

International Journal of Computer Processing of Oriental Languages
© Chinese Language Computer Society &
World Scientific Publishing Company

Classifiers: The Many Ways to Profile ‘one’ A Case Study of Taiwan Mandarin

ONE-SOON HER

*Graduate Institute of Linguistics & Research Center of Mind, Brain, and Learning
National Chengchi University
Taipei, Taiwan
hero@nccu.edu.tw*

WAN-JUN LAI

*Graduate Institute of Linguistics, National Chengchi University
Taipei, Taiwan
98555001@nccu.edu.tw*

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)

Due to the lack of a rigorous methodology and explicit criteria to distinguish between classifiers (C) and measure words (M), previous inventories of Mandarin C's, or *geti liangci* 個體量詞, vary greatly. Based on the insight that an M in a Chinese [Num C/M N] phrase is semantically substantive, while a C is semantically redundant and thus does not block numeral quantification or adjectival modification to the noun, this paper further proposes that while C/M both function as a multiplicand mathematically, with Num as the multiplier, C's value is necessarily 1 and M is not, thus ~ 1 . Cognitively, however, the semantically redundant C serves to profile an inherent semantic feature of N and thus selects a narrow class of N's. With these explicit distinctions between C and M, we then re-examine the inventory of C's put forth in 國語日報量詞典 *Mandarin Daily News Dictionary of Measure Words* and offer a much more reliable list of C's in Taiwan Mandarin.

Keywords: classifier; measure word, profile, multiplicand, Taiwan Mandarin.

1. Introduction

Previous studies on Mandarin Chinese classifiers and measure words have come up with drastically different inventories. One crucial factor for the huge

2 One-Soon Her and Wan-Jun Lai

discrepancies is whether classifiers (C), as in (1), and measure words (M), as in (2), can be meaningfully and accurately distinguished.

(1) a. 一百 尾 金魚
yibai wei jinyu
100 C goldfish
'100 goldfish'

b. 三 根 鉛筆
san gen qianbi
3 C pencil
'3 pencils'

c. 十 匹 馬
shi pi ma
10 C horse
'10 horses'

(2) a. 一百 箱 金魚
yibai xiang jinyu
100 M-box goldfish
'100 box of goldfish'

b. 三 公斤 香蕉
san gongjin xianqian
3 M-kilo pencil
'3 kilos of pencils'

c. 十 群 馬
shi qun ma
10 M-herd horse
'10 herds of horses'

Even for those that do support a formal C/M distinction, such distinctions have not been made explicit, and many works on C/M simply assume that C and M are distinguishable and distinguish the two rather subjectively. An informal but insightful characterization is offered in Tai and Wang [1, p.38]:

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.

For natural language processing, how to make a formal distinction is obviously an important issue, as one of the most distinctive characteristics of Chinese is the C's, not the M's. For example, in a machine translation system for Chinese and English, most C's in Chinese have no counterparts in English, while most M's do, and for the reverse direction, most M's in the English [Num M of N] construction can receive straightforward translations in the Chinese [Num M N] construction, but the English [Num N_{count}] sequence must be turned into [Num C N] in Chinese, where the insertion and choice of C affects grammaticality and facility.

In spite of the same syntactic position C/M occupy in the [Num C/M N] sequence, in this paper we will fully justify the C/M distinction from three perspectives: semantic, mathematical, and cognitive. Section 2 first summarizes the explicit tests developed in [2] based on C/M's semantic distinction. Section 3 then characterizes the C/M distinction in set-theoretic terms. In section 4, we follow [2] and [3] and propose that C/M both function as a multiplicand mathematically, with Num as the multiplier, where C is necessarily of the value 1 and M is not. From a cognitive linguistic point of view, section 5 then approaches the issue from the standpoint of cognitive linguistics and shows that the semantically and mathematically null C nonetheless functions to profile an inherent semantic aspect of the head noun. This strict distinction of C versus M makes it possible to identify true classifiers in a language. A report is given in section 6 on the classifiers identified from the category of 'general measure words' listed in 國語日報量詞典 *Mandarin Daily News Dictionary of Measure Words* [13], a dictionary based on Taiwan Mandarin data in the Sinica Corpus.

2. Formal Tests to Distinguish Classifiers and Measure Words

Her and Hsieh [4] observe that the two formal tests, i.e., *de*-insertion and adjectival modification, which proponents for the C/M distinction proposed previously, have been shown to be unreliable [5][6]. However, based on the insight that M, but not C, constitutes a barrier to numeral quantification and adjectival modification, they refine the previous two tests and come up with much more reliable and accurate formulations (Test A, B). They also restate *ge*-substitution as a heuristic (Test C) and observe that temporary measure words are often restricted to the number *yi* 'one' (Test D).

Test A: Numeral/Adjectival Stacking

(1) If [Num X Num Y N] is well-formed, then X = M and Y = C/M.

