

# Chapter 22

## Beyond one, two, three: Number matters in classifier languages

Cherry Chit-Yu Lam

The Open University of Hong Kong

Chinese has been widely recognised as a classic example of a numeral-licensing classifier language, where the presence of a classifier is obligatory for overt quantification of nouns. This paper presents new data from Mandarin and Hong Kong Cantonese (HKC) to show that the need of classifiers for quantification is not always that absolute. Systematic variation has been found with an extended range of numerals examined (numerals larger than three), and a wider coverage of nouns in terms of animacy. The findings present a consistent pattern that HKC has a stricter requirement for classifiers in enumeration as bare common nouns are not definite in HKC, and it lacks the alternative strategies found in Mandarin.

### 1 Introduction

Chinese, particularly Mandarin, has been an exemplar language with numeral-licensing classifiers. This paper presents new data from mainland Mandarin and Hong Kong Cantonese (HKC) which contradicts such a neat understanding.

It is generally understood that, in Mandarin and HKC, whenever overt quantification is expressed in a noun phrase, whether by quantifiers like *jǐ* (HKC *gei2*) ‘some’, or numerals like *sān* (HKC *saam1*) ‘three’, a classifier must be present, regardless of mass-count distinction (1–2).

- (1) Mandarin (Chierchia 1998: 92; Cheng & Sybesma 1999: 519)
- |                          |                         |
|--------------------------|-------------------------|
| a. liǎng *(zhāng) zhuōzi | b. sān *(píng) jiǔ      |
| two CLF table            | three bottle wine       |
| ‘two tables’             | ‘three bottles of wine’ |



(2) HKC (Sio 2006: 14; Cheng & Sybesma 2005: 272)

- |                       |                       |
|-----------------------|-----------------------|
| a. saam1 *(bun2) syu1 | b. jat1 *(bui1) seoi2 |
| three CLF book        | one cup water         |
| ‘three books’         | ‘a cup of water’      |

This paper focuses only on cases of enumerated common count nouns such as (1a) and (2a), since measure words are necessary to license the counting of mass nouns even in non-classifier languages like English. Indeed, measure words such as those in (1b) and (2b) are termed as “massifiers” in Cheng & Sybesma (1998), which are different from (count-)classifiers as in the (a) sentences. Massifiers are there to *create* a unit of measure, while the count-classifiers, or classifiers in short, are only there to *name* the unit of counting which are inherent to the entity itself.<sup>1</sup> New data present a systematic pattern that the classifier can be optional, sometimes even disfavoured, in a [NUM + CLF + N] structure when the numeral size reaches a certain point. Furthermore, HKC has been found much less permissive with this exception than Mandarin. This new pattern challenges the traditional view (i.a. Krifka 1995; Chierchia 1998; Cheng & Sybesma 1999; 2005; 2012; Doetjes 1996) that numerals in a classifier language like Chinese obligatorily require licensing by the classifier; and forms a consistent picture with the general observation that Cantonese more strictly requires classifiers for individuation than Mandarin.

## 2 Beyond one, two, three: A new perspective

### 2.1 Theoretical background: Krifka (1995) and Chierchia (1998)

Krifka (1995) and Chierchia (1998) offer two classical analyses for Chinese-style classifier languages, where classifiers license enumeration.<sup>2</sup> Krifka suggests that the presence and absence of (the need for) classifiers is determined by whether the numerals in the language have a built-in measure function. In Mandarin, he argues, numerals do not come with such a measure function, hence whether the

---

<sup>1</sup>According to Cheng & Sybesma (1998), massifiers can be used with mass and count nouns, such as, *liǎng bēi shuǐ* ‘two glasses of water’ and *yī qún niǎo* ‘a flock of birds’ – massifiers with count nouns have also been known as “group classifiers” as pointed out by a reviewer.

