Two types of numeral classifiers: Evidence from Shan, Ch'ol (and Chuj)

Mary Moroney*, Carol-Rose Little†, & Justin Royer†
*Cornell University, †McGill University
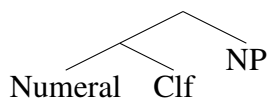
Semantics Babble at UC San Diego
April 6, 2021

. . . .
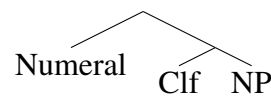
1 INTRODUCTION

What are numeral classifiers for?

- Answer 1: classifiers needed for numerals (Krifka 1995; Bale & Coon 2014); (1).

- Answer 2: classifiers needed for nouns (Chierchia 1998; Cheng & Sybesma 1999); (2).

(1) *Classifiers-for-numerals*   (2) *Classifiers-for-nouns*

- Numeral classifiers can be found in typologically diverse languages.

- For example, in Ch'ol (Mayan) and Shan (Tai) classifiers obligatorily appear with numerals:

(3) CH'OL

ux-**kojty**   ts'i'
three-CLF dog

'three dogs'[1]

(4) SHAN

*mǎa_i sǎam **tǒ**   t_i*
dog   three CLF

'three dogs'

ROADMAP

§2 Two types of theories for classifiers discussed in Bale et al. 2019, which we call '**classifier-for-numeral**' theories and '**classifier-for-noun**' theories;

§3 Background on classifiers in both Shan and Ch'ol

§4 Semantic predictions that the two theories make, situating Ch'ol and Shan w.r.t. to these predictions. Main claim: classifiers fall into two categories, supporting both theories and that the ***variation is in the numeral***

§5 Bring in data from Chuj (Mayan) and draw connections between its two types of classifiers and the classifier systems found in Ch'ol and Shan.

§6 Brief discussion on the mass/count distinction and measure constructions in Ch'ol and Shan

§7 Conclusions

---

[1] CLF = classifier; COMP = complementizer; DEM = demonstrative; IPFV = imperfective aspect; N.CLF = Chuj noun classifier; PROG = progressive aspect; #.CLF = Chuj numeral classifier

2 BACKGROUND

- Using the noun denotation in (6), we show how each theory derives the meaning of *two dogs* in a numeral-classifier language.[2]

(6) $[\![\text{DOGS}]\!] = \lambda x.[\text{DOGS}(x)] = \{a, b, c, ab, ac, bc, abc\}$

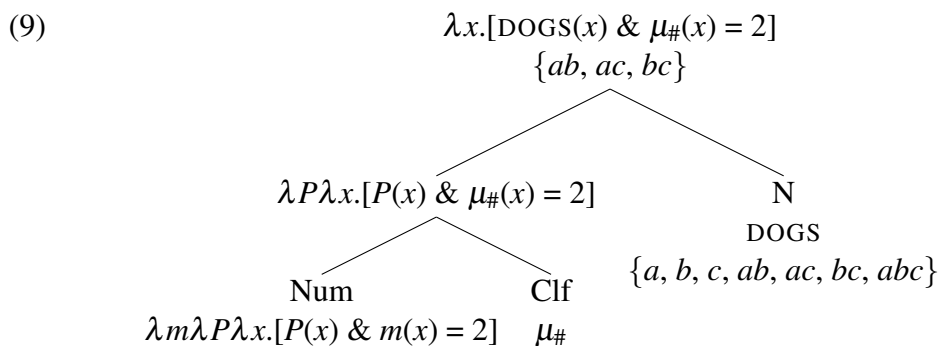- (6) denotes a set containing atomic and plural dog entities.

2.1 Classifier-for-numeral theories

- These accounts argue that classifiers are needed for numerals because the numeral requires an extra semantic argument in order to compose with the noun, as in (7) (Krifka 1995; Bale & Coon 2014; Bale et al. 2019; Hall 2019; Little & Winarto 2019).

- The classifier in (8) saturates the first argument of the numeral in (7), where $\mu_\#$ is a variable over measure functions.[3]

(7) $[\![\text{TWO}]\!] = \lambda m \lambda P \lambda x.[P(x) \ \& \ m(x) = 2]$ 　　　　(8) $[\![\text{CLF}]\!] = \mu_\#$

- (7) gives the set of individuals x such that for predicate P, x has the property of P and the measure of x is 2.

- (8) is a measure function which gives the number of atoms in a plurality x (Wilhelm 2008: 55).

- The numeral in (7) takes the classifier in (8) as its first argument.

- The noun combines directly with the numeral classifier, as in (9), to yield the set of groupings of two dogs:

(9)
$$\lambda x.[\text{DOGS}(x) \ \& \ \mu_\#(x) = 2]$$
$$\{ab, ac, bc\}$$

$\lambda P \lambda x.[P(x) \ \& \ \mu_\#(x) = 2]$ 　　　　　　　 N
　　　　　　　　　　　　　　　　　　　　　　　　DOGS
　　　　　　　　　　　　　　　　　　　$\{a, b, c, ab, ac, bc, abc\}$

Num 　　　　　　 Clf
$\lambda m \lambda P \lambda x.[P(x) \ \& \ m(x) = 2]$ 　 $\mu_\#$

---

[2] (6) could also be written:

(5) $[\![\text{DOG}]\!] = \lambda x.[^*\text{DOG}(x)]$

(5) gives the set of individuals in the complete join-semilattice formed from the atomic set of dogs (Link 1983).

[3] All of the types of classifiers discussed in this talk vary depending on properties of the noun (e.g., whether the noun is human, animal, etc.). For simplicity, we are not representing this in the semantics today.
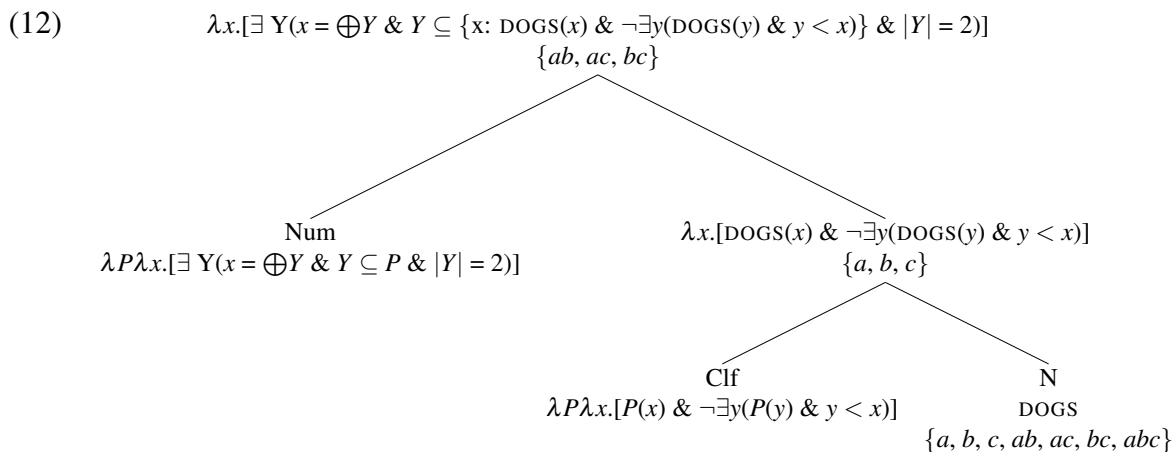
2.2 Classifier-for-noun theories

