Wh-questions and predication in Mayan

[Accepted at *Journal of Linguistics*]

Justin Royer
Département de linguistique et de traduction
Université de Montréal
justin.royer@umontreal.ca

Cristina Buenrostro
Instituto de Investigaciones Antropológicas
Universidad Nacional Autónoma de México
cbuenrostro@iia.unam.mx

Rodrigo Ranero
Department of Spanish & Portuguese
University of California, Los Angeles
ranero@ucla.edu

June 2, 2026

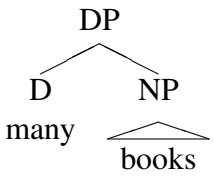
Abstract

This paper examines the syntax of quantificational expressions in Chuj, a Mayan language spoken primarily in Guatemala and Mexico. Building on Royer, Buenrostro, and Jenks (2025a), we argue that Chuj quantifiers fall into two distinct syntactic categories. Some quantifiers function exclusively as *determiners*, while others belong to the well-established category of *nonverbal predicates* (Grinevald and Peake 2012, Coon 2016, Armstrong 2017, Mateo Toledo 2023). This syntactic split is supported by a range of diagnostics, which we systematically apply to another type of quantifier: *wh*-items. The results indicate that Chuj *wh*-items should be treated exclusively as nonverbal predicates, leading us to reanalyze *wh*-questions as constructions that necessarily involve pseudoclefts. Our novel approach challenges the prevailing view in the Mayanist literature, which has treated *wh*-items as components of the extended nominal domain, and supports a less common one instead (Zavala 1992, Tonhauser 2003, 2007). We demonstrate that our proposal also derives three (seemingly idiosyncratic) properties of Mayan *wh*-questions in a unified and theoretically appealing way: (i) the apparent ban on *wh-in-situ* (Caponigro et al. 2021, Coon et al. 2021), (ii) the ban on multiple *wh*-questions (Caponigro et al. 2021, Coon et al. 2021), and (iii) the phenomenon known as "pied-piping with inversion" (Smith-Stark 1988, Aissen 1996).

1 Introduction

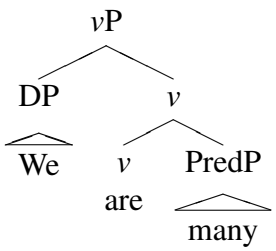
Quantificational items, under standard textbook descriptions, are often regarded as belonging to the extended nominal domain, heading either a Determiner Phrase (DP), Number Phrase (NumP) or Quantifier Phrase (QP) (see e.g., Barwise and Cooper 1981, Dowty et al. 1981, Chierchia and McConnell-Ginet 1990, Heim and Kratzer 1998, Matthewson 2001, Alexiadou et al. 2007, Leu 2015, Jenks 2025). This is a common analysis for some uses of *many* in English. The quantificational expression in (1a) could be represented, in a simplified fashion, with the syntax in (1b).

(1) a. Lou bought [$_{DP}$ **many** books].

b. 

While quantificational expressions are sometimes syntactically instantiated as functional elements of the extended nominal domain, this is not the only possibility. Quantifiers like *many* in English also have "predicative" uses, as in (2a). We could represent such uses (simplifying again) with the syntax in (2b).

(2) a. We are [$_{PredP}$ **many**].

b. 

In some languages, including those that are indigenous to the Americas, quantifiers are *exclusively* expressed via predicative structures such as (2b). This has been discussed for a range of languages spoken in North America (Baker 1995, 1997, Faltz 1995, Jelinek 1995, Matthewson 1998, Lee 2008, Davis and Matthewson 2019). In these languages, one would express the thought in (1a) via a *pseudocleft* paraphrasable as '*The books that Lou bought were many*'. By pseudocleft, we mean a predicational clause where the *wh*-phrase is a predicate that takes as its only argument a relativized nominal.[1] As reflected by our paraphrase, *many* is syntactically employed as a predicate in these languages and the event information (buying a book) is embedded in a relative clause modifying the subject of the predicational clause.

Set against this backdrop, our paper has two overarching goals. The first goal is to show that Chuj possesses a set of quantificational expressions, which, as in other languages of the Americas, are exclusively instantiated as predicates. Specifically, expanding on diagnostics first established in

[1]We are building on work in the Austronesianist literature for our definition; see e.g., Polinsky and Potsdam 2011:108-109. At the end, we will comment briefly on the Austronesian/Mayan parallel regarding our proposal for Chuj. Note that there is an inconsistency in the literature regarding the scope of the term pseudocleft—for instance, authors like Lambrecht (2001) reserve the term for constructions involving subjects that consist of *wh*-free-relatives (e.g., *What you bought is food.*).

Royer, Buenrostro, and Jenks 2025a, we will show that Chuj features separate vocabulary items for "determiner" (Det-Q) versus "predicative" (Pred-Q) quantifiers. The Pred-Qs belong to the well established category of *nonverbal predicates* in Mayan (Grinevald and Peake 2012, Coon 2016, Armstrong 2017, Mateo Toledo 2023). To wrap up this first goal of our paper, we will sketch a basic syntactic analysis of both quantifier types.

The second goal, guided by the results of the first, is to turn our sight to another type of quantifier in the language, namely *wh*-expressions. Having established that non-interrogative quantifiers in Chuj can be instantiated as either Det-Qs or Pred-Qs, with clear consequences for their syntactic distribution, our main research question will be the following: to which quantifier type do *wh*-quantifiers belong? After showing that *wh*-items check the diagnostics for Pred-Qs, we will argue that they are consistently instantiated as nonverbal predicates (as suggested by Zavala (1992) for Akatek and by Tonhauser (2003, 2007) for Yucatec Maya).[2] In other words, *wh*-questions always involve predicational structures that approximate, for example, '*What is it that Lou bought?*'.

Our analytical proposal runs counter to the received wisdom on Mayan *wh*-items, which for the most part (except Zavala 1992 and Tonhauser 2003, 2007), has consistently treated them as elements within the extended nominal spine (see for example Smith-Stark 1988, Aissen 1996, Gavruseva 2000, Broadwell 2005, Coon 2009, 2010, Heck 2009, Kotek and Erlewine 2019, Mendes and Ranero 2021, Ranero 2021, Coon et al. 2014, Clemens and Polinsky 2017, Little 2020, Coon et al. 2021, Ewing 2022, Royer 2021, Aissen and Polian 2024, Alonso-Ovalle and Royer 2024, Coon and Vázquez Álvarez 2025). In these previous works, simple *wh*-questions are not derived via pseudoclefts (or any cleft-like structure), but are instead assumed to involve monoclausal syntax with *wh*-movement of an argument to the left periphery. Our approach is thus a novel contribution that could bear many fruits when extended systematically beyond Chuj. What is more, we argue that our proposal goes further by delivering a unified derivation for a set of particularities of *wh*-questions across the Mayan family that are otherwise mysterious. We account for the three following properties in a unified way:

(3) *Three properties of Mayan wh-questions*

a. Postverbal *wh*-items are ill-formed in questions, which has previously been described as a ban on *wh-in-situ* (see Ch. 8-13 in Caponigro et al. 2021);

b. Multiple *wh*-questions are ill-formed (see Ch. 8-13 in Caponigro et al. 2021);

c. In *wh*-questions with "pied-piped" elements, the *wh*-item surfaces in the opposite order as the one expected from its base-generated position, a phenomenon known as "pied-piping with inversion" (Smith-Stark 1988, Aissen 1996, Coon 2009, Ewing 2022)

Analyses that assume that *wh*-expressions are instantiated within DPs have provided disparate solutions to derive each of the three facts above in a piecemeal fashion. The novel syntax of *wh*-questions we will propose, in contrast, can account for all three in one fell swoop. We take this to be an advancement of our understanding of the syntax of Chuj specifically, Mayan languages more broadly, and the possibilities offered by the grammar to express quantification.

The paper is structured as follows. Section 2 begins with relevant background on Chuj. Section 3 then turns to non-interrogative quantifiers, providing evidence that Chuj distinguishes between

[2] An early precursor for our proposal linking *wh*-items with nonverbal predicates is Zavala (1992: 264), who notes regarding Akateko: "*Tomando en cuenta su comportamiento sintáctico, tal parece ser que algunas partículas interrogativas funcionan como predicados no verbales que anteceden a un predicado complejo*".

Det-Qs versus Pred-Qs. Section 4 then provides an analysis of *wh*-items, arguing that they are consistently instantiated as predicative quantifiers; this will have implications for the syntax of *wh*-questions, which we analyze in this section as well. Finally, section 5 concludes with a short discussion of the implications of this approach for other areas of Mayan grammar and beyond.

2 Chuj background

We will discuss the grammar of Chuj, a Mayan language which belongs to the Q'anjob'alan subbranch of the family (Kaufman 1974, Law 2014). Chuj is primarily spoken in Guatemala and Mexico, but also across diaspora communities in other parts of Mexico, as well as across North America (Maxwell 1993, Buenrostro 2013, Kaplan 2021). Compiling most recent censuses in Guatemala (INE 2018) and Mexico (INEGI 2020), the language has approximately 60,000 speakers in total. There are two main dialects of Chuj: San Mateo Ixtatán and San Sebastián Coatán (Maxwell 1982, García Pablo and Domingo Pascual 2007, Royer et al. 2025b, Ranero et al. 2026). Data in this paper come from theoretically-informed fieldwork (see Matthewson 2004, Bochnak and Matthewson 2020), complemented with corpus analysis (retrieved primarily from transcribed texts in Mateo Pedro and Coon 2018). Although the data presented in this paper come from the San Mateo Ixtatán dialect, the facts described here have been verified across both dialects.

Like many other Mayan languages (England 2001, Coon 2016, Aissen et al. 2017), Chuj is a predicate-initial, ergative-absolutive, head marking language.[3] Crucially for the purposes of this paper, we will follow previous work on Mayan in drawing a distinction among two main types of predicates, based on their ability to occur with tense-aspect marking (see e.g., Grinevald and Peake 2012, Coon 2016 for overviews).

"Verbal predicates" are predicates that generally appear *with* tense-aspect marking. An example is provided in (4).

(4) Ol-ach-w-anht-ej.
PROSP-B2S-A1S-cure-DTV
'I will cure you.'

Following the Mayanist tradition, we refer to ergative/genitive morphemes in the verb stem as "Set A" and we refer to absolutive morphemes as "Set B". The following table lays out these morphemes for San Mateo Ixtatán Chuj (see Royer et al. 2025b):

(5) *Set A and Set B morphemes in San Mateo Ixtatán Chuj*

	Set A (ergative/genitive)		**Set B (absolutive)**
	___C	___V	
1SG	*(h)in-*	*w-*	*(h)in*
2SG	*(h)a-*	*(h)-*	*(h)ach*
3SG	*s-*	*y-*	Ø
1PL	*ko-*	*k-*	*(h)onh*
2PL	*(h)e-*	*(h)ey-*	*(h)ex*
3PL	*s-*	*y-*	Ø

[3]For more extensive information about Chuj, including grammars and grammatical sketches, see e.g., Hopkins 1967, García Pablo and Domingo Pascual 2007, Buenrostro 2021, Royer et al. 2022, Royer et al. 2025b).

When overt third person arguments are present with verbal predicates, the word order is VOS for San Mateo Ixtatán Chuj, as shown in (6).[4] Note that third person Set B has no phonetic realization, and is therefore not represented in glosses.

(6) Ix-s-chi' [$_{Obj}$ nok' kaxlanh] [$_{Subj}$ nok' miston].
PFV-A3-eat CLF chicken CLF cat
'The cat ate the chicken.' (García Pablo and Domingo Pascual 2007: 232)

So-called "nonverbal predicates" are the second type, which will play a central role in this paper (see Grinevald and Peake 2012, Coon 2016, Armstrong 2017, Mateo Toledo 2023 on nonverbal predication across Mayan). These predicates never occur with the tense-aspect markers found on verbal predicates, and they are necessarily intransitive, meaning that they combine with at most one argument. They can be adjectival (7a), nominal (7b), positional (7c), and existential (7d), and generally translate as clauses employing copulas in languages like English. Nonverbal predicates cannot bear any functional material, such as classifiers, quantifiers, or demonstratives, material that is otherwise required when a nominal expression is used as an argument (see Royer 2022b, Royer and Buenrostro 2025a). Note, however, that there is never an overt copula in Chuj predicational clauses.[5]

(7) a. **Wach'** ix.
well CLF.she
'She is well.'

b. **Munlajum** ix.
worker CLF.she
'She is a worker.'

c. **Pitz-an** ix.
awake-STAT CLF.she
'She is awake.'

d. **Ay** jun son t'atik.
EXT one marimba here
'There is a marimba here.'

Zooming in now on the behavior of nominal arguments, two facts will be especially relevant for the purposes of this paper. First, while Chuj shows basic predicate-initial word order, arguments can appear preverbally to indicate contrastive or new information focus (see Royer et al. 2025a) or to indicate topicality. In both cases, they appear with the marker *ha*, which we gloss as a 'preverbal' (PV) marker (see section 5.1). Only topics, however, trigger obligatory resumption in the postverbal argument position. Focused and topicalized objects are provided in (8) for illustration.[6]

(8) a. [**Ha jun te' onh**] ix-s-man ix Malin.
PV INDF CLF avocado PFV-A3-buy CLF Malin
≈ 'It's an avocado that Malin bought.' (focused object)

[4] San Sebastián Coatán Chuj exhibits basic VSO order instead; see Maxwell 1976; Royer et al. 2025b.

[5] An anonymous reviewer asks if the elements that we exemplify being used as adjectival predicates and positional predicates can also be used as nominal modifiers in Chuj. The answer to this is most likely no: positionals never seem to occur as prenominal modifiers and only a small subset of adjectives (mostly color and size terms) can be prenominal modifiers; see Coon 2018, which argues that only prenominal modifiers are actual adjectives. Adjectival and positional roots appearing to the right of a noun are nonverbal predicates inside relative clauses (Coon 2018).

[6] *Wh*-questions, focus, and relativization in Chuj trigger the appearance of "agent focus" morphology on verb stems, which correlates with the absence of Set A marking (for Mayan more broadly, see Stiebels 2006, Coon et al. 2014, Aissen 2017b, Douglas et al. 2017, Coon et al. 2021, Royer 2025, Royer and Coon 2025). We set this aside until the discussion section at the end.

b. [**Ha te' onh**] ix-s-man **te'** ix Malin.
PV CLF avocado PFV-A3-buy CLF.PRON CLF Malin
≈ 'As for the avocado, Malin bought it.' (topicalized object)

Second, it will be important to keep in mind how relative clauses are instantiated. When main arguments (subjects/objects) are relativized, they appear in a preverbal position and there is no obligatory relative marker; an example of an object relative clause is provided in (9) for illustration. As indicated with the brackets, we will assume a head external analysis of relative clauses in this work (Partee 1975, Chomsky 1977, Jackendoff 1977; see De Vries 2018):

(9) Ol-in-man [te' pat [$_{RC}$ ix-a-sikl-ej]].
PROSP-A1S-buy CLF house PFV-A2S-choose-DTV
'I'll buy the house that you chose.'

An example of a non-argument relative clause is provided in (10). Importantly, whereas relativization of core arguments does not necessitate the presence of a relativizer (9), obliques require such an element (e.g., *b'aj* introducing the relative clause targeting a locative in (10)).

(10) Ol-in-man [te' pat [$_{RC}$ ***(b'aj)** ix-ach-kan-i]].
PROSP-A1S-buy CLF house REL.OBL PFV-B2S-stay-IV
'I'll buy the house where you stayed.'

With this background in mind, we are ready to delve into the topic of quantifiers in Chuj.

3 Two types of quantifiers in Chuj

Let us first lay out the core insight developed in this section through an empirical illustration. There are two ways to express in Chuj what would be translated as "Malin bought **many** books":

(11) 'Malin bought many books'

a. Ix-s-man **jantak** ch'anh libro ix Malin.
PFV-A3-buy MANY$_{DET}$ CLF book CLF Malin

b. **Tzijtum** ch'anh libro ix-s-man ix Malin.
MANY$_{PRED}$ CLF book PFV-A3-buy CLF Malin

Note that each version differs in its use of one of two quantifiers, which we have bolded: *jantak* on the one hand and *tzijtum* on the other.

