Classifying classifiers*

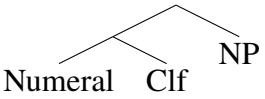
Carol-Rose Little*, Mary Moroney† & Justin Royer*

*McGill University, †Cornell University

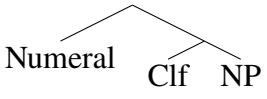
1. Introduction

Drawing on data from three typologically different languages, this paper investigates the semantic function of numeral classifiers. In the literature, there are two main families of proposals for numeral classifiers. The first argues that classifiers are needed for numerals: the classifier first combines with the numeral before combining with the noun (e.g. Krifka 1995, Bale & Coon 2014, Bale et al. 2019), exemplified by the structure in (1). This contrasts with the set of theories that argues that classifiers are needed for nouns, meaning that the classifier combines with the noun before combining with the numeral (e.g. Chierchia 1998, Cheng & Sybesma 1999), with the structure in (2). We will refer to theories represented by (1) as *classifier-for-numeral* theories and those represented by (2) as *classifier-for-noun* theories.

(1) *Classifiers-for-numerals*



(2) *Classifiers-for-nouns*



For instance, in both Ch'ol (Mayan) and Shan (Tai-Kadai) classifiers obligatorily appear with numerals, as in (3) and (4). Specifically, in the Ch'ol example in (3), the numeral

*Unless otherwise noted, all data comes from the authors' fieldwork: Carol-Rose, Ch'ol; Mary, Shan; Justin, Chuj. Thanks to all of our language consultants: the Arcos López family in San Miguel, Chiapas (Ch'ol) and Morelia Vázquez Martínez, *wokolix la'wälä!*; Nan San Hwam in Chiang Mai, Thailand and Sai Noom Hseng in Jacksonville, Florida (Shan) *khɔp tsǎɯ tê tê khaa!*; Matal Torres, Elsa Velasco Torres and Yun Torres in Yuxquen, Guatemala (Chuj) *Yuj wal yos!*. We would also like to thank Scott AnderBois, John Beavers, Jessica Coon, Aron Hirsch, Suzi Lima, Sarah Murray, the Cornell Semantics Group and audiences at the 2020 LSA meeting and SULA 11 for comments and discussion. Carol-Rose's work is supported by the National Science Foundation under grant no. BCS-1852744 and an Engaged Cornell graduate student research grant. Mary's work is supported by an Engaged Cornell graduate student research grant and a Ruchira Mendiones Research Grant through the Southeast Asia Program at Cornell. Justin's work is supported by a Graduate Mobility Award from McGill University's Graduate and Postdoctoral Studies.

Editor names (eds.), SULA 11: 1–15

classifier *-kojty* must accompany the numeral. The same is true of the Shan classifier *tŏ* in (4).

(3) ux-*(**kojty**) ts'i'
three-CLF dog
'three dogs'[1] (Ch'ol)

(4) măa$_i$ săam *(**tŏ**) t$_i$
dog three CLF
'three dogs' (Shan)

Given (3) and (4), it is not immediately clear how to make a principled choice between the two families of proposals in (1) and (2). Here, we investigate predictions of both families of theories and provide evidence that both are correct, but for different languages. We begin by providing some brief background on the different kinds of theories of numeral classifiers (Section 2). Drawing on data from Ch'ol and Shan, we propose in Section 3 that the main point of variation between the two sets of theories should lie in the denotation of the numeral. We argue that Ch'ol classifiers are classifiers-for-numerals and Shan classifiers are classifiers-for-nouns. In Section 4, we bring in data from Chuj (Mayan) which has two types of classifiers, drawing connections with the classifier systems of Ch'ol and Shan.[2]

2. Theoretical background and assumptions

In this section, we briefly review previous literature on semantic analyses of numeral classifiers, which have received various implementations in the literature. We sketch a particular implementation of the distinction between (1) and (2), which hardwires the distinction in the denotation of the numeral. The classifier denotation consequently differs. The denotation of the noun remains the same: we assume for sake of simplicity that the noun is always type $\langle e,t\rangle$. Nouns denote a set containing atomic entities and their sums, as in (5), assuming a denotation for DOGS that contains only three atomic entities.[3] Below, we demonstrate how each family of theories derives 'two CLF dogs'.

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

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

[2] For reasons of space, we focus on so-called sortal numeral classifiers, in contrast to mensurative numeral classifiers. Some analyses, such as Rothstein 2017 and Bale et al. 2019, discuss possible distinctions between sortal versus measure classifier constructions. For instance, Li & Rothstein (2012) have proposed that Mandarin has the structure in (1) for count constructions and the structure in (2) for measure constructions. Though the focus in this paper is on sortal numeral classifiers, the analysis proposed here is not incompatible with the possibility that mensurative numeral classifiers exhibit a different syntax and semantics.