- These accounts argue that classifiers are needed for nouns to mediate between a noun, as in (6), and a numeral, as in (10) (Chierchia 1998; Cheng & Sybesma 1999; Moroney 2020).[4]

- We are assuming a classifier-for-noun theory following Bale et al. (2019).[5]

- The numeral needs to combine with an atomic predicate and returns the set containing all sums with the property that have a cardinality of 2:

(10) $[\![\text{TWO}]\!] = \lambda P \lambda x.[\exists\, Y(x = \bigoplus Y \,\&\, Y \subseteq P \,\&\, |Y| = 2)]$ (based on Bale et al. 2019)

- (10) denotes the x such that x is the sum of entities in each Y that is a subset of the property P that includes exactly two members.

- (10) measures the set. This is different from (7) which measures the entities in the set.

- Since the noun in (6) is not atomic, classifiers, like in (11), are needed to atomize the members in the set denoted by the NP predicate.

(11) $[\![\text{CLF}]\!] = \lambda P \lambda x.[P(x) \,\&\, \neg \exists y(P(y) \,\&\, y < x)]$ (Nomoto 2013; Bale et al. 2019)

- (11) gives the set of x, such that x has the property P and there is no y with the property P that is a subpart of x.

- The classifier in (11) first combines with a noun allowing for the numeral in (10) to then combine with the classifier-noun complex, as shown in (12).

(12)

$\lambda x.[\exists\, Y(x = \bigoplus Y \,\&\, Y \subseteq \{x: \text{DOGS}(x) \,\&\, \neg\exists y(\text{DOGS}(y) \,\&\, y < x)\} \,\&\, |Y| = 2)]$
$\{ab, ac, bc\}$

Num
$\lambda P \lambda x.[\exists\, Y(x = \bigoplus Y \,\&\, Y \subseteq P \,\&\, |Y| = 2)]$

$\lambda x.[\text{DOGS}(x) \,\&\, \neg\exists y(\text{DOGS}(y) \,\&\, y < x)]$
$\{a, b, c\}$

Clf
$\lambda P \lambda x.[P(x) \,\&\, \neg\exists y(P(y) \,\&\, y < x)]$

N
DOGS
$\{a, b, c, ab, ac, bc, abc\}$

In sum: Though derivationally distinct, each theory produces the same meaning for *two dogs*:

(i) For classifier-for-numeral theories, the numeral takes the classifier as a measure function, and then combines with the noun.

(ii) For classifier-for-noun theories, the numeral cannot directly combine with the noun, and so a classifier is needed to individuate the members of the nominal predicate into a set of atoms.

---

[4]$\bigoplus$ comes from Champollion & Esipova 2018, following Sharvy's (1980) analysis of definite descriptions. For $\bigoplus$ to apply to a set, it might be defined: $\lambda Q_{\langle e,t\rangle}.\iota x(\forall y, z(Q(y) \wedge Q(z) \wedge y \neq z \rightarrow y, z \leq x))$.

[5]In some classifier-for-noun theories, such as Chierchia 1998, the noun denotes a kind and the classifier mediates a type mismatch between the classifier and the noun. We assume $\langle e,t\rangle$ type denotations for nouns.

3   BACKGROUND ON CLASSIFIERS IN CH'OL AND SHAN

3.1   Ch'ol classifiers

- Ch'ol is a Mayan language of the Ch'olan-Tseltalan branch, spoken in southern Mexico by approximately 252,000 speakers.

  - There are three main dialects of Ch'ol—Tumbalá, Tila and Sabanilla.

  - The data in this paper comes from the Tumbalá and Tila dialects.

  - No dialectal differences were recorded with respect to the data in this paper.

- Numerals obligatorily appear with classifiers; numerals 1–20 given in Table 1.

  - Numerals 1–19 are given with the generic classifier *-p'ej*.

  - As the Ch'ol numerical system is base-twenty, *-k'al* fills the classifier slot.

  - *-K'al* is a classifier for the base-twenty root 'twenty'. Other classifiers for multiples of twenty include *-bajk* 'four hundred' and *-pijk* 'eight thousand'.

Table 1: Ch'ol numerals (Arcos López 2009: 24)

1	*jum-p'ej*	11	*juñlujum-p'ej*
2	*cha'-p'ej*	12	*lajchäm-p'ej*
3	*ux-p'ej*	13	*uxlujum-p'ej*
4	*chäm-p'ej*	14	*chäñlujum-p'ej*
5	*jo'-p'ej*	15	*jo'lujum-p'ej*
6	*wäk-p'ej*	16	*wäklujum-p'ej*
7	*wuk-p'ej*	17	*wuklujum-p'ej*
8	*waxäk-p'ej*	18	*waxäklujum-p'ej*
9	*bolom-p'ej*	19	*bolomlujum-p'ej*
10	*lujum-p'ej*	20	*juñ-k'al*

- Today, many speakers, including monolinguals, only use Ch'ol numerals up to six and numerals borrowed from Spanish for higher numerals (Vázquez Álvarez 2011: 160).

- As noted in Bale & Coon 2014, the Spanish-based numerals are ungrammatical with classifiers, as seen in (13).

(13)   a.   *ocho(*-p'ej) ja'as*
       SP:eight-CLF banana
       'eight bananas'

   b.   *nuebe(*-p'ej) tyumuty*
        SP:nine-CLF   egg
        'nine eggs'   (Ch'ol)

- Arcos López (2009) identifies at least 180 classifiers, though he notes that this is not an exhaustive list. Examples of common classifiers are found in Table 2.

Table 2: Numeral classifiers in Ch'ol

Form	Used to count	Examples	Translation
-p'ej	Inanimate/generic	*ux-**pe'jl** juñ*	'three books'
-kojty	Animals	*ux-**kojty** mis*	'three cats'
-tyikil	People	*ux-**tyikil** x'ixik*	'three women'
-k'ejl	Flat round objects	*ux-**k'ejl** waj*	'three tortillas'
-ts'ijty	Long things	*ux-**ts'ijty** tye'*	'three trees'
-bujch	Seated/propped up things	*ux-**bujch** bux*	'three (propped up) bottles'

- While the generic classifier *-p'ej* and the classifier for humans *-tyikil* are of unknown origin (Arcos López 2009), many other classifiers are derived from positional and transitive verb roots (see Arcos López 2009 and Bale et al. 2019 and also Haviland 1981 for Tsotsil).

