

# DISTINGUISHING CLASSIFIERS AND MEASURE WORDS: A MATHEMATICAL PERSPECTIVE AND IMPLICATIONS\*

One-Soon Her

Graduate Institute of Linguistics & Research Center of Mind, Brain, and Learning,  
National Chengchi University, Taipei 11605, Taiwan, ROC

## Abstract

Based on the insight that a measure word (M) in [Num (Adj)-C/M N] is semantically substantive, while a classifier (C) is redundant and does not block modification or quantification to N (Her & Hsieh 2010), this paper proposes a distinction of C/M from a mathematical perspective. Synthesizing the concepts of parceler (Landman 2004), divider (Borer 2005), and multiplicand (Au Yeung 2005, 2007), I follow Her (2010) and contend that while C/M both function as a multiplicand mathematically, C's value is necessarily 1 and M's is not, thus  $\neg 1$ . This offers a natural explanation to the semantic tests developed in Her and Hsieh (2010). Implications are discussed for these areas: typology of classifiers and classifier languages, correlations between numeral systems and the employment of C/M, the universal count/mass distinction at the lexical level, and first language acquisition of classifiers and numbers.

Keywords: classifier, measure word, multiplication, multiplicand, 1,  $\neg 1$

## 1. Introduction

Most, if not all, linguists would agree that there is a semantic distinction between the classifier (C) *ben* for books in (1) and the measure word (M) *xiang* 'box' in (2).

---

\* Fax: +996 2 29387466; Mobile: +886 2 926142388.

E-mail address: hero@nccu.edu.tw.

(1) 三 本 書<sup>1</sup>  
*san ben shu*  
3 C book  
'3 books'

(2) 三 箱 書  
*san xiang shu*  
3 M-box book  
'3 boxes of books'

The element between Num and N has been classified into several types<sup>2</sup>; Chao (1968), for example, lists *individual measure* (what we call *classifier*); *group measure* (e.g., 組 *zu* 'group'); *partitive measure* (e.g., 份 *fen* 'portion, share'); *container measure* (e.g., 碗 *wan* 'bowl'); and *standard measure* (e.g., 碼 *ma* 'yard'). However, it is generally agreed that it can be divided into two major groups: (individual) classifiers vs. measure words. Tai and Wang (1990: 38) characterize this C/M dichotomy as follows:

A classifier categorizes a class of nouns by picking out some salient perceptual properties, either physically or functionally based, which are permanently associated with entities named by the class of nouns; a measure word does not categorize but denotes the quantity of the entity named by noun.

Unfortunately, that is where the agreement ends. Terminology is part of the confusion. Terms used for C include 'classifier', 'sortal classifier', 'count-classifier', 'count-noun classifier', and 'qualifying classifier', and those

---

<sup>1</sup> It is difficult to give this or any other genuine classifier an English translation, as there is nothing quite comparable in the English lexicon. However, I shall argue in this paper that the nominal suffix -s marking plurality in English can be viewed as a general classifier, e.g., *ge* in Chinese.

<sup>2</sup> Note that *Num* refers to cardinal numerals only throughout the paper, not to the grammatical category of number in terms of singularity or plurality.

for M include ‘measure word’, ‘mensural classifier’, ‘massifier’, ‘quantifier’, and ‘mass-classifier’. Worse still, some use the term ‘classifier’ or ‘numeral classifier’ for both, while others use ‘measure word’ for both (e.g., Zhang 2007).

Syntactic accounts are likewise contentious. Some studies assign C/M a unified structure, which some, e.g., Li & Thompson (1981: 105), Paris (1981: 105-117), Huang (1982), Tang (1990), Croft (1994: 151), Lin (1997: 419), and Hsieh (2008), argue to be left-branching and others, e.g., Tang (2005), Cheng & Sybesma (1998, 1999), Borer (2005), Watanabe (2006), Huang *et al* (2009), right-branching. Yet, some syntactic accounts, e.g., Zhang (2011) and Li (2011), contend that both kinds of structures are required for C/M. Also, as pointed out in Her and Hsieh (2010), H&H hereafter, previous studies of Mandarin classifiers have suggested very different inventories, ranging from six hundred (Hu 1993), four hundred and twenty-seven (Huang and Ahrens 2003), two hundred (Hung 1996), to as few as just several dozen (Chao 1968, Erbaugh 1986). The major reason for this huge discrepancy is surely the confusion over what counts as a ‘classifier’ (Liang 2006:17).

Following H&H, in this paper ‘classifier’, or C, strictly refers to the kind in (1) and ‘measure word’, or M, refers only to the kind in (2) and all other non-classifier unit words. There are two reasons for doing so. First, H&H have demonstrated with accurate and reliable tests, which will be discussed in section 2, that the C/M distinction is crucial and real. Second, the main purpose of this paper is to further propose a formal and precise C/M distinction from a mathematical perspective and, in doing so, also offer a natural explanation to all the semantic tests developed in H&H. Thus, this paper also aspires to establish the use of the two terms ‘classifier’ and ‘measure word’ by identifying a set of explicit criteria.

The organization of the paper is as follows. Section 2 reviews the semantic characterization of the C/M distinction made by H&H and a set-theoretic interpretation of their insight will be offered. Section 3 then integrates and extends insights gained from Landman (2004), Borer (2005), and Au Yeung (2005, 2007) and looks closely into the mathematical properties of C/M and proposes a precise multiplication-based account for the C/M

distinction. Implications of this account, both within Mandarin Chinese and cross-linguistically, are explored in section 4, and some concluding remarks are given in section 5.

## 2. Semantic Distinction between C/M

This section serves the primary purpose of reviewing H&H, Hsieh (2009), Her (2011), and Her and Lai (2011). A set-theoretic rendition of the semantic distinction of C/M will be offered, which links this section to the discussions in section 3 on the C/M distinction from a mathematical perspective.

### 2.1 Formal Tests for the C/M Distinction

H&H observe several scope phenomena that distinguish C/M. Their first observation relates to the scope of Num, which goes beyond C and covers N and thus refers to the cardinality of a set of N. The example in (3a), referring to a miracle by Jesus, shows that numeral quantification of C scopes over N; C can thus be omitted if stylistically required. Yet, the numerals quantifying the M's in (3b) do not scope over the N and cannot be omitted without changing the meaning of the phrase. A formal test obtains, as in (4).

(3) a. 五 (張) 餅 二 (條) 魚 餵飽 五千 (個) 人  
*wu (zhang) bing er (tiao) yu weibao wuqian (ge) ren*  
 5 C loaf 2 C fish feed-full 5000 C person  
 '5000 people were fed by 5 loaves and 2 fish.'

b. 五 \*(籃) 餅 二 \*(箱) 魚 餵飽 五千 \*(組) 人  
*wu (lan) bing er (xiang) yu weibao wuqian (zu) ren*  
 5 M-basket loaf 2 M-box fish feed-full 5000 M-group person  
 '5000 groups of people were fed by 5 baskets of loaves and 2 boxes of fish.'

#### (4) C/M Distinction in Numeral Quantification Scope

Given a well-formed [Num K N], if Num scopes over N,  
then K = C; otherwise, K = M.

Another observation relates to the scope of adjectival modification in a [Num Adj-C/M N] phrase. Though the bare adjectives allowed here are strictly restricted to size, again the pre-C adjective goes beyond C and modifies N, while the pre-M adjective modifies M only. Compare (5) and (6). As an anonymous reviewer keenly observes, some elements can be C with some nouns, and M with others. The same test used in (5) can be useful in such cases too. 條 *tiao* is thus an M in (7) and C in (8).

(5) 一 大 箱 蘋果                    ≠ 一 箱 大 蘋果 (H&H 13a)  
*yi da xiang pingguo*                    *yi xiang da pingguo*  
one big M-box apple                    one M-box big apple  
'one big box of apples'                    'one box of big apples'

(6) 一 大 顆 蘋果                    = 一 顆 大 蘋果 (H&H 13b)  
*yi da ke pingguo*                    *yi ke da pingguo*  
one big C apple                    one C big apple  
'one big apple'                    'one big apple'

(7) 一 小 條 香菸                    ≠ 一 條 小 香菸  
*yi xiao tiao xiangyan*                    *yi tiao xiao xiangyan*  
one small M-box cigarette                    one M-box small cigarette  
'one small box of cigarettes'                    'one box of small cigarettes'

(8) 一 小 條 鯉魚                    = 一 條 小 鯉魚  
*yi xiao tiao liyu*                    *yi tiao xiao liyu*  
one small C carp                    one C small carp  
'one small carp'                    'one small carp'

Adjectival modification in the [Num C/M N] configuration is also relevant to the C/M distinction, as shown in (9) and (10).

(9) 大顆的 蘋果 = 大 蘋果  
*da-ke-de pingguo da pingguo*  
 big-C-DE apple big apple  
 'big apple(s)' 'big apple(s)'

(10) 大箱的 蘋果 ≠ 大 蘋果  
*da-xiang-de pingguo da pingguo*  
 big-M-box-DE apple big apple  
 'apples that come in big boxes' 'big apple(s)'

Like the pattern seen in (5)-(8), C in (9) does not block the scope of the adjective to cover N, but M in (10) does. Another formal test thus obtains.

**(11) C/M Distinction in Adjectival Modification Scope**

If either [Num A-K N] = [Num K A-N] or [A-K-*de* N] = [A-N] semantically and A refers to size, then K = C, and K ≠ M.

Consequently, whether the adjective modifies C or N, it has the same scope. A pre-C adjective and a pre-N adjective in the same phrase thus cannot contradict each other, as shown in (12a-b), where neither of the examples has a congruent reading, as the apples cannot be big and small at the same time. Yet, a pre-M adjective can contradict a pre-N adjective, as shown in (13a-b), where the box is big and the apples are small.

(12) a.#一 大 顆 小 蘋果 (H&H 15a)  
*yi da C xiao pingguo*  
 one big C small apple

b. #大大的 一 顆 小 蘋果 (H&H 15b)  
*dadade yi ke xiao pingguo*  
big one C small apple

(13) a. 一 大 箱 小 蘋果 (H&H 14a)  
*yi da xiang xiao pingguo*  
one big M-box small apple  
'one big box of small apples'

b. 大大的 一 箱 小 蘋果 (H&H 14b)  
*dadade yi xiang xiao pingguo*  
big one M-box small apple  
'one big box of small apples'

As pointed out earlier, an adjective referring to size can be formed by joining the A and a C/M. Consider the contrast between (14a) and (14b).

(14) a. #大顆的 小 蘋果  
*da-ke-de xiao pingguo*  
big-C-DE small apple

b. 大箱的 小 蘋果  
*da-xiang-de xiao pingguo*  
big-M-box-DE small apple  
'small apples that come in big boxes'

Again, in (14a), the adjectival component in the A-C formation in fact modifies not the internal C, but the external N. Yet, in (14b), the A in A-M formation modifies the internal M, not the external N. Again, a formal test obtains, as in (15).

(15) C/M Distinction in Antonym Stacking

If [Num A<sub>1</sub>-K A<sub>2</sub>-N] or [A<sub>1</sub>-K-*de* A<sub>2</sub>-N] is semantically congruent and A<sub>1</sub> and A<sub>2</sub> are antonyms, then K = M and K ≠

C.

Finally, the respective interaction between the adjectival expressions of the A-C/M-*de* formation and C/M provides further support to H&H's position. Consider (16) and (17) and especially the contrast between (16b) and (17b).

