

yi-wan tang, yi-ge Tang: Classifiers and massifiers*

Lisa L.-S. Cheng, UC Irvine
Rint Sybesma, HIL/Leiden University

1. Introduction

Nouns can be divided into so-called count nouns and mass nouns. Count nouns refer to entities which present themselves naturally in discrete, countable units, while mass nouns refer to substances which do not present themselves in such units. In languages like English, count nouns can be counted by putting the numeral directly in front of the noun (*two books*) but mass nouns can only be counted with the help of a so-called measure word: *a bottle of soy bean milk, two loaves of whole grain bread, every grain of sand*.¹

In Chinese, all nouns are like mass nouns in the sense that, in order for a noun to be countable, a measure word or classifier is always necessary, including for those whose counterparts in English would be count nouns:

- (1) a. san ping jiu
three CL-bottle liquor
'three bottles of liquor'
b. san ba mi
three CL-handful rice
'three handfuls of rice'
c. san wan tang
three CL-bowl soup
'three bowls of soup'
- (2) a. san ge ren
three CL people
'three persons'
b. san zhi bi
three CL pen
'three pens'
c. san ben shu
three CL book
'three books'

The structure of all these phrases is the same: in Chinese, just as it is necessary to say 'three units (of) liquor', we must say 'three units (of) people'. It is precisely the obligatory presence of classifiers in all cases which has led some scholars to claim that all Chinese nouns *are* mass nouns (see Hansen 1972 for such a claim on classical Chinese, and Graham 1989 for more discussion; for recent claims to this

effect for Modern Chinese, see Chierchia 1995 and Krifka 1995).

Even so, there are sets of data which indicate that the cognitively real difference between things in the world that present themselves naturally in discrete, countable units and those which do not is grammatically encoded in Chinese as well.² These data are presented in section 2. One of our conclusions in this paper is that the count-mass distinction *is* indeed relevant in Chinese grammar. The central question of this paper, then, is: if it is true that the count-mass distinction is encoded in Chinese, where is it encoded and is the underlying representation of the phrases in (1) the same as that of those in (2)?

Before investigating the central question, however, we first (in section 2) present the data alluded to above, which make clear that Chinese 'mass nouns' behave differently from 'count nouns' (the terms will be made more precise below) in that 'mass nouns' can undergo certain processes which 'count nouns' cannot. As a second step, we concentrate on the 'mass nouns' and analyze the processes which 'mass nouns', but not 'count nouns', may undergo (section 3). In section 4, we turn to the main question regarding the differences between 'mass nouns' and 'count nouns'.

2. Data reflecting count-mass distinction

2.1. The 'optional' presence of *de*

In this section we present the data which suggest that there is a mass-count distinction in the grammar of the Chinese noun phrase. The data concern the occurrence of the element *de* between the classifier and the noun and the possibility for the classifier to be modified by an adjective. We will review the *de*-data first.

Tang (1990) notes that the modification marker *de* can optionally appear between a classifier and a noun. This fact is also mentioned in Chao (1968) and more extensively discussed in Paris (1981). Consider the examples shown in (3) (from Tang 1990:408, ex. (17a,b)).

- (3) a. san bang (de) rou
three CL-pound DE meat
'three pounds of meat'
b. liang xiang (de) shu
two CL-box DE book
'two boxes of books'

However, as is noted by Chao (1986:555) and Paris (1981:32ff), it is not the case that *de* can appear with every classifier-noun combination. Some illegitimate cases are shown below (from Sybesma 1992:105-106, ex. 96a-c).

- (4) a. ba tou (*de) niu
eight CL-head DE cow
'eight cows'
b. jiu gen (*de) weiba
nine cl DE tail
'nine tails'
c. shi zhang (*de) zhuozi

ten CL DE table
'ten tables'

It is noted that there is a difference between what we may call 'individual' and 'non-individual' classifiers (generalizing the terminology used by Chao 1968:509) in that the non-individual classifiers allow the presence of *de* while the individual ones do not. Are there other differences between individual and non-individual classifiers?

It has been argued (by Tai and Wang 1990 and Croft 1994 among others) that classifiers can roughly be divided into two groups: classifiers that *create* a unit of measure, and those that simply name the unit in which the entities denoted by the noun come naturally. Thus, in (1), *ping* 'bottle', *ba* 'handful' and *wan* 'bowl' create units by which the amount of liquor, rice and soup is measured, but liquor and rice and soup do not come naturally in bottles, handfuls or bowls: they can come in glasses, liters, kilos, cups, portions, etc. For the nouns in (2) and (4) this seems different: for instance, there is a natural unit to count books: books come in separate volumes. Similarly, notions like people, pens, cows, tails and tables provide natural units by which they can be counted: single pens, single persons, etc. The classifiers like *ge*, *zhi*, *ben* and others don't create any unit of measurement, they just name them.³

The latter group of classifiers can be called 'individual' because they single out one countable discrete unit. The former group of measure creating classifiers are the non-individual ones: although they make the noun countable, they do not pick out discrete units. In other words, with non-individual classifiers, even though there is a classifier, we are still left with a mass. For ease of reference, we will refer to the individual classifiers as 'count-classifiers' and to the non-individual classifiers as 'mass-classifiers' or 'massifiers'.

We will discuss these matters more extensively below. For now, all we want to have noted is that there is a difference between count-classifiers and massifiers when it comes to possibility of having *de*. The fact that it is possible with massifiers and not with count-classifiers suggests that the mass-count difference plays a role in the grammar of the Chinese noun phrase.

2.2. The presence of adjectives

The second cluster of facts which indicates that there is a mass-count distinction in Chinese concerns the possible presence of an adjective immediately preceding the classifier. Tang (1990) notes that certain adjectives can occur between the numeral and the classifier, as in (5) (from Tang 1990:418, (32a,b)).

- (5) a. yi da zhang zhi
one big CL-sheet paper
'one large sheet of paper'
b. na yi xiao xiang shu
that one small CL-box book
'that one small box of books'

Tang (1990) observes that these adjectives differ from regular adjectival modification in that they cannot be accompanied by the modificational marker *de*, as shown in (6) (taken from Tang 1990:419, ex. (33a,b)).

