



HAL
open science

Some remarks on the history of Chinese classifiers

Alain Peyraube

► To cite this version:

Alain Peyraube. Some remarks on the history of Chinese classifiers. Santa Barbara Papers in Linguistics, 1991, Asian Discourse and Grammar, 3, pp.106-126. <hal-05525809>

HAL Id: hal-05525809

<https://hal.science/hal-05525809v1>

Submitted on 10 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

SOME REMARKS ON THE HISTORY OF CHINESE CLASSIFIERS*

Alain Peyraube

CNRS, Centre de Recherches Linguistiques sur l'Asie
Orientale, Paris, France

0. Introduction

A classifier (CL) is a word that theoretically 'must occur with a number and/or a demonstrative before a noun' (Li & Thompson, 1981):

- | | |
|-----------------|------------------|
| 1. 一本书 | 这本书 |
| yi ben shu | zhei ben shu |
| one - CL - book | this - CL - book |
| One book | This book |

For purposes of diachronic analysis, it will be useful and desirable to distinguish classifiers from measure words (MW), which are usually also considered as CL. Thus, the following distinction made by Tai & Wang (1990), Tai (1990) will be adopted: 'A classifier categorizes a class of nouns by picking out some salient perceptual properties, either physically- or functionally-based, which are permanently associated with the entities named by the class of nouns; a measure word does not categorize but denotes the quantity of the entity named by a noun' ¹.

These MW can indicate 1) standards for length, weight, volume and area; 2) aggregates; 3) containers ². They all are probably language universals, while the CL are not. CL exist not only in Chinese, of course, but also in many East-Asian languages, both Sino-Tibetan and non-Sino-Tibetan, but they have no equivalents in Indo-European languages.

When did Chinese CL appear? Did they appear in Early Archaic Chinese (11th-6th c. B.C.) or even in Pre-Archaic Chinese (14th-11th c. B.C.) as Dobson (1962), Huang Zaijun (1964), Greenberg (1975) - whose account is based on Dobson -, Pan Yunzhong (1982) and others think? Or did they only begin to appear under the Late Archaic period (5th-3rd c. B.C.) as is supposed by Wang Li (1958) or Ōta Tatsuo (1958)? I will present evidence for asserting that real CL appear not early than the Han times (2nd c. B.C.-3rd c. A.D.) or even later in Early Medieval Chinese, i.e. between the 3rd and the 6th c. A.D., and that they became prominent under the Late Medieval period, 7th-13th c. A.D.

It is generally accepted that there has been in Chinese a word order change of the CL/MW from a post-nominal position to a pre-nominal one:

2. Noun + Numeral + CL/MW > Num + CL/MW + N.

When did this change occur and why, or at least how? This is the second question I will try to answer after having analyzed (and presenting some statistical figures of the various occurrences) the following seven structures which are attested through the different stages of Chinese, from the Pre-Archaic period to the Late Medieval

period: A. Num + Noun, B. Noun + Num, C. Noun1 + Num + Noun2, D. Noun + Num + MW, E. Noun + Num + CL, F. Num + MW + Noun, G. Num + CL + Noun.

Finally, I will assume in the conclusion, in agreement with Erbaugh (1986) that the noun classifier system might be borrowed from Southeast Asian languages, probably Tai languages.

1. Pre-Archaic Chinese, 14th-11th c. B.C.

In Pre-Archaic Chinese, i.e. in the oracle bone inscriptions, the four following patterns A, B, C, D are used for expressing the quantification of the nouns.

A. Num + Noun

This is by far the most common pattern and it is also probably the oldest one, as Huang Zaijun (1964) points out³. Djamouri (1988) has found 871 instances of this order out of 1886 utterances (i.e. 46.18%) involving a NP quantified by a number. Most of the time, the NP is in an object position, thus following a verb.

3. 出一牛祖乙⁴

sacrifice - one - bovine - Ancestor - Yi
(One should) sacrifice one bovine animal to Ancestor Yi.

B. Noun + Num

This order, though less common than A is also widely used: 324 instances out of 1886 (i.e. 17.17%), also used mainly in an object position⁵:

4. 隻虎一鹿卅狐一百六十四麋一百五十九

capture - tiger - one - stag - thirty - fox - one - hundred - six - ten - four - deer - one - hundred - five - ten - nine
(We) captured one tiger, thirty stags, one hundred and sixty-four foxes and one hundred and fifty-nine deers.

Is there any difference between the two orders A and B? Probably not. Takashima (1984-1985) hypothesizes that the "Num + N" is more likely to be used when the noun conveys old information and the numeral new information, but this discourse-level distinction is hard to maintain in many cases where the two orders successively appear in the same sentence, as in:

5. 卅卅甗卅三三

millet wine - thirty - thirty - penned bovine - captive - three - three - silk

Thirty (vases of) wine, thirty penned bovine, three captives (and) three (rolls of) silk.

In fact we often find, as in 5, the two opposite orders successively following each other.

C. Noun1 + Num + Noun2

This pattern is not as common as the forms A and B, but it is also quite widespread. E.g.:

6. 羌十人

Qiang - ten - people
Ten Qiang.

7. 伐一百羌

human victim - one - hundred - Qiang
One hundred Qiang victims.

Here, the numeral determines the N2, and not the N1. It is indeed possible for an extra element X to be inserted between N1 and Num but not between Num and N2. Thus, it might be accurate to derive pattern B "N + Num" from pattern C after deleting the N2 ⁶.

Sometimes, the N2 is equivalent to the N1, as in:

8. 人十出六人

people - ten - and - six - people

Sixteen people.

9. 予于大甲羌一百羌印一牢

sacrifice- to - Dajia - Qiang - one - hundred - Qiang -
immolate - one - penned bovine

(We will) sacrifice one hundred Qiang to Dajia and immolate a
penned bovine (for him).

