

CLASSIFIERS ARE FOR NUMERALS, NOT FOR NOUNS: CONSEQUENCES FOR THE MASS-COUNT DISTINCTION*

Alan Bale & Jessica Coon,
alan.bale@mcgill.ca, jessica.coon@mcgill.ca
December 17, 2013

1 Introduction

In languages with numeral classifiers systems, nouns must appear with one of a series of classifiers in order to be modified by a numeral. This squib presents new data from Mi'gmaq (Algonquian) and Chol (Mayan), arguing that numeral classifiers are required due to the syntactic and semantic properties of the *numeral* (as in Krifka 1995), rather than the *noun* (as in Chierchia 1998). The results are shown to have important consequences for the mass-count distinction.

Mandarin Chinese is a frequently cited example of a language with obligatory numeral classifiers. As shown in (1), classifiers cannot be dropped in the presence of numerals.¹

(1) MANDARIN CHINESE

- a. liǎng *(zhāng) zhuōzi
two CL table
'two tables'
- b. liǎng *(píng) jiǔ
two CL.bottle wine
'two bottles of wine'

Krifka (1995) and Chierchia (1998) provide two very different accounts of the theoretical distinction between languages with obligatory classifiers (like Mandarin) and those without (like English). Chierchia links the distinction to the nominal system, arguing that non-classifier languages have a mass-count distinction among nouns, while classifier languages do not. All nouns in Mandarin are likened to mass nouns in English. Krifka, on the other hand, proposes that the difference lies in the *numeral* system. He argues that classifier languages morphologically separate the semantic measure function (i.e., the classifier) from the numerals, whereas non-classifier languages have a measure function incorporated into the numerals.

Here we bring in new data from Mi'gmaq and Chol—languages which *sometimes* use classifiers—in order to distinguish between the two theories. In both languages, certain numerals obligatorily appear

*Thanks to David Barner, Brendan Gillon, Peter Jenks, David Nicolas, the audience at NELS 43, and two anonymous reviewers for helpful discussion and feedback. We are especially grateful to Janine Metallic, Mary Ann Metallic, and Janice Vicaire for Mi'gmaq, and to Matilde Vázquez Vázquez, Juan Jesús Vázquez Álvarez, and Nicolás Arcos López for Chol. Any errors in data or interpretation are of course our own.

¹There are two basic types of classifiers exemplified in (1): sortal and mensural (Lyons 1977; Aikhenvald 2000). *Sortal* classifiers, like *zhāng* in (1a), rely on an intrinsic “divided reference” (Quine 1960) separate from the classifier. The classifier itself merely permits counting with respect to this division. Such classifiers sometimes vary in form according to inherent properties of the noun, such as animacy, shape, or consistency. *Mensural* classifiers, like *píng* in (1b), specify a way of dividing a reference (e.g., into packages like bottles, or units of measurement like inches or kilos). Such classifiers can combine with any noun whether they have an inherent divided reference or not. This distinction does not figure into the discussion in this paper. However, most of our examples will use sortal classifiers.

Finally, in this paper we restrict our attention to *numeral* classifiers, ignoring systems in which nominals may appear in certain (e.g. definite) environments with a classifier-like element indicating membership to some class, even in the absence of a numeral (cf. discussion of Akatek in Zavala 2000).

with classifiers, while others never do. We show that these idiosyncratic numeral systems—also attested in other languages, discussed below—cannot be accounted for under Chierchia’s influential (1998) proposal. Furthermore, we show that these results have consequences for the mass-count distinction. Krifka’s theory, unlike Chierchia’s, treats the classifier/non-classifier distinction as being *theoretically independent* of the syntactic mass-count distinction (see Wilhelm 2008). We question whether it is meaningful, or even empirically justified, to maintain a mass-count distinction once classifier systems are treated in this way.