The former can appear postverbally in the canonical argument position, as in (11a). However, it may also appear in the preverbal focus position, like any other DP (see also section 3.1.3 for an example of *jantak* within a topic). In such a position, the marker *ha* is observed:

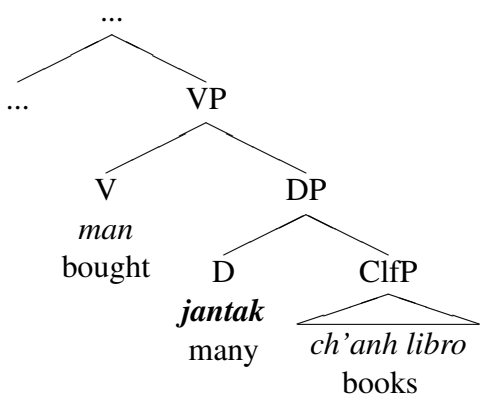
(12) Ha **jantak** ch'anh libro ix-s-man ix Malin.
PV MANY$_{DET}$ CLF book PFV-A3-buy CLF Malin
'It is many books that Malin bought.' (focused object)

Whereas the placement of *jantak* is variable in the way described, the placement of *tzijtum* is limited. It can only appear sentence-initially *without* the presence of *ha*, as in (11b). Furthermore, this quantifier can never appear postverbally, regardless of the information structure status of its DP associate:

(13) * Ix-s-man **tzijtum** ch'anh libro ix Malin.
PFV-A3-buy $\text{MANY}_{\text{PRED}}$ CLF book CLF Malin
Intended: 'Malin bought many books.'

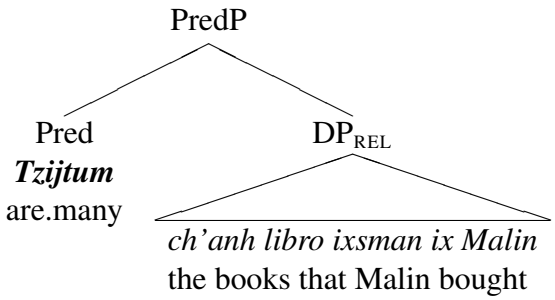
Royer, Buenrostro, and Jenks (2025a) argue that the basic distributional contrast between the two quantifiers can be derived straightforwardly by proposing that each quantifier belongs to a different syntactic category. In particular, quantifiers like *jantak* are "determiner quantifiers" (Det-Q). By this, they mean that they are instantiated somewhere within the extended nominal domain. For simplicity, we assume like Royer et al. (2025a) that Det-Qs sit in the head of the DP (though they could potentially be instantiated higher, in e.g., the head of a QP; see Matthewson 2001 for that type of structure):

(14) *Syntax of* (11a) (*Literal translation*: 'Malin bought *many* books.')



Quantifiers like *tzijtum*, on the other hand, are "predicative quantifiers" (Pred-Qs). Pred-Qs pattern with nonverbal predicates in being able to serve as the main predicate of a sentence, all while never occurring with tense-aspect marking. Examples like (11b) therefore necessarily involve a subject with a relative clause whose verb encodes information about the event described by the sentence (in this case, an event of buying).

(15) *Syntax of* (11b) (*Literal translation*: 'The books that Malin bought *are many*.')



Notice that the proposed structures immediately capture the distributional facts discussed in (11a) and (11b): while Det-Qs are predicted to appear in positions expected of arguments, including the canonical postverbal position, Pred-Qs are predicted to appear in sentence-initial position—namely, the expected position for nonverbal predicates (see (7) above).

In addition to capturing the basic distribution of the quantifiers, this analytical proposal makes several empirical predictions, which the following subsections demonstrate are all borne out. We start by laying out the predicted behavior of a quantifier whose status is that of a determiner (§3.1), showing that *jantak* fulfills all of them, whereas *tzijtum* does not. We then lay out the predicted behavior of a quantifier whose status is predicative (§3.2). In that subsection, we show that *tzijtum* fulfills the predicted behavior, whereas *jantak* does not.

We preview the results below, pointing to the specific subsection where each diagnostic is illustrated. While some of the diagnostics were already discussed in Royer et al. 2025a (see also Royer and Buenrostro 2025b), some of them are novel to this paper. We also foreshadow that in section 4, we will show that *wh*-expressions check all the diagnostics for predicative status, leading to our proposal that they too are Pred-Qs.

(16) *Diagnostics for Det vs Pred status and corresponding sections*

Testing for:	Diagnostic	*Jantak*	*Tzijtum*	§
Det status	Can be part of the comp. of a preposition	✓	✗	3.1.1
	Can be part of the possessor of a DP	✓	✗	3.1.2
	Can be part of a topicalized DP	✓	✗	3.1.3
Pred status	Obligatory predicate-initial syntax	✗	✓	3.2
	Can be the pred. of predicational clause	✗	✓	3.2.1
	Shows signs of relativization	✗	✓	3.2.2
	Can be the pred. of possessive 'have'	✗	✓	3.2.3
	Can be a secondary predicate	✗	✓	3.2.4
	Appearance with predicate modifiers	✗	✓	3.2.5

While we will illustrate the Det-Q vs. Pred-Q contrast by comparing the distribution of *jantak* with that of *tzijtum* specifically, these quantifiers are not alone within their respective categories. The following table provides a list of quantifiers by category in San Mateo Ixtatán Chuj.

(17) *Inventory of quantifiers over entities in San Mateo Ixtatán Chuj*

Det-quantifiers	
jun	'one' / singular indefinite determiner
juntzanh	'some' / plural indefinite determiner
junjun	'some/each'
tzun	'one' (affective/diminutive)
jantak	'many/all'
jab'	'little amount of' (for mass nouns only)
chab'ox-#.CLF	'a few'
jay-#.CLF	'few'
Pred-quantifiers	
tzijtum	'many'
pim	'many' (also means 'thick')
jantaknhej	'many'
ma(nh)jantak(ok)	'many'
niwan	'many' (for mass nouns; also means 'big')
wal	'many'
kennhej	'few'
junjunnhej	'few'
chekelnhej	'few'
kenan	'few'
jay-#.CLF-*nhej*	'few'
jab'tzin	'few' (for mass nouns)
jab'nhej	'few (for mass nouns)'
chab'tzin	'few' (for mass nouns)

Two notes are in order about some of these quantifiers. First, the morphology would suggest that some predicative quantifiers appear to be derived from determiner quantifiers (see some of the quantifiers ending with *-nhej*). We will not attempt to explain this fact here. Second, some quantifiers have ambiguous readings. For example, some existential quantifiers seem to lack an 'upper bound', meaning that they can be used equally in contexts in which a subset of relevant individuals satisfy the existential claim triggered by the quantifier, and also in contexts in which each or all individuals satisfy the existential claim (see Alonso-Ovalle and Royer 2024, example (77) and (78)). For example, *jantak* can mean 'many' or 'all' depending on the context, and importantly for the presentation of diagnostics to come, it can also be used to convey 'how many' in questions. Regarding the latter use, we will show in section 4 that the interrogative quantifier patterns exclusively like a nonverbal predicate, and never as a determiner. For now, however, we will only focus on uses of *jantak* that translate as 'many', and we will show that under such uses, *jantak* is exclusively instantiated as a determiner.

3.1 Predictions for determiner status

The proposal we laid out makes predictions for the behavior of a quantifier depending on its categorial status in the syntax. The following predictions are made if the quantifier is a determiner:

(18) *Three predictions if a quantifier is a determiner (Det-Q)*;
It should:

a. be able to appear within the complement of a preposition (§3.1.1)
b. be able to appear within the possessor of a DP (§3.1.2)
c. be able to appear within a topicalized DP (§3.1.3)

Building on Royer et al. (2025a), we show in the next three subsections that these predictions are fulfilled for *jantak*, but not for *tzijtum*. While not shown here, the remaining quantifiers in table (17) pattern according to their respective category.

3.1.1 Ability to appear in the complement of a preposition

Chuj's sole preposition *t'a* can only combine with nominal expressions (DPs or NPs), and never with verbal or nonverbal predicates. This fact thus predicts that a Det-Q should be readily found within a prepositional phrase (PP) as a complement to the head. As shown below, this prediction is borne out for *jantak*. Note that this is so regardless of the position of the PP: whether the PP is postverbal or preverbal, *jantak* can occur within the complement of *t'a*.[7]

(19) a. Ol-onh-b'at [t'a **jantak** chonhab'].
PROSP-B1P-go PREP MANY_{DET} town
'We will go to many towns.'

b. [T'a **jantak** chonhab'] ol-onh-b'at-ok.
PREP MANY_{DET} town PROSP-B1P-go-IRR
'To many towns, we'll go.'

The predicted distribution is different for Pred-Qs, which, as nonverbal predicates, are expected to never occur in the complement of *t'a*. As shown below, the prediction is borne out for *tzijtum*, regardless of the position of the PP:

(20) a. * Ol-onh-b'at t'a **tzijtum** chonhab'.
PROSP-B1P-go PREP $\text{MANY}_{\text{PRED}}$ town
Intended: 'We will go to many towns.'

b. * T'a **tzijtum** chonhab' ol-onh-b'at-ok.
PREP $\text{MANY}_{\text{PRED}}$ town PROSP-B1P-go-IRR
Intended: 'To many towns, we'll go.'

In section 3.2.2, we delve into the necessary syntax for *tzijtum* to quantify over oblique arguments, which further supports its status as a nonverbal predicate.

[7] Note that when preverbal, PPs in Chuj may but need not co-occur with the 'preverbal' marker *ha*, which is otherwise needed for preverbal arguments. Adding *ha* to the preverbal PPs in (19b) and (20b) changes nothing for the acceptability judgements: (19b) remains well-formed, while (20b) remains ill-formed.

3.1.2 Ability to appear within a possessor

The possessor of a possessed nominal expression is limited to DPs in Chuj. These possessors in possessive DPs are coindexed with Set A (ergative/genitive) morphemes. If *jantak* is indeed a Det-Q, then it should be allowed in the possessor position. This is borne out—again, regardless of the position of the DP (whether preverbal or postverbal):

(21) a. Wojtak [heb' y-et'b'eyum [$_{\text{POSS}}$ **jantak** heb' ix ix]].
I.know PL A3-partner MANY_{DET} PL CLF woman
'I know the partners of many women.'

b. [Ha heb' y-et'b'eyum [$_{\text{POSS}}$ **jantak** heb' ix ix]] wojtak.
PV PL A3-partner MANY_{DET} PL CLF woman I.know
'The partners of many women, I know.

It is important to note that in contrast to other Mayan languages, which have been noted to permit subextraction of possessors out of possessed DPs (Aissen 1979, 1996, Coon 2009, Little 2020; but see Aissen and Polian 2024), Chuj does not generally permit this kind of subextraction.

If *tzijtum* is indeed a nonverbal predicate, it should not be allowed in possessor position. Again, this prediction is borne out, regardless of the position of the possessed phrase:

(22) a. * Wojtak [heb' y-et'b'eyum [$_{\text{POSS}}$ **tzijtum** heb' ix ix]].
I.know PL A3-partner $\text{MANY}_{\text{PRED}}$ PL CLF woman
Intended: 'I know the partners of many women.'

b. * [Ha heb' y-et'b'eyum [$_{\text{POSS}}$ **tzijtum** heb' ix ix]] wojtak.
PV PL A3-partner $\text{MANY}_{\text{PRED}}$ PL CLF woman I.know
Intended: 'The partners of many women, I know.'

To express the intended meanings in (23) using *tzijtum*, the following construction must be employed instead. Here, *tzijtum* can be analyzed as a nonverbal predicate taking a relativized possessed DP as its subject:

(23) **Tzijtum** [heb' y-et'b'eyum [heb' ix ix] [$_{\text{RC}}$ wojtak]].
$\text{MANY}_{\text{PRED}}$ PL A3-partner PL CLF woman I.know
'I know the partners of many women.'
Literally: 'The partners of the women who I know are many.'

3.1.3 Ability to appear in topic position

A third expectation regarding Det-Qs is that they should be able to appear within topicalized DPs. In Chuj, only DPs may be topicalized (see Maxwell 1982, 1987, Bielig 2015, and Royer 2022c on topicalization in Chuj); when topicalized, DPs appear with the marker *ha* and resumption is obligatory. Fulfilling this expectation, the quantifier *jantak* can appear inside a topicalized DP:

(24) [$_{\text{TOP}}$ Ha] **jantak** heb' winh w-et'b'eyum]$_1$ ix-in-y-il heb' winh$_1$.
PV MANY_{DET} PL CLF A1S-friend PFV-B1S-A3-see PL PRON
'As for many of my friends, they saw me.'

On the other hand, the quantifier *tzijtum*, as a nonverbal predicate, should not be able to appear within the topic position. This is indeed borne out:

(25) * [$_{\text{TOP}}$ Ha] **tzijtum** heb' winh w-et'b'eyum]$_1$ ix-in-y-il [heb' winh$_1$].
PV MANY$_{\text{PRED}}$ PL CLF A1S-friend PFV-B1S-A3-see PL PRON
Intended: 'As for many of my friends, they saw me.'

3.1.4 Interim summary

To summarize, we have discussed three diagnostics which lead us to conclude the following: quantifiers like *jantak* have a distribution that is consistent with them being determiners (or at least elements in the extended nominal projection), whereas quantifiers like *tzijtum* have the opposite distribution. This is derived directly if, as proposed at the outset of this section, *jantak* is a Det-Q, while *tzijtum* is a Pred-Q.

In the next section, we will arrive at the same conclusion via a different set of diagnostics that take as their starting point the expected distribution of Pred-Qs: *tzijtum* fulfills the predictions of all such diagnostics, whereas *jantak* does not fulfill any of them.

3.2 Predictions for predicative status

We make the following six predictions for quantifiers in Chuj that are syntactically instantiated as nonverbal predicates:

(26) *Six predictions if a quantifier is a nonverbal predicate*;[8]
It should:

a. not be able to appear postverbally in argument position (examples 11-13)
b. be able to combine with a DP alone in a predicational sentence (§3.2.1)
c. show visible signs of relativization (§3.2.2)
d. be able to participate in possessive 'have' constructions (§3.2.3)
e. be able to serve as a secondary nonverbal predicate (§3.2.4)
f. be able to take morphology reserved for predicates (§3.2.5)

The first prediction above was already shown to hold for *tzijtum*: quantifiers like *tzijtum*, in contrast to those like *jantak*, must generally appear in a sentence-initial position, or at least before verbal predicates (nonverbal predicates can appear, for example, after modifiers like adverbs). Since Chuj is a predicate-initial language, the initial position of quantifiers like *tzijtum* is derived directly from their analysis as nonverbal predicates. Relevant examples are repeated below for convenience:

(27) a. **Tzijtum** ch'anh libro ix-s-man ix Malin.
MANY$_{\text{PRED}}$ CLF book PFV-A3-buy CLF Malin
'Malin bought many books.'

b. * Ix-s-man **tzijtum** ch'anh libro ix Malin.
PFV-A3-buy MANY$_{\text{PRED}}$ CLF book CLF Malin
Intended: 'Malin bought many books.'

[8] A seventh prediction, concerning negation, is addressed in Appendix A.

An anonymous reviewer wonders about our generalization, asking whether *tzijtum* may appear on the surface *non*-initially if the expression that it quantifies over is topicalized. This is indeed true, as shown in (28). Under the relevant topicalization—notice the marker *ha* and the resumption in (28)—the quantifier *tzijtum* appears linearly after a verbal predicate such as *ixwila'*. However, we need not amend our generalization, since the topicalized element is in a peripheral position:

(28) Ha heb' winh ix-w-il-a', **tzijtum** heb' winh.
PV PL CLF PFV-A1S-see-TV many$_{PRED}$ PL PRON
'The ones I saw, they were many.'