(16) a. 一 顆/粒      大顆/粒的      蘋果  
*yi ke/li da ke/li-de pingguo*  
one C      big-C/C-DE      apple  
'one big apple'

b. \*一 顆/粒      大箱/包的      蘋果  
*yi ke/li da-xiang/bao-de pingguo*  
one C      big-M-box/bag-DE      apple

(17) a. 一 箱/包      大箱/包的      蘋果  
*yi xiang/bao da-xiang/bao-de pingguo*  
one box/bag      big-M-box/bag-DE      apple  
'one box/bag of apples in big bags/boxes'

b. 一 箱/包      大粒/顆的      蘋果  
*yi xiang/bao da-li/ke-de pingguo*  
one M-box/bag      big-C-DE      apple  
'one box/bag of big apples'

Example in (16a) and (17a) indicate that doubling for both C/M is fine, as long as each of the doubled elements is well-formed independently and together they do not conflict pragmatically. However, as shown in (16b), the



C-M sequence is intrinsically ill-formed, and yet the M-C order in (17a) is permitted. Thus, another simple test obtains, as in (18).

#### (18) C/M Distinction in Doubling

If [Num  $K_1$  A- $K_2$ -*de* N] is well-formed and  $K_1$  and  $K_2$  are not synonymous, then  $K_1 = M$ ,  $K_1 \neq C$ , and  $K_2 = C/M$ .

All the tests discussed above, whether proposed in H&H or in this study, can indeed be unified under the fact that M blocks the modification by adjectives and quantification by numerals to the noun, while C does not. Thus, metaphorically, M is opaque while C is transparent.

## 2.2 Semantic Characterization of C/M

H&H propose the use of two sets of concepts from philosophy to characterize the C/M distinction: Aristotle's distinction between essential and accidental properties and Kant's distinction between analytic and synthetic propositions. Robertson (2008) offers the definition in (19) for the former dichotomy. Kant's distinction is defined in (20), cited from Rey (2003), along with examples.

#### (19) Essential Property vs. Accidental Property

*P* is an *essential property* of an object *o* just in case it is necessary that *o* has *P* whereas *P* is an *accidental property* of an object *o* just in case *o* has *P* but it is possible that *o* lacks *P*.

#### (20) Analytic Proposition vs. Synthetic Proposition

*Analytic proposition*: a proposition whose predicate concept is contained in its subject concept; e.g., *all bachelors are unmarried*.

*Synthetic proposition*: a proposition whose predicate concept is not contained in its subject concept; e.g., *all bachelors are happy*.

An essential property thus can be translated into an analytic proposition where the subject is the object in question, e.g., *bachelors*, and the predicate

the condition, e.g., *are unmarried*. An accidental property can only be translated into a synthetic proposition where the subject again is the object in question, e.g., *bachelors*, and the predicate the unnecessary condition, e.g., *are happy*. Consequently, a modifier denoting an essential property is redundant, as in *unmarried bachelors*; in other words, the semantic content of *unmarried bachelors* and that of *bachelors* is the same. However, a modifier denoting an accidental property is not redundant, as in *happy bachelors*, where the subtraction of *happy* causes the loss of some semantic substance.

H&H argue for a C/M distinction along these two dichotomies: C denotes an essential property of the noun in [Num C N] and can thus be paraphrased as the predicate concept in an analytic proposition with the noun as the subject concept; M denotes accidental properties of the noun in [Num M N] and can only be restated as the predicate concepts in synthetic propositions with the noun as the subject concept. Consider the C in (21a) and M in (21b).

(21) a. 三百 顆 蘋果

*sanbai ke pingguo*

300 C apple

'300 apples'

b. 三百 噸 蘋果

*sanbai dun pingguo*

300 M-ton apple

'300 tons of apples'

In (21a), *ke*, originally a noun meaning a ball-shaped object, refers to a 3D round shape. Thus, as a C, *ke* can be seen as an essential property of apples, and (21a) can thus be stated as the predicate concept in an analytic proposition, *the 300 apples are 3D and round*. Note that even the so-called general C, 個 *ge*, requires that the N is an inherently discrete unit. However, (21b) can be paraphrased as a synthetic proposition, *the apples' mass is 300 tons* and 300 tons is an accidental property of the apples in question. This characterization echoes Adams and Conklin's (1973:2) insight that C *qualifies*

the head noun while M *quantifies* it. It likewise confirms W. Li's (2000: 1117) observation that *classifiers are semantically redundant* and also Greenberg's (1974: 84) similar but less precise observation that classifiers are redundant when translated into a non-classifier language like English. Thus, crucially, though C does have semantic content, it is redundant in the context of [Num C N].

### 2.3 Distinction of C/M in Set-theoretic Terms

What the above semantic characterization amounts to is that C's semantic attributes constitute a subset of those of N. M, on the other hand, does contribute to the phrase's total semantic value, the same way modifiers do.

#### (22) C/M Distinction in Set-theoretic Terms

Given a well-formed phrase [Num K N],  $X$  the set of semantic attributes denoted by K, and  $Y$  the set of semantic attributes denoted by N, K is C *iff*  $X \subset Y$ ; otherwise, K is M.

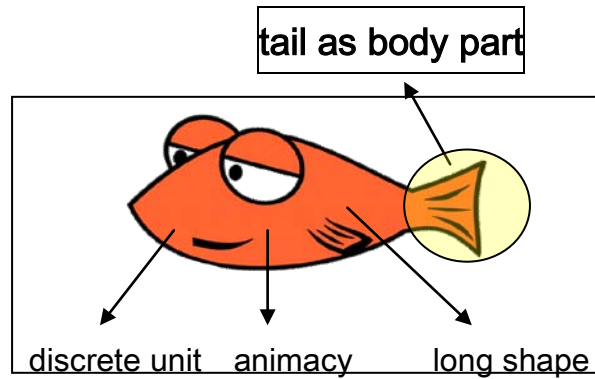
C's semantic content being a subset of that of N's is the reason why any modification or quantification on C is also on N. M, however, does have semantic properties that N does not have; thus, modification and quantification on M do not scope over N.

### 2.4 C as Profiler

What is C's function then, if it is semantically redundant? Hsieh (2009), Her (2011), and Her and Lai (2011) propose that C serves as a profiler, in the sense of Fillmore (1982) and Langacker (1987), highlighting an inherent semantic attribute of N. Three illustrated examples are given in (23a-c).

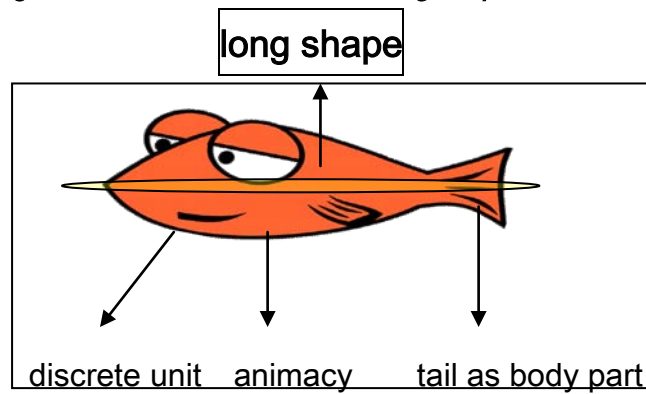
- (23) a. 一 尾 魚  
 yi wei yu  
 1 C-tail fish  
 '1 fish'

Figure 1. *N-fish* as Frame and *C-tail* as Profile



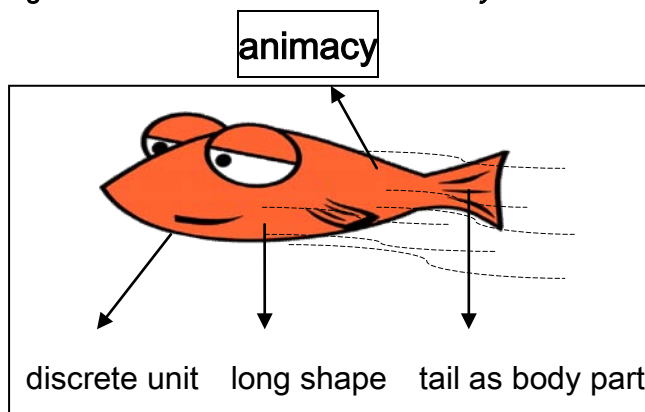
- b. 一 條 魚  
 yi tiao yu  
 1 C-long-shape fish  
 '1 fish'

Figure 2. *N-fish* as Frame and *C-long-shape* as Profile



c. 一 隻 魚  
 yi zhi yu  
 1 C-animacy fish  
 '1 fish'

Figure 3. *N-fish* as Frame and *C-animacy* as Profile



As shown schematically in the three figures, *N*, *yu* 'fish' in this case, provides the base, or frame, within which *C* profiles an inherent feature. This view nicely accounts for the fact that a *C* must select its own class of nouns, for a *C* cannot profile a feature that the noun does not already have. Under this view, *C*'s primary function is profiler, with classification merely as a by-product.

Since *N* is complete in itself with or without a profiling *C*, languages like Chinese in fact need not require *C*, contrary to the conventional view. This is partly the consequence of *C*'s unique mathematical properties, the theme of section 3. Also, in 4.3, we will discuss further this misconceived notion of *C*'s being necessary.

### 3. C/M Distinction in Mathematical Value

In this section I will support the tentative proposal made in Her (2010) that *C* and *M* are both multiplicands, in a mathematical sense, that link the numeral and the noun.<sup>3</sup> The central idea is inspired by the concepts of parceler (Landman 2004), reviewed in 3.1, divider (Borer 2005), discussed in 3.2, and

<sup>3</sup> Using *san da meigui* or *three dozen roses* as an example, in the equation  $3 \times 12 = 36$ , 12 is the multiplicand, or the number in a group, and 3 the multiplier, or the number of groups.

multiplicand (Au Yeung 2005, 2007), summarized in 3.3, and can be seen as a synthesis of them all and will be presented in 3.4.

### 3.1 Landman (2004) : C/M as Parcelers

In exploring the linking between C/M and collectivity, Landman (2004) contends that *time*, as a verb meaning to *multiply*, can be viewed as a parceler, or event classifier, which can bring on a collective interpretation of the head noun. Thus, in (24) *time* seems to function as a C/M, which parcels the boys into three-person groups in this case.

(24) Four times three boys met in the park.

Interpretation (a): The sum of twelve boys met in the park.

Interpretation (b): Four groups of three boys met in the park.

While Landman is interested in the reading of *four groups of three boys*, as a consequence of *four times three boys*, Au Yeung (2007) suggests the possibility of likewise viewing *twelve boys* as a consequence of *one times twelve boys* and hints at fitting such expressions into the syntax of [Num C/M N], as in (25).

(25) [Number four] [C/M times] [three?] [Noun boys]

(26) 四 乘 三 個 男孩

*si cheng san ge nanhai*

four time three C boy

'four times three boys'

Interpretation a: The sum of twelve boys.

Interpretation b: Four groups of three boys.

The Chinese counterpart in (26) thus also has two readings: sum and grouping, neither of which, however, can be seen as a straightforward [Num C/M N] phrase with *cheng* as the C/M. I contend that the two readings of (24)

and (26) arise from the two parses that both phrases receive, as shown in (27) and (28).

