

LICENSING FOCUS CONSTRUCTIONS IN YUCATEC MAYA

Stavros Skopeteas¹ and Elisabeth Verhoeven²

Bielefeld University¹ and Humboldt University Berlin²

Table of contents

- 1 Preliminaries
 - 1.1 Licensing focus
 - 1.2 Grammatical properties of Yucatec Maya
 - 2 Structural properties
 - 2.1 Basic properties
 - 2.2 Clausal layers
 - 2.3 Structure of fronting
 - 2.4 Summary
 - 3 Information-structural properties
 - 3.1 Focus domain
 - 3.2 Focus interpretation
 - 3.3 Occurrence in context
 - 3.4 Summary
 - 4 Prosodic properties
 - 5 Conclusions
- Acknowledgments

Abstract

An important challenge in the study of focus constructions is teasing out the properties of the layers of linguistic structure that are involved, in particular identifying which interpretational properties are associated with the syntactic operation at issue, which properties arise through inferential processes, and which properties can be deduced on the basis of the prosodic structure. This article undertakes this challenge in a language with a structurally identifiable left-peripheral position which is employed for the expression of focus, namely Yucatec Maya. This syntactic configuration comes with a focus interpretation and we show that the occurrence of this construction is not restricted to a subtype of focus corresponding to a truth-conditionally relevant operator. The properties of the syntax-prosody mapping indicate that focus-fronting is a syntactic operation placing the material in focus in the maximally prominent partition of the prosodic constituent that contains the predicate.

Keywords

focus, exhaustivity, word order, prosodic phrasing, information structure

1 Preliminaries

1.1 Licensing focus

A fundamental question in studies on focus is whether the focus interpretation that we identify for certain constructions is an inherent property of particular operations in syntax or an effect of the interaction

between the context and some surface properties of syntactic constructions, in particular their linear order and its relation to prosodic structure.

The first view is explicitly formulated in accounts that assume a form-to-function association between focus and constituent structure (Dik 1997, Rizzi 1997, É. Kiss 1998, 2009, Drubig 2003). For instance, É. Kiss (1998: 267f.) assumes operators such as [+ contrastive] and [+ exhaustive] to be associated with focus positions in different languages. In a different grammatical framework, Dik et al. (1981) and Dik (1997) assume a hierarchy of focus subtypes, such as completive, selective, and corrective, which corresponds to different types of contexts relating to the asserted information, an approach which predicts that particular constructions in the world's languages are associated with particular subsets of this hierarchy.

This view has been challenged in recent years by a number of studies that reject the form-to-function association between syntax and information structure (see Hartmann and Zimmermann 2006, Fanselow and Lenertová 2011, Wedgwood 2006, Zimmermann 2008). Along these lines, several studies on particular languages show that the interpretational properties of focus constructions are implied by the interaction of syntax with prosodic phrasing and prominence within prosodic domains (see, e.g., Szendrői 2001 on Hungarian [HUA], Koch 2008 on Thompson River Salish [THP], Cheng and Downing 2012 on Zulu [ZUL]). Recent focus typologies, such as Büring (2009), show that syntactic operations in several languages can be

accounted for as strategies to achieve prosodic configurations that fit the intended focus domains. These findings motivate a completely different view of the role of information structure in the grammar. The information-structural possibilities of particular constructions are the product of properties of linearization and prosodic structure which exist independently, and not proper features of the constituent structure.

The aim of this article is to give a precise account of the syntactic, prosodic and information structural properties of focus constructions in Yucatec Maya [YUA]. In terms of descriptive adequacy, if a particular syntactic operation is inherently triggered by a discourse feature, e.g., ‘focus’, it is expected that the presence of the ‘focus’ feature is at least a *necessary condition* in order for the operation to take place. Deviations from this generalization may be compatible with the assumption of a feature-triggered operation, if these deviations are accounted for through the application of rules that are independently required. In terms of economy in linguistic descriptions, the assumption of a form-to-discourse association is a reasonable (i.e., minimally stipulative) generalization if there are no grammatical properties of greater generality that can account for the observed facts.

Yucatec Maya is a particularly interesting case for the research question at issue. This language provides a syntactic operation of fronting the focus to a left-peripheral position that can be unambiguously identified through

concomitant morpho-syntactic properties. This is the ideal empirical situation for the assumption of a ‘designated focus position’ in the constituent structure of this language. The challenge of this article is to account for the relation between the syntactic and the information-structural properties of this configuration.

1.2 Grammatical properties of Yucatec Maya

Yucatec Maya is a head-marking language: verbs carry cross-reference markers co-indexed with their arguments. Transitive verbs are accompanied by a set of markers that cross-references the agent (known as ‘set-A’) and a suffix that cross-references the patient (known as ‘set-B’); see (1) (Lehmann 1990: 40; Bohnemeyer 1998: 159). Verb stems are accompanied by a suffix (termed ‘status marker’; *-ah* in (1)) that co-occurs with the tense/aspect/mood auxiliary (*t-* in (1)); in particular, there are three such markers (completive; incomplete; subjunctive) each corresponding to a subset of auxiliaries (see Bohnemeyer 1998: 291).

(1)¹ *T=u hàant-ah òon Pèedróoh.*

PFV=A.3 eat:TRR-CMPL(B.3.SG) avocado Pedro

‘Pedro ate avocado.’

The canonical order is VOS; see (1). The order of postverbal arguments may be influenced by asymmetries in definiteness/givenness, animacy, and weight of the involved constituents (Skopeteas and Verhoeven 2005; Bohnemeyer 2009). Verb-initial orders with two postverbal arguments occur

only rarely in discourse (1% in a corpus query reported in Skopeteas and Verhoeven 2005), which motivates the assumption that Yucatec Maya is an SVO language (Durbin and Ojeda 1978, Gutiérrez Bravo and Monforte y Madera 2007, 2008). Nonetheless, the preference for initial subjects is restricted to clauses involving transitive verbs and two lexically realized arguments; it does not apply to intransitives and passives, nor to transitives with a pronominally realized agent or patient. In light of these facts, we consider the VOS order as the basic syntactic configuration and we assume that there is a constraint against linearizations with two adjacent lexical arguments in the postverbal domain, resulting in a preference for the subject-initial order in the utterances that contain a verb and two lexical NP arguments (Skopeteas and Verhoeven 2009a).

Topicalized constituents occur clause initially and are right-bounded by an enclitic, e.g., the element *=e* ‘D3’ in (2a). This enclitic is the unmarked member of a class of deictic enclitics (Bohnenmeyer 1998: 205; Lehmann 1990; Verhoeven 2007: 105; Skopeteas and Verhoeven 2009b; see further discussion in 3.1).² Narrow focus triggers the displacement of an argument in the immediate preverbal position (Bricker 1979; Tonhauser 2003, 2007; Gutiérrez Bravo 2007); see (2b).

- (2) a. *Pèedróoh=e’ t=u hàant-ah òon.*
 Pedro=D3 PFV=A.3 eat:TRR-CMPL(B.3.SG) avocado
 ‘As for Pedro, he ate (an) avocado.’

- b. òon t=u hàant-ah Pèedróoh.
 avocado PFV=A.3 eat:TRR-CMPL(B.3.SG) Pedro
 ‘Pedro ate an AVOCADO.’

The construction in (2b) is the subject of this article. First of all, we examine its syntactic properties and show that the ‘focus construction’ involves an operation of constituent fronting (Section 2): the constituent in the preverbal position always corresponds to a possible constituent *in situ*. Section 3 shows that narrow focus of any type is a necessary condition for constituent fronting to take place. This observation is crucial because a correlation with a focus type (such as exhaustive or contrastive focus) would indicate that a propositionally-relevant operator is associated with the position of the fronted constituents. Based on this conclusion, we examine the possibility that focus-fronting is the product of independent properties of Yucatec Mayan prosody and syntax. Section 4 shows that preverbal foci are integrated in the prosodic entity that contains the predicate and we observe that the left edge of this prosodic constituent is consistently prominent. In Section 5, we draw conclusions from the presented facts concerning the interplay between syntax, prosody and information structure in Yucatec Maya.

Our account is based on elicited and experimental data collected with native speakers from Quintana Roo (Yaxley and Felipe Carrillo Puerto) in December-January 2004, August 2006, March 2008, and September 2012,

as well as data from a text collection compiled by Christian Lehmann (University of Erfurt).

2 Structural properties

2.1 Basic properties

The focus construction in Yucatec Maya is formed through the preposing of the focused constituent in the immediately preverbal position; see (2b) with a preposed object and (3) with a preposed subject of an intransitive verb (Bricker 1979; Lehmann 1990, 2003: 29; Bohnemeyer 1998: 190f.; 2009; Tonhauser 2003, 2007; Gutiérrez Bravo 2007).

- (3) *Pèedróoh k=u hàan-al.*
 Pedro IPFV=A.3 eat-INC MPL
 ‘PEDRO is eating.’

When the focused constituent is the agent of a transitive verb (and only then), the verb appears in a special morphological form: the tense/aspect/mood auxiliary and the A cross-reference marker are dropped (Bricker 1979: 109; Lehmann 2003: 29; Gutiérrez Bravo 2007; Tonhauser 2007).³ The verb appears with a status marker (see discussion in 2.1), while only two of the otherwise available status markers appear in this construction (incompletive and subjunctive); see (4). With imperfective reference, the verb appears in the incompletive status, marked by the suffix *-ik* ‘INC MPL’, as in (4a). With perfective reference, it appears in the

subjunctive status, which is zero-marked in non-clause final position; see (4b).⁴ The same construction occurs in constituent questions; see (4c).

- (4) a. *Pèedróoh hàant-ik òon.*
 Pedro eat:TRR-INCMPL(B.3.SG) avocado
 ‘PEDRO is eating (an) avocado.’
- b. *Pèedróoh hàant òon.*
 Pedro eat:TRR(SUBJ)(B.3.SG) avocado
 ‘PEDRO ate (an) avocado.’
- c. *máax hàant-ik òon?*
 who eat:TRR-INCMPL(B.3.SG) avocado
 ‘Who is eating (an) avocado?’

Since the morpho-syntactic properties of constituent fronting either for focused constituents or for interrogative pronouns are generally identical, we assume a single syntactic operation for these constructions (following a view established in Bricker 1979; Bohnemeyer 1998: 191; Tonhauser 2003: 208ff.; AnderBois 2012; but see also Gutiérrez Bravo and Monforte y Madera 2011 for several differences between these constructions).

Other constructions that are used for the expression of focus in Yucatec Maya have a bi-clausal structure. A comparison with these constructions will be used to identify the syntactic and interpretational properties of focus fronting in this language. The construction in (5a) is a cleft construction. It consists of a nominal predicate and a headless relative clause formed

through nominalization of the verb phrase via the determiner (Bohnmeyer 2009: 198ff.). Example (5b) shows a linearization that results from the topicalization of the relative clause. The corresponding constructions in English [ENG] would be a pseudocleft for (5b) and a reversed pseudocleft for (5a); however, note that for a verb-initial language, the basic configuration is (5a).

- (5) a. *Pèdróoh le t=u hàant-ah òon-e’.*
 Pedro DEF PFV=A.3 eat:TRR-CMPL(B.3.SG) avocado=D3
 ‘It is Pedro who ate avocado.’
- b. *le t=u hàant-ah òon-e’, Pèdróoh.*
 DEF PFV=A.3 eat:TRR-CMPL(B.3.SG) avocado=D3 Pedro
 ‘Who ate avocado is Pedro.’

The tense/aspect/mood contrasts are reduced with the agent-focus construction through the dropping of the auxiliary. Speakers may use alternative constructions in order to make temporal/aspectual/modal categories explicit. Such a construction is a cleft construction with the future subordinator *kéen* with future time reference (Bohnmeyer 1998: 192ff); see (6).

- (6) *Tèen kéen in hats’ hun-p’éel hit.*
 1.SG SR.FUT A.1.SG beat(SUBJ)(B.3.SG) one-CL.INAN hit
 ‘It is me who will get (lit. beat) a hit.’ (Bohnmeyer 1998:194)

In order to understand the syntactic structure of the focus construction, we will discuss two crucial syntactic properties. First, we compare the structural properties of topics and foci (Section 2.2). Second, we examine the question of whether the focus construction is mono-clausal or bi-clausal (Section 2.3).

2.2 Clausal layers

Since we do not wish to anticipate an association of syntactic positions with pragmatic functions, we refer to the traditionally termed ‘topic position’ in (2) as *left dislocation* and to the traditionally termed ‘focus position’ in (3)/(4) as a *pre-predicate* position. The pre-predicate position is characterized by the fact that fronted agent constituents are accompanied by a particular verb form, which does not hold for left-dislocated agents; compare (4) to (2). Left-dislocated constituents are identified by the fact that an enclitic occurs at their right edge (see =*e*’ ‘D3’ in (2)), which is not possible at the right edge of pre-predicate constituents (see Section 4). The assumption that these two types of left-peripheral constituents occupy different positions is based on their linear order properties that are reported here in brief (see Skopeteas and Verhoeven 2009b for further discussion):

- (a) Whenever both types of constituents are available, the only possible order is ‘left-dislocated XP < pre-predicate YP’.
- (b) Left-dislocated constituents obligatorily precede, while pre-predicate constituents obligatorily follow, the negation marker *ma*’ ‘NEG’.