Expanding on Royer et al. (2025a), we show in the following subsections that the rest of the predictions in (26) are fulfilled for *tzijtum*, but not for *jantak*. Again, while not shown here, the quantifiers in table (17) pattern according to their indicated category.

3.2.1 Ability to combine with a DP alone in a predicational sentence

If a quantifier is a nonverbal predicate, then it should be possible to combine it with a DP to form what would translate as a predicational copular clause in languages like English. This is indeed the case for *tzijtum*: combining it alone with a DP delivers such a sentence. The following example, taken from a text, illustrates this well:

(29) a. **Tzijtum** heb' ix, tz-y-al chi' [...]?
MANY$_{PRED}$ PL CLF IPFV-A3-say DEIX
'So, there were many (women), let's say [...]?'

b. Hi, **tzijtum** heb' ix.
yes, MANY$_{PRED}$ PL PRON
'Yes, they were many.' (txt, CD300715)

Moreover, and as expected, this is not the case with *jantak*. When intended to be used as the quantifier 'many', *jantak* cannot be used predicatively. We show a minimal pair between *tzijtum* and *jantak* below, this time from elicitation. As shown, while *tzijtum* can be used predicatively (30a), *jantak* cannot (30b).

(30) **Context**: Some friends of mine are visiting, and you invite us over for lunch. I want to make sure it's okay for us to come, since we're a lot of people.

a. Tom yel to tz-a-nib'-ej to tz-onh-b'at-i? Yujto **tzijtum** honh.
YNQ certain C IPFV-A2S-desire-DTV C IPFV-B1P-go-IV because MANY$_{PRED}$ B1P
'Are you sure you want us to come? Because we're many.'

b. * Tom yel to tz-a-nib'-ej to tz-onh-b'at-i? Yujto **jantak** honh.
YNQ certain C IPFV-A2S-desire-DTV C IPFV-B1P-go-IV because MANY$_{DET}$ B1P
Intended: 'Are you sure you want us to come? Because we're many.'

3.2.2 Visible signs of relativization

A second diagnostic concerns visible signs of relativization. Recall that building sentences with Pred-Qs often requires a relativization component—for example, the equivalent of '*Malin bought*

many books' translates more literally to '*The books that Malin bought are many*.' Recall as well from section 2 that argument relativization does not exhibit obligatory marking of any kind (9), but that the relativization of oblique arguments does (10): namely, the use of the relative pronoun *b'aj*.

Therefore, if *tzijtum* is indeed a nonverbal predicate, we predict that whenever it quantifies over an oblique argument, we will observe the relative pronoun *b'aj*. As shown below, this prediction is borne out:

(31) **Tzijtum** chonhab' *(b'aj) ol-onh-b'at-ok.
MANY$_{\text{PRED}}$ village REL.OBL PROSP-B1P-go-IRR
'We'll go to many villages.'
Literally: 'The villages where we'll go are many.'

In contrast, whenever *jantak* quantifies over an oblique argument to deliver a meaning such as 'to many towns', it must appear within a prepositional phrase (see example (19) for relevant data).

The data in (31) in conjunction with those in section 3.1.1 provide further support for our claim that quantifiers like *tzijtum* can be nonverbal predicates. Crucially, however, (31) supports the even stronger view that they *must* be so. Otherwise, we could not explain the obligatory presence of the relative pronoun in sentences where *tzijtum* quantifies over an oblique argument.

3.2.3 Ability to participate in possessive 'have' sentences

In Mayan languages, possessive 'have' constructions are formed by combining an existential predicate with a possessed DP (Freeze 1992, Coon 2016). An example with the basic existential predicate *ay* 'EXT' is provided below:

(32) **Ay** heb' winh h-uninal.
EXT PL CLF A2S-son
'You have sons.'
Literally: 'There are your sons.'

Now, since Pred-Qs also involve existential quantification, we might expect them to pattern with the basic existential predicate to form sentences that would translate with possessive 'have' in English. As shown below, this is borne out. Quantifiers like *tzijtum* can combine with a possessed DP to serve as the sole predicate of a possessive 'have' construction—in other words, *tzijtum* is in the same natural class as existential *ay*.

(33) **Tzijtum** heb' winh h-uninal.
MANY$_{\text{PRED}}$ PL CLF A2S-child
'You have many sons.'
Literally: 'Your sons are many'. (txt, CM300715)

We predict a different behavior for the quantifier *jantak*, when used as a Det-Q. Indeed, combining *jantak* with a possessed DP does not deliver the intended translation in (34a).[9] Instead, to utter the equivalent of (33), *jantak* must co-occur with *ay* 'EXT' (34b), as expected if *jantak* is a Det-Q.

[9]Note that *jantak* may be employed as a nonverbal predicate in such constructions, but its only possible interpretation is that of a *wh*-item; i.e., the sentence in (34a) can mean '*How many sons do you have?*'. In section 4, we analyze such uses of *jantak* as Pred-Qs. Crucially for the purposes of this section, though, *jantak* in (34a) cannot mean 'many'.

(34) a. * **Jantak** heb' winh h-uninal.
MANY$_{DET}$ PL CLF A2S-son
Intended: 'You have many sons.' (alternative reading available)

b. Ay **jantak** heb' winh h-uninal.
EXT MANY$_{DET}$ PL CLF A2S-child
'You have many sons.'

For completeness, also note that *ay* and *tzijtum* may not co-occur (no matter the order of *ay* and *tzijtum*) to deliver a possessive 'have' sentence:

(35) a. * **Tzijtum** ay heb' winh h-uninal.
MANY$_{PRED}$ EXT PL CLF A2S-child

b. * Ay **tzijtum** heb' winh h-uninal.
EXT MANY$_{PRED}$ PL CLF A2S-child
Intended: 'You have many sons.'

3.2.4 Ability to serve as a secondary predicate

This diagnostic is specific to the Q'anjob'alan branch of the Mayan languages and their complex secondary predication constructions (see e.g., Mateo Toledo 2012, 2023). These constructions involve the combination of a secondary nonverbal predicate (e.g., an adjective), with an aspectless clause exhibiting nominative-accusative alignment (i.e., an alignment where the subject of an intransitive verb is co-indexed with Set A):

(36) [$_{PRED}$ **Junk'olal**] [$_{VP}$ y-ek' heb' paxyalwum] t'atik.
content A3-pass.by PL visitor here
'The visitors pass by here content.'

If Pred-Qs are nonverbal predicates, we might expect them to be able to serve as secondary nonverbal predicates in constructions similar to the one in (36). As shown below, this expectation proves correct: *tzijtum* can serve as a secondary nonverbal predicate.

(37) [$_{PRED}$ **Tzijtum**] [$_{VP}$ y-ek' heb' paxyalwum] t'atik.
MANY$_{PRED}$ A3-pass.by PL visitor here
'Many visitors pass by here.'
Literally: 'The visitors pass by here (and they are) many.'

As a Det-Q, we do not expect *jantak* to be able to serve as a secondary nonverbal predicate. This is also correct: *jantak* cannot convey a meaning along the lines of the one expressed in (37).[10]

(38) * **Jantak** y-ek' heb' paxyalwum t'atik.
MANY$_{DET}$ A3-pass.by PL visitor here
Intended: 'Many visitors pass by here.' (alternative reading available)

[10]Again, note that when *jantak* is used as a *wh*-item, it behaves like a nonverbal predicate. The crucial point here is that it cannot be used as a nonverbal predicate when its intended meaning is that of 'many'; see section 4 for more details.

3.2.5 Ability to combine with morphology reserved for nonverbal predicates

Our last diagnostic concerns morphology that appears exclusively on predicates, either nonverbal or verbal. This includes morphemes like *-(h)am(a')*, *-(h)amlaj*, *-xom*, and *-xomlaj*, which give rise to various subtle epistemic modal interpretations whose semantics is yet to be completely understood.

We exemplify this with the morpheme *-(h)amlaj* below. When used with nonverbal predicates, *-(h)amlaj* conveys something similar to speaker uncertainty (i.e., it is a possibility modal). As shown in (39-42), this morpheme can attach to all sorts of nonverbal predicates, including nominal, adjectival, existential, and positional predicates. It can never, on the other hand, appear on any morpheme of the extended nominal domain of arguments.

(39) **Doktor**-amlaj ix.
doctor-MODAL she
'Perhaps she's a doctor.'

(40) **Wach'**-amlaj ix.
well-MODAL she
'Perhaps she's well.'

(41) **Ay**-amlaj heb' anima' t'atik.
EXT-MODAL PL person here
'Perhaps there are people here.'

(42) **Pitz-an**-amlaj ix.
awake-STAT-MODAL she
'Perhaps she's awake.'

Since this morphology can only ever appear on predicates, and never on nouns in argument position or on any other element of the extended nominal domain (classifiers, determiners, quantifiers, demonstratives), we make a prediction: quantifiers like *tzijtum* should be able to host morphemes like *-amlaj*, but quantifiers like *jantak* (when used as 'many') should not.[11]

As shown below, this prediction is borne out on both counts, irrespective of the positioning of *jantak* in relation to the verb.

(43) **Tzijtum**-amlaj heb' anima' ol-chanhalw-ok.
$\text{MANY}_{\text{PRED}}$-MODAL PL person PROSP-dance-IRR
'Perhaps many people will dance.'
Literally: 'The people who will dance are perhaps many.'

(44) a. * Ol-chanhalw-ok **jantak**-amlaj heb' anima'.
PROSP-dance-IRR MANY_{DET}-MODAL PL person
Intended: 'Perhaps many people will dance.'

b. * **Jantak**-amlaj heb' anima' ol-chanhalw-ok
MANY_{DET}-MODAL PL person PROSP-dance-IRR
Intended: 'Perhaps many people will dance.' (alternative reading possible)

[11] A reviewer asks if *-amlaj* could be instead analyzed as a second position clitic, in order to explain the unacceptability of (44a). The answer is negative. Example (i) shows that *-amlaj* can appear fourth in the embedded 'because'-clause.

(i) Tz-h-il ha-b'a, yujto tob' **tzijtum**-amlaj heb' anima' ol-jaw-ok t'a ha-k'inh.
PFV-A2S-see A2S-self because REP many-MODAL CLF person PROSP-come-IRR PREP A2S-party
'Careful, because there apparently might be many people that will come to your party.'

3.3 Summary and analysis

To summarize the empirical takeaways of this section, we have expanded on diagnostics originally proposed in Royer et al. 2025a to show that Chuj distinguishes between two types of quantifiers. While quantifiers like *jantak* fulfill all diagnostics for determiner status (they are Det-Qs), quantifiers like *tzijtum* fulfill all diagnostics for nonverbal predicate status (they are Pred-Qs). Crucially, this is different from quantifiers like 'many' in languages such as English, which, as shown at the outset of this paper (1a-2a), exhibit both determiner and predicative uses:

(45) a. Lou bought many books.
b. We are many.

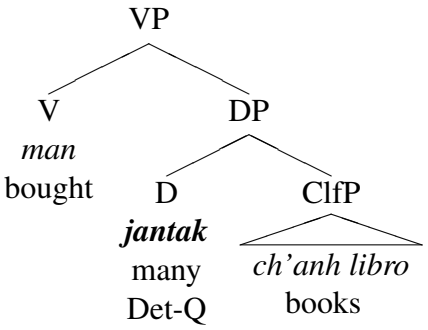
Chuj, instead, employs separate lexical items for each quantifier type: e.g., *tzijtum* is only ever instantiated as a nonverbal predicate, whereas *jantak*—when used to convey 'many'—is instantiated as a determiner. For a summary of diagnostics, see table (16) above.

Following Royer et al. (2025a), we thus conclude that there are two different types of quantifiers in Chuj, requiring distinct syntactic analyses. For Det-Qs, we assume that the quantifier occupies the head of the DP. This builds on Royer (2022b), who locates indefinite determiners in that position. Classifiers are then located in a "Classifier Phrase (ClfP)", which merges with the NP.

(46) *Syntax of Det-Qs*

a. Ix-s-man **jantak** ch'anh libro ix Malin.
PFV-A3-buy MANY_{DET} CLF book CLF Malin
'Malin bought many books.' (repeated from (11a))

b.

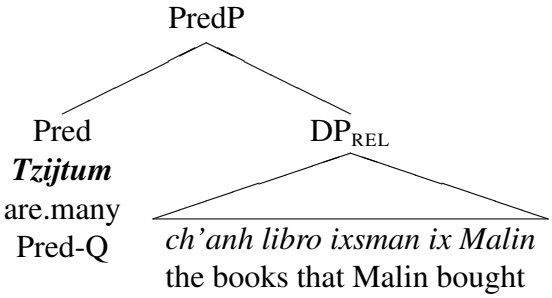


For Pred-Qs, we assume that they occupy the head of a PredP, taking a DP subject as their sole (internal) argument. There is no need for a copula: a quantifier such as *tzijtum* encodes predication itself.

(47) *Syntax of Pred-Qs*

a. **Tzijtum** ch'anh libro ix-s-man ix Malin.
MANY$_{\text{PRED}}$ CLF book PFV-A3-buy CLF Malin
'Malin bought many books.' (repeated from (11b))

b.



We further assume that the PredP head in (47) can also be filled by nominal, adjectival, positional, and existential nonverbal predicates such as those displayed in (7a-7d).

Finally, note that while our treatment of the subjects of predicational clauses as internal arguments follows Coon 2014 and Royer et al. 2025a (instead of external arguments merged in the specifier of PredP; see Bowers 1993, Baker 2003, den Dikken 2006, Lyskawa and Ranero 2022), other approaches to predicational clauses in Mayan like Armstrong 2017 or Coon and Martinović 2023 would also be compatible with the empirical facts discussed here. All that is needed is for nonverbal predicates in general (including Pred-Qs) to consistently surface in sentence-initial position, be it because they are base-generated there (as in the current approach) or as a result of movement (Coon and Martinović 2023; for predicate initiality beyond Mayan, see papers in Lee-Schoenfeld and Ott 2021).

Having now shown evidence that Chuj distinguishes syntactically among two types of quantifiers, we turn to the second goal of this paper: establishing the syntactic status of *wh*-quantifiers, which we will argue instantiate Pred-Qs.

4 On the predicative nature of *wh*-words in Chuj

The previous section used a constellation of diagnostics to differentiate among two types of non-interrogative quantifiers in Chuj. Based on these, this section now turns to Chuj *wh*-items, arguing that they are consistently and exclusively instantiated as nonverbal predicates (i.e., they are Pred-Qs). This proposal will necessitate a novel analysis for the syntax of questions in Chuj, one in which *wh*-questions involve pseudoclefts. This proposal goes against the *status quo* for Mayan, insofar as previous works on the family have overwhelmingly analyzed *wh*-items as part of DPs. Furthermore, we will argue that our approach offers a unified derivation for several properties of the syntax of *wh*-questions in Mayan that had previously been noted but had not been considered to be connected: (i) an apparent ban on *wh-in-situ* (Caponigro et al. 2021, Coon et al. 2021), (ii) a ban on multiple *wh*-questions (Caponigro et al. 2021, Coon et al. 2021), and (iii) a phenomenon known as "pied-piping with inversion" (Smith-Stark 1988, Aissen 1996, Coon 2009, Ewing 2022).

