of discourse principles that are independent from syntax with the properties of prosodic phrasing, that indirectly refer to constituent structure.

---

<sup>1</sup> We are grateful to Artemis Alexiadou, Gisbert Fanselow, Caroline Féry, Ivona Kučerová, Frank Kügler, Christian Lehmann, Ad Neeleman and an anonymous reviewer for their comments concerning the grammar of Yucatec Maya and the interaction between information structure and syntax/phonology. A previous version of this paper was presented at the UCL Workshop on Information Structure, 13-15 September 2008; we thank the organizers and the audience for helpful comments. The empirical study on Yucatec Maya as well as the field work were supported by the research center 632 Information Structure, project D2, at the University of Potsdam for Stavros Skopeteas and by the University of Bremen, project 10/853/05, for Elisabeth Verhoeven.

## **1. Preliminaries**

### 1.1 Theoretical considerations

It has been observed for several languages that constituents in the left periphery bear particular information structural functions such as ‘topic’ and ‘focus’. On the basis of this observation, some accounts assume that these constituents occupy the Specifier positions of functional projections such as *TopP* (topic phrase) and *FocP* (focus phrase) that are dedicated to particular information structural functions (see Rizzi 1997, É. Kiss 1998 among others). This view is summarized in (1).

(1) Discourse configurational hypothesis

The information structural properties of constituents in the left periphery result from the fact that particular structural configurations are associated with information structural concepts.

Recent literature on information structure challenges the isomorphic view between constituent structure and information structure in (1). Proponents of the ‘discourse underspecification hypothesis’ claim that the alleged association is the epiphenomenal result of interface strategies that relate to properties of the linearization and/or the prosodic structure and conclude that reflexes of information structure do not arise from formal features that are inherent in the left peripheral positions, as outlined in (2) (see Fanselow 2006, Fanselow and Lenertová 2008, Hartmann and Zimmermann 2006, Wedgwood 2003, Zimmermann 2008).

(2) Discourse underspecification hypothesis

The information structural properties of constituents in the left periphery result from interface strategies that relate to properties of the linearization and the prosodic structure.

The two hypotheses make different predictions with respect to the systematic occurrence of the information structural properties with particular structural

configurations. In terms of the discourse configurational hypothesis, a formal feature  $F$  is associated with the Head of a functional projection. When a constituent in the numeration bears the same feature  $F$ , e.g. <focus>, then it undergoes a syntactic operation that leads the feature-bearing constituent to the position that allows the matching of the feature at issue with the corresponding functional Head, i.e., the Specifier of the *FocP*. The empirical consequence of this view is that *only* those constituents that bear exactly the feature  $F$  are expected to occur in the position at issue, hence the possibility of constituents that do not bear the critical feature  $F$  to undergo the same syntactic operation is evidence against the fundamental assumptions of the discourse configurational hypothesis (in this vein, Wedgwood et al. 2006 evaluate counterexamples in texts as evidence against the assumption of a *FocP* in Hungarian). Furthermore, in a straightforward application of the discourse configurational hypothesis the association between syntactic position and information structure should be bi-unique, i.e., *all* constituents that bear the feature  $F$  are expected to undergo the operation that leads to the satisfaction of the feature-matching condition in the grammar. The possibility of constituents that bear the critical feature  $F$  to appear in situ may be formally accommodated through additional assumptions (such as a long-distance feature-checking option that renders the syntactic operation optional), but such assumptions reduce the predictive power of the hypothesis that syntactic operations are triggered by feature-matching conditions. The bi-unique association of syntactic operations and pragmatic functions is not implied by the discourse underspecification hypothesis. In this view, syntactic rules are not strictly determined by particular pragmatic or semantic functions, but are rather abstract operations that create an array of alternative linearizations. The linear options that are generated by the syntactic component may be realized through different prosodic structures and these options are selected in discourse in order to satisfy information-structural requirements that hold in an independent component of grammar. This view does not exclude that more than one option may be available for the same information structural configuration nor that the same structural option may occur with alternative information structural functions. Asymmetries in the preference for particular syntactic operations to occur with particular information structural properties may be

accounted for through constraints in the prosodic structure or the interaction of particular formal possibilities with information structural principles.

Furthermore, the two hypotheses are based on different assumptions concerning the involved layers of grammar. The discourse configurational hypothesis in (1) implies that the relation between syntax and information structure is not further decomposable. ‘Topic’ and ‘focus’ are inherent parts of particular positions in the hierarchical constituent structure. In terms of the discourse underspecification hypothesis, these information structural concepts refer to particular properties of the phonological output, see for instance the derivation of focus interpretation from prosodic prominence in Chomsky (1971) and Reinhart (2006: 125-163). Since the possibilities of linear orders and p-phrasing are determined by the hierarchical structure, this does not exclude that some structural possibilities may systematically occur with particular information structural properties. The empirical question here is whether information structural properties may be accounted for by means of the interaction between the phonological output and independent principles of the interface. The background assumption is the criterion of simplicity in grammatical explanations: if information structural properties may be predicted by independent properties of the grammar, then the assumption of functional Heads for Topic and Focus in the constituent structure is an unnecessary complication.

In sum, to the extent that the choice between the discourse configurational and the discourse underspecification hypothesis is an empirical question, two criteria should be examined:

(3) (a) The criterion of bi-unique association

Is the assumed information structural concept a *necessary* and *sufficient* condition for the syntactic operation at issue to take place?

If it is not a necessary condition, the discourse configurationality hypothesis does not make the right predictions. If it is not a sufficient condition, the discourse configurationality hypothesis needs additional assumptions.

(b) The criterion of independent motivation

Is it possible to account for the structure-to-discourse correlation by means of properties of the phonological output and independent conditions at the interface?

If yes, the discourse configurability hypothesis is an unnecessary stipulation.

## 1.2 Yucatec Maya

Mayan languages are particularly interesting with respect to the research question introduced in Section 1.1 because they distinguish between two types of left peripheral constituents that correlate with the expression of topic and focus. In this article, we are dealing with Yucatec Maya, which is a Mayan language spoken in the Mexican states of Yucatán, Quintana Roo, and Campeche, as well as in neighboring parts of Belize and Guatemala (700,000 speakers according to the 1990 census).

The basic facts are introduced in the following examples. *NPs* are not marked for case and relational affixes are attached to the Head (see Lehmann 1990, 1998). With transitive verbs, the so-called set-A clitic refers to the agent constituent and the set-B suffix refers to the patient constituent (with intransitive verbs, the sole argument shows a split marking depending on aspect/mood, see Bohnemeyer 2004). Since there are a few constructions which neutralize the agent of a transitive verb and the sole argument of an intransitive verb, but no construction neutralizing the patient of a transitive verb and the sole argument of an intransitive verb (see Verhoeven 2007, sect. 4.3.3), we refer to the agent constituent as a subject.<sup>2</sup>

Example (4) illustrates the canonical word order, which is VOS (see Durbin and Ojeda 1978, Hofling 1984, Skopeteas and Verhoeven 2009a). Deviations from this order

---

<sup>2</sup> Contrary to Bohnemeyer (2008), we do not assume that the agent focus construction provides evidence for a restricted neutralization of the patient (P) of a transitive verb and the only argument of an intransitive verb (S) in the sense of Van Valin and LaPolla (1997). Since other roles such as recipients, goals, instruments, etc. occur in the same focus construction type as P and S, the respective construction does not identify a syntactic pivot.

of postverbal arguments are possible. Heavy objects are often shifted to the right periphery resulting in a VSO order. Evidence from language comprehension shows that asymmetries in animacy and definiteness may influence the interpretation of two postverbal *NPs* as OS or SO (Skopeteas and Verhoeven 2005, see furthermore Gutiérrez Bravo and Monforte y Madera 2008b, Bohnemeyer 2008).

- (4) *t-u*        *hàant-ah*                    *ha's*    *huan*.  
PFV-A.3    eat:TRR-CMPL(B.3.SG) banana Juan  
'Juan ate (a/the) banana.'

Preverbal occurrence of arguments results in constructions that are structurally marked, and this fact has been the basic argument for assuming that the resulting linear orders are non-canonical in the classical grammatical analyses of this language (see Norman and Campbell 1978: 144, Lehmann 1990: 44). The examples in (5) illustrate the properties of the so-called 'focus position', which we refer to as 'pre-predicate' position in the following, in order to avoid confusion between formal and functional categories. The following examples illustrate this construction for agents, (5a), and patients, (5b). The focused constituent appears in the immediately preverbal position. When the agent of an active transitive verb is in focus, as in (5a), a special morphological form of the verb occurs, which is characterized by the drop of the preverbal aspect marker and the set-A clitic. This construction appears under particular aspectual restrictions and is known as 'agent-focus' construction in Mayan linguistics.<sup>3</sup>

- (5) (a) *huan*        *hàant*                    *ha's*.  
Juan        eat:TRR(SUBJ)(B.3.SG)    banana  
'JUAN ate (a/the) banana.'

---

<sup>3</sup> See Stiebels (2006) for an account of the conditions that determine the occurrence of the agent-focus constructions in Mayan languages.

- (b) *ha's t-u hàant-ah huan.*  
 banana PFV-A.3 eat:TRR-CMPL(B.3.SG) Juan  
 'Juan ate (a/the) BANANA.'