4 One-Soon Her and Wan-Jun Lai

e.g., 一箱十個蘋果, 一箱十包蘋果 vs. *一顆十包蘋果, *一顆十粒蘋果

(2) If $[\text{Num } A\text{-}X \text{ N}] = [\text{Num } X \text{ A-N}]$ semantically, then $X = C$ and $X \neq M$.

e.g., 一大顆蘋果 = 一顆大蘋果 vs. 一大箱蘋果 \neq 一箱大蘋果

(3) Given antonyms A_1 and A_2 , if $[\text{Num } A_1\text{-}X \text{ A}_2\text{-N}]$ is semantically well-formed, then $X = M$.

e.g., *一大顆小蘋果 vs. 一大箱小蘋果

(4) If $[A\text{-}X \text{ de } N]$ is semantically equivalent to $[A\text{-N}]$, then $X = C$.

e.g., 大顆的蘋果 = 大蘋果 vs. 大箱的蘋果 \neq 大蘋果

Test B: De-insertion

Test: $[\text{yi } M/*C \text{ de } N]$

e.g., 一箱(的)蘋果 vs. 一顆(*的)蘋果

Test C: Ge-substitution

Test: If $[\text{Num}_i \text{ X } N_j] = [\text{Num}_i \text{ ge } N_j]$ semantically, then $X = C$.

e.g., 十粒蘋果 = 十個蘋果 \neq 十箱蘋果

3. Semantic Distinction between Classifiers and Measure Words

Her and Hsieh [4] further employ the Aristotelian distinction between essential and accidental properties as well as the Kantian distinction between analytic and synthetic propositions to characterize the C/M distinction: C is semantically redundant; M is semantically substantive. Precisely, C indicates an essential property of the noun, and can be paraphrased as the predicate concept in an analytic proposition with the noun as the subject concept; M indicates an accidental property in terms of quantity, and can be restated as the predicate concept in a synthetic proposition with the noun as the subject concept. Given this characterization, M can be demonstrated to be more of a content word, thus open to innovations, while C is more a function word, thus forms a closed set resistant to innovations.

The semantic distinction of C/M can receive a mathematical interpretation in set-theoretic terms. In short, properties denoted by C do not contribute to the total compositional semantic content of the phrase. M, on the other hand, is semantically substantive in $[\text{Num } M \text{ N}]$ and thus does contribute semantic value specific to M only. This contrast can be made explicit in terms of set theory.

C/M Distinction in Set-theoretic Terms

Given a well-formed phrase $[\text{Num } K \text{ N}]$, X the set of properties denoted by K, and Y the set of properties denoted by N, K is C if $X \subset Y$; otherwise, K is M.

The fact that C does not contribute any semantic value to the semantics of the overall [Num C N] phrase is *not* because it has no semantic value itself; rather, again, it is because it does not contribute any semantic property that the noun does not already possess. This total overlap of semantic properties between C and N is the reason why modification or quantification on C is also on N. M, on the other hand, does contribute semantic properties to the [Num M N] phrase that N does not possess, and any modification or quantification on M thus does not extend to N. Therefore, the following three expressions with the same Num and N share exactly the same truth value, i.e., *three fish*, in spite of the different C's: 三隻魚 *san zhi yu* = 三條魚 *san tiao yu* = 三尾魚 *san wei yu*; yet, with each C replaced with a different M, each expression now has its own unique truth value, e.g., 三磅魚 *san bang yu* 'three pounds of fish' ≠ 三箱魚 *san xiang yu* 'three boxes of fish' ≠ 三打魚 *san da yu* 'three dozens of fish'.

4. Classifiers and Measure Words as Multiplicands

Most importantly, extending and integrating Landman's [7] view of C/M as parcelers, Borer's [7] insight that classifiers in Chinese and the plural suffix /-s/ function the same as dividers, and Au Young's [9][10] findings of the mathematical multiplication basis of classifiers, we propose that there is a multiplication relation between Num and C/M, i.e., [Num C/M] = [multiplier × multiplicand], but the crucial C/M distinction in terms of their mathematical value is that C's value is *necessarily 1*, and M's value is *not necessarily 1*, thus ~ 1 . In other words, an M can have any mathematical value, permanent or temporary, numerical or non-numerical, as long as it is not necessarily 1, while a C must always be translated to the numerical value of 1 and 1 only.

The mathematics proposed here that C is the multiplicand of the value 1 also formalizes the long-standing view that C serves to individuate the following N (e.g., [11] and [12]). Furthermore, the mathematics of C/M also provides another explanation why expressions with the same Num and N shall have the same truth value regardless of the different C's used, e.g., again, 三隻魚 *san zhi yu* = 三條魚 *san tiao yu* = 三尾魚 *san wei yu*, but the same is not true for M, e.g., again, 三磅魚 *san bang yu* 'three pounds of fish' ≠ 三箱魚 *san xiang yu* 'three boxes of fish' ≠ 三打魚 *san da yu* 'three dozens of fish'.

Thus, C, as the (redundant) multiplicand 1, can be omitted, if stylistically required, without affecting the truth value of the nominal phrase, but M cannot.