- (6) a. **yi da-de zhang zhi*
 one big-DE CL-sheet paper
 b. **na yi xiao-de xiang shu*
 that one small-DE CL-box book

In addition, it must be noted that only a very limited number of adjectives can occur in this context; it seems that *da* 'big' and *xiao* 'small' are the only really productive ones; the others are: *man* 'full', *zheng* 'whole', *chang* 'long', *hou* 'thick' and *bao* 'thin' (Wu 1991:85; Lu 1987). Furthermore, it must be observed that the use of the adjective-cum-classifier is severely limited by the numeral used: although they are always good with *yi* 'one', numerals higher than one often lead to ungrammaticality. It has been noted (by Paris and by Chao, among others) that the numeral *yi* 'one' in phrases like (5) means 'whole' rather than just 'one'. This corresponds to what Lu (1987:68) observes in a different context, viz., the construction involving a modified classifier always have a flavor of "exaggeration".

What is important to us, however, is the fact that the presence of these adjectives is sensitive to the kind of classifier that is being used. Consider the examples in (7), to be compared to those in (5).

- (7) a. **yi da zhi gou*
 one big CL dog
 b. **yi da wei laoshi*
 one big CL teacher

As shown in (7), when we use count-classifiers such as *zhi* 'classifier for dogs' and *wei* 'honorific classifier for people', the presence of *da* 'big' leads to ungrammaticality, while massifiers like *xiang* 'box' can be combined with *da* 'big'. Lu lists 117 cases of classifiers which can be combined with an adjective and the vast majority is to be classified as massifiers.⁴ It thus appears that the occurrence of an adjective in front of a classifier is sensitive to the kind of classifier involved: while count-classifiers cannot be modified by an adjective, massifiers can.

2.3. Noun vs. Classifier

We mentioned above that, while massifiers *create* a measure for counting, count-classifiers simply *name* the unit in which the entity denoted by the noun it precedes naturally presents itself. This acknowledges the cognitive fact that some things in the world present themselves in such discrete units, while others don't. In languages like English, the cognitive mass-count distinction is grammatically encoded at the level of the noun (**one water*, *one woman*), in Chinese the distinction seems to be grammatically encoded at the level of the classifier. In any case, the differences mentioned above between count-classifiers and massifiers crucially depend on the type of classifier used, not the type of noun. This can be shown by the following facts. While the examples in (8) with *ren* 'person/people' in combination with a count-classifier involving *de* or a modifier are out, the examples in (9), with a massifier, are in.

- (8) a. *san ge de ren
 three CL DE people
 b. *san da ge ren
 three big CL people
- (9) a. yi qun de ren
 one CL-crowd DE people
 'a crowd of people'
 b. yi da qun ren
 one big CL-crowd people
 'a big crowd of people'

In addition, as these facts suggest, just like in English, the massifiers can create measures for counting with both mass and count nouns: *a group of women*. Note, however, that in English the noun in such cases takes a plural form and it is well known that bare plurals and mass nouns have a lot in common (see Chierchia 1995 for references). In Chinese, there being no Number morphology, the noun does not change in form: *yi ge ren* /one CL person/ 'a person' and *yi qun ren* /one group person/ 'a group of people'.

Furthermore, the two phenomena mentioned in 2.1 and 2.2 (i.e. the appearance of *de* between a classifier and a noun and that of an adjective such as *da* 'big' and *xiao* 'small' between a numeral and a classifier) can be combined. The sentences in (10) show that when we have the right classifier, it is possible to have both *de* and an adjective appearing in the numeral-classifier-noun sequence.

- (10) a. san da wan de tang
 three big CL-bowl DE soup
 'enough soup for three big bowls'
 b. liang da bei de shui
 two big cup DE water
 'enough water for two big cups'

In section 3, we investigate the massifiers with and without *de*. The analysis developed there constitutes the groundwork for section 4, where we first look into the question as to why massifiers and count-classifiers differ when it comes to *de* (i.e., the facts in 2.1), after which we direct our attention to the problem of whether the underlying structure of phrases like those in (1) and (2) differ according to whether the classifier involved is a massifier or a count-classifier. There we also explain the modification facts in 2.2.

3. Massifiers with *de* and massifiers without

3.1. Asymmetries

Given noun phrases of the form Num(eral)-Cl(assifier)-*de*-N(oun), the natural question to ask is whether or not such sequence has the same structure as the more typical Num-Cl-N form, without *de*. We present in this section some interpretational differences and grammatical asymmetries between the Num-Cl-N sequence and the one with *de* (some preliminary discussion can be found

in Paris 1981:32ff). These asymmetries suggest that the Num-Cl-N form might have a different structure from the Num-Cl-*de*-N sequence. In other words, these show that *de* is not simply an optional element which may or may not appear in the Num-Cl-N sequence.

A. Interpretational difference. There are important interpretational differences between the sequence with *de* and the one without. Consider the sentences in (11), showing a difference between *san bei jiu* and *san bei de jiu* 'three glasses of liquor' (taken from Sybesma 1992:107, ex. (100a,b); more examples are given there).

- (11) a. #ta yong xiao-wan he-le san bei jiu
 he with small-bowl drink-LE three CL-cup liquor
 'He drank three glasses of liquor from a small bowl'
 b. ta yong xiao-wan he-le san bei(zhi)-de jiu
 he with small-bowl drink-LE three CL-cup-DE liquor
 'He drank three glassfuls of liquor from a small bowl'

The sentence in (11a) is gibberish, as the # is meant to indicate, but (11b) is not. In (11a), when *bei* 'glass' is used without *de*, the default interpretation is that the wine is consumed from the glass: the actual glass is part of the scene. In contrast, when *bei* 'glass' is used with *de*, as in (11b), the wine need not be consumed from the glass; in this case, *bei* 'glass' merely provides a measure for the amount of liquor that was consumed and it denotes something like: enough liquor to fill three glasses.

Let us look at one more pair of examples:

- (12) a. zhuozi-shang you san wan tang
 table-top there-is three CL-bowl soup
 'There are three bowls of soup on the table'
 b. zhuozi-shang you san wan-de tang
 table-top there-is three CL-bowl-DE soup
 'There is enough soup on the table to fill three bowl'

The picture evoked by (12b) is that the soup is all over the table, spilled, there are no bowls. The default interpretation of (12a) on the other hand is that there are three bowls, filled with soup, standing on the table.

As is discussed in Hoekstra (1988), measure words like *wan* 'bowl' have an interesting property: although they can be used in what we would call an 'abstracted' sense, meaning that they are exclusively used in their function as a container, somehow the 'concrete' container remains present as well. So although in sentences like *She drank three glasses of Vita Boost* it is likely that really only one glass was involved, it is less imaginable that no glass was involved and that she drank it right from the bottle. This explains the default interpretation of (11a) and (12a).