We begin in section 4.1 with some background on the standard analysis of *wh*-items as part of DPs in Mayan, as well as the analysis of *wh*-question formation in general. This section will also establish the three (seemingly) idiosyncratic properties that *wh*-expressions exhibit. Section 4.2 offers our novel analysis of *wh*-items as Pred-Qs, showing that such an approach provides a

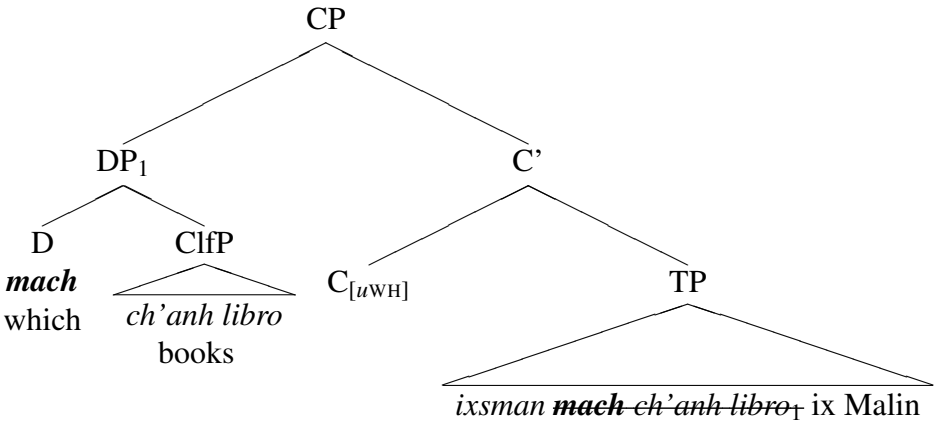
unified derivation for the three properties. Finally, sections 4.3 and 4.4 support the novel approach by showing that *wh*-expressions in Chuj check all diagnostics supporting their analysis as Pred-Qs.

4.1 Background on *wh*-items in Mayan

Previous research into Mayan *wh*-items like 'what', 'which' and 'who', including in Chuj (Kotek and Erlewine 2019, Royer 2021, 2022b), have almost unanimously assumed that they are part of the extended nominal domain (see e.g., Smith-Stark 1988, Aissen 1996, Gavruseva 2000, Broadwell 2005, Coon 2009, 2010, Heck 2009, Douglas et al. 2017, Kotek and Erlewine 2019, Mendes and Ranero 2021, Ranero 2021, Coon et al. 2014, Clemens and Polinsky 2017, Little 2020, Coon et al. 2021, Ewing 2022, Royer 2021, Aissen and Polian 2024, Alonso-Ovalle and Royer 2024, Ranero and Royer 2024, Coon and Vázquez Álvarez 2025). Under these analyses, *wh*-items with an overt nominal associate exhibit the syntax proposed for Det-Qs in this paper, insofar as they head the DP. At the clausal level, *wh*-movement to a clause-peripheral position would account for the word order in (48).

(48) a. **Mach** ch'anh libro ix-s-man ix Malin?
which CLF book PFV-A3-buy CLF Malin
'Which book did Malin buy?'

b. *Possible syntax for* (48a) (to be refuted for Chuj)



Under this kind of approach, *wh*-words exhibit the syntax standardly associated with such items in more widely studied languages, such as English. However, the above-mentioned literature has also established three somewhat surprising observations regarding *wh*-questions in Mayan, suggesting that a different syntax should be explored instead.

The first observation is that *wh*-items in Mayan must always be in sentence-initial position in order for interrogative semantics to arise. The following data exemplify this. The word for 'who/which' in Chuj, *mach*, cannot appear in a postverbal position:

(49) a. **Mach** ix-h-il-a'?
who PFV-A2S-see-TV
'Who did you see?'

b. * Ix-h-il **mach**?
PFV-A2S-see who
Intended: 'You saw *who*?'

Previous works interpreted the pattern in (49) as an indication that *wh-in situ*, even in echo questions, is forbidden, and thus that "focus fronting" in *wh*-questions is obligatory (see Royer 2021 on

Chuj, and Aissen 1996 on Tsotsil; Polian 2013 and Polian and Aissen 2021 on Tseltal; Vázquez Álvarez and Coon 2020 on Ch'ol; Curiel 2017 on Tojol-ab'al; AnderBois and Chan Dzul 2021 on Yucatec Maya; Duncan et al. 2024 on Kaqchikel; Duncan 2010, Velleman 2014 on K'iche'; and Elkins and Brown to appear on Mam).[12] The assumption is thus that *wh*-items are generated in an argument position within a DP, but obligatorily move to the preverbal position. As Coon et al. (2021: 289) put it: "There is broad evidence that the requirement for A'-elements to move is strong in Mayan. Mayan languages prohibit *in-situ wh*-words in interrogative contexts [...]."

The second observation regarding *wh*-items, in Chuj and more broadly in the family, concerns the impossibility for multiple *wh*-questions. In a nutshell, no matter the order of *wh*-items, multiple *wh*-questions are ill-formed (see same references as the one just cited above). This is shown below.

(50) *Intended*: 'Who saw what?'

a. * **Mach** ix-il-an **tas**?
who PFV-see-AF what

b. * **Mach tas** ix-il-an-i?
who what PFV-see-AF-IV

c. * **Tas mach** ix-il-an-i?
what who PFV-see-AF-IV

d. * **Tas** ix-y-il **mach**?
what PFV-see who

Given a DP-based analysis of *wh*-phrases, one could derive the ban on multiple *wh*-questions if there is only one left-peripheral position to which *wh*-phrases may move (as has often been assumed in the literature on Mayan; see e.g., Aissen 1992, Coon et al. 2014, Coon et al. 2021). That is, if focus fronting is obligatory for interrogative meaning to arise (as suggested by data like (49)), and there is only one focus position in the syntax, then multiple *wh*-questions are expected to be impossible.

A final observation regarding Mayan *wh*-questions is what has come to be known as "pied-piping with inversion" (Smith-Stark 1988, Aissen 1996, Coon 2009, Ewing 2022). This phenomenon can be illustrated in Chuj with the relational noun *et'* 'with', which can be characterized as an agreeing preposition, where the complement of the P controls genitive (Set A) agreement (i.e., *y-et'* is decomposable into A3S-with). First consider the data in (51), which shows that postverbal DPs linearly follow the relational noun. As shown in (52), this is in contrast with preverbal *wh*-words, which must linearly *precede* the relational noun (hence the term 'inversion'):

(51) Ol-ach-b'at [yet' **waj Xun**].
PROSP-B2S-go with CLF Xun
'You will go with Xun.

(52) a. **Mach** yet' ol-ach-b'at-ok?
who with PROSP-B2S-go-IRR
'With whom will you go?'
Literally: 'Who with will you go?'

b. * Yet' **mach** ol-ach-b'at-ok?
with who PROSP-B2S-go-IRR
Intended: 'With whom will you go?'

[12]For K'iche' specifically, there is some disagreement. Duncan (2010: 458-462) and Velleman (2014: 73-74) report that *wh-in situ* and multiple *wh*-questions are ungrammatical; Can Pixabaj (2021: 269-270) reports that multiple *wh*-questions are possible (with an echo intepretation) when there is a fronted *wh*-item and a second one *in-situ*.

In work that assumes that *wh*-words are DPs, what must be posited to account for these data is a two-step derivation: (i) an initial movement internal to an oblique/possessive phrase, followed by (ii) movement of the entire phrase (see e.g., Aissen 1996, Coon 2009, Little 2020, Ewing 2022).

In sum, the literature on Mayan languages has long noted certain properties of *wh*-question formation that might seem surprising at first glance. Explaining the ill-formed data in (49) and (50), and the apparent existence of "pied-piping with inversion", all while maintaining a DP analysis of *wh*-expressions, rests on a number of stipulations: (i) the fact that *wh*-fronting is obligatory for interrogative meaning to arise, (ii) the fact that there is only one focus position in the syntax, and (iii) the fact that there is movement internal to oblique/possessive phrases whenever there is A'-extraction of such a phrase. In the next section, we show that a Pred-Q analysis of *wh*-items, in contrast, embodies a way to account for all three restrictions in one fell swoop.

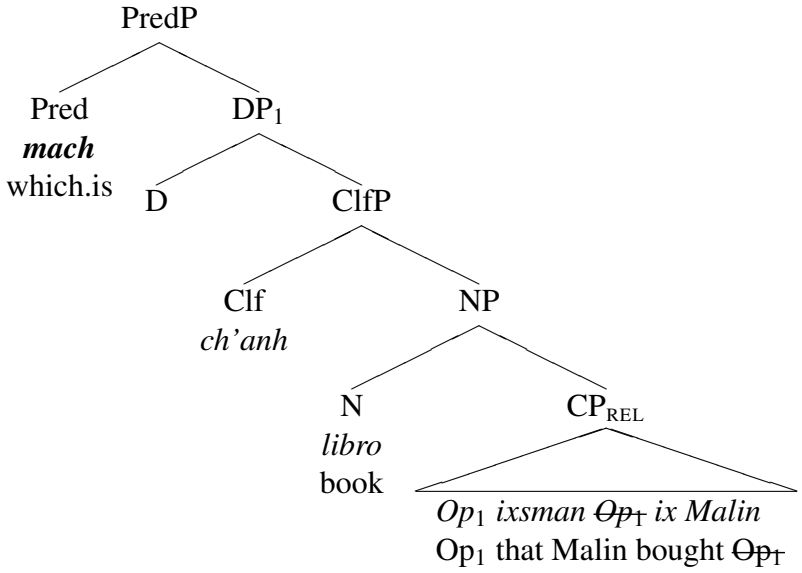
4.2 An alternative: *wh*-items as predicative quantifiers in Chuj

The current paper has established the existence of two types of quantifiers in Chuj: determiner quantifiers (Det-Qs) and predicative quantifiers (Pred-Qs). Since *wh*-expressions are generally regarded as a subtype of existential quantifier (see e.g., Karttunen 1977, Kratzer and Shimoyama 2002, among many others), a plausible approach to Chuj syntax would be to treat *wh*-items consistently as Pred-Qs (as explicitly suggested by Zavala (1992) for Akatek and by Tonhauser, (2003, 2007) for Yucatec Maya; see Verhoeven and Skopeteas 2015 and Gutiérrez-Bravo 2017 for criticisms; see also Ewing (2022), who suggests this analysis for a subset of *wh*-items in Kaqchikel), in contrast with the works cited in section 4.1 that almost unanimously treat them as Det-Qs. We will argue that doing so sheds new light on why *wh*-items exhibit the otherwise three surprising properties that were laid out.

Let us first consider what a *wh*-question would look like in Chuj if the syntax of the *wh*-item were that of a Pred-Q. An example for the same sentence as the one in (48a) is provided in (53a), now bracketed to reflect the alternative analysis we will defend. Again, we are assuming a head-external analysis of relative clauses in the language (see Partee 1975, Chomsky 1977; Jackendoff 1977; see also De Vries 2018). Note, however, that other analyses of relative clauses are equally compatible with our proposal—e.g., head-raising approaches (Schachter 1973, Vergnaud 1974, Bhatt 2002) or those that assume movement inside an internally-headed relative clause (Bianchi 2000, Bodomo and Hiraiwa 2010, Hiraiwa 2017).

(53) a. **Mach** [ch'anh libro [ix-s-man ix Malin]]?
which.is CLF book PFV-A3-buy CLF Malin
Literally: 'Which is the book that Malin bought?'

b.



Under our approach, sentences with *wh*-items belong to a larger class of Pred-Qs, and thereby exhibit the same structure as the one independently needed for quantifiers like *tzijtum*. In particular, the *wh*-item takes a DP subject as its complement, which can be relativized. When relativized, the verb within the relative provides key information about the interrogated event.

Importantly, note that the translation of the example in (53a) showcases a *wh*-item taking an overt nominal complement in English. Translations of questions without nominal complements would thus necessarily map to headless relative clauses in Chuj, which are independently known to exist in the Mayan language family (see references in Caponigro et al. 2021, and Kotek and Erlewine 2019 and Royer 2021 specifically for Chuj). An example, with a schema for the proposed structure, is illustrated below.

(54) a. **Mach** ix-s-man ix Malin?
which.is PFV-A3-buy CLF Malin
'Which did Malin buy?'
Literally: 'Which is it that Malin bought?'

b. [$_{\text{PredP}}$ which.is [$_{\text{DP}}$ *pro*$_1$ [$_{\text{CP}}$ Op$_1$ that Malin bought ~~Op$_1$~~]]]

Regardless of whether the head of the relative is overt or null, this alternative analysis entails that *wh*-questions in Chuj always involve biclausality. Specifically, questions are formed by combining a *wh*-predicate with a single argument, resulting in a pseudocleft strategy as the means to ask a question. The only difference between (53a) and (54a) is that a null head or pronoun heads the relative clause in the latter case.

It is important to note at this juncture that what corresponds to the null pronoun in (54a) can be overtly instantiated by a pronoun in Chuj, resulting in a light-headed relative clause. Like other Q'anjob'alan languages, a set of noun classifiers is employed as third person pronouns (see e.g., Buenrostro et al. 1989, Hopkins 2012, Royer 2022a). A minimal pair for (54a) is provided below:

(55) **Context:** Malin went to the bookstore to buy something, and you know that I know what she bought. Assuming she bought a book, you ask me:

a. Mach **ch'anh** ix-s-man ix Malin?
which.is CLF.PRON.PAPER PFV-A3-buy CLF Malin
'Which is it$_{[paper]}$ that Malin bought?'

b. [$_{PredP}$ which.is [$_{DP}$ **ch'anh**$_1$ [$_{CP}$ Op$_1$ that Malin bought ~~Op$_T$~~]]]

In section 4.3.3, we will provide additional evidence for the analysis that questions like (54a) indeed involve headless relative clauses.

Crucially, if the syntax proposed previously is the *only* strategy available to form questions in Chuj, we have a unified explanation for the three properties exhibited by *wh*-questions in the language and Mayan more broadly (recall section 4.1). Consider first the ban on *wh-in situ*. Under our analysis, the ban is actually an illusion, a misnomer. Since Chuj is a predicate-initial language, *wh*-items must appear in sentence initial position—i.e., they are actually *always in-situ*. What is ruled out is a *postverbal* placement for the *wh*-items, and this follows naturally because nonverbal predicates simply cannot appear in that position.

Turning now to the ban on multiple *wh*-questions, it can also follow from our analysis. Many languages ban cleft/pseudocleft stacking (Kiss 1998, Reeve 2012). This is the case in languages like English:[13]

(56) *Who is what is the thing that you saw?

If the ban on stacking (psuedo)clefts extends to Chuj, then we can understand why multiple *wh*-questions are not permitted, regardless of the placement of each *wh*-item.

Finally, under our approach, "pied-piping with inversion" becomes a misnomer as well. That is, it becomes possible—and analytically necessary—to account for the relevant data without appealing to any inversion at all. As nonverbal predicates, *wh*-items stay *in-situ* (in the preverbal position), and no inversion is needed: an element such as *yet'* below is instead fronted as a byproduct of the movement of the relative operator, and *yet'* is thus expected to surface *after* the *wh*-item.[14]

[13] *Wh*-items cannot be coordinated in Chuj. This follows because predicates, including nonverbal ones, cannot be coordinated in the language; instead, full clauses must be coordinated:

(i) * Saksak k'ank'an ch'anh libro.
white yellow CLF book
Intended: 'The book is white and yellow.'

(ii) Saksak ch'anh libro, k'ank'an ch'anh.
white CLF book yellow CLF
'The book is white and it's yellow.'

If a language had *wh*-predicates but did allow predicate coordination, then we might expect it to exhibit the properties in (3a) and (3c) but not (3b). See Antonio-Ramos 2021 on San Pedro Mixtepec Zapotec for a potential language of this type.

[14] We note that this pied-piping of relational nouns alongside the operator is optional in Chuj; i.e., relational noun stranding is possible, as shown below. Other Mayan languages disallow stranding (see Ranero 2021 for Kaqchikel).

(i) Mach *pro* ix-ach-xit' ek' yet'ok?
who.is PRON PFV-B2S-go DIR.pass with
'Who did you go with?' (*Literally*: 'Who is it that you went with?')

(57) a. Mach *pro* yet' ix-ach-xit' ek'-i?
who PRON with PFV-B2S-go DIR.pass-IV
'With whom did you go?'
Literally: 'Who is it with whom you went?'

b. [$_{\text{PredP}}$ who.is [$_{\text{DP}}$ *pro*$_1$ [$_{\text{CP}}$ with Op$_1$ you left ~~with Op$_1$~~]]]

As was the case with the headless relative clause in (54a), it is important to mention that the covert relative head in (57a) can be instantiated by an overt classifier pronoun. For completeness, a relevant example with context, as well as the proposed syntactic structure, are provided in (58).