Some linguists (Hashimoto 1977, Yau 1986) have argued that the N2 in this C-pattern must be considered not as a noun but as a classifier. They call it echo-classifier, as the same phenomenon exists in Tibeto-Burman languages like Hani, Lahu, Jingpo, Naxi, etc., where the (echo-)classifiers are simply the repetition of the N followed by the number. E.g.:

10. tshoŋ tɕhiŋ tshoŋ (Hani) ⁷

man - one - man

One man.

11. qhâ? ni qhâ? (Lahu) ⁸

village - two - village

Two villages.

However, when the noun is disyllabic, the CL only consists of a partial repetition of the noun, as in:

12. aŋkhwŋ tɕhiŋ khwŋ (Hani)

foot - one - CL

One foot.

13. šú-lè? tē lè? (Lahu)

cigarette - one - CL

One cigarette.

As for the Pre-Archaic Chinese, there is no such situation. Moreover, there are only very few nouns (three, according to Djamouri, 1988) which can be used as N2 as well as N1 in the pattern "N1 + Num + N2". Thus, it will be preferable to consider that the items occurring in N2 position are nouns, as they are in N1 position, and not echo-classifiers.

D. Noun + Num + Measure Word

This pattern is not different from C, as the MW are nouns. Very few MW are attested in the oracle bone inscriptions, probably fewer than ten. E.g.:

14. 鬯六缶

millet wine - six - you

Six you of wine.

You is a noun expressing a container for measuring liquid (wine), as sheng 升 is (see e.g. 17), one you being smaller than one sheng. Yet, you is also a noun meaning "vase", as in:

15. 匚 匚 王 受 又

be - vase - king - receive - assistance
(If it) were a vase, the king would receive assistance (from the gods).

Other examples of the D-pattern:

16. 貝 廿 朋

shell - twenty - coupled string (which was probably a money unit)

Twenty coupled strings of shells⁹.

17. 新 鬯 二 升 一 卣

new - wine - two - sheng - one - you

Two sheng and one you-ful of new aromatic wine.

When a MW is involved in a NP quantified by a number, the order is always the one in D and never the one in F (i.e. "Num + MW + N"). Lastly, neither of the patterns E ("N + Num + CL") or G ("Num + CL + N") is attested in Pre-Archaic Chinese. Wang Li (1958) is right when he asserts that individual classifiers do not exist in the language of the oracle bone inscriptions.

Some linguists like Chen Mengjia (1956), Guan Xiechu (1953), Huang Zaijun (1964), etc. do not agree with Wang Li and claim that individual classifiers do exist. If we exclude from their lists the MW and the so-called "echo-classifiers", which are nouns, the only example they cite which could be analyzed as a CL is bing 丙, as in:

18. 马 五 十 丙

horse - five - ten - bing

Fifty horses.

19. 车 二 丙

carriage - two - bing

Two carriages.

However, bing should rather be considered as a noun and may be a MW, for its meaning is "a set of horses" (Djamouri 1988, Wang Li 1958).

2. Early Archaic Chinese, 11th-6th c. B.C.

The situation in Early Archaic Chinese is not very different from that of the Pre-Archaic period. The four patterns A, B, C, D are all attested. A is still by far the most common one while B is much less used. So is the C-form though we still have the two possibilities of the N2 being different from the N1 (e.g. 20) and the N2 being the same as the N1 (e.g. 21):

20. 子 七 人 (诗经:邶风)

son - seven - people

Seven sons.

21. 牛 三 百 五 十 五 牛 羊 卅 八 羊 (金文, cited in Pan Yunzhong, 1982)

ox - three - hundred - five - ten - five - ox - sheep - thirty - eight - sheep

Three hundred and fifty-five oxen and thirty-eight sheep.

The D-form is more frequent than in Pre-Archaic Chinese. Some forty MW are now found in the "N + Num + MW" pattern, like li 里, cun 寸 for measures of length, dou 斗 or yi 亿 for containers,

etc., which are new:

22. 朱三百亿 (诗:魏风)
grain - three - hundred - yi
Three hundred yi of grain.

As was the case for the Pre-Archaic period, some linguists (Huang 1964, Guan 1981) also think there are some occurrences of "N + Num + CL" (i.e. form E) in Early Archaic Chinese. If we exclude the instances they give which involve a MW or an "echo-classifier" - both nouns and not CL -, they also cite two words pi 匹 and liang 两 which indeed could be possibly interpreted as CL, as in:

23. 马四匹 (尚书:文侯之命)
horse - four - pi
Four horses.
24. 车三百两 (尚:书序)
chariot - three - hundred - liang
Three hundred chariots.

However, liang is probably not a CL because it still has a full lexical meaning, "two wheeled-chariot" ¹⁰; in the same way, pi could mean "a set of horses" (like bing in the Pre-Archaic period) or more likely "single, ordinary (horse)". Anyway, the number of occurrences of liang and pi are too low (there are only very few and scattered examples, the rule still being to have a number immediately followed by the nouns "horse" or "chariot") to allow us to claim that CL already exist during the Early Archaic period.

Lastly, there is no F-form and only one instance of a sentence which could be interpreted as a G-form:

25. 若有一介臣 (尚:秦誓)
if - have - one - CL - servant
If (I) had a servant...

This single example, out of thousands of "Num + N" quantified NP should be, of course, better considered as an interpolated passage, especially when we know that this G-pattern will still not be found in Late Archaic Chinese ¹¹.

Thus, we can conclude that the CL have not yet appeared. There are nonetheless more and more MW, used in pattern D.

3. Late Archaic Chinese, 5th-3rd c. B.C.

In Late Archaic Chinese, which is the Classical language par excellence, the A-pattern is still the most common one and B is becoming rarer and rarer ¹². C also tends to disappear. There are henceforth only some examples with ren 人 'people' in the N2 position, as in:

26. 女五人 (左传:僖二十三)
women - five - people
Five women.

On the other hand, the D-pattern is growing in importance. Numerous new MW are used, for length, weight, military units, surface area, containers, etc. ¹³. For example:

27. 陈文子有马十乘 (论语:公治长)
Chen Wenzi - have - horse - ten - a team of four
Chen Wenzi has ten teams of four horses.