<sup>2</sup>This numeral-licensing function of Chinese-style classifiers contrasts with the classifier system in languages like Japanese (Watanabe 2006), Purepecha (Vázquez-Rojas Maldonado 2012), and Niuean (Massam 2009), where numerals are classifier-licensing, i.e. classifiers can only occur when a numeral is present. This can be seen in the cases of [CLF + N] in argument positions in both Cantonese and Mandarin, though the two varieties differ in terms of whether such noun phrases can appear as subjects or not (cf. Cheng & Sybesma 1999, Sio 2006).

measuring unit is an “object unit” (OU) – a unit that measures the number of specimens of a kind, or a “kind unit” (KU) – a unit that measures subspecies, is left underspecified (Krifka coined that as “object or kind unit” (OKU)). Assuming that OU or KU can only apply to objects but not kinds, the presence of a classifier not only specifies which measuring unit is in use, but also generates an object-referring interpretation for the entity denoted by the noun. The contrary is true in English. English numerals have this measure function inherently, and hence can express what [NUM + CLF] does in Mandarin. This distinction in measure function of numerals has been used to account for typological differences between classifier and non-classifier languages in Krifka (1995); but Bale & Coon (2014) has found in Mi’gmaq (Algonquian) and Chol (Mayan) that such a distinction can appear within a language. In other words, while numerals in different languages can vary in terms of present/absence of measure function hence producing non-classifier and classifier languages respectively, different numerals within a language can also vary in the same way. In the latter case, some numerals can go directly with count nouns, but some cannot. In Mi’gmaq, for instance, Bale & Coon reported that “numerals 1–5 (along with numerals morphologically built from 1–5) do not appear with classifiers, while numerals 6 and higher must” (Bale & Coon 2014: 700), as illustrated in (3–4).

## (3) Mi’gmaq

- a. na’n-ijig ji’nm-ug  
five-AGR man-PL  
‘five men’
- b. \* na’n te’s-ijig ji’nm-ug  
five CLF-AGR man-PL

## (4) Mi’gmaq

- a. \* asugom-ijig ji’nm-ug  
six-AGR man-PL
- b. asugom te’s-ijig ji’nm-ug  
six CLF-AGR man-PL  
‘six men’

On the other hand, Chierchia (1998) explains such difference between Mandarin and English, or rather classifier and non-classifier languages in general, by the inherent properties of their nominals. He suggests that all common nouns in (Mandarin) Chinese are mass nouns; and all mass nouns are inherently plural (a.k.a. “inherent plurality hypothesis”). Chierchia explains that count nouns are

inherently singular, and become pluralised when used to refer to a set of singularities. Singular count nouns form singleton sets and are rudimentary building blocks of all other plural sets (Chierchia termed them *atoms*). Thus, plural count nouns denoting a group of singularities are conceptualised as union relations ( $\cup$ ). To Chierchia, mass nouns are plural-like; only that plural count nouns are sets formed by union of *atoms* while mass nouns are “the closure under  $\cup$  of a set of *atoms*” (Chierchia 1998: 70). In other words, mass nouns denote an enclosed union of all sets, and in that way, neutralize the difference between plural (i.e. sets) and singular (i.e. atoms). Therefore, Chierchia suggests that Mandarin common nouns provide a neat exemplar for the four mass nouns criteria in (5).

- (5) Mass properties of Chinese nouns (Chierchia 1998: 94)
- a. There is no plural marking.
  - b. A numeral can combine with a noun only through a classifier.
  - c. There is no definite or indefinite article.
  - d. Nouns can occur bare in argument position.

Focussing mainly on the second property concerning the distribution of numerals and classifiers, empirical data in §2.2 shows that the claim made in (5b) is too strong to hold. Turning back to Krifka’s alternative, the proposal that the need for classifier stems from the absence of a measure function in numerals seem more plausible, especially with the re-interpretation in Bale & Coon (2014). However, the patterns in Mandarin and HKC are not as clear-cut as that in Mi’gmaq and Chol, which may pose a challenge to an analysis that is purely along the lines of Krifka.

## 2.2 Number size and classifiers

One key observation made from the examples used in existing literature on Chinese classifiers is that most (if not all) examples are confined to the numerals one, two, and three. This study has examined numerals beyond three. Table 22.1 has the list of numerals tested; these are all cardinal numbers.

