On the interpretation of number and classifiers
Lisa Lai-Shen Cheng, Jenny Doetjes, Rint Sybesma & Roberto Zamparelli

Mandarin and Cantonese, both of which are numeral classifier languages, present an interesting puzzle concerning a compositional account of number in the various forms of nominals. First, bare nouns are number neutral (or vague in number). Second, cl-noun combinations appear to have different interpretations depending on contexts. When they occur in isolation (bare cl-noun), they can only be interpreted as singular. When they occur with numerals above one, or with expressions such as hěn duō/hou2 do1 ‘a lot’ they are interpreted as plural. This paper discusses the relevant data, as well as three potential solutions to this puzzle.

In section 1, we present a brief overview of the basic data in both Cantonese and Mandarin, including the co-occurrence of measure words such as hěn duō/hou2 do1 ‘a lot’ and classifiers. In section 2, we discuss the interpretation and syntactic status of the elements di1 (Cantonese) and xiē (Mandarin), which we argue to be a special type of classifiers (contra Iljic 1994). We present three potential analyses to explain the shift in number in Chinese nominals in section 3, and we show that though we can rule out one analysis, further study is needed to determine whether one of the remaining analyses is correct.

1. Background

The basic word order of Chinese noun phrases (both Cantonese and Mandarin) is:

(1) DEM NUM(eral) CL(assifier) MOD+ge/de Noun

We will present below an overview of the interpretation of Chinese noun phrases of various sizes, bare nouns, cl-noun, num-cl-noun, and di1/xiē + noun. In 1.1, we examine the distribution of measure words such as hěn duō/hou2 do1 ‘a lot’, which do not require a classifier, though in Cantonese and some dialects of Mandarin, they allow for the presence of one.

Consider first the examples in !).

(2)  a. wǒ mǎi-le shū  (Mandarin)
    I buy-perf book
    ‘I bought a book/books.’
These examples demonstrate the number neutral property of bare nouns. Rullmann and You (1983) show, moreover, that indefinite bare nouns do not have any quantificational force of their own, as bare indefinites cannot take wide scope. This is compatible with the idea that bare nouns denote a full join semi-lattice (Link 1983). In Mandarin, bare nouns can, under certain conditions, also have a definite interpretation (cf. Cheng & Sybesma 1999 for a discussion of the types of contexts that make this interpretation possible). We assume that this reading results from application of an iota operator (cf. Chierchia 1998), which is located in the head of a Classifier Phrase (ClP) dominating NP and which triggers N to Cl movement (cf. Cheng & Sybesma 1999).

Classifiers in Chinese languages come in two types: sortal and non-sortal classifiers (see among others Cheng and Sybesma 1998). (3a) is an example of sortal classifiers, while (3b,c) are examples of non-sortals. In both cases, we see that numerals precede the classifiers.

(3) a. sān zhī bǐ
three Clbranch pen
‘three pens’

b. liǎng xiāng shū
two box book
‘two boxes of books’

c. yī qún rén
one group person
‘one group of people’

Note also that any numeral can appear before the classifier (in the examples in (3) we see aside from the numeral yī ‘one’ also numerals higher than one).

Classifiers can also appear with nouns without the numeral. We call these ‘bare Cl-N’ combinations.

(4) a. wǒ mǎi-le běn shū
I buy-PERF Cl book
‘I bought a book.’ (one single book)

b. bun² syu¹ hou² cong⁵
Cl book very heavy
‘The book is very heavy.’ (one single book)
As seen in ), in contrast with bare nouns, which are number neutral (or vague in number), bare \textit{cl-n}s are interpreted as singulars. Note that in Cantonese, bare \textit{cl-n}s can have definite and indefinite interpretations in object position, but only definite interpretations in subject position b), while in Mandarin, bare \textit{cl-n}s can only appear in objects position, where they are only interpreted as indefinites (see Cheng and Sybesma 1999 and 2005 for more detailed discussion of this point).

1.2. \textit{di} (Cantonese) and \textit{xīē} (Mandarin)

To express an unambiguous plural, \textit{di} and \textit{xīē} are used in Cantonese and Mandarin respectively.

\begin{enumerate}
\item \textit{a.} \textit{yī \ xīē \ rěn} (Mandarin)
  \begin{itemize}
  \item \textit{one xie person}
  \item ‘some people’
  \end{itemize}
\item \textit{b.} \textit{yat\textsuperscript{1} \ di\textsuperscript{1} \ jan\textsuperscript{4}} (Cantonese)
  \begin{itemize}
  \item \textit{one di person}
  \item ‘some people’
  \end{itemize}
\item \textit{c.} \textit{di\textsuperscript{1} \ jan\textsuperscript{4}}
  \begin{itemize}
  \item \textit{di person}
  \item ‘the people’
  \end{itemize}
\end{enumerate}