Mathematical Distinction of C/M

Given [Num X N], $X = C$ iff $X = 1$; otherwise, $X = M$.

e.g., [五張餅 = 五×1 餅 = 五餅] vs. [五打餅 = 五×12 餅 ≠ 五餅];

[二條魚 = 二×1 魚 = 二魚] vs. [二斤魚 = 二×斤 魚 ≠ 二魚]

Under this view of simple mathematics, the many classifiers in Chinese, while denoting an intricate system of classifying nouns, can be seen as many ways to *profile* some intrinsic semantic aspects of the nouns and ultimately the mathematical value of *one*. The concept of *profile* will be presented in the next section. This mathematical interpretation of C/M further explains why C, as the superfluous multiplicand *I*, may be optional, while M is obligatory, and also why C is semantically *null* and thus *transparent* to numeral quantification and adjectival modification, while M is not. Finally, note that under this mathematical interpretation of C/M, English lacks measure words altogether, given the fact that its multiplicand is restricted to *I* and grammaticalized as the nominal suffix *-s* and thus no longer part of the numeral and must be part of the head noun. Thus, Borer [8], contra to common misconceptions, is exactly right that English plural maker */-s/* is a C. We thus follow through and claim that while English plural suffix */-s/* is a C similar to the Chinese generic C *ge*, English has no measure words and putative M's should in fact be treated simply as common nouns.

Distinction between Chinese and English

Chinese: [Num X N], X=1 (C) or ~1 (M)

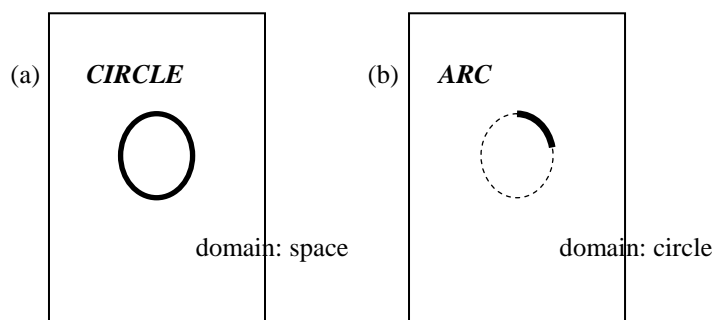
e.g., $3 \times 1\text{book} = \text{三本書}$; $3 \times \text{pile book} = \text{三堆書}$

English: [Num X N], Num>1 and X=1 (C)

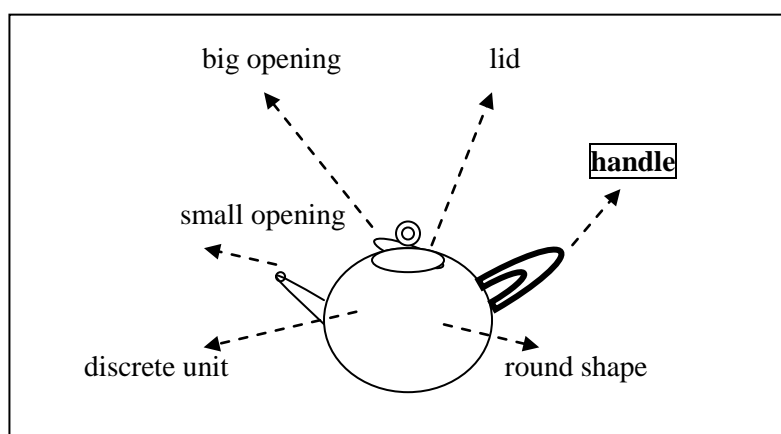
e.g., $3 \times I\text{book} = 3 -s \text{ book} = 3 \text{ books}$; $3 \times I\text{pile} = 3 -s \text{ pile} = 3 \text{ piles}$

5. Classifiers as Profilers

Besides the functions of a pacrler [7], divider [8], and multiplicand [9][10], which C/M have in common, C, being semantically and mathematically redundant, is shown to have the unique function as a profiler [13]. The notion of profile/base segregation has a strong connection to gestalt psychology, a comprehensive model of perception organization. Langacker [14] illustrates the notion of domain/profile by the example of *circle/arc*.



Similarly, Hsieh [6] and Her [3] argue that in a [Num C N] phrase, N can be seen as the base, or likewise the ‘frame’, in the sense of Fillmore’s frame semantics [15], and C the profile. Below, in the example of 一把壺 *yi ba hu* ‘a teapot’, it is shown that the teapot provides the frame or base for *ba* to profile, or to highlight, the handle, an inherent semantic feature of teapot.



Under this view, the classifiers’ function of classification is merely a by-product of their function as profilers. This thus explains at least partially why in Chinese or any other classifier language, there are always numerous idiosyncrasies or gaps in the noun classes categorized by classifier.¹

6. Identifying True Classifiers in Taiwan Mandarin

Based on the discussions above, we are now able to properly define C/M, both occurring between Num and N and serving as the multiplicand mathematically, with Num as the multiplier. However, C serves as an individuating unit, which also must profile a (bundle of) inherent semantic feature(s) of an N that denotes an intrinsically discrete entity; thus, mathematically, C’s value is necessarily numerical and precisely 1. M, on the other hand, whose mathematical value can be anything except 1 and thus may or may not be numerical, provides a measuring unit of N and thus does not profile any inherent semantic feature of

¹ For a full exploration of the profiler analysis of classifiers, refer to the thesis by Hsieh [13], for which the first author of this current paper served as the adviser.

the N. Consequent of these properties, C/M can be distinguished by the formal tests developed in section 2.