These nineteen numerals, ranging from 1 to 11000, are used with eight common count nouns in Mandarin and HKC to form noun phrases which appear as either subject or object in simple declarative sentences. The eight nouns considered are presented in Table 2. They vary in terms of degree of animacy (from human to inanimate) and number of syllables (mono- or disyllabic). (6) and (7) are some sample sentences.

Table 22.1: Chinese numerals

Mandarin	HKC		
<i>yī</i>	<i>jat1</i>	one	1
<i>liǎng</i>	<i>loeng5</i>	two	2
<i>sān</i>	<i>saam1</i>	three	3
<i>sì</i>	<i>sei3</i>	four	4
<i>wǔ</i>	<i>m5</i>	five	5
<i>shí</i>	<i>sap6</i>	ten	10
<i>shí-yī</i>	<i>sap6-jat1</i>	ten-one	11
<i>shí-wǔ</i>	<i>sap6-m5</i>	ten-five	15
<i>èr-shí</i>	<i>ji6-sap6</i>	two-ten	20
<i>èr-shí-yī</i>	<i>ji6-sap6-jat1</i>	two-ten-one	21
<i>sān-shí</i>	<i>saam1-sap6</i>	three-ten	30
<i>sān-shí-yī</i>	<i>saam1-sap6-jat1</i>	three-ten-one	31
<i>sì-shí</i>	<i>sei3-sap6</i>	four-ten	40
<i>wǔ-shí</i>	<i>m5-sap6</i>	five-ten	50
<i>yī-bǎi</i>	<i>jat1-baak3</i>	one-hundred	100
<i>yī-bǎi-líng-wǔ</i>	<i>jat1-baak3-ling4-m5</i>	one-hundred-zero-five	105
<i>yī-qiān</i>	<i>jat1-cin1</i>	one-thousand	1000
<i>yī-wàn</i>	<i>jat1-maan6</i>	one-ten.thousand	10000
<i>yī-wàn-yī-qiān</i>	<i>jat1-maan6-jat1-cin1</i>	one-ten.thousand-one-thousand	11000

Table 22.2: Chinese nouns: Animacy and phonological size

	CLF	Mandarin	HKC
[+HUMAN]	<i>gè/go3</i>	<i>rén</i> ‘person’ <i>xuéshēng</i> ‘student’	<i>jan4</i> ‘person’ <i>hok6saang1</i> ‘student’
[+ANIMATE]	<i>zhī/zek3</i>	<i>gǒu</i> ‘dog’ <i>lánggǒu</i> ‘wolfhound’	<i>gau2</i> ‘dog’ <i>long4gau2</i> ‘wolfhound’
[−ANIMATE]	<i>kē/po1</i>	<i>shù</i> ‘tree’ <i>sōngshù</i> ‘pine tree’	<i>syu6</i> ‘tree’ <i>cung4syu6</i> ‘pine tree’
[−ANIMATE]	<i>běn/bun2</i>	<i>shū</i> ‘book’ <i>zìdiǎn</i> ‘dictionary’	<i>syu1</i> ‘book’ <i>zi6din2</i> ‘dictionary’

(6) Mandarin

- a. *sān-shí-yī* (gè) *rén* *cānjiā-le* *bǐsài*  
*three-ten-one* CLF *person* join-PFV *competition*  
 ‘Thirty-one people joined the competition.’
- b. *wǒ yāoqǐng-le yī-bǎi* ?/\* (ge) *xuesheng*  
 I invite-PFV *one-hundred* CLF *student*  
 ‘I invited one hundred students.’

(7) HKC

- a. *saap6* \*(*po1*) *syu6* *sei2-zo2*  
*ten* CLF *tree* die-PFV  
 ‘Ten trees died.’
- b. *ngo5* *maai5-zo2* *ji6-saap6-jat1* \*(*bun2*) *zi6din2*  
 I buy-PFV *two-ten-one* CLF *dictionary*  
 ‘I bought twenty-one dictionaries.’