(58) **Context:** I just went to collect salt in a salt mine, which usually only happens with men, not women. You want to ask me who I went with.

a. Mach **winh** yet' ix-ach-xit' ek'-i?
who.is CLF.PRON.MALE with PFV-B2S-go DIR.pass-IV
'With whom did you go?'
Literally: 'Who is it$_{[\text{man}]}$ with whom you went?'

b. [$_{\text{PredP}}$ who.is [$_{\text{DP}}$ **winh**$_1$ [$_{\text{CP}}$ with Op$_1$ you left ~~with Op$_1$~~]]]

In sum, this section has offered an analysis of the syntax of *wh*-questions in Chuj that departs from the traditional view that considers them to be DP-internal elements. We extended the independently necessary analysis of Pred-Qs from section 3.3, thus reanalyzing *wh*-questions as involving a pseudocleft structure in which the *wh*-item serves as a matrix nonverbal predicate. Under our approach, *wh*-questions never involve movement of a *wh*-item; instead, they tend to require combination with a relativized DP subject (see section 4.3.2 and 4.3.5 for predicative uses of *wh*-items without relativization).

Moreover, we argued that our analysis provides an appealing way to unify the basic syntax of *wh*-questions with three (only apparent) idiosyncrasies that they exhibit. If interrogatives exhibit the clausal syntax involved in Pred-Qs, then the three properties follow in one fell swoop. In fact, non-interrogative Pred-Qs are subject to the same three syntactic restrictions, which is exactly what we would expect: as shown via the distribution of our Pred-Q quantifier *tzijtum*, (i) Pred-Qs cannot follow verbal predicates (59a), (ii) multiple Pred-Qs in the same sentence are ill-formed (59b), and (iii) Pred-Qs give rise to cases of (perhaps wrongly named) "pied-piping with inversion" (59c).

(59) a. * Ix-s-man **tzijtum** ch'anh libro ix Malin.
PFV-A3-buy MANY$_{\text{PRED}}$ CLF book CLF Malin
Intended: 'Malin bought many books.'

b. * **Tzijtum** heb' ix ix **tzijtum** winh winak ix-il-an-i.
MANY$_{\text{PRED}}$ PL CLF woman MANY$_{\text{PRED}}$ CLF man PFV-see-AF-IV
Intended: 'Many women saw many men.'

c. **Tzijtum** heb' anima' **yet'** ix-ach-lolon-i.
MANY$_{\text{PRED}}$ PL person with PFV-B2S-speak-IV
Literally: 'The people with whom you spoke are many.'

In contrast, under the prevalent DP analysis of *wh*-items, it is not possible to uniformly account for the three properties. As discussed in section 4.1, that approach requires stipulating independent reasons for each.

Now, in order to definitively adjudicate between the DP analysis of *wh*-items and our own predicative analysis, we employ the sets of diagnostics that we used to establish the distinction between Det-Qs and Pred-Qs in section 3. As will become clear, *wh*-items in Chuj exhibit all the signals for predicative status (§4.3) and exhibit none of the testable signals for determiner status (§4.4). In a nutshell, then, we will conclude that *wh*-items in the language are nonverbal predicates and, as a consequence, that *wh*-questions involve the pseudocleft syntax proposed in this subsection, instead of the monoclausal syntax in section 4.1 that is generally assumed in previous work on Mayan *wh*-questions.

4.3 *Wh*-items check predicative diagnostics

If *wh*-items are exclusively Pred-Qs, then they should fulfill the predicative diagnostics discussed in section 3.2. Here, we show that this is correct. While we only provide examples from a handful of *wh*-items below, we note that *wh*-items behave alike. A full list of *wh*-items in San Mateo Ixtatán Chuj is provided below (see Kotek and Erlewine 2019, Royer 2021, and Alonso-Ovalle and Royer 2024 for further descriptions).

(60) List of *wh*-expressions in San Mateo Ixtatán Chuj

	wh-expression
'what'	*tas* (+N)
'who'	*mach*
'which'	*mach* (+N)
'where'	*b'ajt'il/ajt'il*
'when' (future)	*b'ak'inh*
'when' (past)	*b'ak'nitaxo*
'how'	*tas* + light verb
'how much'	*jantak*
'how many'	*jay*-NUM.CLF
'what time'	*janik*
'why'	*tas yuj*

Before delving into the diagnostics, we highlight that there is homophony between the *wh*-items *jantak* and *jay*-NUM.CLF and the non-interrogative determiner quantifiers meaning 'many' (*jantak*) and 'few' (*jay*-NUM.CLF). However, the interrogative versus non-interrogative counterparts for each form are syntactically in complementary distribution. When used as determiners in postverbal positions or within focus constructions (with the marker *ha*), *jantak* and *jay*-NUM.CLF are necessarily non-interrogative, meaning 'many' and 'few' respectively. When used as nonverbal predicates, they necessarily lead to interrogative semantics (meaning 'how many').

The following examples illustrate this. When *jantak* is clearly employed alone in a predicational sentence—where it is necessarily instantiated as a nonverbal predicate—it can only deliver an interrogative meaning 'how many' (61a). In contrast, when *jantak* is employed postverbally (61b), or in a focus construction (61c), it can only deliver the meaning of 'many'.

(61) a. **Jantak** heb' ix ix?
how.many PL CLF woman
'How many women are there?'
Unavailable: 'There are many women.' (clearly Pred → 'how many')

b. Ix-s-man **jantak** ch'anh libro ix Malin.
PFV-A3-buy MANY_{DET} CLF book CLF Malin
'Malin bought many books.'
Unavailable: 'Malin bought how many books?' (clearly Det → 'many')

c. Ha **jantak** ch'anh libro ix-s-man ix Malin.
PV MANY_{DET} CLF book PFV-A3-buy CLF Malin
'It is many books that Malin bought.'
Unavailable: 'It's how many books that Malin bought?' (clearly Det → 'many')

It may be possible to derive Pred-Qs from Det-Qs (or vice-versa), with the goal of explaining this homophony and the observed semantic differences—such an exercise is left for the future. We simply note here that our main point is not affected: as we will show below, *wh*-items systematically behave like nonverbal predicates in Chuj.

4.3.1 Inability to appear in postverbal position

As amply discussed already, *wh*-expressions in Chuj cannot appear postverbally in questions. Examples are repeated below for convenience. While this has traditionally been taken as evidence that *wh-in situ* is not possible, a Pred-Q analysis of *wh*-items in questions predicts the ill-formedness of postverbal *wh*-items—in brief, whatever drives the predicate initial nature of Chuj renders such a configuration impossible.

(62) a. **Mach** ix-h-ila'?
who PFV-A2S-see
'Who did you see?'

b. * Ix-h-il **mach**?
PFV-A2S-see who
Intended: 'Who did you see? / 'You saw *who*?'

As was the case for the distribution of predicate quantifiers like *tzijtum* (see (28) above), when the subject of a *wh*-predicate is topicalized, the *wh*-item appears linearly after a verbal predicate (we thank an anonymous reviewer for pointing this out to us). However, the topicalized element is in a peripheral position, and thus our conclusion need not be amended:

(63) Ha winh ix-h-il-a', **mach** winh?
PV CLF PFV-A2S-see-TV who PRON
'The one you saw, who is he?'

4.3.2 Ability to serve as nonverbal predicate

Chuj *wh*-items can readily be employed as nonverbal predicates, as expected if they belong to that category. Example (61a) already showed this. The following examples further illustrate:

(64) **Mach** winh alkal?
who CLF mayor
'Who is the mayor?'

(65) **Mach** hin?
who B1S
'Who am I?'

4.3.3 Visible signs of relativization

If *wh*-items are nonverbal predicates, then they should show the hallmarks of relativization when a question is complex in involving a verbal event. This follows because the structural correlate of the denoted event will have to be embedded within a relative clause. This prediction is indeed borne out when we consider *wh*-items for oblique arguments. The relevant *wh*-questions require the presence of the (previously discussed) relative pronoun *b'aj*, which we also found with Pred-Qs.

A baseline sentence with an indirect object is provided in (66a), with the corresponding *wh*-question provided in (66b). While indirect objects appear with the preposition *t'a* when postverbal, questioning the indirect object requires a different structure without *t'a* and with a relative pronoun instead (see section 4.4.1 for data on combining *wh*-items with *t'a*).

(66) a. Ix-h-ak' b'at ch'anh karta t'a waj Xun.
PFV-A2S-give go CLF letter PREP CLF Xun
'You sent the letter to Xun.'

b. **Mach** *(b'aj) ix-h-ak' b'at ch'anh karta?
who REL.OBL PFV-A2S-give go CLF letter
'To whom did you send the letter?'
Literally: ≈ 'Who is it where you sent the letter?'

As already hinted at in section 4.2, additional signs of revitalization are present in questions involving "headless relative clauses", which lack a clear nominal head (see Caponigro et al. 2021). The reason is that such clauses are often "light-headed" (Citko 2004) in Chuj: while they lack an overt nominal head, they nonetheless exhibit a determiner, pronoun, or demonstrative. Examples with pronouns as light heads were provided in (55a) and (57a) above. We offer other examples below, one where the light-head is a demonstrative (67a) and another where it is the numeral/indefinite determiner *jun* (67b). For more on these kinds of relative clauses in Chuj, see Royer 2021: §4.2.

(67) a. **Mach** [[ix-h-il] chi']?
who PFV-A2S-see DEM
'Who did you see?'
Lit: 'Who is that, that you saw?'

b. **Mach** [jun [ix-h-il-a']]?
who one PFV-A2S-see-TV
'Who did you see?'
Lit: 'Who's the one that you saw?'

While the light-heads are not obligatory in the examples in (67), they provide additional support for the idea that *wh*-questions in Chuj involve a structural element that is relativized.

4.3.4 Ability to appear in possessive-'have' constructions

If *wh*-items are a subtype of Pred-Q, then we would expect that some of them might appear in possessive 'have' constructions. While this diagnostic is not informative with every *wh*-item in the language (i,e., there might be independent reasons why an incompatibility could arise between *wh*-items and this construction), (68) shows that *wh*-items such as *jay*+NUM.CLF 'how many' can indeed appear as the sole predicate here. Notice as well that an interrogative meaning arises:

(68) **Jay-wanh** h-uninal?
how.many-ANIM A2S-children
'How many children do you have?'
Literally: 'Your children are how many?'

4.3.5 Ability to appear as a secondary predicate

Secondary predicate constructions were another syntactic environment in which Pred-Qs were shown to surface. We might therefore expect some *wh*-items to be licit here as well. This, again, is the case with *jay*-NUM.CLF, which can appear in the position reserved for secondary predicates:

(69) [$_{\text{PRED}}$ **Jay-wanh**] [$_{\text{VP}}$ y-ek' heb' paxyalwum t'atik] ?
how.many-ANIM A3-pass.by PL visitor here
'How many visitors pass by here?'
Literally: ≈ 'The visitors pass by how many here?'

We highlight here that the ability for *wh*-items to appear in complex predicate positions has precedent in the Mayanist literature. The structure in (69) is similar to the structure posited by Can Pixabaj and Aissen (2021) to account for 'how'-questions in K'iche', a configuration where that specific *wh*-item is analyzed as a nonverbal predicate. While the authors do not refer to this configuration as "secondary" predication specifically, there is a clear parallel with our Chuj data.[15] In (70), the *wh*-item is analyzed by the authors as a nonverbal predicate combining with an intransitive nominalization bearing Set A morphology and no aspect marking:

(70) [**Jas**]$_{\text{PRED}}$ [u-tzakab'a-x-ik le sub']$_{\text{SUBJ}}$?
what A3SG-cook-PASS-NMLZ DET tamalito
'How are the tamalitos cooked?' (K'iche', Can Pixabaj and Aissen 2021: 125)

4.3.6 Ability to be modified by predicate modifiers

The final predicative diagnostic dicussed in section 3.2 was the ability for Pred-Qs to co-occur with predicative modifiers, such as *-(h)amlaj*. Again, *wh*-items in Chuj readily co-occur with *-(h)amlaj*. When they do, the result is a meaning paraphrasable as 'Who knows *wh* X':

[15] See Mateo Toledo 2012 for secondary predication in Q'anjob'alan.

(71) **Mach**-amlaj ix-k'ochi-i.
who-MODAL PFV-arrive-IV
'Who knows who arrived.'
Literally: ≈ 'Who possibly is the person who arrived.'

4.3.7 Interim summary

In sum, *wh*-items check all diagnostics indicative of predicative status. Since some of them force a predicative analysis—such as the fact that *wh*-items corresponding to oblique arguments require relativization—this section has already provided support for the stronger claim that *wh*-items are *exclusively* instantiated as nonverbal predicates in Chuj. In the next section, we go a step further and apply the determiner-status diagnostics to Chuj *wh*-items. While only a subset of the diagnostics can be tested, we will definitively conclude that *wh*-items cannot be analyzed as determiners.

4.4 *Wh*-items do not check (testable) diagnostics for determiner status

Since *wh*-items are independently known to be bad candidates for topicality (*wh*-items map to focus instead; see e.g., Aissen 2023), we will only consider as informative the other two diagnostics for determiner status.[16]

4.4.1 Inability to co-occur with prepositions

The first diagnostic we consider is co-occurrence with a preposition. As shown in section 3, only Det-Qs are expected to be found in the complement of the proposition *t'a* in Chuj. If *wh*-items are Pred-Qs, as we maintain, they should never co-occur with a preposition. This prediction is borne out: *wh*-items cannot co-occur with a preposition, irrespective of the order of the elements. Instead, the relativization strategy discussed in section 4.3.3 is used.

(72) a. Ix-h-ak' b'at ch'anh karta t'a waj Xun.
PFV-A2S-give go CLF letter PREP CLF Xun
'You sent the letter to Xun.'

b. * T'a **mach** ix-h-ak' b'at ch'anh karta?
PREP whom PFV-A2S-give go CLF letter
Intended: 'To whom did you send the letter?'

c. * **Mach** t'a ixhak' b'at ch'anh karta?
whom PREP PFV-A2S-give go CLF letter
Intended: 'To whom did you send the letter?'

4.4.2 Inability to appear in the possessor of a DP

The second testable diagnostic for determiner status is the ability to appear in the possessor of a possessed DP. Constructions in which a *wh*-item co-occurs immediately beside a possessum, no

[16]Note, however, that the relevant data has been checked and *wh*-items (unsurprisingly) cannot appear alongside topicalized nominals—we omit the data for brevity.

matter the word order, are indeed unacceptable:[17]

(73) a. * **Mach** s-pat ix-ixta-j b'at-i?
who A3-house PFV-destroy-PASS go-IV

b. * S-pat **mach** ix-ixta-j b'at-i?
A3-house who PFV-destroy-PASS go-IV

c. * Ix-ixta-j b'at s-pat **mach**?
PFV-destroy-PASS go A3-house who

Intended (for all three): 'Whose house was destroyed?'

The only way to convey the intended meaning in (73c) is with the example in (74). In this case, the *wh*-item appears in a sentence-initial position, as expected given the Pred-Q analysis, and the possessed DP remains in its base position. In this case, we assume that a null pronoun appears as the possessor of the possessive phrase.

(74) **Mach** ix-ixta-j b'at s-pat *pro*?
who PFV-destroy-PASS go A3-house PRON
'Whose house was destroyed?'
Literally: Who is it that their house was destroyed?'

A potential counter-analysis could be invoked to derive the example above. While our predicative approach predicts directly that the structure in (74) would be employed, one could alternatively call upon possessor sub-extraction from a DP. In light of similar data, such an analysis is defended by Little (2020) for Ch'ol and Hedding and Yuan (2025) for other Mesoamerican languages exhibiting pied-piping with inversion. We note, however, that sub-extraction analyses have been contested on independent grounds (Aissen and Polian 2024), and that there is no independent evidence for subextraction (of e.g., numerals or adjectives) in Chuj—unlike in Ch'ol, where such independent evidence is actually available.

