Complicating the Oversimplification:
Chinese Numeral Classifiers and True Measures[1]

Chak-Lam Colum Yip
University of Washington

This paper provides evidence that true measures and sortal/mensural classifiers come with different morphosyntactic features, which result in variations in surface syntax. It is argued that each true measure has an [u individual] feature which gets valued by the [i individual] feature in the Num head. Although true measures occupy the classifier head, they cannot appear sentence-initially without a preceding numeral, and they do not individuate and mark definiteness. The case of true measures therefore serves as an argument against the ClP analysis put forth by Cheng and Sybesma (1999), which assumes that all numeral classifiers can perform the functions of D.

1. Introduction

The noun phrase structure of Chinese has been studied quite extensively in recent years. Previous works (Cheng and Sybesma 1999, 2005, Li 1998, Tang 1990, 2005, 2007, Simpson 2005, among others) generally agree that a classifier is hosted by its own functional head which selects an NP complement. In Chinese, numeral classifiers can be split into two types: sortal and mensural classifiers. A sortal classifier as in (1a) 'individuates whatever it refers to in terms of the kind of entity that it is; a mensural classifier as in (1b) 'individuates in terms of quantity'. (Lyons 1977: 463)

Mandarin
(1) a. *yi* **tiao** *xiangjiao*
     one CL banana
     'one banana'
  b. *yi* **tong** *shui*
     one CL water
     'one bucket of water'

---

[1] This paper was presented at the 20th NACCL under the title of "Complicating the oversimplification: Chinese numeral classifiers and measure words". Since the term "measure word" has been used loosely in the literature, I decided to follow Chao (1968) in referring to words like *gongjin* and *mi* as 'true measures', a type of classifiers as distinct from sortal and mensural classifiers.

However, Aikhenvald (2000: 115) has observed that almost all languages (with or without numeral classifiers) have quantifying expressions, for example:

(2) English: *Three **stacks** of books*   (Aikhenvald 2000: 115)
    Hungarian: *egy **csepp** méz* 'one <u>drop</u> of honey'   (Aikhenvald 2000: 115)
    Vietnamese: *một **cân** ru'ō'i cho* 'one and a half <u>pounds</u> of dog (meat)'   (Löbel 2000)

In a classifier language, quantifying expressions may also behave differently from other numeral classifiers, as shown in the Vietnamese data in (3). The numeral classifier in (3b) appears in the slot adjacent to the noun, whereas the quantifying expression *cân* in (3a) does not.

Vietnamese
(3) a. một cân   ru'ō'i cho
        one pound half.of dog
        'one and a half pounds of dog (meat)'
     b. một con   cho ru'ō'i
        one CL:animal dog half
        'one and a half dogs'   (Löbel 2000)

With this in mind, I therefore propose that quantifying expressions in Chinese also exhibit syntactic patterns not found in sortal and mensural classifier constructions. Following the terminology developed in Chao (1968), I will henceforth use the term 'true measures' to refer to quantifying expressions in Chinese. For our purposes, I define 'true measures' as words which represent a unit of measure like dimensions (weight, height, and length), length of time, etc. Examples of true measures include *jin* 'catty', *gongjin* 'kilogram', *mi* 'meter', *limi* 'centimeter', etc.

2. Syntactic Differences Between True Measures and Sortal/Mensural Classifiers
Despite the surface similarities, true measures and sortal/mensural classifiers differ in several ways. First, although a classifier-noun sequence can occur in a post-verbal position as in (4a), a true measure-noun sequence gives rise to ungrammaticality as exemplified in (4b). However, (4b) can be fixed by inserting a numeral before the true measure, resulting in the grammatical (4c).[2]

---

[2] I use the following abbreviations in this paper: CL=Classifier, TM=True Measure, Dem=Demonstrative, NEG=Negative, DE=de(modifying marker), PERF=Perfective, N=Noun, V=Verb.