As shown in a,b), \textit{xīē} and \textit{di\textsuperscript{1}} can be preceded by the numeral \textit{one}, though this still yields an indefinite plural reading. In Cantonese, \textit{di\textsuperscript{1}} can be used like typical classifiers in bare \textit{cl-n} combinations yielding a definite interpretation c). Furthermore, both \textit{xīē} and \textit{di\textsuperscript{1}} can also appear with mass nouns, as illustrated in ) (but see the contrast between ) and (20) below for a semantic difference between Mandarin \textit{xīē} and Cantonese \textit{di\textsuperscript{1}}, suggesting that only Cantonese \textit{di\textsuperscript{1}} is compatible with an interpretation that does not involve a plurality of discrete units).

\begin{enumerate}
\item \textit{a.} \textit{yī \ xīē \ shuǐ}
  \begin{itemize}
  \item \textit{one xie water}
  \item ‘some water’
  \end{itemize}
\item \textit{b.} \textit{yat\textsuperscript{1} \ di\textsuperscript{1} \ seoi\textsuperscript{2}}
  \begin{itemize}
  \item \textit{one di water}
  \item ‘some water’
  \end{itemize}
\end{enumerate}
1.2. Quantifiers

Typical quantifiers do not appear with classifiers, but with the modification particle de, as in (7). However, měi, which is typically translated as ‘every’, and jǐ ‘several’ precede a CL-N, or even a Numeral-CL-N sequence (8).²

(7) a. suǒyǒu de xuéshēng
   all DE student
   ‘all students’

   b. dàbùfēn de xuéshēng
   most DE student
   ‘most students’

(8) a. měi (yī) ge xuéshēng
   every one CL student
   ‘every student’

   b. jǐ ge xuéshēng
   several CL student
   ‘several students’

Interestingly, hěn duō ‘a lot’ in Mandarin and hou² do¹ ‘a lot’ in Cantonese can appear with classifiers:

(9) a. hěn duō (%běn) shū dōu zài tā-de zhuōzǐ
   very many CL book DOU at he-DE table top
   ‘Many books are on his table.’

   b. hou² do¹ (bun²) syu¹ dou¹ hai² keoi⁵-ge³ toi² seong⁶min⁸ (Cantonese)
   good many CL book DOU at he-GE table top
   ‘Many books are on his table.’

The % sign in (9a) indicates that Mandarin speakers do not all agree with respect to the presence of the classifier after hěn duō ‘a lot’. In contrast, in Cantonese, the classifier is optional after hou² do¹ ‘a lot’, as we see in (9b).

If bare nouns are number-neutral while bare CL-NS are singular, the question is how hěn duō/hou² do¹ ‘a lot’ can combine with both. Cross-linguistically, expressions such as hěn duō ‘a lot’ combine with expressions that have cumulative reference (cf. a lot, beaucoup, veel, molt+AGR etc. that all take a plural and/or a mass noun), and typically not with singulars. This question will be taken up in section 3.2 below.
2. Classifier or not? xǐē/dí¹

We have seen above that by using xǐē or dí¹, it is possible to express unambiguous plurality. The question is whether xǐē and dí¹ are (plural) classifiers. A simple argument in favor of classifier status of xǐē is its distribution: it follows the numeral yī ‘one’, just like a typical classifier, and in cases where yī is missing, it can directly follow a demonstrative, just like a typical classifier, as illustrated in ).

(10) a. wǒ mǎi-le yī-běn shū  
    I buy-PERF one-CL book  
    ‘I bought a book.’

b. wǒ mǎi-le yī-xǐē shū  
    I buy-PERF one-xǐē book  
    ‘I bought some books.’

(11) a. zhè běn shū hěn zhòng  
    this cl book very heavy
    ‘This book is very heavy.’

b. zhè xǐē shū hěn zhòng  
    this xǐē book very heavy
    ‘These books are very heavy.’

Furthermore, Cantonese dí¹ also behaves like a typical classifier in that it can appear in bare cl-n combinations and give rise to definiteness:

(12) dí¹ gau² hou² teng¹waa⁶
    dog very obedient
    ‘The dogs are very obedient.’

However, Iljic (1991) presents a number of arguments against xǐē in Mandarin as a plural classifier. First, xǐē can co-occur with the general classifier ge, as in ).

(13) zhème xǐē ge shū nà kàn-de-wǎn?
    so some cl book how read-de-finish
    ‘How can one read through so many books?’
    (from Iljic 1991, citing XHC 1977: 334)

Secondly, even though xǐē can appear with yī ‘one’, it cannot appear with higher numerals:
(14) a. *sān-xiē shū (Intended: ‘three books’)  
three-xie book  
b. *wǔ-xiē shū (Intended: ‘five books’)  
five-xie book

And lastly, when xiē takes the modifier hǎo ‘good’, it can be followed by different classifiers, such as jiān in):

(15) hǎo xiē (jiān) fángzi (adapted from Iljic 1991, ex. 30)  
good xie cl house/room  
‘a good few rooms’

Let us consider each of these arguments in turn. First, xiē can appear with the general classifier ge. Note that this is only limited to ge, and further, it is limited to certain dialects, thus suggesting that xiē-ge is potentially a variant of xiē. Furthermore, this argument cannot be used for Cantonese, since Cantonese never allows the counterpart of xiē, di, to co-occur with any classifier.