2 Theoretical background and previous work

2.1 Chierchia (1998): Classifiers are for nouns

Chierchia (1998) argues that numerals have a uniform interpretation across both classifier and non-classifier languages, but hypothesizes a difference in the nominal systems. In English, there are two categories of nouns: one that is directly compatible with numeral modification (so-called *count nouns*, like *table* and *girl*), and another that is not (so-called *mass nouns*, like *furniture* and *water*). Chierchia proposes that in a classifier language like Mandarin there is only a single category of noun, and, much like English mass nouns, this category is not directly compatible with numeral modification. A simplified version of Chierchia’s nominal interpretations is shown in (2), where \cap is a function from predicates to kinds.² Here the Mandarin noun *zhuōzi* ‘table’ in (2a) denotes a kind, on par with the English mass noun *furniture* in (2b), but not similar to the English count noun *table* in (2c), which denotes a set of atoms.

(2) **Chierchia-style nominals** (simplified)

- a. $\llbracket zhuōzi \rrbracket = \cap \text{TABLE}$ (i.e., the table-kind)
- b. $\llbracket furniture \rrbracket = \cap \text{FURNITURE}$ (i.e., the furniture-kind)
- c. $\llbracket table \rrbracket = \{x : \text{ATOM}(x) \ \& \ \text{TABLE}(x)\}$ (i.e., set of individual tables)

According to Chierchia (1998), numeral modification relies on measure functions that count (stable) atoms. The kinds in (2a) and (2b), in contrast to (2c), contain no such atoms. As a result, they must be converted into atomic sets before combining with numerals. Thus, just as English mass nouns require measure words to combine with numerals (e.g. ‘two *pieces* of furniture’), all nouns in Mandarin require classifiers that convert kinds into atomic sets.

Chierchia-style denotations for numerals and classifiers are provided in (3), where *ATOMIC* is a function true of predicates with atomic minimal parts (i.e., atoms); $\mu_{\#}$ is a measure function from a group to the cardinality of that group; and $*$ is a closure operator from a set of entities to the set of all sums that can be formed from those entities (Link 1983).

(3) **Chierchia-style numerals and classifiers** (simplified)

- a. $\llbracket liǎng \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 2\}$
- b. $\llbracket zhāng \rrbracket = \cup$ (i.e., the function from kinds to sets of atoms)

The numeral *liǎng* in (3a) is a function from atomic sets to sets of groups composed of two members from the atomic set. The classifier *zhāng* in (3b) is a function from kinds to predicates, represented as \cup .

When a classifier like *zhāng* combines with a nominal like *zhuōzi* (as in 1a), the result is denotationally equivalent to an English count noun. This is illustrated in (4).

²Chierchia (1998)’s actual proposal involves coercion operators that freely apply in any language. Critical to the present discussion, one conversion operator (π) maps kinds to complete semi-lattices. The classifier then maps these complete semi-lattices to atomic predicates. Thus, $\llbracket \text{CL} \rrbracket \circ \pi = \cup$. Note also that Chierchia interprets mass nouns in English as complete semi-lattices but has a conversion operator that maps such lattices into kinds. For the sake of exposition, we will ignore this subtlety.

(4) **Equivalences**

$$\llbracket zhāng \rrbracket(\llbracket zhuōzi \rrbracket) = \{x : \text{ATOM}(x) \ \& \ \text{TABLE}(x)\} = \llbracket table \rrbracket$$

2.2 Krifka (1995): Classifiers are for numerals

For Krifka, denotations of nominals in Mandarin are comparable to those in English, shown in the simplified version of his theory in (5).³

(5) **Mandarin Nominals, Equivalent to English Count Nouns**

$$\llbracket zhuōzi \rrbracket = \{x : \text{ATOM}(x) \ \& \ \text{TABLE}(x)\}$$

The difference lies in the numerals. Krifka (1995) hypothesizes that there are two different types of numeral interpretations cross-linguistically (see also Wilhelm 2008). On the one hand, there are numerals in non-classifier languages like English. These have an incorporated measure function, $\mu_{\#}$, and combine directly with nouns, as illustrated for English *two* in (6a) — the function $\mu_{\#}$ maps groups to the number of individuals in that group.⁴ On the other hand, there are Mandarin-like numerals like *liǎng* in (6b). These do not have an incorporated measure function, and thus require classifiers—like *zhāng* in (6c)—in order to introduce a measure.⁵

(6) **Krifka-like numerals and classifiers (simplified)**

- a. $\llbracket two \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 2\}$
- b. $\llbracket liǎng \rrbracket = \lambda m \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ m(x) = 2\}$
- c. $\llbracket zhāng \rrbracket = \mu_{\#}$

Under this account, a Mandarin numeral-plus-classifier is semantically equivalent to an English numeral, shown in (7).