The interesting thing about the interpretation of (11b) and (12b) is that the concrete containers do not force themselves into the picture; the whole phrase with *de* seems to be more like a modifier (see Sybesma 1992): enough wine to fill three glasses, and enough soup to fill three bowls, and the actual glasses and bowls do not play a role.

B. Demonstratives. As we see in (13) and (14), the presence of a demonstrative is allowed in the simple Num-Cl-N sequence but not in the sequence involving *de*. If the two have the same structure, it is unclear where the difference stems from.⁵

- (13) a. na san wan tang
 that three CL-bowl soup
 'those three bowls of soup'
 b. *na san wan de tang
 that three CL-bowl DE soup

- (14) a. zhe wu bang rou
 this five CL-pound meat
 'these five pounds of meat'
 b. *zhe wu bang de rou
 this five CL-pound DE meat

C. Relative clauses. It is well-known that in Mandarin Chinese, a relative clause can precede or follow the Num-Cl sequence (with some interpretational difference, see below); this is illustrated in (15a) and (16a). In (15b) and (16b), we see that regardless of the ordering between the relative clause and the Num-Cl sequence, if *de* is involved, we get a much degraded result.

- (15) a. san wan [wo mama zhu de] tang
 three CL-bowl I mother cook DE soup
 'three bowls of soup which my mother cooked'
 b. ??san wan de [wo mama zhu de] tang
 three bowl DE I mother cook DE soup

- (16) a. [wo mama zhu de] san wan tang
 I mother cook DE three bowl soup
 'three bowls of soup which my mother cooked'
 b. ??[wo mama zhu de] san wan de tang
 I mother cook DE three bowl DE soup

In sum, it appears that the Num-Cl-*de*-N sequence differs from the Num-Cl-N sequence in some fundamental respects.

3.2. Analysis

3.2.1. An alternative: the Noun-Num-Cl order

Given the interpretational differences and other asymmetries between a Num-Cl-N sequence and a Num-Cl-*de*-N sequence, *de* cannot be treated simply as an optional element. We propose that the difference between a Num-Cl-N sequence and the one with *de* is structural. We turn to the Num-Cl-N sequence in section 4; now we will argue that the Num-Cl-*de*-N sequence is derived from a subject-predication structure via relativization, i.e., the Num-Cl acts as a modifier of the noun.

The typical order of elements within a noun phrase, as we have seen, is

Num-Cl-N (as in (17)) (in case there is a demonstrative (Dem), it will precede the Num). There is an alternative order which can be used in some contexts. In this order, the noun precedes the Num-Cl sequence; as is shown in (18).

(17) hufei he-le wu wan tang
 Hufei drink-LE five CL-bowl soup
 'Hufei drank five bowls of soup'

(18) hufei he-le tang wu wan
 Hufei drink-LE soup five CL-bowl
 'Hufei drank five bowls of soup'

Tang (1996) argues convincingly that the N-Num-Cl order in (18) is not the result of movement of the noun from the post-classifier position to a pre-numeral position. She proposes that it has a subject-predicate structure, i.e., the noun is the subject and the Num-Cl constitutes the predicate: [N [Num-Cl]]. We will not discuss her analysis in detail; we discuss two aspects of her analysis which are of special relevance to the present discussion.

First, given the predicate status of the Num-Cl sequence, it is predicted that a demonstrative cannot precede the numeral. As discussed in Stowell (1989) among others, a nominal predicate cannot have definite reference. In Mandarin Chinese, the presence of a demonstrative necessarily yields definite reference. As shown in (19), if the Num-Cl sequence has a demonstrative preceding it, it cannot take the order in which the noun precedes the Num-Cl sequence:

(19) a. hufei he-le na wu wan tang
 Hufei drink-LE that five CL-bowl soup
 'Hufei drank those five bowls of soup'
 b. *hufei he-le tang na wu wan
 Hufei drink-LE soup that five CL-bowl
 'Hufei drank those five bowls of soup'

The ungrammaticality of (19b) thus supports the subject-predicate analysis of the N-Num-Cl sequence: *na wu wan* 'those five bowls (of)', being a definite expression, cannot serve as a predicate.

Secondly, the noun phrases in (17) and (18) have different interpretations and different truth conditions. In most cases, the difference is very subtle. Consider first how nominal small clause [*tang wu wan*] 'soup five bowls' in (18) is interpreted. Given a subject-predicate relationship, for the nominal small clause to be true, the subject *tang* 'soup' has to satisfy what is denoted by the nominal predicate *wu wan* 'five bowls'. Since the predicate here is a quantity/measure predicate, what is being denoted is the quantity which is equal to the amount that would fill five bowls. In other words, the nominal small clause can be interpreted roughly as: a certain amount of soup such that it satisfies the quantificational denotation of five bowls. But it is immaterial (or: left unexpressed) whether Hufei in (18) consumed the soup using a bowl. As long as what he drank satisfies the quantity which is equivalent to five bowls, the nominal small clause will be true. For the Num-Cl-N sequence in (17) this is different. The classifier-noun in such a sequence does not involve a subject-predicate relation. Although in certain contexts (especially when the classifier

used is a non-concrete 'container' like 'kilo'), phrases of the type Num-Cl-N can be used to strictly indicate a measure, the default interpretation is much more straightforward: *wu wan tang* is interpreted as 'five bowls of soup', with the actual bowl playing an actual part (see Cheng and Sybesma 1997 for further discussions about the role of the classifier in mass noun contexts).

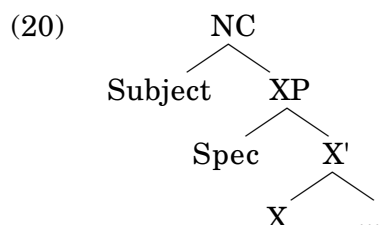
This is reminiscent of the difference between Num-Cl-N and Num-Cl-*de*-N we mentioned in the context of (11). There we also suggested that in the latter case, the case with *de*, the Num-Cl merely denoted a quantity: liquor enough to fill three glasses, while in the former case, the typical Num-Cl-N sequence, the actual glasses were part of the picture too.

We would like to suggest that the similarities are not coincidental. In fact, we claim that the sequence involving *de* is derived from a subject-predicate structure. More particularly, the noun phrase *san wan de tang* 'three bowl *de* soup' is derived from *tang san wan* 'soup three bowl'.