Regarding the classifier–noun pairings in the study, all the common nouns under investigation are paired with the only appropriate classifier in the language. In Mandarin *gǒu* ‘dog’ appears with the classifier *zhí* (e.g. *ten* \**gè/zhí gǒu*), and in HKC *syu6* ‘tree’ with *po1* (e.g. *saap6* \**go3/po1 syu6*). The only “exception” is with the [+HUMAN] nouns, as there are two possible classifiers for the noun *student* – a general classifier *gè/go3* and a specific one *wèi/wai2*. But for better comparison with the monosyllabic [+HUMAN] noun *person*, which cannot go with the specific classifier *wèi/wai2*, the classifier used for both *student* and *person* in this paper is the general classifier *gè/go3*.

In the acceptability judgment task, Mandarin and HKC native speakers<sup>3</sup> were asked to judge the acceptability of these sentences with and without classifiers. The judgement results have revealed several interesting patterns. First, both Mandarin and HKC speakers allow the [+HUMAN] count noun, *person*, to take the

<sup>3</sup>The results reported in this paper are taken from the acceptability judgment questionnaire from 2014. Four native Mandarin speakers and four native Hong Kong Cantonese speakers, aged 25–30, were consulted. Two of the Mandarin speakers were from Guangdong province, and the other two from northern China near Tianjin; samples of both varieties were gender-balanced. Participants were asked to rate sentences on a four-point scale (0–3). By comparing with control sentences, the scale of acceptability was established (in terms of average score): 2.8–3.0 = completely acceptable (✓), 1.8–2.7 = marginally acceptable (?), 1.3–1.7 = unacceptable (?/\*), 0.0–1.2 = absolutely unacceptable (\*). These terminologies will be consistently adopted in this paper. Since little regional variation has been found between southern and northern Mandarin speakers, and for the convenience of exposition, the average judgment scores will be presented in the text.

[NUM + Ø + N] structure, regardless of the value of the numeral. However, more precisely, in Mandarin, tested numerals higher than 10 are all rated acceptable whether in subject or object positions. In HKC, when the noun phrase appears as object, the numerals have to be greater than 30, but when the noun phrase appears as subject, only numerals higher than 100 are rated acceptable. All other sentences with [NUM + Ø + *person*] (as subject or object) are considered marginally acceptable (none completely ill-formed).

Down the scale of animacy, while HKC has a pattern consistent with the traditional understanding, i.e. numerals must be licensed by classifiers; Mandarin speakers allow null-classifier enumeration more liberally, especially with two sets of numerals. The first set involves high numerals 1000, 10000, and 11000. In Mandarin, subject noun phrases allow these three numerals to occur without the mediation of a classifier whenever the noun is animate (object noun phrases require a human noun).<sup>4</sup> Even with nouns of lower animacy, these three numerals consistently show a higher score in Mandarin null-classifier noun phrases. More importantly, in Mandarin, the presence of a classifier is not preferred when the noun *rén* ‘person’ occurs with these three high numerals: those Mandarin sentences are considered marginally acceptable (2.5 for subject, and 2.0 for object) when the classifier is present, and completely acceptable (3.0) when it is not. HKC noun phrases are much more restricted for such exceptions: apart from the noun *jan4* ‘person’, no other nouns can be enumerated without the presence of a classifier, however large the numeral is.

One possible explanation for such unmediated quantification could be that the classifier is still present in the structure but phonologically (partially) covert. An anonymous reviewer has pointed out that there is often a glottal stop between the numeral and the noun whenever the classifier is absent, presumably, where the noun is [+HUMAN] and hence the potential classifier would be *gè* in Mandarin or *go3* in HKC. In the Jin varieties of northern China, for instance, their equivalent of *gè* has been reported to have a final glottal stop in addition to the one in the onset.<sup>5</sup> If the same unmediated quantification is found in the Jin varieties, then what happens there could be that since there are two glottal stops in the classifier *gè*, one of them remains as the “residue” of the classifier and licenses the numeral in the place of the classifier itself.