- (27) English:      a. [Number four times three] [Noun boys]  
                          b. [Number four] times [Number three boys]

- (28) Chinese :    a. [Number si cheng san] [C/M ge] [Noun nanhai]  
                          b. [Number si] cheng [Number san [C/M ge] [Noun nanhai]]

The parses in (27a) and (28a) produce the first reading only, where the multiplicand is *three* or *san* alone, while the parses in (27b) and (28b) give both the first reading and the second reading, where the multiplicand is the entire phrase *three boys* or *san ge nanhai*; in other words, *three boys* can be seen as three individuals or as a three-boy group. This can be demonstrated by reversing the order of the multiplier and the multiplicand, as in (29) and (30).

- (29) English:      [Number three boys] times [Number four]  
 (30) Chinese :    [Number san [C/M ge] [Noun nanhai]] cheng [Number si]

*Time* or *cheng* is thus nothing like a C/M formally, if Chinese C/M serves as a model. However, note that in the second reading, where the multiplicand is taken to be a three-boy unit, the [Num N] phrase can indeed be rephrased as a single nominal unit, as shown in (31b) and (32b). Notice the contrast between the (a) and (b) phrases.

- (31) a. four times one three-boy group                    (4 ×1 3-boy group)  
       b. four three-boy groups                                (4 3-boy groups)

- (32) a. 四    乘        一    個    三人小組                    (4 × 1 C 3-person unit)  
       *si    cheng    yi    ge    san-ren-xiaozu*  
       four   time    one C   3-person-unit  
       ‘4 times 1 3-person unit’

- b. 四 個 三人小組 (4 C 3-boy group)  
*si ge san-ren-xiaozu*  
 four C three-person-unit  
 ‘four three-person units’

The fact that (a) and (b) have exactly the same value means that *times one* or *cheng yi* is redundant, as  $n \times 1 = n$ . The crucial difference between English and Chinese is that while the classifier *ge* is a free morpheme or clitic on Num in Chinese, the suffix *-s* must be bound to N in English. I shall indeed argue in 3.4 that all languages with a multiplication-based number system employ the structure of [Num  $\times 1$  N], where  $\times 1$  can be silent in some languages (e.g., Archaic Chinese), expressed as C’s in classifier languages (e.g., modern Chinese), or as number affixes on N in inflectional languages (e.g., English).

However, in spite of my rebuff of Landman’s (2004) idea to view *time* in English as an event classifier, his thesis that, given the fact that sums and mass cannot be counted, C/M is needed as a parceling device, which partitions a countable parcel from the sum or mass denoted by N, is still insightful. Consider the examples in (33).

- (33) a. 四 瓶/公斤 水 (from mass to count)  
*si ping/gongjin shui* (4  $\times$  bottle/kilo water)  
 four M-bottle/kilo water  
 ‘4 bottles/kilos of water’
- b. 四 組/隊 男孩 (from sum to group)  
*si zu/dui nan-hai* (4  $\times$  group/team boy)  
 four M-group/team boy  
 ‘four groups/teams of boys’
- c. 四 個/位 男孩 (from sum to individual)  
*si ge/wei nan-hai* (4  $\times$  1 boy)  
 four C boy  
 ‘four boys’



Mass, e.g., *shui* ‘water’ in (33a), can only be counted if measured by units, e.g., *kilo*, or contained in containers, e.g., *bottle*. Likewise, *zu* ‘group’ and *dui* ‘team’ parcel a countable unit from the sum of boys. This view thus assumes that Chinese *nanhai* ‘boy’ in (33b-c) is *sum of boys*, thus similar to English *boys* in (33b’-c’). However, an obvious alternative is to see *nanhai* ‘boy’ in (33b-c) as singular. In (34), for example, the noun *zhao* ‘trillion’ is a mathematically defined discrete number.

(34) a. 四 (個) 兆<sup>4</sup>  
*si ge zhao* (4 × 1 trillion)  
 four C trillion  
 ‘4 trillion’

With or without the optional C, the total value of four trillion is arrived at straightforwardly by four instances of a single trillion, implying a multiplication basis in this structure (Au Yeung 2007). This is the position I will advocate in 3.4 for Chinese as well as English. Furthermore, Landman’s insight of C/M as parcelers sheds no light on the C/M distinction or on the fact that each C/M, as a parceler, can be seen as having a mathematical value in parceling. In (35), for example, the C has precisely the mathematical value of *one*, the M in (36) the value of *twelve*, and the M in (37) the value of *kilo*.

(35) 四 (個) 人  
*si ge ren* (4 × 1 person)  
 four C person  
 ‘4 persons’

---

<sup>4</sup> The number *four trillion* is much more commonly expressed in Chinese without the classifier *ge*. However, (34) is quite acceptable, and especially common in China. Note (34) may also indicate four instances of the number *zhao* ‘trillion’ or the character 兆, but such a meaning is irrelevant to our discussion here.

(36) 四 打 玫瑰  
*si da meigui* (4 × dozen roses = 4 × 12 roses)  
 four M-dozenrose  
 ‘4 dozens of roses’

(37) 四 公斤 鹽  
*si gongjin yan* (4 × kilo salt)  
 four M-kilo salt  
 ‘4 kilos of salt’

Note that the only mathematical value allowed to be vacuous, thus optional, as the multiplicand *n* in (*m* × *n*) is *one*. The expression in (35a) is thus equivalent to [*4* (*×1*) *person*]. *C* is of course normally required in Chinese; however, it is stylistically allowed to be absent in certain formal literary contexts, which indicates *C* contributes no additional semantic content that the head noun does not already have.<sup>5</sup> Such is the case for the optional *C* in (35a). In contrast, *da* ‘dozen’ in (36a) and *gongjin* ‘kilo’ in (37a) each contribute crucial information in the equations [*4* × *dozen roses* (=4 × 12 roses)] and [*4* × *kilo salt*] respectively, and thus cannot be optional. They have a value other than 1, and thus are *M*, not *C*. Again, this is the view I will argue for in 3.4.

### 3.2 Borer (2005) : C/M as Dividers

Borer (2005, chapter 4) takes the stand that all nouns in all languages are mass by default. One reason for this is data like those in (38) and (39), where the conventional putative count/mass distinction is shown to be useless in English.

(38) A wine/wines, a love/loves, a salt/salts (on count reading)

(39) There is dog/stone/chicken on the floor (on mass reading)

---

<sup>5</sup> In Khmer, *a.k.a.* Cambodian, however, the absence of the normally required *C*'s is stylistically *less* formal (Greenberg 1990[1972]: 168).

Under this view, a count interpretation can only be obtained when a noun appears in the syntactic configuration projected by a divider, such as a classifier in Chinese or the plural inflection in English. Au Yeung (2007: 841) succinctly illustrates this view with (40-43), where all four nouns are lexically mass and *yan* and *salt* are interpreted as mass in (40-41) given that there is no divider in the syntactic structure but *mao* and *cat* only receive a count interpretation in (42-43) because of the divider *zhi* or *-s*.

- (40) Mass : [Number *henduo* [Noun *yan*]]  
 (41) Mass : [Number *much* [Noun *salt*]]  
 (42) Count : [Number *san* [Divider *zhi* [Noun *mao*]]]  
 (43) Count : [Number *three* [Divider *cat-s* [Noun *cat* ]]]

A similar but more conservative view is found in Chierchia (1998) and Tai (2003: 312) and implied in Wu and Bodomo (2009), where they propose that, while a mass/count distinction does exist in English, all nouns in Chinese are mass. Under this view, C/M share the same function of ‘carving out’ discrete, bounded units from otherwise non-discrete mass. However, as pointed out in Her and Hsieh (2010) and Gebhardt (2011), this view may be problematic. The noun *zhao* ‘trillion’ in (34a), for example, is a number mathematically defined to be discrete, the collective totality of which thus cannot possibly be conceived as non-discrete mass. The reading of four natural units of trillion in (34a) is not accidental but entirely necessary.

For those that see a count/mass distinction in Chinese, only count nouns, i.e., nouns that denote discrete units, co-occur with C. Ahrens (1994: 204) offers this distinction:

Classifiers can only classify over a limited and specific group of nouns, while measure words can be used as a measure for a wide variety of nouns.

C’s thus not only require a count noun, but also one with a specific inherent property, independent of the C. Even the general C, 個 *ge*, in fact

does not select all count nouns. Mass nouns, i.e., nouns denoting non-discrete entities, cannot co-occur with C. Cats, for example, come in discrete forms and are nonhuman animals. The classifier *zhi* requires a complement noun that denotes a nonhuman animate entity, while a measure word like *gongjin* 'kilogram' can take whatever noun, count or mass, denoting something whose weight can be measured perceptually. The view of count/mass unification thus wrongly predicts that C in Chinese can take mass nouns and turn them into count. Unlike the plural suffix *-s*, which is much more liberal in allowing putative mass nouns, as seen in (38), Chinese C's do not freely allow putative mass nouns. Thus, (44) is ill-formed with a mass noun. (45) is also ill-formed, for though *ren* 'person' is count, it denotes human, and yet, *zhi* must select a nonhuman animate nouns.

(44) Count : \*[Number *henduo* 'lots of' [Divider *zhi* [Noun *yan* 'salt']]]

(45) Count : \*[Number *henduo* 'lots of' [Divider *zhi* [Noun *ren* 'person']]]

(46) Count : [Number *henduo* 'lots of' [Divider (*zhi*) [Noun *mao* 'cat']]]

The insight from Borer (2005) is that the English plural marker *-s* functions the same as the Chinese C. Since Borer (2005) makes no distinction between C/M, it is important to stress that it is C, not M, in Chinese that share the same function as the nominal suffix *-s*.

### 3.3 Au Yeung (2005, 2007): C/M as Multiplicands ( $\times n$ token<sub>object</sub> $\times$ unit)

Assuming the necessity of coding the multiplication operation in language, Au Yeung (2005, 2007) argues convincingly for the essential role of the multiplicative identity, *1*, in the emergence of C/M. In the number calling system of both Chinese and English, for example, all multiplicands at ten and above are called. Take the number 6543 as an example.

(47) 六 千      五 百      四 十 三  
 liu qian      wu bai      si shi san  
 six thousand fivehundred four ten three  
 ‘Six thousand five hundred and forty-three’

The Chinese number system is famously regular in its decimal pattern, ( $n \times base$ ) +  $m$ , where  $m < base$  (e.g., Comrie 2006). The number 6543 can thus be derived as shown in (48) and (49).

(48) Derivation of the number 6543 in Chinese (I)

$$(6 \times 10^3) + (5 \times 10^2) + (4 \times 10^1) + (3 \times 10^0)$$

(49) Derivation of the number 6543 in Chinese (II)

$$(6 \times 1000) + (5 \times 100) + (4 \times 10) + (3 \times 1)$$

Note that all operators, i.e., multiplication ( $\times$ ) and addition ( $+$ ), are silent; yet, crucially, all bases, e.g., *qian* ‘thousand’ ( $10^3$ ), *bai* ‘hundred’ ( $10^2$ ), and *shi* ‘ten’ ( $10^1$ ), must be pronounced, thus leaving *ge* ( $10^0$ ) as the sole exception. Thus, Au Yeung (2005) notes an asymmetry between the rightmost and other digits, i.e., *qian*, *bai*, and *shi*. When a number is called in Chinese, the only phonetically null but numerically present slot is *ge*, as shown in Table 1, where the single digit 3 can be consistently represented by the multiplication formula as  $3 \times 1$ .