[3] In some classifier-for-noun theories (e.g. Chierchia 1998) the noun denotes a kind and the classifier mediates a type mismatch between the classifier and the noun. Here we assume an ⟨e,t⟩ denotation for nouns. Also note that (5) could also be written as in (1), with the "star-operator", which gives the set of individuals in the complete join-semilattice formed from the atomic set of dogs (Link 1983).

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

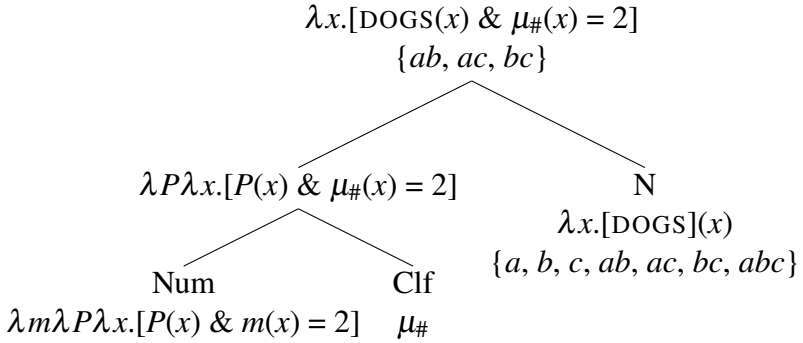
2.1 Classifier-for-numeral theories

Though the implementations vary slightly, classifier-for-numeral theories argue that classifiers are needed for numerals because the numeral in (6) requires an extra semantic argument in order to compose with the noun (Krifka 1995, Wilhelm 2008, Bale & Coon 2014, Bale et al. 2019). The classifier in (7) saturates the first argument of the numeral in (6), where $\mu_{\#}$ is a variable over measure functions. The numeral in (6) denotes the set of individuals x such that for the predicate P, x has the property of P and the measure of x is 2. (7) is a measure function which gives the number of atoms in a plurality x (Wilhelm 2008: 55). As schematized in (8), the noun can directly combine with the numeral-classifier constituent to yield the set of groupings of two dogs.

(6) $[\![2]\!] = \lambda m \lambda P \lambda x.[P(x) \,\&\, m(x) = 2]$

(7) $[\![\text{CLF}]\!] = \mu_{\#}$

(8)



2.2 Classifier-for-noun theories

As with classifier-for-numeral theories, there are different implementations of classifier-for-noun theories. Nevertheless, all argue that classifiers are needed for nouns to mediate between the noun and the numeral (Chierchia 1998, Dayal 2012, Nomoto 2013). For illustration, we provide a version of the classifier-for-noun theory as discussed in Nomoto 2013 and Bale et al. 2019, assuming a numeral denotation from Ionin & Matushansky 2006, and an ⟨e,t⟩ type denotation for nouns. The numeral combines with an atomic predicate and returns the set containing all sums with the predicate property that have a cardinality of 2:

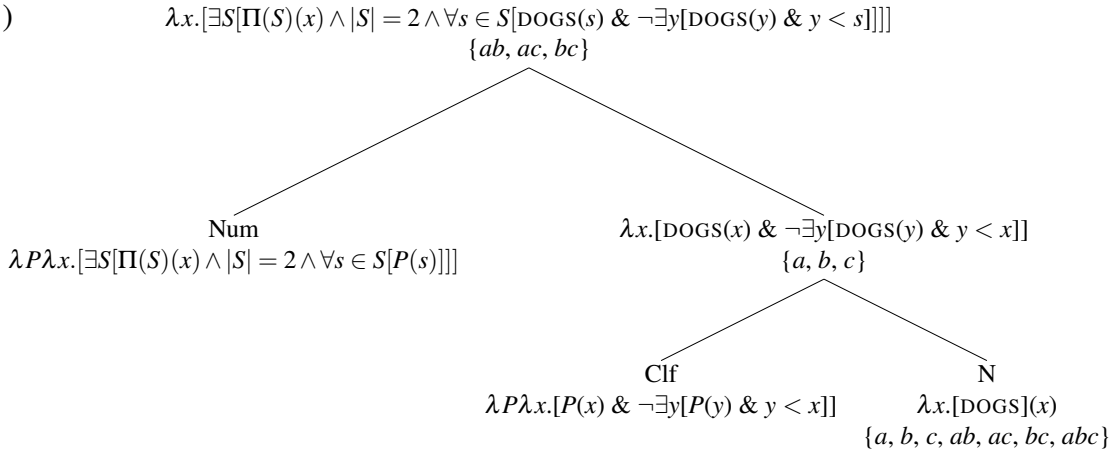
(9) $[\![2]\!] = \lambda P \lambda x.[\exists S[\Pi(S)(x) \wedge |S| = 2 \wedge \forall s \in S[P(s)]]]$ (Ionin & Matushansky 2006)

(9) denotes the entities x such that there is a partition S with two members and every member of S has property P. Crucially, notice that the numeral in (9) looks at a set. This is different from the numeral in (6), which picks out all the entities with a measure of 2. Since the noun in (5) is not atomic, classifiers, like in (10), are needed to atomize the members in the set denoted by the NP predicate.

