THE CLASSIFIER CONSTRUCTION OF NEWARI
AND ITS HISTORICAL SOUTHEAST ASIAN BACKGROUND

Alfons K. Weidert
Kathmandu

There are three characteristic typological features of
languages spoken in Southeast and East Asia: (1) monosyllabic
sound structure, (2) tone systems superimposed on the segmental
sound structure, and (3) noun constructions obligatorily ex-
panded by a special form class called classifiers. The descrip-
tive terms for this form class in the linguistic literature are
inconsistant, such as "numeral auxiliary", "numeral coefficient",
"numerical affix", "numerical determinative", "descriptive
word", "counter-word", or German "Zählawort", "Hilfszählwort",
"Numerative", or French "substantive générique", "spécificatif",
"déterminatif d'espèce", or Chinese "jisuan-danwei". On Nepalese
territory there are two languages with a full-fledged development
of numeral classifiers: Newari and Meche. While Newari is known
to the outside world not only as a literary language of classical
standing but also as the language spoken by the native inhabitants
of the Kathmandu valley, the Meche language of Southeastern Nepal
in the Jhapa District, Mechi Zone, because of the low number of
speakers and its topographical seclusion, has escaped the atten-
tion of linguists and anthropologists so far. Due to the ongoing
research of the Linguistic Survey of Nepal which started its
linguistic field work in the Mechi Zone in April 1981, descriptive
material on the Meche language has now become available and
permits the identification of its closer linguistic affiliation
with the Boro and Garo languages of the neighbouring West Bengal

and Assam provinces of India. I shall not dwell on the peculiarities of the classifier system of that language, however; suffice it to state that Newari is not the only fully developed classifier language of Nepal. While the Meche language has to be viewed as a language situated at the western fringe of a contiguous nucleus of languages all closely related to each other and thus allowing the spread of particular linguistic features through diffusion and areal contact, no such explanatory concept can at first sight be offered for Newari. The immediate linguistic neighbours of Newari are the Tibeto-Burman languages which, in all likelihood, never developed the classifier construction in the course of their history. A very well known example in this context is supplied by Classical Tibetan and all its modern dialectal proxies, such as Lhasa Tibetan, Ladakhi, Balti, Kāgate, Khamba, Bhoke, Sherpa, Sikkimese and Bhutanese, and numerous other dialects of western, central and southeastern Tibet and adjoining southern areas within India, Nepal, and Bhutan. All these languages and dialects are devoid, if not in the most rudimentary way, of the numeral classifier construction which is so characteristic of other Tibeto-Burman languages and those neighbouring them. On the other hand, there are a couple of languages in the geographical neighbourhood of Newari with rudimentary classifier developments, such as the Indo-Aryan Nepali where the class of human beings is usually counted by means of the classifying particle *janā*, such as *ek-janā mānche* (1-clf man), *dui-janā mānche* (2 clf man), etc. The influence exerted by Newari on these languages is easily apprehended.

The intention of this essay is to investigate the peculiar phenomenon of the classifier construction in its historical context. We shall try to find a historical explanation for the apparent spread of a characteristic type of syntactic construction. Special reference to the classifier system of Newari will be made and its likely historical origin sketched.

In order to give the non-initiated reader an understanding of what I intend to investigate, let me outline, in the most general terms, a characterization of the term "classifier". I shall not, however, start with the standard description or definition of "classifier" given by the American linguist Joseph Greenberg. His characterizations of a so-called "true classifier language" have proved to be inaccurate in the subsequent discussions on this subject. While it may be impossible to give an accurate definition of a "classifier language" because of the large amount of disparate phenomena which can all be subsumed one way or another under the general term "classifier", the framework of investigations on classifier systems has shifted in recent years from the purely synchronic-descriptive aspect to others such as the typological one, where the problem of linguistic universals comes into play (i.e., to an aspect which tries to answer the question: must classifiers be considered linguistic universals, despite their geographical limitations?) or to the semantic aspect which tackles the problem of semantic content and intent and which has usually led to analyses in terms of componential analysis or componential features. What I intend to do is to give a general idea of what the classifier construction looks like in Southeast Asian languages, without going too much into descriptive details and without comparing these Southeast Asian examples with similar types of construction of other geographical proveniences. In such a way, a classifier language may be characterized as possessing, in numerical contexts, a class of *obligatory* constituents which

(a)   classify or qualify nouns according to semantic criteria, and

(b)   quantify nouns by forming numerical expressions.