Only with such a precise characterization of C, is it feasible and practical to attempt a comprehensive list of C's in a language. However, we do not intend to provide such a comprehensive list of C's in Taiwan Mandarin and will only attempt a partial list based on *Mandarin Daily Dictionary of Chinese Classifiers* [16] (MDDCC hereafter).² MDDCC was compiled with data from the Academia Sinica Balanced Corpus of Modern Mandarin Chinese, a corpus that largely reflects Taiwan Mandarin. Seven categories of measure words (C/M) are identified, and the first category, called 一般量詞 *yiban liangci* 'general measure words', is intended to be precisely what is defined as C in this paper. Table 1 shows the 173 C's listed in MDDCC.³

Table 1: 173 Classifiers listed in MDDCC

ba3 把	dian3 點	hang2 行	ke1 顆	pi3 匹	tao4 套	zhao1 招
ban1 班	die2 疊	hao4 號	ke4 客	pian1 篇	ti2 題	zhen1 針
ban3 版	ding3 頂	hu4 戶	ke4 課	pian4 片	tiao2 條	zheng4 幀
ban4 瓣	ding4 錠	hui2 回	kou3 口	piao4 票	tie4 帖	zhi1 支
bang1 幫	dong4 棟	huo3 夥	kuai4 塊	pie3 撇	ting3 挺	zhi1 只
ben3 本	du3 堵	ji2 級	kuan3 款	pou2 杯	tou2 頭	zhi1 枝
bi3 筆	duan4 段	ji2 集	kun3 捆	qi2 畦	tuan2 團	zhi1 隻
bing3 柄	dui1 堆	ji2 輯	lan2 欄	qi2 期	tuo2 坨	zhi3 紙
bu4 部	dui4 隊	ji4 記	li4 粒	qi3 起	wan1 彎	zhou2 軸
cai2 槽	dui4 對	ji4 劑	lian2 聯	qu3 曲	wan1 灣	zhu1 株
ce4 冊	duo3 朵	jia1 家	liang4 輛	quan1 圈	wan2 丸	zhu4 柱
ceng2 層	fa1 發	jia4 架	lie4 列	que4 闕	wei3 尾	zhu4 炷
chong2 重	fang1 方	jian1 間	liu3 絡	qun2 群	wei4 位	zhuo1 桌
chu4 處	fang2 房	jian4 件	lu4 路	shan4 扇	wei4 味	zong1 宗
chuan4 串	fen4 分	jie1 階	lun2 輪	shen1 身	xi2 席	zu3 組
chuang2 床	fen4 份	jie2 節	luo4 落	sheng1 聲	xi2 襲	zun1 尊
chuang2 幢	feng1 封	jie2 截	lyu3 旅	shou3 首	xian4 線	zuo4 座
cong2 叢	fu2 服	jie4 介	lyu3 縷	shu4 束	xiang4 項	
cu4 簇	fu2 幅	jin4 進	mei2 枚	shuang1 雙	ye4 頁	
cuo1 撮	fu4 副	jing1 莖	men2 門	si1 絲	ye4 葉	

² For example, in this paper we will not discuss event measure words, another class identified in MDDCC, even though many of them can also be found to be genuine C's. For a more comprehensive inventory of C's in Taiwan Mandarin, the reader is referred to the second author's MA thesis [19], for which the first author again served as the advisor.

³ In the introduction section of MDDCC, it is stated that a total of 174 general measure words are listed in the dictionary. However, we were able to find only 173 of them, in spite of repeated efforts.

da3 打	gan3 桿	ju4 句	mian4 面	sou1 艘	yuan2 員
dai4 代	gen1 根	ju4 具	ming2 名	suo3 所	ze2 則
dai4 帶	ge 個	juan3 卷	pai2 排	tai1 胎	zha1 紮
dang3 檔	gu3 股	juan4 卷	peng3 捧	tai2 台	zhan3 盞
dao4 道	gua4 掛	ke1 科	pi1 匹	tan1 攤	zhang1 張
di1 滴	guan3 管	ke1 棵	pi1 批	tang2 堂	zhang1 章

In one the introduction sections of MDDCC, entitled *The Usage and Classification of Measure Words*, the editors in fact offer no explicit criteria for the selection of these 174 C's. The only relevant characterization is this: 「一般量詞是最典型的量詞...和數詞搭配用來記量物品的數量」 (*General measure words are the most prototypical measure words...they collocate with numbers and are used to count the number of things*) [16, p.10-11]. One of the several examples given in this section, 蔥兩把 *cong liang ba* 'green-onions, two handfuls', in fact involves 把 *ba* 'handful' as an M, not C.⁴ A more serious misconception is the inclusion of measure words that refer to groups or collections: 「...有些一般量詞則是用來記量物品組成集合後的數量，如「一群人、一對手錶、一束花、一批外套」。」 (...*some of the general measure words, however, refer to a collection of entities, for example, 'a group of people', 'a pair of watches', 'a bundle of flowers', and 'a batch of coats'.*) [16, p.11]. These are M's, not C's, as the reader can run the tests and find out, for they all have mathematical values that are not 1 and do not profile any inherent feature of the N. Thus, it seems that, like most, if not all, previous inventories of C's, the selection in this dictionary is also based on rather subjective judgments.