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

(10) gives the set of x, such that x has the property P and there is no y with the property P that is a sub-part of x. When it combines with a noun predicate, this classifier generates a set containing atoms with the property described by the predicate. The classifier in (10) first combines with a noun allowing for the numeral in (9) to then combine with the classifier-noun complex, as shown in (11).

(11)



2.3 Summary

Though derivationally distinct, each kind of theory of numeral classifiers produces the same meaning for *two dogs*. For classifier-for-numeral theories, the numeral takes the classifier as a measure function, and then combines with the noun. 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 to create a set of atoms. Despite producing similar meanings, however, the two theories make different predictions regarding the distribution of numeral classifiers. These predictions can be summarized as in Table 1.[4]

Table 1: Predictions

If a classifier first combines with a:

1	NUMERAL	We might expect idiosyncrasies in whether or not a numeral requires a numeral classifier, as argued in Bale & Coon 2014
2	NOUN	We might expect idiosyncrasies in whether or not a noun requires a classifier, as argued in Simpson 2005, Nomoto 2013.
3	NOUN	We might expect to find the classifier with the noun in places other than with numerals.
4	NUMERAL	We might expect to find the classifier with the numeral when it is not combining with a noun.

[4]Bale et al. 2019 provide a number of *syntactic* diagnostics that favour a classifier-for-numeral theory of Ch'ol numeral classifiers (see sections 4.1 and 4.2 of their paper). Though we do not discuss these diagnostics here, these could also be used to differentiate between classifier-for-noun and classifier-for-numeral theories.

In the next section, we compare two unrelated languages, Ch'ol and Shan, which have both been described as having numeral classifiers. We show that while Ch'ol shows evidence for predictions 1 and 4, Shan shows evidence for predictions 2 and 3, suggesting that both numeral classifier theories sketched above are correct, but for different sets of languages. Diagnostic 4 is, to our knowledge, novel to this paper.

3. Evidence for two types of classifiers

In this section, we discuss in more detail the predictions laid out in Table 1, and how they follow from each theory of numeral classifiers sketched above. We provide data from Ch'ol and Shan providing evidence that both types of theories are in fact needed, supporting the proposal in (12). The overall conclusion will be that Ch'ol classifiers are classifiers-for-numerals, while Shan classifiers are classifiers-for-nouns.[5]

(12) There are two types of numeral classifiers across languages: *classifiers-for-numerals* (CLF-for-NUM) and *classifiers-for-nouns* (CLF-for-N).

3.1 Prediction 1 (NUM): Variation in whether a numeral requires a classifier

In classifier-for-numeral theories, classifiers are needed to satisfy an extra argument required by the numeral, represented as a measure function in (7). Since the appearance of the classifier is contingent on the semantics of the numeral, we might expect to find idiosyncrasies in whether a numeral requires a classifier. That is, as argued in Bale & Coon 2014, some numerals might have the measure function encoded in their lexical semantics, and others not. Such idiosyncrasies are observed in Ch'ol, not in Shan.

As Vázquez Álvarez (2011) notes, while higher Mayan-based numerals exist, speakers tend to use Mayan-based numerals when counting up to six and Spanish-based numerals for those higher than six. As shown in Bale & Coon 2014, Mayan-based numerals require a numeral classifier (13a), whereas Spanish-based numerals cannot combine with one (13b).

(13) 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: all numerals may appear with a classifier.

3.2 Prediction 2 (NOUN): Variation in whether a noun requires a classifier

For classifier-for-noun theories, we may expect to find the opposite of prediction 1. That is, if a classifier is required to atomize the set denoted by the noun, then there might be

[5] Ch'ol is a Mayan language of the Ch'olan-Tseltalan branch, spoken in southern Mexico by approximately 222,000 speakers. The data in this paper come from the Tumbalá and Tila dialects. Shan is a Tai-Kadai language of the Southwestern Tai branch, spoken in Myanmar and surrounding countries by approximately 3 million speakers (Lewis et al. 2016).

variation in whether nouns require atomization (and thus a classifier). For instance, it is possible that a subset of nouns in a language denote only a set of atoms, as argued in Simpson 2005 and Simpson & Ngo 2018 for Vietnamese and other East/Southeast Asian languages. In such cases, the set-counting numeral (see (9) above) could combine directly with a noun, circumventing the necessity for a classifier.

While Ch'ol does not exhibit idiosyncrasies with respect to nouns, Shan does. In Shan, some nouns do not need to combine with a classifier, as shown in (14). In Vietnamese, a language which as far as we can tell patterns with classifier-for-noun languages, some nouns never combine with a classifier, as in (15) from Simpson & Ngo 2018: ex. (7b).