But the most interesting for this Late Archaic period is undeniably the appearance of the F-form which did not exist in Early Archaic. E.g.:

28. 犹以一杯水救一车薪之火也 (孟子:告子上)
be like - with - one - glass - water - rescue - one - cart -
firewood - determinative particle - fire
It is like fighting a fire of a cart-full of firewood with
one glass of water.
29. 与之一筐珠 (左:哀二十)
give - him - one - basket - pearl
(He) gave him a basket of pearls.
30. 一筐食一瓢饮 (论:雍也)
one - basket - food - one - bowl - drink
One basket of food, one bowl of drink.

This form is not yet found with all the MW (we have no example with MW of length, surface area or volume), but it is quite widespread with the MW which are containers and is attested too with the ones which indicate weight, as in:

31. 一鼓铁 (左:昭二十九)
one - gu - iron
One gu of iron.

[One gu was 480 jin 斤 (pounds)].

32. 杀一牛取一豆肉 (韩非子:外储说右上)
kill - one - ox - take - one - dou (bushel?) - meat
(Whenever he) kills one ox, (he) takes (only) one dou of the
(beef) meat.

Although the D-form is still largely predominant, the occurrences of the F-form are far from being rare. Thus, it seems reasonable to date the appearance of pre-nominal MW and the change in 33 in the Late Archaic period:

33. Noun + Num + MW > Num + MW + Noun.

Why did such a change occur? It is difficult to answer this question, but if we admit - and there is much evidence for it - that Chinese has always been a language in which the modifier precedes the head in a NP¹⁴, we can assume that the "Num + MW" phrase has been moved to a pre-nominal position to comply with this order. Why then was the "Num + MW" phrase in a post-nominal position in Pre-Archaic and Early Archaic Chinese? It is because the "Num + MW" phrase of pattern D was probably not a modifier of the preceding noun. This pattern D is not different from pattern C "Noun1 + Num + Noun2", where the number is simply a modifier of the N2 and where "Num + N2" is not at all a modifier of the N1 but a predicate, as supposed by Ota (1958). The gradual disappearance of pattern C has provoked a reanalysis of pattern D: Num + MW has been reanalyzed as a modifier of the preceding noun, and the word order change D > F has thus naturally followed. This word order change has nevertheless not been obligatory, since the D-pattern is still used widely and the two orders will co-occur during a long period of time.

There would have been an intermediary stage between the two orders: 'Num + MW + determinative particle zhi 之 + N', as Pan Yunzhong (1982) supposes. E.g.:

34. 百亩之田 (孟:梁惠王.上)
 hundred - mu (acre?) - det. part. - field
 A hundred mu of field.
35. 以百里之地立为天子 (韩:燕却弑臣)
 with - hundred - li - det. part. - territory - rise - become
 - Heaven - Son
 (He) rose from one hundred (square) li of territory (to)
 become Son of Heaven.

However, it seems to me rash to hypothesize such an intermediate stage between D and F, for 34 and 35 are not different from 36-38 below, where the word following the number is not a special noun MW but a common noun and where the structure is "[Num+N1]NP + det. part. + N2":

36. 万乘之国 (孟:梁惠王.上)
 ten thousand - chariot - det. part. - state
 A state of ten thousand chariots (and not, of course, "ten thousand states").
37. 千室之邑 (论:公冶长)
 thousand - family - det. part. - small town
 A small town of thousand families (and not "a thousand small towns").
38. 用六马之足
 utilize - six - horse - det. part. - leg
 Utilize the legs of six horses.

Thus, 34 and 35 should respectively be understood rather as "a field of a hundred mu" and "a territory of a hundred (square)li".

Lastly, we will admit that there is no E-form nor G-form, i.e. there are no real CL under the Late Archaic period. This statement is obviously contrary to the opinion of most linguists who generally accept that the E-form, if not the G-form, already exists at that time, in addition to which some of them make a strong distinction between the MW, which could be either post-nominal (D) or pre-nominal (F), and the CL, which were always post-nominal (E) and never pre-nominal (G). See for instance Wang Li 1958, Ōta 1958, Huang 1961, etc.

The "classifiers" usually cited in such a E-form are: pi 匹 (for horse), liang 两 (for chariot), zhang 张 (for tent) or even ge 个 (for arrow)¹⁵. However, available examples are very rare. If we look at the following books, Lun yu, Meng zi, Xunzi, Zuo zhuan and Han Feizi - which best represent the Late Archaic period - we find only one occurrence of zhang, one of ge, one of liang, and several cases of pi, all in the Zuo zhuan. However, these occurrences are not certain to be CL. E.g.:

39. 子产以帷幕九张行 (左:昭十三)
 Zichan - with - big tent - small tent - nine - CL - go away
 Zichan went away with nine big and small tents.

In this example, zhang could also be interpreted as a verb "to set up" and the sentence would be "Zichan went away with nine tents to be set up". This interpretation - which is probably the one chosen by Couvreur (1914): "Zichan partit avec des tentures et des baldaquins pour dresser neuf tentes" - seems to me better. As Liu Shiru (1961) notices, it would be odd to have the first occurrence

of the CL zhang used with "tent" and not with "bow", because the original meaning of zhang is "to stretch a bow" and zhang had been commonly used then as a CL for bows a long time before it was used with tents or other objects.

40. 负矢五十个 (荀子: 议兵)

carry on the back - arrow - fifty - ten - CL

He carries on the back fifty arrows.

In this example 40, ge could also have its full lexical meaning "bamboo stem" as arrows are made from bamboo. Ge is to be used later as a general CL, after being used for counting bamboo sticks.

41. 革车三百两 (孟: 尽心下)

leather - chariot - three - hundred - CL

Three hundred leather-covered chariots.

Here too, liang could still have its full lexical meaning "two-wheeled chariot" and the structure would then be that of pattern C.

42. 皆赐玉五鼓马三匹 (左: 庄十八)

all - offer - precious stone - five - pair - horse - three - CL

He offered them all five pairs of precious stones (and) three horses.

