

SYNTAX-PHONOLOGY INTERFACE AND CLITIC PLACEMENT IN MAYAN LANGUAGES

STAVROS SKOPETEAS
UNIVERSITY OF POSDAM

1. Introduction

The study of the syntax-phonology interface is based on the idea that prosodic domains determine segments of the linearization with reference to entities in the hierarchical constituent structure (see Inkelas and Zec 1995, Selkirk 1984, Truckenbrodt 1999, 2007). The observation of phonological domains in the linearization is mainly based on prosodic entities, such as boundary tones and prosodic breaks that occur at the edges of prosodic units (see for instance Féry 2001 on French), the domain of application of phrasal stress rules (see an overview in Zubizarreta and Vergnaud 2005), the height of pitch accents (see Féry and Ishihara 2008) or language-specific constraints on the realization of lexical tones (see discussion on Xiamen Chinese, Tohono O'odham and Chichewa in Truckenbrodt 1999). Mayan languages are particularly challenging for the study of the syntax-phonology interface, because they display a class of enclitics that attach to the edges of intonational units; hence, these languages provide morphological elements that spell out the boundaries of prosodic domains. The challenge is to figure out the prosodic and/or syntactic entities that account for the occurrence of these elements.

Two different phenomena, both entailing an interaction of clitics with prosodic units, are distinguished in this article. In some Mayan languages, particular enclitics occur at the right edge of prosodic units (see Aissen 1992 on Tzotzil and Jacaltec), as illustrated by means of the enclitic *=un* in (1) from Tzotzil Maya. This element cliticizes to a phonological host on its left and does not have any contribution to the propositional content of the clause (see Aissen 1992: 53).

- (1) Tzotzil
Yul s-vatulan taj x-chi'il=un; 'i-bat-ik=un.
 return A.3-shook DET A.3-friend=CL CMPL-go-B.3.PL=CL
 'He returned to shake his friend; they went.' (Aissen 1992: 53)

Furthermore, in some Mayan languages particular interpretable features are spelled out through right edge clitics. For instance, DPs involving the definite article *le* 'DEF' in Yucatec Maya are obligatorily accompanied by an enclitic. This enclitic is a member of a set of elements encoding several deictic concepts. In examples (2a) and (2c), the enclitic =*o*' encodes the distal region of the deictic center/speaker. Example (2b) shows that indefinite DPs and proper nouns are not accompanied by an enclitic (the latter sharing with definite DPs the property of specific reference without being accompanied by a determiner).

- (2) Yucatec
- (a) *k-u xímbat-ik le h-mèen*(=o').*
 IPFV-A.3 visit-INCMPLE DEF M-shaman=D2
 'The shaman (there) comes.'¹
- (b) *k-u táal Pèedróoh / hun-túul h-k'ìin.*
 IPFV-A.3 come Pedro one-CL.AN M-priest
 'Peter comes / a priest comes.'
- (c) *k-u xímbat-ik le h-mèen hun-túul*
 IPFV-A.3 visit-INCMPLE DEF M-shaman one-CL.AN
h-k'ìin=o'.
 M-priest=D2
 'A priest visits the shaman.' (Skopeteas and Verhoeven 2005: 364f.)

In contrast to the enclitic =*un* in Tzotzil, Yucatec Mayan enclitics spell out interpretable features, namely they distinguish among different deictic regions (see details in Section 2). However, their placement in the linearization interacts with the prosodic domains of the utterance: see example (2c), in which the enclitic =*o*' does not attach to the definite DP that licenses it but to the right edge of the intonational phrase.

The aim of this article is to outline the properties of these clitics in different Mayan languages. The languages of this family are genetically classified under five branches: (a) Cholan-Tzeltalan, (b) Yucatecan, (c) Kanjobalan-Chujean, (d) Quichean-Mamean, and (e) Huastecan. The phenomenon at issue appears in languages under the branches (a)-(c).² The properties of the enclitics are not uniform in these languages. Apart from

some languages that do not exhibit this phenomenon at all, the languages that do have enclitics differ with respect to the conditions that license enclitics and to the interaction of enclitics with prosodic units.

The article is structured as follows: Section 2 outlines the properties of enclitics that bear semantic content, such as the enclitics in Yucatec Maya illustrated in (2a-c). Section 3 is devoted to the enclitics that do not have any contribution to the propositional content (such as the Tzotzil enclitic introduced in (1)), but bear a demarcative function determining prosodic domains of the linearization. Section 4 shows that the placement of interpretable clitics in the linearization is determined by the same rules that license the occurrence of demarcative clitics. Section 5 shows some restrictions in the occurrence of multiple adjacent clitics. The conclusions are presented in Section 6.

The data presented in this article is based on my fieldwork on Yucatec Maya (Quintana Roo, Mexico: December 2004 and March 2008), a series of text collections in several Mayan languages, i.e., the text collection of Yucatec Maya at the University of Erfurt (Yuclex database, Christian Lehmann), the collection of Mopán Maya texts edited by Shaw (1971, ed.), the collections of Itzá Maya texts edited by Schumann (1971) and Hofling (1991), and the illustrative texts of Petalcingo Tzeltal provided by Shklovsky (2005), as well as cited examples from articles or grammars of several Mayan languages.

2. Clitics for interpretable features

The facts presented in this Section show that certain interpretable features in the numeration (definiteness, 1st person singular, and deictic categories) are spelled out through right edge clitics in Mayan languages. As it will be shown below, the exact range of elements to which this rule applies is specific to each language. The cross-linguistic variation is lexical in nature and is not relevant for our purposes. The crucial issue is that this subclass of enclitics spells out an interpretable feature F in the numeration, as specified by the C -marking statement in (3).

- (3) C -marking statement I
An interpretable feature F_i is spelled out through a clitic C_j .

Yucatec Mayan enclitics are obligatorily triggered by the definite marker, as already shown in (2a-b), and are optional with further DPs with specific reference (see Lehmann 2003: 95f.). They display a threefold contrast that specifies the relation of the definite DP to the situational and

contextual common ground (see Blair and Vermont-Salas 1965: 31f., Bohmeyer 1998b: 205f., Lehmann 1990: 39, 2003: 25, see also Hanks 1990 for a comprehensive work on the deictic properties of these elements). The enclitic =*a*' 'D1' indicates that the referent is located in the proximity region of the deictic center (i.e., the speaker); the enclitic =*o*' 'D2' indicates that the referent is located away from the speaker, but within the discourse situation. The enclitic =*e*' 'D3' is the unmarked member of the contrast. Native speakers report the intuition that definite referents marked with =*e*' 'D3' are given but outside the visible field of the speaker. Discourse-giveness is encoded through the definite marker *le*, and the interpretation that the referent is not visible can be accounted for as the result of a scalar implicature based on the fact that the speaker does not select a specific option, either =*a*' 'D1' or =*o*' 'D2'.