(14) sǎam mx́ŋ
three country
'three countries' (Shan)

(15) hai chính phủ
two governments
'two governments' (Vietnamese)

3.3 Prediction 3 (NOUN): Classifiers required beyond numerals

For classifier-for-noun theories, 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. That is, it is conceivable that other modifiers or constructions in classifier-for-noun languages require the noun to be atomized, resulting in the obligatory presence of the classifier. This has been noted previously for some classifier languages by Simpson (2005) for example. As shown below, this is the case in Shan, which can have a classifier occur with quantifiers (16), demonstratives (17), and relative clauses (18), even in absence of a numeral. Furthermore, while nouns are number-neutral in Shan, in (17)-(18) the classifier appears to atomize the noun, giving rise to an obligatorily singular interpretation.

(16) mǎa ku **tǒ**
dog every CLF
'every dog'

(17) mǎa **tǒ** nâj
dog CLF DEM
'this dog'

(18) mǎa **tǒ** [$_{RC}$ ʔǎn nɔ́n jù]
dog CLF COMP sleep IPFV
'the dog that is sleeping' (Shan)

This is not the case in Ch'ol. Classifiers only ever appear with numerals or the interrogative numeral *jay-* 'how many' in Ch'ol. They are ungrammatical in other contexts, as shown in the example with a demonstrative in (19).

(19) * ili-kojty ts'i'
DEM-CLF dog
Intended: 'this dog' (Ch'ol)

3.4 Prediction 4 (NUM): Classifiers always obligatory with the numeral

If a classifier is a measure function generally required by a numeral, we would expect it to always appear with that numeral. That is, even when there is no noun, the classifier should always occur if it is required by that numeral. This is the case in Ch'ol: classifiers are *always* required, even when counting, as in (20), or when referring directly to the number, as in (21), which describes a context in which a teacher is pointing to a number on a chalkboard.

(20) 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)

(21) 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 when a noun is not present. For instance, they are optional (though degraded) when counting, as in (22), and are unacceptable when referring to the number itself, as in (23).

(22) CONTEXT: Students are practicing counting.
nɯŋ **(?tǒ)**, sǒŋ **(?tǒ)**, sǎam **(?tǒ)** . . .
one CLF two CLF three CLF
'1, 2, 3' (Shan)

(23) 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)

This diagnostic, to our knowledge, has not been observed before in the literature.

3.5 Summary

We began this section with the proposal in (12), repeated in (24).

(24) There are two types of numeral classifiers across languages: *classifiers-for-numerals* and *classifiers-for-nouns*.

We saw that while Ch'ol shows evidence for classifier-for-numeral theories, Shan shows evidence for classifier-for-noun theories. The predictions each proposal makes are summarized in Table 2.

Table 2: Summary of predictions supporting CLF-for-NUM and CLF-for-NOUN theories

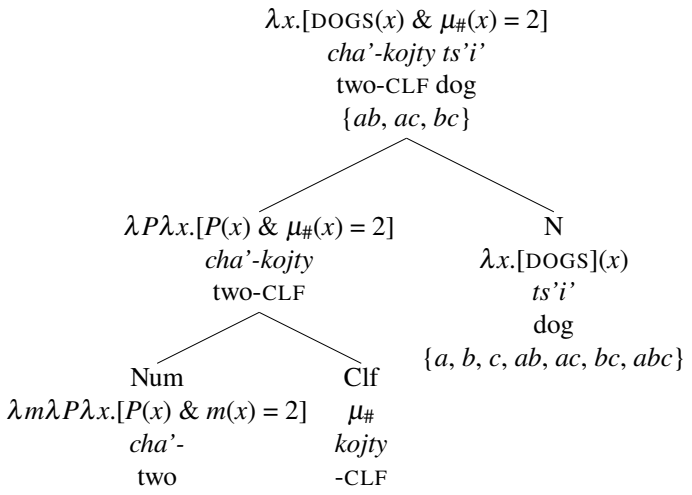
Supporting theory		Ch'ol	Shan
CLF-for-NUM	Prediction 1	✓	✗
CLF-for-N	Prediction 2	✗	✓
CLF-for-N	Prediction 3	✗	✓
CLF-for-NUM	Prediction 4	✓	✗

Clearly, Shan and Ch'ol numeral classifiers behave differently. Ch'ol numeral classifiers are required by the numeral and therefore always appear with numerals. Shan classifiers are dependent on the noun—some nouns do not appear with numeral classifiers and numeral classifiers appear in other environments, even in the absence of a numeral. We believe that the different predictions laid out above should be used as diagnostics in identifying the type of classifier exhibited in a language, and we will use them as such in Section 4 for another language, Chuj, which has two concurrent classifier systems.

For completeness, the full derivations for the Ch'ol and Shan equivalents of 'two dogs' are given below in (25) and (26). While derivationally distinct, both languages end up with equivalent sets after the numeral, classifier, and noun combine.

In (25), a Ch'ol-based numeral first combines with a classifier so that the classifier supplies the measure function. The resulting measure phrase then measures pluralities denoted by the noun, in this case those which have a cardinality of 2.