- (c) A sentence may contain two or more left-dislocated constituents, but not more than one pre-predicate constituent (a property reported for several languages; see Rizzi 1997 on Italian [ITA]).

It is crucial that the pre-predicate constituent is a part of the clause (extracted from the postverbal domain under particular semantic/pragmatic conditions), while this does not hold for left-dislocated constituents (Lehmann 1990: 44; 2003: 28; Bohnemeyer 2009: 190). The pre-predicate constituent but not the left-dislocated constituent *must* have a syntactic relation to an element within the clause (see Aissen 1992 for Tzotzil [TZO]). For instance, example (7a) shows an utterance with a dislocated element that is not a possible constituent of the clause; compare this to (7b). The same element cannot occur in the pre-predicate position, as illustrated by (7c).⁵ (The contrast between left dislocation and pre-predicate constituents is tested by the presence of the enclitic =*e* 'D3'; the verb form does not have distinct properties with objects in the pre-predicate position.)

- (7) a. *ch'úich'-o'b=e' in k'ahóol chen x-k'òok'-o'b.*
 bird-PL=D3 A.1.SG know only F-nightingale-PL
 'As for birds, I know only nightingales.'
- b. **in k'ahóol chen x-k'òok'-o'b ch'úich'-o'b.*
 A.1.SG know only F-nightingale-PL bird-PL
 (intended) 'I know only nightingale birds.'
- c. **ch'úich'-o'b in k'ahóol chen x-k'òok'-o'b.*

bird-PL A.1.SG know only F-nightingale-PL

(intended) ‘BIRDS, I know only nightingales.’

Left-dislocated constituents but not pre-predicate constituents may co-occur with a co-referent pronoun *in situ*. The third person pronoun in (8a) is co-referent with the left-dislocated constituent; the sentence is grammatically well-formed – though it is considered to show redundancy when presented out of the blue. The version in (8b), with the personal pronoun in the pre-predicate position, occurs frequently in discourse and is accepted by speakers without reservations. The crucial data is presented in (8c), which shows that pre-predicate constituents – in contrast to left-dislocated constituents – cannot co-occur with a coreferent pronoun *in situ*. (The configuration with two pre-predicate constituents is excluded, since this position is unique.)

(8) a. *Pèedróoh_{j=e}’ t=u y-il-ah (leti’_{j=e})⁶*

Pedro=D3 PFV=A.3 Ø-see-CMPL(B.3.SG) 3.SG=D3

ba’x h ùuch-ih.

what PFV happen-B.3.SG

‘Pedro_j, he_j saw what happened.’

b. *Pèedróoh_{j=e}’ leti’_j il ba’x*

Pedro=D3 3.SG see-(SUBJ)(B.3.SG) what

h ùuch-ih.

PFV happen-B.3.SG

‘Pedro_j, HE_j saw what happened.’

- c. *Pèedróoh_j il (*let_i'_j=e') ba'x h*
 Pedro see(SUBJ)(B.3.SG) 3.SG=D3 what PFV
ùuch-ih.
 happen-B.3.SG

‘It was Pedro who saw what happened.’

Since subject and object NPs are indexed by the person markers on the verb, they can always be omitted, leaving behind a complete clause. This does not hold for locative complements, which are not cross-referenced through person markers on the verb, as exemplified in (9a). When the locative complement is elided because it is recoverable from context, as in (9b), then a locative enclitic =*i* ‘LOC₂’ obligatorily appears at the right edge of the postverbal domain. This enclitic appears with verbs that govern a locative complement, e.g., *yàan* ‘exist’ in its locative use.

- (9) a. *yàan hun-p'éel mèsáah ichil hun-p'éel nah.*
 exist one-CL.INAN table inside one-CL.INAN house
 ‘There is a table inside a house.’

- b. {What is there inside the house?}
yàan hun-péel mèsáah(=i).*
 exist one-CL.INAN table=LOC₂
 ‘There is a table there.’

If the complement is left-dislocated, the locative enclitic is obligatory, just as when it is elided; compare (10a) and (9b). However, the enclitic is rejected if the locative complement occupies the pre-predicate position, as is the case for the interrogative pronoun in (10b) and for the focused prepositional phrase in (10c).

(10) a. {What is there inside the field?}

ichil le kool=o' yàan hun-túul kolnáal(=i').*

inside DEF field=D2 exist one-CL.AN farmer=LOC₂

‘Inside the field, there is a farmer.’

b. *tu'x yàan hun-túul kolnáal*(=i')?*

where exist one-CL.AN farmer=LOC₂

‘Where is there a farmer?’

c. {Where is a farmer?}

ichil hun-p'él kool yàan

inside one-CL.INAN house exist

hun-túul kolnáal(=i').*

one-CL.AN farmer=LOC₂

‘INSIDE THE FIELD there is a farmer.’

The data presented in (7)-(10) indicate that pre-predicate constituents (but not left-dislocated constituents) correspond to constructions with the same material *in situ*, only differing in the displacement of the constituent in the preverbal position. This difference indicates that pre-predicate

constituents are *necessarily* clause internal, while left-dislocated constituents *can* be external to the core clause. This conclusion does not exclude the possibility that Yucatec Maya displays a lower position for “internal topics”, as argued by Aissen (1992) for Tz’utujil [TZJ]. Indeed, Yucatec Maya shows some properties that support this view (see Gutiérrez Bravo 2011 for a detailed discussion). For the current discussion, the relevant issue is the clear contrast between the possibilities of left-dislocated and pre-predicate constituents.

2.3 Structure of fronting

Previous literature on Yucatec Maya has proposed two different analyses for the constructions involving a pre-predicate constituent. The first option is a *fronting* account, which quite straightforwardly reflects the overt properties discussed in the previous section (see Aissen 1992 on Mayan languages; see Lehmann 2003: 29 on Yucatec Maya). Assuming a core clause containing the verb complex and the thematically determined postverbal positions, the pre-predicate position results from the displacement of a constituent to the left periphery.

Another possible analysis is the assumption of a *cleft construction*, according to which the preverbal constituent is a non-verbal predicate whose subject is a headless relative clause (Bricker 1979: 111; Bohnemeyer 1998: 192; Tonhauser 2003: 212-214). This proposal is based on the observation that focus constructions share some properties with relative clauses, as

illustrated by the examples in (11a-b): in particular, the inflectional properties of the verb are identical in the agent-focus construction and in relative clauses with an extracted agent. Relative clauses do not have an overt complementizer; hence, the complement of the focus constituent in (11b) can be analyzed as a relative clause. Furthermore, a large number of lexical items can serve as predicates (in the spirit of the omnipredicativity hypothesis, see Launey 1994; see Vapnarsky 2013 on Yucatec Maya). For instance, a noun such as *pèek'* 'dog' can be used as predicate, bearing a set-B suffix encoding the subject; see (11c). Crucially, the set-B suffix for 3rd person is zero; hence, the preverbal noun in (11b) can be analyzed as a non-verbal predicate and the VP-constituent as a headless relative clause (Bohnenmeyer 1998, Tonhauser 2003).

- (11) a. *T=in w-il-ah hun-túul pèek'*
 PFV=A.1.SG Ø-see-CMPL(B.3.SG) one-CL.AN dog
hàant-ik lu'm.
 eat:TRR-INCMPL(B.3.SG) earth
 'I saw a dog eating earth.'
- b. *Pèek' hàant-ik lu'm.*
 dog eat:TRR-INCMPL(B.3.SG) earth
 'The / a DOG is eating earth.'
- c. *Pèek'-ech!*
 dog-B.2.SG

‘You are a dog.’

However, the parallels between relative clauses and focus constructions can be equally accommodated under both views. In terms of the clefting account, the agent-focus morphology is a property of relative clauses and the fact that the same morphology appears in focus constructions is evidence that these constructions contain a relative clause (Bohnemeyer 2009: 201). In terms of the fronting account, the morphological properties of the verb in (11a-b) are triggered by the extraction of the agent constituent, which equally applies in relative clauses and focus constructions (Gutiérrez Bravo 2011). Morphological properties are challenging for syntactic questions, but they can only be interpreted if syntactic configurations are established by independent syntactic evidence. Our question in the following is: What do the mono-clausal and the bi-clausal view imply for the syntactic properties of the construction at issue?

The fronting account implies that the structure of the canonical construction must be preserved in the fronting construction; that is, every instance of the fronting construction must correspond to a canonical construction with the same material. Cleft constructions display some differences from the corresponding canonical constructions due to the fact that they contain a headless relative clause (Akmajian 1970: 160ff.). Since the relative clause of a cleft construction is not headed by the clefted constituent,⁷ the verb of the headless relative clause is not necessarily cross-

referred by the clefted constituent; for instance, consider *it's you who is responsible*.⁸ In contrast to cleft constructions, constituent fronting implies that the agreement relations of the basic configuration must be preserved. Examples (12a-b) show that agreement with the pre-predicate subject is obligatory, which is against the view that the material following the pre-predicate constituent is a headless relative clause. Example (12c) shows that this is not the case for cleft constructions: the embedded verb is accompanied by a set-A marker of the third person, co-indexed with the subject variable of the headless relative clause and not with the clefted constituent.

- (12) a. *Tèech k=a bin tak Yaxley.*
 2.SG IPFV=A.2 go as.far.as Yaxley
 ‘YOU are going up to Yaxley.’
- b. **Tèech k=u bin tak Yaxley.*
 2.SG IPFV=A.3 go as.far.as Yaxley
 (intended) ‘YOU are the one that is going up to Yaxley.’
- c. *Tèech le k=u bin tak Yaxley=o’.*
 2.SG DEF IPFV=A.3 go as.far.as Yaxley=D.2
 ‘You are the one that is going up to Yaxley.’

Object cross-reference markers show the same pattern. They agree with the pre-predicate constituent in fronting constructions, see (13a-b), but they do not have to be co-indexed with the clefted constituent in cleft

constructions, as in (13c) (connectivity effects can arise; hence, the version of (13c) with 2nd person is acceptable, but the crucial issue is the non-grammaticality of (13b)).⁹

- (13) a. *Tèech t=in hats’-ah-ech.*
 2.SG PFV=A.1 hit-CMPL-B.2.SG
 ‘I hit YOU.’
- b. **Tèech t=in hats’-ah=eh.*
 2.SG PFV=A.1 hit-CMPL(B.3.SG)=TERM
 (intended) ‘I hit YOU.’
- c. *Tèech le t=in hats’-ah=o’.*
 2.SG DEF PFV=A.1 hit-CMPL(B.3.SG)=D2
 ‘It’s YOU that I hit.’

The same data pattern is observed in the binding possibilities of the pre-predicate constituents: a fronted constituent preserves the binding possibilities of the basic configuration. Reflexive expressions in Yucatec Maya consist of a possessive pronoun and the noun *báah* ‘self’. Subjects bind pronominal expressions in the object constituent in Yucatec Maya (Bohnenmeyer 2009, Skopeteas and Verhoeven 2009a) and this holds also for subjects in the pre-predicate position, as shown in (14a). This property does not hold true for cleft constructions, as illustrated in (14b). The subject of the headless relative clause is a third person operator, and hence these constructions may appear with third person possessors – even if the latter

are co-referent with a clefted local person (cf. English *It is you who cuts himself*; see Akmajian 1970).

- (14) a. *Tèech hats’-ik a/*u báah.*
 2.SG beat:TRR-INCMPL(B.3.SG) A.2/A.3 self
 ‘YOU are hitting yourself.’
- b. *Tèech le k=u hats’-ik*
 2.SG DEF IPFV=A.3 beat:TRR-INCMPL(B.3.SG)
u báah=o’.
 A.3 self=D2
 ‘YOU are the one that is hitting himself.’

The data pattern exemplified in (12) to (14) is clear syntactic evidence that the constituent in the pre-predicate position is extracted out of the postverbal domain, and thus it retains the structural properties that it would have *in situ*: it is co-indexed with the person marker on the verb and binds anaphoric expressions in constituents that are more deeply embedded in the syntactic structure. These facts are clear counterevidence to the analysis of these constructions as clefts.

Another crucial issue is the status of the pre-predicate constituent. An argument in favor of the view that the pre-predicate argument is a predicate comes from the placement of the interrogative particle *wáah* ‘Q’ (Bohnenmeyer 1998: 182, 192; Tonhauser 2003: 211-212; see a similar account of Malagasy [MLG] particles in Potsdam 2006). The basic data are

introduced in (15). The interrogative particle cliticizes to the predicate in polar questions, as in (15a). If a pre-predicate constituent is available, the interrogative particle may cliticize to it (this is the preferred option but not the only one), as in (15b). The interrogative particle cannot follow left-dislocated material, as in (15c), and may not appear at the beginning of the clause under the interpretation presented in (15).¹⁰

- (15) a. *T=a he'-ah=wáah le bentàana=o'?*
 PFV=A.2 open-CMPL(B.3.SG)=Q DEF window=D2
 'Did you open the window?'
- b. *Tèech=wáah he' le bentàana=o'?*
 you=Q open(SUBJ)(B.3.SG) DEF window=D2
 'Did YOU open the window?' (Bohnemeyer 1998: 192)
- c. **Pèedróoh=e'=wáah t=u he'-ah*
 Pedro=D3=Q PFV=A.3 open-CMPL(B.3.SG)
le bentàana=o'?
 DEF window=D2
 (intended) 'Did Pedro open the window?'

