Clitic placement in Mam (Mayan) requires a host restriction*

Tessa Scott

University of California Berkeley

1. Introduction

In San Juan Atitán Mam (SJA Mam), polar questions are expressed by the second position clitic *=m*. This clitic can attach to a plethora of categories including determiners, tense markers, aspectual markers, and negation by means of Local Dislocation (Embick and Noyer 2001). However, *=m* never attaches to the right of another clitic. This pattern raises two important theoretical issues for Distributed Morphology (DM; Halle and Marantz (1993)). The first is whether or not the clitic status of a morpheme is encoded in the grammar. The second is the need for a theory of possible hosts for clitics. Based on the SJA Mam data, I conclude that the linearization instructions for the polar question clitic must reference the clitic status of its host, suggesting that clitics are different in crucial ways from other morphemes in the grammar.

The first relevant theoretical issue for DM arises from the proposal that clitics are not a unique class differentiated in grammar. In an overview article on Distributed Morphology, (Harley and Noyer 1999: 5) say about the status of clitics: "'Clitic' is not a primitive type in DM but rather a behavior which an element may display.' Much subsequent work has maintained this view, with the focus being on how to delineate the behaviors of mobile morphemes like clitics. Embick and Noyer (2001) argue for two post-syntactic merger operations, Lowering and Local Dislocation, that can account for clitic placement, among other empirical facts. Lowering, in (1), operates on syntactic structure and may instruct clitics to lower to the head of their complement.

(1) *Lowering of X^0 to Y^0* (Embick and Noyer 2001: 561)
$[_{XP}\ X^0 \ldots [_{YP} \ldots Y^0 \ldots]] \rightarrow [_{XP} \ldots [_{YP} \ldots [_{Y^0}\ Y^0 + X^0\ ] \ldots]]$

*For valuable comments I would like to thank Amy Rose Deal, Line Mikkelsen, Peter Jenks, Madeline Bossi, and Marie-Luise Popp. I also would like to audiences at UC Berkeley Syntax and Semantic Circle and NELS 49 (Cornell). I must also thank all those involved in the UC Berkeley Field Methods class Fall 2018 and Spring 2019. Lastly, I want to thank Henry Sales and Gladiola Aguilar for being so helpful and patient in sharing their language with me.

Local Dislocation operates on linearized structure and results in movement to the other side of the first linearized element. (2) shows this schematically, where *a* * *b* stands for 'a must linearly precede b.'

(2) *Local Dislocation of X to Z* (Embick and Noyer 2001: 562)
[X ∗ [Z ∗ Y]] → [[$_{Z^0}$ Z + X] ∗ Y]

Under this view, an element's status as a clitic or non-clitic cannot trigger or bleed morphological rules; 'clitic' is simply a name given to a range of placement instructions available for a given morpheme. I will show that in SJA Mam, the clitic status of a potential host determines the behavior of the second position polar question clitic, suggesting that the clitic status of morphemes must be referenced in the grammar.

The second issue for clitics in DM concerns the possible hosts for clitics. While Embick and Noyer (2001) offer a theory on movement instructions, it does not specify restrictions on the features of the element to which a clitic attaches. This might be compared with a view like Anderson (1993), which offers the four parameters in (3) to capture clitic behaviors.

(3) *Clitic parameters* (Anderson 1993)
- a. Scope (sentence, nominal phrase, etc.)
- b. Anchor (the element it attaches to)
- c. Orientation (proclitic or enclitic)
- d. Domain (phonological or syntactic)

The Domain parameter captures the same distinction between Lowering and Local Dislocation: Anderson observes that second position clitics either appear after the first word or after the first constituent. The Anchor parameter is of particular interest because there is no parallel way in DM to specific the host (i.e. anchor) of a clitic, which predicts that the placement of a clitic should not vary based on the host element. I'll show that in SJA Mam, while the second position clitic =*m* usually undergoes Local Dislocation, it cannot undergo this dislocation if the element in first position is a clitic. This suggests not only that we need a theory of potential hosts but also that whether an element is a clitic or not can be specified within such a theory.

The paper is organized as follows: I give an overview of polar questions in SJA Mam in Section 2, showing what counts as first position. In Section 3 I show that when a clitic occupies the initial position, Local Dislocation is blocked. In Section 4 I give an analysis of the SJA Mam facts by stating a host condition in addition to Local Dislocation for =*m*. Section 5 provides more data from enclitics supporting the analysis. Section 6 concludes.