However, empirically, the Mandarin and Cantonese speakers consulted in this study have not displayed such an articulatory feature, and even if it is indeed the case, the phonological reduction process could only be acting as an additional

<sup>4</sup>In any case, the noun concerned has to be disyllabic.

<sup>5</sup>I thank a reviewer for introducing me to the observations in the Jin varieties.

trigger for the omission of the classifier when the noun is [+HUMAN], but not as a sufficient condition to account for the selective permissiveness of [NUM + Ø + N] which is shown to be sensitive to animacy and number size. Otherwise, it would predict that (i) all [+HUMAN] nouns allow [NUM + Ø + N] regardless of number size, and (ii) all nouns that can appear with *gè/go3* (such as, *apple*, *ball*, and other [-ANIMATE] nouns) allow [NUM + Ø + N], but neither is empirically true. In fact, going back to the Mandarin and HKC data, despite the absence of a glottal stop in the coda position of the classifier *gè/go3*, there is one in the onset. So, if, as the phonological reduction hypothesis goes, the glottal stop between the numeral and the noun can act as a reduced form of the classifier, then the glottal stop in the onset may work as well as the one in the code position, but as aforementioned, such an articulatory feature has not been observed and the phonological reduction hypothesis alone would have overgeneralised the pattern of classifier-less enumeration in Mandarin and HKC.

Therefore, the classifier system in the Jin varieties certainly deserves further investigation, but based on the Mandarin and HKC data so far, a more plausible explanation for the observed exception is that big numbers like *yì qiān* ‘one thousand’ and *yí wàn* ‘ten thousand’, like the English *thousands* and *millions*, are not numerals, but measure words (Lisa Cheng, p.c.). It is indeed the case that a measure word cannot co-occur with a classifier, as in (8).

- (8) a. Mandarin  
       \* wǔ jīn kē cài  
       five catty CLF vegetable
- b. HKC  
       \* saap6-jat1 doi6 go3 ping4guo2  
       ten-one bag CLF apple

Nevertheless, it is important to note that even though the presence of a classifier may be disfavoured at times, [NUM + CLF + N] is never an unacceptable structure. In other words, the null-classifier structure is an additional option, but never the only available option. Therefore, I suggest that these high numerals have an inherent measure function emerging in Mandarin (à la Krifka 1995), but has not yet been grammaticalized into a proper measure word. Therefore, when these high numerals occur, the noun can either be individuated by the measure function of the numerals and does not require a classifier, or be individuated by the classifier. The preference for either of the two individuation strategies varies from one speaker to another.

Another exception happens with the numeral *one*. Mandarin speakers consider direct enumeration marginally acceptable when the count noun is disyllabic and

non-human. More specifically, when the noun phrase is a subject, *one* can go directly with any non-human count nouns (the scores range from 2.0 to 2.3); and when it is an object, the count noun must denote an animal or a plant (both scored 1.8) but not a completely inanimate object like *dictionary* (scored 1.3). A possible explanation for this pattern is that Mandarin is developing an indefinite article: the slight subject-object asymmetry in the acceptability of [*one* + N] may be a sign of this being a still ongoing development. Chierchia (1998) suggests that the indefinite article is simply a variant for the first numeral, and this is a well-established grammaticalisation pathway (Heine & Kuteva 2002). Therefore, what Chierchia predicts for Mandarin – there is “no morpheme that combines directly with a noun and means what *a* means in English” (Chierchia 1998: 91) – may not be correct, since the presence of *one* without the mediation of a classifier can be interpreted as an indefinite article (9).

- (9) Mandarin  
 yì sōngshù sǐ-le  
 one pine.tree die-PFV  
 ‘One/a pine tree died.’