We therefore maintain that data like (73c) and (74) indicate that *wh*-items cannot appear in the possessor position of possessive phrases in Chuj, as expected if *wh*-items are exclusively instantiated as nonverbal predicates.

4.5 Summary

In this section, we considered whether *wh*-items, a subtype of quantifier, should be analyzed as Det-Qs or Pred-Qs in Chuj. We applied the diagnostics from section 3 and concluded that *wh*-

[17]The unacceptability of (73a) is perhaps surprising from the perspective of other Mayan languages, for which it has been argued that "pied-piping with inversion" extends to possessive phrases (Aissen 1996, Coon 2009, Polian 2013, Little 2020, Aissen and Polian 2024). This is the case, for instance, in Oxchuc Tseltal:

(i) **Macha** x-nich'an bejk'aj?
who A3-child PFV.be.born
'Whose child was born?' (Oxchuc Tseltal, Polian 2013: 230)

While previous works assume that *wh*-items in languages like Tseltal are part of the DP domain, a predicative analysis like ours could derive these data by extending to them our proposal for the pied-piping of relational nouns in Chuj under relativization (see example (57a) above).

items are exclusively instantiated as nonverbal predicates in the language. The following table compiles the results of all the diagnostics:

(75) *Testing for Pred vs Det status of quantifiers, including wh-items*

Testing for:	Diagnostic	*Jantak*	*Tzijtum*	*Wh-items*
Det status	Can be part of the comp. of a preposition	✓	✗	✗
	Can be part of the possessor of a DP	✓	✗	✗
	Can be part of a topicalized DP	✓	✗	n/a
Pred status	Obligatory predicate-initial syntax	✗	✓	✓
	Can be the pred. of a predicational clause	✗	✓	✓
	Shows signs of relativization	✗	✓	✓
	Can be the pred. of possessive 'have'	✗	✓	✓
	Can be a secondary predicate	✗	✓	✓
	Appearance with predicate modifiers	✗	✓	✓

We also argued that analyzing *wh*-items exclusively as Pred-Qs provides a unified account of three (seemigly) idiosyncratic aspects of *wh*-question formation in Chuj: (i) the apparent ban on *wh-in-situ*, (ii) the ban on multiple *wh*-questions, and (iii) the phenomenon known as "pied-piping with inversion". This conclusion is appealing on a theoretical level and far from trivial, since it runs against the status quo in the Mayanist literature, including previous work on Chuj (Kotek and Erlewine 2019, Royer 2021), which had unanimously assumed a DP analysis for *wh*-items. As far as we can tell, such an analysis cannot account for these restrictions in a unified manner and must invoke independent stipulations.

Moreover, since the idiosyncrasies on *wh*-questions apply across the Mayan family (as far as we know), we suggest that the predicative strategy for question formation may well be widespread—an investigation that we hope will bear many fruits in the future. In the conclusion, we turn to suggestions for such elaborations.

5 Conclusion and discussion

We fulfilled our two main goals in this paper. First, we built on the diagnostics proposed in Royer, Buenrostro, and Jenks 2025a and established that there exists a clear division in Chuj between determiner quantifiers (Det-Qs) and predicative quantifiers (Pred-Qs). Second, we argued that *wh*-items pattern with Pred-Qs in forming questions, a conclusion that departs from the *status quo* in syntactic work on Mayan, which has consistently posited a movement strategy, where the *wh*-item definitionally occupies a position within the DP spine.

Our proposal, however, took us a step further by providing a fresh perspective on patterns that go beyond the basic distribution of *wh*-items in questions. We argued that our analytical conclusions allow for a unification of phenomena that had been noted previously, but had resisted an explanatorily adequate derivation. In a nutshell, our proposal that *wh*-questions involve pseudo-clefting provides us with solid grounding to derive three observations about Chuj *wh*-questions that also appear to be Pan-Mayan in nature: (i) the apparent ban on *wh-in-situ*, (ii) the ban on multiple *wh*-questions, and (iii) the phenomenon known as "pied-piping with inversion".

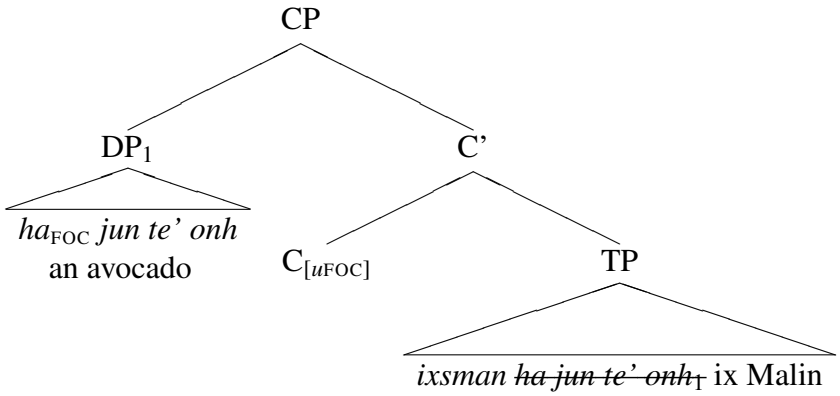
To end this paper, we delve into three areas of extension in order to test our proposal: the first two within Mayan and the other more broadly cross-linguistically.

5.1 Focus constructions and predication

Since this paper focused on the syntactic distribution of *wh*-words within questions, we intentionally set aside a construction whose syntax is generally treated in parallel with that of Mayan *wh*-questions: the construction involving the focusing of nominal expressions. Since Aissen 1992, a standard approach to focus constructions in this language family has been to assume that they involve A'-movement of a DP to the left periphery (see also Aissen 2017a and Royer and Coon 2025 for overviews). Under such an approach, focus constructions are treated as monoclausal, and whatever focus marker is employed (in this case *ha*) is generally viewed as part of the focused DP (see e.g., Royer et al. 2025b).

(76) a. [**Ha jun te' onh**] ix-s-man ix Malin.
PV INDF CLF avocado PFV-A3-buy CLF Malin
≈ 'It's an avocado that Malin bought.' (focused object)

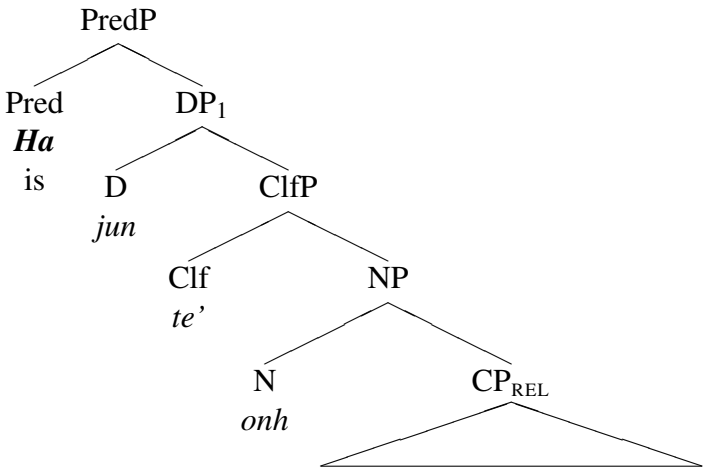
b. *Possible syntax for* (76) (to be refuted for Chuj)



Since *wh*-clauses involve pseudoclefts under our proposal, it is interesting to consider whether Chuj focus constructions could also be analyzed along those lines. This, again, would be in accordance with Tonhauser's (2003, 2007) approach to focus constructions in Yucatec Maya (see also Bricker 1979, Bohnemeyer 1998, and for Ixil, Blunk 2008; see Gutiérrez-Bravo 2017, Verhoeven and Skopeteas 2015 for criticisms). Under this type of approach, *ha* in focus constructions could be reanalyzed as a predicate: a copular morpheme occurring in special kinds of copular clauses (77b).

(77) a. [**Ha jun te' onh**] ix-s-man ix Malin.
PV INDF CLF avocado PFV-A3-buy CLF Malin
≈ 'It's an avocado that Malin bought.' (focused object)

b.



As it turns out, there are various empirical reasons to favour a biclausal analysis of focus constructions in Chuj. The first is that focus constructions exhibit biclausal behaviours. For example, a focused nominal can bear negation in tandem with a verb that also bears negation. We indicate phonological deletion of *ha* in (78), since this segment is expected to be deleted given independent phonological rules governing vowel hiatus in the language; see Royer et al. 2025b. This would follow directly from extending our approach to focus configurations, since an example like (78a) below would be biclausal:

(78) **Context**: Of the ten kids around us, one of them didn't eat their meal. You're under the impression that it was Xun that didn't eat. But I know that it was Petul. I tell you:

a. **Manh**-<ha>-ok-**laj** waj Xun **maj**-wa'-**laj**.
NEG-PV-IRR-NEG CLF Xun NEG.PFV-eat-NEG
'It wasn't Xun who didn't eat.'

b. Ha waj Petul, maj-wa'-laj.
PV CLF Petul NEG.PFV-eat-NEG
'It was Petul who didn't eat.'

Since only predicates may be negated in Chuj, a monoclausal, focus-fronting approach could not provide a straightforward account for the data above.

Several other applications of the diagnostics for predicatehood established in this paper support this. Consider, for instance, the following two examples.

(79) **Ha** [ix Malin [b'aj ix-w-ak' b'at ch'anh karta]].
PV CLF Malin REL PFV-A1S-give go CLF letter
'It's Malin who I sent the letter to.'

(80) **Ha**-xomlaj [jun te' onh [ix-s-man ix Malin]].
PV-MODAL one CLF avocado PFV-A3-buy CLF Malin
'Perhaps it's an avocado that Malin bought.'

In (79), we see that focused oblique arguments behave like predicative quantifiers and *wh*-items in co-occuring with a relative pronoun, suggesting relativization is at play. In (80), we see that

a modal particle that only ever appears on predicates appears on *ha*, suggesting that this element is itself a predicate. In fact, *ha* can never co-occur with predicative quantifiers or *wh*-items in questions, suggesting they indeed belong to the same natural class. These preliminary facts suggest that our analysis of *wh*-questions in Chuj should be extended to focus constructions in general.

Another well known property of constructions involving *wh*-questions, focus, and relativization in Chuj and other Mayan languages is that they exhibit what has been described as syntactic ergativity (Aissen 2017b; for Chuj, Royer and Coon 2025). In particular, when an ergative subject is questioned, focused, or relativized, the verb cannot bear Set A agreement and it must appear with an "agent focus" suffix. Whereas prominent analyses have attributed this phenomenon to a consequence of DP movement in the relevant configurations (see e.g., Coon et al. 2014; Coon et al. 2021), our analysis forces a reconsideration: what unifies the three constructions is the existence of relativization of an ergative subject. We thus open the door to reanalyzing the "agent focus" morpheme as a marker of agent relativization. Much more needs to be said—including what leads to the loss of Set A morphology—but we leave this for future exploration, both in Chuj and beyond.[18]

5.2 Mayan languages beyond Chuj

Applying diagnostics akin to ours in other Mayan languages will put our proposal to the test, thus shedding light (or casting doubt) on whether our unified derivation of the properties of *wh*-items and *wh*-questions replicates beyond Chuj. If such replications are indeed observed, we will have further evidence that the grammar may provide predication as the sole strategy for *wh*-question formation, with cascade effects observed across the syntactic systems of individual languages.

For now, however, let us point out that at least one Mayan language appears to challenge a cross-Mayan unification. Ch'ol has a relativizing morpheme, the clitic *=bä*, that does *not* appear in *wh*-questions (Vázquez Álvarez 2011).[19] Consider for instance the following examples:

(81) Ch'ol *wh*-questions (81a-81b) and relativization (81c)

a. Chu=**ki** ta' a-k'uxu?
what-Q PFV A2-eat
'What did you eat?' (Little et al. 2024: 1089)

b. Majch=**ki** ta' i-k'ele-yety?
who-Q PFV A3-see-B2
'Who saw you?' (Little et al. 2024: 1089)

c. K'am-Ø li wiñik [ta'=**bä** i=ts'äk-ä-Ø aj-wijty]
sick-B3 DET man PFV=REL A3-cure-TV-B3 NCL-shaman
'The man that the shaman cured is sick.' (Vázquez Álvarez 2011: 408)

All else being equal, the above Ch'ol data are surprising if *wh*-questions should always involve pseudoclefts in that language as well; that being the case, we might expect the relativizing clitic *=bä* in (81c) to also appear in *wh*-questions.

[18] A similar reconsideration would be necessary if we sought to extend our analysis to Mayan languages where *wh*-questions of certain adjuncts trigger the appearance of what has been analyzed as the footprint of (successive-cyclic) movement (Mendes and Ranero 2021).

[19] Chontal also has a relative marker *=ba*, which should be explored as well in the future (Osorio May 2016). Thank you to an anonymous reviewer for bringing this to our attention.

While these facts pose a potential challenge to the pan-Mayan unification, they do not rule out a predicative analysis definitively. For example, it could be that Ch'ol employs more widely a secondary predication strategy analogous to the one discussed in section 4.3.5 for *wh*-question formation, or another structure involving a predicative *wh*-item (see e.g., Can Pixabaj and Aissen 2021 on 'how' questions in K'iche'). Another possible avenue might be to reconsider the function of the clitic *=ki*, bolded in (81) and found on all *wh*-items in interrogatives. Indeed, to extend the pseudocleft analysis, we suggest preliminarily that this clitic might be reanalyzable as a relative marker, one sensitive to features of a matrix *wh*-predicate. Finally, note that one *wh*-word in Ch'ol, *ba=ki=bä*, has been claimed to involve the relativizing morpheme *=b'ä* (Vázquez Álvarez and Coon 2020: 371), suggesting that relativization might be involved at some level.

While we can only provide speculative thoughts for the moment, we stress that there is a clear empirical motivation for a pan-Mayan unification: most Mayan languages we are aware of, Ch'ol included, exhibit the three properties identified at the beginning of this paper (3)—the ban on *wh-in situ*, the ban on multiple *wh*-questions, and the existence of "pied-piping with inversion".

5.3 Broader cross-linguistic parallels

Naturally, our search for parallels or extensions should take us beyond the Mayan language family.

For starters, as influentially established by Smith-Stark (1988), "pied-piping with inversion" is an areal feature of Mesoamerica (see also Aissen 1996, Broadwell 2006 and Hedding 2022: chapter 4). Therefore, if we are correct that "pied-piping with inversion" is a misnomer for Chuj, it may also be a misnomer for at least some other Mesoamerican languages. In light of this, we note that the two other properties that we identified at the outset of this paper in (3)—(i) the ban on *wh-in situ* and (ii) the ban on multiple *wh*-questions—are well documented for at least several Mixtec (Otomanguean) languages (Caponigro et al. 2013; Hedding 2022, 2025), as well as for Sierra Popoluca (Mixe-Zoquean; López Márquez 2021: 480-482). Identifying language-specific diagnostics for nonverbal predication in these languages, and applying them to *wh*-words and/or quantificational items in general, should prove fruitful in shaping future analyses of questions, foci, and consequently relativization in the Mesoamerican area. We also note that, in the context of defending the *omnipredicativity hypothesis* for Classical Nahuatl (Launey 1994, 2004), Launey (2004: 8) proposes a predicative analysis of *wh*-items and questions analogous to the one proposed for Chuj in this paper. Again, this may suggest that the types of structures proposed here are more widespread across Mesoamerica.

Second, many of the languages in the Austronesian family have also been argued to employ the pseudocleft strategy for *wh*-questions (see Potsdam 2009, Chen et al. to appear, and references therein). A deeper exploration into our cluster of properties in (3) must be left for the future, but we note a couple of facts. First, in contrast to Chuj, *wh-in-situ* does seem to be available in these languages (see fn. 13 in Potsdam 2009), though subtle distinctions seem to be at play depending on the nature of the *wh*-expression (i.e., subject vs. non-subject in Malagasy; see Paul and Potsdam 2012: 168), and at least some sentence-final *wh*-items may in fact be analyzed as predicates (see Kraft to appear, §3). When it comes to "pied-piping with inversion", some languages in the family such as Sasak do exhibit the phenomenon (Austin 2006). Length considerations preclude us from further discussion, but we hope that our proposal spurs it.