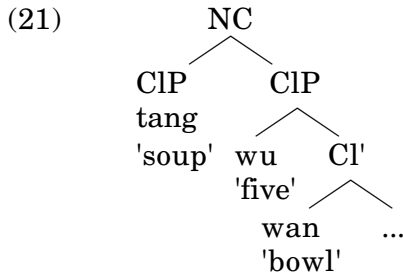
3.2.2. Deriving the Num-Cl-*de*-N sequence

In Mandarin Chinese, *de* appears between a modifier and a noun, such as with possessives, relative clauses and adjectives. Since *de* normally appears with modifiers, it seems likely that the Num-Cl-*de*-N sequence involves modification, with the Num-Cl acting as a modifier of the noun. We have suggested above that the Num-Cl-*de*-N sequence is derived from the subject-predicate structure [N-[Num-Cl]]. We propose that the noun in the Num-Cl-*de*-N sequence is a relativized noun and the Num-Cl-*de* sequence is a relative clause.⁶

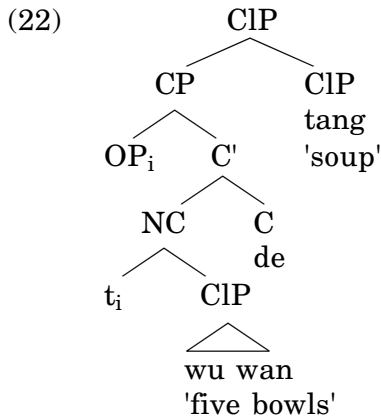
Consider again the subject-predicate relationship in the [N [Num-Cl]] sequence. We follow Stowell (1983, 1989) and others in assuming that the subject-predicate constitutes a nominal small clause. The structure of the nominal small clause is as in (20) (from Stowell 1989:253, ex. (39)).



We thus consider the whole small clause to be a [+N] category. Following Stowell, we use the label NC (Nominal Clause) for the projection of the small clause. Further, we assume, following Tang (1990), that classifiers head their own projection (CIP) and they take NP as their complements. Thus, the structure of the NC of (18) is as in (21).⁷ (See Tang 1996 for a different structure to represent the subject-predicate relation.) (We'll discuss the internal structure of CIP shortly.)⁸



The Num-CI-*de*-N sequence is derived from a structure such as (21) by relativization. The structure is as in (22).



In short, the Num-CI-*de*-N sequence is an instance of subject relativization. The subject of the small clause, in this case, *tang* 'soup', has been relativized.

3.2.3. Explaining the asymmetries

We now show how the subject-predicate structure as well as the relativization analysis account for the asymmetries between massifier structures involving *de* (Num-CI-*de*-N) and those without (Num-CI-N). In section 4, we examine the differences between the 'count nouns' and the 'mass nouns' and see how they follow from our analysis.

A. Interpretational differences. The interpretational difference noted is that in the Num-CI-*de*-N sequences, Num-CI is like a modifier: it denotes the amount of N, while the CI in Num-CI-N forms is more concretely present. In the preceding section we argued that the sequence with *de* is a modificational constituent, a relative clause. In other words, the N in the Num-CI-*de*-N phrase is the head noun, which is modified by a modifier (Num-CI). As we will indicate below, in the Num-CI-N sequence, as in Tang (1990), we consider the classifier head (CI) to be the head of the noun phrase. There is no modification relationship. The interpretational difference therefore follows from a structure difference.

B. Demonstratives. We have shown that when *de* is present in a Num-CI-N string, it is not possible to have a demonstrative. This follows if we derive it from the subject-predicate structure. In (19), we have shown that the predicate in a nominal small clause must be indefinite, therefore disallowing the

canonical structure of the ungrammatical (15b) (repeated here as (25)) is (26).

(25) ??san wan de wo mama zhu de tang
three CL-bowl DE my mother cook DE soup

(26) [_{CIP} [_{Rel1} three bowls DE] [_{CIP} [_{Rel2} my mother cooked DE] [_{CIP} soup]]]

The structure in (26) is similar to a multiple adjunction structure: the relative clause *wo mama zhu de* 'my mother cooked DE' is adjoined to the noun phrase *tang* 'soup'. This complex is further adjoined to by another relative *san wan de* 'three bowls DE'. If our analysis is correct, (25) is similar to a stacking relative structure. It is thus essentially the same as (27).

(27) ?wo xihuan [wo gangcai he-de] [wo-mama zhu-de] tang
I like I just drink-DE my-mother cook-DE soup
'I like the soup that my mother cooked that I just drank'

(27) appears to be better than (25). This is probably related to the fact that in (25) the relativized noun is the 'subject' of first relative clause and the 'object' of the second one, while it is the 'object' of both relative clauses in (27). This is supported by (28), which is similar to (25) in that the relativized noun is the 'subject' of the first relative clause and the 'object' of the second. The result is much less acceptable than (27) and much more on a par with (25).

(28) ??wo xiang-kan [xihuan wo de] [wo-gege hen de] nei-ge ren
I want-see like me DE my-brother hate DE that-CL person
'I would like to see the guy who likes me who my brother hates.'

Though it is not entirely clear how (25) and (28) can be ruled out, the comparable effect found in regular relativization cases further supports our analysis of the Num-Cl-*de*-N sequence.

For all clarity, let us emphasize that the Num-Cl-N and the Num-Cl-*de*-N sequences are *not* derivationally related.

4. Mass nouns and count nouns

In section 2, we pointed out that the presence of *de* between the classifier and the noun is not always possible. We saw that it is only possible to have *de* intervening between the classifier and the noun in case the classifier is of the massifier type (see (3)); in contrast, with count-classifiers, *de* is barred from appearing in the sequence, as was illustrated in (4). In the preceding section we analyzed the forms with *de* as relative clauses, deriving from a small clause in which the N is the subject of a predicate consisting of the numeral and the classifier. The other difference between massifiers and count-classifiers we noted in section 2 had to do with the fact that the former can be modified by adjectives like *da* 'big', while for the latter this was not an option. These facts were presented in (5) and (7).

In this section we would like to offer an explanation for the different grammatical behavior of massifiers and count-classifiers and extrapolate our conclusions in analyzing the Num-Cl-N sequences.

Before we move on, we have to make clear what our view is on the structure of NP. We argue elsewhere (Cheng and Sybesma 1997), that bare nouns in Chinese are in fact ClPs: the bare noun is embedded in a ClP with an empty head. We present two types of arguments. First, we point out that there are a number of similarities between D in European languages and Cl in Chinese; specifically we argue that the referential, deictic function D is supposed to have in European languages is performed by Cl in Chinese. The second argument concerns the distribution and interpretation of bare nouns in Chinese, which can be explained if we follow Longobardi (1994), in assuming that in some cases N-to-D (in Chinese: N-to-Cl) takes place (for Chinese, notably in case bare NPs have a definite interpretation). In other words, the bare noun that is relativized in the cases under discussion is a fullfledged ClP; hence the label 'ClP' in (22). This should be kept in mind when we discuss more relativization cases below.