### 2.3 More than numbers

The data presented in §2.2 boils down to one general conclusion: classifiers can be optional in licensing a numeral, especially in Mandarin, depending on the size of the numeral. This observation points to two issues: (i) numeral size can determine the necessity of classifiers for individuation – *one* and high numerals behave differently, and (ii) HKC classifiers are much more obligatory for individuation than Mandarin classifiers. The first issue has been discussed in the previous section, thus this section is devoted to discussing the cross-linguistic variations in the use of classifiers.

The difference between Mandarin and HKC in permitting [NUM +  $\emptyset$  + N] structures is consistent with a more general pattern that HKC more strictly requires the presence of classifiers for individuation. Figures 22.1 and 22.2 summarise the Mandarin and HKC classifier paradigms.

On the one hand, §2.2 has shown that HKC only allows null-classifier enumeration with the noun *jan4* ‘person’ and when the numeral is greater than 100 (for subject) or 30 (for object); on the other hand, Cheng & Sybesma (1999) have famously identified that HKC allows [CLF + N] as both subject and object, whereas Mandarin only allows it as object. What appears to be two separate issues, can be rethought as one if we take another perspective on the second issue. HKC, in fact,

PN	<i>Xiǎomíng</i> 'Xiaoming'		<i>Lǐsī</i> 'Lisi'	PN
common noun used as PN	<i>lǎobǎn</i> 'boss'		<i>lǎoshī</i> 'teacher'	common noun used as PN
CN		<i>kànjiàn</i> see		CN
*CLF+N	<i>mífēng</i> '(the) bee'	'sees'	<i>dàngāo</i> '(the) cake'	CLF+N
<i>one+CLF+N</i>	<i>*zhī mífēng</i> CLF bee 'the bee'		<i>gè dàngāo</i> CLF cake 'the/a cake'	
	<i>yì zhī mífēng</i> one CLF bee 'a bee'		<i>yī gè dàngāo</i> one CLF cake 'a cake'	<i>one+CLF+N</i>

Figure 22.1: Mandarin classifier paradigm

PN	<i>Siu2ming4</i> 'Siuming'		<i>Daai6man4</i> 'Daaiman'	PN
CN used as PN	<i>lou5ban2</i> 'boss'		<i>lou5si1</i> 'teacher'	CN used as PN
*CN		<i>gin3-dou2</i> see-COMPL		?CN
*CN	<i>*mat6fong1</i> 'bee'	'saw'	? <i>dan6go1</i> 'cake'	
CLF+CN	<i>zak3 mat6fong1</i> CLF bee 'the bee'		<i>go3 dan6go1</i> CLF cake 'the/a cake'	CLF+CN
<i>one+CLF+N</i>	<i>jat1 zak3 mat6fong1</i> one CLF bee 'a bee'		<i>jat1 go3 dan6go1</i> one CLF cake 'a cake'	<i>one+CLF+CN</i>

Figure 22.2: HKC classifier paradigm

does not allow bare common nouns in subject position (*mat6fong1* ‘bee’ in Figure 22.2), except when they act as proper names (*lou5ban2* ‘boss’ in Figure 22.2). Therefore, instead of viewing the second issue as Mandarin disallowing [CLF + N] as subjects, it is more appropriate to see it as HKC requires a classifier for subject noun phrases with a common count noun. In that case, the two issues are unified to a general cross-linguistic variation that HKC more obligatorily requires the presence of a classifier for individuation, regardless of the need for enumeration. To account for this requirement in HKC, Cheng & Sybesma (1999) have suggested that classifiers express definiteness like the English determiner *the*, hence a classifier phrase (CLFP) is projected whenever a definite reading arises. Since they report that both HKC [CLF + N]s and Mandarin bare common nouns have a definite reading, the difference between the HKC strategy and the Mandarin one is that the former has an overtly articulated CLF<sup>0</sup> while the latter has an empty CLF<sup>0</sup>. In contrast, since bare common nouns in HKC are not definite, the classifier phrase which encodes definiteness is not projected in HKC bare common nouns. Therefore, assuming that Chinese requires a definite subject, bare common nouns cannot be subjects in HKC.