**Table 1. Asymmetry of the rightmost digit**

Number 6543	6	5	4	3
Position naming	千位( $10^3$ )	百位( $10^2$ )	十位( $10^1$ )	個位( $10^0$ )
	<i>Qian-wei</i>	<i>Bai-wei</i>	<i>Shi-wei</i>	<i>Ge-wei</i>
Digit value calling	<i>Liu-qian</i>	<i>Wu-bai</i>	<i>Si-shi</i>	<i>San-GE<sub>silent</sub></i>
Number calling	六 千    五 百    四 十    三(*個)			
	(6 qian) (5 bai) (4 shi) (3 × ge)			
	<i>Liu qian wu bai si shi san GE<sub>silent</sub></i>			

Though due to the scope of the paper we will not look into the historical development of C/M in Chinese, it is apparent that this silent *ge* 個 in number calling has precisely the phonological shape of that of the general classifier *ge* 個 in cannot be accidental. As Au Yeung (2005: 201) points out: ‘The silent classifier in the form of 1<sub>GE</sub> in the CL slot could serve as a seed for the noisy sortal classifier to grow’.<sup>6</sup>

Surprisingly however, Au Yeung (2005) does not follow through this simple mathematical value of *ge* as C, which is quite simply the multiplicand 1; instead, he further pursues a more complicated formula and takes a C as having a numerical value ‘one token<sub>object</sub> per unit’ while an M as ‘*n* token<sub>object</sub> per unit’. Au Yeung (2007) further interprets ‘token<sub>object</sub>’ as the size of the ‘unit’, or the set. Let’s look at two examples.

(50) 三 個 球  
*san ge qiu* ( $3 \times (1 \times 1set) \times qiu$ )  
 3 C ball  
 ‘3 balls’

(51) 三 對 球  
*san dui qiu* ( $3 \times (2 \times 1set) \times qiu$ )  
 3 M-pair ball  
 ‘3 pairs of balls’

The universal *1set* in the likewise universal formula ( $n \times 1set$ ) is said to express Borer’s (2005) notion of C/M as dividers and that of Landman’s (2007) as parcelers. Thus, for our purpose, we can reasonably claim that in Au Yeung’s formula the distinction between C and M rests on the value of *n*, i.e., C, if  $n=1$ , and M, if  $n \neq 1$ .

(52) Au Yeung’s (2005, 2007) Formula

[Num K N] = [Num  $\times (n \times 1set) \times N$ ], where K=C if  $n=1$  and K=M if  $n \neq 1$ .

---

<sup>6</sup> Note that Au Yeung (2007) does not distinguish C and M and uses ‘classifier’ for both, while Au Yeung (2005) does distinguish ‘sortal classifier’ (C) and ‘non-sortal classifier’ (M).

Au Yeung (2005, 2007) is the first to make such a mathematically precise C/M distinction. Next, I shall discuss Her's (2010) simpler proposal in the same spirit and demonstrate that Au Yeung's idea of C/M as ( $n \times 1set$ ) is unnecessarily complicated.

### 3.4 Her (2010): C/M as Simple Multiplicands

Her (2010) tentatively proposed that, if C/M is to be interpreted as having a mathematical value, then the only possible mathematical function linking Num and C/M is multiplication, where C as the multiplicand is necessarily of the value 1. This explains why C is semantically redundant in [Num C/M N]. M, on the other hand, is semantically substantive, and thus mathematically must have a value that is not 1. This idea of C as ' $\times 1$ ' was first explicitly suggested by Greenberg (1990[1972]:172): "all the classifiers are...merely so many ways of saying 'one' or, more accurately, 'times one'." an idea endorsed of late by Yi (2009, 2011). Her (2010) takes this idea further and sees C/M as ' $\times x$ '. The precise C/M distinction is stated as (53). Note that this  $x=1$  condition for C is both necessary and sufficient, thus the use of *iff*, not *if*.

#### (53) Her's (2010) Formula

[Num K N] = [Num  $\times x$  N], where  $K = C$  *iff*  $x = 1$ , otherwise  $K = M$ .

While all C's equal 1, and M's express all other, indeed infinite, values, which can be numerical or non-numerical, as long as it is not 1. A numerical M can denote a specific number, e.g., 雙 *shuang* 'pair', 對 *dui* 'pair', and 打 *da* 'dozen'; or an unspecified number, e.g., 群 *qun* 'group', 套 *tao* 'set', and 組 *zu* 'team'. A non-numerical M can in turn have a fixed value, such as standard measures, e.g., 磅 *bang* 'pound' and 公尺 *gongchi* 'meter', or a variable value, e.g., 箱 *xiang* 'box' and 瓶 *ping* 'bottle'. The crucial distinction is  $C=1$ ,  $M \neq 1$ .

This mathematical C/M distinction explains why C is a closed set and a functional category, while M an open set and a lexical category. Seeing C as multiplicand 1 also offers a simple explanation for the fact that in Chinese, as

seen in (34) and (35), the normally required C may be omitted for stylistic considerations. This omission is possible because C as multiplicand 1 is mathematically vacuous, and semantically it merely serves to highlight a certain feature that the N already has. Two examples are given in (54).

(54) a. 兩 人 三 腳 的 比賽<sup>7</sup>  
*liang ren san jiao de bisai*  
2 person 3 foot DE race  
'a two-person-three-foot race'

b. 一 人 一 信 運 動  
*yi ren yi xin yundong*  
1 person 1 letter campaign  
'a 1-man-1-letter campaign'

In Beijing Mandarin, classifiers are more freely omitted than in Taiwan Mandarin. In (55) are listed three examples from the popular movie *非誠勿擾 Fei Cheng Wu Rao (If You are the One)* and its sequel; all three examples are uttered by characters supposedly from Beijing.

(55) a. 登 一 徵 婚 廣 告  
*deng yi zhenghun guanggao*  
publish 1 seeking-marriage ad  
'put out a seeking-marriage ad'

b. 捐 兩 器 官  
*juan liang qiguan*  
donate 2 organ  
'donate a couple of organs'

---

<sup>7</sup> Different from Taiwan, the same game in China is more commonly referred to as 二人三足 *er ren san zu* '2 person 3 foot'.



c. 辦垮 過 三 公司  
*bankua guo san gongsi*  
 run-fail ASP 3 company  
 ‘ran and bankrupted 3 companies’

Though routinely overlooked by formal linguists, the fact that Mandarin C’s can be omitted in some contexts has been duly noted by some pedagogical grammarians to distinguish C’s and M’s.

...因為個體量詞不表量，故可省略，“一個杯子”=“一杯子”...  
 ... *Yinwei getiliangci bu biao liang, gu ke shenglue, “yi ge beizi”=“yi beizi”...*  
 (because classifiers do not express quantity, they can be omitted, “1 C cup”=‘1 cup’...) (Ma 2011)

個體量詞：一張床（一床）、一頭牛（一牛）、一個人（一人），省略後語意不變。*Getiliangci: yi zhang chuang (yi chuang), yi tou niu (yi niu), yi ge ren (yi ren), shenglue hou yuyi bu bian.* (Classifiers: 1 C bed (1 bed), 1 C ox (1 ox), 1 C person (1 person), can be omitted without any change in meaning. (Wang 2004: 113)

In teaching Chinese for the special purpose of technology, Chu (1994) notes that in a 1.67 million-character corpus of science textbooks, 1731 instances of [Num N] are found. In fact, in languages with a less developed classifier system, for example, the Tibeto-Burman languages Tibetan, Jingpho, and Cuona Monpa, the use of C is often optional (e.g., Jiang 2006: 18).

Another interesting fact related to optional C as *x1* is that when Num is 1, it can be omitted, in the right environment. The same is true for M. See (56) and (57).

(56) 這 (一) 朵 玫瑰  
*zhe (yi) duo meigui (this 1×1 rose)*  
 this 1 C rose  
 ‘This rose’

- (57) 這 (一) 箱 玫瑰  
*zhe yi xiang meigui* (*this 1×box rose*)  
 this 1 M-box rose  
 ‘This box of roses’

C/M is thus best treated as a clitic in Chinese, as convincingly argued for by Yang (2001). Therefore, details aside, for the numeral *1* to be optional, the following C/M requires a proper host in lieu of the missing Num. Thus, unlike Cantonese, Mandarin Chinese does not allow C/M in sentence-initial positions (Au Yeung 2005).

Compared with Au Yeung’s (2005, 2007) formula, the obvious difference is that he sees C/M uniformly as (*n×1set*), while I take it uniformly as a simple value, numerical or non-numerical. Let’s look at an example with M and a mass noun.

- (58) 四 瓶 水  
*si ping shui* (Au Yeung: *4×(n× 1bottle)×water*), (Her: *4×bottle×water*)  
 4 M-bottle water  
 ‘4 bottles of water’

Under my view, *ping* has the simple value of *bottle*. Whatever precise value it may be further reduced to according to the context, *bottle* is not *1* and thus not a C.<sup>8</sup> Yet, under Au Yeung’s view, every C/M must be seen as a number of countable tokens of a set, e.g., *n× 1bottle*, where *n* is the countable tokens of the one set, the bottle. This is precisely how he sees the example in (59):

...one bottle of water may be imagined as containing only three water molecules. Under this circumstance, when water molecules are counted in terms of bottles, what is actually counted in one counting action is that every

---

<sup>8</sup> Under this view, *4×bottle×water* can of course be seen, rather redundantly, as *4×1bottle×water*. However, it is important to note that ‘*1bottle*’ is not numerical and does *not* equal the numerical *1*.

3-token molecule is put aside or this 3-token<sub>water-molecule</sub> unit is put into a bottle.  
 As a result, four bottles of water in this case means that there are four  
 3-token<sub>water-molecule</sub> units... (Au Yeung 2005: 24-25)

This view suggests that all nouns, in languages with C/M, are count, never mass. Or there is simply no distinction between count and mass at the lexical level, each seen as a collection of countable tokens in [Num C/M N]. It does take considerable imagination, as well as scientific knowledge of course, to see a bottle of water as a certain number of water molecules. However, a bottle of water can also be seen as a collection of various atoms, quarks, or strings (as in the super string theory). The possibilities are infinite. Consider another example with a count noun.

(59) 四 瓶 彈珠  
*si ping danzhu* (Au Yeung:  $4 \times (n \times 1 \text{ bottle}) \times \text{marbles}$ ), (Her:  $4 \times \text{bottle} \times \text{marble}$ )  
 4 M-bottle marble  
 '4 bottles of marbles'

Here the only interpretation is that there are four bottles of marbles, and thus a certain number of marbles, whose materials and composition are unknown, and more importantly entirely irrelevant. Yet, Au Yeung's view again opens up infinite other possibilities, in terms of molecules, atoms, quarks, etc., depending on the materials that the marbles in question are made of. Such a view bears no psycholinguistic reality and also fails to offer a sensible interpretation in some cases. Consider (60).

(60) 四 躡 報酬  
*si pa baochou* (Au Yeung:  $4 \times (n \times 1\%) \times \text{remuneration}$ ),  
 4 M-% remuneration (Her:  $4 \times \% \times \text{remuneration}$ )  
 '4% of remuneration'

There is no telling in what form this remuneration refers to would come; it could be in the form of a currency or indeed anything else concrete or abstract that is perceived as something with value in the specific culture. There is thus

no way to determine the value of  $n$  in  $(n \times 1\%)$ , when there is no way to determine what exactly the head noun refers to. Abstract mass nouns present an even more serious problem.