Consider next the argument that xiē only appears with the numeral yī ‘one’. This is actually expected if we assume that xiē combines a classifier and a measure or amount expression (like the English ‘number/quantity’ in a number of people/a quantity of water), as it is expected that it could not combine with greater numerals (see *two numbers of people). The co-occurrence with unstressed a but not with numerals is common to other measure words such as couple, lot, etc.:

(16) a. {*two lots / a lot} of books  
b. {*3 couples / a couple} of apples

As for hào-xiē (good-xie), which can appear with other classifiers, we suggest that this is similar to hèn-duō ‘a lot’, and seems to have obtained the status of a quantifying expression.

Aside from the simple distributional argument in favor of xiē / di as a classifier, both xiē and di behave like typical classifiers in that they can also license N-ellipsis. In (17), we provide two Mandarin examples with typical classifiers (both the general classifier ge and the classifier bèn), where the nouns following them are elided (data from Cheng and Sybesma 2009). In '), we see that xiē in Mandarin can also licenseellipsis just like typical classifiers.

(17) a. tā gāngcái chī-le yī-ge píngguǒ, nǐ yě yīnggāi chī yī-ge  
3s just-now eat-pref one-cl apple, 2s also ought eat one-cl  
‘he just ate an apple, you should also eat one’
On the interpretation of number and classifiers

b. tā bù xīhuān nèi-běn shū, tā xīhuān zhěi-běn
3s NEG like that-CL book, 3s like this-CL
‘he does not like that book, he likes this one’

(18) a. tā mài zhè -xiē shū, wǒ mài nèi-xiē
he buy this -XIE book I buy that-XIE
‘He buys these books, and I buy those.’

Iljic points out that xiē, when combining with mass nouns, yields discrete units. He illustrates it with the contrast in ).

(19) a. yī-diār shuǐ
one-bit water
‘a bit of water’

b. zhè-xiē shuǐ
this-XIE water
‘trickles of water/qualitative varieties of water’

The same however cannot be said about di¹ in Cantonese as illustrated by the various uses of di¹ in (20).³

(20) a. keoi⁵ jam²-zo² di¹ seoi²
he drink-PERF di water
‘He drank some water.’

b. keoi⁵ baai²-zo² di¹ mat⁶-tong⁴ hai² di¹ ca⁴ dou⁶
he put-PERF di honey be.at di tea in
‘He put some honey in the tea.’

For the sake of completeness, in the following table we give a full summary of the interpretations available across positions for various sizes of Chinese nominals:⁴

Table 1. Number and (in)definiteness in various Mandarin / Cantonese nominals

<table>
<thead>
<tr>
<th>Visible words</th>
<th>Mandarin</th>
<th>Cantonese</th>
<th>Mandarin</th>
<th>Cantonese</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 N</td>
<td>Def (sg/plur/mass)</td>
<td>*Def</td>
<td>*Def (sing/plur)</td>
<td>*Def (sing/plur)</td>
</tr>
<tr>
<td></td>
<td>*Indef</td>
<td>*Indef</td>
<td>Indef (sing/plur)</td>
<td>Indef (non spec)</td>
</tr>
<tr>
<td>2 CL N</td>
<td>*Def</td>
<td>Def (sing)</td>
<td>*Def</td>
<td>Def (sing)</td>
</tr>
<tr>
<td></td>
<td>*Indef</td>
<td>*Indef</td>
<td>Indef (sing)</td>
<td>Indef (sing)</td>
</tr>
<tr>
<td>3 NUM CL N</td>
<td>*Def</td>
<td>Def (plur)</td>
<td>*Def</td>
<td>Def (pl/mass)</td>
</tr>
<tr>
<td></td>
<td>*Indef</td>
<td>*Indef</td>
<td>Indef (pl/mass)</td>
<td>Indef (pl/mass)</td>
</tr>
<tr>
<td>4 CLPLUR N</td>
<td>*Def</td>
<td>Def (pl/mass)</td>
<td>*Def</td>
<td>Def (pl/mass)</td>
</tr>
<tr>
<td>(xiē-NMmandarin di¹-NCantonese)</td>
<td>*Indef</td>
<td>*Indef</td>
<td>Indef (pl/mass)</td>
<td>Indef (pl/mass)</td>
</tr>
<tr>
<td>5 <em>one</em> CLPLUR N</td>
<td>*Def</td>
<td>*Def</td>
<td>*Def</td>
<td>*Def</td>
</tr>
<tr>
<td></td>
<td>*Indef (pl/mass)</td>
<td>*Indef (pl/mass)</td>
<td>Indef (pl/mass)</td>
<td>Indef (pl/mass)</td>
</tr>
</tbody>
</table>
Given the above picture, the puzzle is how to derive a compositional account of number in these nominals: plural or singular in bare Ns, singular in bare CL-N, plural or singular in Num CL-N (depending on Num), only plural with \(xîê/dî^1\).

3. Three possible analyses

3.1 The Covert Numeral Hypothesis

The first and simplest hypothesis to consider is that a CL-N is actually number-neutral, but in the absence of an overt numeral, it is preceded by a phonetically non-overt numeral meaning ‘one’.