- Positional roots are large and distinct class of roots that convey information about the position or configuration of an object (see e.g. England 1983, Haviland 1994, and Henderson 2019).

  - The numeral classifier *-bujch*, used to count things which are things propped up, seated, or leaning against something, is derived from the positional root *buch* 'seated'.

  - Animal classifier *-kojty*, is derived from the positional root *koty* 'standing on four legs'.

- The position or shape of a noun is relevant for the choice of classifier, meaning that the same noun could be counted with more than one classifier:

(14) a. *juñ-ts'ijty tye'*
         one-CLF   tree
         'one long tree'

      b. *juñ-jäjl tye'*
         one-CLF tree
         'stretched out tree'

      c. *juñ-bujch tye'*
         one-CLF tree
         'one fallen tree'

      d. *juñ-bujñ tye'*
         one-CLF tree
         'one fat tree'   (Ch'ol)

- Finally, numeral classifiers in Ch'ol only occur with numerals and the interrogative quantifier *jay-* 'how many', which we take to be the interrogative version of a numeral.

  - Ungrammatical with other quantifiers (15a), demonstratives (15b), or modifiers (15c).

(15) a. * *kabäl-k'ej waj*
         many-CLF tortilla
         Intended: 'many tortillas'

      b. * *ixä-kojty ts'i'*
         DEM-CLF dog
         Intended: 'that dog'

      c. * *säsäk-kojty ak'ach*
         white-CLF turkey
         Intended: 'a white   turkey'
         (Ch'ol)

3.2 Background on Shan.

- Shan is a Tai-Kadai language of the Southwestern Tai branch, spoken in Myanmar and surrounding countries by approximately 3 million speakers (Lewis & Fennig 2016).

- While classifiers in Thai, a related Tai language, have been investigated in detail (e.g., Iwasaki & Ingkaphirom 2005; Piriyawiboon 2010; Jenks 2011), there have been very few descriptions or analyses of Shan numeral classifiers.

- Cushing (1887) first identified classifiers in Shan, calling them 'numeral auxiliaries', which denotes "some rank of being, some form of object or some quality in the noun to which it belongs".

- The numerals in Shan from one to twenty can be seen in Table 3.

Table 3: Shan numerals

1	*nɯɯŋ*	11	*síp ʔét*
2	*sɔ̌ŋ*	12	*síp sɔ̌ŋ*
3	*sǎam*	13	*síp sǎam*
4	*sì*	14	*síp sì*
5	*haa*	15	*síp haa*
6	*hók*	16	*síp hók*
7	*tsét*	17	*síp tsét*
8	*pɛ̀t*	18	*síp pɛ̀t*
9	*kaw*	19	*síp kaw*
10	*síp*	20	*sáaw*

- Unlike Ch'ol numeral classifiers, which are derived from verb-like elements, Shan classifiers appear to be derived from nominal elements.

- For example, *tǒ*, the classifier for animals, also means 'body'.

- For some nouns that are typically found as compounds, there is a connection between the form of the classifier and one part of the compound, as in (16).

- Here the noun compound *ton-mâj* 'tree' and the classifier *ton* both contain *ton*.

- Classifiers obligatorily appear in the presence of a numeral, as shown in (17).

- Which classifier appears depends on properties of the nominal referent. Table 4 gives some examples of properties that determine which classifier is used.

(16) **ton**-*mâj* *sǎam* **ton**
plant-tree three plant
'three trees' (Shan)

(17) *mǎa sǎam *(tǒ)*
dog three CLF
'three dogs' (Shan)

Table 4: Basic numeral classifiers in Shan

Form	Used to count	Examples	Translation
ʔăn	Inanimates	*tsɔ̂ săam* **ʔăn**	'three forks'
tŏ	Animals	*méw săam* **tŏ**	'three cats'
kɔ̂	People	*kón săam* **kɔ̂**	'three people'
hòj	Round objects	*màak-khɔ̌ săam* **hòj**	'three jujube'
ton	Plants, trees	*ton-mâj săam* **ton**	'three trees'
lăŋ	Buildings	*hɤ́n săam* **lăŋ**	'three houses'

- One interrogative numeral, *lăaj* 'how many' in (18), obligatorily appear with classifiers, just like the numerals.

- (19), a similar interrogative, cannot appear with classifiers.

- As the contrast between (18) and (19) shows, even interrogative quantifiers of quantity, such as *lăaj* 'how many' and *kaa-hŭɪ* 'how many' differ in whether they must appear with a classifier.

- (20) is another expression of quantity that does not appear with a classifier.

- Moreover, the word *táŋ* 'all' can appear in addition to a numeral phrase, indicating that this word is in a different syntactic position than the numeral.

(18) *măa lăaj*   **tŏ**
   dog how.many CLF
     'How many dogs?'

(19) *măa kaa-hŭɪ*   *(*tŏ)*
     dog much-how
     'How many dogs?'

(20) *măa táŋ mót*
     dog all wholly
     'all dogs'

(21) *măa táŋ săam* **tŏ**
     dog all three CLF
   'all three dogs'   (Shan)

- The basic word order of the extended nominal domain in Shan is Noun-Numeral-Classifier.

- Simpson (2005) proposed that this word order in Thai, another Southwestern Tai language, comes as a result of the noun moving from its base position to a position above the numeral and classifier.

- This is the same structure argued for by Jenks (2011) for Thai.

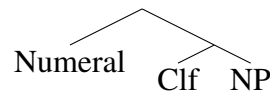- We assume the same movement happens in Shan, as schematized in (22).

(22)
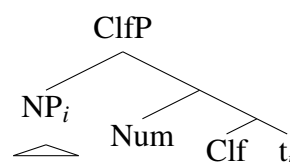
4 TWO TYPES OF CLASSIFIERS

(23) *Classifiers-for-numerals*

```
        /\
       /  \
      /    NP
  Numeral  Clf
```
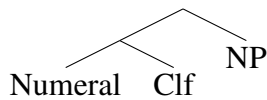
(24) *Classifiers-for-nouns*

```
        /\
       /  \
  Numeral  /\
          Clf NP
```

Despite producing similar meanings, the two theories make different predictions:

PREDICTIONS   (spelled out in more detail in §4.2)

If a classifier first forms a constituent with the...

1. NUMERAL, we might expect to find idiosyncrasies in whether or not a numeral requires a classifier, as argued in Bale & Coon 2014.
2. NOUN, we might expect to find idiosyncrasies in whether or not a noun requires a classifier, as argued in Simpson 2005 and Simpson & Ngo 2018.