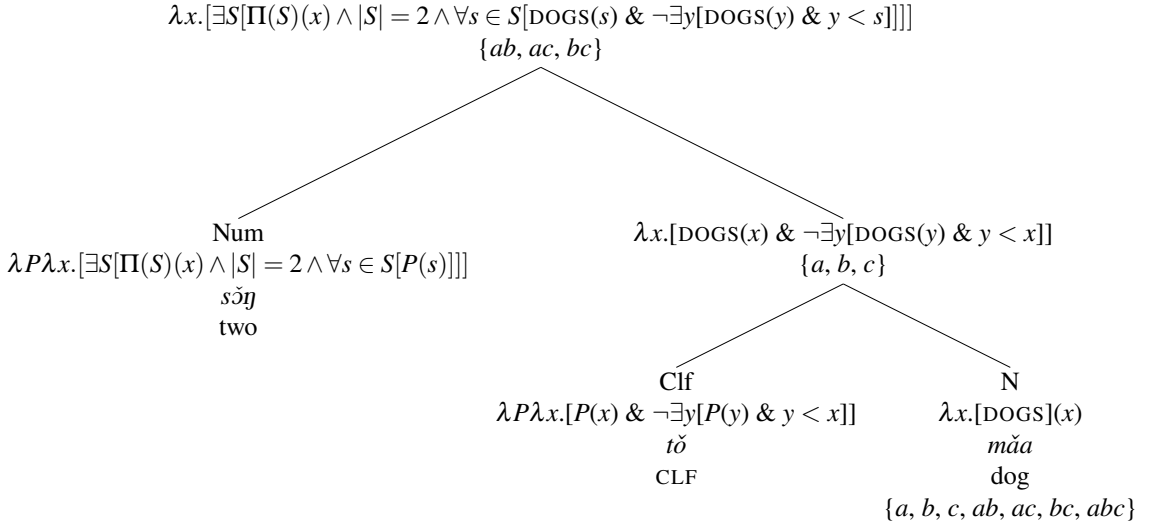
(25) Ch'ol



In Shan, the proposed semantics is given in (26). A classifier first combines with a noun. This is required given the denotation of the numeral. Numerals in Shan, unlike Ch'ol, do not count pluralities. Therefore, the classifier is needed to atomize the noun.[6]

[6]Note that we assume that the surface position of the nominal in Shan is derived by movement of the NP to a higher position, as proposed for Thai by Simpson (2005). However, for the purpose of semantic composition, the noun is interpreted in its base position.

(26) Shan



4. Discussion: Can a language have both kinds of classifiers?

We have seen evidence that there are at least two kinds of numeral classifiers across languages: classifiers-for-numerals and classifiers-for-nouns. We argued above that the type of classifier selected depends on semantic properties related to the numeral. On the one hand, numerals in classifier-for-numeral languages measure pluralities and must combine with a measure function argument, provided by the classifier-for-numerals. On the other hand, numerals in classifier-for-noun languages measure sets of atoms, and so the classifier-for-nouns is needed to extract the atoms denoted by the nominal predicate.

In the remaining subsections, we address a question raised by our proposal: Should we expect to find languages with both kinds of classifiers at the same time? At first glance, Chuj, another Mayan language, may be a good candidate, as it has two classifier systems:[7]

(27) ox**-e'** **ch'anh** libro
three**-**#.CLF N.CLF book
'three books' (Chuj)

In (27), there are two classifying morphemes: *-e'*, which we refer to as a "numeral classifier" ("#.CLF"), is suffixed to the numeral and classifies the noun's referent as inanimate; *ch'anh*, which we refer to as a "noun classifier" (glossed "N.CLF") immediately precedes the noun *libro* 'book' and classifies its referent in the set of paper entities.

Though Chuj does have two classifier systems, we will see that while *-e'* patterns with classifiers-for-numerals, *ch'anh* does not quite seem to pattern with classifiers-for-nouns. We will ultimately conclude, based on the semantics of both kinds of numeral classifiers in section 2, that no language should feature both kinds of numeral classifiers at the same time, at least semantically. The reason is that the semantics of classifiers-for-numerals is

[7]Chuj is an understudied Q'anjob'alan Mayan languages spoken in Guatemala and Mexico by approximately 70,000 speakers (Piedrasanta 2009).

incompatible with that of classifiers-for-nouns, and vice-versa. We therefore conclude that while a language might show overt exponents of the syntactic heads associated with both kinds of numeral classifiers, no language should feature morphemes that exhibit the semantics of both kinds of numeral classifiers in the same extended nominal projection.

4.1 Chuj numeral classifiers as classifiers-for-numerals

Chuj's numeral classifiers are obligatory with most Mayan-based numerals and with the interrogative numeral *jay* 'how many'. Chuj numeral classifiers abide by the diagnostics established in section 3.1 for classifiers-for-numerals.