The examples in (6) illustrate the properties of the so-called 'topic position'. In order to avoid terminological circularity, we refer to this construction as left dislocation. In contrast to the construction involving an agent constituent in the pre-predicate position in (5a), the left dislocated agent in (6a) does not trigger any special morphological form of the verb (see Lehmann 1990, Bohnemeyer 1998b, 2004). Rather left dislocation is characterized by the occurrence of a clitic at the right edge of the left peripheral constituent, as illustrated by the enclitic *-e'* in (6) (see details in Lehmann 1990, Bohnemeyer 1998a, 1998b). These elements cannot occur at the right edge of a pre-predicate constituent.

- (6) (a) *huan-e' t-u hàant-ah ha's.*  
 Juan-D3 PFV-A.3 eat:TRR-CMPL(B.3.SG) banana  
 'Juan ate (a/the) banana.'
- (b) *ha's-e' t-u hàant-ah huan.*  
 banana-D3 PFV-A.3 eat:TRR-CMPL(B.3.SG) Juan  
 '(A/the) banana, Juan ate it.'

The aim of this paper is to give an outline of the interaction of these left-peripheral configurations with information structure. The data presented in this paper are restricted to the information structural properties of arguments. Further categories, such as adverbs, undergo similar operations of focusing and topicalization.<sup>4</sup> Section 2 introduces the structural properties of these constructions. Section 3 outlines the occurrence of these constructions in discourse and addresses the issues related to the criterion of bi-unique association, see (3a), concluding that the information structural properties are not uniform. On the basis of this conclusion, section 4 addresses the issues related to the

---

<sup>4</sup> The reader is referred to Bohnemeyer (1998b: 189-217).

criterion of independent motivation in (3b) and suggests an alternative account for the observed correlations, in which information structural conditions are independent from constituent structure.

## **2. Structural properties of left peripheral constituents**

The examples (5) and (6) show that morphological properties suggest the distinction between two types of left peripheral constituents in Yucatec Maya. This section provides evidence that these left peripheral configurations correspond to different syntactic operations.

An examination of the possibilities in the linear order of the preverbal constituents reveals that the left peripheral constituents are strictly ordered. When both a left dislocated and a pre-predicate constituent are present in a sentence, then the only possible order is ‘XP<sub>LD</sub> < YP<sub>PR</sub>’ (whereby LD=left dislocated, PR=pre-predicate constituents), as exemplified in (7a) (see Lehmann 1998: 28). The order ‘XP<sub>PR</sub> < YP<sub>LD</sub>’ is categorically excluded, as shown in (7b) which contains a left dislocated object (accompanied by a right edge clitic) and a verb that is marked for agent-focus.

- (7) (a) *huan-e’ ha’s t-u hàant-ah.*  
Juan- D3 banana PFV-A.3 eat:TRR-CMPL(B.3.SG)  
‘Juan ate (a/the) BANANA.’
- (b) *\*huan ha’s-e’ hàant-eh.*  
Juan banana-D3 eat:TRR-SUBJ(B.3.SG)  
(intended) ‘JUAN ate (a/the) banana.’

Additional evidence for the analysis that the two positions belong to different layers in the hierarchical constituent structure comes from negation. As shown in (8a-b), the negation particle must follow the left dislocated constituent. As shown in (9a-b), the same particle must precede the pre-predicate constituent.

- (8) (a) *huan-e' ma' t-u hàant-ah ha's-i'.*  
 Juan-D3 NEG PFV-A.3 eat:TRR-CMPL(B.3.SG) banana-D4  
 'Juan did not eat (a/the) banana.'
- (b) \**ma' huan-e' t-u hàant-ah ha's-i'.*
- (9) (a) *ma' huan hàant ha's-i'.*  
 NEG Juan eat:TRR(SUBJ)(B.3.SG) banana-D4  
 'JUAN did not eat (a/the) banana.'
- (b) \**huan ma' hàant ha's-i'.*

From these facts we conclude that left dislocated constituents are projected by a higher projection than pre-predicate constituents. Following previous accounts on Mayan languages (see Aissen 1992 for Mayan languages, see Gutiérrez-Bravo and Monforte y Madera 2008a for Yucatec Maya), we assume that the pre-predicate constituent occupies the Specifier position of a functional projection *IP* (=inflectional phrase). The Head I of this projection hosts the aspect/mood auxiliary that precedes the lexical verb. Left dislocated constituents are projected within a layer of the clause structure that is higher than the *IP*. We assume that it occupies a Specifier position within the *CP* (=complementizer phrase) layer.

- (10) [<sub>CP</sub> XP [<sub>IP</sub> YP [<sub>I'</sub> ... ]]].

Further distributional properties of the two possibilities in the left periphery show that the pre-predicate position is a unique position in the clause structure, while left dislocation is an operation that can be recursively used. Hence, it is possible to have more than one left dislocated constituent, as shown in (11a), while it is only possible to have one focus constituent, as shown in (11b).

- (11) (a) *huan-e' ha's-e' t-u hàant-ah.*  
 Juan-D3 banana-D3 PFV-A.3 eat:TRR-CMPL(B.3.SG)  
 'As for Juan, as for banana, he ate it.'

- (b) \**huan ha's hàant-eh.*  
 Juan banana-D3 eat:TRR-SUBJ(B.3.SG)  
 (intended) 'JUAN ate (a/the) BANANA.'

The question is which syntactic operation underlies the constructions displaying a left peripheral constituent. The available evidence shows that the occurrence of a constituent in the pre-predicate position corresponds to a gap in situ, while left dislocation does not necessarily do so. The effect of this difference is that a constituent extracted to the pre-predicate position cannot co-occur with a co-referent pronoun in situ, see (12a), while this is possible with left dislocated constituents, as shown in (12b). The presence of this emphatic pronoun is structurally possible, however pragmatically marked in (12), since Yucatec Maya is a prodrop language in which the person markers on the verb are the usual means of pronominal reference. However, native speakers have a clear intuition about the asymmetry in the acceptability of (12a) and (12b).

- (12) (a) \**pèdróoh<sub>j</sub> táan u bin màan leti'<sub>j</sub>.*  
 Pedro PROG A.3 go buy:INTRV that.one  
 (intended) 'PEDRO<sub>j</sub>, he<sub>j</sub> goes shopping.'
- (b) *pèdróoh<sub>j</sub>-e' táan u bin màan leti'<sub>j</sub>.*  
 Pedro-D3 PROG A.3 go buy:INTRV that.one  
 'As for Pedro<sub>j</sub>, he<sub>j</sub> goes shopping.'

A configuration that occurs very frequently in spontaneous discourse is the occurrence of an emphatic pronoun in the pre-predicate position that is co-referent with the left dislocated constituent, as illustrated in (13). These examples support the view that the left dislocated constituents do not necessarily correspond to a gap in situ.

- (13) *le ah kòonol<sub>i</sub>-o' leti'<sub>i</sub> tíun y-áalkab.*  
 DEF M seller-D2 that.one PROG:A.3 0-run  
 'The seller, HE is running.'

Moreover, pre-predicate constituents must be possible constituents of the clause, while there is no such restriction concerning left dislocated constituents. Hence, in example (14a) the preverbal constituent is an *NP* having a (not structurally coded) local relation to the event. Local adjuncts are generally *PPs* (apart from some lexically determined collocations with typical toponyms and verbs of motion), which has the effect that a local adjunct in the pre-predicate position has to be a *PP*, as shown in (14b).

- (14) (a) *in nah-il-e' yan u yàantal hun-p'éel cha'n*  
 A.1.SG house-REL-D3 DEB A.3 EXIST:PROC one-CL.INAN spectacle  
*sáamal-i'*.  
 tomorrow- D4  
 'As for my house, there will be a party tomorrow.'
- (b) *\*(t-)in nah-il yan u yàantal hun-p'éel cha'n*  
 LOC-A.1.SG house-REL DEB A.3 EXIST:PROC one-CL.INAN spectacle  
*sáamal-i'*.  
 tomorrow- D4  
 'In my HOUSE there will be a party tomorrow.'

The facts that pre-predicate constituents (a) cannot co-occur with co-referent material in situ and (b) must have the form of possible in situ constituents is evidence that they result from movement out of a postverbal position, see (15a). Additional evidence is provided by Norcliffe (2008), who shows that extraction to spec,*IP* is sensitive to island constraints. On the contrary, the facts that left dislocated constituents (a) can co-occur with co-referent material in situ and (b) may not have the form of possible in situ constituents is evidence that at least a subset of the left dislocated constituents are base generated in their surface position and do not correspond to a trace in the postverbal domain, see (15b).

- (15) (a) [<sub>IP</sub> YP<sub>j</sub> [<sub>I'</sub> ... t<sub>j</sub> ... ] ] ]  
 (b) [<sub>CP</sub> XP<sub>j</sub> [<sub>C'</sub> ... (proj) ... ] ] ]

However, the evidence in (12)-(14) shows that some left dislocated constituents may not correspond to possible constituents of the clause, but examples such as (6) show that this is not necessarily the case. Hence, we cannot exclude the possibility that a subset of left dislocated constituents results from movement.<sup>5</sup> Binding facts provide a further asymmetry between left dislocation and movement to the pre-predicate position. It has been shown that Yucatec Mayan subjects in VOS order bind into their antecedent objects (see Bohnemeyer 2008) and not vice versa, which provides evidence that subjects asymmetrically c-command objects in the VOS order. The binding possibility of the postcedent subject is exemplified in (16a). The binding possibilities remain the same, when the object constituent is placed in the pre-predicate position as shown in (16b), which is expected under the view that a constituent in this position is linked to the trace in situ. However, these binding possibilities do not apply to left-dislocated constituents as shown in (16c). The reading in which the possessor of the left dislocated constituent is bound by the subject is excluded, which suggests that left dislocated constituents in Yucatec Maya cannot be linked to a trace in situ which could be bound by the c-commanding subject.