(4) Yucatec

Táan u wen-el le xibpal(=a'/=o'/=e').*
 PROG A.3 sleep-INCMPLE DEF man:child=D1/=D2/=D3
 'The boy (here/ there/ afore mentioned) is sleeping.'

In some further Mayan languages, definite DPs are accompanied by enclitics, but lacking the contrast of deictic regions, which is only attested in Yucatec Maya. This pattern is exemplified by the enclitic =*eh* 'CL' in Itzá Maya (see Schumann 1971: 39, Hofling 1991: 14) which also belongs to the Yucatecan branch, see (5a).³ A further example is the enclitic =*e* 'CL' in Tzeltal, which – at least in the majority of its occurrences – is licensed by a set of determiners that encode definiteness/specificity (see Shklovsky 2005: 71), see (5b). Finally, example (5c) shows that Tzotzil Maya displays an enclitic =*e* 'CL.DEF' that is licensed by a set of determiners, some complementizers, and some deictic adverbs (see Aissen 1987: 8, 1992: 61) and is different from the enclitic =*un* 'CL' in (1) (which cliticizes to hosts of several categories).

(5) (a) Itzá

He'l-oh a' pek'=eh.
 DEM-DIST DEF dog=CL
 'There is the dog.' (Hofling 1991: 15)

- (b) Petalcingo Tzeltal⁴
te chij=e laj s-kuch bajel te
 DEF deer=CL PFV A.3-carry(B.3) DIR:away DET
wits' kerem=e.
 small boy=CL
 'The deer carried away the boy.' (Shklovsky 2005: 148)
- (c) Tzotzil
He'-em li na=e.
 close-PF DEF house=CL.DEF
 'The house is closed.' (Aissen 1987: 94)

Definite DPs in Mopán Maya, which is a further member of the Yucatecan branch, are also right enclosed by an enclitic, however this enclitic does not have a fixed phonological form. Apart from some cases of adverbs with a lexicalized enclitic (e.g., *wa'ye* 'here', < *wa'y=e*'), productive enclitics copy the vowel of the last syllable of their phonological host, as shown in (6).

- (6) Mopán (examples found in Shaw 1971, ed.: 411-430)
- | | | | |
|-----|----------------------|-----|------------------------|
| (a) | <i>a nooch'up=u</i> | (d) | <i>a chuyubac'=a</i> |
| | DEF old.woman=CL.DEF | | DEF carrot.vine=CL.DEF |
| (b) | <i>aj t'iiw=i</i> | (e) | <i>a che'=e</i> |
| | DEF eagle=CL.DEF | | DEF tree=CL.DEF |
| (c) | <i>aj coj=o</i> | (f) | <i>a nüc'=ü</i> |
| | DEF cougar=CL.DEF | | DEF stomach=CL.DEF |

Jacaltec Maya displays a different licensing condition for right edge clitics, as discussed in detail in Grinevald Craig (1977: 276-286) and Aissen (1992: 61-68, 200). The enclitic =*an* 'CL.1' in the examples under (7) is licensed by the presence of a first person in the clause, as illustrated in (7a) for a first person agent and in (7b) for a first person possessor (see further examples in Grinevald Craig 1977: 276f.).

- (7) Jacaltec
- (a) *Ch-in to=an.*
 INCMPL-B.1 go=CL.1
 'I go.' (Grinevald Craig 1977: 276)
- (b) *X-cam hin-cheh=an.*
 CMPL-(B.1)die A.1-horse=CL.1
 'My horse died.' (Grinevald Craig 1977: 277)

In sum, this section has shown that a subset of Mayan languages has a class of enclitics that is licensed by a language-specific set of triggers (determiners, complementizers, and deictic elements). Yucatec differs from the other languages in that it displays a threefold contrast of deictic regions {proximal, distal, not available in the discourse situation}. Jacaltec is special in that the enclitics at issue are triggered by a pronominal element (first person pronoun).

3. Demarcative clitics

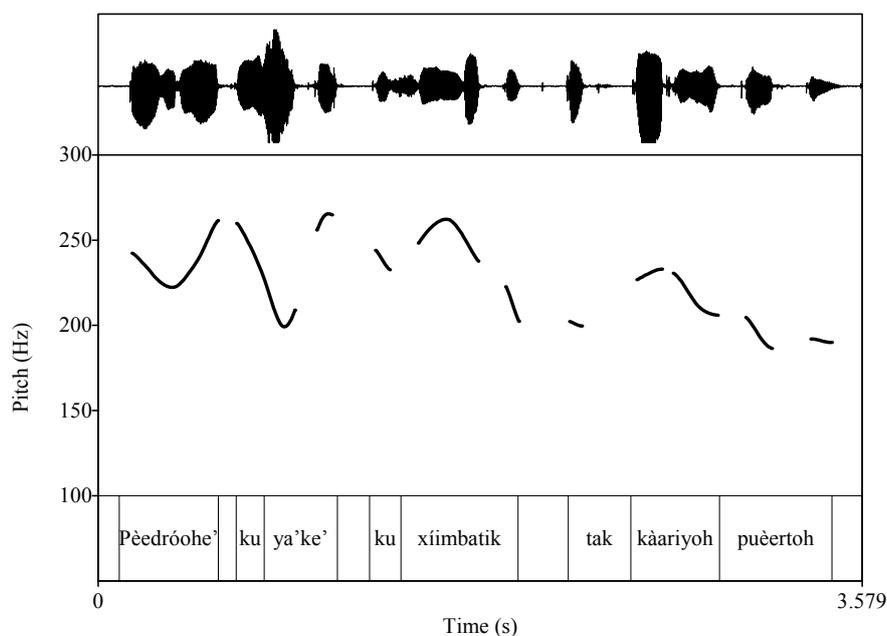
Another subset of enclitics in Mayan languages does not have any contribution to the propositional content but bears a demarcative function in separating particular prosodic domains. They occur in diverse environments sharing in common that they are exactly the environments in which we expect an intonation-phrase edge to occur. There is no systematic account about their exact prosodic properties, but all descriptions report that these elements are associated with a boundary tone. Aissen (1992: 57) reports that the Tzotzil enclitic *=un* ‘CL’ coincides with “phrase-final contour” and is followed by a significant pause and she mentions that the Jacaltec suffix shows similar properties (see Aissen 1992: 61 and Day 1973: 57). By means of illustration, we present the prosodic properties of the enclitic *=e*’ in Yucatec Maya, see example (8) and pitch track in Fig. 1. The enclitic *=e*’ occurs twice in this example, once following the topic constituent and once before the complement clause.