4.1. *Num-Cl-*de*-N

We have indicated in the beginning that not every classifier can be in the sequence Num-Cl-*de*-N. In particular, count-classifiers cannot. Given our analysis, the direct prediction is that count-classifiers cannot appear in the nominal small clause serving as the predicate. However, it is apparently the case that both massifiers and count-classifiers can be the predicate in nominal small clauses, as is shown in the following sentences (see Tang 1996):

- (29) a. wo mai-le shu san ben
 I buy-LE book three CL
 'I bought three books'
 b. ta mai-le bi san zhi
 he buy-LE pen three CL
 'He bought three pens'

These examples show that count-classifiers can constitute the predicate of the nominal small clause; they also show that nouns like *shu* 'book' and *bi* 'pen', whose counterparts in Indo-European languages would count as count nouns, can appear as the subject of a nominal small clause with a Num-Cl as the predicate. Given the relativization analysis of the Num-Cl-*de*-N sequence, the question that arises is why the nominal small clauses in (29a) and (29b) resist relativization. Let us note, first of all, for the sake of completeness, that outside of the realm of nominal small clauses, nouns whose counterpart would be count nouns in Indo-European languages, can be relativized:

- (30) ta mai-le zhangsan xie de shu
 he buy-LE Zhangsan write DE book
 'He bought (some) book(s) that Zhangsan wrote'

Note further that, as we can see from the following minimal pair, it is the classifier that makes the difference:

- (31) a. *ta mai-le san zhi de bi
 he buy-LE three CL DE pen
 'He bought three pens'
 b. ta mai-le san xiang de bi
 he buy-LE three box DE pen
 'He bought three boxfuls of pens'

Thus, the reason why we do not have the Num-Cl-*de*-N sequence with Chinese count noun counterparts is not related to the nature of the noun.

This is in accordance with what we noted earlier: all nouns need classifiers once one starts counting and that the fact that the class of classifiers can be divided into two types (count-classifiers and massifiers) strongly suggests that the cognitive difference between things in the world that present themselves in naturally countable units and those which do not (which, of course, is *semantically* encoded in the noun) is *grammatically* encoded in Chinese, not at the level of the NP, as is the case in Indo-European languages, but at the level of the CIP. Because it is the classifier which determines whether we are dealing with individual elements or whether, in the case of a massifier, even in the presence of a measure, we are still left with a mass, we may say that in Chinese the mass-count distinction is grammatically encoded at the classifier level.

This suggests that when we want to account for the contrasts noted in section 2, we should consider the classifiers and not the nouns. So let us have a closer look at the classifiers in Chinese.

Besides the division mentioned above, dividing the classifiers into two groups, viz., count-classifiers and massifiers, there is another way to divide the classifiers into two classes. Even though all classifiers are nominal, one group of classifiers consists of elements which are completely grammaticalized as classifiers. They form a closed class, and they cannot occur as independent nouns: this class includes classifiers such as *ge* 'classifier for person and other general objects' and *zhi* 'classifier for animals'. The other group of classifiers do not constitute a closed class in the sense that, just like in English, any noun which can be seen to create a unit for measuring mass nouns can be used (notably all sorts of container words or words that can be interpreted as containers; e.g. *wan* 'bowl', *bei* 'cup'). So the latter group consists of elements which also occur as independent nouns.

Interestingly, *grosso modo* this division happens to coincide with the division in count-classifiers and massifiers: the former form a closed class and they solely classify nouns that are cognitively singularizable, i.e., count nouns, such as pens, dogs, etc. The latter, the massifiers, do not form a closed class, and they classify (create counting units for) nouns that are cognitively masses, such as water and sand, (plural) pens and dogs.

In what follows we argue that it is this difference (a classifier is either a full nominal element or a grammaticalized classifier) that yields the different grammatical behavior with regard to the question as to whether or not the subject of a nominal predicate can be relativized.

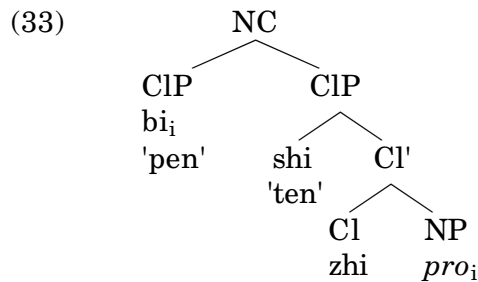
Consider the two nominal small clauses in (32), (32a) involving a count-classifier, (32b) a massifier.

- (32) a. bi [shi zhi]
 pen ten CL

- b. tang [san wan]
 soup three CL-bowl

The predicate in these nominal small clauses consists of a numeral and the classifier head. There is no phonologically overt noun phrase serving as the complement of the head. In regular noun phrases, however, the complement of a classifier is obligatory (excluding cases involving topicalization of a noun phrase stranding a classifier). Thus, the absence of the complement noun phrase in the nominal predicate must be accounted for.

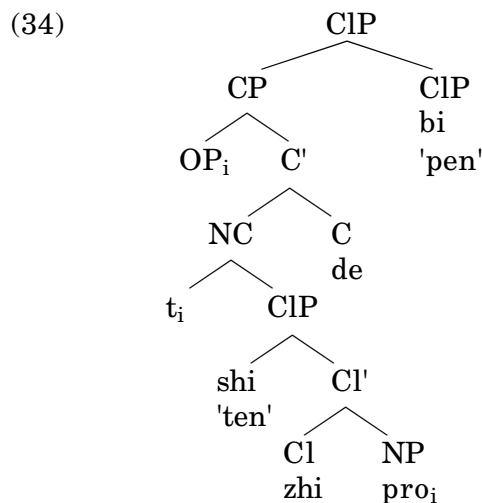
Consider first nominal predicates involving a count-classifier. We suggest that the complement noun phrase of the classifier is a null pronoun:



By positing a null pronoun in the complement position of the classifier, we ensure that all classifiers take a complement. *Pro* in such a structure can be controlled by the subject of the predicate. Note that the structure in (33) raises the question as to why control by the subject noun phrase does not yield a binding violation. We appeal to the fact that *pro* is an NP (rather than a full-fledged CIP) selected by a classifier (classifiers select NPs not CIPs). We assume further that CIP constitutes a binding domain and, since *pro* is within the CIP and there is nothing within the CIP which binds it, no binding violation occurs.