Mandarin
(4) a. wo xiang mai ba   dao.
      I   want buy CL knife
      'I want to buy a knife'.
  b. wo xiang zou   *li   lu.
      I   want walk TM road
      'I want to walk a mile.'
  c. wo xiang zou   yi li   lu.
      I   want walk one TM road.
      'I want to walk one mile'

Second, it is well-known in the literature that Cantonese classifiers mark definiteness[3] (see Cheng and Sybesma 1998). (5a) shows a classifier-noun sequence in the sentence initial position. In Cantonese, when the classifier occurs in this position, it must be definite in reference. However, when a true measure appears in the sentence-initial position in Cantonese as in (5b), the sentence is not licensed.

Cantonese
(5) a. bui1 caa4 hou2 jit6
      CL tea   very hot
      'The tea is very hot'
  b. *cek3 dei6 hou2 gwai3
       TM land very expensive
      'The square feet of land is very expensive'

Third, classifiers can appear directly after a demonstrative and the insertion of a numeral between the demonstrative and the classifier is completely optional (see 6a). However, in (6b) we see that the Dem-TM sequence is not licensed by the grammar. Again, the ungrammatical example can be saved by adding a numeral before the true measure, as in (6c).

---

[3] Cheng and Sybesma (1998) argue that the functional head Cl is determiner-like and a D layer needs not be projected. However, in this work I follow the Cl-to-D movement analysis put forth in Simpson (2005). According to Simpson, Cantonese classifiers can be [+definite] because Cl moves to D in this language. I feel that this analysis better fits the standard assumption that nominal arguments across languages should be DPs. Please see Simpson (2005) for more arguments against the ClP proposal (Cheng and Sybesma 1998).

Mandarin

(6) a. na   (san) ba dao hen fengli
        Dem   (three) CL knife very sharp
        'That knife is very sharp.' (without numeral)
        'Those three knives are very sharp.' (with numeral)

   b. *na li   lu   hen nan   zou
        Dem TM road very difficult walk
        'That mile is hard to walk.'
   c. na   san li   lu   hen nan   zou
        Dem three TM road very difficult walk
        'Those three miles are hard to walk.'

Lastly, most classifiers in Chinese can reduplicate to create the "each/every" meaning, as in (7a). Since the function of classifiers is to individuate, it can be argued that the reduplication intensifies the individuation. In (7b) we see that the reduplication of a true measure immediately renders the sentence ungrammatical.[4]

Mandarin

(7) a. zheli de dao,   ba ba dou hen fengli.
        Here DE knife   CL CL dou very sharp
        'Every one of the knives here is very sharp.'
   b. zhe chengshi de lu,   *li   li   dou hen pingtan.
        Dem   city   DE road   TM TM dou very flat
        'Every mile of the road in this city is very flat.'

To sum up, we have shown that true measures do not appear post-verbally, do not mark definiteness, cannot appear directly after a demonstrative, and do not reduplicate.

3. The Syntax of True Measures
3.1 Numeral-True Measure Sequences as Adjuncts?
If we assume that true measures are not numeral classifiers and they appear in a different position in the tree, we have to decide what the syntactic relationship between the true

---

[4] In fact, I hold a more radical view in assuming that all sortal/mensural classifiers can reduplicate. However, reduplication of classifiers that are not actively used in daily speech creates odd-sounding expressions and is therefore avoided. For example, classifiers only used in idioms or literal, old-style, bookish expressions cannot be reduplicated.
(i) yi   xian xiwang
    one   ray hope
    'A ray of hope.'
(ii)* xian xian xiwang

measure and the noun may be. As observed in Tang (1996), the numeral-true measure sequence sometimes appears in the modifier position. Compare (8a) with (8b):

Mandarin
(8) a. san   gongjin   rou
       three TM   meat
       'Three kilogram of meat'
    b. san   gongjin de rou
       three TM   DE meat
       'Three kilogram of meat'
    c. yi bao san   gongjin*(de) rou
       one CL three   TM   DE meat
       'One pack of meat that weighs 3kg.'