(7) **Equivalences**

$$\llbracket liǎng \rrbracket(\llbracket zhāng \rrbracket) = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 2\} = \llbracket two \rrbracket$$

As noted by Krifka,⁶ there is very little evidence internal to English or Mandarin that would favour one proposal over another. Both theories succeed in capturing the fact that Mandarin requires classifiers for counting, while English does not. For Chierchia, the requirement for classifiers is due to a deficiency of the nouns: they do not denote countable entities. For Krifka, classifiers are necessary because of a problem with

³Krifka (1995)’s actual theory treats nouns in English and Mandarin as kinds. These kinds serve many purposes, including fixing how measure functions count. For simplicity, and to make the separation between his theory and Chierchia’s, we have change the kind denotations to atomic sets. What is critical to the present discussion is that Krifka makes distinctions in the numeral system rather than the nominal system.

⁴We follow the convention of using the symbol $\mu_{\#}$ to represent the measure function that maps groups of a certain category to natural numbers, specifically the number of minimal parts (a.k.a., individuals) contained within the group. This measure function is always relative to a category or kind. We do not address how this relativization is implemented, but see Bale and Barner (2009) and Krifka (1995) for a discussion. Krifka (1995) uses a different symbol to represent this measure function, namely OU .

⁵In our simplified version, both types of numerals introduce a presupposition that their nominal arguments are atomic (i.e., they have atomic minimal parts, whether those minimal parts are inherent in the nominal denotation itself or induced by mensural classifiers). This requirement is represented by $\text{ATOMIC}(P)$. Such presuppositions would predict the infelicity of DPs like *two firewood(s)*. The noun *firewood* does not have atomic minimal parts and hence fails to satisfy the presupposition.

⁶In his discussion of English, Krifka (1995) compares a theory where the measure function is incorporated into the noun versus the numeral. Thus, his comparison does not directly involve kinds and kind conversions, like Chierchia’s (1998). Rather, his alternative more resembles Cresswell (1976), where count nouns have a built in measure but mass nouns require a measure function as an argument. Still, the empirical consequences of Cresswell’s and Chierchia’s theory are rather similar, namely that the presence of measure terms is dependent on the noun rather than the numeral.

the numerals. Informally speaking, they do not come pre-specified with information about which types of things they count.⁷

2.3 Case study: Western Armenian

In Western Armenian the presence or absence of a classifier is completely optional, shown in (8).

- (8) *yergu* (**had**) *dəgha*
 two CL boy
 ‘two boys’ (c.f., Donabédian 1993)

The two theories described above offer two possible explanations for this variation. Under Chierchia’s account, the noun *dəgha* ‘boy’ would be ambiguous, having one meaning that permits the noun to combine directly with numerals (a “count” denotation, as in 9a), and another that requires a classifier (a “mass” denotation, as in 9b). Numerals and classifiers have denotations as in (9c–d), as in (3) above.

- (9) a. $\llbracket d\acute{e}gha_1 \rrbracket = \{x : \text{BOY}(x)\}$
 b. $\llbracket d\acute{e}gha_2 \rrbracket = \cap \text{BOY}$
 c. $\llbracket yergu \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 2\}$
 d. $\llbracket had \rrbracket = \cup$

Krifka, in contrast, could hypothesize that the noun *dəgha* ‘boy’ has a consistent count-type interpretation, but the numeral *yergu* is ambiguous. One meaning incorporates a measure function, as in (10b). The second meaning does not, as in (10c). See Borer 2005 for a similar proposal.