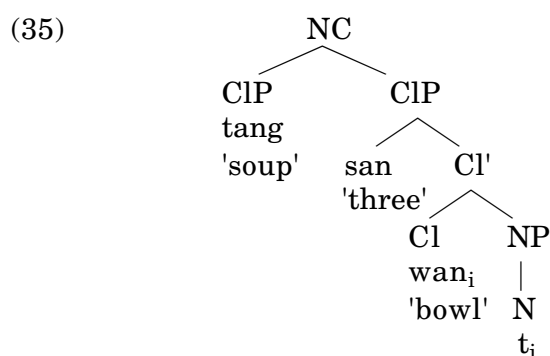
For the sake of clarity, we would like to note that in the system of Cheng and Sybesma (1997), in which noun phrases are really full-fledged CIPs, in regular cases of overt and null pronouns, though generated in NP, they undergo N-to-Cl movement. As a consequence, in regular cases of pronominals, the binding domain of a pronominal is outside CIP.

Consider now the structure after relativization of (33):



Being an empty pronominal, *pro* in (34) (just as in (33)) needs an antecedent/controller. After relativization, *pro* is in a configuration that leads to a bound variable interpretation (since *pro* cannot have independent reference to establish simple coreference). In particular, it is coindexed with the subject of the predicate, which in turn is bound by an operator (see Higginbotham 1980). However, we suggest that this is not possible because *pro* in (34) (or even in (33)) is not a full-fledged pronoun because it stays in the NP rather than moves to Cl. In other words, we hypothesize that to be interpreted as a bound variable, the pronominal has to be a CIP (like a regular pronoun in Chinese). However, in (34), *pro* stays inside the NP due to the presence of a classifier. Thus, (34) is ruled out.¹⁰

As for the massifiers, we noted above that the classifiers which belong to this category are nouns. We propose that when they function as predicates, they start out in N and then undergo N-to-Cl movement, as shown in (35).

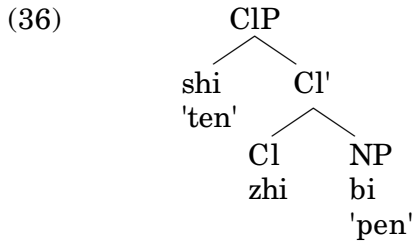


The classifier *wan* 'bowl' provides a measure which is independent of the noun that it is measuring. Thus, it differs from count-classifiers such as *zhi* 'classifier for pens' which is solely a pen-counter. As shown in (35), the noun *wan* 'bowl' undergoes N-to-Cl movement (we come back to this N-to-Cl movement shortly). Thus, massifiers, due to their noun-based property, do not lead to the problem in relativization that is generated for count-classifiers. This difference explains the asymmetry we noted between count-classifiers and massifiers.

4.2. The structure of Num-Cl-N sequences

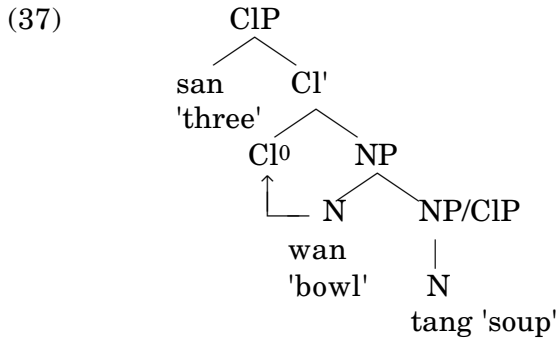
The analysis of the Num-Cl sequences when they occur as predicates has consequences for the analysis of the more typical Num-Cl-N sequences. If Num-Cl as a predicate is analysed differently according to whether the Cl is a count-classifier or a massifier, the question arises whether or not Num-Cl in Num-Cl-N sequences must all be analyzed in the same way.

We propose that the Num-Cl-N sequence *is* analysed differently depending on whether the Cl is a count-classifier or a massifier. For count-classifiers, the ones belonging to the closed class, we propose the following structure, in many ways similar to the structure proposed by Tang (1990) (but there are differences too: for instance we put the numeral in the SpecCIP (but see fn. 8), while Tang treats it as part of the head of CIP together with Cl⁰); cf. (33).



So we have a CIP, the head of which selects an NP which it 'agrees' with (so it cannot select a CIP).

For massifiers, the situation is slightly more complicated. We tentatively propose the basic structure in (37):



As was the case in (35), we take into account the fact that massifiers are really nouns: they start out in the complement of the Cl. The head of this CIP is empty. But, although they are nouns, they are nouns of a certain type: they can be used as a measure and in that capacity they can select another noun. One way to formalize this is by appealing to an idea proposed by Hoekstra (1988), who argues that nouns of this type have a feature, let's call it *CONT* of 'content', and this feature enables the noun to thematically select (and Case-mark) another noun: the content.¹¹ Furthermore, as was the case in (35), at some stage during the derivation, the measure noun moves to fill the head of CIP.

So when it comes to the Chinese Classifier Phrase, we can say that it always selects an NP-complement; the head of the complement may in some cases be phonologically empty, i.e., *pro* (see (33)). Furthermore, the head of CIP may be filled in any of three different ways. In case there is a numeral (that is, all the cases we have dealt with in this paper so far), the Cl-head can be filled either by inserting a count-classifier from the closed class of classifiers, or by moving the complement head N into Cl⁰, which, however, is only possible in case the N is of a certain type, let's say, involving the feature *CONT*. The third way in which the CIP head can be filled is treated in Cheng and Sybesma (1997). As we mentioned above, we suggest there that bare nouns in Chinese are CIPs and, following Longobardi (1994) in spirit, we suggest that in certain cases the head N moves to the head Cl. This movement takes place at LF and is not limited to just *CONT* nouns. Only in case there is an overt numeral do we see the restriction that the raising N have the feature *CONT*: numerals must be in (spec-head) agreement with a 'counter' and only count-classifiers and nouns that can function as classifiers, i.e., the *CONT* nouns qualify as 'counter'.

With regard to these *CONT* nouns in relation to the structure in (37), we have to rule out strings like the following:

- (38) *san ge wan tang
 three CL bowl soup

Apparently, when we insert nouns with the CONT feature into the head of the complement of Cl in (37), the only way to fill the Cl head slot is by raising; as (38) shows, inserting a classifier is not an option. This may be related to the fact that nouns exploiting their CONT feature are somehow in between nouns and classifiers in this position; eventually, it may have something to do with economy (movement to the classifier position is preferred over insertion), although the exact formulation in minimalist terms needs to be worked out.