Within this characterization, emphasis has first to be

placed on the *obligatory* use of classifiers. The non-obligatory use of classifying or characterizing particles occurs most probably in all languages as a "technical and exceptional" usage which may even grow "into a mere fashion and artificially developed" written style (cf. Hobson-Jobson, s.v. NUMERICAL AFFIXES). Allow me to quote, in this context, English expressions such as *20 head of cattle, 10 file of soldiers, 100 sail of ships, 20 pieces of cannon, a dozen stand of rifles.*

Such a characterization of what a classifier system is like implies both a qualifying and a quantifying aspect. Let us see how these two aspects function in a language like Newari. In order to count nouns, the major prerequisite will be whether it is countable or not. Nouns can be divided according to whether they are countable or measurable. In the first case we speak of *count nouns* or *enumerated nouns,* in the second of *measure nouns* or also *mass nouns.* Countable nouns are human beings, animals, fruits, utensils or any individual object. Measurable nouns are liquids such as water, milk, oil, or any substances whose counting would either take too much time or require the use of containers. For example, it is possible to count stones individually, saying *1 stone, 2 stones, 3 stones,* etc., but if there are too many they may be collected together by a container, enabling us to say *1 basket of stones, 2 lorries of stones,* etc., or they may be abstractly conceived of as non-countable, enabling us to say *1 lump of stones, 2 heaps of stones,* etc. Now let us go back to countable nouns and quote some examples from Newari.

Animate beings are obligatorily counted by means of the classifying particle *-mha,* such as

manuu cha-mha	one person
macā cha-mha	one child
khicā cha-mha	one dog

The two similar sounding classifiers *-ga*(l) and *-gwa*(l) can be interchangeably used for a class of nouns whose inherent property is that of roundness, such as potatoes or balls, e.g,

laddu cha-gaa/gwaa	one bread ball
ālu cha-gaa/gwaa	one potato

But there is also the exclusive use of *-ga*(l) in occurrence with containers, house parts, or as a particulate with grains and granules, e.g,

thala cha-gaa	one pot
thā̃q̃ cha-gaa	one pillar

The two classifiers *-pa* and *-pā*(t) are both used in connection with generally flat objects such as a mat, coat, dish, face, fried egg, pastry, etc. but their distribution with reference to individual nouns is non-interchangeable. Such peculiarities in the use of individual classifiers show that classifiers and their classified nouns are in a constant process of realignment and readjustment. The indiscriminate use of either *-ga*(l) and *-gwa*(l) for the same class of nouns as pointed out above most probably occurred after the two classifiers developed into similar sounding syllables, thus creating a case of near-homonymy.

Amongst the special usages of classifying particles in Newari, two separate cases are noteworthy. There are classifiers with only one head noun; they can only be connected with one noun. Examples are the classifier *-khā* used exclusively for counting houses:

che cha-kha	one house;

or the classifier *-kuu* for counting wounds:

ghāā cha-kuu	one wound;

or the classifier *-ti* for counting arrows:

balā cha-ti	one arrow.

And there is also a group of classifiers which are not classifiers in themselves but simply consist of the reduplication of the counted noun. These are usually called reduplicative classifiers or repetitive classifiers. Examples are

dhwaa cha-dhwaa	one line
haa cha-haa	one leaf
sala-khwaa cha-khwaa	one horse's hoof
pā cha-pā	one feather
palāā cha-palāā	one step
kicaa cha-kicaa	one shadow

Note that from the foregoing examples three important observations can be made. First, regardless of which classifying particle is used for an individual noun, the numerical construction constituted by the three separate parts of speech: noun, numeral and classifier *must* contain the classifier, or else the whole nominal construction will be ungrammatical. In Newari, the usual order of the three obligatory elements which constitute the syntagmatic noun construction is Noun + Numeral + Classifier. Other orders do occur in other languages; but the noun may on no account be placed *between* the numeral and the classifier. In this context, numerical interrogative pronouns such as 'how many?', or indefinite quantifying pronouns such as 'some', 'a few', 'many' can be substituted for the class of numerals represented by positive integers. We shall see afterwards that, once a classifier system has been established for a subsection of countable nouns, the positional contrast created by the obligatory usage of these classifiers may exert a strong *system pressure* on the not yet classified or classifiable part of the countable noun vocabulary so that, at a later stage of the classifier development, the position of the classifier becomes obligatory for every noun countable. System pressure will be made responsible for the existence of the special case quoted above of reduplicative/repetitive classifiers which are also called 'dummy classifiers"

Secondly, we have no way of knowing beforehand whether a countable noun is actually a member of the class of nouns classified by a certain classifier. Though the sensory characteristics of a countable noun (e.g, visual, spatial, or tactile dimensions) may straightaway qualify it to be combined with a certain classifier, that individual noun may unpredictably take another classifier. For example, whereas nouns with the meaning of 'leaf' are usually grouped within the class of flat or 2-dimensional objects, the above-quoted Newari *haa* 'leaf' selects its own repeated form as classifier (*haa cha-haa* '1 leaf'). This observation is usually generalized by asserting that classifiers impose an *arbitrary classification* on the classified nouns.