There are five cases of pi with a numeral and ma 马 "horse", all in the Zuo zhuan, four of which being in the same passage (chapter zhao. 6) ¹⁶.

It is not impossible either that pi still has a full lexical meaning "single, ordinary", as the following example, where pi is not following a number and is obviously a noun or an adjective, seems to confirm:

43. 匹马隻轮无及者 (左: 僖三)

single - horse - single - wheel - negation - oppose - that which

Not one single horse nor a single wheel was against (him).

In any case, these "classifiers" cited above all have controversial interpretations. Moreover, there are very few occurrences, out of hundreds of quantified NP involving numbers immediately followed by the nouns "tent", "arrow", "chariot", "horse", etc. without anything between the number and the noun. Thus, it seems reasonable to assume that the CL have not yet appeared in the Late Archaic period.

4. Pre-Medieval (Han) Chinese, 2nd c. B.C.- 3rd c. A.D.

In Pre-Medieval Chinese, the A-pattern (e.g. 44) is still the most current one and B (e.g. 45) is by now very rare:

44. 以二女妻舜 (史记: 五帝本纪)

object marker - two - daughter - marry - Shun

(He) married (his) two daughters to Shun.

45. 骑数百 (史: 高祖本纪)

horseman - several - hundred

Several hundred horsemen.

Form C has almost disappeared, except in some instances where the N2 is the noun ren "people", as in:

46. 分予文君僮百人 (史:司马相如列传)
 distribute - give - Wenjun - slave - hundred - people
 (He) distributed a hundred slaves to Wenjun.
 The pattern D is still attested, as in:
 47. 献白璧一双...孔雀二双 (汉书:南粤王赵佗传)
 offer - white - jade - one - pair - peacock - two - pair
 (He) offered one pair of white jade, ... (and) two pairs of peacocks.

However, since the change occurred in Late Archaic Chinese from D ("N + Num + MW") to F ("Num + CL + N"), the F-pattern gradually supersedes the form D which becomes quite obsolete. The normal position of a MW in a NP now seems to be between the numeral and the noun, i.e. before the noun. Examples of F:

48. 则幸分我一椀羹 (史:项羽本纪)
 then - fortunately - distribute - I - one - bowl - soup
 Then (he) fortunately gave me a bowl of soup.
 49. 一尺布尚可缝 一斗粟尚可舂 (史:淮南衡山列传)
 one - chi - cotton - still - can - sew - one - dou - grain - still - can - pound
 One chi of cotton could still be sewn, one dou of grain still could be pounded.

But the most important question for this Pre-Medieval period is certainly the following: do the individual CL, either in the E-form ("N + Num + CL") or in the G-form ("Num + CL + N") already exist? Scholars agree that they do, at least for the structure with a post-nominal CL, and this seems to be the case if we look at the following examples, which involve not only liang or pi:

50. 骑马三匹 (史:匈奴列传)
 ride - horse - three - CL
 Ride three horses.
 51. 羊万头 (史:货殖列传)
 sheep - ten thousands - CL
 Ten thousands sheep.
 52. 牛车千两 (史:冬)
 ox - chariot - thousand - CL
 A thousand ox-driven chariots.
 53. 竹竿万个 (史:冬)
 bamboo - stick - ten thousand - CL
 Ten thousand bamboo sticks.
 54. 弓一张 矢四发 (汉:匈奴传)
 bow - one - CL - arrow - four - CL
 One bow (and) four arrows.

It would be hard, by this time, to claim that pi 匹, liang 两, ge 个, or zhang 张 still retain their original meaning and are not CL¹⁷, even if we can find examples like 55, where pi is a subject and consequently a noun in the second half of the sentence:

55. 马三匹 匹二吏 (居简引页)
 horse - three - pi - pi - two - officer
 Three individual (?) horses, each one for two officers.

However, these cases involving CL are rare, though a little more frequent than in Late Archaic Chinese. Usually, the number is still directly followed by the noun without any CL. From the scarcity of

such CL, one could legitimately ask if these scattered examples are not simply interpolated passages. The answer is apparently negative if we look at the list of CL found in Han bamboo strips given by Huang Shengzhang (1961). Relating to this, Huang puts forward an interesting hypothesis according to which, in Han times, there are probably many more CL in the spoken language than in the written texts, the only documents available to us. He then concludes that the quantified NP occur without CL in the written language but with CL in the spoken language. This hypothesis, however, seems to be overstated, as in Late Han vernacular Buddhist translations, the CL are practically absent; the only "CL" I have found in these texts is zhong 种 "sort of, type of" which is obviously not a CL but a MW¹⁸. The common pattern for the quantified NP in Late Han vernacular Buddhist texts is: Num + Noun.

Lastly, the G-form is not attested in Han times, except in a very few cases; these exceptions are probably interpolated passages. E.g.¹⁹:

56. 五千匹马 (史:大宛列传)

five - thousand - CL - horse

Five thousand horses.

57. 数十辆车 (史:淮南衡山列传)

several - ten - CL - chariot

Several tens of chariots.

In conclusion, two changes occurred in Han times: 1) the pattern F (with pre-nominal MW) becomes widespread and supersedes the D-form (with post-nominal MW); 2) the real CL seem to arise, but in a very timid manner. As we will see in the next paragraph, they probably still retain many noun properties.

5. Early Medieval Chinese, 3rd-6th c. A.D.

During the Six Dynasties period (220-589), the CL become more numerous, but they are still not very frequent. Zhuang Zhengrong (1980) gives the following figures found in the Shi shuo xin yu, written by Liu Yiqing (403-444) and which is one of the best documents of the period in the vernacular language: 300 occurrences of the A-pattern; 10 occurrences of the B-pattern (usually for counting age and money); no occurrence of C; 20 occurrences of both D and E (Zhuang does not distinguish CL from MW); 20 occurrences of both F and G. These figures are revealing: they show that the most common pattern used is still A. E.g.:

58. 袁彦道有二妹 (世说新语:任诞)

Yuan Yandao - have - two - young sisters

Yuan Yandao had two young sisters.