The data in (15) motivate the hypothesis that *wáah* 'Q' is a post-predicate particle (compare Potsdam 2006: 2164 on Malagasy). This assumption may explain why this particle accompanies the verb complex in (15a), and may lead to the conclusion that the material in the pre-predicate position is a predicate, as in (15b), but not so for the left-dislocated material, as in (15c).

However, the idea that this particle follows a syntactically determined unit (i.e., the ‘predicate’) does not account for the fact that it may occur in several positions in the clause; see (16).

- (16) a. *T-a xok-ah=wáah óox-p’él áanalte’-o’b*
 PFV-A.2 read-CMPL(B.3.SG)=Q three-CL.INAN book-PL
ho’lyak?
 yesterday
 ‘Did you read three books yesterday?’
- b. *T-a xok-ah óox-p’él=wáah áanalte’-o’b ho’lyak?*
- c. *T-a xok-ah óox-p’él áanalte’-o’b=wáah ho’lyak?*
- d. *T-a xok-ah óox-p’él áanalte’-o’b ho’lyak=wáah?*

The particle *wáah* ‘Q’ has scope over the constituent on its left: the focus of the question falls on the verb in (16a), on the numeral in (16b), on the noun phrase in (16c) and is ambiguous between a local reading (adverb focus) and a reading in which the particle has the entire sentence in its scope (truth value focus) in (16d). The hypothesis that the particle *wáah* ‘Q’ is predicate-final can be rejected on the basis of these examples. But let’s examine the possibility that this particle is predicate-final when it appears in the pre-predicate position. An illustration of alternative scopal readings in the pre-predicate position is given in (17a-b). In (17a), the particle follows the pre-predicate constituent, while in (17b) the particle follows the

quantifier. The former example is interpreted as a question about the entire noun phrase, while the latter as a question about the quantifier.

- (17) a. *óox-p'éel mèsáah-o'b=wáah ts'o'k u*
 three-CL.INAN table-PL=Q TERM A.3
man-ik Pèedróoh?
 buy-INCMPL Pedro
 'Did Pedro buy THREE TABLES?'
- b. *óox-p'éel=wáah mèsáah-o'b ts'o'k u*
 three-CL.INAN=Q table-PL TERM A.3
man-ik Pèedróoh?
 buy-INCMPL Pedro
 'Did Pedro buy THREE tables?'

The example in (17b) shows that the particle *wáah* 'Q' does not necessarily appear at the right edge of the pre-predicate constituent. An account of *wáah* 'Q' as a predicate-final particle should show that the material after the particle is a relative clause. This possibility is completely excluded on the basis of examples in which the material following the interrogative particle is not a possible constituent. If the preposition in (18) were a predicate, then the material following *wáah* should be the subject constituent. However, the alleged head *nah* 'house' of the relative clause is not a possible argument of the embedded verb.

- (18) *iknal=wáah le nah yàan hun-túul m̀is=o'?*

near=Q DEF house=D2 exist one-CL.AN cat=D2

‘Is there a cat NEAR the house?’

Interestingly, some elements are not possible predicates in Yucatec Maya (Vapnarsky 2013). The preposition *tuméen* ‘by’ cannot be used with a Set-B suffix as a predicate. This preposition can serve as a phonological host for the interrogative enclitic; see (19).¹¹ This possibility expresses focus on the preposition *tuméen* ‘by’ in contrast to other alternatives (e.g., ‘by the girl’ or ‘for the sake of the girl’).

- (19) *Tuméen=wáah hun-túul xch’úupal túun*
 by=Q one-CL.AN F:woman:child=Q PROG:A.3.SG
 léench’in-t-a’l le chan mèsásah=o’?
 push-TRR-PASS.INCMPL DEF small table=D2
 ‘Is the small table pushed BY a girl?’

The crucial point for our analysis is that the interrogative particle *wáah* ‘Q’ does not strictly follow the predicate, and hence it is not compelling evidence that the pre-predicate constituent is a predicate. The particle may occur in several positions in the utterance depending on the intended scopal interpretation, as illustrated in (16)-(17), and may occur within the pre-predicate constituent in positions that cannot be predicate-final; see (18)-(19). Hence, the facts provided through this particle do not challenge the view that the pre-predicate position is the landing site of a fronting operation. A last question is what determines the position of this particle; we

will come back to this question in Section 4 after introducing the necessary facts about the intonational domains to which the placement of this enclitic refers.

2.4 Summary

This section has shown that the left periphery in Yucatec Maya contains at least two distinct syntactic configurations, i.e., dislocation to the left side of the clause and fronting to a pre-predicate position. Purely syntactic facts indicate that left-dislocated material is outside the clause while pre-predicate constituents are part of the clause (Section 2.2). The fronting construction and the relative clauses share some inflectional properties that appear when an agent constituent is extracted. These properties motivate the idea that the fronting operation is a cleft construction. However, the crucial syntactic properties (namely, the cross-reference and binding possibilities) show that pre-predicate constituents display exactly the restrictions that are expected for clausal constituents, which excludes a bi-clausal analysis. A further argument in favor of the cleft analysis relates to the placement of the question particle under the assumption that this particle must follow a predicate; we have presented rich evidence that this is not the case. Another argument relates to the use of a subordinator under future time reference. The presence of a subordinator indicates that the construction is bi-clausal, but this fact does imply that the instances of fronting without subordinator are also bi-clausal (in particular since there is syntactic evidence to the

opposite). A last relevant issue is a restriction on phrases with a definite marker, which is discussed in Section 3.1.

3 Information-structural properties

The terms ‘focus construction’ or ‘focus position’ in previous research on Yucatec Maya are based on the observation that constituent fronting occurs in contexts in which the constituent at issue is focused. In view of current discussion on information structure (Section 1), the critical question is whether focus is associated with the syntactic construction or arises through inferences that apply in particular contexts. The first question is whether focus is a necessary condition for a constituent to be fronted; see Section 3.1. The next question is whether focus-fronting relates to a particular subtype of focus (e.g., contrastive focus); this question will be examined with intuition and corpus data in Section 3.2 and with speech production data in Section 3.3.

3.1 Focus domain

In general, our facts show that focus is a necessary condition for a constituent to appear in the pre-predicate position. The information structural distinction between pre-predicate and left-dislocated constituents is clear-cut: the former but not the latter contains the focus of the utterance (with the exception of a limitation on definite noun phrases, to be discussed below). While the share of burden in the left periphery is very robust, the functional delimitation between preverbal and postverbal focus is less clear.

The pre-predicate constituents always contain focused arguments or adjuncts, i.e., instances of narrow focus.¹² There are three limitations to this generalization that are discussed in the following: (a) the failure of co-extensivity, (b) the constraints on definite noun phrases, and (c) the role of ‘informative-presupposition’ constructions.

The focus domain is not always co-extensive with the constituent in the pre-predicate constituent. Violations to strict co-extensivity of the pre-predicate constituent with the focus domain frequently arise through limitations on the extraction possibilities. For instance, it is not possible to extract subconstituents of the noun phrase, even if these subconstituents are focused, as exemplified in (20). A discontinuous noun phrase with the quantifier in the pre-predicate position is not grammatical, as illustrated in (20a), due to a restriction on extraction that is independent of the contextual properties.¹³ The well-formed answer to this question is (20b), in which the entire noun phrase is pied-piped to the pre-predicate position (Aissen 1999 on Tzotzil).

(20) {How many languages do you know?}

- a. **ka'-p'éel* *in* *w-ohel* *t'àan*.
 two-CL.INAN A.1 0-know speech
 (intended) ‘I know TWO languages.’
- b. *ka'-p'éel* *t'àan* *in* *w-ohel*.
 two-CL.INAN speech A.1 0-know

‘I know TWO languages.’/‘I know TWO LANGUAGES.’

Thus, the focus of the utterance can be *a part of* the pre-predicate constituent, if restrictions on extraction do not allow for fronting the exact unit in focus. A further deviation from the association of the left-peripheral positions with information structure arises from a constraint on clitic placement. Definite noun phrases are obligatorily accompanied by a right-edge clitic, which is a member of a set of three elements, =*a* ‘D1’ (localization of the referent in the proximal region of the deictic center), =*o* ‘D2’ (distal region), and =*e* ‘D3’ (deictically empty). Crucially, these enclitics are not contiguous with the noun phrase but appear at the right edge of the intonational phrase (see Section 4). In the left-periphery, we observe that they occur at the right edge of a left-dislocated constituent – see (2) and (41) – but not at the right edge of a pre-predicate constituent (Lehmann 1990: 44; 2003: 28); see (21a-b).

(21) a. {Who ate the avocado?}

**Le ah kòonol=o’ hàant òon.*

DEF M seller=D2 eat:TRR(SUBJ)(B.3.SG) avocado

(intended) ‘The SELLER ate the avocado.’

b. {Where is a cat?}

#*t=u y-iknal le nah=o’ yàan hun-túul mìis.*

LOC=A.3 Ø-near DEF house=D2 exist one-CL.AN cat

(intended) ‘There is a cat NEAR THE HOUSE.’

The restriction at issue does not relate to the referential properties of the noun phrase, since proper nouns, as in (4a-b), and personal pronouns, as in (8b), may occur in the pre-predicate position. The restriction only applies in the subset of definite descriptions that involve a right-edge clitic, indicating that the crucial factor is the enclitic and not the referential properties. Indeed these enclitics delimit intonational phrases; the right edge of a left-dislocated constituent is compatible with an intonational boundary while the right edge of a pre-predicate constituent is not (see further discussion in Section 4). A solution to this conflict is to realize the enclitic at the end of the clause, as illustrated in (22a-b) (Aissen 1992: 56 on Tzotzil). Native speakers accept this construction when the pre-predicate constituent is a prepositional phrase, as in (22b), while they are reluctant to accept the same construction with noun phrases, as in (22a), with the explanation that “this is not a complete sentence”. We speculate that this intuition comes from the obvious similarity of (22a) with a relative clause (note that a relative clause interpretation is also possible for (22b)).

(22) a. {Who ate the avocado?}

**Le ah kòonol hàant òon=o’.*

DEF M seller eat:TRR(SUBJ)(B.3.SG) avocado=D2

(intended) ‘The SELLER ate the avocado.’

b. {Where is a cat?}

t=u y-iknal le nah yàan hun-túul mìis=o’.

LOC=A.3 Ø-near DEF house=D2 exist one-CL.AN cat=D2

‘There is a cat NEAR THE HOUSE.’

Native speakers use a different construction in expressions of focus on definite noun phrases; see (23). This construction involves left dislocation of the definite noun phrase and a co-referent third person pronoun in the pre-predicate position.¹⁴ This construction resolves the conflict between focus fronting and the constraint on clitic placement. The result is a construction in which focused material is left-dislocated instead of being fronted.

(23) {Who ate the avocado?}

Le ah kòonol=o’ leti’ hàant òon.

DEF M seller=D2 3.SG eat:TRR(SUBJ)(B.3.SG) avocado

‘The seller, HE ate the avocado.’

A further construction occurring in this context involves the third person pronoun at the beginning of the clause accompanied by a relative clause headed by the referent in focus; see (24). This construction has a different syntax. It does not allow an analysis as focus-fronting, i.e., it is unambiguously bi-clausal. The pronoun *leti’* ‘3.SG’ is the predicate of the matrix clause.

(24) {Who ate the avocado?}

leti’ le ah kòonol hàant òon=o’.

3.SG DEF M seller eat:TRR(SUBJ)(B.3.SG) avocado=D2

‘It is the seller who ate the avocado.’

A challenge for the generalization that the pre-predicate constituent is a focus domain relates to the occurrence of fronting in contexts in which the complement of the pre-predicate position is informative. This phenomenon has already been observed for English *it*-clefts (termed ‘informative presupposition *it*-clefts’; Prince 1978: 898-903; Hedberg 2000: 902f.; Delin 1992) and also in languages that express focus through fronting operations (e.g., Hungarian in Wedgwood 2009). Example (25) shows that fronting is attested in similar contexts in Yucatec Maya. The pre-predicate position contains a third person pronoun relating to the referent that is the running discourse topic (the devil H-Wayá’s), while the complement of the pre-predicate position expresses information that is not yet introduced. These constructions present information that is new to the hearer as being a ‘fact’, i.e., as being presupposed information – even if it is not part of the established common ground (Prince 1978: 899). The use of a construction in this context invokes the interpretation that potential alternatives of the given constituent are excluded (Hartmann 2012), which is compatible with the analysis of this constituent as narrowly focused, although the partition of given and new information deviates from the typical case.

- (25) ‘It is said that there lived a snake with wings. Its name was Hapai Kan. (...) H-Wayá’s is guarding it. That H-Wayá’s is a devil whose hair is dirty (...)’

leti' kaxant-ik mehen pàal-al u bis

3.SG seek-INCMPL small child-COLL A.3 carry

ti' u háap-eh.