- (61) 一 半            快樂  
*yi ban            kuaile*            (Au Yeung:  $1 \times (n \times 1 \text{ half}) \times \text{happiness}$ ),  
 1 half            happiness            (Her:  $1 \times \text{half} \times \text{happiness}$ )  
 ‘half of the happiness’

There are no conceivable countable units that the notion of ‘happiness’ or ‘compassion’ can be reduced to. After all, there are no molecules or atoms that happiness or compassion can be reduced to. Thus, splitting *ban* ‘half’ into  $(n \times 1 \text{ set})$  in fact allows no sensible interpretation for (61).

To summarize, Au Yeung’s (2005, 2007) view that C/M be seen as multiplicands is of great insight, but treating C/M as  $(n \times 1 \text{ set})$  is unnecessarily complicated and also unworkable in many cases. Under my view, C/M enter a multiplication relation with Num, where  $C=1$  and  $M \neq 1$ . Thus, *bottle*, *percent*, and *half* in (59), (60), and (61) respectively, each denote precisely a simple variable value, regardless of the referent  $N$ , thus *bottle*, *percent*, and *half*, which may or may not be numerical. Syntactically, there is precisely one slot for C/M, which also favors the interpretation of C/M as a simple  $x$ . In the following section, I will discuss some of the implications of this mathematical account.

## 4. Discussions and Implications

### 4.1 Typological Implications

Out of the 400 languages surveyed and reported in Gil (2011), in 260 numeral classifiers are absent, in 62 they are optional, and in 78, obligatory (see Table 2 below). Consistent with most conventional classifications, Chinese is described as a typical example of languages with obligatory C and English as a prime example for languages without C.

Table 2. Numeral Classifiers in 400 Languages (Gil 2011)

Absent	260
Optional	62
Obligatory	78

However, this typology may be seen in a new light, once we consider Borer's (2005: 94) insight that plurality is not a number specification and plurals are in fact classifiers. Since Borer (2005) makes no C/M distinction, Au Yeung (2005: 265) thus quite appropriately takes the English plural suffix *-s* to be a C, not M. Now, since it applies to all count nouns in the language, I claim that *-s* can be more revealingly and accurately seen as a general classifier, much like the Chinese *ge*, the Korean *kay*, or the Japanese *-tsu*. Let's see a point-by-point comparison below.

- (62) a. Chinese: [Num 3 [c *ge* [N *cup*]]] => [Num 3 [c *ge* [N *cup*]]]  
 b. Japanese: [Num 3 [c *-tsu* [N *cup*]]] => [Num 3-*tsu* [c *-tsu* [N *cup*]]]  
 c. English: [Num 3 [c *-s* [N *cup*]]] => [Num 3 [c *cup-s* [N  $\emptyset$ ]]]

Note that in [Num C/M N], the only feasible C/M for English is the suffix *-s*, a C. Chinese and Japanese, on the other hand, have both C and M. Thus, contrary to common misconceptions, English has C, but no M, not the other way around. All the putative measure words in English behave exactly like common nouns, e.g., *cup* in (62c). Note also that when the numeral has the value of 1, the general C *-s* is not allowed, as shown in (63c). This explains why it is *0.5 apples* and *0 apples* and not *\*0.5 apple* and *\*0 apple*, as *0.5*, *1.0*, and *0* are not straightforwardly 1; *-s* is thus required. This differs from the Chinese *ge* in (63a), which, like all other C's in the language, allows the multiplier 1 to be optional (cf., 3.4 and (56)), but is precisely the same as the Persian general C *ta*, which can appear with any number except 1, as in (63b) (Gebhardt 2009: 212, Hamedani 2011: 139). Likewise, in Amis (Formosan) (Tang 2004: 389), Tat (Southwestern Iranian) and Khasi (Austro-Asiatic) (Greenberg 1990[1972]: 168), only numerals larger than 1 are marked with

classifiers.<sup>9</sup>

- (63) a. Chinese: [Num 3/(1) [C *ge* [N *cup*]]]  
b. Persian: [Num 3/\*1 [C *ta* [N *cup*]]]  
c. English: [Num 3/\*1 [C *-s* [N *cup*]]]<sup>10</sup>

This parallel between C's and plural markers can be extended to explain the use of bare plural nouns as kinds. In (64a), for example, is a generic statement referring to *cats* and *dogs* as kinds. Given that English forbids the number 1 to co-occur with *-s*, the presence of *-s* means that the covert Num can refer to any number except 1, thus allowing the interpretation of Num to be kind (64a) or an unspecified but fixed number larger than 1 (64b). The opposite is true in Mandarin and Cantonese, where a bare [C N] phrase has only a singular reading, as in (65) and (66), respectively.<sup>11</sup>

- (64) a. Cats run faster than dogs.  
b. I saw cats and dogs running around in the house.

- (65) 我 去 台北 看 個 朋友  
*Wo qu taipei kan ge pengyou* (Mandarin)  
I go Taipei see C friend  
'I'm going to Taipei to see a friend.'

---

<sup>9</sup> Another relevant typological variation is that between the numeral 1 and C/M, it is the C/M that is not allowed. Khasi, an Austro-Asiatic language of the Meghalaya state in India, for example, does not allow C with the numeral 1 (Temsen 2007: 6). Another way of viewing this is that 1 requires a silent C.

<sup>10</sup> An anonymous reviewer suggests that it is not the English plural *-s per se*, but rather the English number in general, i.e., singular and plural, that functions like Chinese C. The difference between this view and the one presented in the paper, however, is trivial, as in the latter view the English singular can be seen as having a silent *-s*, e.g., *one -s book*. Refer also to Khasi in fn. 9.

<sup>11</sup> Note also that the bare C phrase in Mandarin is indefinite but the one in Cantonese is definite. See Tang (2011) for an account based on a noun movement parameter.

(66) 隻 狗 想 過 馬路 (Cheng and Sybesma 2005: 9 (24c))

*Zek gau soeng gwo maalou.* (Cantonese)

CL dog want cross road

'The dog wants to cross the road.'

Under the view that plurality marking is a general C, it would be interesting to find out if C's in the conventional sense and plurality marking are indeed in complementary distribution, as first claimed by T'sou (1976) and later most strongly advocated by Borer, emphatic that plural morphology and classifiers do not co-occur (Borer 2005:6, 10, 95), proposes that plurality marking occupies the same syntactic position as C, serving the same function as a divider (Borer 2005:21-22) (cf., 3.2). To what degree is this claim justified is of course an empirical question. Haspelmath (2011) examines 291 languages, 114 of which are included in Gil's (2011) 400, and find that the majority, over 90%, employs nominal plurality (see Table 3).

**Table 3. Nominal Plurality in 291 Languages (Haspelmath 2011)**

Absent (28)	Human nouns only, optional	Human nouns only, obligatory	All nouns, always optional	All nouns, optional in inanimates	All nouns, always obligatory
28	20	40	55	15	133

However, Gil (2011) and Haspelmath (2011), though each a chapter in *The World Atlas of Language Structures Online*, make no reference to each other and contain no discussion on whether or to what degree C and nominal plurality are in complementary distribution. Table 4 shows the detailed search results of the two combined features: numeral classifiers and nominal plurality, among the 114 languages covered by both. Table 5 shows a simplified version where only the presence and absence of each feature is considered.

**Table 4. Numeral Classifiers and Nominal Plurality in 114 Languages**

		Occurrence of Nominal Plurality (114/291)					
		Absent (28)	Human nouns only, optional (20)	Human nouns only, obligatory (40)	All nouns, always optional (55)	All nouns, optional in inanimates (15)	All nouns, always obligatory (133)
Numeral Classifiers (114/400)	Absent (260)	8	8	12	12	8	40
	Optional (62)	2	1	0	5	0	3
	Obligatory (78)	2	2	2	6	1	2

**Table 5. Numeral Classifiers and Nominal Plurality in 114 Languages (Simplified)**

		Nominal Plurality (114/291)	
		Absent (28)	Present (263)
Numeral Classifiers (114/400)	Absent (260)	8	80
	Present (140)	4	<b>22</b>

There are 8 languages that have neither C nor plurality and thus do not come into play. Complementary distribution between C and plurality is verified in 84 languages; among them, surprisingly, only 4 are classifier languages: Abun and Kana with obligatory C and Maybrat and Tidore with optional C, while 80 are with plurality only. 22 languages are found to have both C and plurality, where Mandarin Chinese is listed as with obligatory C and optional plural marking on human nouns.

Note that these 22 languages do not necessarily falsify T'sou's (1976) and Borer's (2005) generalization, for C and plurality within such languages may still be in complementary distribution. Data from Chinese are unfavorable to this generalization, given the fact that, again contrary to common misconceptions, in [Num C N], a human N can indeed be marked by plural suffix *-men*. In one of the most prominent works on plurality and the suffix *-men*, Li (1999: 77) makes this remark:



Unlike those languages with a true plural morpheme, the occurrence of a quantity [number+classifier] expression is not compatible with the occurrence of *-men*.

- (2) \**san-ge xuesheng-men*  
three-CI student-MEN  
'three student+*men*'

Counterexamples, nonetheless, abound. Hsieh (2008: 8) shows several examples from the Sinica Corpus. In a Google search within the Taiwan (.tw) domain, there are 51 exact matches for (67a). Dozens of more matches are found with *ji* 'several' replaced with an exact number, e.g., *liang* '2', *san* '3', *si* '4', etc. And even (67b), which is nearly identical to Li's (1999: 77) own example in the above quote, has 8 exact matches and thus serves as a direct counterexample.

- (67) a. 幾 位 老師們  
*ji wei laoshi-men*  
several C teacher-PL  
'several teachers'

- b. 幾 個 學生們  
*ji ge xuesheng-men*  
several C student-PL  
'several students'

Traditionally, however, *-men* is treated as a collective marker when attached to nouns (e.g., Chao 1968, Norman 1988, Ijic's 1994, 1998, among others). For example, *N-men* would refer to a group of people anchored by *N*. Nonetheless, Li (1999) demonstrates convincingly in spite of her judgment that [*Num C N-men*] is ill-formed, that *-men* shows some plurality properties and there are some facts that run counter to *-men*'s being a collective marker. In Taiwan Mandarin, in fact, *N-men*, where *N* cannot be a proper noun, does not have the collective reading, but [*N tamen*] 'N they' does. Japanese, likewise,

according to Ueda and Haraguchi (2008), has C and plural markers not in complementary distribution.

One important difference between *-men* and *-tati* is that the Japanese counterpart of (7) is grammatical, as shown in (11).

- (11) san-nin-no      gakusei-tati  
      three-CL-NO    student-TATI  
      ‘three students’

Thus, it suffices to say that Chinese and Japanese, among others, pose a challenge to the claim that C and plurality do not co-occur. Gebhardt (2009), for example, contra Borer (2005), proposes a feature-based syntactic analysis that permits plurality marking and C to co-occur and claims that Persian is such a language. However, Gebhardt’s (2009) analysis would imply that the co-occurrence of plurality and C could in principle be as common as the mutual exclusiveness of the two features. The numbers in Table 4 and 5 show that this implication is not borne out. I thus contend that this claim by T’sou and Borer is in principle on the right track. Languages that violate this generalization are indeed uncommon, and language change may provide an explanation. Massam (2009), for example, while confirming Borer’s (2005) claim on the mutual exclusiveness of C and plurality in Niuean, suggests that the language, with neither C nor number being canonical, is undergoing a change from a classifier language to one with a full-fledged number system. The Chinese plural *-men* can likewise be seen as a residue in the language’s change from a synthetic language to an isolating, analytic language. This is precisely what Li and Shi (2000) have discovered in their study of the historical development of *-men* and C’s in early Chinese. They conclude that between the 10<sup>th</sup> and 15<sup>th</sup> century *-men* gradually acquired the properties of a full-fledged plurality maker; however, after the 15<sup>th</sup> century it regressed and lost many of these properties due to the rise and establishment of the classifier system.