When *rou* is only preceded by *san gongjin*, both (8a) and (8b) are possible. However, *de* becomes obligatory when the classifier *bao* selects its complement. In that case, only (8b) can be the complement of *bao*. One may postulate that all numeral-true measure sequences are always adjoined to NPs, whereas classifiers are always heads. However, topicalization tests (9b-c) show that *yi gongjin* in a nominal like (9a) cannot be an adjunct:

Mandarin
(9) a. wo yao yi gongjin   mi
       I   want one kilogram rice
       'I want one kilogram of rice'
    b. mi   wo yao yi   gongjin pro.
        rice I   want one kilogram
        'Rice, I want one kilogram.'
    c. mi   wo yao yi   gongjin   de pro.
        rice I   want one kilogram DE
        'Rice, I want one kilogram./Rice, I want the kind that weighs one kilogram'

Out of (9b-c), only (9b) is the logical output of topicalization. Even though (9c) is perfectly grammatical, it creates a second reading which means that there exists a kind of rice that weighs one kilogram and the speaker only wants that particular kind. Furthermore, all modifiers (*de* or *de*-less[5]) **must** occur with *de* in sentence-final position.

---

[5] Paul (2004) argues that de-less modification is possible at the level of the syntax. Please refer to the paper for more on the matter.

Mandarin
(10) a. wo yao   mu   zhuozi!
          I   want wooden table
        'I want a wooden table!'
    b. zhuozi, wo yao   mu   de!
        Table   I   want wooden DE
        'Table, I want a wooden one!'

Since (9b), the logical output of topicalization, does not contain a sentence-final *de*, the possibility that numeral-true measure sequences are adjuncts appears to be dwindling. Instead, the fact that in (8c) *bao* only selects (8b) and not (8a) as its complement suggests that *bao* in (8c) and *gongjin* in (8a) are both numeral classifiers. However, since the syntax only allows one classifier within a DP, the grammar does not license (8a) as the complement of *bao*. Therefore, it does seem that true measures and sortal/mensural classifiers occupy the same position in the tree.

3.2 Towards a Solution
The differences between true measures and sortal/mensural classifiers can be summed up in the following chart:

(11)

	Sortal/Mensural classifiers (CL)	**True Measure (TM)**
a.	**V-CL-N** **V-Num-CL-N** wo xiang mai (yi) ba dao. I want buy (one) CL knife 'I want to buy (one)/a knife'.	***V-TM-N** **V-Num-TM-N** wo xiang zou *li lu. I want walk TM road 'I want to walk a mile.'
b.	**CL-N** bui1 caa4 hou2 jit6 CL tea very hot 'The tea is very hot'	***TM-N** *cek3 dei6 hou2 gwai3 (Cantonese) TM land very expensive 'The square feet of land is very expensive'
c.	**Dem-CL-N** **Dem-Num-CL-N** na (san) ba dao hen fengli Dem (three) CL knife very sharp 'That knife is very sharp.' or 'Those three knives are very sharp.'	***Dem-TM-N** **Dem-Num-TM-N** na (yi) li lu hen nan zou Dem (one)TM road very difficult walk 'That (one) mile is hard to walk.'

d.	**CL-CL**	***TM-TM**
	zheli de dao, ba ba dou hen	zhe chengshi de lu, *li li dou
	Here DE knife CL CL dou very	Dem city DE road TM TM dou
	fengli.	hen pingtan.
	Sharp	very flat
	'Every one of the knives here is	'Every mile of the road in this city is very
	very sharp.'	flat.'

From (11), it seems that true measures have to be preceded by numerals, while Cl-NP sequences can occur alone. Nonetheless, this assumption too is not true when we run into an example like the following:

Mandarin
(12) mei   gongjin   rou dou   hen gui
    Every kilogram meat DOU very expensive
    'Every kilogram of meat is very expensive.'

If the function of classifiers is to individuate, then one possible explanation is that true measures do not individuate in our cognitive system. This explanation seems quite logical since a classifier like *li* in *san li tang* 'three candies' picks out three individuated candies out of the world of possible candies, while a true measure like *mi* in *san mi lu* 'three miles' assigns the measurement the numeral specifies to the noun. Furthermore, notice that *san mi lu* 'three miles' refer to a three-mile-long road, not three individuated roads which are one-mile-long each.