LOC A.3 gulp-SUBJ

‘It’s him who is looking for the little children in order to bring them to it for eating.’ (HAPAIKAN_014)

Summing up the facts discussed in this section, we can maintain the generalization that narrow focus is a necessary condition for fronting to the pre-predicate position. The lack of co-extensivity is not counterevidence, but shows that discourse-licensed operations respect syntactic restrictions (on extraction of constituents). The choice of different constructions in order to avoid violations of the rules of clitic placement is a further phenomenon of the same type. The facts from informative-presupposition constructions are an expected possibility of a focus-related operation.

3.2 Focus interpretation

The question of this section is whether fronting to the pre-predicate position may appear with any focused constituent or is associated with a particular type of focus (as proposed for several languages: É. Kiss 1998 for Hungarian, Aissen 1992 for Tzotzil, Trechsel 1993 for K’iche’ [QUC]). The question under discussion has important implications for the further analysis. If fronting is triggered by a particular subtype of focus (such as contrastive focus), we will conclude that the pre-predicate position is associated with an operator with propositionally relevant content, i.e., an

identificational operator. If an underspecified concept of focus is at issue, then it may be that the discourse partitions of the utterance (focus and background) can be deduced on the basis of properties of greater generality – related to the syntactic and prosodic properties of the pre-predicate position. The aim of this section is to examine the hypothesis in (26) for Yucatec Maya.

(26) Hypothesis of association with an identificational operator

A constituent α occurs in the pre-predicate position iff the referent of α (or of a part of it) is identified to the exclusion of all relevant alternatives.

We compare the interpretational properties of fronting with cleft constructions associated with an exhaustive interpretation. In these constructions, the presupposed information is introduced by the definite determiner. The definite description gives rise to an existential presupposition (of a set of referents of this description) and the clefted constituent is equated to the set of referents of this description via the zero copula.¹⁵ Let's assume a simple context which introduces a set of referents in discourse, as illustrated in (27).

- (27) *Ichil le nah=o' yàan bu'l, yàan ixi'm*
in DEF house=D2 EXIST bean EXIST corn
yéetel papas=i' ...
and potatoes=LOC₂

‘There are beans, corn, and potatoes in the house...’

As soon as these entities are introduced into the common ground, the set $R = \{\text{beans, corn, potatoes}\}$ is part of the set of contextually available referents in discourse. In this context, the sentences in (28a-b) illustrate two possible assertions about an individual which is already part of the implicit common ground such that it may be referred to by a proper noun. The sentences in (28a-b) contain a member of R *in situ*, as in (28a), and in the pre-predicate position, as in (28b). These versions indeed invoke different interpretations: when the constituent is realized *in situ*, as in (28a), native speakers do not exclude that the referent is not the exhaustive subset of the relevant referents in R for which the presupposition holds: it is possible that ‘Deysi ate something else, too’. However, if the same constituent is placed in the pre-predicate position, the preferred interpretation is that the referent is the exhaustive subset of the set R for which the presupposition holds true. Hence, the focus construction in example (28b) induces the interpretation that ‘beans’ is the only member of R that fills the variable of the presupposition ‘Deysi ate x’.

(28) In the context of (27):

a. ... *Deysi=e*’ *k=u* *hàant-ik* *bu’l*.

Deysi=D3 IPFV=A.3 eat:TRR-INCMPL(B.3.SG) bean

‘... Deysi is eating beans.’ (native speaker: ‘it is possible that she is eating other things, too.’)¹⁶

b. ... *bu'l k=u* *hàant-ik* *Deysi.*

bean IPFV=A.3 eat:TRR-INCMPL(B.3.SG) Deysi

‘...Deysi is eating BEANS.’ (native speaker: ‘this time, she eats only beans.’)

The minimal pair in (28a-b) reveals an interpretational difference with respect to the exhaustive identification. However, the exhaustive interpretation of (28b) may well be the result of a pragmatic inference arising through the fact that a constituent is in focus. The hearer of the utterance seeks a functional motivation that may account for the fact that the speaker selected an expression in which the object is placed in a prominent position in the clause, as in (28b), instead of an unmarked expression, as in (28a). In this view, the exclusion of alternatives is only a possibility among an array of interpretations of the speaker’s intention. If this is the case, there should be some contexts in which this interpretation does not arise.¹⁷ This hypothesis is borne out, as shown in example (29). Yucatec Maya speakers share the knowledge that the proposition ‘a turkey eats beans’ is less likely than the proposition ‘a turkey eats corn’. If the less likely proposition is the case, the fact that beans is highly salient in this particular situation is an effective motivation to license the placement of the object in the pre-predicate position, as exemplified in (29).

(29) In the context of (27):

... *bu'l k=u* *hàant-ik* *le úulum=o’.*

bean IPFV=A.3 eat:TRR-INCMPL(B.3.SG) DEF turkey=D2

‘The turkey is eating BEANS.’ (native speaker: ‘it may have eaten corn too.’)

The important point is that the inference of exhaustivity does not arise in this context. In line with the rationale sketched out above, the interpretational properties are inferences of a hearer seeking a functional motivation for the fact that the speaker has chosen an expression involving a narrow focus domain. If this motivation is provided by the fact that the involvement of a certain referent is salient for the event at issue, as in (29), then the inference of exhaustivity does not arise. The hearer’s rationale is the following: “The speaker preposed the ‘beans’ since turkeys do not normally eat beans, and hence this does not imply that the ‘turkey’ did not eat anything else”.

The intuition reported in (29) differs from the interpretation of cleft constructions with a relative clause introduced by the definite determiner. For our purposes, this construction serves as a control condition which confirms that the informants consider the semantic properties of the construction at issue and do not just draw conclusions from the context. The crucial finding is that the exhaustive interpretation of the cleft construction is not affected by the context; compare (29) with (30).

(30) In context of (27):

... *bu’l* *le* *k=u* *hàant-ik*

bean DEF IPFV=A.3 eat:TRR-INCMPL(B.3.SG)

Deysi=e’/ le úulum=o’.

Deysi=D3 DEF turkey=D2

‘...Beans is what Deysi/the turkey is eating.’ (native speaker: ‘this time, Deysi/the turkey eats only beans.’)

In conclusion, evidence from interpretation shows that placement in the pre-predicate position may invoke an exhaustive interpretation, but this effect depends on particular contextual conditions. We have identified contexts that do not give rise to an exhaustive interpretation, and hence we conclude that this interpretational effect is not an inherent property of the syntactic configuration.

Further evidence for the association of a syntactic configuration with a semantic feature comes from contradiction effects. If the pre-predicate position in Yucatec Maya were associated with an exhaustive operator, we would expect a distributional restriction on the use of *also*- and *even*-phrases in this position, since these particles involve the presupposition that the focused referent is a member of a set of alternatives for which the presupposition holds true (É. Kiss 1998: 251-253).¹⁸ Yucatec Maya does not display any restriction on the occurrence of *also*- and *even*-phrases in the pre-predicate position, as shown in (31). Both the particles *xan* ‘also’ and *tak xan* ‘even’ give rise to an existential implicature that there is another referent (beyond the asserted one) for which the proposition holds true. This

is also the crucial difference between these particles and the focus fronting construction: the existence of an alternative is not coercive in the latter case.

- (31) a. *wàah xan k=u k'áat-ik le*
 tortilla also IPFV=A.3 ask-INCMP(L(B.3.SG) DEF
h-mèen=o'.
 M-curer=D2
 'The curer also is asking for TORTILLA.'
- b. *tak xan Pèedróoh k'áat-ik wàah=o'*.
 as.far.as also Pedro ask-INCMP(L(B.3.SG) tortilla=D2
 'EVEN PEDRO is asking for tortilla.'

The effects of the distributional restriction at issue can be observed in cleft constructions, as in (32). The inherent exhaustivity of this construction is not compatible with the semantics of *also* and *even*.

- (32) a. **wàah xan le k=u k'áat-ik*
 tortilla also DEF IPFV=A.3 ask-INCMP(L(B.3.SG)
le h-mèen=o'.
 DEF M-curer=D2
 (intended) 'It is also tortilla that the curer is asking for.'
- b. **tak xan Pèedróoh le k'áat-ik wàah=o'*.
 as.far.as also Pedro DEF ask-INCMP(L(B.3.SG) tortilla=D2
 (intended) 'It is even Pedro that is asking for tortilla.'

In conclusion, the examination of speakers' intuitions has shown that the interpretation of the pre-predicate constituent as excluding alternatives depends on the context – compare (28b) and (29) – while this is not the case for cleft constructions, as in (30). Distributional evidence supports the view that the pre-predicate position is not associated with an exhaustivity operator, since no contradiction arises from the use of particles like *xan* 'also' and *tak xan* 'even' in this position (in contrast to the compared cleft constructions). Based on the presented facts, we reject the hypothesis that the pre-predicate position is inherently associated with a semantic operator that leads to the exclusion of relevant alternatives, as stated in (26). We maintain the generalization in Section 3.1: any type of narrow focus can trigger fronting to the pre-predicate position – insofar this does not lead to a violation of syntactic rules.

3.3 Occurrence in context

The aim of this section is to examine the factors that determine the choice of focus constructions in speech production. We discuss two factors that are known to influence the choice of focus operations in general: the different focus types and the different focus domains. An array of studies indicate that non-canonical constructions for the expression of focus are more likely to occur if focus entails a stronger revision of the common ground, e.g., it is more likely in contrastive than in new information contexts, as summarized in (33).

(33) Strength of the common ground revision

The likelihood of a constituent α to appear in the pre-predicate position correlates with the strength of the revision of the common ground that is implied by asserting α .

A further correlation reported for several languages is an asymmetry relating to different focus domains. Focus on subjects is more likely to be expressed through non-canonical syntactic constructions than focus on non-subjects (French [FRA] in Lambrecht 2001; Spanish [SPA] in Buring and Gutiérrez Bravo 2001; Hausa [HAU] in Hartmann and Zimmermann 2006; several West African languages in Fiedler and Schwarz 2005, Fiedler et al. 2010; Northern Sotho [NSO] in Zerbian 2006, etc.).

(34) Influence of focus domains

The occurrence of a constituent α in the pre-predicate position depends on the syntactic properties of α , such that focus on subjects is more frequently realized *ex situ* than focus on non-subjects.

The examination of the hypotheses in (33)-(34) is crucial for our purposes. If such asymmetries apply in Yucatec Maya, then we should ask where such phenomena come from; that is, whether they imply that particular discourse features are associated with constituent structure and how they relate to the intuitions reported in 3.2.

3.3.1 Method

The methodological approach reported in the following is part of a fieldwork tool developed for the investigation of information structure, namely the *Questionnaire on Information Structure* (Skopeteas et al. 2006). Each informant was shown four printed pictures and (s)he was instructed to observe the presented scenes carefully. When (s)he was ready, the pictures were taken away, and four pre-recorded questions related to the pictures were played. The speakers were instructed to give “full” answers to the questions at the beginning of the field session. This task was repeated four times in each experimental session: each informant was shown 4 sheets (containing 4 pictures each) and gave 16 answers in total.

The questions were designed to induce different types of context and different focus domains, as exemplified in (35). The factor CONTEXT TYPE contains three contexts that are part of the classification of Dik et al. (1981) and Dik (1997): (a) ‘completion’ refers to an answer to a constituent question, (b) ‘selection’ refers to an answer to an alternative question, and (c) ‘correction’ refers to the rejection of a presupposition contained in a polar question. A fourth context, namely ‘confirmation’, serves as a control condition: it offers a baseline corresponding to the behavior of the speaker if s/he does not wish to revise the assumptions of the addressee. The factor CONTEXT TYPE is crossed with two FOCUS DOMAINS (subject focus and object focus), which gives a set of $2 \times 4 = 8$ permutations; see (35).

(35) Experimental conditions

(C=completion, S=selection, R=correction; F=confirmation;
s=subject; o=object)

Stimulus: 'In front of a well, a man is pushing a car.'

C/s: 'In front of the well, who is pushing the car?'

C/o: 'In front of the well, what is the man pushing?'

S/s: 'In front of the well, is a man or a woman pushing the car?'

S/o: 'In front of the well, is the man pushing a car or a bicycle?'

R/s: 'In front of the well, is a woman pushing the car?'

R/o: 'In front of the well, is the man pushing a bicycle?'

F/s: 'In front of the well, is a man pushing the car?'

F/o: 'In front of the well, is it a car that the man is pushing?'

The translations of all these questions in Yucatec Maya involved a constituent in the pre-predicate position, either the noun phrase of the polar and alternative questions or the interrogative pronoun of the constituent questions. Hence, differences between the conditions in (35) cannot be traced back to possible effects of structural priming by the form of the question.

If the proportions of *ex situ* focus depend on the strength of the common ground revision, as stated in (33), then we expect that the frequency of *ex situ* focus will correspond to the scale in (36). The contribution of the answer is minimal if the answer only confirms a hypothesis that the utterer of the question already introduced to the common ground. The contribution

of the answer is higher in selective contexts, in which the question introduces a set of two alternatives, and even higher in completive contexts since the set of possible alternatives is larger (contains all relevant referents in discourse). The contribution of the answer is maximal in corrective contexts, in which a member of the set of relevant referents is asserted (as in the completive context) and additionally the hypothesis of the utterer of the question has to be revised.