- (8) Yucatec
Pèedróoh=e’ k-u ya’k=e’ k-u xímbat-ik
 Pedro=D3 IPFV-A.3 say:CMPL-D3 IPFV-A.3 walk-INCMPL
tak kàariyo’ puèerto’.
 up.to Carrillo Puerto
 ‘Pedro says that he is walking up to Carrillo Puerto.’

The crucial property is that enclitics are associated with a high F_0 target, as may be observed at the right edge of the topic constituent and of the matrix verb in Fig. 1 and they are often accompanied by a prosodic break. A further property is that they determine prosodic domains within which the tonal events are progressively downstepped. This property may be observed in Fig. 1: the two first intonational phrases, enclosed by the enclitic *=e*’ at their right edge, are aligned with boundary tones that reach the same pitch level. The larger prosodic domain following the second

clitic contains two prosodic phrases, a first one which is aligned with the edges of the verb complex and a second one which is aligned with the edges of the PP constituent. The highest pitch level is reached by the first prominent (i.e., tone bearing) syllable within the phrase, which is the case for the syllable *xíim* of the first phrase, and for the syllable *kàa* of the second phrase. Crucially, since both prosodic phrases are parts of the same intonational phrase, the pitch scaling of the prominent syllables shows the effects of downstep.

Fig. 1. F₀ excursion of example (8)



Assuming that these enclitics are aligned with the right edge of prosodic domains and adopting the view that prosodic domains refer to syntactic projections, the occurrence of enclitics in particular points in the linearization is expected to be informative for the rules of syntax/phonology interface in Mayan. The distributional properties of the corresponding enclitics in Tzotzil are studied in detail in Aissen (1992: 57). The enclitic =*un* 'CL' (see (1)) may occur: (a) sentence finally but not sentence initially, (b) sentence internally before adverbial clauses and CP complements but not before relative clauses and IP/VP complements, and

(c) following left dislocated material, either sentence-initial adverbs or topics, but not following the focus.

This distribution is straightforwardly accounted for if we assume a phase-based account on the generation of linearization statements (Chomsky 2001, 2005, 2007). In terms of this framework, syntactic entities are incrementally transferred to the interfaces in stages containing fragments of clausal structure that are complete, i.e., not subject to further valuation processes during the next stages of the derivation. The critical points for the phenomenon we are observing in this article are: (a) that CP and vP but not TP/IP is a phase and (b) that the leftmost constituent of the phase, i.e., the specifier of the phase's head is transferred in a later step, such that it provides an "escape hatch" for movement operations.

The linearization statements that are incrementally transferred to the phonological component form separate prosodic domains (see Kratzer and Selkirk 2007, Ishihara 2007). The critical layer for our purposes is the CP-Phase: we assume that linearization statements at this layer are mapped onto the highest order prosodic entities, namely intonational phrases (IntP), see (9a). This implies that the head of the phase (i.e., spec-CP for the CP phase) and the spellout domain of the phase, i.e., the complement of the phase's head (TP/IP for the CP phase) are mapped on separate IntP's at the PF component.⁵ In order to stipulate the licensing of demarcative clitics, a rule of morphological marking of the right IntP edge is needed, which is formulated in (9b).

- (9) (a) Syntax-phonology interface
Each linearization statement at the *CP*-Phase forms an *IntP*.
(b) *C*-marking statement II
The right edge of *IntP* licenses a *C*-element.

The rules in (9) make a number of predictions about the occurrence of demarcative enclitics in the left periphery (see Section 3.1) and the occurrence of enclitics at clausal boundaries (see Section 3.2). Demarcative clitics occur in Yucatec, Tzeltal, Mopán, and Itzá. Tzotzil differs from these languages in having different morphological exponents for the interpretable function, i.e., =*e* 'CL.DEF' in (5c), and for the demarcative function, i.e., =*un* 'CL' in (1). In Tzeltal and Itzá, it is the same morphological form that is used for both functions. Yucatec Maya displays a threefold deictic contrast that appears with definite DPs but is absent from the enclitic with demarcative function, in which case we find only the unmarked member of the set, namely =*e*' (see example (4) above). In Mopán Maya, demarcative clitics are a copy of the last

syllable's vowel, just as the clitics that spell out an interpretable feature in (6). There is no evidence for corresponding demarcative clitics in Jacaltec (based on the data presented by Day 1973 and Grinevald Craig 1977).

3.1 Left periphery

The languages at issue are V-initial, whereby Tzotzil, Tzeltal, Yucatec, Mopán, and Itzá are considered as VOS languages (cf. Aissen 1992: 48 for Tzotzil, Robinson 2002 and Shklovsky 2005: 32 for Tzeltal, and Hofling 1984 for the Yucatecan languages) and Jacaltec as VSO (see Day 1973: 64, Grinevald Craig 1977: 8). All these languages share the property of having two positions at the left periphery, a lower position that hosts focused constituents and a higher position that is used for topics and other left dislocated material. In line with previous accounts (Aissen 1992), I assume that the lower position is the specifier of an IP (=inflectional phrase) and I refer to it as a “pre-predicate position” (following Aissen 1987:18); the higher position is the specifier of a CP (=complementizer phrase) and I refer to it as left dislocated, see (10).⁶

(10) [CP XP [IP YP [VP ...]

The rules in (9) and the assumption that CP but not IP constitutes a phase straightforwardly predict that a demarcative enclitic will surface between the specifier of the CP and its complement, but not after the specifier of the IP. This difference is reported for Tzotzil in Aissen (1992: 57). The same difference holds for Yucatec Maya (see Skopeteas and Verhoeven 2009): left dislocated constituents are enclosed by a demarcative enclitic, see (11a), which is not licensed by properties of the numeration (in contrast to the enclitics licensed by definiteness in (4)).⁷ In line with the phase-based predictions, the pre-predicate position cannot be enclosed by an enclitic, see (11b). This position is diagnosticized through the verb morphology: the verb occurs in subjunctive which appears whenever the agent constituent occupies the pre-predicate position (also termed “agent-focus” construction, see Stiebels 2006).