- (10) a. $\llbracket d\acute{e}gha \rrbracket = \{x : \text{BOY}(x)\}$
 b. $\llbracket yergu_1 \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 2\}$
 c. $\llbracket yergu_2 \rrbracket = \lambda m \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ m(x) = 2\}$
 d. $\llbracket had \rrbracket = \mu_{\#}$

There is no clear way to decide between the two theories language-internally in Western Armenian.⁸ However, this optionality raises an interesting consideration, namely, the possibility of variation within a single language. The two theories make different predictions with respect to cross-linguistic variation: Krifka’s numeral-based theory predicts the possibility of a language with idiosyncratic behaviour among the numerals, whereas Chierchia’s theory is inconsistent with such a pattern. In the sections below, we provide examples of languages which show idiosyncratic patterns in the numeral domain and show that these data are uniquely compatible with Krifka’s account of classifiers.

⁷As an anonymous reviewer points out, languages like Cantonese have a broader range of patterns than languages like Mandarin (see Cheng and Sybesma 1999). For example, classifiers in Cantonese can be used without numerals to signal singular definite interpretations (see footnote 1). On the surface, such facts seem to favour Chierchia’s account. According to Krifka (1995), classifiers serve as arguments to the numerals, and hence it is not expected for classifiers to appear without numerals. However, to account for the Cantonese data, Chierchia (1998) requires some kind of coercion operator to map the atomic set denotation to a unique individual (e.g., a covert application of the supremum operator). A similar hypothesis could be maintained in Krifka’s system, except that the coercion operator would apply to measure functions rather than atomic sets. Such a function could be represented as follows: $\lambda m. \lambda P. \iota(\{x : P(x) \ \& \ m(x) = 1\})$. In the end, the broader pattern can be integrated into either account of classifiers.

⁸The facts in Western Armenian have been simplified for the purposes of exposition. As discussed in Donabédian 1993, there is a plural marker in Western Armenian that cannot co-occur with classifiers. However, numerals can combine directly with either bare nouns or plural forms. Borer (2005) argues that this pattern is evidence that the plural marker and classifiers compete for the same syntactic head. However, Bale and Khanjian (2009) demonstrate that semantic reasons alone can account for the lack of co-occurrence. Furthermore, Doetjes (2012) demonstrates that for some languages co-occurrence is possible.

3 Idiosyncratic numerals

3.1 Mi'gmaq and Chol

In Mi'gmaq, an Eastern Algonquian language, numerals 1–5 (along with numerals morphologically built from 1–5) do not appear with classifiers, while numerals 6 and higher must. In (11a) we observe that the numeral *na'n* ‘five’ combines directly with the noun; the classifier *te's* is impossible, as shown in (11b).

- (11) a. *na'n-ijig ji'nm-ug*
 five-AGR man-PL
 ‘five men’
 b. **na'n te's-ijig ji'nm-ug*
 five CL-AGR man-PL

In contrast, the numeral *asugom* ‘six’ in (12a) cannot combine directly with a noun. It must instead appear with the classifier *te's*, shown in (12b).

- (12) a. **asugom-ijig ji'nm-ug*
 six-AGR man-PL
 b. *asugom te's-ijig ji'nm-ug*
 six CL-AGR man-PL
 ‘six men’

Chol, a Mayan language of southern Mexico, also demonstrates idiosyncratic behaviour in the numeral system. Mayan languages have a vigesimal (base 20) numeral system. Many speakers today, however, generally know and use Chol numerals only for numbers 1–6, 10, 20, 40, 60, 80, 100, and 400 (Vázquez Álvarez 2011, 180); otherwise they use number words borrowed from Spanish.

As shown in (13), the traditional Mayan numerals, like *ux* ‘three’, require a classifier.

- (13) a. *ux-p'ej tyumuty*
 three-CL egg
 ‘three eggs’
 b. **ux tyumuty*
 three egg

In contrast, the Spanish-based numerals, like *nuebe* ‘nine’, cannot be used with classifiers, as shown in (14). This contrast is consistent across all Spanish vs. Mayan-based numerals in the language and cannot be reduced to other factors like phonological size: multi-syllabic Chol numerals like *waxäk* ‘eight’ still require classifiers, and Spanish-based numerals like *ses* ‘six’ still prohibit them.