There is some evidence for this idea. First, in both Cantonese and Mandarin, when the demonstrative is present, the numeral one (\(yî\) (Mandarin), \(yat^1\) (Cantonese)) is optional, as shown in 21).

(21) a. zhè (\(yî\)bèn shū)  
this one CL_ book  
‘this book’

b. li^1 (\(yat^1\)) bun^2 syu^1  
this one CL_ book  
‘this book’

Further, in colloquial Mandarin, both the numeral one and the classifier can be omitted:

(22) zhè (\(yî\))(bèn) shū  
this one CL_ book  
‘this book’

This hypothesis works well for the post-verbal, ‘indefinite’ cases of CL-N and N, which would receive the minimal structure in 23):

(23) \([\text{NumP } yî/yat^1 [\text{CLP bèn/bun}^2 [\text{NP shū/syu}^1]]]\)

However, this hypothesis falls short of accounting for all the data. The problematic case is Cantonese: bare CL-N’s in Cantonese can have a definite interpretation. This is completely unexpected if bare CL-N’s are supposed to have a (covert) numeral \(yî\) ‘one’, since numerals block the definite reading (see Table 1, row 3). Thus, even though this analysis is plausible (and indeed, hard to rule out) for the ‘indefinite’ Mandarin cases, we need to search further to explain the singularity of the Cantonese definite cases.
3.2 Singularity is due to CL

If it is not an empty numeral one that turns the number neutral noun to a singular expression, it seems plausible that the classifier itself is the source of the singular interpretation of CL-N, as argued by Cheng & Sybesma (1999). While the bare noun denotes an atomic semi-lattice, the classifier reduces it to the set of atoms (cf. also Chierchia 1998). This means that a bare noun in itself is number neutral, as illustrated in ) above, while CL-N is a singular, as illustrated in (4).

However, this analysis predicts that any expression that combines with CL-N combines with a singular expression. For cases where the CL-N combines with the numeral yi or yat ‘one’ in Mandarin and Cantonese respectively, this is straightforward. However, in view of the examples in (3), where the classifier is present in the context of numerals higher than one and (9), where the classifier is optionally present with hén duól hou² do¹ ‘a lot’, the idea that the classifier creates a singular denotation needs to be discussed in more detail.

Let’s first turn to the numerals higher than one. The example in (3a) is repeated below as ):

(24) sán zhī bǐ (Mandarin)
    three CLbranch pen
    ‘three pens’

If zhī bǐ (CLbranch-pen) in ) denotes a set of atoms (that is, a singular denotation), this has consequences for the interpretation of the numeral. In most languages with a singular-plural opposition in the nominal morphology, the numeral (higher than one) combines with the plural and not with the singular noun. The numeral is usually seen as a filter: given a lattice structure, it filters out all plural individuals with a cardinality lower than the one indicated by the numeral. If one adopts the idea that CL-N is a singular expression, this type of analysis for numerals is not possible.

Recently, Ionin & Matushansky (2006) have proposed that even in languages such as English, the denotation of nouns that are modified by numerals (i.e., ‘books’ in three books) is singular rather than plural. The plural morphology that is found on the noun results from agreement with the numeral rather than from the presence of a plural interpretation. They assume that the meaning of the numeral is such that, on the basis of a set of atoms, it creates plural individuals the cardinality of which corresponds to the one indicated by the numeral.
As such, their analysis results in the same interpretation for two books as an analysis which takes the plurality of books as a starting point, and which selects from the set of plural individuals denoted by books the ones that have a cardinality of two.

However, Ionin and Matushansky’s analysis cannot account for all numerals cross-linguistically. Across languages, numerals are usually found with plural nouns (e.g., English), with number neutral nouns (e.g., Tagalog), or, in classifier languages, in combination with classifiers and number neutral nouns (e.g., Mandarin). In many classifier languages, the classifier is optional. An example is Khmer, an Austroasiatic language spoken in Cambodia (Jacob 1965), but also Armenian, where the numerals are found with number neutral nouns, with classified nouns and with plurals (see Borer 2005, Bale & Khanjian 2008). If classifiers are analyzed cross-linguistically as an expression that creates a singular on the basis of a number neutral expression, optional classifiers are potentially problematic (cf. also Doetjes 2012). One might say that the optionality of the classifier reflects the presence of two types of numerals: one that selects a singular expression (or CL-N) and one that selects a number neutral expression (N). The alternative would be to assume that the language has empty classifiers, which create the set of atoms needed by the numeral without a visible reflection of this meaning change.

This brings us to the second type of example that needs to be considered in this context. As we indicated above, expressions such as hěn duō / hou² do¹ ‘a lot’ allow for insertion of a classifier. Consider again the data in (9), repeated in (25):

(25) a. hěn duō (běn) shū dōu zài tā-de zuǒzì shàng
   very many CL book DOU at he-DE table top
   ‘Many books are on his table.’

b. hou² do¹ (bun²) syu¹ dōu¹ hai² keoi³-ge³ toi² seong⁶min⁶
   good many CL book DOU at he-GE table top
   ‘Many books are on his table.’