59. 尝以一珊瑚树高二尺许赐椹 (世:汰修)

once - object marker - one - coral - branch - high - two - chi - about - offer - Kai

(He) once offered a branch of coral of about two chi tall to Kai.

Examples of pattern B:

60. 有錢數千萬 (世: 億萬)

have - money - several - thousand - ten thousand
(He) got thousands and thousands of money.

61. 王平子年十四五 (世: 規箴)

Wang Pingzi - age - ten - four - five
Wang Pingzi was fourteen or fifteen.

Actually, the pattern C is not non-existent, as is claimed in Zhuang (1980). We have a few cases of C with ren 人 "people" in the N2 position, but Zhuang probably considers ren as a CL. E.g. :

62. 婢子百餘人 (世: 汰侈)

maid-servant - hundred - more than - people
More than a hundred maid-servants.

As for the 20 instances of mingled D and E, there are as many E (with CL) as D (with MW). Example of D:

63. 米千斛 (世: 雅量)

rice grain - thousand - hu
A thousand hu of rice grain.
[One hu = ten dou (bushel?)]

Example of E:

64. 因下玉鏡台一枚 (世: 假譎)

then - send - jade - mirror - frame - one - CL
(He) then sent (to her) a jade mirror (with) frame.

Besides this CL mei 枚, which was at that time a sort of general classifier which could be combined with many nouns (something equivalent to ge 个 in Contemporary Chinese)²⁰, we can find in various documents of the period the following CL: kou 口 (for both people and animals), zhu 株, tiao 条, ling 頤, tou 头, gen 根, ben 本, zhang 張, ge 个, zhi 只, etc.²¹. Some examples:

65. 以捕鹿二千口供厨 (世: 惑溺)

with - catch - deer - two - thousand - CL - supply - kitchen
(He) supplied the kitchen by catching two thousand deers.

66. 忽生草三株 (拾遺記)

suddenly - grow - herb - three - CL
Suddenly three shoots of herb grew.

67. 獻其國弓三十張 (魏志: 陳瑒王紀)

offer - his - State - bow - three - ten - CL
(He) offered his State thirty bows.

Where do these CL come from? Each CL has its own history. Each comes generally from a noun (or from a verb in very few cases) through a process of grammaticalization which bleaches it and makes it lose its full lexical meaning transforming it into an empty word which has no reality outside the numeral expression. For instance, the general CL mei comes from a noun meaning "tree-trunk", kou from the noun "mouth", zhu and gen from the noun "root", tou from the noun "head", ge from the noun "bamboo-trunk", zhang from the verb "to stretch (a bow)", zhi from the noun "bird", tiao from the noun "branch", etc.²².

This process of grammaticalization is gradual, discrete, discontinuous, beginning probably under the Han period. By the Early Medieval period, it only concerns a small number of items, and the CL, when attested, are not obligatory when used in the

quantified NP; the most common pattern is still "Num + Noun" without any classifiers. Finally, the CL are found in a post-nominal position.

There are nevertheless some occurrences of the pattern G, i.e., with a CL in a pre-nominal position, but this is even rarer: only two instances of the G-form (and eighteen of the F-form) out of the twenty instances of mingled F and G in the Shi shuo xin yu. Some examples of the F-form are:

68. 送一船米遗之 (世:方正)
offer - one - boat - rice - give - he
(He) offered him one boat-full of rice.
69. 送一车枝与和公 (世:俭音)
send - one - chariot - branch - to - He Gong
(They) sent a chariot of branches to He Gong.
70. 长星欢尔一杯酒 (世:雅量)
comet - invite - you - one - cup - wine
Comet, (I) invite you (to drink) one cup of wine.

The pre-nominal MW are now more frequent than the post-nominal ones.

Examples of the G-pattern are:

71. 斋前种一株松 (世:言语)
studio - in front of - plant - one - CL - pine tree
(He) planted a pine tree in front of his studio.
72. 有二百五十条牛 (百喻经)
there is - two - hundred - five - ten - CL - ox
There are two hundred and fifty oxen.

However, these occurrences are too rare to allow us to admit that the G-form already existed in Early Medieval Chinese.

Thus, in conclusion, the CL appear in Early Medieval Chinese but their number is still limited and they only occur in a post-nominal position. On the other hand, the MW are now more likely to be pre-nominal than post-nominal.

6. Late Medieval Chinese, 7th-13th c. A.D.

At the beginning of the Late Medieval period, the CL become frequently used. However, the use of pattern A ("Num + N") is still quite widespread. It represents about 30% of all the occurrences of quantified NP in the Dunhuang bianwen ("transformational texts") dating from ca. 850 to 1015 A.D.²³. E.g.:

73. 我有一女在家 (佛说阿弥陀经)
I - have - one - daughter - at - home
I have a daughter at home.
74. 空中有一神人 (叶净能诗)
sky - in the middle of - there is - one - divine being
There is one divine being in the sky.

We can also find the pattern B, especially for expressing age, as in:

75. 始年十七 (韩朋赋)
just - age - ten - seven
(She) was just seventeen.

As for pattern C, it is only attested when the N2 is ren

"people", as in:

76. 大比丘众千八百人俱 (长兴四年...)
with - big - monk - crowd - thousand - eight - hundred -
people - together
With a crowd of one thousand eight hundred big monks
together.

When a MW is involved in the quantified NP, the common order is F ("Num + MW + N") but some instances of form D ("N + Num + MW") still occur, as in:

77. 美人一对 (韩偓诗)
beautiful women - one - couple
A couple of beautiful women.
78. 罇中有酒五升 (叶净能诗)
Wine bottle - in - there is - wine - five - sheng (liter?)
There was five sheng of wine in the wine bottle.

Examples of the F-pattern are:

79. 取三两粉与这个上座 (祖堂集. 6. P119)
take - three - liang (ounce?) - powder - give - this - CL -
Reverend
Take three liang of powder (and) give (it to) this
Reverend.
80. 万种歌中悦爱情 (维摩诘经)
ten thousand - sort of - song - in - find pleasure - love -
passion
To find pleasure of love in ten thousand songs.