Thirdly, classifiers have no other meaning than the class meaning extracted as the most salient characteristic feature of the class of nouns that are classified by the classifier in question. Thus, the class meaning of the classifier *-pu* in Newari is 'long-and-thin object' because the most salient characteristic feature of individual objects such as pen, road, shawl is their being long and thin. Sometimes no separate or individual class meaning can be elicited for a class of nouns classified by the same classifiers because the nouns in question have no overt physical similarity. Such is the case for example in Vietnamese where the classifier *thô·t* is used in connection with elephant, garden, and raft; or the classifier *tãm* is used for a bolt of cloth, a board or plank, a hide, a photograph, a ticket, a mirror, a heart, or an example. In Newari, there is an example of the classifier *-gu*(li) with a relatively wide range of combinable nouns. According to the description of the Newari classifier system by Austin Hale and Iswaranand Shreshthacarya, *-gu*(li) is used as a true classifier with nouns referring to locations and geographic features, to abstract states (such as language, thought, matter or topic), and to activities (such as task, trick, guess). Besides these uses, there are miscellaneous items which are not classifiable

under such broad semantic distinctions, such as book, branch, bicycle, body etc. *(sarir cha-guu* = 1 body, *bāiskal cha-guu*= 1 bicycle, *saphuu cha-guu* = 1 book, etc.). Such large semantic classes are usually said to be counted or classified by the *general classifier* if it can be substituted for other classifiers without changing the whole meaning. The class meaning of the general classifier would then merely be 'generally classifiable object'.

On the other hand, in the case of reduplicative classifiers the class meaning of the classifier is the meaning of the noun itself; the meaning of *-haa* in the expression *haa cha-haa* 'one leaf' may be said to have the tautological or tautosemantic meaning content of 'leaf-like object'.

By keeping all these descriptive observations on classifier systems in mind we may now approach the question of historical origin. In order to answer such a question, we have to start from the observation that classifiers are not needed to count countable individual objects. This is shown by those languages which do not possess classifier systems as an obligatory device for numerical noun construction. Most of the world's languages simply use the juxtaposition of noun + numeral in order to express nominal quantification. The order of elements differs from language to language so that in the Germanic languages the order is numeral + noun, whereas in those Tibeto-Burman languages which have not yet developed classifiers the order is noun + numeral, instead. The addition of a third element of construction must be viewed as a superfluous or unnecessary unit of expression, since the meaning content of noun quantification is conveyable by the juxtaposition of noun + numeral alone.

Another observation pertaining to our problem concerns distinguishing unit counting classifier systems from the quantifiers used for noncountable nouns. Quantification of

of non-countable nouns or mass nouns is achieved by measuring
units in the widest possible sense. Though measuring presupposes
counting, it does not necessarily follow that the classifier noun
construction is the historical antecedent of quantification ex-
pressions such as 'a stack of books', 'a bunch of vegetables',
'a piece of meat', 'a lump of clay', 'a mouthful of rice', 'one
cubit of cloth' *(kāpaa ku-chi),* etc. In our investigation, the
reverse result will turn out to be true. It has also been
observed and made into an implicative universal by Joseph
Greenberg that a language is a unit-counting classifier language
if and only if that language also has non-unit counters or mass
quantification units. In contrast to classifiers, non-unit
counters such as 'a stack of', 'aheap of', 'a mouthful of', etc.
are of almost universal occurrence in the world's languages. It
has been suspected by scholars such as Austin Hale that the prior
existence of non-unit counters or measure nouns "provides a
pattern which can be taken as part of the analogical basis for
the development of true classifier constructions". We shall
see that the analogy, or, as we already said above, the system
pressure caused by the mass quantification construction with
its three obligatory constituents: Noun, Numeral, Quantifier
is one of three reasons for the origin and development of
classifier systems.