Note that Krifka’s theory can be easily adapted to account for the Western Armenian pattern. The bare noun would be interpreted as a complete semi-lattice while the plural noun would be a strict plural without atomic parts (e.g., $\llbracket d\acute{e}gha \rrbracket = \{x : *BOY(x)\}$ and $\llbracket d\acute{e}gha+PL \rrbracket = \llbracket d\acute{e}gha \rrbracket - BOY$; see Bale et al. 2011). There would still be an ambiguity among the numeral modifiers such that one would take a measure function as an argument whereas the other would not. Also, one of the modifiers would presuppose that its nominal argument is a complete semi-lattice whereas the other would not (e.g., $\llbracket yergu_1 \rrbracket = \lambda P. \{x : P(x) \ \& \ \mu_{\#}(x) = 2\}$, $\llbracket yergu_2 \rrbracket = \lambda m. \lambda P. COMPLETE(P). \{x : P(x) \ \& \ m(x) = 2\}$, where COMPLETE is true only of predicates closed under join and meet). Thus, one type of numeral can combine with either plural or non-plural nouns but prohibits classifiers. The other type of numeral requires classifiers but can only combine with non-plural nouns. It is difficult for Chierchia’s theory to account for the Western Armenian plural pattern without hypothesizing either a phonologically null classifier (to convert plural nouns to singular denotations), or an ambiguity in the numeral system (one numeral applying to singulars denotations, the other to plural denotations).

- (14) a. * *nuebe-p'ej tyumuty*
 nine-CL egg
 b. *nuebe tyumuty*
 nine egg
 ‘nine eggs’

It should be noted that this is true not just of bilingual Spanish–Chol speakers, but also of speakers who are essentially monolingual in Chol. Regardless of degree of fluency, age, or level of bilingualism, speakers consistently find classifiers on Spanish-borrowed numerals to be ungrammatical. Furthermore, this variation is *not* found within the nominal system. Nominals borrowed from Spanish require classifiers when they are used in conjunction with a Chol numeral, as shown with the Spanish loan *mansana* ‘apple’ in (15a). When appearing with numerals of Spanish origin, no classifier is possible, as in (15b).⁹

- (15) a. *Tyi k-mäñä ux-p'ej mansana*
 ASP 1ERG-buy three-CL apple
 ‘I bought three apples.’
 b. *Tyi k-mäñä nuebe mansana*
 ASP 1ERG-buy nine apple
 ‘I bought nine apples.’

3.2 Discussion

Both Mi’gmaq and Chol have some numerals that require classifiers, and some numerals that cannot appear with classifiers. This is consistent with an approach in which nominals have a consistent denotation and variation is found within the numerals themselves, i.e., Krifka’s analysis. This is illustrated below with Chol lexical items, but is readily transportable to Mi’gmaq.

Under Krifka’s analysis, nominals like *tyumuty* ‘egg’ have denotations equivalent to their English counterparts. The noun *tyumuty* is a predicate true of eggs, as in (16).

- (16) $\llbracket tyumuty \rrbracket = \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\}$

The requirement for a classifier is dependent, not on the noun, but on the syntax and semantics of the numeral. In Chol, the interpretation of Spanish-origin *nuebe* ‘nine’ is a nominal modifier that has a cardinality measure ($\mu_{\#}$) built into its meaning, as shown in (17).

- (17) **Denotation of numeral which does not permit classifier**
 $\llbracket nuebe \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 9\}$

In contrast, the interpretation of *ux* (Chol ‘three’) is a function that takes a measure function as an argument, such as the cardinality measure *p’ej*, and yields a numeral modifier. This is illustrated in (18).

⁹An anonymous reviewer suggests that this could be a problem with attaching Chol inflectional morphology to Spanish loans. This again cannot be right: Spanish loans appear with inflectional morphology (e.g. plural, possession) across the language. Even Spanish-based numerals may be inflected for person, as in (ia). Compare with the Chol numeral plus classifier in the same construction in (ib).