The issue of referentiality or definiteness can be a plausible explanation for the [CLF + N] and bare noun distinction in HKC and Mandarin, but it does not provide an answer for the difference in numeral-licensing function of classifiers in the two Chinese varieties, since both [NUM + CLF + N] and [NUM + ∅ + N] are indefinite.<sup>6</sup> The answer to this cross-linguistic variation in classifier use can be found in three related phenomena in Mandarin (none attested in HKC): (i) the development of *one* as an indefinite article (see §2.2); (ii) the presence of special forms for *two* and *three* – *liǎ* ‘two/two of’ and *sā* ‘three/three of’ (10); (iii) the use of plural marker *-men* for animate nouns/noun phrases (11).

---

<sup>6</sup>Huang (2015) views this [CLF + N] pattern from another perspective: numeral requirement (more specifically, *one* requirement). He interprets that Cantonese allows bare classifier phrases in both subject and object positions, Mandarin restricts their occurrences to environments with a governing verb or preposition, and generally prohibits them in subject position. This observation is captured in the null numeral ‘one’ micro-parameter (i).

- (i) a. In Mandarin, [<sub>one</sub> e] is [–strong], triggering Agree with CLF.
- b. In Cantonese, [<sub>one</sub> e] is [+strong], triggering Move of CLF.

In short, Huang claims that Cantonese has a [+strong] number head, and Mandarin a [–strong] one. This interpretation of the classifier paradigms is insightful, but still fails to capture the new data on null-classifier enumeration presented in this paper.

(10) Mandarin

- a. wǒ-men *liǎ* / *sā* (\**gè*) shì hǎo pínghou  
1PL two three CLF be good friend  
'We two/three are good friends.'
- b. *liǎ* (\**kē*) shù sǐ-le  
two CLF tree die-PFV  
'Two trees died.'

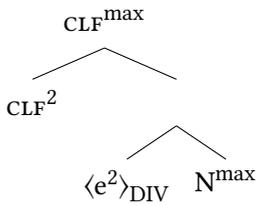
(11) Mandarin

- a. xuésheng-*men* xǐhuan chī miàn  
student-PL like eat noodles  
'The students like to eat noodles.'
- b. shí-èr *gè* xuésheng xǐhuan chī miàn  
ten-two CLF student like eat noodles  
'Twelve students like to eat noodles.'
- c. \*shí-èr *gè* xuésheng-*men* xǐhuan chī miàn  
ten-two CLF student-PL like eat noodles  
intended: 'Twelve students like to eat noodles.'

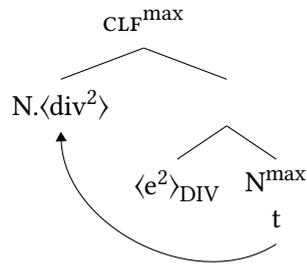
All three developments have one common property: the presence of classifiers become either optional or disallowed. The development of *one* as indefinite article in Mandarin allows the classifier to be optional when *one* appears with non-human (disyllabic) count nouns. The two special forms for *two* and *three* in Mandarin cannot occur with classifiers, because they themselves mean 'two of' and 'three of' respectively, meaning that they have inherent measure functions, just as the three high numerals 1000, 10000, and 11000. Finally, the fact that the Mandarin plural *-men* is much more developed than its HKC counterpart (*-dei6*) which can only suffix on pronouns, is another piece of evidence showing that Mandarin enumeration is less dependent on the use of classifiers. However, this only suggests that plural-marking and classifiers are competing strategies for the enumeration function, but not that they are morpho-phonological competitors, as they take up different structural positions. In Mandarin and Cantonese, for instance, classifiers are in pre-nominal position, while plural markers are post-nominal. Borer formalises the difference as: "the plural marker is a spell-out of an abstract head feature <div> [divided] on a moved N-stem, while the classifier is an independent f[unction]-morph occurring in the left-periphery of the N" (2005: 95), as represented in (12a,b) below:<sup>7</sup>

<sup>7</sup>Adapted from Borer (2005: 95), the open value <e><sub>DIV</sub> is the classifier head, and <div> is the plural head feature. The co-superscripts (e.g. *max*) indicate range assignment relations.