The fact that classifier languages are observable over a
contiguous area of Mainland Asia leads us to the hypothesis of
area diffusion. If there was ever a single center of origin,
then classifier systems must have spread in all directions,
with only special factors impeding such a spread. The dif-
fusional centre is usually thought to be the Chinese cultural
heart land, though without further justification. In a recent
study of Southeast Asian classifier systems Robbins Burling
has challenged this hypothesis by suggesting that the Tai group
of languages were "a possible source of influence in the spread

of the use of classifiers in Southeast Asia and China" (Burling 1970: 11). His hypothesis is based on the geographical distribution of the structural similarities of the classifier construction. We need not dwell on the further elaboration of Burling's hypothesis, however, because it leaves the development of classifier systems during the recorded history of Chinese unaccounted for. Whereas the Chinese language displays a recorded history of as much as 3000 years, most other languages spoken in Southeast and East Asia do not possess written records of more than only a 1000 years. Such is the case for Tai, the oldest written record of which dates back to Siamese inscriptions of ca. AD 1300. The history of the classifier development from Archaic to Ancient to Middle to Modern Chinese makes it quite clear, however, that classifier developments started from the time of the earliest written records, implying that without outside structural influence, or, to use anthropological terminology, without *stimulus diffusion,* the development into full-fledged classifier systems may require several thousands of years. On the other hand, periods of less than 500 years are quite sufficient for the development of a classifier system if structural innovation or stimulus diffusion is strong enough to spread into adjacent areas, regardless of whether the affected languages are genetically or even structurally related or not to the diffusing language. The Chinese evidence has in fact only been studied in a very superficial way so far. What also has been neglected in the historical study of classifier systems so far, is the political and economic importance of the Chinese territory and its attested historical impact on neighbouring Southeast Asian countries.

The oldest stage of recorded Chinese language history is called "Archaic Chinese" and dates back to the second and first millenia BC. The most famous literary work of that period is the Shy-Jing or *Book of Songs* (cf. Dobson 1968) which spans the period from the 11th to the 7th century B.C. There

are numerous older documents from the 11th and 10th centuries
B.C., with Chinese characters engraved on bronze vessels, animal
bones, or turtle shells. These documents allow us to trace back
modern Chinese classifier systems to their earliest beginning.
The syntax of all these documents is devoid of any classifier
system in its characterization given above. An analysis of
counting and measuring contexts in these documents reveals
some interesting results.

As far as pure counting of individual objects is concerned,
the order within the noun construction usually is *noun followed
by numeral*. If the order N + Num is considered within a topic-
comment or a theme-rheme framework, then the noun must be
viewed as the topic and the numeral as its comment. The
sequence of the two Chinese characters for 'ox' + '2' can
therefore be rendered as: "As far as oxen are concerned,
there are two'. On the other hand, the inverse order of
Numeral + Noun yields semantic structures of quite a different
nature. Here attention is focused on the numeral and the whole
construction results in an adverbial temporal construction,
such as the combination of the two characters for '3' + 'year'
resulting in the adverbial meaning of 'for three years', but
also in the sense of 'in the third year'. Other examples are:

5 – 6 – day	=	'for five or six days'
10 – generation	=	'for ten generations'
10000 – year	=	'for ten thousand years'

The remaining cases of the order sequence Num + N are also
cases of topicalization of the numeral so that the meaning of
the whole construction refers to a particular abbreviate ex-
pression or flowery phrase which is known to the reader or
hearer. Examples are:

3 – gate	=	'the third gate'

$4 - book$ = 'the four books' (= the four classical books of Confucianism)

$4 - barbar$ = 'the four barbaric peoples' (those in the north, south, east and west)

The quantification of enumerated nouns is usually expressed by the order sequence N + Num as was said before. There are examples such as the following ones:

$ox - 2$ = two oxen

$red - cow - 1$ = one red cow

$bow - 1 - arrow - 100$ = one bow and 100 arrows

$cow - 1 - sheep - 1 - pig - 1$ = a cow, a sheep and a pig

If we compare this Archaic Chinese evidence with those Tibeto-Burman languages that have not developed classifier systems to date, the position of the numeral following the noun presents itself as an *overwhelming body of synchronic evidence for the original order sequence Noun + Numeral* of the ancestral language of Chinese and Tibeto-Burman, usually called Proto-Sino-Tibetan.