(36) confirmation < selection < completion < correction

Twelve native speakers of Yucatec Maya, all inhabitants of Quintana Roo, Mexico, participated in the experiment (5 female; mean age 29.8; age range 17-57 years). Each native speaker was confronted twice with each condition in (35). The field sessions contained pseudo-randomized tasks from several speech production tasks (total session duration: approximately 45 min.).

3.3.2 Results

The obtained dataset contains $12 \text{ (speakers)} \times 8 \text{ (conditions)} \times 2 \text{ (answers)} = 192$ answers. Thirty-seven answers (19.3%) were coded as “non-valid”, because they did not correspond to the intended condition, either because the informant misinterpreted some stimuli or because (s)he failed to retrieve the intended scene from memory (“non-valid” in Table 1). Furthermore, we elicited thirty elliptical answers (19.4% of 155 valid) without an overtly realized verb that do not provide evidence for the position of the focused

constituent (“elliptical” in Table 1). A further subset of ten answers (8% of 125 “full” answers) involved a biclausal construction, as exemplified in (37): the presupposed part of this answer is nominalized through the definite article and the NP-final enclitic *=o* ‘D2’. The clefted constituent is always the focused argument in our data; however, since this construction occurs rarely in general (“biclausal” in Table 1), we cannot draw reliable conclusions about its correlation with the contexts at issue.

(37) Q. ‘Is the man carrying a girl?’

A. *bèey hun-túul xch’úuppal le k=u*
 thus one-CL.AN F:woman:child DEF IPFV=A.3
bis-ik=o.
 carry-INCMPL(B.3.SG)=D2

‘(It is) like a GIRL what he is carrying.’ (F/o: “biclausal”)

The remaining subset of 115 answers is the dataset with which we will test the hypotheses in (33)-(34). A large proportion of the answers in all examined conditions involved placement of the narrow focused argument in the pre-predicate position, as exemplified in examples (38)-(41). The argument which is part of the background is either realized postverbally, as in (38)-(39), or elided, as in (40), or left-dislocated, as in (41) (all instances of left-dislocation in Table 1 are preverbal constituents accompanied by an enclitic). The alternative construction is to realize the constituent under

question *in situ*, as illustrated in (42). The findings in Table 1 show that this option is only attested in object questions.

(38) Q. ‘Is a woman pushing the man?’

A. *ma’, hun-túul máak tul-ik le xib=o’.*
 NEG one-CL.AN person push-INCMP(LB.3.SG) DEF man=D2
 ‘No, a PERSON is pushing the man.’ (R/s: “S_PVO”)¹⁹

(39) Q. ‘Is it a table that the man is pushing?’

A. *óolbey hun-p’éel k’àan-che’*
 seemingly one-CL.INAN hammock-wood
k=u kohche’kt-ik le máak=o’.
 IPFV=A.3 push:foot:TRR-INCMP(LB.3.SG) DEF person=D2
 ‘It seems that it is a CHAIR that the person is pushing.’
 (R/o: “O_PVS”)

(40) Q. ‘Is a man or a woman cutting the melon?’

A. *hun-túul xìib xot-ik.*
 one-CL.INAN man cut-INCMP(LB.3.SG)
 ‘A MAN is cutting it.’ (S/s, “S_PV”)

(41) Q. ‘Is the woman hitting a window?’

A. *ma’, le xch’úup=o’ hun-p’éel k’àax k=u*
 NEG DEF F:woman=D2 one-CL.INAN wood IPFV=A.3
lox-ik.
 box-INCMP(LB.3.SG)

‘No, the woman, she is hitting a (piece of) WOOD.’

(*S/o*: “S_LO_PV”)

(42) Q. ‘What is the man pushing?’

A. *túun tul-ik hun-p’éeł kamyòon.*

PROG:A.3 push-INCMPL(B.3.SG) one-CL.INAN pick.up

‘He is pushing a pick-up.’ (*C/o*: “VO”)

Deictic enclitics cannot intervene between the pre-predicate constituent and the clause (see Section 3.1 and prosodic analysis in Section 4). Focus on definite noun phrases leads to a conflict, since definite noun phrases must be accompanied by a deictic enclitic; see (21). As already shown in (23) and (24), this conflict can be resolved by alternative constructions in which the definite noun phrase does not appear in the pre-predicate position. Example (43) illustrates such a construction, in which the referent in focus is introduced in a first clause, while the target utterance contains a pronoun that is coreferent with the introduced noun phrase. This strategy is attested in our experimental data in three utterances (found in the conditions *C/s*, *F/s*, and *S/s*); see (43). The utterances at issue are coded as having a pre-predicate constituent in Table 1.

(43) Q. ‘Who is looking at the girl?’

A. *leti’ le máak=o’ leti’ pakt-ik*

3.SG DEF person=D.2 3.SG watch-INCMPL(B.3.SG)

le x-ch’úupal=o’.

DEF F-woman:child=D.2

‘This person is it, HE is looking at the girl.’ (C/s: “S_PVO”)

Table 1. Results of the production experiment

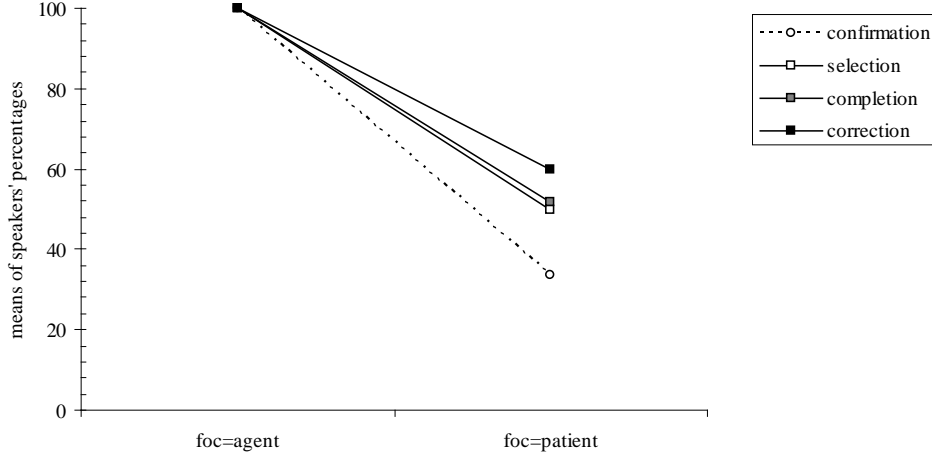
(X_P = constituent in the pre-predicate position; X_L= left-dislocated constituent)

focus = subject										
	COMPLETION		SELECTION		CORRECTION		CONFIRMATION		TOTAL	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	24		24		24		24		96	
non-valid	1		6		10		9		26	
valid	23		18		14		15		70	
elliptical	6		2		5		5		18	
full	17		16		9		10		52	
biclausal	3		0		1		1		5	
monoclausal	14		16		8		9		47	
S _P VO	7	50.0	9	56.3	6	75.0	5	55.6	27	57.4
S _P V	7	50.0	7	43.8	2	25.0	4	44.4	20	42.6
focus = object										
	COMPLETION		SELECTION		CORRECTION		CONFIRMATION		TOTAL	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	24		24		24		24		96	
non-valid	0		3		7		1		10	
valid	24		21		17		23		85	
elliptical	3		2		2		5		12	
full	21		19		15		18		73	
biclausal	2		2		0		1		5	
monoclausal	19		17		15		17		68	
O _P VS	7	36.8			2	13.3	1	5.9	10	14.7
O _P V	3	15.8	8	47.1	6	40.0	3	17.6	20	29.4
S _L O _P V					1	6.7	1	5.9	2	2.9
S _L VO	3	15.8	4	23.5	1	6.7	5	29.4	13	19.3
VO	6	31.6	5	29.4	5	33.3	7	41.2	23	33.8

The distribution of word order possibilities in Table 1 reveals a clear distinction in the contextual conditions of the left-peripheral configurations. Left-dislocated constituents (X_L) do not host the narrow focused constituent, while pre-predicate constituents (X_P) only occur with narrow focused constituents. However, we observe two alternative realizations of the constituent under question: either in the pre-predicate position (S_P in subject focus and O_P in object focus) or postverbally (O in object focus). Thus, the material under question may appear *in situ* only if it is an object, which indicates that fronting to the pre-predicate position depends on the FOCUS DOMAIN.

The next question is whether the use of the pre-predicate position is sensitive to the factor CONTEXT TYPE. In the case of subject focus, the result is categorical (100% $S_PV(O)$), and hence the potential contrast between context types is neutralized. In the case of object focus, we observe that the sum of proportions of answers with the focused constituent in the pre-predicate position (i.e., $O_PVS + O_PV + S_L O_PV$) slightly differs depending on CONTEXT TYPE. The proportions of utterances with the a pre-predicate focus – aggregated per speaker – are the following: (a) correction 60%, S.E.: ± 14.5 , (b) completion 51.9%, S.E.: ± 12.2 , (c) selection 50%, S.E.: ± 15.1 , and (d) confirmation 33.8%, S.E.: ± 15.7 (visualized in Figure 1). The obtained hierarchies descriptively confirm the prediction in (36).

Figure 1. Placement in the pre-predicate position (aggregated per speaker)



We carried out a generalized linear mixed model with FOCUS DOMAIN and CONTEXT TYPE as fixed factors and SPEAKER and ITEM as random factors (using the *glmer* function from R's *lme4* library, Bates et al. 2012). We started from a maximal model containing all permutations of both random factors with the fixed factors and their interaction effects and we stepwise reduced the model by comparing the model fit (AIC) with a maximum likelihood test. Our final model contained the fixed factors and the effect of speakers. Removing the interaction effect from the model does not have a significant impact and results in a model with better fit: AIC = 97.6 (without interaction effect) vs. AIC = 103.6 (with interaction effect) (a likelihood test results in a $\chi^2(3) = .001$, associated with a non-significant *p*-value). Likewise, removing the effect of CONTEXT TYPE does not have a significant impact (likelihood test on the model fit: $\chi^2(3) = 2.7$, non-significant *p*-value). The analysis reveals that the only crucial effect is the effect of FOCUS

DOMAIN on the choice of an answer with a pre-predicate focus: removing this effect from the model has a significant impact (likelihood test on the model fit: $\chi^2(1) = 51.3, p < .001$). Hence, these findings show that we do not have evidence that the influence of the different types of focus have a significant impact on the choice of the fronting operation in Yucatec Maya. The finding that needs an interpretation is the highly significant effect of FOCUS DOMAIN.

3.3.3 Discussion

All types of narrow focus examined induce an answer with the focused argument in the pre-predicate position. In the case of subject focus, the *ex situ* strategy is the only attested pattern. In the case of object focus, we obtained answers with the constituent at issue *ex situ* and *in situ*. The proportions of fronting in object focus differ across CONTEXT TYPE; however, these differences are not significant, although the descriptive pattern corresponds to the expectations in (36). The absence of a significant effect of CONTEXT TYPE is informative if we compare it with the corresponding results in American English, Québec French and Hungarian. Native speakers of these languages participated in the same experiment (same instructions, same experimental material, data reported in Skopeteas and Fanselow 2010). For American English, we obtained cleft constructions only in the contexts inducing subject focus. Crucially with respect to the CONTEXT TYPE, American English speakers produced cleft constructions

only in corrective contexts. In contrast to American English, clefts in Québec French (subject focus) and fronting in Hungarian (either on the subject or on the object) do not show an influence of CONTEXT TYPE. This comparison shows that fronting in Yucatec Maya belongs to the constructions that are triggered by any type of narrow focus. Thus, speech production data strengthen the conclusions of Section 3.2, that there is no evidence that fronting to the pre-predicate position is associated with identificational properties.

The effect of FOCUS DOMAIN involves an asymmetry between subject and object focus, which is reminiscent of the reported findings in a large number of languages (see references at the beginning of this section) and confirms the expectations of the hypothesis in (34). A different type of asymmetry between subject and object focus is obtained through the same experimental procedure in American English and Québec French cleft constructions (Skopeteas and Fanselow 2010). In the same contexts, American English and Québec French speakers never use cleft constructions in object focus and use a proportion of cleft constructions in subject focus (which in the context ‘correction’ is 28.5% for American English and 74% for Québec French). The data pattern of Yucatec Maya is different: questions inducing subject focus are always answered with a fronting construction, but questions inducing object focus can be also answered with the material under question in the postverbal domain.

In order to understand this asymmetry, we have to take into account the set of possible paradigmatic alternatives for each answer type. Given that the constituents under question may be fronted to the pre-predicate position and that the background constituents may be left-dislocated, four paradigmatic possibilities compete in both cases, as illustrated in (44).