4.3. Adjectival modification

We saw in section 2.2 that the adjectival modification of a classifier is sensitive to the type of classifiers involved. In particular, massifiers can be modified while count-classifiers cannot. This fact may be seen to follow from the different structures underlying Num-CountCl-N and Num-Massifier-N. It was pointed out that the adjectival modification is furthermore restricted in the following two ways: (a) the element *de*, which typically links the modifier and the modifiee, cannot occur between the adjective and the classifier; (b) only a very small number of adjectives can be inserted. Let us see how our analysis can help us understand these restrictions.

It is well known (see Hoekstra 1988 and references cited there) that container words used as measure phrases cannot be freely modified: the only type of modification that is allowed is modification of the container in its 'abstracted' function as a container. For instance, to use an example from Hoekstra, when the word *wheelbarrow* is used purely as measure expression (as in *He carted four wheelbarrows of dung into the fields*, meaning, of course, that he used one wheelbarrow four times) it is possible to modify the wheelbarrow with words like *big* but not with modifiers like *wooden*: *big* modifies the wheelbarrow in its abstracted function as a container, as a measure, while *wooden* would modify the concrete thing.

If it is true that in configurations like the one we propose in (37), the NP which is the complement of Cl⁰ must necessarily be an NP of a certain type, i.e., marked with the feature CONT, this means that we already use these nouns in their abstracted meaning. So this explains why the range of adjectives that can be used is so very limited.

The fact that classifiers of the closed class cannot be modified while massifiers can, can be explained if we assume that only pure nouns can be modified by adjectives.

We think the impossibility of having *de* as a modification marker here is explained as follows. N-to-Cl is an instance of head movement. There are reasons to assume that in some cases in Chinese combinations of the adjective and a noun form a complex head, both semantically and syntactically. For instance, *hao-ren* 'good-man' is a good person, someone who is inherently good. Combinations of an adjective, *de* and a noun never form a such a complex head (see also Tang 1990). To quote T'ung and Pollard (1982:58): when there is no *de* "the modifier and the modified fuse together into one concept or one mental image: if one says *dà yú* one is thinking of 'big-fish' as a form of life; if one says *dà de yú* one is thinking of a

fish that in this instance is big".

5. Conclusion

In this paper we have investigated the grammar of the classifiers in Chinese, paying special attention to the question of whether the count-mass distinction which is common in the noun system of Indo-European languages plays a role in Chinese grammar too.

We first observed that there are certain phenomena, in particular *de*-insertion and adjectival modification of the classifier, which occur with some but not all classifiers. A closer look revealed that neither phenomenon occurs with what we called count-classifiers: classifiers that enable the noun to be counted by simply naming the unit which is a natural way by which the elements denoted by the noun are counted. On the other hand, both phenomena are possible with what we called massifiers, classifiers which create a measure of counting.

Both phenomena and the difference between count-classifiers and massifiers in this respect can be explained if we assume that the underlying structure of CIPs with a massifier and CIPs with a count-classifier are different. Whereas massifiers are nouns (be it of a special type – they have the feature *CONT*), which are selected by the head of CIP and then move up to fill the *Cl*⁰ slot, as in (39a), count-classifiers belong to a closed class of elements and they are inserted in the *Cl* head position directly, see (39b); in (39) the numeral is disregarded.

- (39) a. [_{CIP} N_i [_{NP-CONT} *t*_i [_{NP} N]]]
b. [_{CIP} Cl [_{NP} N]]

These are the structures involved in all constructions with massifiers and count-classifiers respectively: the Num-Cl-((*)*de*)-N orders as well as the N-Num-Cl orders. As we have seen, adopting these structures, we have been able to explain, not only why CIPs with massifiers and those with count-classifiers display the different forms that they do, but also why they are different.

We have also noted that in all the grammatical processes we have investigated, the noun following the classifier did not seem to play any formal role. That is why we suggest that the cognitively real distinction between things that present themselves in discrete countable units and those that do not is grammatically encoded at the level of the noun in Indo-European languages (there are count nouns and mass nouns count nouns can undergo processes and occur in contexts that mass nouns do not undergo or occur in and vice versa), while in Chinese, the count-mass distinction is grammatically encoded at the level of the classifier: there are count-classifiers and massifiers, and they do not behave the same grammatically. In all languages, the semantic encoding is part of the semantics of the noun, but the grammatical reflexes are found in the nominal domain of the grammar in Indo-European and in the classifier domain in Chinese.

References

- Akmajian, A. and A. Lehrer (1976). NP-like quantifiers and the problem of determining the head of an NP, *Linguistic analysis* 2:395-413.
- Bennis, H. (1979). Appositie en de interne structuur van de NP, *Spektator* 8:209-228.
- Chao, Y.R. (1968). *A grammar of spoken Chinese*. University of California Press, Berkeley.
- Cheng, L.L.-S. and R.P.E. Sybesma (1997). Bare and not-so-bare nouns and the structure of NP. Ms., UC Irvine and Leiden University.
- Chierchia, G. (1995). Plurality of mass nouns and the notion of semantic parameter. Ms., University of Milan.
- Croft, W. (1994). Semantic universals in classifier systems, *Word* 45/2:145-171.
- Doetjes, Jenny (1996). Mass and count: syntax or semantics? Ms. HIL/Leiden University. [To appear in: Proceedings of Meaning on the HIL.]
- Graham, A.C. (1989). *Disputers of the Tao*. Open Court, La Salle, Ill.
- Hansen, C.D. (1972). *Philosophy of language and logic in ancient China*. Doctoral dissertation, University of Michigan.
- Higginbotham, J. (1980) Pronouns and Bound Variables, *Linguistic Inquiry*, 11.4: 679-708.
- Hoekstra, J. (1988). *-fol, fol* en de sintaksis fan mjitteoantsjuttingen, in: S. Dyk and G. de Haan (eds.), *Wurdfoarried and wurdgrammatika*. Fryske Akademy, Leeuwarden/Ljouwert, p. 74-114.
- Huang, C.-T.J. (1982). *Logical relations in Chinese and the theory of grammar*, Doctoral dissertation, MIT.
- Krifka, M. (1995). Common nouns: a contrastive analysis of Chinese and English, in: G. Carlson and F. Pelletier (eds.), *The generic book*. The University of Chicago press, Chicago, p. 398-411.
- Li, C.N. and S.A. Thompson (1981). *Mandarin Chinese. A functional reference grammar*. University of California Press, Berkeley.
- Longobardi, G. (1994). Reference and proper names, *Linguistic Inquiry* 25/4: 609-665.
- Lu, J.M. (1987). shuliangci-zhongjian charu xingrongci qingkuang kaocha [Investigation into the insertion of adjectives with classifiers]. *Yuyan Jiaoxue yu Yanjiu* 1987/4: 53-72.
- Paris, M.-C. (1981). *Problèmes de syntaxe et de sémantique en linguistique Chinoise*. Collège de France, Paris.
- Peyraube, A. (1997). On the history of classifiers in archaic and medieval Chinese. . In: B. T'sou, A. Peyraube and L. Xu (eds.), *Studia Linguistica Sinica*, City University, Hong Kong.
- Selkirk, E. (1977). Some remarks on noun phrase structure, in: P. Culicover, T. Wasow and A. Akmajian (eds.), *Formal syntax*. New York, p. 285-325.
- Stowell, T. (1983). Subjects across categories, *The Linguistic Review* 2:285-312.
- Stowell, T. (1989). Subjects, specifiers, and X-bar theory, in Mark R. Baltin and Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*. The University of Chicago Press, Chicago.
- Sybesma, R.P.E. (1992). *Causatives and accomplishments. The case of Chinese ba*. Doctoral dissertation, Leiden University.
- Tai, J. and L.Q. Wang (1990). A semantic study of the classifier *tiao*. *Journal of the Chinese Language Teachers Association* 25/1:35-56.
- Tang, C.-C.J. (1990). *Chinese phrase structure and the extended X'-theory*.