(11) Yucatec

(a) Left dislocation

Raul=e' sáamal=e' yan u xobt-ik
 Raul=D3 tomorrow=D3 DEB A.3 whistle:TRR-INCMPL
u suku'n.
 A.3 elder.brother

'Tomorrow, Raul will whistle at his elder brother.'

(b) Pre-predicate position

Raul(=e') hàant òon.*
 Raul=D3 eat:TRR(SUBJ)(B.3.SG) avocado
 'RAUL ate avocado.'

Observational evidence from texts suggests the same distinction for Mopán Maya (negative evidence is not available). The topical first person pronoun in (12a) is accompanied with an enclitic that copies the last vowel of the pronoun. In contrast to the examples in (6), this enclitic is not licensed by properties of the numeration. The assumption that the pronoun occupies the specifier of the CP projection comes from the context, which licenses a contrastive topic. Example (12b) shows that the same pronoun in the specifier IP is not accompanied by an enclitic. Evidence that this constituent occupies the specifier of the IP comes from the fact that it follows the negative particle that is projected above spec,IP and below spec,CP in these languages (see Aissen 1992: 46).

(12) Mopán

(a) Left dislocation

Context: 'Try it and see! It is very good. It is very sweet.'

Inen=e wi'ij-en.
 1.SG=CL hungry-B.1.SG
 'As for me, I am hungry.' (Shaw ed., 1971: 422)

(b) Pre-predicate Position

Ma' inen a c'ülüjen-e'ex=e.
 NEG 1.SG A.2.SG lock-PL=CL
 'You did not locked ME up.' (Shaw ed., 1971: 418)

The same pattern holds for Itzá Mayan. The enclitic *=eh* 'CL' that right encloses definite DPs, see (5a), also occurs after left dislocated pronouns, possessed nouns and indefinite nouns (see Hofling 1991: 14 that accounts for these elements as topics). The contrast between left dislocation and pre-predicate position is already observed by Hofling (1991: 39), see the examples in (13).⁸

- (13) Itzá
- (a) Left dislocation
Layti'-oo'=eh yan-u-b'et-ik-oo' han-al.
 3-PL=CL OBLIG-A.3-make-PPM-PL eat-NR
 'They have to make food.' (Hofling 1991: 39)
- (b) Pre-predicate Position
pwes layti'-oo' yan-u-chäk-ik-oo' u-b'u'ul.
 well 3-PL OBLIG-A.3-cook-PPM-PL A.3-bean
 'Well, THEY have to cook their beans.' (Hofling 1991: 39)

In the available data from Petalcingo Tzeltal (Shklovsky 2005) and Tenejapa Tzeltal (Robinson 2002), left dislocated adverbs or arguments are generally not accompanied by an enclitic (apart from those that are licensed by definiteness), as exemplified in (14a). However, a few instances with an enclitic occur in contexts that do not involve focus of the preverbal constituent, see (14b). The available data do not provide evidence for a contrast between the two left peripheral positions.

- (14) Petalcingo Tzeltal
- (a) *Patil namij-ik bajel ta s-na.*
 afterwards move.away-PL DIR:away LOC A.3-house
 'Later, they moved away from the house.'
 (Shklovsky 2005: 145)
- (b) *witch kerem=e muh a te ta te'=e.*
 small boy=CL climb PT DET LOC tree=CL
 'The small boy climbed the tree.' (Shklovsky 2005: 149)

In sum, the data from the Yucatecan languages, Yucatec, Mopán, and Itzá, as well as the data from Tzotzil (see Aissen 1992: 57) show that the specifier of CP and not the specifier of IP is accompanied by a non-interpretable enclitic that surfaces at its right edge. The available data from Tzeltal data show that the occurrence of enclitics at the right edge of the specifier of CP is scarce.

3.2 Clausal boundaries

A second environment in which non-interpretable clitics occur is the right edge of clausal constituents. Sentence-initial subordinate clauses are right enclosed by a demarcative enclitic, see (15) (see detailed account for

Yucatec in Bohnemeyer 1998a, for Itzá Hofling 1991: 43-45, for Tzotzil Aissen 1992: 53).

- (15) (a) Yucatec
u ti'a'l a lúuk'-bal xàan=e'
 A.3 property A.2 absolve-DEAG also=CL
k-u mach-ik hun-hek' sip-che'
 IPFV-A.3 seize-IPFV one-branch (shrub)-tree
 'In order for you to be absolved, he seizes a branch of the Sip-tree.' (CHAAK_077)
- (b) Mopán
Abix ti c'üla'an-en ich so'oy=o'
 just LOC lock-1.SG in chicken.coop=CL
balo' ca' in c'üxe-ech ich so'oy
 thus CNJ A.1.SG lock-B.2 in chicken.coop
 'Just as I was locked in the chicken-coop, thus I will lock you in the chicken-coop.' (Shaw, ed., 1971: 418)

The occurrence of enclitics in (15) could also be accounted for as an effect of the left dislocation of the adverbial clauses, however this generalization does not account for the full range of data. Example (8) above shows that enclitics also occur at the right edge of matrix clauses when they are followed by a clausal complement. Two subtypes of such complements have to be distinguished (as already shown by Aissen 1992: 57 for Tzotzil). The verb 'say' (see (8)) takes a CP complement, that also can contain a left dislocated constituent. The verb *k'áat* 'want' corresponds to a modal element and is accompanied by a lower clausal complement (presumably IP), which (a) displays a verb in the subjunctive mood and (b) cannot contain a left dislocated constituent, see (16). The critical point for our assumptions is that CP-complements but not IP-complements are separated from the matrix verb by an enclitic, compare (8) and (16).

- (16) Yucatec
Pèdróoh=e' u k'áat(=e') u xok hun-p'éel analte'.*
 Pedro=CL A.3 wish=CL A.3 read(SUBJ) one-CL.INAN book
 'Pedro wants to read a book.'

Finally, right dislocated material may be separated from the clausal domain with an enclitic as illustrated in (17) from Itzá. This data pattern is in line with the generalization in (9), if we assume that right dislocated constituents are material that is added to the CP through adjunction.