Therefore, a weaker claim may prove to be more revealing, i.e., that the more developed a classifier system is in a language, the weaker its plurality marking system, and vice versa. Chinese, for example, now has a highly developed and thus strong classifier system and a very weak plurality system,

one that marks only human nouns, and only optionally. English, in contrast, has a strong plurality system and no conventional C at all. Tang (2004), in a similar spirit, claims that classifier languages come in two types: *rich-classifier languages*, e.g., Chinese, where C's are generally required with numerals and do not seem to freely co-occur with the plurality marker, and *poor-classifier languages*, e.g., several Formosan languages: Paiwan, Bunun, Kavalan, Tsou, and Amis, where C's tend to be optional with numerals and do co-occur with plural nouns. It will thus be a meaningful typological study to closely examine the 22 languages in Table 4 and 5, which have both C's and plurality.

Another typological issue is whether a language has C/M, again C here includes plurality. Logically, there are four possibilities: (1) C only, (2) M only, (3) both C and M, and (4) neither C nor M. English can be seen as type (1), Chinese, type (3). Canglo Monpa, a Tibeto-Burman language in Tibet, reported to have a [N M Num] construction only and thus no C's (Jiang 2006: 50), is of type (3), and possibly also Bulgarian (Cinque and Krapova 2007) and Khasi (Temsen 2007), with merely two C's. The 8 languages in Table 4 and 5 that are without plurality nor C/M belong to type (4), and so does Archaic Chinese (e.g., Norman 1988: 120).

The final typological issue is C's grammatical status: free morpheme or bound morpheme, again under the view that C includes plurality. For C as a free morpheme, there are in turn two possibilities: words, e.g., Cantonese (e.g., Au Yeung 2005), or clitics, e.g., Mandarin (e.g., Yang 2001). For C as bound morphemes, there are also two possibilities in term of the stem: Num, e.g. Japanese *-tsu*, or N, e.g., English *-s*.

#### 4.2 C/M and Number Systems

Taking C/M as multiplicands suggests that there is a strong connection between the employment of C/M in a language and its number systems. According to Comrie (2006, 2008), there are around 20 languages whose number systems are with little or no internal structure or with addition only. For example, there are no numerals in Pirahã, only 1 to 3 in Mangarayi, 1 to 5 in Yidiny, and 1-5 and then 10 in Hixkaryana, and 1, 2, 1+2, and 2+2 in Haruai.

Also, a small number of languages of Highland New Guinea languages employ extended body-part systems, which again are rather limited in the range of numbers that they can express. A logical and falsifiable prediction that can be derived from C/M as multiplicands is that, in languages with such restricted number systems where multiplication is not used, there will be no C/M.

On the other hand, the majority of number systems employed in languages follows the general pattern in (68).

(68) General Pattern of Number Systems in Languages (Comrie 2006)

For base  $b$ :  $(n \times b) + m$  (where  $m < b$ )

Thus, another logical and likewise easily falsifiable prediction is that if a C/M system is justified in a language, then it must also have a multiplication-based number system. The opposite is not necessarily true, however, that if a language's number system involves multiplication, then it must have C/M.

Based on a cross reference of Comrie's (2008) database on number systems and Gil's (2011) database on classifiers, both predications are verified. As shown in Table 6, no C/M is found in any of 12 languages with a restricted number system, and without exception, all 30 languages with C/M also employ a multiplication-based number system.

**Table 6. Correlation between Numeral Bases and Numeral Classifiers in 106 Languages**

	C/M	ABSENT	PRESENT
<i>Number System</i>			
<i>Multiplication-based</i>		64	30
<i>Restricted</i>		12	0

4.3 Count/mass Distinction and the Misconceived Necessity of C

Viewing C, including both the conventional C and the plural makers, as a multiplicand with the simplest value, 1, with the cognitive function to profile an inherent feature of the frame or base provided by N requires that the N refer to

a concrete or abstract discrete entity, i.e., one with clearly definable boundaries conceptually; in other words, C requires a count noun. C does not take a mass noun and turn it into count. Consequently, all languages that employ C, again including plural marking, distinguish count and mass, at the lexical level. That certainly would cover the majority of languages in the world. The small number of languages, e.g., the 8 languages in Table 4 and 5, that have neither C nor overt inflectional plural marking, do use other means to express plurality. Archaic Chinese (1000B.C.-200A.D.), for example, is well-known for its lack of classifiers and plurality makers and allows count nouns to co-occur with numerals directly. The obvious implication is thus that the count/mass distinction is lexical and universal.

This position, which is akin to Cheng and Sybesma's (1998) claim that C's select count and M's select either count or mass, runs counter to the widespread view that putative count nouns in classifier languages like Chinese may be understood as either mass or count (Gil 1987) or that all nouns in such languages are mass (e.g., Allan 1977: 293, Krifka 1995, Chierchia 1998, Tai 2003: 312, Borer 2005). Similarly, Landman (2004) considers bare nouns as sums or mass. As Gil (2011) aptly points out, this generalization has led to a widespread belief that C is necessary in such languages in order to individuate the noun, which is mass or sum, and thus provide the necessary units for counting. However, this widespread view is based on the misconception that C in Chinese is required in [Num C N].<sup>12</sup>

Thus, according to this view, the Mandarin \*sān píngguǒ 'three apple(s)' is semantically ill-formed for the same reason that the English \*three water(s) is: just as English water requires an explicit mensural classifier before it can be quantified, as in three glasses/ounces/drops of water, so Mandarin píngguǒ requires a sortal classifier before it can be successfully enumerated, as in sān gè píngguǒ. (Gil 2011)

---

<sup>12</sup> A related issue in terms of typology is thus this: how many languages that have been considered to have obligatory C in [Num C N] like Chinese in fact do allow their C to be optional?

Recall that examples in (54) and (55), in 3.4, clearly show that in certain literary styles in Chinese, C, as *x1*, need not appear. Let's see one more example from Mandarin. The idiomatic expression in (69) refers to a trade that looks good superficially but in substance disfavors the trader. A Google search in the *.tw* domain has turned up 3,670 matches.

(69) 五 馬 換 六 羊  
*wu ma huan liu yang*  
5 horse trade 6 goat  
'Trading 5 horses for 6 goats' (an apparent good trade)

In fact, there are as many as 62 languages with optional C's in Gil's (2011) survey (see Table 2). Also, in Table 5, out of the 114 languages examined for both plurality (Haspelmath 2011) and C (Gil 2011), there are 8 languages with neither. Pidgins and creoles are also known to have neither. Greenberg (1990[1972]: 168), thus goes as far as saying "...it is not excessive to state that there are no numeral classifier languages".

Recall also that Num as the multiplier can also be optional when Num = 1, as in (70), which has 128 exact Google matches. Clearly, the interpretation of the nouns in (69) can only be count and plural, and in (70) count and singular.

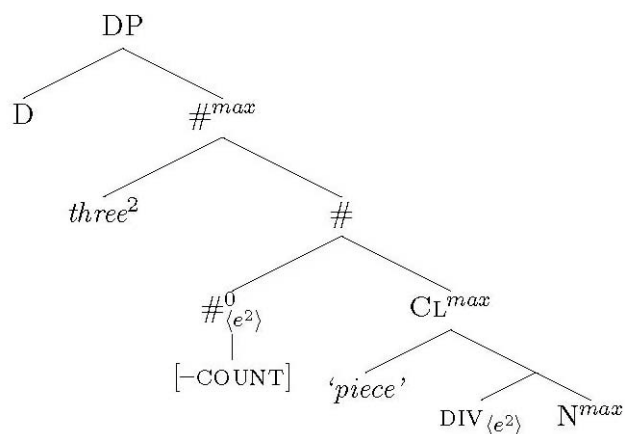
(70) 買 本/箱 書  
*mai ben/xiang shu*  
buy C/M-box book  
'buy a book/a box of books'

Unlike English mass nouns *honesty* and *kindness*, the Chinese *ma* 'horse', *yang* 'goat', and *shu* 'book' do come as inherently countable units and thus lexically can only refer to a single level of atomic singular individuals. Li, Barner, and Huang (2008), in their empirical studies, indeed find support in adult Mandarin speakers for Cheng and Sybesma's (1998) count/mass distinction, which indirectly supports my position that C is not required.

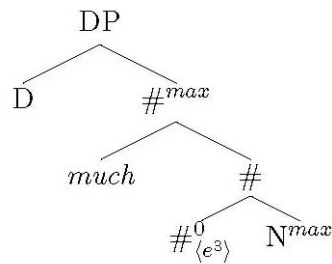
However, the most serious problem for the hypothesis that count/mass distinction exists in languages like English which mark plurality but does not exist in classifier languages like Chinese is that it allows no gradient, no middle ground. That is, given any language at any point in time, it either has this distinction or it does not. Such a view cannot possibly explain how languages change typologically between analyticity and syntheticity, which is always a *gradual* process, never *abrupt*. This means that the count/mass distinction must be universal.

The remaining issue is whether this distinction is universally lexical or syntactic. This brings us to Borer's (2005) theory that nouns in all languages are lexically mass by default and countability is created in the syntax by the projection of a functional structure above the noun. (71) is a lucid adaptation by Mitrovic (2011) of the different structures for count and mass. The Chinese C and the English plural *-s* function as a mass *divider*, i.e., the CL 'piece', in the structure of (71a) and affords the noun a count interpretation, while the absence of C or *-s* in the DP of (71b) means the noun's default mass interpretation must remain.

(71) a. Count Noun



b. Mass Noun



Even though Borer rules out (incorrectly, as pointed out in 4.3) the co-occurrence of C and plurality, because they occupy the exact same syntactic position in (71a), her theory does not rule out the possibility of C's and plurality marking co-existing within the same language. Thus, languages' increasing analyticity (thus more use of C's) or syntheticity (thus more use of inflectional plurality marking) can still be explained.

It is without controversy that the count/mass distinction must be flexible as it is well-recognized that mass nouns can turn into count and vice versa. As pointed out in 3.2, Borer's (2005) theory incorrectly predicts that all nouns can appear in either (71a) or (71b) and receive the count or mass interpretation accordingly. It allows the maximal flexibility between count and mass and overgenerates as a consequence. It is well-known that each C in Chinese selects not only a count noun but also a count noun with a specific semantic feature in terms of human, animacy, shape, etc. The fact that only nouns denoting vehicles, e.g., those in (72a), can be profiled by *liang* means these nouns must be marked as [count] as well as [vehicle] *lexically*. In the case of *fute* 'Ford', metonymy is involved and the term still must be interpreted as [vehicle] and not [company]. In (72b), the C *liang* cannot coerce the incompatible nouns into [count] and [vehicle]. Metonymy is futile: Acer does not make cars, diesel cannot be interpreted as vehicles that run on diesel (even though *he drives a diesel* is perfect in English), and the plastic material does not give rise to the reading of a plastic toy car.

- (72) a. 一 輛 福特/巴士/玩具車  
*yi liang fute/bashi/wanjuche*  
 one C Ford/bus/toy car  
 'one Ford/bus/toy car'



b.*一	輛	宏碁/柴油/塑膠
<i>yi</i>	<i>liang</i>	<i>honggi/chaiyou/sujiao</i>
one	C	Acer/diesel/plastic

Borer's theory allows no nouns in any language to be count lexically. However, it can be easily demonstrated that at least in strong classifier languages, or rich-classifier languages in Tang's (2004) term, it is necessary for at least some nouns to be lexically marked [count], among other features, to ensure the grammatical compatibility between the C and the nouns it selects. Thus, the thesis that there are no count nouns in the lexicon of any language cannot be justified. Note that this does not in any way negate the possibility of [count] also being assigned syntactically.