2. Mam polar questions

2.1 Mam background: initial position

SJA Mam is a dialect of Mam (Mayan) spoken in San Juan Atitán, Guatemala.[1] The neutral word order in SJA Mam is VSO. The verbal complex is largely the same across Mayan languages, which is illustrated for Mam in Table (4). The main locus of variation is the position of absolutive marking (Coon 2016): absolutive is either marked before ergative (high-abs) or at the end of the verb (low-abs). SJA Mam is a 'high-abs' language, shown by the placement of absolutive before ergative.[2]

(4) *Mam verbal complex*

(NEG)	ASPECT	DIR	(ABS)	ERG	ROOT	(VOICE)	PERS
	o	txi		n-	wan		=i
	PFV	DIR		1SGA	eat		=i

(5) O txi n-wan=i wab'j.
PFV DIR 1SGA-eat=PERS tortilla
'I ate tortillas.'[3]

In addition to basic VSO word order, focus and topic movement to initial position is extremely common, like many Mayan languages. Relative clauses and wh-words may also appear in initial position, shown in (6).

(6) a. *Focus*
[A Henry] ma tz'ok b'yon-t a Xwan.
[DET Henry] PROX DIR hit-AF DET John
'HENRY hit John.'

b. *Relative clause*
[A=j xuj o b'aj sch'in-t u'j] o tz'ok b'iyon-t a Eric.
[DET=REL woman PFV DIR read-AF book] PFV DIR hit-AF DET Eric
'The woman who read the book hit Eric.'

[1] All data here come from SJA Mam speakers living in Oakland, California.

[2] Absolutive is only marked in intransitives in SJA Mam, in contrast with Ixtahuacán Mam (England 1983a,b) and most Mayan languages (Coon 2016) in which both ergative and absolutive is marked in transitive clauses.

[3] Abreviations: 1,2,3 = first, second, third person, A = set A agreement (ergative and possessive), B = set B agreement (absolutive), AF = agent focus, CLF = classifier, DET = determiner, DIR = directional (lexically specified), DS = directional suffix, INDF = indefinite determiner, LOC = locative PERS = local person marker, NEG = negation, PFV = perfective aspect, PL = plural, PROG = progressive aspect, PROX = proximate aspect, Q = polar question, SG = singular, REL = relative marker, RN = relational noun.

c. *Wh-question*
[Al] o kub' qes-n-t t-e Noah?
[who] PFV DIR cut-DS-AF 2/3SGA-RN Noah
'Who cut Noah?'

These facts indicate that the word or constituent in initial position in SJA Mam can come from a variety of syntactic categories and take a range of phonological shapes. This gives us a starting ground for testing the nature of the second position clitic, *=m*.

2.2 Distribution of *=m*

The polar clitic *=m* appears in all polar questions. When the first morpheme is a stand-alone word like in (7), *=m* appears as an enclitic to it. In (7a) the first word is the directional, *kxe'l*. (7b) is the polar question counterpart where *=m* is an enclitic to the directional.

(7) a. Kxe'l t-kxun q'a i'ẍ.
DIR 2/3SGA-chew boy corn
'The boy will eat corn.'

b. Kxe'l =m t-kxun q'a i'ẍ?
DIR =Q 2/3SGA-chew boy corn
'Will the boy eat corn?'

This pattern is robust and can be seen in the following examples with the determiner, aspect, negation, and locative predicate, as well as many other categories not shown here.

(8) a. *Determiner*
A =m qini o t-il jun ch'it?
DET =Q 1SG PFV 2/3SGA-see one bird
'Did I see a bird?'

b. *Aspect*
Ma =m t-il Xuan a Liy?
PROX =Q 2/3SGA-see Xuan DET Liy
'Did Xuan see Liy?'

c. *Negation*
Nti' =m ma w-il=i a q'a Eric?
NEG =Q PROX 1SGA-see=PERS DET CLF Eric
'Didn't I see Eric?'

d. *Locative predicate*
At =m q'a?
LOC Q boy?
'Is he here?'

In all of the examples above, the initial word is mono-morphemic. In order to understand the syntactic size of the initial element to which the polar clitic attaches, we need to know whether this element can be multi-morphemic. Relational nouns, functioning like prepositions across the Mayan language family (Coon 2016), are multi-morphemic, showing set A agreement with the noun they introduce.