- (i) a. *Nuebe-j-oñ-loñ.*
 nine-EPEN-1ABS-PL.EXCL
 ‘We are nine.’ (i.e. a group of nine people)
 b. *Cha’-tyikil-oñ-loñ.*
 two-NC.people-1ABS-PL.EXCL
 ‘We are two.’ (i.e. a group of two people)

(18) a. **Denotaion of numeral which requires classifier**

$$\llbracket ux \rrbracket = \lambda m \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ m(x) = 3\}$$

b. **Denotation of the classifier**

$$\llbracket p'ej \rrbracket = \mu_{\#}$$

As illustrated in (19a), *nuebe* can combine directly with nouns like *tyumuty* to yield a set of groups where each group consists of 9 individual eggs. However, the combination of *nuebe* with a classifier leads to a type mismatch and presupposition failure.

(19) a. $\llbracket nuebe \rrbracket(\llbracket tyumuty \rrbracket) = \{x : x \in * \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\} \ \& \ \mu_{\#}(x) = 9\}$

$$\llbracket nuebe \rrbracket(\llbracket p'ej \rrbracket) \rightarrow \text{type mismatch.}$$

b. $\llbracket ux \rrbracket(\llbracket tyumuty \rrbracket) \rightarrow \text{type mismatch.}$

$$(\llbracket ux \rrbracket(\llbracket p'ej \rrbracket))(\llbracket tyumuty \rrbracket) = \{x : x \in * \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\} \ \& \ \mu_{\#}(x) = 3\}$$

The opposite pattern holds for *ux*, as illustrated in (19b). The combination of *ux* directly with *tyumuty* leads to a type mismatch, whereas combination with the classifier *p'ej* and then *tyumuty* yields a set of groups where each group consists of 3 individual eggs.

Unlike Krifka's account, Chierchia's theory cannot account for the patterns illustrated in (19). To account for acceptable forms where numerals combine directly with nouns, as in *nuebe tyumuty*, as well as forms where classifiers intervene, as in *ux-p'ej tyumuty*, Chierchia would need to hypothesize that nouns in Mi'gmaq and Chol are ambiguous. Under this account, all nouns would have two interpretations: one interpretation that requires classifiers, and another that does not, as shown in (20a). The numerals would have interpretations that were independent of the classifier, whereas the classifier would be a function from kinds to sets, as shown in (20b–c).

(20) **Chierchia-inspired Interpretations of Chol**

a. **NOMINAL INTERPRETATIONS**

$$\llbracket tyumuty_1 \rrbracket = \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\},$$

$$\llbracket tyumuty_2 \rrbracket = \cap \text{EGG}$$

b. **NUMERAL INTERPRETATIONS**

$$\llbracket ux \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 3\},$$

$$\llbracket nuebe \rrbracket = \lambda P : \text{ATOMIC}(P). \{x : *P(x) \ \& \ \mu_{\#}(x) = 9\}$$

c. **CLASSIFIER INTERPRETATION**

$$\llbracket p'ej \rrbracket = \cup$$

Critically, if nouns like *tyumuty* in (20a) are ambiguous in this respect then the *ungrammatical* forms are unexpected. Nothing would prevent a classifier-less Mayan numeral from combining with the interpretation of *tyumuty* 'egg' which denotes an atomic set. Similarly, nothing rules out the the possibility that the kind-denoting variant of *tyumuty* could combine with the Spanish-based numeral *nuebe*, requiring a classifier.

(21) **FALSE PREDICTIONS**

a. $\llbracket ux' \rrbracket(\llbracket tyumuty_1 \rrbracket) = \{x : x \in * \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\} \ \& \ \mu_{\#}(x) = 3\} \rightarrow \text{well defined (c.f., 13b)}$

b. $(\llbracket nuebe \rrbracket(\llbracket p'ej \rrbracket))(\llbracket tyumuty_2 \rrbracket) = \{x : x \in * \cup \cap \text{EGG} \ \& \ \mu_{\#}(x) = 9\} \rightarrow \text{well defined (c.f., 14a),}$
 where $\cup \cap \text{EGG} = \{x : \text{ATOM}(x) \ \& \ \text{EGG}(x)\}$

However, these combinations of numerals and classifiers are not acceptable.