Finally, a case study by Semenza *et al* (1997) of an Italian-speaking patient with brain damage may prove to be rather significant to this debate. The study shows that the patient was able to retrieve words of the category mass and yet her grammatical performance regarding the properties of mass and non-countable nouns, in both production and reception, is severely and selectively damaged, her use of grammar otherwise perfect. This shows how a set of specific grammatical rules, stored at the lemma level of lexical retrieval, is independently represented and accessible. And here is the authors' conclusion that is directly relevant to our debate here:

This conclusion leads to another point of theoretical interest. As has been already mentioned in many languages, like, for instance, Chinese and Japanese, all nouns are mass nouns. ...However a loss in aphasia of the rules concerning mass nouns vis-a-vis a preservation of the rules concerning count nouns may be taken as evidence that, at least in Italian, *unmarkedness is borne upon count nouns...* (Semenza *et al* 1997: 674, *emphasis added*)

#### 4.4 Implications for Acquisition Studies

Viewing C's as a multiplicand 1 and a cognitive profiler also has significant implications for studies on children's acquisition of C's (again, including

plurality markers) and numbers. First, in classifier languages, all C's being equally a multiplicand *1*, the acquisition of the general C surely comes before that of specific C's. This is confirmed by many studies, the largest of which is Tse *et al* (2007), a study of Cantonese-speaking children between 3-5 years of age. The extensive overuse of general C is also predicted and confirmed (e.g., Japanese: Sanches 1977, Matsumoto 1985, Naka 1999; Mandarin: Erbaugh 1986, Liu 2008; Hokkien: Ng 1989, Cantonese: Tse *et al* 2007). Besides the usual linguistic, cognitive, and contextual reasons (cf., Tse *et al* 2007: 514 for a summary), the fact that the general C is much like a universal plurality maker, e.g., English *-s*, is also likely to lead the child to misconceive it as a plurality marker, hence the extensive overuse.

Also, C is mathematically simpler than M, which can be of any value besides *1*. M is thus much more complex mathematically. C, therefore, should come before M in acquisition. However, once C and M are both established, the number of M's, an open class, should be larger than that of C, a closed class. Again, this is exactly what Tse *et al* (2007: 512) have discovered: types of M used by Cantonese-speaking children significantly outnumber those of C, just like adult language, but the top ten most frequent C/M are nearly all C, except one M.

Recall that in Mandarin [Num C/M N], if Num is *1*, then it can be omitted in the right syntactic contexts, for example when preceded by a demonstrative *zhe* 'this' or *na* 'that'. This leaves C/M the multiplicand without an overt multiplier and is thus more complex cognitively than a straightforward overt multiplier *1*. Thus, it can be expected that the acquisition of this well-formed omission should come later than that of the C/M with numbers. I am not aware of any research on this specific issue and will thus leave it for further research.

Our analysis also predicts that the acquisition of the count/mass distinction embedded in English plurality marking *-s* is much easier than the acquisition of Mandarin specific C's, each of which, besides being a multiplicand *1* just like *-s*, profiles a certain cognitive aspect and thus selects a range of nouns that are not always homogenous and can even be arbitrary. This prediction is vindicated by Li, Barner, and Huang's (2008) study, where

they find Mandarin-speaking children's mastery of count classifiers comes much later than English-speaking children's control of count syntax.

However, given the fact that a general classifier is much like a plurality marker, it can also be predicted that the acquisition of C and its overgeneralization by children comes before that of the specific C's. If this line of thinking is correct, then the acquisition and overgeneralization of English plurality inflection and that of Chinese general C should be around the same time. This is confirmed by Meyers and Tsay (2000) study of pre-school Taiwanese-speaking children: the onset of the overgeneralization of the Taiwanese general C and English inflection both occur between 2-3 years of age, which is among their several findings on the parallels between the early acquisition of Taiwanese C's and English inflection.

Given the underlying mathematical structure of  $(n \times base) + m$  in a language's number system, it is also logical to assume that the acquisition of larger numbers involving multiplication, e.g., *san-shi san* or *thirty-three*  $((3 \times 10) + 3)$ , should come after the mastery of either plurality marker, e.g., English *-s*, or the general C, e.g., Mandarin *ge*, depending on the typology of the language. Again, as no existing research on this specific issue is found, confirmation of this prediction is left for further research.

## 5. Concluding Remarks

Classifiers have been seen as one of the most salient features of Chinese and other similar languages, in sharp contrast with non-classifier languages. This paper has explored the universality of classifiers in all human languages and claimed that the plurality marker in non-classifier languages, e.g., English nominal suffix *-s*, may be seen as a general C, much like the Mandarin *ge*, except the former is even more general. Specifically, this unification is based not on syntax but on semantic and mathematical grounds. Classifiers, unlike measure words, only serve to *profile* an essential or inherent feature of the head noun in [Num C/M N] and thus contribute no additional meaning to the head noun. Based on the different semantics and mathematics between

classifiers and measure words, this paper has also provided concrete grammatical tests for the C/M distinction in Chinese.

Implications of these research findings have also been discussed for the following areas of further research: typology of classifiers and classifier languages, correlations between number systems and the employment of classifiers and measures, the universal count/mass distinction at the lexical level, and first language acquisition of classifiers and numbers.

### **Acknowledgements**

I give my heartfelt thanks to the four anonymous *Lingua* reviewers for their professionalism and generosity. Their insightful and constructive comments have played a significant role in improving the paper from its original form. Different parts of the research reported in the paper have been presented in Chinese Lexical Semantics Workshop 2011 (CLSW2011), The 16th International Lexical Functional Grammar Conference (LFG 2011), The Third Symposium on Linguistics Research Methods, and The International Conference on Language Evolution: Origin and Change of Language in Descriptive and Formal Linguistic Theories; I thank the audiences for their questions and comments. Research reported in the paper was partly funded by two NSC grants, 99-2410-H-004-190-MY2 and 101-2410-H-004-184-MY3. However, I am solely responsible for the content of the paper.

### **References**

- Adams, K., Conklin, N. F. 1973. Toward a theory of natural classification. Paper presented at the 9th Regional Meeting of the Chicago Linguistic Society, 1-10.
- Ahrens, K. 1994. Classifier production in normals and aphasics. *Journal of Chinese Linguistics* 22(2), 203-248.

- Anderson, S. W., Damasio, A. R., Damasio, H. 1990. Troubled letters but not numbers. *Brain*, 113, 749-766.
- Allan, K. 1977. Classifier. *Language* 53, 285-311.
- Au Yeung, W. H. B. 2005. An interface program for parameterization of classifiers in Chinese. PhD. Dissertation, Hong Kong University of Science and Technology.
- Au Yeung, W. H. B. 2007. Multiplication basis of emergence of classifier. *Language and Linguistics* 8(4), 835-861.
- Borer, H. 2005. Structuring Sense, vol. 1: In Name Only. Oxford: Oxford University Press.
- Butterworth, B., Cappelletti, M., Kopelman, M. 2001. Category specificity in reading and writing: the case of number words. *Nature Neuroscience* 4, 784-785.
- Chao, Y.-R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Carpenter, K. 1991. Later rather than sooner: extralinguistic categories in the acquisition of Thai classifiers. *Journal of Child Language* 18, 93-113.
- Cheng, L. L.-S., Sybesma, R. 1998. Yi-wan Tang, Yi-ge Tang: Classified and Massifiers. *Tsing Hua Journal of Chinese Studies, New Series* 28(3), 385-412.
- Cheng, L. L.-S., Sybesma, R. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4), 509-542.
- Cheng, L. L.-S., Sybesma, R. 2005. Classifiers in four varieties of Chinese. In: Cinque, G., Kayne, R. S. (Eds), *The Oxford Handbook of Comparative Syntax*. Oxford University Press, Oxford, pp. 259-292.
- Chierchia, G. 1998. Plurality of mass nouns and the notion of semantic parameter. In: Rothstein, S. (Ed.), *Events and Grammar*. Kluwer, Dordrecht, pp. 53-103.
- Chu, F. 1994. Keji hanyu zhong de shuming jiegou (The Num-N construction in Chinese for technology). *Hanyu Xuexi (Learning Chinese)* 1994(3), 40-41.
- Cinque, G., Krapova, I. 2007. A note on Bulgarian numeral classifiers. In: Alboiu, G., Avram, A. A., Avram, L., Isac, D. (Eds.), *Pitar Moș: A Building*

- with a View. Papers in Honour of Alexandra Cornilescu. Editura Universităţii din Bucureşti, Bucharest, pp. 45-51.
- Cipolloti, L., Butterworth, B., Denes, G. 1991. A specific deficit for numbers in a case of dense acalculia. *Brain* 114, 619-637.
- Comrie, B. 2006. Numbers, language, and culture. Paper presented at the Jyvaskyla Summer School, Jyvaskyla, Finland.
- Comrie, B. 2008. Numeral bases. In: Haspelmath, M., Dryer, M. S., Gil, D., Comrie, B. (Eds.), *The World Atlas of Language Structures Online*. Max Planck Digital Library, Munich, chapter 131. Available online at <http://wals.info/feature/131>. Accessed on 2011/4/23.
- Croft, W. 1994. Semantic universals in classifier systems. *Word* 45, 145-171.
- Domahs, F., Bartha, L., Lochy, A., Benke, T., Delazer M. 2006. Number words are special: Evidence from a case of primary progressive aphasia. *Journal of Neurolinguistics*, 19(1), 1-37.
- Erbaugh, M. S. 1986. Taking stock: The development of Chinese noun classifiers historically and in young children. In: Craig, C. (Ed.), *Noun Classes and Categorization*. John Benjamins Publishing Company, Amsterdam, pp. 399-436.
- Fillmore, C. 1982. Frame semantics, in *Linguistic Society of Korea. Linguistics in the Morning Calm*, ed. by Linguistic Society of Korea. Seoul: Hanshin Publishing, pp.111-137.
- Gebhardt, L. 2009. Numeral classifiers and the structure of DP. Ph.D. Dissertation, Northwestern University.
- Genhardt, L. 2011. Classifiers are functional. *Linguistic Inquiry* 42, 125-130.
- Gil, D. 1987. Definiteness, noun-phrase configurationality, and the Count-Mass distinction. In: Reuland, E. J., ter Meulen, A. G. B. (Eds), *The Representation of (In)definiteness*. MIT Press, pp. 254-269.
- Gil, D. 2011. Numeral classifiers. In: Haspelmath, M., Dryer, M. S., Gil, D., Comrie, B. (Eds), *The World Atlas of Language Structures Online*. Max Planck Digital Library, Munich, chapter 55. Available online at <http://wals.info/feature/55>. Accessed on 2011/12/13.
- Greenberg, J. 1974. Studies in numerical system: Double numeral system. *Working Papers on Language Universals* 14, 75-89.