According to [17], a paper co-authored by the first editor of the dictionary, there are two ways to distinguish C/M that can be found in the relevant literature. One is by way of the *de*-insertion test, where a [Num C] sequence resists *de*-insertion, while *-de* can be freely inserted after [Num M]. Recent works, most notably [4], [5], and [18], have proven this test highly unreliable, using both corpus data and solicited data. What we have adopted in this paper is Test B in section 2, where only the Num 1, thus [yi C], is found to be resistant to *de*-insertion. The other criterion advocated in [17] is based on the informal characterization by Tai and Wang [1], quoted in section 1, i.e., C categorizes a particular type of N and also picks out a salient property of N, but M does not. These informal criteria can be rather subjective. In comparison, our methodology, while maintaining Tai and Wang's [1] conception of C/M distinction and reinterpreting it in terms of the concept of profiling, employs the three sets of formal tests developed in section 2 and also the mathematically precise test that C has the exact value of numeral 1 and M does not.

⁴ 把 *ba* can indeed be a C and profiles the part of an article that functions as a handle, e.g., 一把刀 *yi ba dao* 'one knife'.

We have therefore carefully re-examined the 173 items and come up with a revised list of true classifiers. In order to be more accurate, Table 2 shows the 76 items that are C's and C's only, and Table 3 shows the 21 items that can function as both C's and M's. Thus, out of MDDC's 173 putative C's, all together only 97 are confirmed to be genuine C's. A total of 76 items in MDDC's 173 C's are in fact M's, not C's, as shown in Table 4.

Table 2: 76 Classifiers out of the MDDCC's 173

ben3 本	yi1 ben3 shu1 一本書	jian4 件	yi1 jian4 da4yi1 一件大衣	ting3 挺	yi1 ting3 ji1qiang1 一挺機槍
bi3 筆	yi1 bi3 shou1ru4 一筆收入	jie4 介	yi1 jie4 shu1sheng1 一介書生	tou2 頭	yi1 tou2 da4xiang4 一頭大象
bing3 柄	yi1 bing3 fu3tou2 一柄斧頭	jing1 莖	yi1 jing1 bai2fa3 一莖白髮	wan1 彎	yi1 wan1 ming2yue4 一彎明月
ce4 冊	yi1 ce4 shu1 一冊書	ju4 句	yi1 ju4 kou3hao4 一句口號	wan1 灣	yi1 wan1 liu2shui3 一灣流水
chu4 處	yi1 chu4 shang1kou3 一處傷口	ju4 具	yi1 ju4 shi1ti3 一具屍體	wan2 丸	yi1 wan2 yao4wan2 一丸藥丸
chuang2 床	yi1 chuang2 mian2bei4 一床棉被	juan3 卷	yi1 juan3 lu4yin1dai4 一卷錄音帶	wei3 尾	yi1 wei3 yu2 一尾魚
chuang2 幢	yi1 chuang2 lou2fang2 一幢樓房	ke1 棵	yi1 ke1 song1shu4 一棵松樹	wei4 位	yi1 wei4 lao3shi1 一位老師
dang3 檔	yi1 dang3 gu3piao4 一檔股票	ke1 顆	yi1 ke1 xi1gua1 一顆西瓜	xi2 席	yi1 xi2 dong3shi4 一席董事
dao4 道	yi1 dao4 zhuan1qiang2 一道磚牆	li4 粒	yi1 li4 hong2dou4 一粒紅豆	xi2 襲	yi1 xi2 bo2sha1 一襲薄紗
ding3 頂	yi1 ding3 mao4zi 一頂帽子	liang4 輛	yi1 liang4 jing3che1 一輛警車	yuan2 員	yi1 yuan2 da4jiang4 一員大將
ding4 錠	yi1 ding4 yuan2bao3 一錠元寶	mei2 枚	yi1 mei2 jiang3zhang1 一枚獎章	ze2 則	yi1 ze2 xiao4hua4 一則笑話
dong4 棟	yi1 dong4 da4 lou2 一棟大樓	mian4 面	yi1 mian4 jing4zi 一面鏡子	zhan3 盞	yi1 zhan3 deng1 一盞燈
du3 堵	yi1 du3 qiang2 一堵牆	ming2 名	yi1 ming2 xue2sheng1 一名學生	zhang1 張	yi1 zhang1 chunag2 一張床
duo3 朵	yi1 duo3 mei2gui1 一朵玫瑰	pi1 匹	yi1 pi1 ma3 一匹馬	zhao1 招	yi1 zhao1 ce4lue4 一招策略
fa1 發	yi1 fa1 zi3dan4	pian1 篇	yi1 pian1	zheng4	yi1 zheng4