Cross-linguistically this type of modifiers usually combines with expressions that have cumulative reference, such as mass nouns, plurals and number neutral nouns and typically not with singulars (cf. Doetjes 1997). Again, the optional presence of the classifier makes the problem even more interesting in light of the hypothesis that classifiers create a singular denotation. It implies that hěn duō / hou² do¹ ‘a
lot’ should be ambiguous in the sense that it either selects a singular or a number-neutral expression. It is quite unlikely that when the classifier is not present in ), the sentences contain empty classifiers. In the first place, if generalized empty classifiers existed in Mandarin and Cantonese, we would expect them to show up with numerals as well, contrary to fact. Moreover, as already indicated above, there are dialects of Mandarin that do not allow for insertion of the classifier, which shows that hěn duō in these dialects combines only with number-neutral expressions.

On the other hand, there do exist expressions meaning a lot which are restricted to singulars. An example is Dutch menig which has an interpretation similar to that of ‘many a’ in many a boy. Contrary to the standard Dutch high degree modifier veel ‘a lot’, menig only triggers distributive readings (cf. again many a boy):

(26) a. Menig studenti had zijn/*hun i huiswerk te laat ingeleverd.
   many-a student had-sg his/their homework too late handed-in
   ‘Many students handed in their homework too late’

b. *Menig student was samengekomen voor de demonstratie.
   many-a student was gathered for the demonstration
   ‘Many students came together for the demonstration’

c. Veel studenteni hadden hun/*zijn i huiswerk te laat ingeleverd.
   many students had-pl their/his homework too late handed-in
   ‘Many students handed in their homework too late.’

d. Veel studenten waren samengekomen voor de demonstratie.
   many students were gathered for the demonstration
   ‘Many students came together for the demonstration.’

Whereas menig student binds a singular pronoun, veel studenten only binds a plural pronoun. Collective predicates such as to gather cannot be used with menig, but do occur with veel. This shows that the number properties of the noun affect the availability of a collective interpretation. The collective interpretation is only possible when the noun is plural.

Turning back to Mandarin and Cantonese, one might argue that in as far as these languages allow for optional insertion of the classifier, there are two instances of hěn duō/ hou2 do1 ‘a lot’: one is similar to veel in Dutch, which combines with number neutral nouns, and the other is similar to menig in Dutch, which combines with singulars. If this is on the right track, we predict that insertion of the classifier should block a collective interpretation, on a par with menig, which combines with singulars, as in ). The Cantonese counterparts are given in ):
(27) a. hou² do¹ go³ hok⁶saang¹ dou¹ gau¹-zo² keoi⁵,*(dei⁶) ge³ gong¹fo³
good many cl student dou hand-in-perf he-(pl) ge homework

‘Many students handed in *his/their homework.’

b. *hou² do¹ go³ jan⁴ dou¹ jat'cai⁴ lei⁴
good many cl person dou together come

‘Many people came in together.’

c. hou² do¹ hok⁶saang¹ dou¹ gau¹-zo² keoi⁵-* (dei⁶) ge³ gong¹fo³
good many student dou hand-in-perf he-(pl) ge homework

‘Many students handed in *his/their homework.’

d. hou² do¹ jan⁴ dou¹ jat'cai⁴ lei⁴
good many person dou together come

‘Many people came in together.’

On the one hand, the contrast between a) and c) shows that both in the presence and in the absence of the classifier, the plural pronoun is used. On the other hand, predicates such as to gather seem to be incompatible with the form that includes the classifier (hou² do¹ go³ jan⁴ ‘many people’ in b)) as predicted. The two tests seem to be contradictory at first, but the sentence in ) shows that there is no real variable binding in Cantonese, in the sense that even truly distributive quantifiers do not allow for binding of a singular pronoun:

(28) go³-go³ hok⁶saang¹ dou¹ daai³-zo² keoi⁵-* (dei⁶) ge⁵ gong¹fo³ lei⁴
cl-cl student dou bring-perf he-(pl) ge homework come

‘Every studenti brought hisi/theiri homework.’

This shows that the binding properties cannot serve as diagnostics. Moreover, as the presence of the classifier results in a clearly singularizing reading, the impossibility of b) might not be due to singularity of CL-N but rather to this singularizing effect. Further data have to be investigated in order to see whether one can maintain that hên duō/ hou² do¹ ‘a lot’ are ambiguous between a singular and a number neutral selecting expression.

This section examined the hypothesis that the classifier is responsible for the singular interpretation of CL-N combinations. This hypothesis offers a very simple account of the denotations of bare CL-N’s and bare N’s in Mandarin and Cantonese, as the former have a singular denotation while the latter are number neutral. As shown above, the hypothesis has consequences for the analysis of numerals and of degree expressions such as hên duō/ hou² do¹ ‘a lot’. As insertion of the classifier is optional in case of hên duō/ hou² do¹ ‘a lot’, one has to
assume that these expressions are ambiguous in order to account for the combination with CL-N’s (singular) and N’s (number neutral). There might be evidence in favor of such ambiguity, but more data have to be considered. Notice, in addition, that numerals are always followed by a classifier, but do allow a cumulative reading, as in (29):

(29) jau⁵ ng⁵ go³ jan⁴ jat¹cai⁴ lei⁴
    have five cl person together come
    ‘Five students came in together.’