Syntactic facts also favour Krifka's analysis. In Chol, classifiers morphologically attach as suffixes to numerals. Although Mi'gmaq classifiers are separate words, word-order effects provide similar evidence

that numerals and classifiers form a constituent independent of the noun. As shown in (22a–b), the numeral and classifier can be separated as a unit from the noun. However, as shown in (22c), the classifier and noun cannot be separated from the numeral. This suggests that there is a tighter connection between the numeral and classifier than between the classifier and noun.

- (22) a. Etlenm-ultijig **asugom te's-ijig jinm-ug**
 laugh.PRES-PL six CL-AGR man-PL
 'Six men are laughing.'
- b. **Asugom te's-ijig** etlenm-ultijig **jinm-ug**
 six CL-AGR laugh.PRES-PL man-PL
 'Six men are laughing.'
- c. * **Asugom** etlenm-ultijig **te's-ijig jinm-ug**
 six laugh.PRES-PL CL-AGR man-PL
 'Six men are laughing.'

Li and Thompson (1981) propose that the numeral and classifier form a constituent in Mandarin Chinese; see Zhang 2011 for further references and discussion.

Finally, it is worth noting that the choice of classifier does not depend on the specific lexical noun with which it combines, as illustrated by the data in (23). A single noun, like *ja'as* 'banana', can combine with a number of classifiers which serve to specify the units being counted. That is, the classifier affects how the denotation of the noun is counted—the noun does not determine which classifier is used.

- (23) a. cha'-**ts'ijty** ja'as
 two-NC.long.skinny banana
 'two bananas'
- b. cha'-**pajl** ja'as
 two-NC.bunch banana
 'two bunches of bananas'
- c. cha'-**tyek** ja'as
 two-NC.tree banana
 'two banana trees'

Note that the evidence above only demonstrates that classifier systems in *some* languages are uniquely compatible with Krifka's theory. It has not been demonstrated that all languages have the same kind of classifier system. It is possible that there are two types, one like Krifka's and another that patterns as Chierchia's theory would predict. Indeed, the investigation of Mi'gmaq and Chol provides a template for the kind of pattern one would need to find to establish the existence of this other classifier system. Unlike Krifka's theory, Chierchia's theory predicts that it should be possible to have a lexical numeral that requires a classifier when modifying one noun, yet prohibits one when modifying another.

(24) Chierchia's Predicted Pattern

- a. NUMERAL NOUN₁, *NUMERAL CL NOUN₁
 b. *NUMERAL NOUN₂, NUMERAL CL NOUN₂

Such a pattern would demonstrate that the presence or absence of a classifier depends on the noun being modified rather than on the numeral. On the surface, one might think that English has such patterns, as shown in (25).

- (25) a. one chair, *one item of chair(s)
 b. *one furniture, one item of furniture

However, the status of this as an example of the Chierchia's predicted pattern rests on the classification of *item* and the use of the partitive preposition *of*. Are measure words like *item* and *kilo* classifiers? Unlike classifiers in other languages, these words share the same distributions with regular nouns and take nominal morphology such as plural marking. In other words, the surface evidence suggests that these words do not belong to the same type of category as classifiers (c.f., [Cheng and Sybesma 1999](#)).

Whether Chierchia's predicted pattern exists or not is an empirical matter, one that will not be resolved in this paper. However, the mere existence of Krifka-style classifiers, even if they are not universal, has some consequences for the study of syntax and semantics cross-linguistically.

4 Implications

Mi'gmaq and Chol demonstrate that, at least in some languages, the factors governing the appearance of classifiers are independent of the existence of a syntactic distinction between mass nouns and count nouns (cf., [Wilhelm 2008](#)). A weak implication of this finding is that the presence or absence of a rich classifier system is not a reliable diagnostic for whether a language has count nouns or not. However, this separation of classifier systems from nominal distinctions brings into question whether it is useful to classify languages in terms of mass-count.

As discussed in ([Bloomfield 1933](#)), what makes the mass-count distinction interesting are the corresponding semantic and syntactic patterns that are, in principle, separable from the ontological divide between "countable things" and "uncountable stuff" (see also [Bunt 1985](#); [Gillon 1992](#); [Chierchia 1998](#); [Bale and Barner 2009](#)). For example, consider the following grammatical properties associated with count syntax.