(44) a. Paradigmatic alternatives for subject-focus:

focus (S)			
		fronted	<i>in situ</i>
given (O)	left-dislocated	O _L S _P V	O _L V _S
	<i>in situ</i>	S _P V _O	V _O S
attested: S _P V _O (see Table 1)			

b. Paradigmatic alternatives for object-focus:

focus (O)			
		fronted	<i>in situ</i>
given (S)	left-dislocated	S _L O _P V	S _L V _O
	<i>in situ</i>	O _P V _S	V _O S
attested: S _L V _O , O _P V _S , S _L O _P V (see Table 1)			

There is ample evidence that V-initial orders with two postverbal NPs are strongly dispreferred in discourse (see discussion in Section 1.2). These observations are in line with the fact that the VOS option in (44a-b) does not occur at all in our dataset. The assumption of a constraint against two postverbal noun phrase arguments predicts that subjects *in situ* will occur if

the object is not realized in the postverbal domain, which is the case in Table 1: O_PVS is the most frequent option of utterances with an object in the pre-predicate position and a lexically realized S.

Our findings reveal an asymmetry in the X_LVY possibilities: we found many instances of left-dislocated subjects (S_LVO: 15 tokens; S_LO_PV: 2 tokens), but no instance of the corresponding construction with left-dislocated objects (O_LVS/O_LS_PV: 0 tokens). Left-dislocation is possible for any argument and adjunct, i.e., both S_LV and O_LV are grammatical possibilities. However, they involve an asymmetry: S_LVO is the most frequent configuration in discourse, while O_LVS is a very rare construction. When the left-dislocated constituent is the patient of a transitive verb, the speakers most frequently use passive voice (Skopeteas and Verhoeven 2009b). The asymmetry between left-dislocated subjects and objects in Table 1 is in line with this phenomenon.

We speculate that this difference in left dislocation is also the source of the asymmetry in focus fronting in our data. There is no independent motivation for the asymmetry between fronting the subject and fronting the object, but there are independent reasons against O_LV and VOS, which implies that S_PVO is the only possibility for focusing a subject.

Based on these considerations, we conclude that the subject/object asymmetry that underlies the significant main effect of FOCUS DOMAIN reported for the data in Figure 1 is the result of an interaction with

independent syntactic properties of the language at issue. The obligatory fronting of focused subjects is not the effect of a discourse rule applying to focused subjects, but results from the sub-optimality of the paradigmatic alternatives that are available in order to express this information-structural configuration, i.e., the suboptimality of VOS and the suboptimality of left-dislocated objects in configurations with two non-local arguments.

3.4 Summary

The data presented in this section show that fronting to the pre-predicate position in Yucatec Maya does not depend on a particular type of focus. Intuition data shows that the possible identificational interpretation depends on the context, i.e., it does not arise in all contexts; speech production data shows that fronting occurs in all contexts involving narrow focus, without being sensitive to the exact type of focus.

We conclude that the fronting operation is an expression of narrow focus in Yucatec Maya. Restrictions arise in conflicting situations with independent syntactic rules, as shown by the instances of pied-piping and the particular constructions for the expression of focus on definites in Section 3.1. A challenging fact is a subject/object asymmetry observed in the frequencies of fronting in speech production. We argued that this asymmetry does not reflect a discourse asymmetry between subjects and objects, but results from constraints on the set of available structural options for the expression of focus.

4 Prosodic properties

Section 3 concludes that the pre-predicate position hosts the focus of the utterance. A straightforward account of the observations in the left periphery would be to assume that the two syntactic configurations in this domain are associated with distinct information-structural features, as indicated in (45):

- (45) a. left-dislocated XP → –focus
b. pre-predicate XP → +focus

The statements in (45) are violated if independent syntactic rules apply (see the restrictions discussed in Section 3.1). However, we should examine the question of whether the generalizations in (45) can be derived from underlying properties of Yucatec Maya; that is, whether they can be traced back to statements of greater generality. Recent research on the prosody-syntax interface shows that focus-driven deviations from canonical word order can be explained if one takes into account properties of prosodic structure (Szendrői 2001 on Hungarian, Koch 2008 on Thompson River Salish). As a starting point, we assume a premise established in several theories of focus, as stated in (46); see Truckenbrodt (1995), Büring (2009).

- (46) Focus targets the maximally prominent position within the relevant prosodic domain.

The statement in (46) is uni-directional: focus is ideally realized in the maximally prominent position in the prosodic domain, but not vice-versa (it is not the case that the maximally prominent position in the domain is

always interpreted as focused). Languages differ with respect to the strategies that they employ in order to satisfy the requirement in (46) (Büring 2009): ‘boundary languages’ (e.g., Chichewa [NYA], Bengali [BEN]) add prosodic boundaries that modify the prosodic domain in which the prominence asymmetry applies, while ‘edge languages’ (e.g., Spanish, Hungarian) use deviations from the canonical order such that the focused constituent appears in the prominent part of the prosodic domain.

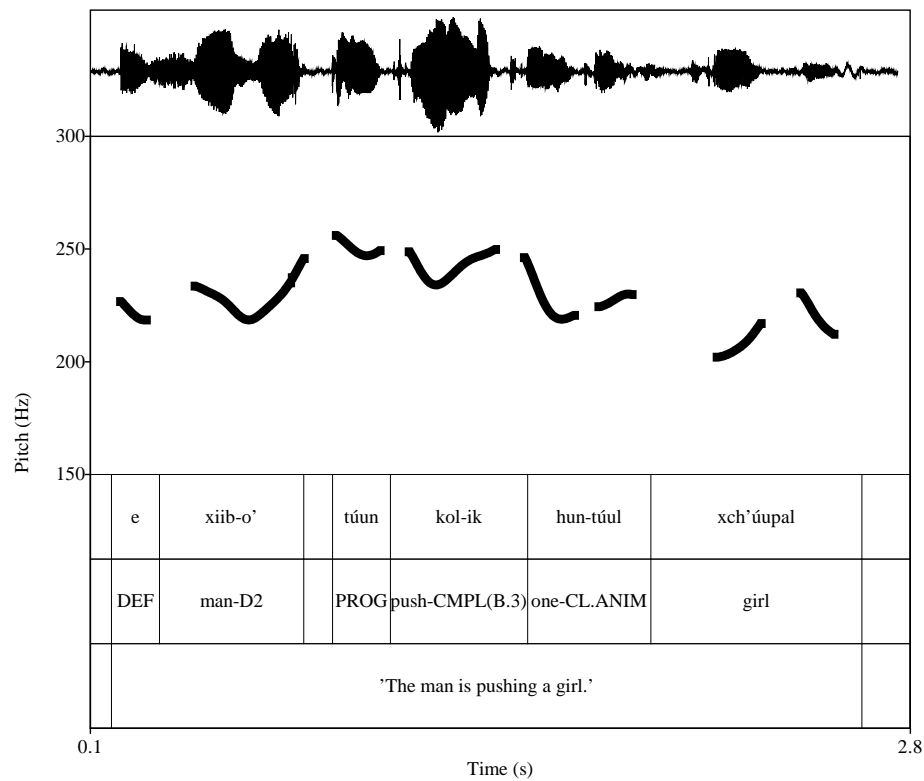
Two properties of Yucatec Maya establish the relevance of the focus-to-prominence correspondence in (46) for the data pattern in (45): (a) the relevant prosodic domain is the ι -domain (= intonational phrase) and (b) the leftmost constituent within the intonational phrase is maximally prominent (see evidence in Figure 4 below). The crucial property of the prosody-to-syntax mapping in Yucatec Maya is that the core clause including the pre-predicate constituent (if any) is mapped onto a single intonational phrase, while left-dislocated constituents are mapped onto an intonational phrase of their own. Hence, left-dislocated and pre-predicate constituents differ with respect to prosodic phrasing, as illustrated in (47) (Selkirk 2011). Assuming a left-dislocated constituent α , a pre-predicate constituent β , and a predicate phrase γ , the left-dislocated constituent forms its own intonational phrase, while the pre-predicate constituent is part of the same intonational phrase with the predicate.

(47) Syntax-prosody mapping in the left periphery of Yucatec Maya

$$\begin{array}{c}
 [\text{dislocated } \alpha \quad \quad \quad [\text{pre-predicateP } \beta \quad \quad [\text{predicateP } \gamma \quad \quad]]] \\
 (\quad \quad \quad)_i \quad (\quad \quad \quad)_i
 \end{array}$$

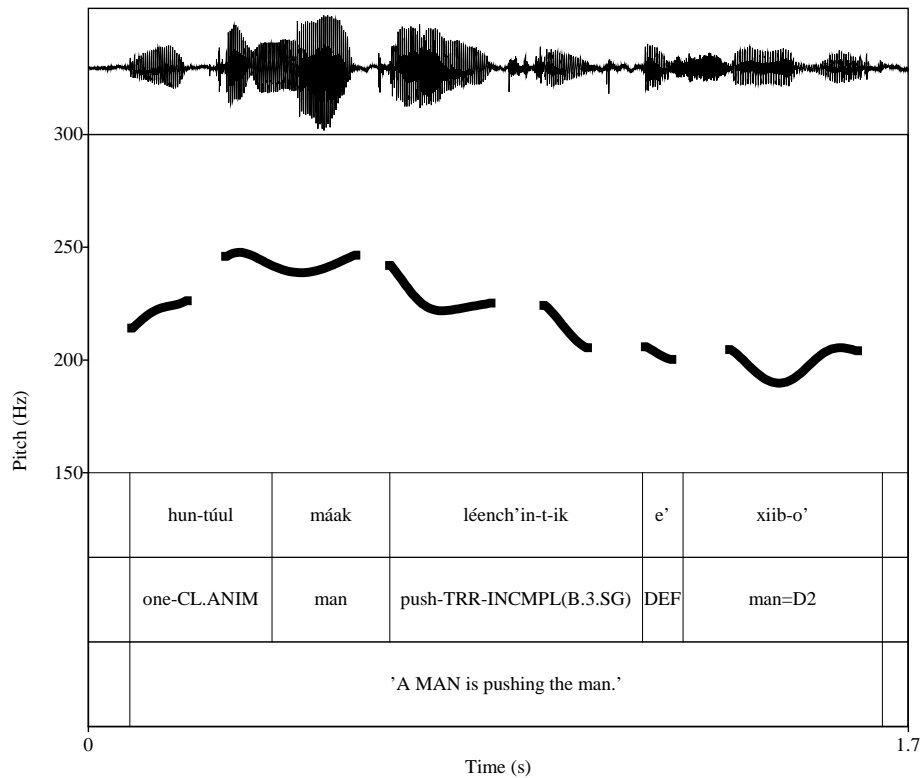
Phonetic evidence for intonational phrases in Yucatec Maya is provided by the following phenomena: (a) an intonational phrase determines the phonological domain within which the tonal events are downstepped; and (b) the right edge of an intonational phrase is associated with a tonal target (which is high for non-final intonational phrases) and is frequently accompanied by a prosodic break. These phenomena are exemplified through illustrative utterances collected in the experiment reported in Section 3.3. Figure 2 illustrates a construction involving a left-dislocated definite noun phrase (see pitch track in Figure 2). The left-dislocated constituent forms an intonational phrase on its own: the pitch contour targets a high tonal target reached at the right boundary of the phrase and is followed by a short prosodic break (95 msec). A further intonational phrase is mapped onto the predicate phrase: this prosodic entity is a downstep domain (see the scaling of the high lexical tones of *túun*, *-túul*, and *xch'úup*).

Figure 2. Left-dislocated XP: answer to the question ‘Who is the man pushing?’ (speaker SUE, feminine, b. 1987)



The answer in Figure 3 involves a pre-predicate constituent (the noun phrase *hun-túul máak*); the verb appears in the agent-focus form. The preverbal material is phrased differently from the left dislocated constituent in Figure 2. It is contained within the intonational phrase of the predicate and the entire utterance forms a single downstep domain (see the scaling of the high tones in *-túul*, *máak* and *léen*). The subject is not separated by a high tonal target or by a prosodic break from the predicate phrase.

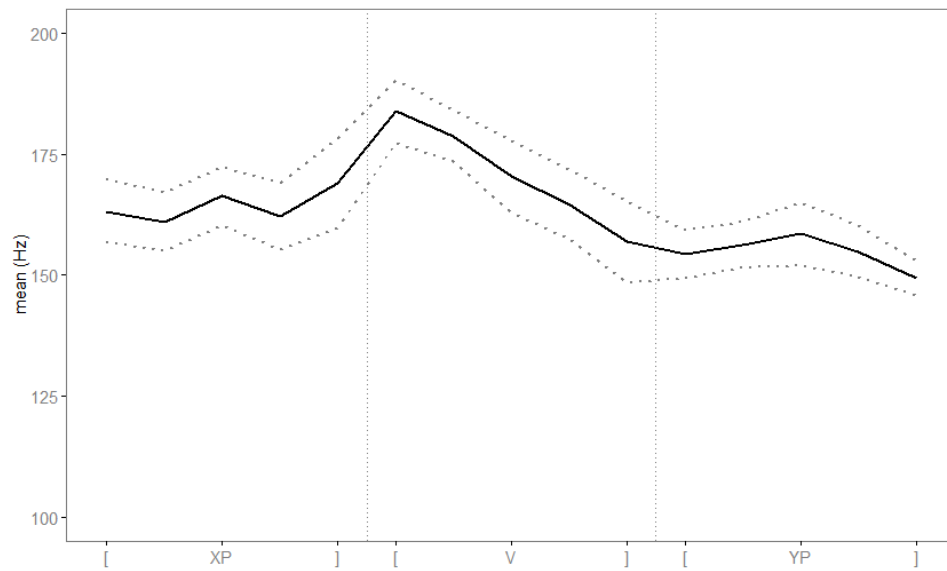
Figure 3. Pre-predicate XP: answer to the question ‘Who is pushing the man?’ (speaker ROX, feminine, b. 1987)



The prosodic difference between left-dislocated and pre-predicate constituents is visible in the F_0 means of utterances with a single constituent in the left periphery. Figure 4 presents the average of the F_0 means of five equal intervals of the noun phrase in the left periphery (XP), of the verb complex (V), and of the postverbal material (YP). A one-way analysis of variance was carried out on the difference in F_0 mean between the first interval of the preverbal noun phrase and the first interval of the verb. This analysis reveals that the pitch movement differs significantly for left-dislocated and pre-predicate constituents ($F_{1,17} = 28.3, p < .001$).