- (17) Itzá
Chok'-o' a k'ab'=ih=eh, t-u-ka' p'el-il.
 put.in-IMP A.2 hand=CL.LOC=CL LOC-A.3-two CL.INAN-POS
 'Put in your paws, the two.' (Hofling 1991: 109)

Tzeltal differs again from the other languages in that it does not display any demarcative clitics between clauses. The typical examples in which we find such enclitics are sentence-initial subordinate clauses, as exemplified for Yucatec Maya in (15). Even in this context, Petalcingo Tzeltal texts do not display any enclitics, as exemplified in (18).

- (18) Petalcingo Tzeltal
Te me x-bah-at ta Majasil x-mil-ot-at.
 COMP EMPH INCMPL-go-B.3 LOC Majasil INCMPL-kill-PASS-B.2
 'If you go to Majasil, you will be killed.' (Shklovsky 2005: 68)

4. Placement of interpretable clitics

The previous sections have shown that Mayan languages display a class of enclitics that have a semantic effect (see Section 2) and a class of semantically vacuous enclitics that occur at particular intonational boundaries (see Section 3). The current section shows that enclitics of the former class do not attach to the licensing head but surface in positions of the linearization that may host an intonational boundary, i.e., their placement is conditioned by exactly the same rules that determine the placement of demarcative enclitics. In order to account for this data pattern, we assume that a rule that is inverse to (9b) applies, to the effect that any enclitic is bi-uniquely associated to an IntP boundary:

- (19) C-projection rule
 A C-element creates the right edge of an *IntP*.

In the following subsections, we examine three critical types of linearizations in which phonological material intervenes between the licensing head of an interpretable enclitic and the next intonational boundary on the right: (a) linearizations in which an enclitic is licensed by a head that is not adjacent to the right clause boundary, (b) linearizations involving a left peripheral constituent and its clausal complement, and (c) linearizations involving a head licensing an enclitic and a relative clause.

4.1 Right clause boundary

When an interpretable clitic is licensed within the clausal domain (i.e., not in the left periphery), then it surfaces at the right edge of the clause. In Yucatecan languages, this generalization may be observed through postverbal but non-clause final definite DPs. Compare the examples (20a-b) from Mopán Maya. The definite DP licenses a right edge clitic that appears right-adjacent to the licensing DP when the latter occurs clause finally, as in (20a). However, when the licensing DP is followed by additional material, as in (20b), the enclitic appears at the end of the clause. This is the dominant pattern in Yucatecan languages, as already illustrated for Yucatec Maya in (2c) and (29b) (see similar examples in Itzá, Hofling 1991: 239, sentence 224).

(20) Mopán

- (a) *c'ülbi ich so'oy aj coj=o.*
 lock-PASS in chicken.coop DEF cougar=CL.DEF
 'The cougar was locked in the chicken-coop.'
 (Shaw 1971, ed.: 418)
- (b) *u müch-aj aj coj ich so'oy=o.*
 A.3 grab-CMPL DEF cougar in chicken.coop=CL.DEF
 'She grabbed the cougar in the chicken-coop.'
 (Shaw 1971, ed.: 418)

The interpretable enclitic =*an* 'CL.1' in Jacaltec Maya shows similar properties. Apart from the case that is licensed by an element in the left periphery (see section 4.2), it always surfaces at the right clause boundary, as shown in (21).

(21) Jacaltec

- laŋan hintx'ahni xil kape=an, yet xcach huli.*
 I.was washing clothes=CL.1 when 2.SG came
 'I was washing clothes when you came.'
 (Grinevald Craig 1977: 280)

A few exceptions to the clause-final occurrence of clitics appear in contexts that suggest a right-dislocation of the material following the enclitic, as in example (17) with a demarcative clitic (see similar examples with interpretable clitics in Itzá, Hofling 1991: 110, sentences 113-115).

4.2 Left periphery

As shown by the contrast between (11a) and (11b), left dislocated material in Yucatec Maya is separated from its complement by a demarcative enclitic, while pre-predicate constituents cannot be separated from the predicate. Hence, the critical question is what happens when an interpretable enclitic is licensed by the left peripheral constituent. In line with the observations from non-interpretable enclitics, a left dislocated definite DP which is enclosed by the interpretable enclitic is grammatical, see (22a), but a pre-predicate definite DP with an enclitic on its right edge is not, as shown in (22b).⁹

- (22) Yucatec
- (a) *e xibpàal=o' k-u hàant-ik òon.*
 DEF man.child-D2 IPFV-A.3 eat-ICMPL avocado
 'The boy ate avocado.'
- (b) **e xibpàal=o' hàant òon.*
 DEF man.child-D2 eat:TRR(SUBJ)(B.3.SG) avocado
 'The BOY ate avocado.' (intended)

In order to express narrow focus on a definite DP, Yucatec Mayan speakers use a construction of contrastive left dislocation, as exemplified in (23). The available data from Itzá and Mopán Maya contain ample evidence that left dislocated definite DPs are right enclosed by an enclitic, but do not contain examples of definite DPs in the pre-predicate position.

- (23) Context: 'Who is eating avocado?'
- e xibpàal=o' leti' hàant òon.*
 DEF man.child-D2 3.SG eat:TRR(SUBJ)(B.3.SG) avocado
 'The boy, IT is running.'

The interpretable enclitic =*an* 'CL.1' in Jacalteco shows the same distribution, see (24a-b). Both examples involve a locative adjunct in the left periphery containing a first person possessive pronoun which licenses the first person enclitic. When the locative adjunct is left dislocated as exemplified in (24a), in which the locative expression may co-occur with a co-referent adverbial that occupies the pre-predicate position, the enclitic =*an* 'CL.1' surfaces on the right edge of the left periphery. When the locative adjunct has a focus interpretation and by hypothesis occupies the pre-predicate position as exemplified in (24b), then the enclitic =*an* 'CL.1' surfaces on the right edge of the clause.

- (24) Jacaltec
- (a) *w-atut an (hat) chuhuj kiŋ hecal.*
 1-house CL.1 there will.happen fiesta tomorrow
 ‘In my house, a fiesta will happen there tomorrow.’
 (Grinevald Craig 1977: 280)
- (b) *w-atut chuhuj kiŋ hecal an.*
 1-house will.happen fiesta tomorrow CL.1
 ‘IN MY HOUSE will happen a fiesta tomorrow.’
 (Grinevald Craig 1977: 279)

The available data from Tzotzil and Tzeltal Maya are inconclusive with respect to the behavior of enclitics that are licensed by definite DPs in the pre-predicate position.