	一發子彈		wen2zhang1 一篇文章	幀	jie2hun1zhao4 一幀結婚照
fang1 方	yi1 fang1 yin4zhang1 一方印章	qi2 畦	yi1 qi2 dao4tian2 一畦稻田	zhi1 只	yi1 zhi1 jiu3tan2 一只酒罈
feng1 封	yi4 feng1xin4 一封信	qu3 曲	yi1 qu3 liu2xing2ge1 一曲流行歌	zhi1 枝	yi1 zhi1 shu4zhi1 一枝樹枝
fu2 幅	yi1 fu2 hua4 一幅畫	que4 闕	yi1 que4 gu3ci2 一闕古詞	zhi1 隻	yi1 zhi1 mao1 一隻貓
gan3 桿	yi1 gan3 qiang1 一桿槍	shan4 扇	yi1 shan4 men2 一扇門	zhi3 紙	yi1 zhi3 qie4jie2shu1 一紙切結書
gen1 根	yi1 gen1 tou2fa3 一根頭髮	sheng1 聲	yi1 sheng1 jian1jiao4 一聲尖叫	zhou2 軸	yi1 zhou2 hua4 一軸畫
ge 個	yi1 ge ren2 一個人	shou3 首	yi1 shou3 er2ge1 一首兒歌	zhu1 株	yi1 zhu1ying1hua1 一株櫻花
guan3 管	yi1 guan3 mao2bi3 一管毛筆	sao1 艘	yi1 sao1 chuang2 一艘船	zhu4 柱	yi1 shu4 dian4xian4gan1 一柱電線杆
ji4 記	yi1 ji4 zuo3gou1quan2 一記右勾拳	suo3 所	yi1 suo3 da4xue2 一所大學	zhu4 炷	yi1 shu4 xiang1 一炷香
ji4 劑	yi1 jie4 qiang2xin1ji4 一劑強心劑	ti2 題	yi1 ti2 xuan3ze2ti2 一題選擇題	zun1 尊	yi1 zun1 fo2xiang4 一尊佛像
jia4 架	yi1 jia4 fei1ji1 一架飛機	tiao2 條	yi1 tiao2 wei2jing1 一條圍巾	zuo4 座	yi1 zuo4 shan1 一座山
jian1 間	yi1 jian1 shu1dian4 一間書店				

Table 3: 21 dual status C/M out of the MDDCC's 173 Classifiers

ba3 把	ba3 把 c	yi1 ba3 dao1zi 一把刀子	ba3 把 m	yi1 ba3 tong2ban3 一把銅板
ban1 班	ban1 班 c	yi1 ban1 fei1ji1 一班飛機	ban1 班 m	yi1 ban1 xue2sheng1 一班學生
ban4 瓣	ban4 瓣 c	yi1 ban4 hua1ban4 一瓣花瓣	ban4 瓣 m	yi1 ban4 ju2zi 一瓣橘子
bu4 部	bu4 部 c	yi1 bu4 qi4che1 一部汽車	bu4 部 m	yi1 bu4 shu1 一部書
dian3 點	dian3 點 c	yi1 dian3 zhi4 一點痣	dian3 點 m	yi1 dian3 qian2 一點錢

12 One-Soon Her and Wan-Jun Lai

fen4 分	fen4 分 c	yi1 fen4 bao4gao4 一分報告	fen4 分 m	yi1 fen4 qing2yi4 一分情意
fen4 份	fen4 份 c	yi1 fen4 bao4gao4 一份報告	fen4 份 m	yi1 fen4 qing2yi4 一份情意
jia1 家	jia1 家 c	yi1 jia1 gong1si1 一家公司	jia1 家 m	yi1 jia1 ao4zhou1ren2 一家澳洲人
jie2 節	jie2 節 c	yi1 jie2 che1xiang1 一節車廂	jie2 節 m	yi1 jie2 gan1zhe4 一節甘蔗
kou3 口	kou3 口 c	yi1 kou3 jing3 一口井	kou3 口 m	yi1 kou3 zhu4ya2 一口蛀牙
kuai4 塊	kuai4 塊 c	yi1 kuai4 zhuan1tou2 一塊磚頭	kuai4 塊 m	yi1 kuai4 di4 一塊地
lun2 輪	lun2 輪 c	yi1 lun2 ming2yue4 一輪明月	lun2 輪 m	yi1 lun2 bi3sai4 一輪比賽
lü3 縷	lü3 縷 c	yi1 lü3 xian4 一縷線	lü3 縷 m	yi1 lü3 qing1yan1 一縷清煙
men2 門	men2 門 c	yi1 men2 da4pao4 一門大砲	men2 門 m	yi1 men2 sheng1yi4 一門生意
pian4 片	pian4 片 c	yi1 pian4 shu4ye4 一片樹葉	pian4 片 m	yi1 pian4 nai3you2 一片奶油
qi3 起	qi3 起 c	yi1 qi3 yi4wai4 一起意外	qi3 起 m	yi1 qi3 ren2ma3 一起人馬
tai2 台	tai2 台 c	yi1 tai2 dian4shi4 一台電視	tai2 台 m	yi1 tai2 ge1zai3xi4 一台歌仔戲
xian4 線	xian4 線 c	yi1 xian4 che1dao4 一線車道	xian4 線 m	yi1 xian4 xi1wang4 一線希望
ye4 葉	ye4 葉 c	yi1 ye4 pian1zhou1 一葉扁舟	ye4 葉 m	yi1 ye4 shu1 一葉書
zhi1 支	zhi1 支 c	yi1 zhi1 ge1 一支歌	zhi1 支 m	yi1 zhi1 chun2mao2sha1 一支純毛紗
zong1 宗	zong1 宗 c	yi1 zong1 yi4wai4 一宗意外	zong1 宗 m	yi1 zong1 huo4wu4 一宗貨物