First, as per prediction 1, there are idiosyncrasies in whether or not a numeral requires a numeral classifier. While most Mayan-based numerals require a classifier, the Mayan-based numeral *jun* 'one', as well as all numerals borrowed from Spanish, cannot combine with a numeral classifier, as shown in (28). Note that Chuj contrasts with Ch'ol with respect to its treatment of the numeral *jun* 'one', which unlike in Ch'ol, cannot combine with a numeral classifier, further supporting the view that the requirement for a numeral classifier is an idiosyncratic property of numerals.

(28) jun-**(*e')** … wentiyuno-**(*e')**, wentitres-**(*e')**
one-#.CLF SP:21-#.CLF, SP:22-#.CLF
'1 … 21, 22.' (Chuj)

The same type of idiosyncrasy is observed with Chuj's use of interrogative numerals (i.e. words that mean 'how many'). Chuj has two interrogative numerals, *jay*, which can only introduce count nouns, and *jantak*, which can introduce both count and mass nouns. While *jay* must obligatorily combine with numeral classifiers, *jantak* cannot:

(29) a. Jay-***(wanh)** kaxlan ix-a-man-a'?
HOW.MANY-#.CLF chicken PFV-A2S-buy-TV
'How many chickens did you buy?'

b. Jantak-**(*wanh)** kaxlan ix-a-man-a'?
HOW.MANY-#.CLF chicken PFV-A2S-buy-TV
'How many chickens did you buy?' (Chuj)

As already discussed in section 3.1 for Ch'ol, we propose, following Bale & Coon (2014) and Bale et al. (2019), that this observed idiosyncrasy follows from the fact that some Chuj numerals do not encode a measure function, while others do. Those that lack the measure function require an additional measure-function argument, realized as a numeral classifier.

Chuj numeral classifiers also show evidence for prediction 4: classifiers are required in all environments in which numerals are found, including in counting (30a) and when referring directly to the numeral (30b):

(30) a. CONTEXT: Students are practicing counting.
ox-***(e')**, chanh-***(e')**, hoy-***(e')**
three-#.CLF, four-#.CLF, five-#.CLF
'3, 4, 5.' (Chuj)

b. 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)

Finally, Chuj's numeral classifiers do not pattern with predictions 2 and 3, which we identified as diagnostics of classifiers-for-nouns. Indeed, contrary to Shan, the requirements on the presence of numeral classifiers do not change based on the choice of noun (contra prediction 2). That is, all Mayan-based numerals above 'one' require a numeral classifier, irrespective of the choice of the noun. Moreover, numeral classifiers are never used in environments other than with numerals or *jay* 'how many' (contra prediction 3).

We take the above facts as evidence that the numerals that require a numeral classifier in Chuj exhibit the same semantics as the Mayan-based numerals found in Ch'ol, and that Chuj numeral classifiers are the realization of variables over measure functions.

4.2 Chuj noun classifiers as classifiers-for-nouns?

As shown in example (31), Chuj has a second type of classifying morpheme, referred to as "noun classifiers" in previous literature (e.g. Hopkins 2012). These sometimes co-occur with numerals, though optionally, and they are used to mark distinctions of specificity and definiteness (see e.g. Buenrostro et al. 1989 and Royer 2019). Examples with different noun classifiers are provided below:

(31) a. ox-e' **lum** chen
three-#.CLF N.CLF pots
'three pots'

b. ox-e' **anh** pajich
three-#.CLF N.CLF tomato
'three tomatoes'

Given that Chuj's numeral classifiers exhibit the same distribution as Ch'ol classifiers-for-numerals, as in (7), it is natural to consider whether Chuj's noun classifiers could exhibit the semantics of classifiers-for-nouns. At first glance, Chuj noun classifiers seem to share a number of properties with Shan classifiers-for-numerals, not shared with Ch'ol and Chuj classifiers-for-numerals. In fact, they satisfy all of the diagnostics from section 3.1 for classifiers-for-nouns.

As per prediction 2, not all Chuj nouns combine with a classifier. For instance, in the closely related language Akatek, Zavala (1992, p. 157) lists several categories of nouns that cannot combine with noun classifiers. These include nouns denoting body parts (e.g. *nhi* 'nose'), abstract entities (e.g. *ib'* 'strength'), natural phenomena (e.g. *'asun'* 'cloud'), place names (e.g. *chonhab'* 'village'), measure terms (e.g. *ja'* 'armful'), and other entities that have been recently introduced in the language. The same restrictions also hold for Chuj.

Chuj noun classifiers also abide by prediction 3 in appearing in environments other than with numerals, as in the examples in (32)–(34).