- (26) PROPERTIES OF "COUNT LANGUAGES"
- a. plural marking (e.g., -s in English)
 - b. direct numeral modification
 - c. lack of a rich classifier system
 - d. quantifier allomorphy (e.g., *many* vs. *much*)
 - e. semantically singular denotations for lexical nouns

Mandarin does not allow numerals to combine directly with nouns, has a rich classifier system, does not have a productive plural marker, and lacks allomorphy amongst its quantifiers. English, in contrast, has two lexical noun categories (mass and count), no classifier system, a productive plural, allows numerals to combine directly with nouns, and permits quantifier allomorphy. Linguists influenced by [Bloomfield \(1933\)](#) have explored the hypothesis that the clustering of these properties were in some-way connected: that non-count languages patterned like Mandarin, whereas count languages, for the large part, patterned like English.

However, previous work has shown that plural marking does not always cluster with the other properties ([Borer 2005](#); [Bale and Barner 2012](#)). Mi'gmaq and Chol demonstrate further that classifiers are independent of the nominal distinction in some languages. The fact that the first three properties in (26) do not reliably cluster together weakens the utility of classifying languages in terms of whether they have a mass-count distinction or not. Since the only correlation remaining is the relatively minor connection between quantifier allomorphy and singular denotations, one wonders whether it is better for investigative purposes to give up on term "mass-count language," which carries with it the burden of being defined with respect to all of the

properties in (26), and instead concentrate on the individual properties independent of whether they correlate or not in any given language.

References

- Aikhenvald, Alexandra. 2000. *Classifiers: A typology of noun categorization devices*. Oxford Studies in Typology and Linguistic Theory. Oxford University Press.
- Bale, Alan, and David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26:217–252.
- Bale, Alan, and David Barner. 2012. Semantic triggers, linguistic variation and the mass-count distinction. In *Count and mass across languages*, ed. Diane Massam. Oxford University Press.
- Bale, Alan, Michaël Gagnon, and Hrayr Khanjian. 2011. On the relationship between morphological and semantic markedness: the case of plural morphology. *Morphology* 21:197–221.
- Bale, Alan, and Hrayr Khanjian. 2009. Classifiers and number marking. In *Proceedings of Semantics and Linguistic Theory (SALT) XVIII*.
- Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
- Borer, Hagit. 2005. *Structuring sense*. Oxford: Oxford University Press.
- Bunt, Harry C. 1985. *Mass terms and model-theoretic semantics*. New York: Cambridge University Press.
- 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. Plurality of mass nouns and the notion of “semantic parameter”. In *Events and grammar*, ed. Susan Rothstein, 53–104. Kluwer Academic Publishers.
- Cresswell, M. J. 1976. The semantics of degree. In *Montague grammar*, ed. B. Partee, 261–292. New York: Academic Press.
- Doetjes, Jenny. 2012. Count/mass distinctions across languages. In *Semantics: an international handbook of natural language meaning*, ed. C. Maienborn, K. von Stechow, and P. Portner, volume 3, 2559–2580. De Gruyter.
- Donabédian, Anaïd. 1993. Le pluriel en arménien moderne. *Faits de Langues* 2:179–188.
- Gillon, Brendan. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15:597–640.
- Krifka, Manfred. 1995. Common nouns: A contrastive analysis of English and Chinese. In *The generic book*, ed. G. N. Carlson and F. J. Pelletier, 398–411. Chicago: Chicago University Press.
- Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. R. Baeuerle, C. Schwarze, and Arnim von Stechow. DeGruyter.
- Lyons, John. 1977. *Semantics*. London: Cambridge University Press.
- Quine, W. V. O. 1960. *Word and object*. MIT Press.
- Vázquez Álvarez, Juan J. 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 Sųlíné. *Natural Language Semantics* 16:39–68.
- Zavala, Roberto. 2000. Multiple classifier systems in Akatek (Mayan). In *Systems of nominal classification*, ed. Gunter Senft, 114–146. Cambridge, UK: Cambridge University Press.
- Zhang, Niina Ning. 2011. The constituency of classifier constructions in Mandarin Chinese. *Taiwan Journal of Linguistics* 9:1–50.