- (16) (a) *k-u kol-ik u<sub>i/j</sub> kòol káadah hun-túul kolnáal<sub>j</sub>.*  
 IPFV-A.3 cut-INCMPL A.3 milpa every one-CL.AN farmer  
 ‘Every farmer clears his milpa.’
- (b) *chen u<sub>i/j</sub> kòol k-u kol-ik káadah hun-túul kolnáal<sub>j</sub>.*  
 just A.3 milpa IPFV-A.3 cut-INCMPL every one-CL.AN farmer  
 ‘Every farmer clears a MILPA.’
- (c) *u<sub>i/\*j</sub> kool-e’ k-u kol-ik káadah hun-túul kolnáal<sub>j</sub>.*  
 A.3 milpa-D3 IPFV-A.3 cut-INCMPL every one-CL.AN farmer  
 ‘As for his milpa, every farmer clears it.’

---

<sup>5</sup> Evidence for island violations in this configuration is not available; however, since left dislocation is always an option, island violations are not expected to occur, which would not shed light on the question at issue.

Based on the difference in binding properties, we assume that constituents in the left periphery instantiate exactly the configurations given in (15). Constituents land to the pre-predicate position as result of A-bar movement, while left dislocated constituents are base-generated in their surface position.

A final question relates to the structure of the thematic layer of the clause in Yucatec Maya. The assumption that Yucatec Maya is a V-initial language is based on the observation that preverbal constituents are only possible in morphologically marked configurations, see (5) and (6), and the assumption that the canonical order is VOS is based on the observation that the object-subject order is more likely to occur when the two postverbal arguments are symmetric with respect to discourse status, animacy, and weight. However, the assumption of a basic VOS order is not obvious. Antisymmetric accounts on constituent structure argue that word order permutations are universally derived by a basic Specifier-Head-Complement order; VOS order results from predicate movement past the Specifier (see Kayne 1994: 36).

The possible SVO orders in Yucatec Maya cannot be assumed to be basic: SVO order appears either when the subject is realized in the pre-predicate position (spec,*IP*) and triggers a special morphological form of the V (under particular circumstances) or when the subject constituent is left dislocated in which case it is separated from the predicate through a right edge clitic. Both configurations involve functional positions above the thematic layer of the clause. In order to justify the assumption of a basic SVO order, we would like to see the possibility of preverbal subjects to surface in a position lower than the pre-predicate position, a configuration that simply does not exist in this language. Hence, the only possibility to derive the Yucatec Mayan VOS order from a basic SVO is to assume an obligatory predicate-fronting operation from a basic configuration that never surfaces.

An empirical diagnostic for this operation is proposed in the work of Chung (2005, 2006): fronting a constituent implies that its subconstituents are inaccessible to extraction. Non-subjects in predicate-fronting languages, such as Malagasy or Seediq, cannot be relativized and are not accessible for *wh*- movement (see examples and discussion in Chung 2006: 693-697). In this view, the possibility of object-extraction in other V-initial languages (notably among else Tzotzil Maya) is evidence against

predicate-fronting in these languages. Yucatec Maya patterns with the V-initial languages that allow for object extraction, as exemplified in (17). Evidence that the pre-predicate position in (17) involves movement comes from the binding possibilities illustrated in (16b), as well as from the fact that pre-predicate constituents correspond to a gap in situ, see (12a).

- (17) *ba'x k-u tul-ik le máak-o'?*  
what IPFV-A.3 push-INCMPL DEF person-D2  
'What is the man pushing?'

In sum, the possible merits of a predicate-fronting account for Yucatec Maya are that it would accommodate the empirically attested deviation from the assumption of a universal Specifier-Head-Complement order and it would straightforwardly account for the fact that subjects bind into object constituents. However, to the extent that predicate-fronting is an empirical question and in line with the argumentation in Chung (2005, 2006), this language does not display positive evidence for predicate-fronting.

Our research question, i.e. the identification of discourse properties of preverbal constituents, relates to the conditions that determine the choice among grammatically possible word orders in this language. Hence, the empirical question is to determine the pragmatic difference between the configurations in (15) and the canonical configuration in (18), which is adopted from Aissen (1992: 46f.).

- (18) [<sub>VP</sub> [<sub>VP</sub> V Obj ] Sbj ]

### **3. Information structural properties of left peripheral constituents**

The previous section established that the constituents at the left periphery of Yucatec Maya correspond to two different structural operations: left-dislocation and movement to the pre-predicate position. The empirical question of this section is whether these structural operations are associated with particular discourse properties (see empirical question in (3a)).

### 3.1 Left dislocated constituents

We assume that the information structural function of topics is to identify the entity under which the asserted information has to be stored in the common ground (Krifka 2008). The question of this section is whether the feature [+topic] is an inherent property of the spec,*CP*. Following the argumentation in Section 1.1., the critical question for the discourse configurability hypothesis is whether all left dislocated constituents bear the information structural properties of topics.

Corpus studies show that the choice of left dislocation in discourse depends on the interaction of pragmatic and syntactic factors (see Gutiérrez-Bravo and Monforte y Madera 2008a, 2008b; Skopeteas and Verhoeven 2009b). Left dislocation of subjects of intransitive verbs is sensitive to the context. When the subject constituent is new information as illustrated in (19), which constitutes the beginning of a narrative, the predominant order is VS (88.2% out of 68 clauses with an intransitive verb and a lexically realized argument in a corpus of narrative texts, see Skopeteas and Verhoeven 2009b).

- (19) *h líub hun-p'éel che',*  
PFV fall one-CL.INAN branch  
*káa h tàal hun-tíul x-ch'úup-e' ...*  
CNJ PFV come one-CL.AN F-woman-D3  
'A branch fell down, then a girl came, ...'

When the subject constituent is part of the given information and not a continuing topic, it appears most often as left dislocated (60.7% out of 28 clauses with an intransitive verb and a lexically realized argument). Example (20) illustrates a typical context which licenses left dislocation: the subject referent of the second clause is a member of the set of given referents in discourse. It is crucial that the subject referent is not identical to the subject of the preceding clause (continuing topic), in which case it would be rather not lexically realized.

- (20) Context: ‘Near the water, there is a woman. The door of the house is open and a man is going there ...’ (further 12 clauses containing actions of the two individuals)

*káa t-u sut-ah u báah-o’b-e’*,  
CNJ PFV-A.3 turn-CMPL(B.3.SG) A.3 self-3.PL-D3  
*le máak-o’ káa h wa’l-lah yéetel u cubo*,  
DEF person-D2 CNJ PFV stand-CMPL(B.3.SG) with A.3 pot  
‘(...) and they turned around, the man stopped with his pot (...)’

The contrast between (19) and (20) shows that left dislocation is sensitive to contextual conditions. Based on the descriptive generalization that left dislocation is significantly more frequent when the constituent is related to a referent that is a member of the set of given referents in the discourse, we may assume that left dislocation signals an anaphor to the common ground. The above examples illustrate the facts from intransitive verbs. The same distribution may be observed for the single argument of different classes of intransitives (unergatives and unaccusatives), the subject of passive verbs, the single lexical argument of reflexives, as well as the subject of transitive verbs when the patient constituent is either not lexically realized or is moved into the pre-predicate position (see Skopeteas and Verhoeven 2009a for details and examples).

The contextual conditions for left dislocation differ when the linearization involves two lexical arguments. In this case, the subject constituent (almost always) surfaces as left dislocated. In the few occurrences of clauses with a transitive verb and two lexically realized arguments in the corpus mentioned above, this order is dominant with either given subject constituents (88.9% of 9 clauses) or new subject constituents (87.5% of 8 clauses) (see Skopeteas and Verhoeven 2009a, see similar findings in Gutiérrez-Bravo and Monforte y Madera 2008b). The crucial issue is that left dislocation of the subject of transitive verbs is the dominant pattern even when the subject is part of the new information. Example (21) has been elicited through picture description and illustrates a context in which the subject is new information. Such examples are in line with the observation in diverse corpus studies that V-initial sentences with two postverbal arguments occur very rarely in discourse (0.5% of transitive clauses in Skopeteas and

Verhoeven 2005, 0.8% of transitive clauses in Gutiérrez Bravo and Monforte y Madera 2008b).

(21) Context: {There is a ball on the table.}

Target: *hun-túul máak-e' túun hats'ik le bòoláah (...)*  
one-CL.AN man-D3 PROG hit-INC.MPL DEF ball  
'A man hits the ball (...)' (J 42.271)

Examples of the type illustrated in (21) are counterevidence to the putative generalization that left dislocation is licensed by the contextual conditions of topics in Yucatec Maya. This finding in the behavioral data is supported by the possibility to left dislocate non-referential *NPs*, as shown through the elicited example in (22). Native speakers confirm that (22) is acceptable and moreover that it is the preferred option to express the intended propositional content at issue (compared to the VOS order).