This shows that the ungrammaticality of b) cannot be attributed to the singularizing effect of the cl alone, but it must be specific to the interpretation of hěn duō/ hou² do¹ ‘a lot’ in combination with a singular expression as opposed to hěn duō/ hou² do¹ followed by a number neutral expression.

3.3 The pragmatic scale hypothesis

A third possibility that we would like to entertain to explain the singular/plural shift of CL+N is based on ‘pragmatic scales’ (Horn 1968, Levinson 1983). Suppose that, contrary to the assumption in Section 3.2, CL-N always denotes a full join semi-lattice (minus the empty set). This denotation of CL-N will thus contain both atoms and proper pluralities, regardless of whether the ClP is preceded by a numeral or bare. The exception is a ClP introduced by the plural classifier, CLPLUR N. Below we restrict the discussion to Cantonese since Mandarin CL-NS only appear postverbally, and are more easily accounted for under the empty numeral hypothesis.

As we have seen, the Cantonese CLPLUR seems to incorporate a measure expression, which excludes a singular meaning for count Ns. Thus, its denotation will be a proper plurality (much as the denotation Chierchia 1998 assigned to plural nouns). With three books, a, b, and c, book’ = {a, b, c} and PL the closure under sum (+), we have:

(30) a. \([_{\text{CLP}} \text{bun}^2 [_{\text{N}} \text{syu}^1]]\) → PL(book’) = {a, b, c, a+b, a+c, b+c, a+b+c}

b. \([_{\text{CLP}} \text{di}^1 [_{\text{N}} \text{syu}^1]]\) → PL(book’) – book’ = {a+b, a+c, b+c, a+b+c}

Note that the denotation of CLPLUR N is a proper subset of the one produced by other CL-NS. Numerals applied to CL-Phrase are interpreted as a function from a full semi-lattice to the subset containing pluralities with the appropriate number of atoms (e.g. exactly 3, in ), cf. Heycock and Zamparelli (2005)).
Given the number neutral interpretation of $\text{CL-N}$ defined in a), the fact that a bare $\text{CL-N}$ receives a ‘singular’ interpretation must be explained. Under the pragmatic hypothesis, the explanation is cast not in terms of denotations, but rather as a result of scalar implicatures: a speaker who intends to utter a bare $\text{CL-N}$ combination can choose between a number-neutral classifier like the one in a) and a plural-only classifier as in b). The two classifiers are in a scale of informativeness: $\text{bun}^2 \text{syu}^1 (\text{CL book})$ could in principle be used in contexts with individual books as well as in contexts with multiple books, while $\text{di}^1 \text{syu}^1 [\text{CLPLUR book}]$ can only be used when multiple books are present; therefore the latter is more informative (i.e. its excludes more contexts). Assuming Grice’s maxim of Quantity, a hearer who hears $\text{bun}^2 \text{syu}^1$ ‘CL book’ should assume that a speaker would not have evidence for the presence of more than one book, for in that case it would have been more informative, and thus more cooperative, to use the plural form, $\text{di}^1 \text{syu}^1$. Thus, the hearer concludes that the speaker intended to convey that there is only one book, since this is the only case not covered by the $\text{di}^1$ form. This derives the pragmatic singularity of $\text{bun}^2 \text{N}$ and all other semantically number-neutral forms.

For this pragmatic explanation to run its course, the forms contrasted in the scale must be in free variation, otherwise the hearer could not infer that the speaker’s choice among them was motivated by the principle of cooperation alone. This assumption is not satisfied when the $\text{CL-N}$ is preceded by a numeral, since numerals from zero up obligatorily select for the number-neutral $\text{CL-N}$ (see section 2). As discussed above, this could be due to the fact that Cantonese $\text{di}^1$ embeds a quantity expression which, like the English words quantity/number (of), is intrinsically vague, and thus not countable b). Note that a) would be perceived as false if only 1 person came, despite the fact that 1 is a number.5

(32) a. A number/quantity of people came.
    b. *Two numbers/quantities of people came.

A pragmatic analysis along these lines can work only under certain conditions, which must be carefully verified. First, one could wonder whether the implicature on $\text{CL-N}$ can be cancelled in downward monotonic contexts. In order to interpret the data in the right way, let us first look at the interpretation of a $\text{N}$ as opposed to exactly $\text{N}$ in English.
In a), the indefinite *a fly* is set in a downward entailing environment (the antecedent of a conditional), and is compatible with a situation in which there are multiple flies in the soup, but this is not so in b). In c), an upward entailing environment, *a fly* conversationally implicates that (as far as the cooperative speaker knows) there are no additional flies in the soup.6

(33) a. If there is a fly in my soup, the waiter will change it.

   *Implies*: if there are two flies, the waiter will still change the soup

b. If there is exactly one fly in my soup, the waiter will change it.