Some examples for this order sequence may be adduced from Tibeto-Burman languages:

Modern Lhasa Tibetan:

[¯ta ¯$t\check{s}ik$] horse - 1 = one horse

[¯ta ¯$\tilde{n}i$] horse - 2 = two horses

Limbu:

[$si\eta bu\eta$ $naasi$] tree - 5 = five trees

Tangsa:

[$m\Lambda i?$ ${}^1\Lambda{}^2 n\Lambda i$] man - 2 = 2 men

[${}^2 lik-{}^2 d\Lambda p$ ${}^1\Lambda{}^2 n\Lambda i$] book - 2 = 2 books

[${}^1 d\check{z}\Lambda\eta{}^1 s\Lambda i$ ${}^1\Lambda{}^1\check{s}i?$] knife - 1 = 1 knife

Kachin:

$[^1lai^2ka^1buk\ ^3l\partial^{31}nai]$ *book – 1 = 1 book*

$[^2la\ ^3l\partial^{31}\eta ai]$ *man – 1 = 1 man*

$[^2la\ ^1m\partial^2\eta a]$ *man – 5 = 5 men*

Such a hypothesis on the syntactic behaviour of the ordinary Sino-Tibetan cardinal number gains momentum if compared with the syntactic behaviour of measurable mass nomina. Note that so far only two components have been compulsory for expressing the desired count noun quantification, i.e, a single noun plus a single numeral. On the other hand, a third syntactic component becomes compulsory for expressing mass noun quantification. This component is called the *quantifier* and expresses the container, or, more generally, the reference frame required to allow the measurement of the substance in question. In literary Archaic Chinese the order sequence of this construction type is Noun + Num + Quantifier, which means that the noun is under focus, whereas the rest of the construction makes some quantitative comment on the noun. Examples are as follows (cf. Dobson, 1959 and 1962):

bey shyr perng   cowrie shell – 10 – chain ('10 chains of cowrie shells', lit. "as far as cowrie shells are concerned, they are strung into ten strings')

chanq i yeou   sacrificial wine – 1 – yu ('unit of measure') (= '1 yu of sacrificial wine')

ju chyi ell ling   scarlet – flag – 2 – bell (= 'a scarlet banner with 2 bells')

her san-bae yih   grain – 3 – 100 – *yih*   (= '300 yih of grain')

7

sheng chuu i shuh	fresh - hay - 1 - truss (= '1 truss of fresh hay')
ger leu wuu leang	cloth - shoes - 5 - pair (= '5 pairs of cloth shoes')

Regarding the quantification of nouns in the period of Archaic Chinese, we thus encounter the two construction types

   (a) N + Num

for enumerated nouns, and

   (b) N + Num +

for measured nouns. Because of this structural difference the semantic difference in the two construction types must have been clearly felt, at least within lists of counted and measured objects. If this semantic difference is felt, however, as a difference of degree rather than of kind, a structural rapprochement of the two types is probable, and a third structural position will be created for the construction of countable nouns.

   Such a semantic rapprochement also requires a parallelism of sound structure. Translating these two structural formulae into the then prevalent sound structure, an imbalance in the length of the two utterances results, due to the monosyllabism of Archaic Chinese. In the ordinary colloquial language the shortest possible words are used. This means that the item referred to is expressed by means of a monosyllabic noun; the numerals are monosyllabic from 1 to 10 plus powers of 10; and the quantifier within the measured noun construction is always monosyllabic. There is thus no parallelism in the number of syllables of the two construction types, the enumerated noun construction having two syllables altogether, but the measured noun construction having three syllables altogether. It is quite clear that within a pragmatic, situational, every-day

context, especially in the language of commerce and trade (e.g, while enumerating lists of merchandise), the asymmetry in the number of syllables may be removed if appropriate means of expression have become available through other semantic sources.

However, the hypothetical third position of the enumerated noun construction is occasionally occupied, as documented by texts of the Western Chou dynasty (1122 - 771 B.C.). It consists in the repetition of the noun after the numeral, e.g,

> *ch'iang shih ch'iang* = ch'iang - 10 - ch'iang = '10 Ch'iangs, ten people of the Ch'iang tribe'
>
> *t'ian qi t'ian* = field - 7 - field = 'seven fields'
>
> *ren shi you wu ren* = man - 10 - and - 5 - man = '15 people'

We can now distinguish between the latitudinal and longitudinal spread of classifiers as proposed by the Japanese Sino-Tibetanist Mantaro Hashimoto (cf. Hashimoto 1977). The *latitudinal* spread of classifiers means that a certain lexeme will generalize as an inherent class feature of a group of lexemes, e.g. for the class of human objects the lexeme *ren* 'man' is generalized as the inherent class feature of all constructions counting human objects. The above example

> *ch'iang shih ch'iang*

is replaced in a later developmental stage by

> *ch'iang shih ren* = '10 people of the Ch'iang tribe'.