(22) *hu'hun-túul-il káala'n-e' u k'áat servesa.*  
RDP:one-CL.AN-REL drunk:RSLV-D3 A.3 wish beer  
'Every drunk person likes beer.'

It is crucial that the preference for left dislocation does not relate to the syntactic properties of agents of transitive verbs, but only applies if both verbal arguments are lexically realized. When the patient is a local (=1<sup>st</sup> or 2<sup>nd</sup>) person (and is only realized through the pronominal affix on the verb, see (23)), VS is the dominant pattern (occurs at 60% out of 5 clauses with a new subject and at 35.7% out of 14 clauses with a given subject). Note, furthermore, that subjects surface preferably in the postverbal domain of clauses that involve movement of the object constituent to the pre-predicate position (see Skopeteas and Verhoeven 2009b). This evidence suggests that the observed word order patterns do not relate to positional properties licensed by different types of verbal Heads (transitives vs. intransitives) but are determined by some constraint on the possible configurations of postverbal arguments.

- (23) *ts'u*            *y-áant-ik-en*            *le*    *xòok-a'* ...  
TERM:A.3    0-help-INCMPL-B.1.SG    DEF    story-D1  
'This story has helped me, ...' (HIJO\_043)

The behavioral data presented so far show that there are two independent licensing conditions for left dislocation in Yucatec Maya. In a subset of the data, which corresponds to clauses with one lexically realized argument, left dislocation occurs more frequently when the argument at issue is part of the given information. In another subset of the data, which corresponds to clauses with two lexically realized arguments, left dislocation of the subject constituent is the default option independently of information structure.

There are two different accounts that may provide an answer to the question what triggers left dislocation in clauses with two lexically realized argument *NPs*. The first theoretical possibility is to assume that the structural configuration at issue is excluded by a filter in the syntactic component. Alternatively, we may assume that the phenomenon at issue relates to conditions on the output linearization.

Alexiadou and Anagnostopoulou (2001, 2007) observe that in a set of environments whenever a clause contains a subject and an object constituent, one of them must vacate the VP. They account for this phenomenon in terms of a constraint that does not allow a VP with more than one overt argument to reach the overt component: "By Spell-Out VP can contain no more than one argument with an unchecked Case feature". This constraint accounts for a number of constructional restrictions cross-linguistically that involve a data pattern very similar with the Yucatec Maya data presented so far. For instance, subject inversion in constructions containing an expletive in English and French does not apply to transitive verbs with two lexically realized arguments. The fact that subject left dislocation in Yucatec Maya occurs across contexts may be perfectly explained through the same constraint that bans VPs with two unchecked argument *NPs* from Spell-Out.

An alternative view on the phenomenon at issue is provided by attempts to capture the impact of dissimilating processes in syntax (see Neeleman and Van De Koot 2005, Richards 2006). The intuition behind these accounts is that adjacent syntactic units of the same category that have an asymmetric relation in hierarchical syntax are difficult to

parse. Hence, linearization statements of this type are suboptimal and are occasionally banned in particular grammars. In this vein, Richards (2006) postulates a distinctness condition that bans linearization statements which contain two nodes of the same type belonging to the same derivational phase. By means of this condition, Richards (2006) accounts for a long set of syntactic phenomena, such as quotative inversion in English, stylistic inversion in French, the obligatory A-bar movement of *DP* subjects in the context of *DP* predicates in Tagalog, the ban of two *PPs* with the same prepositional Head in nominalization, constraints in sequences of adjacent verbs in several languages, the ban of multiple postverbal *DPs* in Chol, etc. The concept of the distinctness condition applies straightforwardly to our data. Yucatec Maya is a head-marking language, hence arguments do not bear case affixes, which implies a significant amount of ambiguity in a V-initial utterance with a chain of two *NPs*  $\langle NP_\alpha, NP_\alpha \rangle$  that bare identical overt features  $\alpha$ . The two adjacent postverbal *NPs* can be interpreted as two distinct arguments (with a further SO vs. OS ambiguity) or as parts of a complex *NP* containing a possessor *NP* and a possessed *NP*. The processing cost associated with the potential ambiguity of such a chain is the functional basis for a constraint against such configurations in the phonological output.

The two alternative accounts mentioned above equally explain the data pattern presented so far. We tend to see the phenomenon at issue as an effect of distinctness, because it involves gradience in that it is sensitive to the amount of features that the two postverbal *NPs* are sharing. A previous experimental study on language comprehension has shown that it is more likely that speakers avoid to interpret two postverbal *NPs* as thematically distinct arguments (and select an appositional reading instead) when the two postverbal *NPs* share the same animacy properties (see Skopeteas and Verhoeven 2005: 355). Furthermore, the avoidance of two postverbal arguments is not categorical in Yucatec Maya: VOS clauses are grammatical in this language, but they are suboptimal, as evinced by their rare occurrence in the corpus. The grammaticality of VOS structures in Yucatec Maya implies that we are dealing with a preference and not with a grammatical constraint and this aspect of the phenomenon at issue differs from the type of phenomena reported in Alexiadou and Anagnostopoulou (2001, 2007) and Richards (2006). However, it is easier to imagine a weak version of a condition relating to the

phonological output than a weak version of a rule restricting the possibilities of syntactic structure. Hence, the distinctness condition in our data renders  $\langle NP_\alpha, NP_\alpha \rangle$  linearizations suboptimal but not ungrammatical (as in the case of the phenomena reported in Richards 2006).

Section 2 has shown that there is no evidence that left dislocation involves A-bar movement in Yucatec Maya. The corpus evidence discussed in this section is based on the intuition that predicate-initial and subject-initial clauses are alternative ways to encode the same propositional content, however this assumption does not imply a derivational relation between the involved structures. The syntactic evidence suggests that left dislocation and predicate-initial structures are derivationally independent from each other. Hence, we assume that they are two independent structural possibilities resulting in alternative linearizations. Depending on the context and the involved structural properties (i.e., two or less than two lexically realized arguments) speakers select the option for left dislocation either because it satisfies the interface condition to front the topic of the sentence (in case of linearizations with one lexically realized argument) or because it does not violate the distinctness condition (in case of linearizations with two lexically realized argument).

### 3.2 Pre-predicate position

The evidence presented in Section 2 suggests that constituents in the pre-predicate position bind a trace in situ. Assuming that constituents land in this position after A-bar movement, the question is which condition licenses this operation. Previous literature on Yucatec Maya agrees that the pre-predicate position is associated with a focus interpretation (see Durbin and Ojeda 1978, Bricker 1979, Lehmann 1990, 1998, Bohmeyer 1998, Tonhauser 2003). Thus, the question arises which is the source of this interpretation. The first possibility is that the pre-predicate position is associated with a propositional operator that specifies the relation between the referent of the pre-predicate constituent and the contextually available referents, such as [+identificational] (see É. Kiss 1998, Horvath 2008). The alternative hypothesis is that movement to the

pre-predicate position is not triggered by a propositional operator but by interface properties relating to the realization of the focused part of the utterance.

The assumption of an identificational operator means that the constituent in the pre-predicate position identifies the exhaustive subset of the set of contextually and situationally given and relevant referents for which the predicate phrase holds (see É. Kiss 1998:245). The association of a position in the constituent structure with an identificational operator may be diagnosed through distributional restrictions. For instance, the particles *also* and *even* denote that the referent in their scope is a non-exhaustive subset of the set of relevant referents for which the predicate holds (see É. Kiss 1998: 251-253). Hence, the denotation of these particles contradicts the denotation of an identificational operator that is associated with the constituent structure, which results to ungrammaticality. Example (24) illustrates the effects of this restriction in English cleft constructions (see similar evidence for Hungarian focus positions in É. Kiss 1998: 252).

- (24) *It was ?also John/\*even John that Mary invited to her birthday party.* (É. Kiss 1998: 253)

The distributional properties of the pre-predicate position in Yucatec Maya differ from the properties of English clefts as exemplified in (24) and the properties of the focus position in Hungarian (reported in É. Kiss 1998: 252). The examples in (25) illustrate the fact that the Yucatec Mayan particles *xan* ‘also’ (object focus in (25a)) and *tak xan* ‘even’ (subject focus in (25b)) are grammatical in the pre-predicate position (the same holds for *also*-phrases in subject focus and *even*-phrases in object focus).

- (25) (a) *wàah xan k-u k'áat-ik le h-mèen-o'.*  
 tortilla also IPFV-A.3 ask-INCMPL(B.3.SG) DEF M-curer-D2  
 ‘The curer asks also for tortilla.’
- (b) *tak xan Pèedróoh k'áat-ik wàah-o'.*  
 as.far.as also Pedro ask-INCMPL(B.3.SG) tortilla-D2  
 ‘Even Pedro asks for tortilla.’

The distributional restrictions of particles that contradict the exhaustive interpretation may be observed in Yucatec Maya reversed pseudo-clefts, illustrated in (26). Pseudo-cleft constructions compositionally encode that the set of referents in the matrix clause is equal to the set of referents that are denoted by the variable of the relative clause, which results in the logical inference of the exhaustive interpretation (see Hedberg 2000).