3. NOUN, we might expect to find it with the noun in places other than with numerals.
4. NUMERAL, we might expect to find it with the numeral when it is not combining with a noun. This point, to our knowledge, has not been observed before in the literature on classifiers.

We compare two unrelated languages, Ch'ol and Shan, which have been described as having numeral classifiers, and show that while Ch'ol shows evidence for predictions 1 and 4, Shan shows evidence for predictions 2 and 3, supporting our proposal in (25):

(25) PROPOSAL
      We argue that there are two types of numeral classifiers across languages: *classifiers-for-numerals* (CLF-for-NUM) and *classifiers-for-nouns* (CLF-for-N)—and the difference lies in the semantics for the *numeral*

4.1 Quick review

- In (26), Ch'ol classifiers are suffixed to numerals; numerals are *pre*nominal

(26) Ch'ol nominal structure 'two dogs'

cha'	*-koty*	*ts'i'*
two	CLF.ANIMAL	dog
Numeral	Classifier suffix	Noun

- In (27), Shan classifiers follows the numeral; numerals are *post*nominal

(27) Shan nominal structure 'two dogs'

măa	*sɔ̌ŋ*	*tǒ*
dog	two	CLF.ANIMAL
Noun	Numeral	Classifier

4.2 Predictions

> PREDICTION 1 (CLF-for-NUM): If a classifier is a measure function required by a numeral, there might be idiosyncrasies in whether or not a numeral requires a classifier (some numerals might have the measure function in their lexical semantics, others not).

- This is the case in Ch'ol, as shown in Bale & Coon 2014. Mayan-based numerals require a classifier, whereas those borrowed from Spanish do not:

(28) a. *ux*(-kojty) ts'i'*
      three-**CLF** dog
      'three dogs'

    b. *ocho(*-kojty) ts'i'*
      SP:eight-**CLF** dog
      'eight dogs'   (Ch'ol)

- No such idiosyncrasies are found in Shan.

> PREDICTION 2 (CLF-for-N): If a classifier is used to create an atomic set from the noun predicate, there might be idiosyncrasies in whether or not a noun must combine with it (e.g. some nouns might only denote a set of atoms), as argued in Simpson 2005 and Simpson & Ngo 2018 for Vietnamese and other East/Southeast Asian languages.

- In Shan, some nouns do not need to combine with a classifier:

(29) *săam wán*
     three day
     'three days'

(30) *săam mɤ́ŋ*
     three country
     'three countries'   (Shan)

- This is different from Ch'ol, which always requires classifiers to combine with numerals, irrespective of the noun.

> PREDICTION 3 (CLF-for-N): If a classifier is used to create an atomic set from the noun predicate, we might expect to find it in environments other than with numerals.

- This is the case in Shan, which can have a classifier occur with quantifiers (31), with demonstratives (32), and with relative clauses (33), even in the absence of a numeral.

- In (31), the quantifier functions like a numeral. In (32)-(33), the classifier atomizes the noun, giving rise to a singular interpretation.

(31) *măa ku tŏ*
     dog every CLF
     'every dog'

(32) *măa tŏ nâj*
     dog CLF DEM
     'this dog'

(33) *măa tŏ [RC ʔăn nɔ́n jù]*
     dog CLF   COMP sleep IPFV
     'the dog that is sleeping'   (Shan)

- This is not true in Ch'ol. Classifiers only ever occur in the presence of a numeral (or with the quantifier *jay-* 'how many').

> PREDICTION 4 (CLF-for-NUM): If a classifier is a measure function required by a numeral, we would expect it to *always* appear with that numeral. This point, to our knowledge, has not been observed before in the literature on classifiers.

- This is the case in Ch'ol: classifiers are *always* required, even when counting (34) and referring to directly to the number (35).

(34) CONTEXT: Students are practicing counting.
jum-*(p'ej), cha'-*(p'ej), ux-*(p'ej) ...
one-CLF   two-CLF   three-CLF
'1, 2, 3'   (Ch'ol)

(35) CONTEXT: A teacher is pointing at the number three and says:
Ili   jiñ   ux-*(p'ej).
this DET three-CLF

'This is three.'   (Ch'ol)

- In contrast, Shan classifiers are not *always* required with numerals. They cannot appear with numeral when counting (36) or referring to the number (37).

(36) CONTEXT: Students are practicing counting.
nɯɯŋ (?tǒ), sɔ̆ŋ (?tǒ), sǎam (?tǒ) ...
one CLF   two CLF   three CLF
'1, 2, 3'   (Shan)

(37) CONTEXT: A teacher is pointing at the number three and says:
nâj pěn mǎaj   sǎam (*tǒ) .
this COP number three CLF
'This is the number three.'   (Shan)

In sum: While Ch'ol shows evidence for the classifier-for-numeral theories, Shan shows evidence for the classifier-for-noun theories.

Table 5: Summary

		Ch'ol	Shan
CLF-for-NUM	Prediction 1	✓	✗
CLF-for-N	Prediction 2	✗	✓
CLF-for-N	Prediction 3	✗	✓
CLF-for-NUM	Prediction 4	✓	✗

Conclusion: There is evidence that both kinds of classifiers exist, as proposed in (25).

5   TWO KINDS OF CLASSIFIERS IN ONE LANGUAGE? THE CASE OF CHUJ

> QUESTION: Can a language have both classifier types at the same time?

Chuj is a Mayan language of the Q'anjob'alan branch, spoken in Guatemala and Mexico by $\approx$ 70,000 speakers (Piedrasanta 2009, Buenrostro 2013). It has two types of classifiers we call (following the literature) "numeral classifiers" (#.CLF) and "noun classifiers" (N.CLF):

- They can co-occur, but N.CLFs are always optional when they occur with numerals (see Craig 1986, Buenrostro et al. 1989, Zavala 2000, Hopkins 1970, 2012).

(38) *ox-wanh   (nok') tz'i'*
      three-#.CLF N.CLF dog
      'three dogs'   (Chuj)

- Chuj's #.CLFs pattern with Ch'ol CLF-for-NUMs, and not with CLF-for-Ns (see appendix).

What about Chuj's "noun classifiers"?