Third, we find it thought-provoking that languages from a diverse range of language families

have also been analyzed as involving pseudocleft structures for question formation, including Salish (Kroeber 1991, Davis et al. 1993, Jelinek and Demers 1994, Jelinek 1998, Davis 2008, 2009), Algonquian (Bloomfield 1946, Blain 1997) and Tsimshianic (Beck 2002, Davis and Brown 2011, Brown 2024; though see Brown 2026 on Sm'algyax), and that all such languages seem to exhibit some or all of the same restrictions on *wh*-questions as the ones discussed in (3).

Finally, another potentially relevant language family to consider is Celtic, which has also been claimed to exhibit identical properties as the ones listed in (3) (see e.g., McCloskey 1979, Oda 2002, Adger and Ramchand 2005), including what Thoms (2025: section 4.2) recently describes as "pied-piping with inversion" (see Adger 2025 for an analysis of this phenomenon in Scottish Gaelic). This is in addition to the well known fact that *wh*-questions in Celtic languages appear with the complementizer otherwise employed for nominal relativization (see morpheme glossed as "C.REL" in (82b)), a fact that led previous authors such as McCloskey (1979) and Oda (2002) to propose that *wh*-questions in languages like Irish consistently involve cleft-like syntax. To illustrate, consider the following example from Scottish Gaelic, zooming in on the order of the preposition and DP in the declarative sentence in (82a) versus the order of the preposition and the *wh*-item in the interrogative in (82b).

(82) Scottish Gaelic

a. Thug Màiri leabhar do **dh'Iain**.
give.PST Màiri book to Iain
'Màiri gave a book to Iain.' (Thoms 2025: 142)

b. **Cò** dha **a** thug i ball?
who to.him C.REL give.PST she ball
'Who did she give a ball to?' (Thoms 2025: 159)

It would be encouraging to uncover more evidence that the phenomena discussed across our paper indeed cluster together. Again, one potentially productive way of going about this would be to first establish language-internal diagnostics that single out predication specifically.

To conclude, note that many of the languages in the families mentioned above are described as "predicate initial". It is worth continuing to explore how tight a link exists between predicate initiality and the range of the inventory of predicative *wh*-words in specific grammars (see Potsdam 2009 for an exploration of this nature in Austronesian). It seems natural to us that acquirers would integrate *wh*-expressions as predicates if their target grammar reserves clause initiality precisely to predicates in baseline declaratives, with cascade effects arising as a direct consequence.

Acknowledgments We are very grateful to our Chuj collaborators: Matin Pablo, Matal Torres, Petul Gómez, Neper Diego López García, Xapin Torres, and Elsa Velasco. We also thank editor Kersti Börjars, three anonymous reviewers, Judith Aissen, Colin Brown, Jessica Coon, Henry Davis, Hamida Demirdache, Peter Jenks, Carol Rose Little, Gilles Polian, and Roberto Zavala for comments. Audiences at colloquia held at the Universidad Nacional Autónoma de México, University of Washington, McGill University, and University of Wisconsin, Madison provided key feedback. Royer is first author; Buenrostro and Ranero share second-authorship and are listed alphabetically. All errors are our own.

Funding statement This study was funded by the Social Sciences and Humanities Research Council of Canada and the Fonds de Recherche du Québec.

List of abbreviations *: ungrammatical construction; <>: elided segment; 1: first person; 2: second person; 3: third person; ABS: absolutive marker; AP: antipassive suffix; AF: agent focus suffix; ASP: aspect marking; CLF: noun classifier; C: complementizer; DEIX: deictic particle; DTV: derived transitive status suffix; ERG: ergative marker; EXT: existential; IPFV: imperfective aspect; IRR: irrealis clitic; IV: intransitive status suffix; NCL = noun class prefix; NEG: negation; P: plural person in agreement markers; PASS: passive; PFV: perfective; PL: plural marker; PREP: preposition; PRON: pronoun; PV: preverbal marker; S: singular agreement marking; TV: transitive verb; YNQ: yes-no question particle.

References

Adger, David. 2025. Mind your Ps and Ks! Quirky subjects and nominal licensing. In *Footprints of Phrase Structure: Studies in Syntax in Honour of Tim Stowell*, ed. María J. Arche, Jan-Wouter Zwart, Hamida Demirdache, and Hagit Borer, 111–134. Amsterdam: John Benjamins.

Adger, David, and Gillian Ramchand. 2005. Merge and move: Wh-dependencies revisited. *Linguistic Inquiry* **36**, 161–193.

Aissen, Judith. 1979. Possessor ascension in Tzotzil. In *Papers in Mayan linguistics*, ed. Laura Martin, 89–108. Columbia, Missouri: Lucas Brothers.

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

Aissen, Judith. 1996. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* **14**, 447–491.

Aissen, Judith. 2017a. Agent Focus and passive in Tsotsil. In *Asking the right questions: Essays in honor of Sandra Chung*, ed. Jason Ostrove, Ruth Kramer, and Joseph Sabbagh, 143–161. UC Santa Cruz.

Aissen, Judith. 2017b. Correlates of ergativity in Mayan. In *Oxford handbook of ergativity*, ed. Jessica Coon, Diane Massam, and Lisa Travis. Oxford: Oxford University Press.

Aissen, Judith. 2023. Documenting topic and focus. *Language Documentation and Conservation* **26**.

Aissen, Judith, Nora C. England, and Roberto Zavala Maldonado. 2017. *The Mayan languages*. Routledge Language Family Series. New York: Routledge.

Aissen, Judith, and Gilles Polian. 2024. Possessor extraction and categorical subject in Tseltalan. *Natural Language & Linguistic Theory* **43**, 63–113.

Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.

Alonso-Ovalle, Luis, and Justin Royer. 2024. Modal indefinites and semantic variation: Lessons from Chuj. *Semantics and Pragmatics* **17**, 1–54.

AnderBois, Scott, and Miguel Óscar Chan Dzul. 2021. Headless relative clauses in Yucatec Maya. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 444–474. Oxford: Oxford University Press.

Antonio-Ramos, Pafnuncio. 2021. Headless relative clauses in San Pedro Mixtepec Zapotec. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 221–259. Oxford: Oxford University Press.

Armstrong, Grant. 2017. The syntax of non-verbal predication in Yucatec Maya. *Cuadernos de Lingüística* **4**, 137–212.

Austin, Peter. 2006. Content questions in Sasak, Eastern Indonesia: An optimality theoric syntax account. In *Insular Southeast Asia: Linguistic and Cultural Studies in Honour of Bernd Nothofer*, ed. Fritz Schulze and Holger Warnk, 1–12. Harrassowitz Verlag.

Baker, Mark. 1997. Thematic roles and syntactic structure. In *Elements of grammar*, ed. Liliane Haegeman, 73–137. Dordrecht: Kluwer.

Baker, Mark C. 1995. On the absence of certain quantifiers in Mohawk. In *Quantification in natural language*, ed. Emmon Bach, Eloise Jelinek, and Angelika Kratzer, 21–58. Dordrecht: Kluwer.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, ed. J. Kulas, J. H. Fetzer, and T. L. Rankin, volume 2, 241–301. Dordrecht: Springer.

Beck, David. 2002. Tsimshianic from a Central Northwest areal perspective. In *ICSNL XXXVII*, 35–60.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics* **10**, 43–90.

Bianchi, Valentina. 2000. The raising analysis of relative clauses: A reply to Borsley. *Linguistic Inquiry* **31**, 123–140.

Bielig, Louisa. 2015. Resumptive classifiers in Chuj high topic constructions. BA Thesis, McGill University.

Blain, Eleanor Marie. 1997. Wh-constructions in Nêhiyawêwin (Plains Cree). Doctoral Dissertation, University of British Columbia.

Bloomfield, Leonard. 1946. Algonquian. In *Linguistic structures of Native America*, ed. Harry Hoijer, volume 6, 85–129. Viking Fund Publications in Anthropology.

Blunk, William B. 2008. The agent-focus construction in Ixil Maya: A descriptive/formal analysis. Master's thesis, University of Texas Austin.

Bochnak, Ryan, and Lisa Matthewson. 2020. Techniques in complex semantic fieldwork. *Annual Review of Linguistics* **6**, 261–283.

Bodomo, Adams, and Ken Hiraiwa. 2010. Relativization in Dàgáárè and its typological implications: Left-headed but internally-headed. *Lingua* **120**, 953–983.

Bohnemeyer, Jurgen. 1998. Time relations in discourse: Evidence from a comparative approach to Yukatek Maya. Doctoral Dissertation, Tilburg University.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* **24**, 591–656.

Bricker, Victoria R. 1979. Wh-questions, relativization, and clefting in Yucatec Maya. In *Papers in Mayan linguistics*, ed. Victoria R. Bricker and William Norman, 109–138. Columbia, Missouri: Lucas Brothers.

Broadwell, Aaron. 2005. Pied-piping and optimal order in Kiche (K'iche'). Ms. SUNY Albany.

Broadwell, George Aaron. 2006. Alignment, precedence, and the typology of pied-piping with inversion. In *CSLA Publications*, ed. Miriam Butt and Tracy Holloway King.

Brown, Colin. 2024. Questions and their relatives in Sm'algyax. *International Journal of American Linguistics* **90**, 277–326.

Brown, Colin. 2026. Bipartite question formation in Sm'algyax. *Glossa: A journal of general linguistics* **11**, 1–49.

Buenrostro, Cristina. 1995. La negación en chuj de San Mateo Ixtatán. In *Vitalidad e influencia de las lenguas indígenas en Latinoamérica: Memorias del segundo coloquio m swadesh*, ed. Ramón Arzápalo Martín and Yolanda Lastra, 382–391. México, DF: Universidad Autónoma de México.

Buenrostro, Cristina. 2013. La voz en Chuj de San Mateo Ixtatán. Doctoral Dissertation, El Colegio de México, Mexico City.

Buenrostro, Cristina. 2021. *Relaciones gramaticales y disminución de transitividad en Chuj de San Mateo Ixtatán*. Mexico City: Universidad Nacional Autónoma de México.

Buenrostro, Cristina, José Carmen Díaz, and Roberto Zavala. 1989. Sistema de clasificación nominal del Chuj. In *Memorias del segundo coloquio internacional de Mayistas*, volume II. Mexico City: UNAM.

Can Pixabaj, Telma. 2021. Headless relative clauses in K'iche'. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado. Oxford: Oxford University Press.

Can Pixabaj, Telma, and Judith Aissen. 2021. Nominalization and the expression of manner in K'iche'. *International Journal of American Linguistics* **87**, 109–146.

Caponigro, Ivano, Harold Torrence, and Carlos Cisneros. 2013. Free relative clauses in two Mixtec languages. *International Journal of American Linguistics* **79**, 61–96.

Caponigro, Ivano, Harold Torrence, and Roberto Zavala Maldonado. 2021. *Headless relative clauses in Mesoamerican languages*. Oxford: Oxford University Press.

Chen, Victoria, Maria Polinsky, and Eric Potsdam. to appear. Austronesian syntax. In *Oceania*, ed. Bill Palmer. Berlin: Mouton.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and grammar*. Cambidge, MA: MIT Press.

Chomsky, Noam. 1977. On wh-movement. In *Formal Syntax*, ed. Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York, NY: Academic Press.

Citko, Barbara. 2004. On headed, headless, and light-headed relatives. *Natural Language & Linguistic Theory* **22**, 95–126.

Clemens, Lauren Eby, and Maria Polinsky. 2017. Verb-initial word orders (Primarily in Austronesian and Mayan languages). In *The companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk. Wiley-Blackwell, 2nd edition.

Coon, Jessica. 2009. Interrogative possessors and the problem with pied-piping in Chol. *Linguistic Inquiry* **40**, 165–175.

Coon, Jessica. 2010. VOS as Predicate Fronting in Chol. *Lingua* **120**, 354–378.

Coon, Jessica. 2014. Predication, tenselessness, and what it takes to be a verb. In *Proceedings of the 43rd meeting of the Northeast Linguistics Society (NELS 43)*, ed. Hsin-Lun Huang, Ethan Poole, and Amanda Rysling, 77–90. CreateSpace.

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

Coon, Jessica. 2018. Distinguishing adjectives from relative clauses in Chuj (with help from Ch'ol). In *Heading in the right direction: Linguistic treats for lisa travis*, ed. Laura Kalin, Jozina van der Klok, and Ileana Paul, 90–99. Montreal, QC: McGill Working Papers in Linguistics.

Coon, Jessica, Nico Baier, and Theodore Levin. 2021. Mayan Agent Focus and the ergative extraction constraint: Facts and fictions revisited. *Language* **97**, 269–332.

Coon, Jessica, and Martina Martinović. 2023. Predication, specification, and equation in Ch'ol. In *Proceedings of NELS 53*, ed. Suet-Ying Lam and Satoru Ozaki.

Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* **14**, 179–242.

Coon, Jessica, and Juan Jesús Vázquez Álvarez. 2025. Embedding, extraction, and clause-piping in Ch'ol. *Glossa: A journal of general linguistics* **10**, 1–28.

Curiel, Alejando. 2017. Tojolabal. In *The Mayan languages*, ed. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 570–609. Dordrecht: Routledge.

Davis, Henry. 2008. Wh-in-situ: Japanese meets Salish. In *Individual languages and language universals*, ed. Michiko Takeuchi, 17–58. Yokohama, Japan: Kanagawa University Studies in Language.

Davis, Henry. 2009. Cross-linguistic variation in anaphoric dependencies: Evidence from the Pacific Northwest. *Natural Langauge and Linguistic Theory* **27**, 1–43.

Davis, Henry, and J. Brown. 2011. On A'-dependencies in Gitksan. International Conference on Salish and Neighbouring Languages.

Davis, Henry, Dwight Gardiner, and Lisa Matthewson. 1993. A comparative look at wh-questions in Northern Interior salish. In *29th International Conference on Salish and Neighbouring Languages*.

Davis, Henry, and Lisa Matthewson. 2019. Quantification. In *The Routledge Handbook of North American Languages*, ed. Daniel Siddiqi, Michael Barrie, Carrie Gillon, Jason Haugen, and Éric Mathieu, 310–328. Dordrecht: Routledge.

De Vries, Mark. 2018. Relative clauses in syntax. *Oxford Research Encyclopedia of Linguistics* .

den Dikken, Marcel. 2006. *Relators and linkers*. Cambridge: MIT Press.

Douglas, Jamie, Rodrigo Ranero, and Michelle Sheehan. 2017. Two types of syntactic ergativity in Mayan. In *Proceedings of GLOW in Asia XI: Volume 2*, 41–56. Cambidge, MA: MIT Working Papers in Linguistics.

Dowty, David R., Robert E. Wall, and Stanley Peters. 1981. *Introduction to Montague Semantics*. Dordrecht: Reidel.

Duncan, Lachlan. 2010. The syntactic structure of K'ichee' Mayan. Doctoral Dissertation, SUNY Albany.

Duncan, Philip T., Harold Torrence, and Pedro Mateo Pedro. 2024. Indeterminate pronouns in Kaqchikel. *Linguistic Variation* **24**, 116–163.

Elkins, Noah, and Colin Brown. to appear. Free relative clauses in mam. In *Proceedings of WSCLA 26*.

England, Nora. 2001. *Introducción a la gramática de los idiomas Mayas*. Guatemala: Cholsamaj.

Ewing, Caleb. 2022. Pied-piping and wh-movement: A syntactic analysis and comparison of Kaqchikel and K'iche'. Doctoral Dissertation, University of Florida.

Faltz, Leonard. 1995. Towards a typology of natural logic. In *Quantification in natural language*, ed. Emmon Bach, Eloise Jelinek, and Angelika Kratzer, 321–362. Dordrecht: Kluwer.

Freeze, Ray. 1992. Existentials and other locatives. *Language* **68**, 553–595.

García Pablo, Gaspar, and Pascual Martín Domingo Pascual. 2007. *Stzolalil sloloni-spaxtini heb' chuj: Gramática descriptiva chuj*. Academia de Lenguas Mayas de Guatemala.