- (26) (a) \*wàah xan le k-u k'áat-ik le h-mèen-o'.  
tortilla also DEF IPFV-A.3 ask-INCMP(L)(B.3.SG) DEF M-curer-D2  
(intended) 'What the curer asks for is also tortilla.'
- (b) \*tak xan Pèedròoh le k'áat-ik wàah-o'.  
as.far.as also Pedro DEF ask-INCMP(L)(B.3.SG) tortilla-D2  
(intended) 'The one asking for tortilla is even Pedro.'

The contrast between (25) and (26) provides counterevidence to the hypothesis that the pre-predicate position is associated with an identificational operator. The grammaticality of the examples in (25) shows that A-bar movement to this position is not restricted to cases in which an exhaustive relation of the referent at issue to the relevant alternatives holds true. Presumably, the examples in (25) may be licensed in discourse when the *also*-phrase in (25a) or the *even*-phrase in (25b) is the asserted part of the utterance, while the complement of the pre-predicate position contains the presupposed information. On the basis of this evidence, we conclude that movement to the pre-predicate position is not triggered by an identificational operator but relates to conditions concerning the partition between focus and background information of the utterance.

The following examples show the interaction between the occurrence of clauses with a pre-predicate constituent and different focus domains. When the focus domain encompasses the entire clause, the structure that frequently occurs is the presentational cleft illustrated in (27). This construction consists of a matrix clause with the existential predicate *yàan* 'EXIST' and a relative clause headed by the lexical verb.

- (27) *ti' lela' yàan hun-tíul xibpal yéetel*  
 LOC it:D1 EXIST one-CL.AN man:child with  
*hun-tíul chan x-ch'úuppal báax-t-ik esten... bòoláah*  
 one-CL.AN small F-girl play-TRR-INCMPL HESIT ball  
 'In this (picture) here, there is a boy and a small girl playing ...hmm... ball'

Furthermore, movement to the pre-predicate position is not observed when the focused argument is part of a wide focus domain. A study on language production reported in Skopeteas and Verhoeven (2009a) shows that when the focus domain consists of the verb and the patient constituent, the most frequently produced structure involves left dislocation of the subject/agent and the expression of the object/patient constituent in situ, see (28). When the focus domain consists of the verb and the agent constituent, the most frequently produced structure involves left dislocation of the patient and passivization, see (29). In both discourse conditions, movement to the pre-predicate position is not attested at all.

- (28) Context: 'A girl is running on the stairs...'  
*le x-ch'úuppal he'l-a' t-u léench'int-ah*  
 DEF F-woman:child PRSV-D1 PFV:A.3 push:TRR-CMPL(B.3.SG)  
*hun-p'éel k'áanche'*  
 one-CL.INAN chair  
 'this girl here pushed a chair'
- (29) Context: 'There is a snake that has opened the mouth...'  
*le kàan-o' ts'o'k u chu'k-ul tuméen le pèek'-o'*  
 DEF snake-D2 TERM A.3 catch:PASS-INCMPL by DEF dog-D2  
 'the snake has been caught by the dog'

Instances of pre-predicate constituents are frequently attested in cases of narrow focus on an argument, as exemplified in (30). The constituent *leti* 'that.one' in the pre-predicate position expresses the focused information of the utterance. The

complement to this position represents the presupposed information ‘that something makes the addressee tired’.

- (30) *déependyèenteh-ech? he'l-o'! leti' bèet-ik túun ka'n-a'n*  
employee-B.2.SG PRSV-D2 that.one do-INCMPL then tire-RSLTV  
*a w-òok-o'b.*  
A.2 0-foot-PL  
‘...You’re a clerk? So it is! That’s what makes your feet tired!’ (BVS\_15.01.27)

The view that narrow focus on an argument induces movement to the pre-predicate position is empirically supported by evidence from language production. In a production study we elicited the semi-spontaneous answers of 16 native speakers to different question types, which were related to visual stimuli.<sup>6</sup> The examples in (31) illustrate answers to subject questions of different types. Example (31a) illustrates an answer to a *wh*- question, example (31b) illustrates a corrective answer to a truth value question, and example (31c) illustrates an answer to an alternative question. In all three answers, the subject constituent is realized in the focus position, as shown by the fact that (a) the aspectual auxiliary and the set A cross-reference marker are dropped and (b) the subject constituent is not accompanied by a right edge clitic (see diagnostics in Section 1.2, example (5)).

- (31) (a) Q: {Who is looking at the girl?}  
*hun-tíul xibpal pak-t-ik le x-ch'úuppal-o'.*  
one-CL.AN man:child see-TRR-INCMPL(B.3.SG) DEF F-woman:child-D2  
‘It is a boy that is looking at the girl.’

---

<sup>6</sup> This production study is part of a larger field work agenda developed by the project *Typology of Information Structure* (part of the SFB 632 on *Information Structure*, University of Potsdam and Humboldt University Berlin). A full documentation of the experiments including stimuli, instructions, and a description of the experimental procedure may be found in Skopeteas et al. (2006).

- (b) Q: {Is a woman pushing the man?}  
*ma', hun-tíul máak tul-ik le xib-o'.*  
NEG one-CL.AN person push-INCMPL(B.3.SG) DEF man-D2.  
'No, a man is pushing the man.'
- (c) Q: {Is a man or a woman cutting the melon?}  
*hun-tíul xib xot-ik.*  
one-CL.INAN man cut  
'A man cuts it.'

All three types of context illustrated in (31) invoked almost exclusively answers with the subject constituent in the pre-predicate position, without revealing a significant effect of focus type.<sup>7</sup> This result gives further support to the view that movement to the pre-predicate position is equally likely in different focus types, notably also in contexts that do not require an exhaustive answer, as e.g. in the context of *wh*- questions. *Wh*-questions do not require an answer involving an expression of exhaustive identification, since exhaustivity is already conveyed through conversational implicature in this context (under the assumption that the speaker is cooperative and obeys the maxim of quantity).

In conclusion, the evidence from the distributional properties as well as the behavioral evidence from different context types show that this syntactic operation is not restricted to a particular propositional relation between the focus constituent and the set of relevant referents in the common ground. The comparison between different focus domains shows that movement to the pre-predicate position occurs in cases of narrow focus.<sup>8</sup> Hence, in contrast to left dislocation (see Section 3.1), the licensing conditions of the pre-predicate position are uniform.

---

<sup>7</sup> Next to the answers to subject questions that involve a constituent in the pre-predicate position, we elicited a small number of pseudo-clefts.

<sup>8</sup> The constituent in the pre-predicate position is not always identical to the focused part of the utterance. It is known that pied-piping is possible in Yucatec Maya (see Lehmann 1998: 29), hence focus movement is also licensed when the focus domain encompasses a part of the moved constituent.

#### **4. Discourse-interface strategies**

The data presented in Section 3 lead to different conclusions with respect to the correlation of the left peripheral configurations with information structural functions. The licensing conditions of left dislocation are not uniform, i.e. in a subset of the data they relate to the information structural function of topics and in another subset of the data they relate to a structural constraint that bans linearizations with a subject and an object constituent in situ. On the other hand, movement to the pre-predicate position is uniformly licensed by narrow focus.

In *discourse configurational* terms, this data pattern can be accounted for through the structure in (32), whereby XP corresponds to the left-dislocated constituent and YP to the pre-predicate constituent, see É. Kiss (1998:256) for Hungarian and Aissen (1992:47) for several Mayan languages (compare with (10)).

(32) [<sub>TopP</sub> XP [<sub>FocP</sub> YP [<sub>F'</sub> ... ]]]

The approach in (32) is based on the assumption that the functional projections in the left periphery involve Heads bearing the formal features [+topic] and [+focus]. When a goal category with the same feature occurs in the c-command domain of these Heads, then it moves to the corresponding Specifier position in order to enter into a feature-matching relation with the Head. This approach makes too strong predictions about the occurrence of left dislocation in discourse, i.e. it does not account for the instances of left dislocation that are licensed by structural conditions and do not bear the information structural properties of topics. The observations with respect to the pre-predicate position may be perfectly accounted for in terms of (32), since the information structural properties of this position are uniform. In contrast to the analysis of the Hungarian data by É. Kiss (1998), the relevant feature for Yucatec Maya is not the propositional relation of ‘exhaustive identification’, but the occurrence of a narrow focus domain. This difference is crucial, since evidence for a propositional operator cannot be reduced to the interaction of the syntax with information structural articulations.

In terms of the *discourse underspecification hypothesis* in (2), the mapping of syntactic configurations to information structure is not part of the constituent structure.

Following the syntactic evidence presented in Section 2, the left peripheral constituents occupy Specifier positions in the structural layers of *CP* and *IP* (see (10)), as repeated in (33). The constituent structure in (33) does not specify more than the structural possibilities in the left periphery. The occurrence of these possibilities in discourse depends on an array of interface conditions that independently hold. For instance, left dislocation in Yucatec Maya is licensed by the preference to front the topic constituent in a subset of our data and by the preference against violations of distinctness in another subset.<sup>9</sup>

(33) [<sub>CP</sub> XP [<sub>IP</sub> YP [<sub>I'</sub> ... ]]]

Information structural properties are determined by independent principles. Neeleman and Van De Koot (2008) outline the information structural articulations as presented in (34a-b). The representations in (34) imply an asymmetry in the embedding possibilities: a focus-background articulation may be part of the comment, but a topic-comment articulation cannot be part of the background. The following discussion will examine whether the generalizations in (34a-b) have to be postulated as independent rules.