- Doctoral dissertation, Cornell University.
- Tang, C.-C.J. (1996). *ta mai-le bi shi-zhi* and Chinese phrase structure, *BIHP* 66.
- T'ung, P.C. and D.E. Pollard (1982). *Colloquial Chinese*. Routledge, London.
- Wu, B.T. (1991). *xiandai hanyu yufa tansuo [An exploration of contemporary Chinese grammar]*. Haiyang Daxue chubanshe, Qingdao.

Notes

* We received grants from the International Institute for Asian Studies, Leiden, The Netherlands and from LOT, the Netherlands Graduate School of Linguistics, Utrecht, The Netherlands. The second author's research is financed by the Foundation for Language, Speech and Logic, which is funded by NWO, the Dutch Organization for Scientific Research. All the financial support is gratefully acknowledged.

We have presented parts of the material in this paper in Hong Kong, Paris, Tilburg and Groningen and we thank the audiences for their input and feedback. We would furthermore like to thank Jenny Doetjes, Teun Hoekstra, Utpal Lahiri, Li Xiaoguang, Jane Tang, Luther Liu as well as two anonymous reviewers for comments, discussion and help with the data.

1. English uses the preposition *of* between the measure word and the noun to be measured. Other languages, like Dutch, don't: like in Chinese, nothing intervenes between the measure word and the noun. For discussion of measure expressions and their structure in English, Dutch and Frisian, see Akmajian and Lehrer (1976), Bennis (1979), Selkirk (1977) and Hoekstra (1988).
2. It is important to note that we concentrate on the grammatical reflexes of the mass-count distinction. We believe that semantically the distinction is still part of the semantic make-up of the noun.
3. See Peyraube (1997) for the historical perspective on the difference between the different types of classifiers. For discussion of the function of the individual classifiers, see Doetjes (1996).
4. One of the exceptions is *ben* 'volume' as in *yi da ben shu* /one big volume book/ 'a big book', but it may be the case that *shu* means 'writing' rather than 'book'. Another interesting case is *tiao* when it is used as the classifier for fish: *yi da tiao yu* /one big CL fish/ 'a big fish' is fine. Is 'fish' just seen as one (edible) mass? More difficult to explain is *yi chang bu yingpian* /one long CL film/ 'a long movie'. The same applies to *yi xiao duo huar* /one small CL flower/ 'a small flower' and *yi da zhang ditu* /one big CL map/ 'a big map'. It should be noted however that some speakers do not consider it grammatical to have classifiers such as *tiao*, *bu*, and *duo* co-occurring with an adjective such as *da* 'big' and *xiao* 'small'.
5. The sentence in (13b) does have a grammatical reading. The bracketing then is: [*na san wan de*] *tang*, with the entire *na san wan* 'those three bowls' as the modifier of the *tang* 'soup', and the demonstrative modifying *wan* 'bowl': 'the

soup in those three bowls'. The reading we refer to in the text involves the demonstrative as modifying *tang* 'soup'. We discuss this in more detail below.

6. Tang (1990) argues against a modification structure of the Num-Cl-*de*-N sequence. Leaving aside arguments that relate to count-classifiers, the only potential problem for our analysis is the fact that in the Num-Cl-*de*-N sequence, the N cannot be null while null head nouns are certainly possible with adjectives and typical relative clauses (from Tang 1990):

- (i) wo mai-le xin-de/ [ni-xie]-de e
I buy-LE new-DE you write-DE
'I bought new ones/the ones you wrote.'
- (ii) *wo mai-le [san-bang]-de e
I buy-LE three-CL-DE

One possible source of the ungrammaticality of (ii) under our analysis is that the subject of the nominal small clause is null. We leave this question open.

7. One may have reservations with regard to the nominal small clause analysis because of the semantic selection of the verb, as a reviewer points out. However, it should be noted that the small clause here is not meant to have a clausal interpretation. See for example, Kayne's (1983) analysis of double object construction. See also Tang's (1996) discussion of the structure of the nominal small clause.

8. Despite the fact that elsewhere (Cheng and Sybesma 1997) we argue that the numeral heads its own projection we put it in SpecCIP here, since it is not material to the issues discussed here.

In the examples (21) and (22) *tang* 'soup' is labeled with 'CIP'. We will explain this label in section 4.

9. The sentence in (24a) appears to be more or less acceptable for some speakers. We think that it is possible that the speakers parse the sequence *na guo san wan de tang* /that pot three bowl-DE soup/ as involving appositives. That is, the structure will be as in (i):

- (i) [na guo] [san wan de tang]
that pot three bowl DE soup
'that pot, three bowls of soup'

10. This explanation implies that a noun phrase has to be a CIP to be a bound variable. But an NP (rather than a CIP) can be a controlled element. This certainly leads to many interesting questions and consequences, which we will not be able to address here.

11. The noun selected must be bare. This may have to do with the requirement that the noun must be able to 'totally affect' the container measure. Whether the noun is an NP or a CIP needs to be worked out.

Lisa Lai-Shen Cheng
Department of Linguistics
3151 Social Science Plaza
University of California, Irvine
Irvine, CA 92697-5100, U.S.A.

Rint Sybesma
HIL/Leiden University
Sinological Institute
P.O. Box 9515
2300 RA Leiden
The Netherlands