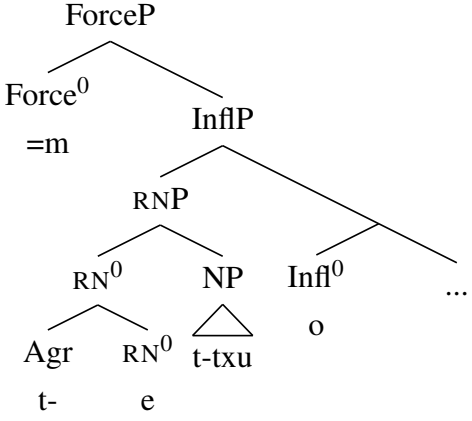
(9) O txi t-q'o-'n Gladiola u'j [**t-e** xna'tzal].
PFV DIR 2/3SGA-give-DS Gladiola book [2/3SGA-RN teacher]
'Gladiola gave the book to the teacher.'

When a relational noun phrase is focused, the relational noun appears in initial position. When the polar question clitic is added, it attaches as an enclitic to the relational noun. The placement of the clitic after the entire relational noun shows that the clitic does not simply attach to the first morpheme, but to some other unit.

(10) [T-e =*m* t-txu] o txi t-q'on Noah jun ne tal
[2/3SGA-RN =Q 2/3SGA-mother] PFV DIR 2/3SGA-give Noah INDF CLF small
txian?
dog
'Did Noah gave a puppy TO HIS MOTHER ?'

The placement of =*m* in (10) also shows that this second position clitic does not attach to the first major constituent, indicated with brackets. Following Rizzi (1997), I assume the polar question clitic orginates in Force^0. Following Aissen (1992), I take focus constructions across Mayan to reflect movement to a Spec,InflP. In (11), *te ttxu* is the focused relational noun phrase in Spec,InflP. If the polar clitic placement were the result of Lowering, we would expect it to attach to Infl^0, creating *o=m*. However, in (10), the clitic attaches to the relational noun, *t-e*, the first linearized element in (12). The data strongly indicate the placement of =*m* is derived by Local Dislocation (13).

(11) *Syntactic structure*



(12) *Linearization*
[=m [t-e * t-txu] * [o] ...

(13) *Local Dislocation*
[t-e=m * t-txu] * [o] ...

3. Progressive aspect triggers *tzu* support

While the perfective and proximate aspect markers are phonological words (*o* and *ma* respectively), the progressive aspect marker *n*= does not contain a vowel and leans on the following element. In (14a), it is a proclitic to the set B marker and in (14b) it is a proclitic to the verb.

(14) a. N=chin wan=i.
PROG=1SGB eat=PERS
'I am eating.'

b. N=kxun q'a i'ẍ.
PROG=chew boy corn
'The boy is chewing corn.'

In progressive constructions, the polar question clitic does not invert via Local Dislocation with the linearly first element (i.e. the proclitic *n*=). Instead, a morpheme is added (*tzu*) to host the polar question clitic =*m*. In (15), the polar question clitic attaches to *tzu*.

(15) Tzu =m n=kxun q'a i'ẍ?
tzu =Q PROG=chew boy corn
'Is the boy chewing corn?'

I argue here that *tzu* is the epenthetic syllable to host =*m*. The first argument for this is that *tzu* is not a set B marker co-referencing *q'a* in (15). This is a possible hypothesis because the second and third person set B marker has two allomorphs: it is typically null, though it appears as *tz*= before vowels.[4]

(16) a. Ma ∅ pon q'a.
PROX 2/3SGB arrive boy
'He arrived (there).'

b. Ma tz=ul q'a.
PROX 2/3SGB=arrive boy
'He arrived (here).'

The *tzu* in (15) at first might appear to be the realization of *tz*= in (16b). However, this is not the case; in (17), the only argument is third person plural and the set B morpheme tracking it, *chj*=, appears as a proclitic on the verb. The clitic *n*= attaches to the resulting stem. The polar clitic still attaches to *tzu* in initial position, suggesting that *tzu* is not an agreement morpheme.

[4]When 2/3SGB is null, I do not gloss it elsewhere in this paper.

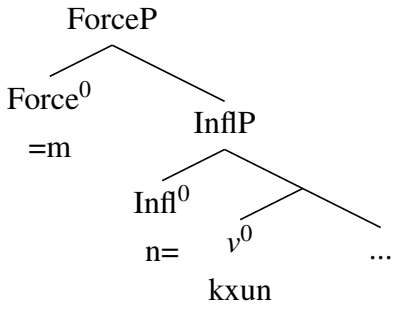
(17) Tzu =m n=chj=ajqelan qa jil chej?
tzu =Q PROG=2/3PLB=run PL CLF horse?
'Are the horses running?'