Table 4: 76 Measure Words out of the MDDCC's 173 Classifiers

ban3 版	yi1 ban3 xin1wen2 一版新聞	hui2 回	ba1shi2hui2 hong2lou2meng4 八十回紅樓夢	piao4 票	yi1 piao4 sheng1yi4 一票生意
bang1 幫	yi1 bang1 gong1ren2 一幫工人	huo3 夥	yi1 huo3 qiang2dao4 一夥強盜	pie3 撇	yi1 pie3 hu2xu1 一撇鬚鬚
cao2 槽	yi1 cao2 ya2	ji2 級	yi1 ji2 shi2jie1	pou2 杯	yi1 pou2 tu3

	一槽牙		一級石階		一抔土
ceng2 層	yi1 ceng2 lou2 一層樓	ji2 集	yi1 bai3 ji2 lian2 xu4 ju4 一百集連續劇	qi2 期	za2 zhi4 di4 yi1 qi2 雜誌第一期
chong2 重	wan4 chong2 shan1 萬重山	ji2 輯	cong2 shu1 di4 yi1 ji2 叢書第一輯	quan1 圈	yi1 quan2 liu3 shu4 一圈柳樹
chuan4 串	yi1 chuan4 fo2 zhu1 一串佛珠	jie1 階	yi1 jie1 lou2 ti1 一階樓梯	qun2 群	yi1 qun2 peng2 you3 一群朋友
cong2 叢	yi1 cong2 ye3 cao3 一叢野草	jie2 截	yi1 jie2 zhu2 zi 一截竹子	shen1 身	yi1 shen1 yi1 shang 一身衣裳
cu4 簇	yi1 cu4 mei2 gui1 一簇玫瑰	jin4 進	yi1 jin4 fang2 zi 一進房子	shu4 束	yi1 shu4 xian1 hua1 一束鮮花
cuo1 撮	yi1 cuo1 mao2 fa3 一撮毛髮	juan4 卷	za2 zhi4 di4 yi1 juan4 雜誌第一卷	shuang1 雙	yi1 shuang1 xie2 一雙鞋
da3 打	yi1 da3 qian1 bi3 一打鉛筆	ke1 科	ying1 wan2 yi1 ke1 英文一科	si1 絲	yi1 si1 rou4 一絲肉
dai4 代	shang4 yi1 dai4 ren2 上一代人	ke4 客	yi1 ke4 niu2 pai2 一客牛排	tai1 胎	yi1 tai1 xiao3 gou3 一胎小狗
dai4 帶	yi1 dai4 yu2 cun1 一帶漁村	ke4 課	yi1 ke4 shu4 xue2 一課數學	tan1 灘	yi1 tan1 shui3 一灘水
di1 滴	yi1 di1 yan3 lei4 一滴眼淚	kuan3 款	di4 yi1 kuan3 gui1 ding4 第一款規定	tang2 堂	yi1 tang2 jia1 ju4 一堂傢具
die2 疊	yi1 die2 chao1 piao4 一疊鈔票	kun3 捆	yi1 kun3 dao4 cao3 一捆稻草	tao4 套	yi1 tao4 can1 ju4 一套餐具
duan4 段	yi1 duan4 gan1 zhe4 一段甘蔗	lan2 欄	yi1 lan2 xin1 wen2 一欄新聞	tie4 帖	yi1 tie4 zhong1 yao4 一帖中藥
dui1 堆	yi1 dui1 tu3 一堆土	lian2 聯	er4 lian2 shou1 ju4 二聯收據	tuan2 團	yi1 tuan2 shi4 bing1 一團士兵
dui4 隊	yi1 dui4 shi4 bing1 一隊士兵	lie4 列	yi1 lie4 luo4 tuo2 一列駱駝	tuo2 坨	yi1 tuo2 nai3 you2 一坨奶油
dui4 對	yi1 dui4 fu1 qi1 一對夫妻	liu3 絡	yi1 liu3 tou2 fa3 一絡頭髮	wei4 味	hun1 cai4 wu3 wei4 葷菜五味
fang2 房	yi1 fang2 er2 sun1	lu4 路	yi1 lu4 ren2 ma3 一路人馬	xiang4 項	xing2 fa3 di4 ti1 xiang4

	一房兒孫				刑法第一項
fu2 服	yi1 fu2 zhong1yao4 一服中藥	luo4 落	yi1 luo4 bao4zhi3 一落報紙	ye4 頁	yi1 ye4 shu1 一頁書
fu4 副	yi1 fu4 kuai4zi 一副筷子	lü3 旅	yi1 lü3 bu4dui4 一旅部隊	zha1 紮	yi1 zha1 zhi3hua1 一紮紙花
gu3 股	yi1 gu3 xiang1qi4 一股香氣	pai2 排	yi1 pai2 shi4bing1 一排士兵	zhang1 章	di4yi1zhang1 nei4rong2 第一章 內容
gua4 掛	yi1 gua4 fo2zhu1 一掛佛珠	peng3 捧	yi1 peng3 sha1 一捧沙	zhen1 針	yi1 xhen1 qiang2xin1ji4 一針 強心劑
hang2 行	yi1 hang2 liu3shu4 一行柳樹	pi1 批	yi1 pi1 huo4 一批貨	zhuo1 桌	yi1 zhuo1 cai4 一桌菜
hu4 戶	yi1 hu4 nong2min2 一戶農民	pi3 匹	yi1 pi3 bu4 一匹布	zu3 組	yi1 zu3 ren2yuan2 一組人員
hao4 號	di4yi1hao4 dao4lu4 第一號道路				