- Greenberg, J. 1990[1972]. Numerical classifiers and substantival number: problems in the genesis of a linguistic type. In: Denning, K., Kemmer, S. (Eds.), *On Language: Selected Writings of Joseph H. Greenberg*. Stanford University Press, Stanford, CA, pp. 166-193. [First published 1972 in *Working Papers on Language Universals* 9, 1-39. Department of Linguistics, Stanford University, Stanford, CA.]
- Hamedani, L. 2011 *The function of number in Persian*. Ph.D. Dissertation, University of Ottawa.
- Haspelmath, M. 2011. Occurrence of nominal plurality. In: Haspelmath, M., Dryer, M. S., Gil, D., Comrie, B. (Eds), *The World Atlas of Language Structures Online*. Max Planck Digital Library, Munich, chapter 34. Available online at <http://wals.info/feature/34>. Accessed on 2011/4/23.
- Her, O.-S. 2010. Distinguishing classifiers and measure words. Paper presented at The 4th Conference on Language, Discourse and Cognition (CLDC2010), National Taiwan University, Taipei, May 1-2, 2010.
- Her, O.-S. 2011. Structure of classifiers and measure words: A lexical functional account. Paper presented at the 2011 Annual Conference of the International Lexical-Functional Grammar Association, University of Hong Kong, July 16-19, 2011.
- Her, O.-S., Hsieh, C.-T. 2010. On the semantic distinction between classifiers and measure words in Chinese. *Language and linguistics* 11(3), 527-551.
- Her, O.-S., Lai, W.-J. 2011. Classifiers: The many ways to profile 'one'. Paper presented at the 12th Chinese Lexical Semantics Workshop, National Taiwan University, Taipei, Taiwan, May 3-5, 2011.
- Hsiao, J. C. 2007. *The acquisition of classifiers in Taiwan Southern-Min*. MA Thesis, National Chung Cheng University.
- Hsieh, C.-T. 2009. *A frame-based approach to classifiers: A case study of Taiwan Mandarin*. National Chengchi University MA Thesis, Taipei.
- Hsieh, M.-L. 2008. *The internal structure of noun phrases in Chinese*. *Taiwan Journal of Linguistics: Book Series in Chinese Linguistics*. No. 2. Crane Publishing, Taipei.
- Hu, Q. 1993. *The acquisition of Chinese classifiers by young Mandarin-speaking children*. Boston University Dissertation, Boston.

- Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. MIT PhD Dissertation.
- Huang, C.-T. J., Li, A. , Li. Y. 2009. The Syntax of Chinese. Cambridge University Press.
- Huang, C.-R., Ahrens, K. 2003. Individuals, kinds and events: classifier coercion of nouns. *Language Sciences* 25, 353-373.
- Hung, F.-S. 1996. Prosody and the Acquisition of Grammatical Morphemes in Chinese Languages. Indiana University Linguistics Club.
- Iljic, R. 1994. Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32, 91-116.
- Iljic, R. 1998. Number and person. Paper presented at IACL-7/NACCL-10, Stanford University.
- Jiang, Y. 2006. Hanzang Yuxi Mingliangci Yanjiu 'A study of nominal classifiers in Sino-Tibetan'. PhD dissertation, The Central University of Nationalities.
- Krifka, M. 1995. Common nouns: a contrastive analysis of Chinese and English. In: Carlson, G., Pelletier, F. (Eds), *The Generic Book*. University of Chicago Press, pp. 398-411.
- Lai, W.-J. 2011. Identifying true classifiers in Taiwan Mandarin. MA Thesis. National Chengchi University, Taipei.
- Landman, F. 2004. *Indefinites and the Type of Sets*. Blackwell, Malden.
- Langacker, R. W. 1987. *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*. Stanford, California: Stanford University Press.
- Li, C., Thompson, S. 1981. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press, Berkeley.
- Liang, Y.-C. 2006. Nominal phrases in English and Japanese speakers' L2 Mandarin grammars. Cambridge University Dissertation, Cambridge.
- Li, P., Barner, D., Huang, B. H. 2008. Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese. *Language, Learning, and Development* 4(4), 1-42.
- Li, W. 2000. The pragmatic function of numeral-*classifiers* in Mandarin Chinese. *Journal of Pragmatics* 32(8), 1113-1133.
- Li, Y.-H. A. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8(1), 75-99.



- Li, Y.-H. A., Shi, Y. 2000. Hanyu lingci xitong-de jianli yu fushu biaoji “men”-de fazhan (The establishment of the Chinese classifier system and the development of the plurality marker -men). *Contemporary Linguistics* 2(1), 27-36.
- Li, X.-P. 2011. On the semantics of classifiers in Chinese. Ph.D. Dissertation, Bar-Ilan University. Available online at: <http://www.semanticsarchive.net/Archive/mY3YWYzO/>. Accessed on 2011/09/018.
- Lin, J.-W. 1997. Chinese noun phrase structure: DP or NP? *Chinese Languages and Linguistics* 3, 401-434.
- Liu, H. 2008. A case study of the acquisition of Mandarin classifiers. *Language Research* 44(2), 345-360.
- Ma, A. 2011. Hanyu jiaoxue zhong ying zhuyi mingliangci de fenlei wenti. (Nominal classifiers should be properly classified for pedagogical purposes.) *Xiandai Yuwen (Modern Chinese)* 2011.3.
- Massam, D. 2009. On the separation and relatedness of classifiers, number, and individuation in Niuean. *Language and Linguistics* 10(4), 669-699.
- Matsumoto, Y. 1985. Acquisition of some Japanese numeral classifiers: the search for convention, *Stanford Papers and Reports in Child Language Development* 24, 79-86.
- Messina, G., Denes, G., Basso, A. 2009. Words and number words transcoding: A retrospective study on 57 aphasic subjects. *Journal of Neurolinguistics* 22, 486-494
- Mitrovic, M. 2011. On Count/Mass Distinction in Slovene. In: Greenberg, M. L., Snoj, M. (Eds.), *Slovenski jezik/Slovene Linguistic Studies*, vol. 8. Ljubljana & Lawrence: Slovenian Academy of Arts and Sciences & Hall Center for Humanities publication, pp. 115-167
- Myers, J. 2000. Rules vs. analogy in Mandarin classifier selection. *Language and Linguistics* 1(2), 187-209.
- Myers, J., Tsay, J. 2000. The acquisition of the default classifier in Taiwanese. In: *Proceedings of The Seventh International Symposium on Chinese Languages and Linguistics*, Chia-Yi: National Chung Cheng University, pp. 87-106.

- Naka, M. 1999. The acquisition of Japanese numerical classifications by 2–4-year-old children: The role of caretakers' linguistic inputs. *Japanese Psychological Research* 41(1), 70-78.
- Ng, B. C. 1989. The acquisition of numeral classifiers in Hokkien, a Southern Min language. PhD Dissertation, La Trobe University.
- Norman, J. 1988. *Chinese*. Cambridge University Press, Cambridge.
- Paris, M.-C. 1981. *Problèmes de Syntaxe et de Sémantique en Linguistique Chinoise*. Collège de France, Institut des Hautes Etudes Chinoises, Paris.
- Rey, G. 2003. The analytic/synthetic distinction. *The Stanford Encyclopedia of Philosophy*, ed. by Edward Zalta. Accessed online, December 1, 2009. <http://plato.stanford.edu/entries/analytic-synthetic/>
- Robertson, T. 2008. Essential vs. accidental properties. *The Stanford Encyclopedia of Philosophy*, ed. by Edward Zalta. Accessed online, December 1, 2009. Available online at: <http://plato.stanford.edu/entries/essential-accidental>
- Sanches, M. 1977. Language acquisition and language change: Japanese numeral classifiers. In: Blount, B. G., Sanches, M. (Eds.), *Sociocultural Dimensions of Language Change*. Academic Press, New York, pp. 51-62.
- Semenza, C., Mondini, S, Cappelletti, M. 1997. The grammatical properties of mass nouns: An aphasia case study. *Neuropsychologia* 35(5), 669- 675.
- Tai, H-Y. J. 2003. Cognitive relativism: resultative construction in Chinese. *Language and Linguistics* 4(2), 301-316.
- Tai, H-Y. J., Wang, L. 1990. A semantic study of the classifier Tiao. *Journal of the Chinese Language Teachers Association* 25(1), 35-56.
- Tang, C.-C. J. 1990. A note on the DP analysis of Chinese noun phrases. *Linguistics* 28, 337-354.
- Tang, C.-C. J. 2004. Two types of classifier languages: A typological study of classification markers in Paiwan noun phrases. *Language and Linguistics* 5(2), 377-407
- Tang, C.-C. J. 2005. Nouns or classifiers: A non-movement analysis of classifiers in Chinese. *Language and Linguistics* 6(3), 431-472.
- Tang, S.-W. 2011. A parametric approach to NP ellipsis in Mandarin and Cantonese. *Journal of East Asian Linguistics* 20(2), 107-115.

- Temsen, G. 2007. Characteristics of Khasi numeral classifiers. In: Proceedings of the National Seminar on Perspectives in Linguistics, Department of Linguistics, University of Kashmir, Srinagar, India. 1-15. Available online at:  
[http://uohyd.academia.edu/GraceTemsen/Papers/955834/Characteristics\\_of\\_Khasi\\_Numeral\\_Classifiers](http://uohyd.academia.edu/GraceTemsen/Papers/955834/Characteristics_of_Khasi_Numeral_Classifiers)
- Thioux, M., Pillon, A., Samson, D., De Partz, M.-P., Noel, M. P., Seron, X. 1998. The isolation of numerals at the semantic level. *Neurocase* 4, 371-389.
- Tien, Y.-M., Tzeng, O. J. L., Hung, D. L. 2002. Semantic and cognitive basis of Chinese Classifiers: A functional approach. *Language and Linguistics* 3(1), 101-132.
- Tsai, S.-H. 2009. Processing of Chinese classifier-noun agreement: An event-related potential study. MA Thesis, National Cheng Kung University.
- Tse, Shek Kam, Li, Hui, & Leung, Shing On. 2007. The acquisition of Cantonese classifiers by preschool children in Hong Kong. *Journal of Child Language* 34, 495-517.
- T'sou, B. K. 1976. The structure of nominal classifier systems. In: Jenner, P. N., Starosta, S., Thompson, L. C. (Eds.), *Austroasiatic Studies*, vol. 2. University Press of Hawaii, Honolulu, pp. 1215-1247.
- Ueda, Y. Haraguchi, T. 2008. Plurality in Japanese and Chinese. *Nanzan Linguistics: Special Issue* 3, vol. 2, 229-242.
- Wang, H. 2004. Liangci de fenlei yu duiwai hanyu liangci jiaoxue. (Classification of classifiers and teaching classifiers in Chinese as a second language.) *Jinan Xuebao (Journal of Jinan University, Humanities and Social Sciences Edition)* 2004(2), 113-116.
- Watanabe, A. 2006. Functional projections of nominals in Japanese: syntax of classifiers. *Natural Languages and Linguistic Theory* 24, 241-306.
- Wu, Y. Bodomo, A. 2009. Classifiers ≠ determiners. *Linguistic Inquiry* 40, 487-503.
- Yang, R. 2001. Common nouns, classifiers, and quantification in Chinese. Rutgers, State University of New Jersey, Dissertation.

- Yi, B. U. 2009. Chinese classifiers and count nouns. *Journal of Cognitive Science* 10, 209-225.
- Yi, B. U. 2011. What is a numeral classifier? *Philosophical Analysis* 23, 195-258.
- Zhang, H. 2007. Numeral classifiers in Mandarin Chinese. *Journal of East Asian Linguistics* 16(1), 43-59.
- Zhang, N. N. 2011. The constituency of classifier constructions in Mandarin Chinese. *Taiwan Journal of Linguistics* 9(1), 1-50.