Figure 4. Prosodic realization of preverbal XPs: Time-normalized F_0 means and confidence intervals (95%)²⁰

(a) Left-dislocated XPs



(b) Pre-predicate XPs

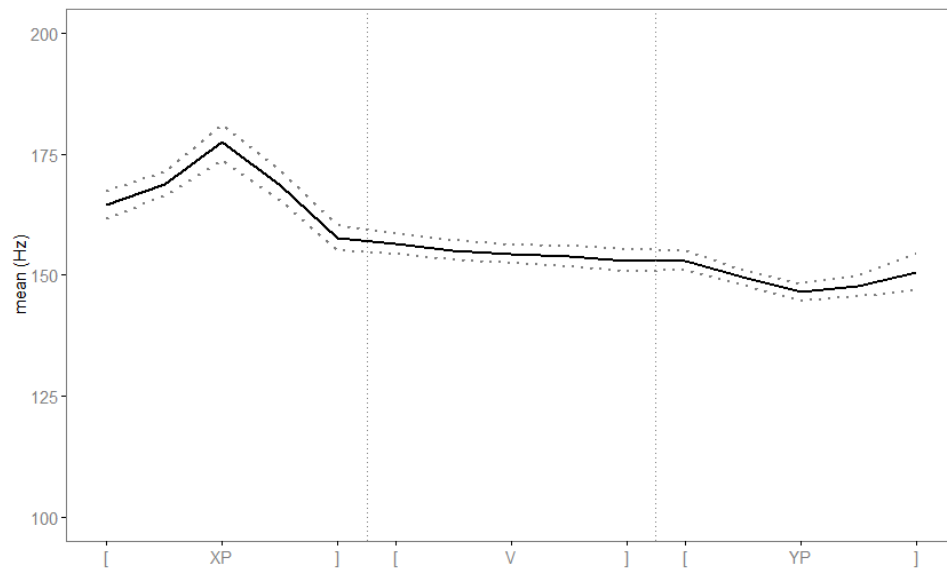


Figure 4 reveals a difference in intonational prominence between left-dislocated constituents and pre-predicate constituents. In both cases, the maximally prominent constituent is the leftmost part of the intonational phrase that contains the predicate. The difference lies in the syntactic entity that is mapped on this intonational phrase. Since left-dislocated XPs form an intonational phrase of their own, the maximally prominent area of the intonational phrase that contains the predicate is the beginning of the verb complex; tonal events after this point are gradually downstepped. Since pre-predicate XPs are part of the intonational phrase that contains the predicate, the maximally prominent area of the intonational contour is the preverbal constituent.

It is crucial that the relation between prominence and focus is not bi-unique (see (46)); that is, the maximally prominent part of an intonational phrase is not necessarily focused. The focused constituent in Figure 4 is the postverbal argument, but the maximally prominent part of the utterance is the initial part of the predicate phrase. Yucatec Maya is a tonal language and previous empirical studies on prosody show that focus does not correlate with phonologically-determined tonal events – such as pitch accents signaling focus – in this language; see details from phonetic studies in Kügler and Skopeteas (2006, 2007), Gussenhoven and Teeuw (2007).²¹ Hence, there are no phonetic effects of prominence that can be used independently of the syntactic construction in Yucatec Maya. Thus, we

conclude that the S_LVO utterances in object focus (Figure 4, left panel) are possible answers in an object-focus question but they do not contain any phonological indicator of object focus (since no evidence for such signals is found in the instrumental phonetic studies cited above). This implies that these utterances are ambiguous with respect to their focus domain: they allow for the possibilities of object focus and VP focus. With this background, we can now hypothesize which functional motivation drives a speaker's choice between focus *in situ* and focus fronting in this language: focus fronting is a syntactic strategy to realize the focus in the maximally prominent position of the utterance avoiding a configuration with ambiguous focus interpretations. In terms of the focus typology in Büring (2009), Yucatec Maya is an 'edge language' using a non-canonical constituent order in order to achieve the placement of the focus in the maximally prominent position.

The syntax-prosody mapping in (47) allows for a prosodic account of two phenomena that have been mentioned in previous sections. We discuss them only briefly here, since a complete account of the syntax-prosody mapping in Yucatec Maya is not the issue of this article. The set of deictic enclitics discussed in Section 3.1 have particular intonational properties: they are always associated with a high target in the pitch contour and are frequently associated with a prosodic break (Kügler and Skopeteas 2006: 87; Avelino 2009: 11; see also the prosodic realization of the clause-final enclitic in

Figure 3 that triggers a rise at the end of the utterance in which final lowering would be expected). These phonetic properties have a demarcative function: they determine boundaries of prosodic constituents. The observed facts in Section 3.1 are accounted for by assuming the generalization in (47): the enclitics may appear at the right edge of a left-dislocated constituent but not at the right edge of a pre-predicate constituent (see Aissen 1992 for a prosodic account along these lines concerning the enclitics in Tzotzil). That is, the deictic enclitics are associated with the boundary of an intonational phrase, which may be the right edge of a dislocated constituent but not the right edge of the pre-predicate constituent since the latter is integrated in the intonational phrase of the core clause. Similar phenomena can be observed in the right periphery: the deictic enclitics appear at the right edge of the prosodic constituent that is mapped onto the thematic layer of the clause (which is not necessarily the end of the sentence, since it may be followed by right-dislocated material).

A further phenomenon that is related to the syntax-prosody mapping in (47) is the placement of the interrogative particle *wáah* ‘Q’. This element (a) cannot appear at the beginning of the sentence and (b) cannot follow a left-dislocated constituent (Section 2.3). This particle is an enclitic, i.e., it attaches to a phonological host at its left. This explains why it cannot appear at the left edge of the intonational phrase. Eligible hosts are phonologically non-bound elements, which excludes elements such as the definite article *le*

‘DEF’, the preposition *ti* ‘LOC’ or the auxiliary preclitics, e.g., *t=u* ‘PFV=A.3’. The fact that *wáah* ‘Q’ does not attach to left-dislocated constituents implies that the question particle must be realized within a particular domain, namely the intonation phrase of the core clause; see (47). Such a behaviour has been observed for enclitics in other languages, too (see Kahnemuyipour and Megerdumian 2011 for Armenian [HYE]).

This section presented evidence that pre-predicate constituents – but not left-dislocated constituents – are part of the same intonational phrase with the clause; see (47). As shown in Figure 4, this asymmetry has implications for information structure since the nucleus of the prosodic structure (i.e., the most prominent constituent) has to be realized within the intonational phrase of the core clause. Our conclusion is that the information-structural difference between left-dislocated and pre-predicate constituents introduced in (45) can be deduced on the basis of the properties of the syntax-prosody mapping in Yucatec Maya. Left-dislocated constituents are adjoined material to the clause, which implies that they are extrametrical (Szendrői 2001: 46 on Hungarian). Pre-predicate constituents are the leftmost part of the core clause and host the maximally prominent material in the pitch contour of the utterance.

5 Conclusions

The aim of this article was to examine the interplay between syntax, prosody and information structure in a particular type of focus construction

in Yucatec Maya. Based on syntactic evidence we have shown that the construction at issue involves fronting of the focused constituent into a pre-predicate position, which is left-adjacent to the thematic layer of the clause. Our claims about the discourse properties of this construction are summarized in (48). We distinguished between two left peripheral configurations on the basis of syntactic evidence and we observed that this distinction is crucial for prosodic phrasing: Left-dislocated constituents are phrased separately from the clause and do not host the intonational nucleus. Pre-predicate constituents are contained by the intonational phrase of the core clause and exactly occupy the maximally prominent part of this prosodic domain.

(48) a. **CLAUSE STRUCTURE**

Yucatec Maya is a VOS language with two left-peripheral options: fronting to the pre-predicate position and left dislocation outside the core clause.

b. **PROSODY-SYNTAX MAPPING**

The pre-predicate position is part of the intonational phrase containing the predicate; left dislocated material is mapped onto a separate intonational phrase.

c. **PROMINENCE ASYMMETRY**

The left edge of prosodic constituents is maximally prominent.

d. **FOCUS PROMINENCE**

Narrow focus targets the maximally prominent position in intonational phrase containing the predicate.

The statement in (48d) is uni-directional: ‘narrow focus targets the maximally prominent position’ and not vice-versa. The fronting operation is motivated (a) if a narrow focus is available (which is not always the case) and (b) if it is not already in the left edge of the core clause (which is the case for verbs). The fronting operation is subject to independent syntactic limitations determining which types of constituent can be hosted by the pre-predicate position (see discussion about co-extensivity in 3.1). As a result, focus is not necessarily co-extensive with the fronted constituent.

It is crucial for this account that we did not encounter evidence that the fronting correlates with a specific subtype of focus. Identificational properties may occasionally arise depending on context, which implies that this property is not a necessary concomitant of this syntactic construction. This view is strengthened by the findings of a production experiment as well as by the fact that no contradiction effects arise through the use of *also*- and *even*-phrases in the pre-predicate position. This finding is crucial because it implies that we do not need to assume a propositional operator triggering fronting to the pre-predicate position.

Speech production data show an interesting contrast between subject and object foci. While subjects in question are always in the pre-predicate position, objects in question are frequently in situ. This asymmetry is the

result of differences in the paradigmatic alternatives for the expression of focus on subjects and focus on objects. We know from phonetic studies that postverbal material cannot be (prosodically) marked for focus in this language, i.e., it is not possible to pronounce postverbal arguments in a way suggesting that it is in focus.

The rationale of our account is based on the idea that information-structural effects on word order are the result of properties of linearization and prosodic structure. Syntactic configurations are mapped onto prosodic domains with prominence asymmetries, which cannot be deliberately modified (especially in a tonal language without free pitch accent placement). The product of this mapping is a set of information-structural possibilities that correlate with the available syntactic options.

Acknowledgments

We are grateful to Gisbert Fanselow, Caroline Féry, Frank Kügler, and Malte Zimmermann for comments on presentations of this study. We received detailed comments by Judith Aissen, Jürgen Bohnemeyer, the Editor and Associate Editor of this journal, which contributed substantially to the final account that we present in this article. We are grateful to Joseph P. DeVeugh-Geiss for editing the final draft. The main part of the data collection was supported by the SFB *Information Structure* at the University of Potsdam/Humboldt University Berlin (sponsored by the German Research Foundation).

REFERENCES

- Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68: 43-80.
- , 1999. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14: 447-491.
- Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1: 149-168.
- AnderBois, Scott. 2012. Focus and un informativity in Yucatec Maya questions. *Natural Language Semantics* 20.4: 349-390.
- Avelino, Heriberto. 2009. Intonational patterns of topic and focus constructions in Yucatec Maya. *MIT Working Papers in Linguistics, Special Issue on New Perspectives in Mayan Linguistics*. eds. Heriberto Avelino, Jessica Coon, and Elisabeth Norcliffe. 1-22.
- Bates, Douglas, Martin Maechler, and Ben Bolker. 2012. lme4: Linear mixed-effects models using S4 classes. Version 0.999999-0 (2012-06-23).
- Bohnenmeyer, Jürgen. 1998. *Time Relations in Discourse: Evidence from a Comparative Approach to Yucatec Maya*. PhD Dissertation, Tilburg University.
- , 2009. Linking without grammatical relations in Yucatec: alignment, extraction, and control. *Form and Function in Language Research*,

- eds. Johannes Helmbrecht, Yoko Nishina, Yong-Min Shin, Stavros Skopeteas, and Elisabeth Verhoeven, pp. 185-214. Berlin: Mouton de Gruyter.
- Bricker, Victoria Reifler. 1979. Wh-questions, relativization, and clefting in Yucatec Maya. *Papers in Mayan linguistics*, vol. 3, ed. Laura Martin, pp. 107-136. Columbia, MO: Lucas Brothers.
- Büring, Daniel. 2009. Towards a typology of focus realization. *Information Structure: Theoretical, typological, and experimental perspectives*, eds. Malte Zimmerman and Caroline Féry, pp. 177-205. Oxford: Oxford University Press.
- , and Rodrigo Gutiérrez Bravo. 2001. Focus-related constituent order variation without the NSR: A prosody-based crosslinguistic analysis. *Syntax and semantics at Santa Cruz 3*, ed. James McCloskey, pp. 41-58.
- Cheng, Lisa Lai-Shen, and Laura J. Downing. 2012. Against FocusP: arguments from Zulu. *Contrasts and Positions in Information Structure*, eds. Ivona Kučerová and Ad Neeleman, pp. 247-266. Cambridge: Cambridge University Press.
- Delin, Judy. 1992. Properties of it-cleft presupposition. *Journal of Semantics* 9: 289-306.