   *Does not imply*: if there are two flies, the waiter will change it.

c. There is a fly in my soup!

   *Conversationally implicates*: There is exactly 1 fly in the soup.

On the basis of this, we do not want to imply that *a fly* is number neutral in English. Rather, contrary to *exactly one fly*, *a fly* allows for a number neutral interpretation. To illustrate this point, it is useful to have a look at a sentence such as *John saw Mary this morning.*

This sentence does not exclude that John saw both Mary and someone else, for instance Sue, simply because the sentence is not exhaustive. Similarly, (33a) does not imply that the fly is the only thing there is in the soup, and as such there is no way to exclude that there are other flies in the soup. For pragmatic reasons, however, one would use a plural if the speaker knows that there is more than one fly in the soup, even though the sentence is perfectly fine in a context where the speaker ignores whether there are more flies in the soup (see also footnote 6).

(34) jyu⁴gwo² jau⁵ zek³ wu²jing¹ hai² wun² tong¹ leoi⁵-min⁶,

   if have CL fly at CL soup in

   go³ fok⁶mou⁶saang¹ wu⁵ wun⁶-zo² keoi⁵

   CL waiter will change-PERF it

‘If there is a fly in the soup, the waiter will change it.’

This shows that the interpretation of CL-N does not allow us to decide between the three theories discussed in this paper, which are all compatible with the data in (34), given that the effect can be attrib-
uted to non-exhaustivity (as one has to assume under the analyses described in the previous sections) or to the underlying number neutral interpretation of CL-N.

In order to pursue the pragmatic analysis, we need to make sure that $di^1 N$ behaves as a proper plural even in downward monotonic environments. As is well known, English or Romance bare plurals do not preserve proper plurality in such environments, while interestingly a number of $N$ in English does (contrast a) with b)).

(35) a. If there are flies in the soup, the waiter will change it.
   Implies: even if there is only one fly, he still will.

   b. If there are a number of flies in the soup, the waiter will change it.
   Does not imply: if there is only one fly, he will.

(36) shows that CLPLUR $N$ in Cantonese behaves more like a number of than as a bare plural: $di^1 N$ preserves its plural meaning even in downward entailing environments. For the pragmatic analysis, this is crucial, as this analysis relies on a competition between the number neutral CL-N on the one hand and a real plural CLPLUR $N$ on the other. As a result of this competition, the number neutral form CL-N is used for the singular in those contexts where it is in competition with the more specific, plural form CLPLUR $N$.

(36) jyu²gwo² jau⁵ $di^1$ wu¹jing¹ hai² wun² tong¹ leoi¹-min⁶,  
   if have CLPLUR fly at CL soup in  
   go³ fok¹mou¹saang¹ wui⁵ wun²-zo² keoi⁵  
   CL waiter will change-perf it  
   ‘If there are (a number of) flies in the soup, the waiter will change it.’
   Does not imply: if there is only one fly, he will.

The second important point for the analysis concerns the comparison between the CL-N and the bare N meaning. Recall that, Cantonese bare Ns appear to be number neutral, just as CL-NS in the analysis discussed in this section. We want to make sure that the pragmatic proposal just outlined does not force bare Ns to a singular meaning. In other words, we do not want bare Ns to be in the same scale as CL-NS and CLPLUR $N$. There are various ways to explain why this should happen, all of which are based on the idea that scalar inferences hold only if the items contrasted are in free variation, i.e., if the hearer can assume that the only reason for the speaker to use one or the other is the meaning aspect which determines their order in the scale, here number. But this is clearly not the case with
Cantonese cl-n{s} and ns: in this language, argumental bare Ns are possible only as non-specific indefinites (in object position, and in episodic sentences, see footnote 1), while bare cl-N can be definites, or (non)-specific indefinites (see Cheng and Sybesma to appear). So there is no reason why number scalar implicatures should apply to the two forms.

As the other accounts discussed above, the present account has pros and cons. An advantage of this account is that it allows us to assume that numerals combine with number neutral forms rather than with singular forms. The assumption that numerals combine with singulars runs against the typological tendency for languages to have numbers apply to plural or number-neutral forms (see section 3.2).

On the other hand, the idea crucially rests on the hypothesis that clplur N is directly comparable to cl-N, but not with the bare noun. Given the syntactic similarities between the former two, this might be reasonable. However, from a semantic point of view, the comparability of the two forms is less straightforward. In particular, in the present account, $di^1$ seems to have a more complex semantic content than other classifiers, which could make it non-comparable. As indicated in section 2 above, $di^1$ is similar to vague quantity terms such as quantity, which is further confirmed by the fact that it is compatible with both mass nouns and count nouns. When used with a mass noun, no plural interpretation obtains (see (20) above). As such, $di^1$ is not a real plural marker, and one might even wonder where the plural interpretation comes from. One option might be to assume that the plural interpretation is due to competition with a ‘singular’ classifier on the one hand and with a number neutral bare noun on the other, which would bring us back to the previous solution. The data in ) above, which suggest that the source of the plural interpretation is not pragmatic but semantic in nature, are hard to interpret, also because of the independent existence of bare number neutral nouns in the language (languages with real plurals seem to have number neutral bare nouns as well, cf. for instance Bale and Khanjian 2008).