Similar syntactic distribution: Chuj *noun classifiers*, which mark definiteness and specificity (see Buenrostro et al. 1989, García Pablo & Domingo Pascual 2007, Royer 2019), pattern like Shan and unlike Ch'ol classifiers in their syntactic distribution:
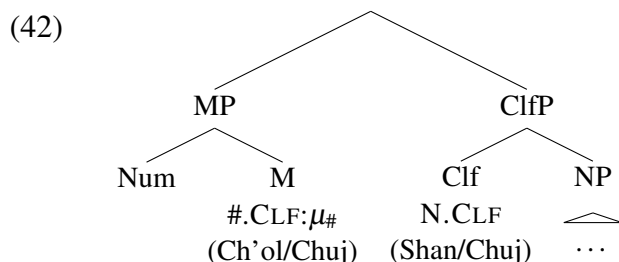
- Prediction 2 (CLF-for-N): Like Shan, not all Chuj nouns can combine with a noun classifier, such as abstract nouns like *ib'* 'strength'.

- Prediction 3 (CLF-for-N): Appear in environments other than with numerals, such as alone with nouns (39), with demonstratives (40), and with relative clauses (41):

(39) ***nok'** tz'i'*   (40) ***nok'** tz'i' **chi***
      CLF dog   CLF dog that
      'the dog'   'that dog'

(41) *nok' tz'i' [_{RC} **nok'** lan   s-way-i   ]*
      CLF dog   CLF PROG A3-sleep-IV
      'the dog that is sleeping.'   (Chuj; cf. similar Shan examples above)

- Perhaps Chuj N.CLFs share the same syntactic position as Shan classifiers, and Chuj #.CLFs share the same syntactic position as Ch'ol classifiers:

(42)

```
                          MP                    ClfP
                       /      \              /       \
                    Num        M          Clf        NP
                              #.CLF:μ#    N.CLF       /\
                            (Ch'ol/Chuj) (Shan/Chuj)  ···
```
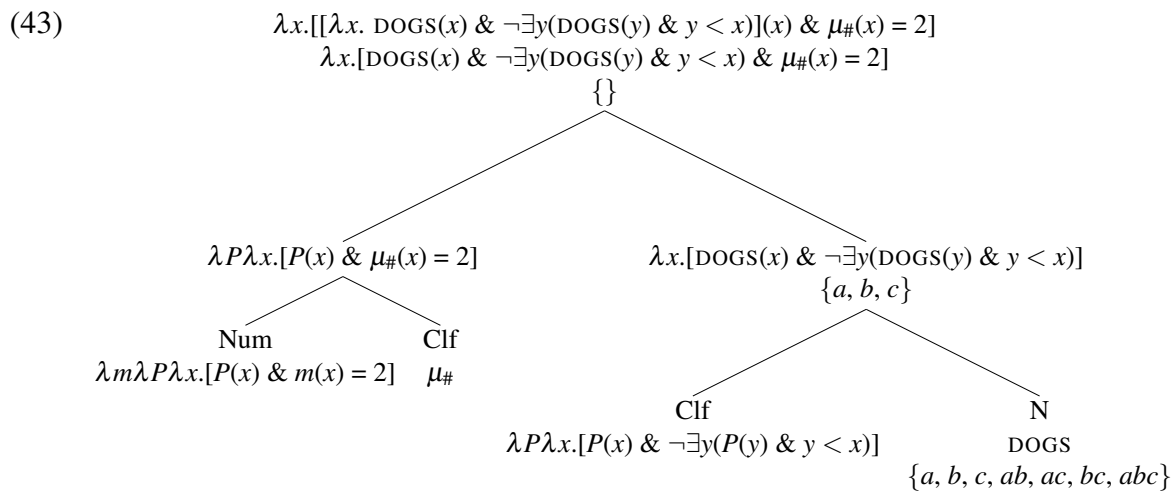
- The structure in (42) follows Bale et al. (2019) in positing a Measure phrase (MP) that contains the numeral and CLF-for-NUM.

- And follows Cheng & Sybesma (1999), Simpson (2005), a.o. in positing a Classifier Phrase (ClfP) that takes the nominal as complement.

Different semantics: But even though Chuj noun classifiers seem to pattern with Shan classifiers, they cannot fulfill the exact same semantic function.

- Crucially, unlike Shan classifiers, Chuj's *noun* classifiers are **never** required in presence of numerals (only Chuj's *numeral* classifiers are).

- Therefore, it seems that while Chuj has Ch'ol-style numeral classifiers, the noun classifier is semantically distinct from the Shan-style classifier.

What would it mean semantically for a language to have both types of classifiers?

Answer: The result would be semantically ill-formed. CLFs-for-Ns generate a set of atoms, so when the CLF+N combines with the numeral, there are no pluralities in the set to measure.

(43)

$$\lambda x.[[\lambda x. \text{DOGS}(x) \ \& \ \neg\exists y(\text{DOGS}(y) \ \& \ y < x)](x) \ \& \ \mu_\#(x) = 2]$$
$$\lambda x.[\text{DOGS}(x) \ \& \ \neg\exists y(\text{DOGS}(y) \ \& \ y < x) \ \& \ \mu_\#(x) = 2]$$
$$\{\}$$

$\lambda P \lambda x.[P(x) \ \& \ \mu_\#(x) = 2]$ $\lambda x.[\text{DOGS}(x) \ \& \ \neg\exists y(\text{DOGS}(y) \ \& \ y < x)]$
$\{a, b, c\}$

Num   Clf
$\lambda m \lambda P \lambda x.[P(x) \ \& \ m(x) = 2]$ $\mu_\#$

Clf
$\lambda P \lambda x.[P(x) \ \& \ \neg\exists y(P(y) \ \& \ y < x)]$

N
DOGS
$\{a, b, c, ab, ac, bc, abc\}$

Result: We don't expect to find a language with the semantics of CLFs-for-NUMs (which measure pluralities) and CLFs-for-Ns (which atomize) at the same time.

But we might expect to find a language with exponents of both syntactic heads. Perhaps Chuj is this kind of language.

- The presuppositions associated with Chuj noun classifiers are shared among other cases of classifier-for-nouns across languages.

- For instance, it is well-known that Cantonese (Cheng & Sybesma 1999) and Vietnamese (Simpson 2005) classifiers (which are CLF-for-Ns) are associated with definiteness.

- This isn't surprising if Chuj noun classifiers and CLFs-for-Ns share a syntactic position.[6]

---

[6]Aikhenvald (2000) observed that in many Southeast Asian languages, noun classifiers appear to be a subtype of numeral classifiers, whereas in some Mayan languages, numeral and noun classifiers are separate lexical items. This fits with the current study.