At this stage, there is one important issue we have to consider given the syntactic differences of true measures and sortal/mensural classifiers. If true measures are a distinct type of classifiers, what modifications are needed in order for the grammar to fit them into the Cl head but still produce the surface differences mentioned above? The solution can be a quick fix. Since true measures and sortal/mensural classifiers occupy the same position in the tree, the differences must lie in their morphosyntactic features when they enter into the numeration.

I propose that true measures come with an [u individual] feature which needs to be valued. The corresponding [i individual] feature can be found in quantifiers like *mei* 'every/each' or numerals. Failure to value the [u individual] feature results in ungrammaticality, as shown in (13):

(13)

```
                    DP
                   /  \
                  D    NumP
                      /    \
                    Num     ClP
                     |      /  \
                    san   Cl    NP
              [i individual]     |
                  [u Cl]         |
                        |        |
                     gongjin     |
                      [i Cl]     mi
                  [u individual]
```

When the Num head is merged with ClP, the [u Cl] feature in the Num head looks for a valued version of the same feature in its c-command domain. Since [i Cl] can be found in the Cl head, the [u Cl] feature in Num gets valued. While the valuation of [u Cl] is taking place, the [u individual] feature in the Cl head gets valued by the [i individual] feature in the Num head concomitantly. The feature structure above correctly rules out the ungrammatical data in the (b) examples in (4-7), repeated below. Whenever a numeral or a quantifier is not present, the [u individual] feature in the true measure will not be valued, and the sentence becomes ungrammatical.

(14) (i) A sortal classifier 'individuates whatever it refers to in terms of the kind of entity
      that it is.' (Lyons 1977: 463)
  (ii) A mensural classifier 'individuates in terms of quantity'. (Lyons 1977: 463)
  (iii)A true measure does not individuate and only assigns to the noun the
      measurement that the numeral specifies.

Mandarin
(4)b. wo xiang zou　*li　　　　　lu.
　　　　　　　　　　[u individual]
　　I　want　walk　TM　　　　road
　　'I want to walk a mile.'

Cantonese
(5)b. *cek3　　　　dei6　hou2 gwai3
　　　[u individual]
　　　TM　　　　land　very　expensive
　　　'The square feet of land is very expensive'

292

Mandarin
(6)b. na   *li   lu   hen nan   zou
          [u individual]
      Dem TM   road very difficult walk
      'That one mile is hard to walk.'

Mandarin
(7)b. zhe chengshi de lu,   *li   li   dou hen pingtan.
                                        [u individual] [u individual]
        Dem   city   DE road TM   TM dou   very flat
        'Every mile of the road in this city is very flat.'

Since there is not a matching [i individual] feature for the [u individual] feature in each example, the derivation crashes.

4. Implications

Our analysis of true measures provides arguments against Cheng and Sybesma (1999)'s ClP hypothesis. One of the major reasons why a ClP (as opposed to a DP) is proposed in Cheng and Sybesma (1999) is due to the existence of the following Cantonese construction in (14):

Cantonese
(15) bui1 caa4 hou2 jit6
      CL   tea   very hot
      'The tea is very hot'

Cheng and Sybesma (1999) argue that an example like (14) shows that classifiers in Chinese denote definiteness, individuation, and number, which are the functions of determiners in languages with overt Ds like English. Therefore, following their train of thought, argument nominals in Chinese should be represented as ClPs. However, the ClP analysis does not predict the ungrammaticality of (15):

Cantonese
(16) *sing1 seoi2 hou2 cung5
      TM   water very heavy
      '*The liter of water is very heavy.'

If our analysis is on the right track, true measures are numeral classifiers that cannot appear in the sentence-initial position. The derivation crashes as the [u individual] feature in true measures fails to be valued. Therefore, we see that not all classifiers can be in sentence-initial position and perform the functions of D, as suggested in Cheng and Sybesma (1999). Hence, our analysis gives fairly solid arguments against the ClP analysis.

5. Conclusion

This paper provides evidence that true measures and sortal/mensural classifiers come with different morphosyntactic features which result in variations in surface syntax. It is argued that each true measure has an [u individual] feature which gets valued by the [i individual] feature in the Num head. Although true measures occupy the classifier head, they cannot appear sentence-initially without a preceding numeral, and they do not individuate and mark definiteness. The case of true measures therefore serves as an argument against the ClP analysis put forth by Cheng and Sybesma (1999), which assumes that all numeral classifiers can perform the functions of D.