Zhong "sort of, kind of" is by far the most frequently used MW in the Dunhuang bianwen.

When it is a CL and no longer a MW which is involved in the quantified NP, the two orders E ("N + Num + CL") and G ("Num + CL + N") are now widely attested. Thus, the CL can also occur in a pre-nominal position. Let us look first at some examples of pattern E, with post-nominal CL:

81. 女王遂取夜明珠五颗白马一匹赠 (大唐三藏取经诗话)
female - king - then - take - night - shine - pearl - five -
CL - white - horse - one - CL - offer
Then, the queen took five pearls which shine at night (and)
one white horse (to) offer (them).
82. 送龙腿一只 (叶净能诗)
offer - dragon - leg - one - CL
(He) offered one dragon leg.
83. 汗衫子一领 (文)
shirt - one - CL
One shirt.
84. 圣贤罗列百千张 (维摩诘经)
saint - spread out - in row - hundred - thousand - CL
One hundred thousand saints spread out in row.

These CL are now much more numerous than they have been under the Early Medieval period. They are henceforth undoubtedly part of the system concerning the expression of quantified NP. Almost 70% of them involve classifiers in the bianwen and this proportion increases as time goes on.

Another important characteristic of the Late Medieval period is

that the post-nominal CL only represent 10% of the CL used in the bianwen whereas pre-nominal CL, which barely exist under the Early medieval period, now represent 90%. Examples of pre-nominal CL (G-pattern) are:

85. 皇帝与高力士见一条紫气 (叶净能诗)
emperor - and - Gao Lishi - see - one - CL - purple - wind
The emperor and Gao Lishi saw a purple wind.
86. 乘一朵黑云 (韩擒虎)
ride - one - CL - black - cloud
(They) rode on one black cloud.
87. 红叶开时一朵花 (下女(夫)词)
red - leaf - bloom - moment - one - CL - flower
When red leaves bloom, (it is like) a flower.
88. 一件袈裟挂在身 (佛说阿弥陀经)
one - CL - cassock (monk's robe) - hang - on - body
To wear a cassock.
89. 五百头牝牛并犊子 (不知名变文)
five - hundred - CL - cow - and - calf
Five hundred cows and calves.

The most common CL, by far, is now the general CL ge 个, which is commonly used for persons 人. E.g.:

90. 臣有一个问头 (唐太宗入冥记)
I - have - one - CL - question
I have one question.
91. 便变作一个道士 (叶净能诗)
then - transform - make - one - CL - buddhist hermit
(He) then transforms (it) into a buddhist hermit.
92. 我有端严一个女 (佛说阿弥陀经)
I - have - dignified - serious - one - CL - daughter
I have one dignified (and) serious daughter.
93. 教学八万个徒弟 (不知名变文)
teach - eight - ten thousand - CL - disciple
(He) taught eighty thousand disciples.

This possibility for the CL to be in a pre-nominal position and, above all, the frequency of this new pattern G, compared to the old one where the CL was post-nominal (E), should convince us that we are henceforth dealing with real CL; they form CL phrases together with the numerals (or demonstratives) they follow. These CL phrases are directly associated with the nouns they modify, without any possible insertion of any other element between the CL phrases and the nouns.

This new situation should also convince us that before this time the CL we encounter in Early Medieval Chinese or even in Han times - and whose figures given above show that they were very rare - have not probably acquired a real status of CL, i.e., have not been completely divested of their full lexical meaning to become, through the grammaticalization process, empty words. In other words, I am tempted to think that the possibility of being moved from a post-nominal position to a pre-nominal position has been given to the CL, as soon as they acquire the status of CL.

Why, or at least how, does this moving occur? It occurs probably by analogy with the F-form involving a pre-nominal MW and which has

already been predominant for many centuries, because the "Num + CL" phrase is a modifier of the head noun.

Moreover, besides being used in NP involving numerals, some CL (mostly ge) can also begin to be used in NP involving demonstratives, which is a new situation. E.g. (see also the second half of example 79):

94. 这个修行是道场 (维摩诘经)

this - CL - cultivation - be - truth-plot (way to become Buddha)

This cultivation is the truth-plot.

95. 这个地狱中有一青提夫人已否 (大目干连冥间救母)

this - CL - hell - in - there is - one - Qingti - madam - yes - no

Is there a Madam Qingti in this hell?

Sometimes, the demonstrative followed by the classifier ge is a subject, i.e. it does not modify a noun, as in:

96. 这个名为真道场 (维摩诘经)

this - CL - call - to be - real - truth-plot

This is called the real truth-plot.

Lastly, sometimes, there is no numeral or demonstrative, but a single CL modifying a noun. It is obvious, in that case, that a numeral yi "one" has been deleted. E.g.:

97. 路上见个师子 (妙法莲华经)

road - on - see - CL - lion

(He) saw one lion on the road.

Thus, in conclusion, the situation of the CL, beginning in Tang times, is pretty much the same as the one which prevails in Contemporary Chinese.

Conclusion

If a CL 'has no reality outside of the numeral expression', as Greenberg (1972) says, I think it is reasonable to date the common use of Chinese CL only in the Tang dynasty, some time around the 9th or the 10th century A.D., and not in the Archaic period, where there are only MW²⁵. The MW are first used in post-nominal position and then moved in pre-nominal position. This word order change occurs some time in Late Archaic Chinese, between the 5th and 3rd centuries B.C.

As for the individual CL, they probably begin to appear under the Han (1st c. B.C.). However, at that time, they still retain many characteristics of the nouns from which they have issued and they are always post-nominal. The grammaticalization process they undergo to become real CL deprived of their original properties is very long and is completed only in Late Medieval Chinese for a great majority of these CL. As soon as this process is over, the CL are moved to a pre-nominal position and widely used for quantified NP. They also begin to be used between a demonstrative and a noun and not only between a numeral and a noun.