6 A SHORT ASIDE: MASS/COUNT AND MEASURE EXPRESSIONS

- We have concentrated on sortal classifiers and constructions with count nouns

- Before concluding we'd like to briefly discussion measure constructions and mass nouns, focusing on container measure expressions, rather than standard measures

6.1 Ch'ol

- Ch'ol has both mass and count nouns; for instance distributive adjectives and numerals with the generic classifier are only grammatical with count nouns such as in (44) and not substance nouns in (45)

(44) a. *ch'och'ok mesa*
        small   table

        'a small table'

    b. *cha'-p'ej mesa*
        two-CLF table

        'two tables'

(45) a. * *ch'och'ok ch'ich'*
        small   blood

        intended 'small (unit of?) blood'

    b. * *cha'-p'ej ch'ich'*
        two-CLF blood

        'two bloods'

- Measure terms can appear in the same position as the numeral classifier in Ch'ol, as in (46a), though measure phrases may also be expressed in other ways
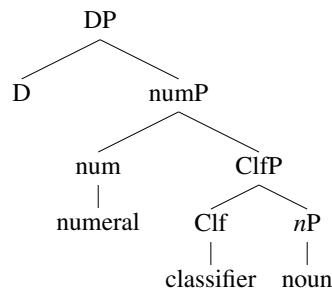
(46) a. *cha'-lujch ja'*
        two-CLF   water

        'two ladles of water'

    b. *cha'-p'ej lujch-ib ja'*
        two-CLF ladle-NML water

        'two ladles of water'

- Bale et al. (2019) has argued that in Ch'ol sortal and measure constructions have the same syntax, that in (47) where CLF can be a sortal or measure classifier (i.e., it could be *p'ej*, sortal, or *lujch*, measure)

(47) Measure structure
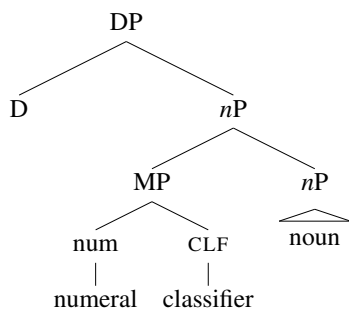
(48) Partition structure

- Sortal classifiers and measure classifiers appear in the same syntactic slot, but as per sortal classifiers, may only appear with Ch'ol-based numerals (PREDICTION 1)

(49) a. * *ocho-lujch ja'*
        SP:eight-CLF water

        Intended 'eight ladles of water'

    b. *ocho lujch-ib ja'*
        eight ladle-NML water

        'eight ladles of water'

6.2 Shan

- Measure terms in Shan are homophonous with the noun they are related to.

- They appear in the same position as numeral classifiers, as shown in (50a).

(50) a. *nâm săam kɔ́k*
      water three cup
      'three cups of water'

    b. *kɔ́k nâm sɔ̌ŋ kɔ́k*
      cup water two cup
      'two water cups'   (Shan)

- The head noun—*nâm* 'water' in (50a) or *kɔ́k nâm* 'water cup'/'cup of water' in (50b)—determines whether the expression refers to the container or contents.

- Shan also has both mass and count nouns, which can be demonstrated with distributive adjective data.

- For example, the distributive adjective *lêk* 'small' can combine with the count noun (51a) but not not substance nouns in (51b).

(51) a. *màakhŏ lêk*
      hat   small
      'a small hat'

    b. * *lɤt   lêk*
      blood small
      intended: 'a small (unit of) blood'

- Therefore, what we find is that mas nouns in Shan and Ch'ol, unlike Yudja (Lima 2014) and Nez Perce (Deal 2017), cannot directly combine with distributive adjectives, suggesting a clear mass/count distinction.

- Count nouns in Shan can combine with a generic classifier, as (51a) shows, but the noun specific classifier is better.

- The sortal (52a) and measure expressions (52b) have the same apparent syntax, with modifiers in the same position.

(52) a. *màakhŏ sɔ̌ŋ hòj   / ˀʔăn*
      hat   two CLF.RND / CLF.GEN
      'two hats'

    b. *lɤt   sɔ̌ŋ tɔ̀m /* Ăn*
      blood two drop / CLF.GEN
      'two drops / *units of blood'

- As shown above in (29), measure expressions of time can have the form NUM N, supporting prediction PREDICTION 2 for CLF-FOR-N languages.

(53) *lɤt   AB sɔ̌ŋ thŏŋ*
     blood AB two bag
     'two bags of AB blood'

(54) *lɤt   AB thŏŋ nân*
     blood AB bag that
     'that bag of AB blood'

- The container term *thŏŋ* 'bag' can be used with a numeral to measure a substance as in (53) or it can be used with a demonstrative as in (54)—in line with PREDICTION 3 for CLF-FOR-N languages.

6.3   Measure phrases and our predictions

- Given that there is a mass/count distinction, we need to investigate whether classifier expressions with mass nouns work any differently than classifier expressions with count nouns.

- Our analysis predicts that the numeral is responsible for the differences between these two languages

    - Since our analysis is based on the numeral and not the noun we do not necessarily predict a distinction between count and measure expressions in whether they function as classifiers-for-numerals or classifiers-for-nouns.

    - Bale et al. (2019) argue that measure and sortal classifier syntax are the same, so on that basis there is no reason to have a difference.

- Based on a preliminary survey, we find that the mass/count distinction does not affect count/measure expression with numeral-classifiers in either language, summarized in Table 6.

Table 6: Summary (preliminary)

		Ch'ol MEASURE	Shan MEASURE
CLF-for-NUM	Prediction 1	✓	✗
CLF-for-N	Prediction 2	✗	✓
CLF-for-N	Prediction 3	✗	✓
CLF-for-NUM	Prediction 4	✓	✗

- However, different syntactic structures and semantics have been proposed for Mandarin Chinese count and measure expressions by Li & Rothstein (2012).

    - The semantic analysis proposed to account for this contrast requires different numeral semantics for the count and measure interpretations. This is compatible with our analysis.

    - Additionally, they argue that nouns denoting substances can have either a count or measure interpretation, meaning that the noun denotation by itself does not determine whether the classifier is a CLF-for-NUM or CLF-for-N.

7 CONCLUSIONS

In this talk, we reviewed two types of theories on the cross-linguistic status of classifiers, as described in Bale et al. 2019:

1. Classifier-for-numeral theories (Krifka 1995, Bale & Coon 2014, Bale et al. 2019, Hall 2019, Little & Winarto 2019, a.o.)

2. Classifier-for-noun theories (Chierchia 1998, Simpson & Ngo 2018, Moroney 2020, a.o.)

- We concluded, based on predictions that each theory makes and on the distribution of classifiers in three languages from two language families (Ch'ol, Chuj, Shan), that both theories are needed.

- There are at least two kinds of classifiers that mediate between numerals and nouns across languages: *classifiers-for-numerals* and *classifiers-for-nouns*.

- Drawing on previous work, we've provided some diagnostics to distinguish between these different kinds of classifiers, including a new counting diagnostic.

- We also proposed that the semantics of numerals can vary across languages, and that this is connected to which classifier is needed:

  - *Classifiers-for-numerals* arise to saturate an extra argument of the numeral required to count; Such numerals measure (atomic and plural) entities denoted by the noun.