- Dik, Simon C. et al. 1981. On the typology of focus phenomena. *Perspectives on Functional Grammar*, ed. Teun Hoekstra, pp. 41-74. Dordrecht: Foris.
- Dik, Simon. 1997. *The Theory of Functional Grammar, Part 1: The Structure of the Clause*, ed. Kees Hengeveld. Berlin, New York: Mouton De Gruyter.
- Drubig, Bernhard H. 2003. Towards a typology of focus and focus constructions. *Linguistics* 41.1: 1-50.
- Durbin, Marshall and Fernando Ojeda. 1978. Basic word-order in Yucatec Maya. *Papers in Mayan Linguistics*, vol. 2, ed. Nora C. England, pp. 69-77. Columbia: University of Missouri.
- É. Kiss, Katalin. 1998. Identificational vs. information focus. *Language* 74.2: 245-273.
- , 2009. Structural focus and exhaustivity. *Information Structure from different perspectives*, eds. Malte Zimmerman and Caroline Féry, pp. 64-88. Oxford: Oxford University Press.
- Fanselow, Gisbert, and Denisa Lenertová. 2011. Left peripheral Focus: Mismatches between syntax and information structure. *Natural language and linguistic theory* 29.1: 169-209.
- Féry, Caroline and Vieri Samek-Lodovici. 2006. Focus projection and prosodic prominence in nested foci. *Language* 82.1: 131-150.

- Fiedler, Ines, and Anne Schwarz. 2005. Out-of-focus encoding in Gur and Kwa. *ISIS (Interdisciplinary Studies on Information Structure)* 3: 111-142.
- Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz, and Malte Zimmermann. 2010. Subject focus in West African languages. *Information Structure from different perspectives*, eds. Malte Zimmerman and Caroline Féry, pp. 234-257. Oxford: Oxford University Press.
- Gussenhoven, Carlos and Renske Teeuw. 2007. A moraic and a syllabic H-tone in Yucatec Maya. In *Fonología instrumental: patrones fónicos y variación lingüística*, eds. E. H. Z. and P. M. Burtogeno, pp. 47-71. Colegio de México, Mexico City, 21-27 October 2006.
- Gutiérrez Bravo, Rodrigo. 2007. Foco de agente y cláusulas relativas en maya yucateco. Ms.
- , 2011. External and internal topics in Yucatec Maya. *Representing Language: Linguistics essays in honour of Judith Aissen*, eds. Rodrigo Gutiérrez Bravo, Line Mikkelsen, and Eric. Potsdam, pp. 105-119. Santa Cruz: University of California.
- Gutiérrez Bravo, Rodrigo, and Jorge Monforte y Madera. 2007. La alternancia sujeto inicial/verbo inicial y la teoría de la optinidad. Ms.

- , 2008. On the nature of unmarked word order in Yucatec Maya. Handout presented in the Symposium on Mayan Languages, Workshop at the 2008 meeting of the SSILA, Chicago, 3-4 January 2008.
- , 2011. Focus, agent focus, and relative clauses in Yucatec Maya. *New Perspectives in Mayan Linguistics*, ed. Heriberto Avelino, pp. 257-274. Newcastle: Cambridge Scholars Publishing.
- Hartmann, Jutta. 2012. The focus structure of it-clefts. Ms., University of Tübingen.
- Hartmann, Katharina, and Malte Zimmermann. 2006. In place – out of place? Focus strategies in Hausa. *On Information Structure, Meaning and Form: Generalizations across Languages*, eds. Kerstin Schwabe and Susanne Winkler, pp. 365-403. Amsterdam: Benjamins.
- Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76.4: 891-920.
- Kahnemuyipour, Arsalan and Karine Megerdumian. 2011. Second-position clitics in the vP phase: The case of the Armenian auxiliary. *Linguistic Inquiry* 42.1: 152-162.
- Koch, Karsten. 2008. Intonation and focus in Nl̥əʔkepmxcin (Thompson River Salish). Ph.D. dissertation, Vancouver: The University of British Columbia.

- Kügler, Frank, and Stavros Skopeteas. 2006. Interaction of lexical tone and information structure in Yucatec Maya. *Proceedings of the 2nd International Symposium on Tonal Aspects of Languages*, La Rochelle, pp. 83-88.
- , 2007. On the universality of prosodic reflexes of contrast: the case of Yucatec Maya. *ICPhS XVI*, 1025–1028.
- Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions, *Linguistics* 39.3: 463-516.
- Launey, Michel. 1994. *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique*. CNRS éditions. Paris.
- Lehmann, Christian. 1990. Yukatekisch. *Zeitschrift für Sprachwissenschaft* 9.1-2: 28-51.
- , 2003. *Possession in Yucatec Maya* (second, revised edition). ASSIDUE (Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt), Nr. 10. Erfurt: Seminar für Sprachwissenschaft der Universität.
- Potsdam, Eric. 2006. More concealed pseudoclefts in Malagasy and the clausal typing hypothesis. *Lingua* 116: 2154-2182.
- Prince, Ellen F. 1978. A comparison of wh- clefts and it-clefts in discourse. *Language* 54.4: 883-906.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of Grammar*, ed. Liliane Haegeman, pp. 281-337. Dordrecht: Kluwer.

- Selkirk, Elisabeth. 2011. The syntax-prosody interface. *The Handbook of Phonological Theory*, 2nd edition, eds. John Goldsmith, Jason Riggle and Alan Yu, pp. 435-484. Oxford: Blackwell Publishing.
- Skopeteas, Stavros, Ines Fiedler, Sam Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry, and Manfred Krifka. 2006. Questionnaire on Information Structure (QUIS). *Interdisciplinary Studies on Information Structure (ISIS)*, Vol. 4. Potsdam: Universitätsverlag Potsdam.
- Skopeteas, Stavros, and Elisabeth Verhoeven. 2005. Postverbal argument order in Yucatec Maya. *Sprachtypologie und Universalienforschung* 53.1: 71-79.
- , 2009a. Distinctness effects on VOS order: evidence from Yucatec Maya. *MIT Working Papers in Linguistics*, Special Issue on Mayan Languages, eds. Heriberto Avelino, Jessica Coon, and Elisabeth Norcliffe, pp. 135-152.
- , 2009b. The interaction between topicalization and structural constraints. *Linguistic Review* 26: 239-259.
- Skopeteas, Stavros and Gisbert Fanselow. 2010. Focus types and argument asymmetries: a cross-linguistic study in language production. *Contrastive information structure analysis*, ed. Carsten Breul, pp. 169-197. Amsterdam: Benjamins.

- , 2011. Focus and the exclusion of alternatives: On the interaction of syntactic structure with pragmatic inference. *Lingua* 121: 1693-1706.
- Stiebels, Barbara. 2006. Agent focus in Mayan languages. *Natural language and linguistic theory* 24: 501-570.
- Szendrői, Kriszta. 2001. Focus and Syntax Phonology Interface. Ph.D. dissertation, UCL.
- Tonhauser, Judith. 2003. F-constructions in Yucatec Maya. *The Proceedings of SULA 2*, eds. Jan Anderssen, Paula Menéndez-Benito, and Adam Werle, pp. 203-223. Vancouver, BC: GLSA, Amherst: University of Massachusetts.
- , 2007. Agent focus and voice in Yucatec Maya. *Proceedings of the 39th Meeting of the Chicago Linguistic Society*.
- Trechsel, Frank R. 1993. Quiché focus constructions. *Lingua* 91: 33-78.
- Truckenbrodt, Hubert. 1995. Phonological Phrases: Their Relation to Syntax, Focus, and Prominence. Ph.D. thesis, MIT.
- Vapnarsky, Valentina. 2013. Is Yucatec Maya an omnipredicative language? Predication, the copula, and focus constructions (to appear in *Language Typology and Universals*).
- Verhoeven, Elisabeth. 2007. Experiential constructions in Yucatec Maya. Amsterdam: Benjamins.

- Wedgwood, Daniel. 2006. Identifying inferences in focus. On Information Structure, Meaning and Form: Generalizations across Languages, ed. Kerstin Schwabe, and Susanne Winkler, pp. 207-227. Amsterdam: Benjamins.
- , 2009. Variation in focus. Focus at the Syntax-Semantics Interface, ed. Arndt Riester, and Edgar Onea. Working Papers of the SFB 732, Vol. 3.
- Zerbian, Sabine. 2006. The subject/object asymmetry in Northern Sotho. On Information Structure, Meaning and Form: Generalizations across Languages, eds. Kerstin Schwabe and Susanne Winkler, pp. 323-345. Amsterdam: Benjamins.
- Zimmermann, Malte. 2008. Contrastive focus and emphasis. *Acta Linguistica Hungarica* 55.3-4: 347-360.

NOTES

- ¹ *Orthographical conventions.* We adopt a slightly modified version of the local orthography developed in the project *Typology of Yucatec Maya* (Christian Lehmann, University of Erfurt). The following letters do not correspond to the IPA conventions: $x = /ʃ/$, $ch = /tʃ/$, $y = /j/$. The letter h refers to a voiceless glottal fricative (the local orthography uses j

instead). The digraphs *aa*, *ee*, *oo*, *uu*, *ii* stand for long vowels. There is a contrast between high and low tones in long vowels, while short vowels do not bear tone. Tones are indicated on the first letter of long vowels: *áa* stands for a high tone, *àa* for a low tone.

- ² The definiteness marker triggers an enclitic; see (10a) and (15). Apart from definite markers, these enclitics demarcate the right edge of non-final intonational phrases, such as the topicalized constituent in (2).
- ³ The ‘agent-focus construction’ appears in many Mayan languages with differences in form and function; see Stiebels (2006).
- ⁴ In clause-final position, the verb form is accompanied by the suffix *-eh* ‘SUBJ’.
- ⁵ We owe these examples to the collaboration with Gisbert Fanselow and Caroline Féry within the framework of the project *Discontinuous noun and prepositional phrases* at the University of Potsdam.
- ⁶ The enclitic *=e* ‘D3’ in this example demarcates the right boundary of a non-final clause and is not triggered by the pronoun *leti* ‘that one’.
- ⁷ Depending on framework, the relative clause is either analyzed as the subject or as an attribute to a pronominal subject.
- ⁸ The crucial issue is that a third person cross-reference marker in the headless relative clause is possible. Further options can arise through

connectivity effects; see, for example, *it's you who are responsible* (J. Aissen p.c.). Such effects also occur in Yucatec Maya: native speakers accept a version of (12c) with a second person cross-reference affix.

⁹ An example with the configuration in (13b) is reported as grammatical in Tonhauser 2003 (33b). Our informants reject this example.

¹⁰ In sentence-initial position, we find the subordinative conjunction *wáah* ‘if’. The conjunction *wáah* is homonym with the interrogative particle.

¹¹ Furthermore, see *Tuméen hun-túul=wáah xch'úupal túun léench'in-t-a'l le chan mèsáah=o'?* (with focus on the numeral), also excluding an analysis of the material at the left side of the particle as a predicate.

¹² ‘Broad focus’ refers to a focus domain containing a higher constituent (the VP or the entire clause). ‘Narrow focus’ is a focus domain restricted to any constituent below these layers (e.g., focus on V, NP, N, PP, P, etc.).

¹³ It would be grammatical to left-dislocate the noun and to front the quantifier into the pre-predicate position (restricted to contexts in which the noun serves as contrastive topic).

¹⁴ The pronoun *leti* ‘3.SG’ does not require an enclitic; compare with (25).

¹⁵ See also Hedberg (2000) for a similar view on English *it*-clefts.

- ¹⁶ Speaker intuitions have been elicited with Ernesto May Balam (2006), Ramón May Cupul (2006), and Amedee Colli Colli (2008, 2012).
- ¹⁷ See Skopeteas and Fanselow (2011) for an experimental study on the exhaustive interpretation of similar constructions in several languages.
- ¹⁸ Compare: *It was ?also John/*even John that Mary invited to her birthday party* and similar examples in Hungarian in É. Kiss (1998: 252f.).
- ¹⁹ The noun *máak* ‘person’ is not specified for sex; the speaker uses this noun in order to refer to a ‘man’ in the stimulus.
- ²⁰ The F_0 averages of left-dislocated constituents relate to 12 (out of 13) utterances elicited in contexts inducing focus on the object (all S_LVO; see Table 1). The F_0 averages of pre-predicate constituents relate to 68 (out of 77) utterances (23 S_PVO and 19 S_PV tokens in subject focus; 9 O_PVS and 17 O_PV tokens in object focus). The remaining tokens were excluded from pitch analyses due to disfluencies.
- ²¹ Weak phonetic reflexes of focus are reported in Gussenhoven and Teeuw (2007) (effect of corrective focus on peak alignment) and in Kügler and Skopeteas (2007) (lowering effect of contrast on the scaling of L-tones).