(34) (a) topic\* [<sub>COMMENT</sub> FOCUS [<sub>BACKGROUND</sub> ... ]]  
(b) \*FOCUS [<sub>BACKGROUND</sub> topic [<sub>COMMENT</sub> ... ]]

The representations in (34) refer to parts of the phonological output and not to constituent structure (see Neeleman and Van De Koot 2008). The question is how these information structural articulations are mapped onto the constituent structure in (33). Taking a radical discourse-underspecification view, we may assume that the rules in (34) relate to purely linear partitions of the output of the syntactic rules without any reference to hierarchical structure. This view obviously makes wrong predictions about the data we presented so

---

<sup>9</sup> The crucial point for our argument is that there are at least two distinct licensing conditions for left dislocation, but we do not need to assume that the two factors illustrated by our data set represent the exhaustive set (e.g., another potential licenser could be the animate-first principle).

far. In particular, it fails to explain why focus-fronting and topic-fronting have different landing sites in Yucatec Maya. The distributional properties of *only*-phrases illustrate that the two left peripheral configurations are not interchangeable. Example (35a) involving a left dislocated *only*-phrase is judged as unacceptable when presented out of the blue (viz. it can only be accommodated as answer to the question ‘what does only Pedro eat?’). The acceptable version with the *only*-phrase in the pre-predicate position is presented in (35b). The contrast between (35a) and (35b) shows that the licensing of the left-peripheral options in Yucatec Maya cannot be accounted for if (34) relates to the output linear order alone, i.e., without any reference to hierarchical syntax.

- (35) (a) *ʔʔchen pèdróoh-e’ k-u hàant-ik bu’l.*  
           only Pedro-D3 IPFV-A.3 eat:TRR-INCMPL(B.3.SG) bean  
           ‘Only Pedro eats beans.’
- (b) *chen pèdróoh hàant bu’l.*  
       only Pedro eat:TRR-INCMPL(B.3.SG) bean  
       ‘Only PEDRO eats beans.’

These facts lead to the possibility of a one-to-one mapping between the information structures in (34) and the constituent structure in (33). Such a mapping implies the correspondence rules in (36), that are equivalent to the discourse configurational view in (32) to the exception that the association is not bi-unique, hence it does not exclude the existence of further triggers. As a consequence of the correspondence rules, the restriction in (34b) does not need to be postulated independently of syntax, since it follows from the fact that focus is projected through a lower position in the hierarchical structure, see (7). In sum, the account in (36) explains the asymmetry in the landing sites of topics and foci, does not exclude alternative triggers for the structural operations at issue and predicts a restriction against embedding topics in the background.

- (36) (a) topic           → spec,CP  
       (b) COMMENT       → C’  
       (c) FOCUS           → spec,IP

(d) BACKGROUND  $\rightarrow$  I'

The correspondence rules in (36) are descriptively adequate, but they are stipulative. In the following, we argue that the generalizations in (36) may be captured, if we take the mapping rules between syntax and phonology into account. The advantage of this approach is that it accounts for the occurrence of the syntactic configurations in discourse by means of grammatical properties that independently hold. In this view, the information structures in (34) refer indirectly to hierarchical syntax, i.e. via the mediation of prosody, and in particular the mediation of phonological phrasing.

The relevant issue is that while a left dislocated constituent forms a major prosodic phrase (*MajP*) that is separated from its complement, a pre-predicate constituent is part of the same prosodic phrase with its complement as indicated in (37).

- (37) (a)  $[_{IP} XP [_{I'} VP ]]$   $\rightarrow$   $(XP VP)_{MajP}$   
 (b)  $[_{CP} XP [_{C'} VP ]]$   $\rightarrow$   $(XP)_{MajP} (VP)_{MajP}$

Evidence for the generalizations in (37) comes from the distribution of enclitics and the properties of the phonetic realization of the corresponding utterances. Right-edge enclitics in Yucatec Maya are associated with a high  $F_0$  target and are followed by a prosodic break. These elements are attached to the right boundary of non-final major prosodic phrases (see Aissen 1992 for Tzotzil). The interaction of these enclitics with the positions in the left periphery is shown in (5) and (6). An enclitic that is associated with the right edge of a major phonological phrase follows left dislocated constituents but cannot follow pre-predicate constituents.

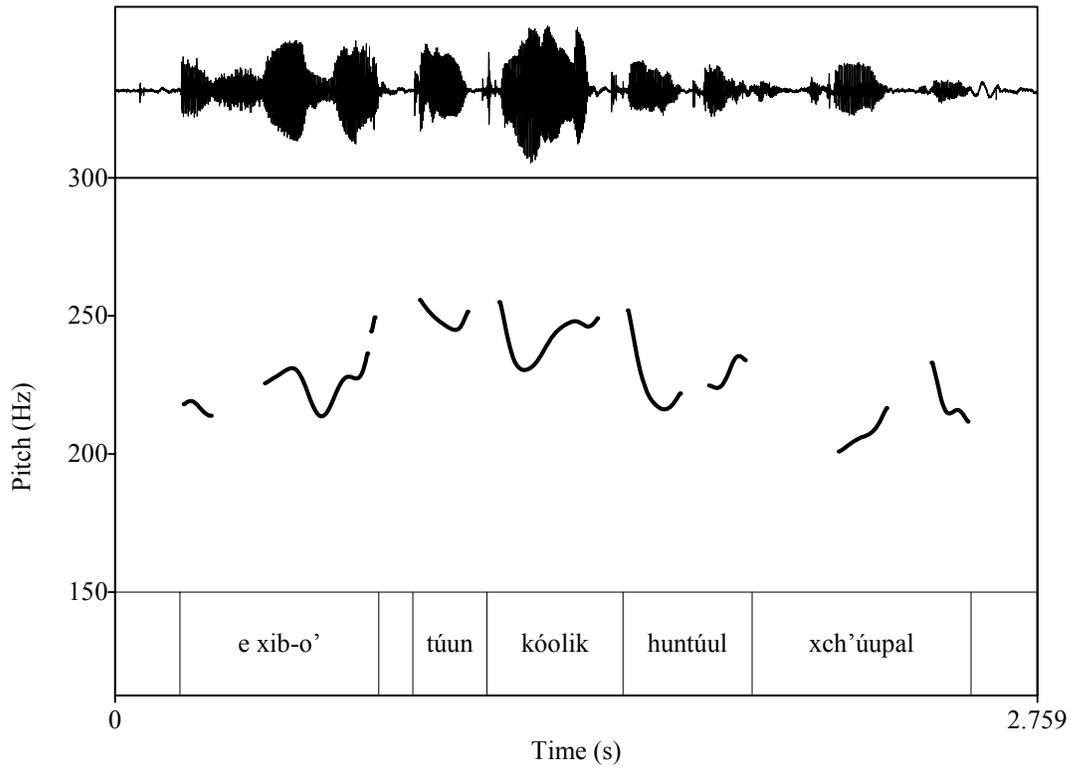
Further evidence for the difference in phrasing comes from the observation of the pitch excursions of sentences with left peripheral constituents. Yucatec Maya is a tonal language and it does not display a focus-to-accent association, as is usual in some intonation languages (see accounts on the reflexes of information structure on the Yucatec Mayan tones in Kügler and Skopeteas 2006, 2007, Gussenhoven 2007). The phonetic realization of (37a) and (37b) differs in the following respects: (a) the right edge

of a non-sentence-final *MajP* is associated with a high  $F_0$  target, which is expected to occur at the right edge of a left dislocated XP but not of a pre-predicate XP; (b) the left edge of the *MajP* is characterized by an upwards pitch reset, which occurs at the beginning of the verb complex after a left dislocated constituent but not after a pre-predicate constituent; (c) the tonal events within a *MajP* show a stepwise reduction of the pitch range: the domain of this reduction encompasses the pre-predicate constituent but not the left dislocated constituent.

These differences are illustrated in the following two figures, that show the contrast between an utterance with a left dislocated constituent (Figure 1) and an utterance with a pre-predicate constituent (Figure 2). The left peripheral constituent in Figure 1 is delimited by a high target at its right edge which is not the case for the corresponding constituent in Figure 2. The major phrase of the predicate in Figure 1 starts at the auxiliary *túun* which shows the expected upwards pitch reset signaling the beginning of a major phrase. This reset is absent in Figure 2, since the pre-predicate constituent belongs to the same major phrase with the predicate. Furthermore, it may be observed that the scaling of the high tones within the predicate in Figure 1 shows a stepwise reduction beginning from the predicate. The domain of the pitch reduction in Figure 2 starts from the beginning of the utterance and contains the pre-predicate constituent.