Gavruseva, Elena. 2000. On the syntax of possessor extraction. *Lingua* **110**, 743–772.

Grinevald, Colette, and Marc Peake. 2012. Ergativity and voice in Mayan: A functional-typological approach. In *Ergativity, valency and voice*, ed. Gilles Authier and Katharina Haude, 15–49. Berlin: Mouton de Gruyter.

Gutiérrez-Bravo, Rodrigo. 2017. Clefts and focus in Yucatec Maya. *Cuadernos de Lingüística* **4**, 5–47.

Heck, Fabian. 2009. On certain properties of pied-piping. *Linguistic Inquiry* **40**, 75–111.

Hedding, Andrew. 2022. How to move a focus: The syntax of alternative particles. Doctoral Dissertation, UC Santa Cruz.

Hedding, Andrew. 2025. On the representation of wh-words and foci: Evidence from Mixtec. *Accepted at Linguistic Inquiry* .

Hedding, Andrew, and Michelle Yuan. 2025. Distinct pathways to possessor A'-extraction in Mesoamerican languages. In *Proceedings of NELS 55*.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell Publishers.

Hiraiwa, Ken. 2017. Internally headed relative clauses. In *The Wiley Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk, volume 2, 1–32. Hoboken, NJ: Wiley-Blackwell.

Hopkins, Nicholas. 1967. The Chuj language. Doctoral Dissertation, University of Chicago, Chicago, IL.

Hopkins, Nicholas. 2012. Noun classifiers of the Cuchumatán Mayan languages: A case of diffusion from Otomanguean. *International Journal of American Linguistics* **78**, 411–427.

INE. 2018. *Censo nacional de población y vivienda en 2018*. Instituto Nacional de Estadística, Guatemala.

INEGI. 2020. *Censo general de población y vivienda en 2020*. Instituto Nacional de Estadística, Geografía e Información.

Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Jelinek, Eloise. 1995. Quantification in straits Salish. In *Quantification in natural language*, ed. Emmon Bach, Eloise Jelinek, and Angelika Kratzer, 487–540. Dordrecht: Kluwer.

Jelinek, Eloise. 1998. Prepositions in Northern Straits Salish and the noun/verb question. In *Salish languages and linguistics: Theoretical and descriptive perspectives*, ed. Ewa Czaykowska-Higgins and M. Dale Kinkade, 325–346. Berlin: Mouton de Gruyter.

Jelinek, Eloise, and Richard A. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* **70**, 697–736.

Jenks, Peter. 2025. The comparative syntax of nominal quantifiers. In *The cambridge handbook of comparative syntax*, ed. Sjef Barbiers, Norbert Corver, and Maria Polinsky. Cambridge: Cambridge University Press.

Kaplan, Emily. 2021. A Guatemalan town remakes itself in Indiana. *National Geographic* .

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and philosophy* **1**, 3–44.

Kaufman, Terrence. 1974. *Idiomas de mesoamerica*. Guatemala: Editorial José De Pineda Ibarra.

Kiss, Katalin É. 1998. Identificationanal focus versus information focus. *Language* **74**, 245–273.

Kotek, Hadas, and Mitcho Erlewine. 2019. *Wh*-indeterminates in Chuj (Mayan). *Canadian Journal of Linguistics* **64**, 62–101.

Kraft, Austin. to appear. Pseudoclefts in Simpakng (Southern Land Dayak). In *The proceedings of AFLA 32*.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*, ed. Y. Otsu, 1–25.

Kroeber, Paul D. 1991. Comparative syntax of subordination in Salish. Doctoral Dissertation, University of Chicago.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* **39**, 463–516.

Launey, Michel. 1994. *Une grammaire omniprédicative: Essai sur la morphosyntaxe du nahuatl classique*. Paris: CNRS éditions.

Launey, Michel. 2004. The features of omnipredicativity in Classical Nahuatl. *STUF 57* **1**, 49–69.

Law, Danny. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya lowlands*. Amsterdam: John Benjamins.

Lee, Felicia. 2008. On the absence of quantificational determiners in San Lucas Quiavini Zapotec. In *Quantification: A cross-linguistic perspective*, ed. Lisa Matthewson, 353–383. Leiden: Brill.

Lee-Schoenfeld, Vera, and Dennis Ott. 2021. *Parameters of Predicate Fronting*. Oxford: Oxford University Press.

Leu, Thomas. 2015. *The architecture of determiners*. Oxford University Press.

Little, Carol Rose. 2020. Left branch extraction, object shift, and freezing effects in Tumbalá Ch'ol. *Glossa: A journal of general linguistics* **5**.

Little, Carol Rose, Scott AnderBois, and Jessica Coon. 2024. Type-shifting in headless relative clauses. *Natural Language & Linguistic Theory* **42**, 1083–1123.

López Márquez, Wendy. 2021. Headless relative clauses in Sierra Popoluca. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 475–508. Oxford University Press.

Lyskawa, Paulina, and Rodrigo Ranero. 2022. Optional agreement as succesful/failed Agree: Evidence from Santiago Tz'utujil (Mayan). *Linguistic Variation* **22(2)**, 209–267.

Mateo Pedro, Pedro, and Jessica Coon. 2018. Chuj oral tradition: Collection of Pedro Mateo Pedro and Jessica Coon. Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Accessed 2018-2022.

Mateo Toledo, Eladio. 2012. Secondary predication in Q'anjob'al (Maya): Structure and semantic types. *International Journal of American Linguistics* **78**, 139–174.

Mateo Toledo, Eladio. 2023. *Complex predicates in Q'anjob'al*. Brill.

Matthewson, Lisa. 1998. *Determiner systems and quantificational strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.

Matthewson, Lisa. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* **9**, 145–189.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* **70**, 369–415.

Maxwell, Judith. 1982. How to talk to people who talk *chekel* 'different': The Chuj (Mayan) solution. Doctoral Dissertation, University of Chicago, Chicago, IL.

Maxwell, Judith. 1987. Some aspects of Chuj discourse. *Anthropological Linguistics* **29**, 489–506.

Maxwell, Judith. 1993. Chuj. In *Encyclopedia of world cultures: Middle America and the Caribbean*, ed. David Levinson, James W. Dow, and Robert Van Kemper, volume VIII, 70–74. New York, NY: G. K. Hall Company.

Maxwell, Judith M. 1976. Chuj intransitives: Or when can an intransitive verb take an object? *Journal of Mayan Linguistics* **1**, 128–140.

McCloskey, James. 1979. *Transformational syntax and model theoretic semantics: A case study in Modern Irish*. Dordrecht, Holland: Reidel.

Mendes, Gesoel, and Rodrigo Ranero. 2021. Chain reduction via Substitution: Evidence from Mayan. *Glossa* **6**.

Oda, Kenji. 2002. Wh-questions in V-initial languages. Master's thesis, University of Toronto.

Osorio May, José del Carmen. 2016. Temas de la sintaxis del yokt'an "chontal" de Tecoluta, Nacajuca, Tabasco. Doctoral Dissertation, Centro de Investigaciones y Estudios Superiores en Antropología Social.

Partee, Barbara H. 1975. Comments on C. J. Fillmore's and N. Chomsky's papers. In *The scope of American linguistics: Papers of the 1st Golden Anniversary of the Linguistic Society of America*, ed. Robert Austerlitz, 197–209. Lisse: Peter de Ridder Press.

Paul, Ileana, and Eric Potsdam. 2012. Sluicing without wh-movement in Malagasy. In *Sluicing: Cross-Linguistic Perspectives*, ed. Jason Merchant and Andrew Simpson, 164–182. Oxford: Oxford University Press.

Polian, Gilles. 2013. *Gramática del tseltal de Oxchuc*. Publicaciones de la Casa Chata. México: Centro de Investigaciones y Estudios Superiores en Antropología Social.

Polian, Gilles, and Judith Aissen. 2021. Headless relative clauses in Tseltalan. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 403–443. Oxford University Press.

Polinsky, Maria, and Eric Potsdam. 2011. Austronesian Syntax. In *Topics in Oceanic morphosyntax*, ed. Claire Moyse-Faurie and Joachim Sabel, 107–134. Berlin: de Gruyter Mouton.

Potsdam, Eric. 2009. Austronesian verb-initial languages and *wh*-question strategies. *Natural Language & Linguistic Theory* **27**, 737–771.

Ranero, Rodrigo. 2021. Identity conditions on ellipsis. Doctoral Dissertation, University of Maryland.

Ranero, Rodrigo, and Justin Royer. 2024. Deep and surface anaphora: A Mayan reappraisal. In *NELS 54: Proceedings of the Fifty forth Annual Meeting of the North East Linguistic Society*, ed. Shaunak Phadnis, Carla Spellerberg, and Brynne Wilkinson, 127–136.

Ranero, Rodrigo, Justin Royer, and Diego Neper López García. 2026. Split ergativity isn't always epiphenomenal. In *NELS 56: Proceedings of the Fifty sixth Annual Meeting of the North East Linguistic Society*.

Reeve, Matthew. 2012. *Celfts and their relatives*. Amsterdam: John Benjamins Publishing Company.

Royer, Justin. 2021. Headless relative clauses in Chuj. In *Headless relative clauses in languages of Mesoamerica*, ed. Ivano Caponigro, Harold Torrence, and Roberto Zavala, 327–361. Oxford, UK: Oxford University Press.

Royer, Justin. 2022a. Decomposing definiteness in Chuj. *Canadian Journal of Linguistics* **67**, 383–418.

Royer, Justin. 2022b. Elements of (in)definiteness and binding: A Mayan perspective. Doctoral Dissertation, McGill University.

Royer, Justin. 2022c. Prosody as syntactic evidence: The view from Mayan. *Natural Language & Linguistic Theory* **40**, 239–284.

Royer, Justin. 2025. Binding and anti-cataphora in Mayan. *Linguistic Inquiry* **56**, 247–310.

Royer, Justin, and Cristina Buenrostro. 2025a. Demostrativos nominales en Chuj de San Mateo Ixtatán. *Anales de Antropología* **59**, 115–138.

Royer, Justin, and Cristina Buenrostro. 2025b. Tres tipos de cuantificadores en Chuj de San Mateo Ixtatán. *Lingüística Mexicana* **7**, 7–40.

Royer, Justin, Cristina Buenrostro, and Peter Jenks. 2025a. The syntax of quantifiers in Chuj. *To appear in the Canadian Journal of Linguistics* .

Royer, Justin, and Jessica Coon. 2025. Correlates of object raising in Mayan. *Language and Linguistic Compass* **19**, 1–12.

Royer, Justin, Pedro Mateo Pedro, Elizabeth Carolan, Jessica Coon, and Matal Torres. 2022. *A*tz'am k'ik' atz'am: The story of Xuwan and a grammatical sketch of Chuj. *Tlalocan* **27**, 215–286.

Royer, Justin, Rodrigo Ranero, Petul Federico Felipe Gómez, and Diego Neper López García. 2025b. El cuento de kokolom/chi'jolom: Bosquejo gramatical y comparación dialectal del idioma Chuj. *Tlalocan* **30**, 187–289.

Schachter, Paul. 1973. Focus and relativization. *Language* **49**, 19–46.

Smith-Stark, Thomas. 1988. Pied-piping con inversión en preguntas parciales. Ms. Colegio de México y Seminario de lenguas indígenas.

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

Thoms, Gary. 2025. Reassessing Oehrle effects: Evidence from Scottish Gaelic. In *Foundational approaches to Celtic linguistics*. Language Science Press.

Tonhauser, Judith. 2003. F-constructions in Yucatec Maya. In *Proceedings of semantics of under-represented languages of the Americas (sula) 2*, ed. Jan Anderssen, Paula Menéndez-Benito, and Adam Werle, 203–223. Amherst, MA: GLSA.

Tonhauser, Judith. 2007. Agent focus and voice in Yucatec Maya. In *Proceedings of the 39th Meeting of the Chicago Linguistics Society*, 540–558. Chicago, IL.

Vázquez Álvarez, Juan J. 2011. A grammar of Chol, a Mayan language. Doctoral Dissertation, University of Texas Austin.

Vázquez Álvarez, Juan Jesús, and Jessica Coon. 2020. Headless relative clauses in C'hol. In *Headless relative clauses in Mesoamerican languages*, ed. Ivano Caponigro, Harrold Torrence, and Roberto Zavala Maldonado, 362–402. Oxford: Oxford University Press.

Velleman, Leah Bridges. 2014. Focus and movement in a variety of K'ichee'. Doctoral Dissertation, University of Texas Austin.

Vergnaud, Jean-Roger. 1974. French relative clauses. Doctoral Dissertation, MIT.

Verhoeven, Elisabeth, and Stavros Skopeteas. 2015. Licensing focus constructions in Yucatec Maya. *International Journal of American Linguistics* **81**, 1–40.

Zavala, Roberto. 1992. *El Kanjobal de San Miguel Acatán*. México: UNAM.

A Negation as a diagnostic for nonverbal predication

Another diagnostic for nonverbal predicate status, one that involves negation, was not discussed above. To be concrete, there is a specific way to negate nonverbal predicates, and this negation indeed appears on predicative quantifiers like *tzijtum* 'many', but not on quantifiers like *jantak* 'many'. To illustrate, first consider the following examples without quantifiers:

(83) a. Manh-**wach'**-ok-laj winh.
NEG-well-IRR-NEG CLF.PRON
'He is not well.' (negated nonverbal predicate)

b. Maj-**in-y-il**-laj winh.
NEG.PFV-B1S-A3-see-NEG CLF.PRON
'He didn't see me.' (negated verbal predicate)

What we find is that the form of negation in Chuj depends on whether the negated element is a nonverbal or verbal predicate (Maxwell 1982, Buenrostro 1995). Negating nonverbal predicates (in San Mateo Ixtatán Chuj) requires the use of three morphemes: *manh-* prefixes to the predicate, irrealis *-ok* suffixes closest to the root of the predicate, and *-laj* suffixes in the outermost position (83a). Negated verbal predicates, on the other hand, do not show the irrealis marker; moreover, the negative prefix fuses with some aspect markers, leading to a different surface exponent (83b). Now, if quantifiers like *tzijtum* are nonverbal predicates, we expect they should be negated as such. This is the case: *tzijtum* can be negated, and when it is, we observe the negative morphology corresponding to nonverbal predicates (84). In contrast, it is not possible to negate *jantak* 'many' in this way (85).

(84) Manh-**tzijtum**-ok-laj heb' ix ix ix-w-il-a'.
NEG-MANY$_{\text{PRED}}$-IRR-NEG PL CLF woman PFV-A1S-see-TV
'I didn't see many women.'
(*Lit*: 'The women that I saw weren't many.')

(85) * Ix-w-il manh-**jantak**-ok-laj heb' ix ix.
PFV-A1S-see-TV NEG-MANY$_{\text{DET}}$-IRR-NEG PL CLF woman
Intended: 'I didn't see many women.'

We did not apply this diagnostic in the main body of the text since it is not clearly extendable to *wh*-items. As a baseline, it seems odd to negate *wh*-items altogether, even when used predicatively (??It's not who that arrived?/??It's not where that you went?). This was reflected in speaker judgments, which varied in reaction to examples like the following:

(86) % Manh-**mach**-ok-laj winh k'ayb'um?
NEG-who-IRR-NEG CLF teacher
Intended: 'Who isn't the teacher?'

However, when rejected, the resulting sentence that is offered as an alternative translation is one which is entirely consistent with our analysis:

(87) **Mach** [winh [manh -k'ayb'um- ok-laj]]?
who CLF.PRON NEG-teacher-IRR-NEG
'Who isn't the teacher?'
(*Lit*: 'Who is the one who is not a teacher?')

Under our analysis, there are two predicates in (87): a predicative *wh*-word, which takes a light-headed relative subject headed by a pronoun, which in turn is the subject of the nonverbal predicate (*not a teacher*) in the embedded clause. Such structures are exactly the type of structures we expect to find in languages employing the pseudocleft strategy for *wh*-questions.