(12) a. [+CLF -PL]



b. [-CLF +PL]



Borer's explanation suffices for the complementary distribution of classifiers and plural markers but does not account for differences in the distribution of Cantonese and Mandarin classifiers.

### 3 Implications

This paper has presented new empirical data from mainland Mandarin and HKC, and a new perspective in viewing the classifier paradigms of the two Chinese varieties, particularly regarding the variation in distribution of bare classifier phrases in subject position. While previous studies have examined the issue from the angle of definiteness-encoding (Cheng & Sybesma 1999) – bare nouns vs. bare classifier phrases, and strength of numeral head (Huang 2015) – numeral phrases with *one* vs. bare classifier phrases, neither can account for empirical cases where classifiers are optional in licensing numerals in Chinese (especially Mandarin). Therefore, this paper opens a new way to rethink this puzzle by showing (i) how numeral size, animacy, and phonological size can determine classifier obligatoriness, and (ii) three related phenomenon that happened exclusively in Mandarin which weaken the need for classifiers in its individuation function – *one* as indefinite article, special forms for *two* and *three*, and plural marker with animate count nouns. These together should offer a more unified picture for the use of classifiers in Mandarin and HKC.

### Abbreviations

1 first person  
AGR agreement  
CLF classifier

CN common noun  
COMPL completive aspect  
HKC Hong Kong Cantonese

KU	kind unit	PFV	perfective
NUM	numeral	PL	plural
OKU	object or kind unit	PN	predicate nominal
OU	object unit		

## References

- Bale, Alan & Jessica Coon. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry* 45(4). 695–707. DOI: 10.1162/ling\_a\_00170.
- Borer, Hagit. 2005. *The normal course of events* (Structuring sense 2). Oxford: Oxford University Press.
- Cheng, Lisa Lai-Shen & Rint Sybesma. 1998. Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing-Hua Journal of Chinese Studies New Series* 28(3). 385–412.
- Cheng, Lisa Lai-Shen & Rint Sybesma. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4). 509–542. DOI: 10.1162/002438999554192.
- Cheng, Lisa Lai-Shen & Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In Guglielmo Cinque & Richard S. Kayne (eds.), *Handbook of comparative syntax*, 259–292. Oxford: Oxford University Press.
- Cheng, Lisa Lai-Shen & Rint Sybesma. 2012. Classifiers and DP. *Linguistic Inquiry* 43(4). 634–650. DOI: 10.1162/ling\_a\_00109.
- Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of ‘semantic parameter’. In Susan Rothstein (ed.), *Events and grammar*, 53–103. Dordrecht: Kluwer.
- Doetjes, Jenny. 1996. Mass and count: Syntax or semantics? In Ana Arregui & Crit Cremers (eds.), *Proceedings of Meaning on the HIL*, 34–52. Leiden: HIL.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Huang, C.-T. James. 2015. On syntactic analyticity and parametric theory. In Audrey Li, Andrew Simpson & Wei-Tien Dylan Tsai (eds.), *Chinese syntax in a cross-linguistic perspective*, 1–48. Oxford: Oxford University Press.
- Krifka, Manfred. 1995. Common nouns: A contrastive analysis of English and Chinese. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The generic book*, 398–411. London: University of Chicago Press.
- Massam, Diane. 2009. On the separation and relatedness of classifiers, number, and individuation in Niuean. *Language and Linguistics* 10(4). 669–699.
- Sio, Joanna U.-S. 2006. *Modification and reference in the Chinese nominal*. Leiden University. (Doctoral dissertation).

- Vázquez-Rojas Maldonado, Violeta. 2012. *The syntax and semantics of Purepecha noun phrases and the mass/count distinction*. New York University. (Doctoral dissertation).
- Watanabe, Akira. 2006. Functional properties of nominals in Japanese. *Natural Language & Linguistic Theory* 24(1). 241–306. DOI: 10.1007/s11049-005-3042-4.