As a whole, even though the analysis is attractive in that it explains why cl-N combinations can be found with numerals and with expressions such as hou² do¹ ‘a lot’, we did not find conclusive evidence for this analysis either. In particular, we did not find independent evidence for the number neutral status of cl-N other than its compatibility with numerals and with expressions such as hou² do¹ ‘a lot’. More data will be needed to reach a firm conclusion.
4. Conclusion

In this paper we discussed the interpretation of CL-N in Mandarin and Cantonese. As shown above, there is both evidence for treating CL-NS as singular expressions and for treating them as being number neutral. In the first part of the paper we gave an overview of the relevant data in both Mandarin and Cantonese. In both languages bare CL-NS have a singular interpretation, while CL-NS also occur with numerals and with quantity expressions such as hěn duō/ hou² do¹ ‘a lot’ contexts that typically trigger insertion of a plural expression in number marking languages such as English. In the second part of the paper, we discussed three possible solutions to this puzzle. In the first place, the singular interpretation of the ‘bare’ CL-N could be triggered by insertion of an empty numeral one. This solution might work for Mandarin, but offers problems in Cantonese, where CL-N has a larger distribution and can be used under conditions that are not compatible with the presence of an empty numeral. A second solution assumes that CL-N is singular in nature, which has consequences for the semantic interpretation of numerals and hěn duō/ hou² do¹ ‘a lot’. Alternatively, one might assume that CL-N is number neutral in nature. This allows us to keep a standard type analysis for numerals and expressions such as hěn duō/ hou² do¹ ‘a lot’, but forces an alternative analysis of the singularity of bare CL-N. A possible account for this would be competition with the ‘plural’ classifier xie/ di. The first analysis seems to be the only one that one can show to be not right, at least for Cantonese. As for the other analyses, more data have to be investigated in order to make a principled choice between the two.

Addresses of the Authors

Lisa Lai-Shen Cheng: Department of Linguistics, Leiden University, P.O. Box 9515, 2300 RA Leiden, The Netherlands
< L.L.Cheng@hum.leidenuniv.nl >

Jenny Doetjes: Department of French Language and Culture, Leiden University, P.O. Box 9515, 2300 RA Leiden, The Netherlands
< J.Doetjes@hum.leidenuniv.nl >

Rint Sybesma: Department of China Studies, Leiden University, P.O. Box 9515, 2300 RA Leiden, The Netherlands
< R.P.E. Sybesma@hum.leidenuniv.nl >

Roberto Bamparelli: Università di Trento, Centro Interdipartimentale di Mente e Cervello, Dipartimento di Scienze della Cognizione e della Formazione, Corso Bettini, 31, 38068 Rovereto (TN), Italy
< Roberto.zamparelli@unitn.it >
Notes

1 For reasons of space, in this paper we only consider object-level denotations for nominals, excluding those cases where noun phrases refer to ‘kinds’ (in the sense of Carlson 1977). In Cantonese, a bare noun such as si1zi2 ‘lion(s)’ can be used in preverbal position in contexts such as lions will soon be extinct. This case is compatible with analyzes that see kind nominals as a class of proper names or definite descriptions, but we do not discuss this any further here.

2 Měi ‘every’ can also appear with numerals higher than one, as in (i).

(i) bān-shàng měi wù ge xuéshēng jiù yǒu yi-ge qù-guò zhòngguó
class-up every 5 cl student then have one-cl go-exp China

‘In the class, one in 5 has been to China.’

3 Di1 is reminiscent of the distribution, if not the meaning, of Italian molt-AGR ‘much/many’, a vague amount expression which combines with mass nouns in the singular and with count nouns in the plural.

4 The table gives the definite and indefinite status of nominals, which will not be discussed in this paper (see Cheng & Sybesma 1999, 2012, Cheng, Heycock and Zamparelli forthcoming). It does not, however, give the readings available in ‘characterizing sentences’ (see footnote 1).

5 Of course, English quantity has another, non-grammaticalized reading in which it refers to a specified amount, or to quantities along different dimensions. These can of course be counted: Three quantities are used to describe a wave: amplitude, speed, and wavelength or frequency.

6 The strength of the conversational implicature also depends on whether the situation is one in which numbers matter, or can be easily established. So There is a fly in the room does not seem to grant the conclusion that there is just one fly in the room with the same force as c) (flies are easier to spot in soups than in rooms; two flies are a bigger scandal in a soup than in a room, etc.).

Bibliographical References


Cheng Lisa L.-S., Caroline Heycock & Roberto Zamparelli Forthcoming. Definitely Predicative?


HEYCOCK Caroline & Roberto ZAMPARELLI 2003. Coordinated Bare Definites, Linguistic Inquiry 34. 443-469.


JACOB Judith 1965. Notes on the numerals and numerical coefficients in Old, Middle and Modern Khmer. Lingua 15. 143-162.