(32) **nok'** tz'i'
N.CLF dog
'the dog'

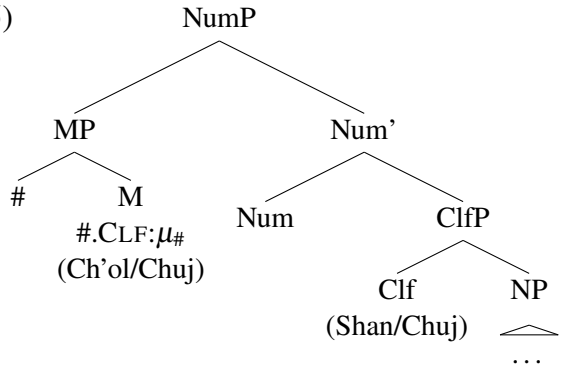
(33) **nok'** tz'i' **chi'**
N.CLF dog that
'that dog'

(34) [$_{RC}$ **nok'** ix-way-i]
N.CLF PFV-sleep-IV
'the one (animal) that slept.' (Chuj)

The examples in (32) to (34) show a number of environments in which noun classifiers are found without a numeral. In (32), the noun classifier for animals, *nok'*, appears alone with the noun. In such cases, noun classifiers generally lead to definite interpretations (see Royer 2019 for more details). In (33), the noun classifier is also required in the presence of a demonstrative (32), similarly to Shan, which optionally allows the presence of numeral classifiers with demonstratives. Example (34) shows that noun classifiers can introduce relative clauses, also observed for Shan above in (18).

Given the overall similarity in the syntactic distribution of Chuj noun classifiers and Shan classifiers-for-nouns, we contend that both are exponents of the same syntactic head. Combining the proposed structures for classifiers-for-numerals and classifiers-for-nouns, we propose the following structure for the numeral-noun complex in Chuj.

(35)



The above structure and its labels follow Bale et al. (2019) in positing a Measure phrase (MP) that contains both the numeral and the classifier-for-numerals. Following Li (1999), we locate the numeral in the specifier of NumP, where NumP heads plural marking (see e.g. Valois 1991, Ritter 1995 for similar approaches to numerals and number marking across languages). All of the heads seen in (35) can be overtly realized in Chuj, including the NumP head, which has an obligatory exponent (*heb'*) with nouns that denote humans:

(36) ox-wanh **heb'** winh winak
three-#.CLF PL N.CLF man
'three men' (Chuj)

Finally, as is standard in the literature on East Asian and South East Asian languages (see e.g. Li 1999, Cheng & Sybesma 1999, Simpson 2005, a.o.), we posit a Classifier Phrase (ClfP) that takes the nominal as its complement. We propose that this projection hosts both Chuj's noun classifiers, and Shan's classifiers-for-nouns.

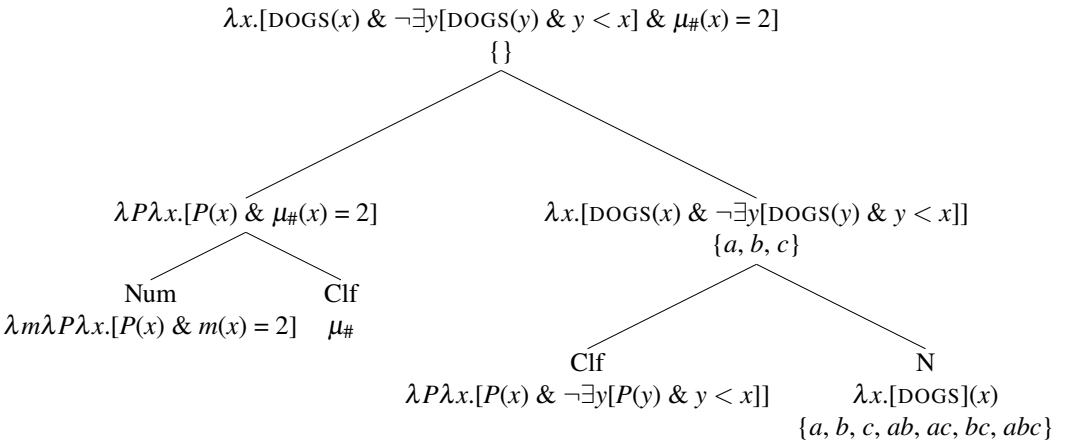
4.3 Chuj noun classifiers ≠ classifiers-for-nouns

We have proposed that Chuj noun classifiers and Shan's classifiers-for-nouns occupy the same syntactic head, labelled the Classifier Phrase. However, we argue that Chuj noun classifiers cannot fulfill the exact same semantic function as classifiers-for-nouns. This is because the denotations we have provided for numerals in classifier-for-numeral languages and numerals in classifier-for-noun languages are not semantically compatible.

It is important to mention at this juncture that Chuj noun classifiers are different from Shan's classifiers-for-nouns, and other clear cases of classifiers-for-nouns across languages, in one crucial respect: they are optional in the presence of a numeral. This suggests that, contrary to classifiers-for-nouns, Chuj's noun classifiers do not serve to extract the atomic entities from the set of atomic and plural entities denoted by the nominal predicate. Otherwise, we would expect them to appear whenever a noun appears with a numeral. As we will argue, this is a welcome prediction.