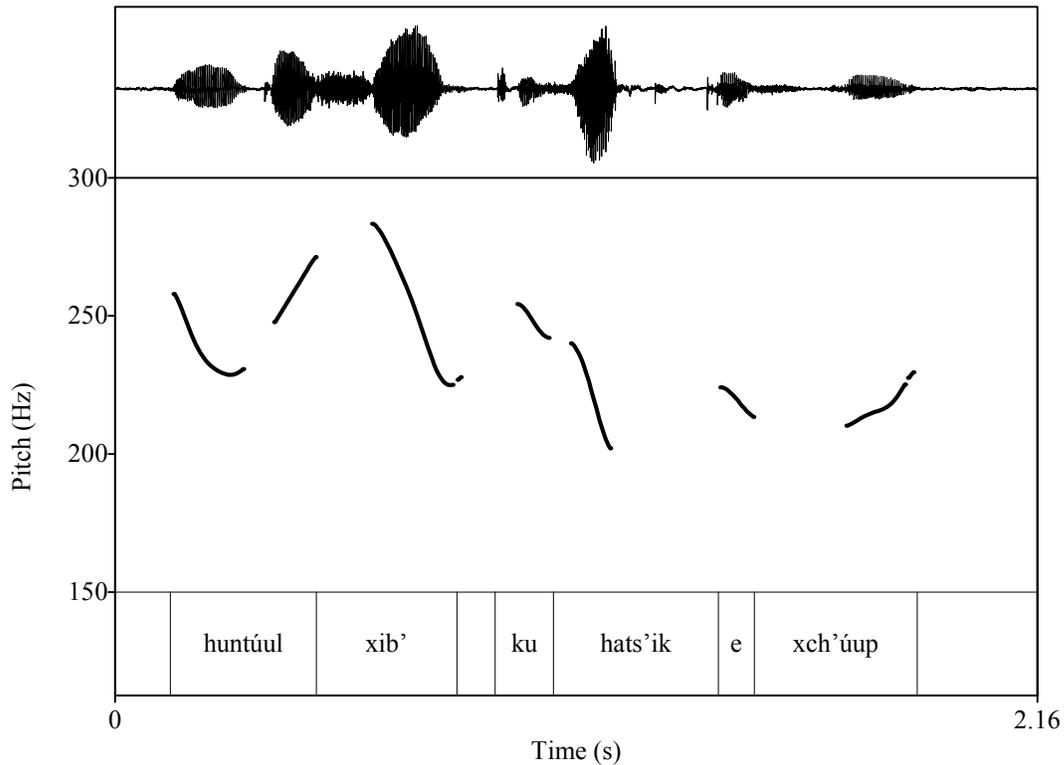
- (38) [[<sub>CP</sub> *e xib-o'*] [<sub>VP</sub> *túun kóol-ik hun-túul xch'úuppal*]]  
DEF man-D2 PROG pull-INCMPL one-CL.AN F:woman:child  
'The man is pulling a girl.'

Figure 1. Illustrative  $F_0$  excursion of a topic construction



- (39) [ [IP *hun-túul* *xib* ] [VP *k-u* *hats'-ik* *e* *x-ch'úupal* ]  
 one-CL.AN man IPFV-3.SG beat-INC MPL DEF F-woman  
 'It is a man that the woman hits.'

Figure 2. *Illustrative F<sub>0</sub> excursion of a focus construction*



The phonological evidence discussed so far shows that pre-predicate and left dislocated constituents differ in phrasing: the former but not the latter are integrated to the *MajP* that contains the predicate. The assumed interaction between these facts and the information structural concepts is sketched in the following. Research on focus has established that focused constituents target the most prominent position in the clause (see Szendrői 2001:47, Büring 2009 among others). A detailed account of the prominence rules in the Yucatec Mayan prosodic structure is not yet available, but there is evidence that Yucatec Maya belongs to the languages in which the most prominent constituent within a phonological is the leftmost one.<sup>10</sup> Following Szendrői (2001:46) on Hungarian,

---

<sup>10</sup> Yucatec Maya is a tonal language and as shown in previous research does not display pitch accents for marking focus (see Kügler and Skopeteas 2006, 2007, Gussenhoven 2007). However, there is evidence for leftmost prominence already in the lower layers of prosodic constituency: the first tone bearing unit within the prosodic word is realized with maximal prominence, independently of whether it is part of a lexical or a functional unit.

we assume that narrow focus targets the leftmost position within the *MajP* that contains the predicate, while the *MajP* that contains the left dislocated constituent is extrametrical, i.e., it contains adjoined material that is not visible for the rules that assign prosodic prominence. Hence, assuming (40a) and taking into account the generalization in (37) implies that focused constituents will target the pre-predicate condition and not the configuration of left dislocation. For topicalized constituents, we need to assume a rule that forces the realization of the topic and the comment partitions in separate prosodic units. The generalization in (37b) shows that the corresponding phrasing option is made available by left dislocation.

- (40) (a) FOCUS → max. prominent  
(b) topic [COMMENT ...] → (...)<sub>MajP</sub> (...)<sub>MajP</sub>

The assumptions in (40) relate the information structural properties to the properties of the phonological output without direct reference to hierarchical structure. The interaction with hierarchical structure is predicted by the fact that the latter determines the possibilities to realize a propositional content in phonological phrases. The assumption of correspondence rules between information structure and prosody, see (40), and mapping rules between constituent structure and prosodic structure, see (37), accounts for the full range of data reported in this article. Finally, the asymmetry in the embedding possibilities in (34) is the epiphenomenal result of the principles that relate information structural articulations to phonological phrases, since topics target a structural position that is phrased separately from the predicate and foci target a position that is phonologically integrated to it.

In the light of the criterion in (3b), the argumentation in this section shows that the data pattern of left peripheral constituents in Yucatec Maya can be accounted for without resort to the assumption that structural positions are inherently associated with information structural properties. The part of the data that involves a 1:1 correspondence between syntax and discourse, i.e. the uniform licensing of movement to the pre-predicate position by narrow focus, can be fully predicted by the mapping conditions between hierarchical structure and prosodic phrasing.

## **5. Conclusion**

This paper presented evidence that the left periphery of Yucatec Maya displays two structurally distinct configurations: left dislocated constituents are base-generated and are realized at the beginning of the utterance; pre-predicate constituents result from movement to a position that is left adjacent to the verb complex. Behavioral and distributional evidence has shown that there are at least two different types of licensors for left dislocation, the first of which is information structural (topicalization) and the second is syntactic (avoidance of distinctness violations). Movement to the pre-predicate position is pragmatically uniform, i.e., it is licensed by narrow focus.

Based on the bi-uniqueness criterion, we argued that the discourse configurational approach can only account for the pre-predicate position and not for the discourse properties of left dislocation. After this conclusion, we sketched an alternative account based on the idea of discourse interface strategies that do not directly refer to hierarchical clause structure. The observed phenomena may be accounted for by the assumption of discourse templates that capture the linear properties of a topic-comment and a focus-background articulation. Rules referring exclusively to the linear order do not capture all facets of interaction of information structure with hierarchical syntax, in particular they do not account for the fact that narrowly focused constituents are not left dislocated. In order to account for this difference, we have shown that the two structural options in the left periphery license different prosodic structures. Crucially, it is only the pre-predicate position that is realized in a phonological phrase with the predicate and this explains why focus-material targets exactly this position.

The consequences of the discourse underspecification hypothesis led to an account on the information structural phenomena based on the independent properties of the grammar. Apart from the fact that this theoretical possibility accounts for a wider range of data (the critical case being the functions of left dislocation), it provides an explanation of the interplay between syntax and phonology for the realization of information structural configurations without the stipulative assumption of discourse features that are associated with syntactic configurations.

Our account suggests that at least for a subset of the languages that have been explained in terms of the discourse configurational approach, the cartography of the left

periphery is an unnecessary complication of the syntactic component. The extension of our account to further languages is an empirical question that requires detailed examination of the exact conditions that license left peripheral configurations (see the bi-uniqueness criterion in (3a)) and the possibility to decompose the information structural properties by means of interface conditions (see the criterion of independent motivation in (3b)). It is reasonable to assume that the observed correlations between the concepts of topic and focus and positions in the left periphery can be decomposed in a similar manner in further languages that have been dealt with in the framework of discourse configurational approaches. However, this does not exclude the possibility that structural configurations in some languages display semantic or information structural properties that cannot be derived from interface conditions. For instance, the asymmetry of Yucatec Mayan left peripheral positions with respect to prosodic phrasing that is discussed in Section 4, is very similar to the facts from Hungarian, as reported in Szendrői (2001). However, Hungarian differs from Yucatec Maya in that *also*-phrases and *even*-phrases are banned from the pre-predicate position (see É. Kiss 1998: 252). Such distributional restrictions cannot be straightforwardly accounted for through interface conditions, hence they are strong evidence for the existence of a semantic operator, such as [+identificational]. In view of such cross-linguistic differences, we conclude that the issue whether word order is associated with discourse-related concepts is not a matter of conceptual decision but a language-specific empirical question.

### Abbreviations

A= person affix, class A; AN= animate; B= person affix, class B; CL= class; CMPL= completive; CNJ= conjunction; D1= 1<sup>st</sup> person deixis; D2= 2<sup>nd</sup> person deixis; D3= 3<sup>rd</sup> person deixis; D4= locative/negative particle; DEB= debitive; DEF= definite; EXIST= existential; F= feminine; HESIT= hesitant; INAN= inanimate; INCMPL= incomplete; INTRV= introversive; IPFV= imperfective; LOC= locative; M= masculine; NEG= negation; PASS= passive; PFV= perfective; PL= plural; PROC= processive; PROG= progressive; PRSV= presentative; RDP= reduplicative; REL= relationalizer; RSLV= resultative; SG= singular; SUBJ= subjunctive; TERM= terminative; TRR= transitivizer; 0= meaningless element; 1= 1<sup>st</sup> person; 3= 3<sup>rd</sup> person.