  - *Classifiers-for-nouns* arise because the numeral is set counting, and so the classifier is needed to extract the atoms out of the nominal predicate (and thus exclude plural entities).

- Finally, we have argued that there is a count/mass distinction in both Ch'ol and Shan and that preliminary results suggest that measure expressions in these languages are the same as other numeral-classifier expressions in being classifiers-for-numerals (Ch'ol) and classifiers-for-nouns (Shan).

- We conclude that there is *crosslinguistic variation in the syntax and semantics of numerals and classifiers*, which makes sense especially if we consider the diversity of counting systems found across the world's languages.

REFERENCES

AIKHENVALD, ALEXANDRA Y. 2000. *Classifiers: A typology of noun categorization devices: A typology of noun categorization devices*. OUP Oxford.

ARCOS LÓPEZ, NICOLÁS. 2009. *Los clasificadores numerales y las clases nominales en ch'ol*. Master's thesis, CIESAS, México.

BALE, ALAN, and JESSICA COON. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry*.

BALE, ALAN; JESSICA COON; and NICHOLÁS ARCOS. 2019. Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa: a journal of general linguistics* 4.

BUENROSTRO, CRISTINA. 2013. *La voz en Chuj de San Mateo Ixtatán*. Mexico City: El Colegio de México dissertation.

BUENROSTRO, CRISTINA; JOSÉ CARMEN DÍAZ; and ROBERTO ZAVALA. 1989. Sistema de clasificación nominal del Chuj. *Memorias del segundo coloquio internacional de Mayistas*, vol. II. Mexico City: UNAM.

CHAMPOLLION, LUCAS, and MARIA ESIPOVA. 2018. Integrating compositional semantics and event semantics. Lecture notes for the NASSLLI 2018 course, June 23-29, 2018.

CHENG, LISA LAI-SHEN, and RINT SYBESMA. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30.509–542.

CHIERCHIA, GENNARO. 1998. Reference to kinds across language. *Natural Language Semantics* 6.339–405.

CRAIG, COLETTE GRINEVALD. 1986. Jacaltec noun classifiers. *Lingua* 70.241–284.

CUSHING, JOSIAH NELSON. 1887. *Grammar of the Shan language*. American Baptist Mission Press.

DAYAL, VENEETA. 2012. Bangla classifiers: Mediating between kinds and objects. *Rivista di Linguistica* 24.195–226.

DEAL, AMY ROSE. 2017. Countability distinctions and semantic variation. *Natural Language Semantics* 25.125–171.

ENGLAND, NORA. 1983. *A grammar of Mam, a Mayan language*. University of Texas Press.

GARCÍA PABLO, GASPAR, and PASCUAL MARTÍN DOMINGO PASCUAL. 2007. *Stzolalil sloloni-spaxtini heb' chuj: Gramática descriptiva chuj*. Academia de Lenguas Mayas de Guatemala.

HALL, DAVID. 2019. Licensing D in classifier languages and "numeral blocking". *Definiteness across languages*, ed. by Ana Aguilar-Guevara, Julia Pozas Loyo, and Violata Vázquez-Rojas Maldonado, 221-257. Language Science Press.

HAVILAND, JOHN B. 1994. "Ta xa setel xulem" [The buzzers were circling] – categories of verbal roots in (Zinacantec) Tzotzil. *Linguistics* 32.691–741.

HAVILAND, JOHN BEARD. 1981. *Sk'op sotz'leb: El Tzotzil de San Lorenzo Zinacantán*. UNAM.

HENDERSON, ROBERT. 2019. The roots of measurement. *Glossa: A journal of general linguistics* 4.1–31.

HOPKINS, NICHOLAS. 2012. Noun classifiers of the Chuchumatán Mayan languages: A case of diffusion from Otomanguean. *International Journal of American Linguistics* 78.411–427.

HOPKINS, NICHOLAS A. 1970. Numeral classifiers in tzeltal, jacaltec and chuj (mayan). *Papers from the sixth regional meeting of the chicago linguistic society*, 23–35.

IWASAKI, SHOICHI, and PREEYA INGKAPHIROM. 2005. *A Reference Grammar of Thai*. Cambridge University Press.

JENKS, PETER. 2011. *The hidden structure of Thai noun phrases*. Harvard University dissertation.

KRIFKA, MANFRED. 1995. Common nouns: A contrastive analysis of English and Chinese. *The generic book*, ed. by Gregory Carlson and Francis Jeffry Pelletier, 398–411. Chicago: University of Chicago Press.

LEWIS, GARY F. SIMONS, M. PAUL, and CHARLES D. FENNIG. 2016. Online: http://www.ethnologue.com.

LI, XUPING, and SUSAN ROTHSTEIN. 2012. Measure readings of Mandarin classifier phrases and the particle de. *Language and Linguistics* 13.693.

LIMA, SUZI. 2014. *The Grammar of Individuation and Counting*. University of Massachusetts Amherst dissertation.

LINK, GODEHARD. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Formal semantics: The essential readings*, 127–146.

LITTLE, CAROL-ROSE, and EKARINA WINARTO. 2019. Classifiers and the definite article in Indonesian. *Proceedings of the 49th North East Linguistics Society*, ed. by Maggie Baird and Jonathan Pesetsky, vol. 2, 209–220. Amherst, MA: GLSA.

MORONEY, MARY. 2020. Taking the measure of the Shan plural morpheme. Poster at the Linguistic Society of America, New Orleans, LA.

NOMOTO, HIROKI. 2013. *Number in classifier languages*. University of Minnesota dissertation.

Piedrasanta, Ruth. 2009. *Los chuj, unidad y rupturas en su espacio*. Guatemala City, Guatemala: Amrar Editores.

Piriyawiboon, Nattaya. 2010. *Classifiers and determiner-less languages: The case of Thai*. University of Toronto Doctoral dissertation dissertation.

Royer, Justin. 2019. Domain restriction and noun classifiers in Chuj (Mayan). *Proceedings of NELS 49*, ed. by Maggie Baird and Jonathan Pesetsky.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The philosophical review* 89.607–624.

Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. *The Oxford handbook of comparative syntax*, 806–838.

Simpson, Andrew, and Binh Ngo. 2018. Classifier syntax in Vietnamese. *Journal of East Asian Linguistics* 27.211–246.

Vázquez Álvarez, Juan Jesús. 2011. *A grammar of Chol, a Mayan language*. Austin, TX: University of Texas Austin dissertation.

Wilhelm, Andrea. 2008. Bare nouns and number in Dëne Sųłiné. *Natural Language Semantics* 16.39–68.