The internal mechanisms of the birth of the noun classifier system described above contradicts Erbaugh (1986)'s assumption that this 'system is not native to Chinese', but she nevertheless might

be right when she hypothesizes that Chinese CL have been borrowed from the Thai, and that they never become a fully automatic part of the grammatical level. Indeed such a late development could well have been caused by external borrowing. This could explain why we have now, in Contemporary Chinese, many more CL in Southern dialects - which have been in closer contact with Tai languages - than in Mandarin (Northern dialects), in contact with Altaic languages known not to have any classifiers.

* A preliminary version of this paper was presented at the UCSB East Asian Linguistics Workshop on December 12, 1990. I wish to thank the participants of this workshop who gave me valuable comments, in particular Charles N. Li, Sandra A. Thompson and Tao Hongyin whose kindness in providing me with relevant information in the Dunhuang bianwen is especially appreciated.

Footnotes

1. The function of a CL is also to quantify the noun which follows it. As Paris (1982) points out, it might indeed be the main function of the CL while a secondary function would be of qualification. See also T'sou (1973).

2. They correspond to what the Chinese linguists call 1) duliangheng 度量衡 (measure classifiers), 2) jiti liangci 集体量词 (collective classifiers), 3) rongliang 容量 (classifiers of content). These linguists call our CL geti liangci 个体量词 (individual classifiers).

3. For a different opinion, see Yau (1986).

4. The sentences from the oracle bone inscriptions cited here are quoted from the books and articles by Djamouri (1988), Serruys (1981) and Takashima (1984-1985) listed in the bibliographical references.

5. Serruys (1981) considers that the number - either before the noun (pattern A) or after the noun (pattern B) - must be considered as a verb. Thus he translates 生于祖辛五伐 (sacrifice-to-Ancestor-Xin-five-human victim) as: "When sacrificing to Ancestor Xin, make it five human victims" and 出岁于祖戊年三 (sacrifice-kill-to-Ancestor-Wu-penned bovine-three) as: "When sacrificing, the penned bovines which we shall kill to Ancestor Wu will be three". This interpretation is controversial, for there are many occurrences of "Noun + Num + you 又 +Num", where you is a coordinating conjunction which can never coordinate two verbs but only links two nominals

(see Djamouri 1988).

6. For a different point of view, see Yau (1986) who thinks that the B-form is the origin of the C-form. He asserts that there is a need to repeat the noun after the number if it was a large number. This opinion is very debatable, as we have small numbers involved in form C and large numbers in form B.

7. See Li Yongsui (1979).

8. See Matisoff (1982).

9. Like you "vase", peng is not necessarily used with a number, e.g.:

易多女出贝朋

offer - many - woman - and - shell - string

(We will) offer many women and coupled strings of shells.

This shows that MW are nouns and not CL because "classifiers have no reality outside of the numeral expression" (Greenberg, 1972).

10. See Liu Shiru (1961) who cites the following example, where liang could not mean anything else but "chariot":

百两卸之 (诗:召南.鹊巢)

hundred - chariot - carry - she

A hundred chariots carry her.

Another example is often cited to assume that the CL already existed in Early Archaic:

葛屨五两 (诗:齐风.南山)

linen - shoe - five - pair

Five pairs of linen shoes.

However, liang means here "a pair"; thus, it is a MW and certainly not a CL.

11. Couvreur (1890) thinks that jie has a full lexical meaning: "ordinary". Thus, he translates "If there were a minister of an ordinary talent...". The translation of 25 is taken from Dobson (1962).

12. An example of the A-form is given in 32 (in the first half of the sentence); an example of the B-form is:

与之粟九百 (论:雍也)

give - him - grain - nine - hundred

(Confucius) gave him nine hundred (measures of) grain.

13. For an impressive list of these new MW, see Yang Xiaomin (1988).

14. This was already the case in Pre-Archaic Chinese. The sole difference in this regard between the language of oracle bone inscriptions and Contemporary Chinese is that in the OB inscriptions the patronymic names follow the kinship terms or the titles.

15. The lists given by Yang Xiaomin (1988) or He Leshi (1989) compiled from the Zuo zhuan are longer but they include numerous MW.

16. There is also the following example in the Mengzi, often cited:

力不能胜一匹雞 (孟: 告子. 下)
strength - negation - can - raise - one - small - chicken
His strength is unable to raise one small chicken.
(Couvreur (1895) translates "duck" instead of "chicken").
However, as Wang Li (1958) pointed out, pi is likely not a CL here but an adjective meaning "small". See also Liu (1961).

17. In example 51, tou 头 can be interpreted as a noun "head", the translation then being "ten thousand head of sheep".

18. I checked the eight following texts, out of the 29 texts that Zurcher (1977) identifies as unquestionably from the Late Han period (the translations were done at Luoyang between 150 and 220): Chang ahan shi baofa jing (Taisho 13, vol. I, 233-241); Ren benyu sheng jing (T. 14, I, 241-246); Yiqie liu she shou yin jing (T. 31, I, 813-814); Si di jing (T. 32, I, 814-816); Ben xiang yi zhi jing (T. 36, I, 819-820); Shifa feifa jing (T. 48, I, 837-839); Lou fenbu jing (T. 57, I, 851-853); Pu fayi jing (T. 98, I, 922-924).

19. Some linguists also cite the following example:

千足羊 (史: 货殖列传)
thousand - CL - sheep

A thousand sheep.

But it would be better to consider zu 足 - whose original meaning is "foot" as a MW - like tou 头 "head"- for counting cattle and not as a CL; the translation would be "A thousand head of sheep".

20. There are however only two mei in the Shi shuo xin yu.

21. For a longer list of CL in Early Medieval Chinese, see Liu Shiru (1961).

22. For the origin of other CL attested during the Early Medieval period, see Liu Shiru (1961).

23. I checked the quantified NP in the collection compiled by Wang Zhongmin (1957).

24. For an analysis of the CL ge under the Tang, see Wang Shaoxin (1989).

25. This contradicts Hashimoto (1977) who claims that 'it is obvious that they (CL) are primarily there for avoiding the ambiguity of monosyllabic nouns'. If it were the monosyllabism of Chinese which caused the emergence of the CL, they should have developed earlier, when the language was really monosyllabic and

not at a time when the monosyllabism was no longer obvious. For a detailed critic of Hashimoto's assumption, see also Paris (1982).