7. Concluding Remarks

The precise distinction of classifiers (C) and measure words (M) in a classifier language like Chinese is an important issue for natural language processing, as one of the most distinctive characteristics of Chinese is its C's, not its M's. For example, in a machine translation system for Chinese and English, most C's in Chinese have no counterparts in English, while most M's do, and in the reverse direction, while most M's in English can receive straightforward translations in Chinese, [Num N_{count}] sequences must be turned into [Num C N] in Chinese. The insertion of a semantically appropriate C is crucial for grammaticality as well as facility. Based on Her and Hsieh's [4] insight that M in a Chinese [Num C/M N] phrase is semantically substantive, while C is semantically redundant and thus does not block the numeral quantification or the adjectival modification to the noun, this paper further proposes a formal distinction of C/M from a mathematical perspective. Synthesizing the concepts of parceler [7], divider [8], and multiplicand [9][10], we follow [2] [3] and propose that while C/M both function as a multiplier mathematically, C's value is necessarily 1 and M is not, thus ~1. The semantically null C nonetheless functions to profile an inherent semantic aspect of N. Finally, based on these strict distinctions of C versus M, a

report is given on the true classifiers identified in 國語日報量詞典 *Mandarin Daily News Dictionary of Measure Words* [16].

Acknowledgments

We thank the anonymous reviewers and the audience of CLSW2011 as well as the guest editors and the two anonymous reviewers of this IJCPOL special issue for their insightful comments. Comments, whether supportive or critical, from the latter reviewers were especially constructive and helpful in further clarifying the central issues discussed in the paper. Yet, we are solely responsible for any remaining errors. The findings reported here on the inventory of classifiers in Taiwan Mandarin have been presented as part of the second author's MA thesis [19], for which the first author served as the supervisor.

References

- [1] Tai, J. and L. Wang. 1990. A semantic study of the classifier *Tiao*. *Journal of the Chinese Language Teachers Association* 25(1) (1990) 35-56.
- [2] Her, O. 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.
- [3] Her, O. Structure of classifiers and measure words: A Lexical Functional account, invited talk given at the 16th International Lexical Functional Grammar Conference (LFG2011), University of Hong Kong, July 16-19, 2011.
- [4] Her, O. and C. Hsieh. On the semantic distinction between classifiers and measure words, *Language and Linguistics* 11(3) (2010) 527-550.
- [5] Zhang, N. The constituency of classifier constructions in Mandarin Chinese, *Taiwan Journal of Linguistics* 9 (1) (2011) 1-50.
- [6] Hsieh, M. *The Internal Structure of Noun Phrases in Chinese* (Taiwan Journal of Linguistics: Book Series in Chinese Linguistics. No. 2. Crane Publishing, Taipei, 2008).
- [7] Landman, F. *Indefinites and the Type of Sets* (Blackwell, Malden, MA, 2004).
- [8] Borer, H. *Structuring Sense. Vol. 1. In Name Only.* (Oxford University Press, Oxford 2005).

- [9] Au-Yeung, B. *An Interface Program for Parameterization of Classifiers in Chinese* (PhD Thesis, Hong Kong University of Science and Technology, 2005).
- [10] Au Yeung, B. Multiplication basis of emergence of classifier, *Language and Linguistics* 8(4) (2007) 835-861.
- [11] Aikhenvald, A. *Classifiers: a typology of noun categorization devices* (Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press, 2000).
- [12] Bisang, W. Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages, *Studies in Language* 20(3) (1996) 519-597.
- [13] Hsieh, C. *A Frame-based Approach to Classifiers: A Case Study of Taiwan Mandarin* (MA Thesis, National Chengchi University, 2009).
- [14] Langacker, R. *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*, (Stanford University Press, 1987).
- [15] Fillmore, C. Frame semantics, in *Linguistics in the Morning Calm*, ed. Linguistic Society of Korea, (Hanshin Publishing, Seoul, 1982) 111-137.
- [16] Huang, C., K. Chen, and C. Lai (eds.) *Mandarin Daily Dictionary of Chinese Classifiers* (Mandarin Daily Press, Taipei, 1997).
- [17] Huang, C. and K. Ahrens. Individuals, Kind and Events: Classifier Coercion of Nouns. *Language Sciences* 25(4) (2003) 353-373.
- [18] Li, X. 2011. *On the semantics of classifiers in Chinese..* (Ph.D. Dissertation, Bar-Ilan University, 2011). Available online at: <http://www.semanticsarchive.net/Archive/mY3YWYZO/>. Accessed on 2011/09/018.
- [19] Lai, W. *Identifying True Classifiers in Taiwan Mandarin* (MA Thesis, National Chengchi University, 2011).