Zavala, Roberto. 2000. Multiple classifier sytems in Akatek (Mayan). *Systems of nominal classification*, ed. by Gunter Senft, 114–146. Cambridge: Cambridge University Press.
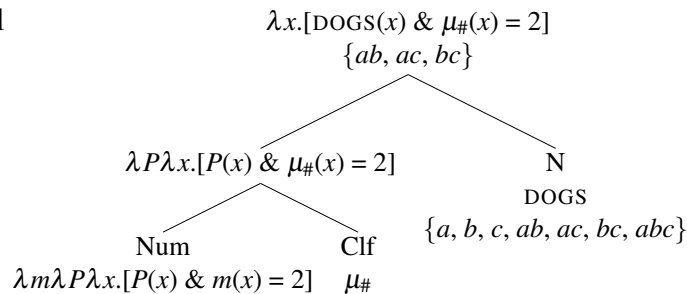
Contact for authors:
Carol-Rose Little: carolrlittle@gmail.com
Mary Moroney: mrm366@cornell.edu
Justin Royer: justin.royer@mail.mcgill.ca

A DERIVATIONS

(55) Ch'ol

$$\lambda x.[\text{DOGS}(x) \,\&\, \mu_{\#}(x) = 2]$$
$$\{ab, ac, bc\}$$

$\lambda P\lambda x.[P(x) \,\&\, \mu_{\#}(x) = 2]$

N
DOGS
$\{a, b, c, ab, ac, bc, abc\}$

Num
$\lambda m\lambda P\lambda x.[P(x) \,\&\, m(x) = 2]$

Clf
$\mu_{\#}$

(56) Shan

$$\lambda x.\,[\exists Y\,(x = \textstyle\bigoplus Y \,\&\, Y \subseteq \{x : \text{DOGS}(x) \,\&\, \neg\exists y(\text{DOGS}(y) \,\&\, y < x\}) \,\&\, |Y| = 2]$$
$$\{ab, ac, bc\}$$

Num
$\lambda P\lambda x.[\exists Y\,(x = \bigoplus Y \,\&\, Y \subseteq P \,\&\, |Y| = 2)]$

$\lambda x.[\text{DOGS}(x) \,\&\, \neg\exists y(\text{DOGS}(y) \,\&\, y < x)]$
$\{a, b, c\}$

Clf
$\lambda P\lambda x.[P(x) \,\&\, \neg\exists y(P(y) \,\&\, y < x)]$

N
DOGS
$\{a, b, c, ab, ac, bc, abc\}$
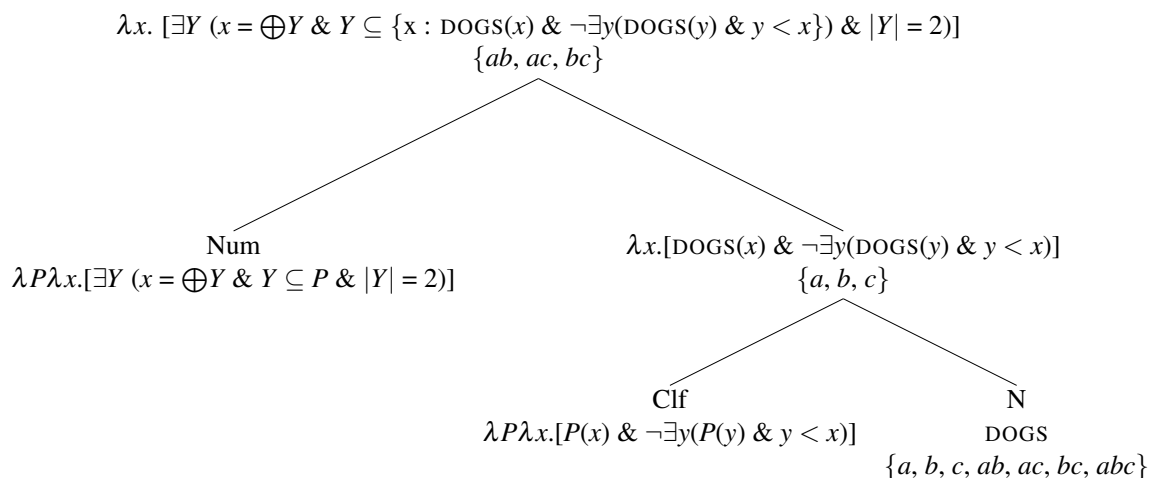
B CHUJ'S NUMERAL CLASSIFIERS AS CLASSIFIERS-FOR-NUMERALS

Chuj's numeral classifiers (#.CLF) pattern according to the classifier-for-numeral diagnostics established in §3.2 (i.e. like Ch'ol's classifiers).

- Prediction 1 (CLF-for-NUM): Obligatory when counting or referring directly to numeral:

(57) CONTEXT: Students are practicing counting.
ox-*(e'),   chanh-*(e'), hoy-*(e')
three-#.CLF, four-#.CLF, five-#.CLF

'3, 4, 5.'   (Chuj)

(58) CONTEXT: A teacher is pointing at the number three and says:
ha   jun tik   ox-*(e').
TOP one DEM three-#.CLF

'This is three.'   (Chuj)

- Prediction 2 (CLF-for-N): Some numerals can't combine with a numeral classifier, including Mayan-based *jun* 'one' and numerals borrowed from Spanish:

(59) *jun-(*e')* ... *wentiyuno-(*e'), wentitres-(*e')*
one-#.CLF   21-#.CLF,   22-#.CLF

'1 ... 21, 22.'   (Chuj)

C PREVIOUS SEMANTIC ANALYSES OF CLASSIFIERS

Table 7: Summary of previous semantic analyses of classifiers

	Noun type	Classifier type	N-at-Num	Num
Krifka 1995	e^k	$\langle n, \langle k, \langle e,t \rangle \rangle \rangle$	$\langle e,t \rangle$	n
Wilhelm 2008	$\langle e,t \rangle$	$\langle n, \langle \langle e,t \rangle, \langle e,t \rangle \rangle \rangle$	$\langle e,t \rangle$	n
Bale & Coon 2014	$\langle e_{AT},t \rangle$	μ	$\langle e_{AT},t \rangle$	$*$
Bale et al. 2019	$\langle e,t \rangle$	μ	$\langle e,t \rangle$	$|\text{ENTITY}|$
Chierchia 1998	e^k	$\langle k, \langle e,t \rangle \rangle$	$\langle e,t \rangle$	-
Jenks 2011	e^k	$\langle k, \langle e,t \rangle \rangle$	$\langle e_{AT},t \rangle$	-
Dayal 2012	e^k	$\langle k, \langle e,t \rangle \rangle$	$\langle e_{AT},t \rangle$	$|\text{SET}|$
Nomoto 2013	$\langle e,t \rangle$	$\langle \langle e,t \rangle, \langle e,t \rangle \rangle$	$\langle e_{AT},t \rangle$	$|\text{SET}|$