BIBLIOGRAPHICAL REFERENCES

- Chen Mengjia (1956), *General Study of the Divinatory Inscriptions of the Yin*. Beijing: Kexue chubanshe. [In Chinese].
- Couvreur S. (1890), *Dictionnaire de la langue classique chinoise*. Ho Kien Fou: Imprimerie de la mission catholique.
- Couvreur S. (1895), *Les Quatre livres (avec un commentaire abrégé en chinois, une double traduction en français et en latin et un vocabulaire des lettres et des noms propres)*. Ho Kien Fou: Imprimerie de la mission catholique.
- Couvreur S. (1914), *Tch'ouen Tsieou et Tso Tchuan. Texte chinois avec une traduction française*. Ho Kien Fou: Imprimerie de la mission catholique.
- Djamouri R. (1988), *Etudes des formes syntaxiques dans les écrits oraculaires gravés sur os et carapaces de tortue*. Paris: Thèse de l'EHESS.
- Dobson W.A.C.H. (1962), *Early Archaic Chinese*. Toronto: Toronto University Press.
- Erbaugh M.S. (1986), "Taking Stock: The development of Chinese Noun Classifiers Historically and in Young children". C. Craig ed., *Noun Classes and Categorization*. John Benjamins Publ. Comp. 399-436.
- Greenberg J.H. (1972), "Numeral Classifiers and substantival number: problems in the genesis of a linguistic type". *Working papers on language universals* 9. 1-39.
- Greenberg J.H. (1975), "Dynamic aspects of word order in the numeral classifier". C.N. Li ed., *Word order and word order change*. Austin: University of Texas Press. 27-46.
- Guan Xiechu (1953), *Researches on the Grammar of the Yin Oracle Bone Inscriptions*. [In Chinese].
- Guan Xiechu (1981), *Researches on the Grammar of Bronze Inscriptions of the Western Zhou*. Shanghai: Shangwu yinshuguan. [In Chinese].
- Hashimoto M. (1977), "The Genealogy and the Role of Classifier in Sino-Tibetan". *Computational Analyses on African and Asian Languages* 7. 69-78.
- He Leshi (1989), "Numerals and Classifiers in *Zuozhuan*". *Collected essays on Language and Literature*. Shanghai: Zhishi chubanshe. 89-107. [In Chinese].
- Huang Shengzhang (1961), "Classifiers during Han Times". *Zhongguo yuwen* 8. 21-28. [In Chinese].
- Huang Zaijun (1964), "The Origin and the Development of Classifiers in Chinese". *Zhongguo yuwen* 6. [In Chinese].
- Li C.N. & Thompson S.A. (1981), *Mandarin Chinese. A functional Reference Grammar*. Berkeley & Los Angeles: University of California Press.
- Li Yongsui (1979), "A Sketch of the Hani Language". *Minzu yuwen* 2. [In Chinese].
- Liu Shiru (1961), "Research on the Individual Classifiers during the Wei-Jin-Nan-Bei-Chao Period". *Zhongguo yuwen* 10-11. 26-40. [In Chinese].

- Matisoff (1982), *The Grammar of Lahu*. Berkeley: University of California Press.
- Ōta Tatsuo (1958), *Historical Grammar of Chinese*. Tokyo: Konan Shoin. [In Japanese].
- Pan Yunzhong (1982), *Precise Account of Chinese Historical Grammar*. Henan: Zhongzhou shuhua she. [In Chinese].
- Paris M.-C. (1982), "Fonction et fonctionnement des classificateurs en Mandarin". R.P. Kramers ed., *Papers of the XXVIIth Congress of Chinese Studies*. Zurich University. 291-310.
- Schafer E. Jr. (1948), "Noun Classifiers in Chinese". *Language* 24. 408-413.
- Serruys P.L.M. (1981), "Towards a Grammar of the Language of the Shang bone inscriptions". *Proceedings of the 1st International Congress of Sinology*. Taipei: Academia Sinica. 313-364.
- Tai J. H.-Y. (1990), "Variation in Classifier Systems Across Chinese Dialects: Towards a Cognitive-Based Semantic Approach". *First International Symposium on Chinese Languages and Linguistics*. Taipei, Academia Sinica, July 20-22, 1990.
- Tai J. H.-Y. & Wang Liangqing (1990), "A Semantic Study of the Classifier tiao". *Journal of the Chinese Language Teachers Association* 25-1, 35-36.
- Takashima K. (1984-1985), "Noun Phrases in the Oracle-Bone Inscriptions". *Monumenta Serica* 36. 229-302.
- T'sou B.K. (1973), "The Structure of Nominal Classifier Systems". L. Thompson & P. Jenner ed. *Proceedings of the First International Conference on Austro-Asiatic Linguistics*. Honolulu.
- Wang Li (1958), *A Preliminary History of Chinese*. Beijing: Kexue chubanshe. [In Chinese].
- Wang Shaoxin (1989), "Development of the Classifier ge ca. the Tang period. *Yuyan jiaoxue yu yanjiu* 2. 98-119. [In Chinese].
- Wang Zhongmin ed. (1957), *Collected bianwen documents from Dunhuang*. Beijing: Renmin wenzue chubanshe. [In Chinese].
- Yang Xiaomin (1988), "Classifiers in Zuo zhuan". *Journal of Chinese Linguistics (Peking)* 3. 83-95. [In Chinese].
- Yau Shun-chiu (1986), "Classifiers bis". Paper delivered at the 2nd Workshop in East-Asian Linguistics. Paris, CRLAO.
- Zhuang Zhengrong (1980), "Numerals in the Shi shuo xin yu". *Zhongguo yuwen* 3. [In Chinese].
- Zurcher E. (1977), "Late Han Vernacular Elements in the Earliest Buddhist Translations". *Journal of the Chinese Language Teachers Association* 12-3. 177-203.