4.3 Relative clauses

In the languages at issue, relative clauses are right adjacent to their head. Aissen (1992) observes that in Tzotzil the demarcative clitic =*un* ‘CL’ cannot occur at the left edge of a relative clause. The examples cited by Aissen (1992) suggest that this constraint also applies to the interpretable enclitic =*e* ‘CL.DEF’, which encloses definite DPs, see (25). In this example, an interpretable clitic =*e* ‘CL.DEF’ is licensed by the definite DP ‘the house’, but it surfaces clause-finally.

- (25) Tzotzil
- I-s-tuki-ik la li na (li)*
 CMPL-A.3-destroy-A.2/3.PL EVID DEF house COMP
j-meltzan-oj-tikotik xa ox=e.
 A.1-make-PF-A.1.PL.EXCL already just=CL.DEF
 ‘They destroyed the house that we had already just built (they say).’
 (Aissen 1992: 56)

The same difference is observed in Mopán, see (26a), in which the enclitic of the definite DP is realized at the end of the relative clause copying the last vowel of its phonological host and not of the licensing head (compare with (20)). In (26b) from Yucatec Maya, the interpretable enclitic =*o* ‘D2’ denoting that the referent is localized in a distal region is unambiguously licensed by the head of the relative clause (adverbs do not license interpretable enclitics in Yucatec Maya). Examples (26c-d) illustrate the same phenomenon in Itzá and Jacaltec.

- (26) (a) Mopán
 ... *in choic aj coj a watac=a.*
 A.3 trick DEF cougar DEF come=CL
 ‘...I am going to trick the cougar that is coming.’
 (Shaw 1971, ed.: 425)
- (b) Yucatec
táan u ko'-kol-ik-o'b le ch'uy-u'b
 PROG A.3 RED-tear-INCMPL-3.PL DEF hang-NR
yàan teh ka'nal=o'.
 exist LOC above-D2
 ‘They are tearing down the ch'uyu'b that is above.’
 (CHAAK.067)
- (c) Itzá
a' ha' t-in-wuk'-ah=eh ...
 DEF water PFV-A.1-drink-CMPL=CL.DEF
 ‘the water that I drank ...’ (Hofling 1991:14)
- (d) Jacaltec
wohtaj naj xul ewi=an.
 A.1-know 3.SG came yesterday=CL.1
 ‘I know the one that came yesterday.’
 (Grinevald Craig 1977: 279)

Tzeltal differs again from the further languages, in that interpretable enclitics may intervene between the relative clause and its head, see (27).

- (27) Petalcingo Tzeltal¹⁰
me sak=e' mach'a yakal ta we'-el.
 DET white=CL who PROG(B.3) LOC eat-PART
 ‘The white one that’s eating.’ (Shklovsky 2005: 70)

5. Haplology

Section 4 shows that interpretable clitics surface at the right edge of a prosodic domain that is identified as intonational phrase. Whenever more than one interpretable clitic is licensed within the same intonational phrase, then a rule of haplology bans the concatenation of multiple clitics on the right edge. This is illustrated by the following examples from Tzeltal and Mopán. Compare example in (14b) with (28a) below: the definite DP ‘the tree’ that is enclosed with an enclitic in (14b) is not accompanied by an enclitic in (28a), since in the latter case its right edge does not coincide with a right intonational phrase boundary. We may

expect that this enclitic will surface at the right edge of the clause, at which the enclitic of the following DP appears. However, the enclitics of the two postverbal DPs are never concatenated at the phrase-final boundary. In all these cases, a single enclitic appears instead. Exactly the same phenomenon is illustrated in (28b) from Mopán Maya (see similar examples in Itzá in Hofling 1991:291, sentence 132).

- (28) (a) Tzeltal
 ... *och pahk-aj-uk ta y-ahkolal te te' te*
 enter flat-VR-IRR LOC A.3-top DEF tree DEF
witch kerem=e.
 small boy=CL
 '... the small boy flattened himself on the tree'
 (Shklovsky 2005: 149)
- (b) Mopán
 ... *a noochwinik etel a nooch'up=u.*
 DEF big.man with DEF big.woman=CL
 '... the big man with the big woman.' (Shaw ed., 1971: 418)

The same phenomenon is already observed for Jacalteco: whenever more than one first person pronoun occurs within the same clausal domain, a single first person enclitic occurs at the right edge of the clause (see Grinevald Craig 1977: 284, also presenting negative evidence for this generalization). The haplological omission is not obvious in Yucatec Maya in which the enclitics contrast for the encoding of different deictic features. Example (29a) illustrates a possibility for two different enclitics to co-occur in a sentence: the first enclitic occurs at the right edge of the left dislocated DP and the second enclitic at the right edge of the clause. However, when both DPs are in situ as in (29b), then the occurrence of a clause medial enclitic is judged as non-grammatical by the native speakers.¹¹

- (29) Yucatec
 (a) *le chàan tùunich=a' náach yàan ti le*
 DEF small stone=D1 far exist LOC DEF
nojoch tùunich=o'.
 big stone=D2
 '... this small stone is far from that big stone.'

- (b) *náach yàan le chàan tìunich*(=a') ti*
 far exist DEF small stone=D1 LOC
le nojoch tìunich=o'
 DEF big stone=D2
 '... this small stone is far from that big stone.'

Haplological omission does not apply in Tzotzil Maya: the interpretable clitic =*e* 'CL.DEF' and the non-interpretable clitic =*un* 'CL' may co-occur at the right phrase edge, as illustrated by the right edge of a left dislocated constituent in (29). The order of the two clitics is fixed when they co-occur (=un=*e*, see Aissen 1987: 8), though it is not the order that we would expect in syntactic view (the interpretable enclitic should be closer to the licensing head than the demarcative one).

- (30) Tzotzil
li rey=un=e, chak' kastiko=un.
 DEF king=CL=CL.DEF gave punishment=CL
 '... the king meted out punishment' (Aissen 1992: 54)

In conclusion, we assume that a rule of haplology applies in all contexts and bans any occurrence of multiple enclitics in Yucatec, Itzá, Mopán, and Tzeltal. This rule is not restricted to enclitics belonging to the same stratum (i.e., multiple interpretable clitics), but bans the possibility of concatenating enclitics on the same phrase boundary. We assume that in all these languages interpretable and demarcative clitics are members of the same morphological class. In Itzá, Tzeltal, and Mopán, interpretable and non-interpretable clitics have the same form (=eh, =e, or copy of the last vowel respectively). In Yucatec Maya, the demarcative enclitic =*e*' is the unmarked member of a set of enclitics with deictic properties. Tzotzil is special in that it displays two formally distinct enclitics for the interpretable and the demarcative function (=e and =un respectively). The non-application of the rule of haplology for these enclitics suggests that these two clitics are treated as members of different classes, such that complementary distribution does not apply.