Based on its presence in (17), I conclude */tzu/* is the epenthetic syllable which hosts the polar clitic when it does not have a suitable host, similar to 'do-support' in English.[5] In Section 4 I argue that the host would have been *n=* 'PROG' and it is not suitable because it is a clitic.

4. Restriction against clitic hosts

An analysis of the placement of the second position polar question clitic in SJA Mam requires an explanation beyond simple Local Dislocation (Embick and Noyer 2001). The explanation should capture the generalization that *=m* does not undergo Local Dislocation in the presence of the progressive proclitic *n=*. The generalization I propose is that *=m* cannot attaches to another clitic. Following Aissen (1992), I assume aspect markers are in Infl^0 and following Coon and Carolan (2017)'s analysis of Chuj, I also assume that the verb moves to the head at the edge of the verbal projection,[6] resulting in the structure in (18). With this structure, the linearization is given in (19).

(18) *Syntactic structure*



(19) *Linearization of (15)*
[=m [n=] * [kxun] ...

(20) a. *Local Dislocation blocked:*
*[[n==m] * [kxun] ...
b. *Tzu insertion:*
[tzu=m * [n=] * [kxun] ...

To rule out (20a), the placement instructions for *=m* must reference the clitic status of *n=*. I interpret the two morphological merger operations Lowering and Local Dislocation from Embick and Noyer (2001) as only partial placement instructions for morphemes. In addition, clitics need to be specified as leaning left (enclitics) or right (proclitics) (Orientation from Anderson (1993)). What I propose, based on the *tzu* insertion in (15), is that the polar question clitic in Mam has a third specification regarding its host: it cannot attach to another clitic. As we have seen, the repair is that it does not undergo Local Dislocation at all: phonological *tzu* support is inserted to host the enclitic. The three placement instruc-

[5]However, note that in Chuj, the imperfective morpheme is itself *tz* (Carolan et al. 2015). In simple imperfective declaratives in SJA Mam, *tz* is not observed. It is possible that the *tzu* here is an older form of the imperfective marker and only used in polar questions. To test this, we can test polar questions with other proclitics. I leave this question to future investigation.

[6]Coon and Carolan (2017),Clemens and Coon (2018), and others label this head *ss*^{0} for 'status suffix'. I give the label here as v^0 for simplicity.

tions needed for the polar question clitic *=m* in SJA Mam are listed below. (21a) and (21b) occur in order; if the restriction is not met, a repair must occur.

(21) *Clitic placement instructions for =m in SJA Mam*

a. First, undergo Local Dislocation.
b. Second, lean left.
c. Host restriction: Host must not be a clitic.

In addition to the placement instructions in (21), I propose that, like all clitics in SJA Mam, the polar question morpheme must be specified as a clitic itself. This must be stipulated as a diacritic that simply carries the label 'clitic' which the placement instructions of other clitics can see. While this adds a stipulative and perhaps unwanted diacritic to the grammar, we already need a similar type of feature to stipulate the leaning specification in the placement instructions. In DM, Lowering and Local Dislocation provide algorithms that produce the desired outputs for the Domain of clitics. However, we still need to stipulate whether a clitic leans left or right. The data from clitic placement in SJA Mam discussed here show that an additional stipulation is required; an element's clitic status must be accessible in the grammar, since clitics can impose the requirement that their host not be another clitic.

5. Predictions in Mam

The conclusion for SJA Mam is that the polar question clitic *=m* cannot attach to another clitic. This conclusion makes the prediction that if the element in initial position ends in an enclitic, *=m* cannot attach to it. This prediction is born out. In the following examples, *=m* always attaches inside of other enclitics, regardless of the type of enclitic.

If the subject of a verb or the noun introduced by a relational noun is local person (except first plural inclusive), the enclitic *=i*/*=ni* appears on the verb or relational noun, shown in (22).[7] When the verb or relational noun appear in first position in polar questions, the question clitic *=m* attaches inside of the PERS clitic:

(22) [W-uk' =m =i] o kub' t-qes-n=i chi'b'j?
[1SGA-RN =Q =PERS] PFV DIR 2/3SGA-cut-DS=PERS meat
'Did you cut the meat WITH ME?'

(23) T-aj =m =ni jun lo'b'j?
2/3SGA-want =Q =PERS INDF banana
'Do you want a banana?'

The same local person enclitic that appears on relational nouns and verbs also appears on local person pronouns. The second person singular pronoun is made up of a determiner and

[7] Among Mayan varieties, this enclitic is unique to Mamean and appears in other Mam dialects with slightly different phonological shapes (see England (1983b) for Ixtahuacán Mam).

the person enclitic: *a=i*, written orthographically as *ay*. When *ay* is in initial position, *=m* attaches inside of *=ni*.