If noun classifiers in Chuj were to exhibit the semantics of classifiers-for-nouns, the noun classifier would retrieve only the set of atoms denoted by the nominal predicate, as classifiers-for-nouns do. If this were to happen, there would be no remaining pluralities for the numeral to measure, meaning that applying the noun to the numeral would always yield the empty set, clearly an undesired consequence. This is schematized in (37).

(37)



The result is that we do not expect to find a language with the semantics of classifiers-for-numerals and classifiers-for-nouns at the same time. This follows from the denotation assigned to numerals in classifier-for-numeral languages, which must measure plural entities, instead of sets of atomic entities. We conclude that while Chuj noun classifiers are syntactically like classifiers-for-nouns, insofar as they seem to realize the same syntactic head, they are semantically unlike classifiers-for-nouns, since they are not used to atomize.

Nevertheless, Chuj noun classifiers convey similar presupposition triggers with what we suspect are cases of classifiers-for-nouns in other languages. For example, it is well known that classifiers in some Southeast Asian and East Asian languages are associated with definiteness. q Cantonese (Cheng & Sybesma 1999, a.o.) and Hmong (Simpson 2005, a.o.) classifiers trigger a definiteness presupposition when appearing alone with a noun (see also Jaisser 1987; Bisang 1993; Aikhenvald 2000; and Simpson et al. 2011, a.o.):

(38) a. **Gaa** ce zo-zyu **go** ceot-hau.
CLF car block-CONT CLF exit
'The car is blocking the exit.' (Cantonese: Cheng & Sybesma 1999, p. 521)

b. **Tus** tsov tshaib tshaib plab.
CLF tiger hungry hungry stomach
'The tiger is very hungry.' (Hmong: Jaisser 1987, p. 171)

As mentioned above, noun classifiers in Chuj also trigger presuppositions related to definiteness and specificity (see, e.g., Royer 2019). We take the shared presuppositions between classifiers-for-nouns and Chuj noun classifiers as additional evidence that these morphemes originate in the same syntactic position.

In sum, classifiers across languages do not behave uniformly. We have proposed that classifiers which occur with numerals can be divided into the two categories argued for here. We hope future lines of inquiry will continue to investigate distinctions in classifier languages and use the diagnostics laid out in this paper to do so.

References

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. OUP Oxford.

Bale, Alan, & Jessica Coon. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry* .

Bale, Alan, Jessica Coon, & Nicholás Arcos. 2019. Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa: a journal of general linguistics* 4.

Bisang, Walter. 1993. Classifiers, quantifiers and class nouns in Hmong. *Studies in Language* 71:1–51.

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

Cheng, Lisa Lai-Shen, & 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.

Dayal, Veneeta. 2012. Bangla classifiers: Mediating between kinds and objects. *Rivista di Linguistica* 24:195–226.

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

Ionin, Tania, & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23:315–360.

Jaisser, Annie. 1987. Hmong classifiers: A problem set. *Linguistics of the Tibeto-Burman area* 10:169–176.

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

Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig. 2016. URL `https://www.ethnologue.com/language/shn`.

Li, XuPing, & Susan Rothstein. 2012. Measure readings of Mandarin classifier phrases and the particle *de*. *Language and Linguistics* 13:693.

Li, Yen Hui Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8:75–99.

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

Nomoto, Hiroki. 2013. Number in classifier languages. Doctoral dissertation, University of Minnesota.

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

Ritter, Elizabeth. 1995. On the syntactic category of pronouns and agreement. *Natural Language and Linguistic Theory* 13:405–443.

Rothstein, Susan. 2017. *Semantics for Counting and Measuring*. Key Topics in Semantics and Pragmatics. Cambridge: Cambridge University Press.

Royer, Justin. 2019. Domain restriction and noun classifiers in Chuj. In *NELS 49: Proceedings of the Forty-Ninth Annual Meeting of the North East Linguistic Society*, ed. Maggie Baird & Jonathan Pesetsky, volume 3, 87–97.

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

Simpson, Andrew, & Binh Ngo. 2018. Classifier syntax in Vietnamese. *Journal of East Asian Linguistics* 27:211–246.

Simpson, Andrew, Hooi Ling Soh, & Hiroki Nomoto. 2011. Bare classifiers and definiteness: A cross-linguistic investigation. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 35:168–193.

Valois, Daniel. 1991. The internal syntax of DP and adjective placement in French and English. In *NELS 21: Proceedings of the 21st Conference of the North East Linguistic Society*, 367–382.

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

Wilhelm, Andrea. 2008. Bare nouns and number in Dëne Sųłiné. *Natural Language Semantics* 16:39–68.

Zavala, Roberto. 1992. El Kanjobal de San Miguel Acatán. Doctoral dissertation, Universidad Nacional Autónoma de México, Ciudad de México.

Carol-Rose Little, Mary Moroney, Justin Royer
crl223@cornell.edu, mrm366@cornell.edu, justin.royer@mail.mcgill.ca