6. Conclusions

Two classes of enclitics are observed in several Mayan languages. *Demarcative enclitics* do not have a contribution to the propositional content and are used to delimit the right boundary of non-final intonational phrases. *Interpretable enclitics* spell out interpretable features of the

numeration. These enclitics are also associated with particular prosodic properties and are aligned with the right boundary of an intonational phrase.

We accounted for the distributional properties of these enclitics by means of the rules outlined in (31). The statements in (31c-d) cannot be subsumed under a single rule of bi-unique association between enclitics and IntP boundaries (C-element \leftrightarrow right edge of IntP), since (31c) refers to the licensing of demarcative enclitics and (31d) to the interaction of demarcative and interpretable enclitics with phrasing. Hence, the statement in (31d) refers to a superset of the elements referred to by (31c).

- (31) (a) C-marking statement I
An interpretable feature F_i is spelled out though a clitic C_i , that belongs to the set of interpretable enclitics in the language.
- (b) Syntax-phonology interface
Each linearization statement at the CP-Phase forms an IntP.
- (c) C-marking statement II
The right edge of IntP licenses a clitic C_d , which is the sole member of the set of demarcative enclitics in the language.
- (d) C-projection rule
Any C_j belonging to the superset of interpretable/demarcative enclitics creates the right edge of an IntP.

The Mayan languages examined in this article differ with respect to the features of the numeration that license enclitics: definite DPs license an enclitic in Yucatecan languages and Tzotzil first person pronouns license an enclitic in Jacaltec. Such differences are situated in the lexicon and relate to the range of interpretable features to which the feature F_i in (31a) applies.

Petalcingo Tzeltal displays systematic differences to the further languages with respect to the properties for which the available evidence is conclusive. The evidence for a demarcative enclitic is weak (see (14b)), and there is no evidence that the definite enclitic attaches to the next available intonational boundary, see counterevidence in (27). This data pattern suggests that enclitics in this language are reanalyzed as DP-final markers that do not attach to the intonational phrase boundary.

Tzotzil Maya is special in providing two morphologically distinct enclitics: the C-marking statement I in (31a) licenses an enclitic =*e*, which is subject to the C-projection rule in (31d), while the C-marking statement II in (31c) licenses an enclitic =*un*. The fact that these enclitics are not in

complementary distribution shows that they belong to different morphological classes.

The classes of Mayan enclitics presented in this article are morphological markers that either associate with intonational phrase boundaries or – in the case of demarcative enclitics – just spell out intonational phrase boundaries. Hence, they present two different possibilities of interaction of morphology with prosodic structure and indirectly with syntax. In conclusion, this paper presented a typologically striking phenomenon, i.e., a class of morphemes that surface outside the projection of their licensing heads. Though the data pattern is highly complex, the assumption that they associate with prosodic entities and the rules of syntax-phonology interaction offer a compositional account that adequately explains the peculiarities of their placement in the utterance.

Acknowledgments

I am grateful to Elisabeth Verhoeven for her close collaboration in all aspects of this study, as well as to Gisbert Fanselow, Caroline Féry, Shinichiro Ishihara, and Barbara Stiebels for their discussion on the Yucatec Mayan enclitics. The main ideas of this article were presented at the workshop “Meeting Clitics” at the Linguistics Institute of Barcelona, August 2008. Thanks go to the organizers and the audience for their discussion and in particular to Arsalan Kahnemuyipour and Lisa Brunetti for their comments, as well as to an anonymous reviewer for his/her comments on a previous manuscript. This article is part of the project D2 “Typology of Information Structure” at the University of Potsdam (part of the research institute “Information Structure”, sponsored by the DFG).

Abbreviations

A= person affix, class A; AN= animate; B= person affix, class B; C= clitic; CL= class; CMPL= completive; COMP= complementizer; CNJ= conjunction; D1= deictic clitic 1; D2= deictic clitic 2; D3= deictic clitic 3; DEB= debitive; DEF= definite; DEM= demonstrative; DET= determiner; DIR= directional; DIST= distal; EMPH= emphatic; EVID= evidential; EXCL= exclusive; F= feminine; INAN= inanimate; INCMPL= incompletive; INTRV= introversive; IPFV= imperfective; IRR= irrealis; LOC= locative; M= masculine; NEG= negation; NR= nominalizer; OBLIG= obligative; PART= participle; PASS= passive; PF= perfect; PFV= perfective; PL= plural; PPM= proximal patient marker; PROG= progressive; PT= particle; RED= reduplication; REL= relationalizer; SG= singular; SUBJ= subjunctive status;

TERM= terminative; TRR= transitivizer; VR= verbalizer; 0= meaningless element; 1= 1st person; 3= 3rd person.

References

- Aissen, J.L. 1987, *Tzotzil Clause Structure*. Dordrecht: D. Reidel.
 —. 1992, Topic and focus in Mayan. *Language* 68: 43-80.
 Blair, R.W. and Vermont-Salas, R. 1965, *Spoken (Yucatec) Maya*. Chicago, Il: University of Chicago.
 Bohnemeyer, J. 1998a, “Sententiale Topics im Yukatekischen”. In D. Zaefferer (ed.), *Deskriptive Grammatik und allgemeiner Sprachvergleich*. Tübingen: Niemeyer, 55-85.
 —. 1998b, *Time Relations in Discourse: Evidence from a Comparative Approach to Yukatek Maya*. Ph.D. Dissertation, Katholieke Universiteit Brabant.
 Chomsky, N. 2001, “Derivation by phase”. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1-52.
 —. 2005, Three factors in language design. *Linguistic Inquiry* 36.1, 1-22.
 —. 2007, “Approaching UG from below”. In U. Sauerland and H.-M. Gärtner (eds.), *Interface + Recursion = Language?* Berlin: Mouton De Gruyter, 1-29.
 Day, C. 1973, *The Jacaltec Language*. Bloomington, In: Indiana University.
 Dürr, M. 1987, *Morphologie, Syntax und Textstrukturen des (Maya) Quiche des Popol Vuh*. Bonn: Holos.
 Edmonson, B. W. 1988, *A descriptive grammar of Huastec (Potosino dialect)*. Ph.D. Dissertation, Tulane University.
 Féry, C. 2001, “Focus and Phrasing in French”. In C. Féry and W. Sternefeld (eds.), *Audiat Vox Sapientiae: A Festschrift for Arnim von Stechow*. Berlin: Akademie Verlag, 153-81.
 —. 2007, “The prosody of topicalization”. In K. Schwabe and S. Winkler (eds.), *On Information Structure, Meaning and Form: Generalizations across Languages*. Amsterdam: John Benjamins, 69-86.
 Féry, C. and Ishihara, S. 2008, “How focus and givenness shape prosody”. In M. Zimmerman and C. Féry (eds.), *Information Structure from Different Perspectives*. Oxford: Oxford University Press.
 Grinevald Craig, C. 1977, *The Structure of Jacaltec*. Austin: University of Texas Press.
 Hanks, W.F. 1990, *Referential practice: language and lived space among the Maya*. Chicago, Il: The University of Chicago Press.