(24) [A =m =ni] o t-il jun ch'it?
[DET =Q =PERS] PFV 2/3SGA-see INDF bird
'Did YOU see a bird?'

Further evidence that *=m* never attaches to enclitics comes from the relative clause marker, *a=j*, which is made up of the determiner *a* and the enclitic *=j*. The question clitic attaches to the determiner.

(25) [A =m =j xuj o b'aj wan-t wab'j] o tz'ok b'yon-t a Eric?
[DET =Q =REL woman PFV DIR eat-AF tortilla] PFV DIR hit-AF DET Eric
'Did the woman who ate the tortilla hit Eric?'

The relative marker *a=j* is homophonous with *aj* 'when'. Compare (25) to (26), in which *=m* cliticizes to the outside of *aj* because it is monomorphemic.

(26) Aj =m tz-ul Liy, kxe'l *=m* t-kxun q'a i'ẍ?
when =Q 2/3SGA-arrive Liy, will =Q 2/3SGA-chew boy corn
'When Liy arrives, will the boy chew corn?'

The data in this section further point towards a host restriction on the placement of the polar question clitic *=m* in SJA Mam. In Section 3 we saw that *=m* never attaches to proclitics and here we see that it never attaches to enclitics, regardless of the phonological shape of the enclitic or the structure of the sentence. The full range of data can be captured by the Host Restriction that the host must not be a clitic.

6. Conclusion

In this paper I have shown that the polar question clitic *=m* in SJA Mam undergoes Local Dislocation (Embick and Noyer 2001) and attaches as an enclitic to the linearly first element in the sentences. If the element in initial position is a clitic or ends in a clitic, *=m* does not attach to the clitic. To account for this, I propose a Host Restriction in the placement instructions for *=m*: The host must not be a clitic. This conclusion has two major implications for accounting for clitics in DM.

First, it is clear that clitics need more in their placement instructions than just Lowering and Local Dislocation as proposed by Embick and Noyer (2001). Clitics need to be stipulated to lean left (enclitics) or right (proclitics). In addition, clitics also need to clarify what is and is not a potential host, resulting in three types of information that must be specified in placement instructions. A second implication is that the clitic status of morphemes can be referenced as such. This appears to be in direct conflict with Harley and Noyer's 1999

proposal that clitic status is not a primitive in DM. Just as we need a diacritic for 'lean left' versus 'lean right', we also need one to reference the clitic status of morphemes.

If clitic status is a primitive in the grammar, this raises questions about the cross-linguistic status of 'clitics' as this label often refers to a diverse class of phenomena. However, the polar question clitic in SJA Mam may not be the only instance of a clitic that alters its typical placement to avoid attaching to other clitics (Marie-Luise Popp, p.c. regarding Kuna), warranting further research on the topic. Another question concerns language internal labeling: if one clitic must come with a clitic diacritic in a language, must all clitics have the same diacritic, or can there be language internal clitic diversity? Future work on modeling clitics in DM requires understanding the variation within the class of clitics, as well as how clitics differ from other types of morphemes.

References

Aissen, Judith L. 1992. Topic and focus in Mayan. *Language* 68:43–80.

Anderson, Stephen R. 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language* 68–98.

Carolan, Elizabeth, Marc Pomerleau, and Eve-Marie Gendron-Pontbriand. 2015. An exploration of tense in Chuj. *Revue de l'édition 2014 de VocUM* .

Clemens, Lauren, and Jessica Coon. 2018. Deriving verb-initial word order in Mayan. *Language* 94:237–280.

Coon, Jessica. 2016. Mayan Morphosyntax. *Language and Linguistics Compass* 10:515–550.

Coon, Jessica, and Elizabeth Carolan. 2017. Nominalizations and the structure of progressives in Chuj Mayan. *Glossa* 2:1.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

England, Nora C. 1983a. Ergativity in Mamean (Mayan) languages. *International Journal of American Linguistics* 49:1–19.

England, Nora C. 1983b. A grammar of Mam, a Mayan language. Austin. *University of Texas Press.* 57:446–86.

Halle, Moriss, and Alec Marantz. 1993. *Distributed morphology and the pieces of inflection*. The View from Building 20. Hale, K. & SJ Keyser (eds.) MIT Press: Cambridge, MA.

Harley, Heidi, and Rolf Noyer. 1999. Distributed morphology. *Glot International* 4:3–9.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, 281–337. Springer.

Tessa Scott
tessa_scott@berkeley.edu