Appendix A: Dialectal Differences

In this paper, I argue that true measures cannot appear sentence-initially and be definite in reading. Some counterexamples, however, can be found in Cantonese.

Cantonese
(1) bong6   min6baau1 faat3 zo2   mou1
   TM? CL?   bread   grow PERF mold
   'The loaf of bread is molded
   Literal: '*The pound of bread is molded'

In here, *bong6* is both sentence-inital and definite. The literal meaning of the sentence should be 'the pound of bread is molded'. However, in Hong Kong Cantonese, *bong6* when used with *min6baau1* 'bread' doesn't necessarily mean 'pound'. The person who says (1) might not even know the actual weight of the bread. *bong6* 'pound' in (1) is used so excessively in Cantonese that now it can mean 'bag' or 'loaf' when used with bread in daily speech. Frequent usage sometimes bleaches its true measure status and turns it into a classifier. Thus, *bong6* can have two copies in the lexicon in Cantonese – a mensural classifier copy and a true measure copy.

Cantonese
(2) bong6 bun3 juk6 m4   gau3   so2jau5 jan4 sik6
    TM   half   meat NEG enough every   one eat
  'A pound and a half of meat is not enough for everyone to eat'

In (2), *bong6* appears sentence-initially but it is not definite. I argue that *bong6* in (2) is still a true measure. It has been observed across languages that the numeral 'one' always behaves differently from other numerals. Here, the numeral 'one' is probably deleted at PF and the cause of this deletion is left for further research. However, this explanation is warranted because *bong6* in (2) is not [+definite], as we would expect. Following Simpson (2005), Num-Cl-N sequences are indefinite because the Num head blocks the movement from the Cl head to the D head (Head Movement Constraint). I therefore believe that the movement from Cl-to-D is still blocked in (2) and the deletion of the numeral 'one' only happens later at PF.

REFERENCES

Aikhenvald, A. Y. 2000. *Classifiers*. New York: Oxford University Press.

Bowers, J. 1993. The syntax of predication. *Linguistic Inquiry* 24: 591-656.

Chao, Y-R. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.

Cheng, Lisa & Rint Sybesma. 2005. Classifiers in Four Varieties of Chinese. In *Handbook of Comparative Syntax*. Ed. By Cinque, G. and R. S. Kayne (eds.) Oxford University Press.

Cheng, Lisa & Rint Sybesma. 1999. Bare and Not-So-Bare Nouns and the Structure of NP. *Linguistic Inquiry* 4, 509-542.

Chierchia, Gennaro. 1998. Reference to kind across languages. *Natural Language Semantics* 6.339–405.

Li, Yen-Hui Audrey. 1998. Argument Determiner Phrases and Number Phrases. *Linguistic Inquiry* 29:693-702.

Löbel, E. 2000, *Classifiers versus genders and noun classes: A case study in Vietnamese*, in B. Unterbeck, M. Rissanen, T. Nevalainen and M. Saari (eds.), *Gender in Grammar and Cognition*, Berlin – New York, Mouton de Gruyter, 259–319.

Lyons, J. 1977. Semantics. 2vols. Cambridge: Cambridge University Press.

Simpson, Andrew. 2005. Classifiers and DP structure in southeast Asian languages. In *Handbook of Comparative Syntax*, Ed. By Cinque, G. and R. S. Kayne (eds.) Oxford University Press.

Tang, Jane C-C. 1990. *Chinese Phrase Structure and the Extended X' Theory*. Doctoral Dissertation. Cornell University.

Tang, Jane C.-C. 1996. Ta maile bi shizhi and Chinese Phrase Structure. *Bulletin of the Institute of History and Philology* 67.3:445-502.

Tang, Jane C.-C. 2005. Nouns or Classifiers: A Non-movement Analysis of Classifiers in Chinese. *Language and Linguistics* 6.3:431 – 472.

Waltraud, Paul. 2004. Adjectival modification in Mandarin Chinese and related issues. *Linguistics* 43.4:757-793.