- Hofling, C.A. 1984, On proto-Yucatecan word order. *Journal of Mayan Linguistics* 4.2, 35-64.
- . 1991, *Itzá Maya Texts: With a Grammatical Overview*. Salt Lake City, Ut: University of Utah Press.
- Kratzer, A. and Selkirk, E. 2007, Phase theory and prosodic spellout: The case of verbs. *Linguistic Review* 24, 93-135.
- Inkelas, S. and Zec, D. 1995, "Syntax-phonology interface". In J.A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 535-549.
- Ishihara, S. 2007, Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO). *Linguistic Review* 24, 137-167.
- Lehmann, C. 1990, Yukatekisch. *Zeitschrift für Sprachwissenschaft* 9.1-2: 28-51.
- . 2003, *Possession in Yucatec Maya*. ASSIDUE: Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt, Nr. 10.
- Robinson, S. 2002, Constituent order in Tenejapa Tzeltal. *International Journal of American Linguistics* 68.1, 51-80.
- Schumann Gálvez, O. 1971, *Descripción estructural del Maya Itzá del Petén, Guatemala C.A.* México: Universidad Autónoma de México.
- . 2000, *Introducción al Maya Itzá*. Universidad Autónoma de México, Instituto de Investigaciones Antropológicas.
- Selkirk, E. 1984, *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Shaw, M. (ed.) 1971, *According to our Ancestors: Folk Texts from Guatemala and Honduras*. Guatemala: Summer Institute of Linguistics.
- Shklovsky, K. 2005, Person marking in Petalcingo Tzeltal. Ph.D. Dissertation, Reed College.
- Skopeteas, S. and Verhoeven, E. 2005, Postverbal argument order in Yucatec Maya. *Sprachtypologie und Universalienforschung* 53.1: 71-79.
- . 2009, "Left peripheral arguments and discourse interface strategies in Yucatec Maya". In I. Kucerova and R. Vermeulen (eds.), *Interface-Based Approaches to Information Structure*, Oxford: Oxford University Press.
- Stiebels, B. 2006, Agent focus in Mayan languages. *Natural Language and Linguistic Theory* 24: 501-570.
- Tonhauser, J. 2003, "F-constructions in Yucatec Maya". In J. Anderssen, P. Menéndez-Benito and A. Werle (eds.), *The Proceedings of SULA 2*. Vancouver, BC: GLSA, University of Massachusetts, Amherst, 203-223.

- Truckenbrodt, H. 1999, On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30.2: 219-255.
- . 2007, “The syntax-phonology interface”. In P. de Lacy (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 435-456.
- Zubizarreta, M.L. and Vergnaud, J.-R. 2005, “Phrasal stress and syntax”. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*. Oxford: Blackwell, 522-568.

Notes

¹ Whenever the source of the cited example is not indicated, it is elicited during my fieldwork on Yucatec Maya (Yaxley and Felipe Carrillo Puerto, Quintana Roo, México: December 2004 and March 2008).

² The class of enclitics presented here is not available in Huastec, the sole member of the Huastecan branch (see Edmonson 1988), and in the languages of the Quichean-Mamean branch, already in the oldest documented stages, dating back to the 16th century (see Colonial Quiché in Dürr 1987). Similarly, no instances of this type of enclitics are found in the available materials for Chol (Cholan-Tzeltalan branch), Kanjobal and Acatec (Kanjobalan-Chujean branch), and Lacandon (Yucatecan branch).

³ The contrast between the clitics =*ah* ‘proximal’ and =*oh* ‘distal’ appears in Itzá only with demonstratives (see Hofling 1991: 15, Schumann Gálvez 2000: 80f.). In this case, the former enclitics are reanalyzed as part of the demonstrative, which is supported by the fact that they can be combined with the definite final clitic, e.g., *a' kol he'l-a'=eh* (DET milpa DEM-PROX=CL.DEF) ‘this milpa here’, whereby the reanalyzed former enclitic *-a'* is part of the demonstrative and hence does not contrast to the enclitic =*eh*.

⁴ See also Robinson (2002: 55) who observes that definite DPs are accompanied by an enclitic in Tenejapa Tzeltal.

⁵ We assume that the prosodic layer at issue is an intonational phrase (IntP), since there is evidence for a lower layer of prosodic constituency, i.e. a major phrase (MaP) that corresponds to lower phase statements (vP). The consequence of this choice is that IntP's are prosodic constituents that enclose either clauses or topics (see previous use of the concept of IntP for these types of constituents in Féry 2007).

⁶ The assumption of a TP (=tense phrase) for Mayan languages faces the problem that these languages do not encode tense through verb morphology or auxiliaries. The head I' of the IP projection is occupied by an auxiliary, encoding aspectual and/or modal categories.

⁷ Observation in texts reveals that left dislocated material is enclosed by an enclitic more often than not; enclitics may be missing due to performance factors (e.g., quick speech tempo).

⁸ Left dislocated adverbs are also enclosed by demarcative enclitics in Itzá, see Hofling (1991: 208, example 112).

⁹ See also the same observation under a different account in Tonhauser (2003).

¹⁰ The relative clause contains a periphrastic progressive, formed through an inflected progressive auxiliary and a nominalized form of the lexical verb headed by a preposition (see Shklovsky 2005: 97).

¹¹ According to the discussion in section 3.2, the version with a clause-medial enclitic should be grammatical when the second DP is right dislocated. I assume that native speakers judge this example as non-grammatical when presented out-of-the-blue, since they do not reconstruct the contextual conditions of right dislocation in the field situation.