## References

- Aissen, J. 1992, Topic and focus in Mayan. *Language* 68: 43-80.
- Alexiadou, A. and Anagnostopoulou, E. 2001, The subject in situ generalization, and the role of Case in driving computations. *Linguistic Inquiry* 32: 193-231.
- Alexiadou, A. and Anagnostopoulou, E. 2007, The subject in situ generalization revisited. In: H.-M. Gärtner and U. Sauerland (eds.), *Interfaces + recursion = language*. Berlin: Mouton de Gruyter, 31-59.
- Bohnemeyer, J. 1998a, Sententiale Topics im Yukatekischen. In Dietmar Zaefferer (ed.), *Deskriptive Grammatik und allgemeiner Sprachvergleich*. Tübingen: Niemeyer, 55-85.
- Bohnemeyer, J. 1998b, *Time Relations in Discourse: Evidence from a Comparative Approach to Yucatek Maya*. Ph. D. Dissertation, Katholieke Universiteit Brabant.
- Bohnemeyer, J. 2004, Split intransitivity, linking, and lexical representation: the case of Yucatek Maya. *Linguistics* 42: 67-107.
- Bohnemeyer, J. 2008, Linking without grammatical relations in Yucatec: alignment, extraction, and control. Ms., University of Buffalo.
- Bricker, V. R. 1979, Wh-questions, relativization, and clefting in Yucatec Maya. In L. Martin (ed.), *Papers in Mayan Linguistics*. Columbia, Miss.: Lucas Brothers, 107-136.
- Büring, D. 2009, Towards a typology of focus realization. *Information Structure from different perspectives*, M. Zimmerman & C. Féry (eds). Oxford: Oxford University Press.
- Chomsky, N. 1971, Deep structure, surface structure, and semantic interpretation. In D. Steinberg and L. Jakobovits, (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge: Cambridge University Press, 183-216.
- Chung, S. 2005. What fronts? On the VP-raising account of verb-initial order. In: Andrew Carnie, Heidi Harley, and Sheila Ann Dooley (eds.), *Verb first: On the syntax of verb initial languages*. Amsterdam, Philadelphia: John Benjamins, 9-29.
- Chung, S. 2006. Properties of VOS languages. In: Everaert, Martin and van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax*, vol. 3. Oxford: Blackwell, 685-720.

- Durbin, M. and Ojeda, F. 1978, Basic word-order in Yucatec Maya. In Nora C. England (ed.), *Papers in Mayan Linguistics, Vol. 2*. Columbia: University of Missouri, Department of Anthropology, 69-77.
- É. Kiss, K. 1998, Identificational focus versus information focus. *Language* 74.2: 245-273.
- Fanselow, G. 2006, On pure syntax (uncontaminated by information structure). In: Brandt, Patrick, Fuss, Erik (eds.), *Form, Structure and Grammar*. Akademie Verlag, Berlin, 137-157.
- Fanselow, G. & Lenertová D. 2008. Left peripheral focus. Mismatches between syntax and information structure (submitted).
- Gussenhoven, C. 2007, A moraic and a syllabic H-tone in Yucatec Maya. *Proc. In Fonología instrumental: patrones fónicos y variación lingüística*, Colegio de México, Mexico City, 21-27 October 2006.
- Gutiérrez-Bravo, R. and Monforte y Madera, J. 2008a. La alternancia sujeto inicial/verbo inicial y la teoría de la optimidad. In: R. Gutiérrez Bravo y E. Herrera Zendejas (eds.), *Teoría de la optimidad: estudios de sintaxis y fonología*. El Colegio de México, Centro de Estudios Lingüísticos y Literarios-Laboratorio de Estudios Fónicos, México, 61-90.
- Gutiérrez Bravo, R. and Monforte y Madera, J. 2008b, On the nature of unmarked word order in Yucatec Maya. Paper presented at the 2008 Meeting of SSILA, Chicago, January 4, 2008.
- Hartmann, K., Zimmermann, M. 2006, In place – out of place? Focus strategies in Hausa. In Schwabe, K., Winkler, S. (eds.), *On Information Structure, Meaning and Form: Generalizations across Languages*. Amsterdam/Philadelphia: Benjamins, 365-403.
- Hedberg, N. 2000, The referential status of clefts. *Language* 76.4: 891-920.
- Hofling, C. A. (1984). On proto-Yucatecan word order. *Journal of Mayan Linguistics* 4.2: 35-64.
- Horvath, J. 2008, “Discourse features”, syntactic displacement and the status of contrast, *Lingua* (to appear).
- Kayne, R. S. 1994. *The antisymmetry of syntax*. Cambridge/Mass., London: MIT Press.

- Krifka, M., 2008, Basic notions of information structure. In: Féry, C., Fanselow, G., Krifka, M., (Eds.), *The Notions of Information Structure*, Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS), 6, 14-42.
- Kügler, F. and Skopeteas, S. 2006, Interaction of Lexical Tone and Information Structure in Yucatec Maya. *Second International Symposium on Tonal Aspects of Languages, La Rochelle (France), April 27-29<sup>th</sup>, 2006*, 380-388.
- Kügler, F. and Skopeteas, S. 2007, On the universality of prosodic reflexes of contrast: the case of Yucatec Maya., *16<sup>th</sup> International Congress of Phonetic Sciences, 6-10 August 2007, Saarbrücken, Germany*, 1025–1028.
- Neeleman, A. and Van De Koot, H. 2005, Syntactic haplology. In M. Everaert, H. van Riemsdijk, R. Goedemans, and B. Hollebrandse (eds.), *The Blackwell Companion to Syntax: Volume IV*. London: Blackwell, 684-710.
- Neeleman, A. and Van De Koot, H. 2008, Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11.2: 137-189.
- Norcliffe, E. 2008. Filler-Gap dependencies in Yucatec. Paper presented at the Symposium *New Perspectives in Mayan Linguistics*, held at the 2008 SSILA Annual Meeting, January, 3-4, 2008, Chicago.
- Norman, William M. and Campbell, Lyle 1978, Toward a proto-Mayan Syntax: A comparative perspective on grammar. In: England, Nora C. (ed.), *Papers in Mayan Linguistics*, vol. 2. Columbia: University of Missouri, Department of Anthropology (University of Missouri Miscellenous Publications in anthropology, 6), 136-155.
- Lehmann, C. 1990, Yukatekisch. *Zeitschrift für Sprachwissenschaft* 9.1-2: 28-51.
- Lehmann, C. 1998, *Possession in Yucatec Maya*. Unterschleissheim: LINCOM Europa.
- Reinhart, T. 2006, *Interface Strategies: Optimal and Costly Derivations*. Cambridge, Mass.: MIT Press.
- Richards, N. 2006, A distinctness condition on linearization. Unpublished manuscript. MIT.
- Rizzi, L., 1997, The fine structure of the left periphery. In: Haegeman, L., (ed.), *Elements of Grammar*. Kluwer, Dordrecht, 281-337.

- Skopeteas, S. and Fanselow, G. 2008, Effects of givenness and constraints on free word order. In Malte Zimmerman and Caroline Féry (eds.), *Information Structure from different perspectives*. Oxford: Oxford University Press.
- Skopeteas, S. and Verhoeven, E. 2005, Postverbal argument order in Yucatec Maya. *Sprachtypologie und Universalienforschung* 53.1: 71-79.
- Skopeteas, S. and Verhoeven, E. 2009a, Distinctness effects on VOS order: evidence from Yucatec Maya. In Avelino, H., Coon, J., and Norcliffe, E. (eds.), *New Perspectives in Mayan Linguistics*. Cambridge, MA: MIT (MIT working papers in Linguistics), 157-173.
- Skopeteas, S. and Verhoeven, E. 2009b, The interaction between topicalization and structural constraints: Evidence from Yucatec Maya. *The Linguistic Review*, in press.
- Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., Féry, C. and Krifka, M. 2006, *Questionnaire on Information Structure: Reference Manual*. Special issue in *Interdisciplinary Studies on Information Structure*, Vol. 4. Potsdam: Universitätsverlag Potsdam.
- Szendrői, K. 2001, *Focus and Syntax Phonology Interface*. Ph.D. dissertation, UCL.
- Szendrői, K. 2003, A stressed-based approach to the syntax of Hungarian focus. *The Linguistic Review* 20, 37-78.
- Tonhauser, J. 2003, F-constructions in Yucatec Maya. In Jan Anderssen, Paula Menéndez-Benito and Adam Werle (eds.), *The Proceedings of SULA 2*. Vancouver, BC: GLSA, University of Massachusetts, Amherst, 203-223.
- Van Valin, R. D. and LaPolla, R. J. 1997, *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Verhoeven, E. 2007, *Experiential constructions in Yucatec Maya*. Amsterdam: Benjamins.
- Wedgwood, D. J. 2003, Predication and Information Structure: A Dynamic Account of Hungarian Pre-verbal Syntax. Ph. D. dissertation, Univ. of Edinburgh.
- Wedgwood, D., Pethő, G., and Cann, R. 2006, Hungarian 'focus' position and English *it*-clefts: the semantic underspecification of focus readings. Ms., University of Edinburgh.

Zimmermann, M., 2008, Contrastive focus. In Féry, C., Fanselow, G., and Krifka, M. (Eds.), *The notions of information structure*. Ms., University of Potsdam.