As for the *latitudinal* spread of classifiers, the lack of a suitable classifying particle is felt when compared to the quantifier construction; the empty structural slot becomes occupied by any means even if there is no suitable classifier lexeme, so that the respective noun will have to be reduplicated. Even nowadays it is still possible to observe the emergence of a

classifier system spreading unequally along the latitudinal and
longitudinal axes. Such is the case of certain tribal languages
of Southwestern China, either Miao-Yao or Tibeto-Burman languages,
such as the Na-Khi language, where (a), in the latitudinal direc-
tion, a small class of lexemes has been generalized as the
classifying feature of certain groups of words, and (b), in the
longitudinal direction, system pressure exerted by the three
obligatory structural slots of the quantifier construction has
not yet led to the obligatory usage of classifying particles,
albeit certain nouns are repeated as "repetitive classifiers".

The Archaic Chinese examples given above represent the
longitudinal spread of classifiers by means of the repetition
of the count noun. As far as the latitudinal spread of clas-
sifiers in Archaic Chinese is concerned, there are two examples
of the utmost importance to be found in the inscriptional docu-
ments. Right from the earliest existing documents, the two
words *ma* 'horse' and *che* 'chariot' are connected with the two
classifying suffixes *pi* and *liang*, respectively. Both of the
siffixes have been preserved up to the modern day as classifiers
for 'horse' and 'chariot', e.g, in Mandarin Chinese:

> *ma sz-pi*   horse 4-clf (= 'four horses')
>
> *che sz-liang*   car   4-clf (= 'four cars')

In Modern Mandarin, they have lost all inherent meaning content.
Whereas the classifier *pi* is usually combined with the single
head noun 'horse', thus being similar in usage to our Newari
-khā classifier in the construction *che cha-khā* 'one house',
the syllable *liang* is the ordinary classifier for any kind of
wheeled vehicle.

In the language of the oldest documents the syllables *pi*
and *liang* had their own individual meaning, however. The use of
pi in connection with an animal name implied that originally

horses or other domestic animals were kept in pairs for breeding purposes (cf. Erkes 1956: 88n.). In ancient times riding on horseback was unknown; horses were kept in pairs of two for pulling the war chariot. Originally, the full meaning of an expression such as

> *ma sz-pi* horse 4-clf

was '4 horses, kept in pairs of two (for breeding, or pulling the chariot)', and not just '4 horses' as is the meaning in modern Mandarin.

As far as the syllable *liang* is concerned, its original sense is that of 'a pair, both, two, double'. This is amply demonstrated by the shape of the Chinese logograph itself which shows a scale with two arms and weights attached on both sides. In connection with 'chariot', *liang* implies the sense of a two-wheeled carriage which represents the most important cultural and technical innovation of the Shang dynasty (cf. Dewall 1964). The expression

> *che san-shyr leang*   chariot – 30 – chariot

occurring in the inscriptions must therefore not be translated with '30 chariots' as would be done in Mandarin, but rather '30 two-wheeled chariots'.

In order to show how such an incipient classifier system functions in practice, one half of an inscription is quoted. Note that all the classifying particles still preserve their inherent semantic content.

Inscription No. 14 (SYD, cf. Dobson 1962: 226f.):

'I captured

> *chyou er-ren* 'chief 2-person' = 'two chiefs'

captured
>   *gwo syh-chian ba-bae* 'scalp/4000/800' = '4800 scalps'
>   (no clf !) ...

seized
>   *ren wann san-chian ba-shyr-i-ren* 'man 10000/3000/81 man' =
>   '13081 men'

seized
>   *ma ... pi*   '... horses'

seized
>   *che san-shyr leang*   'chariots - 30 - two-wheelers'

seized
>   *niou san-bae wuu-shyr-wuu niou* 'cattle - 300 - 55 - cattle' =
>   '355 oxen'

>   *yang san-shyr-ba yang* 'sheep - 38 - sheep'.

The results of our investigation of the Archaic Chinese classifier system may be summarized as follows. The emergence of a classifier system in Chinese depends on three peculiarities of the Chinese linguistic structure:

(1) the monosyllabic base structure,

(2) the position of the numeral following the noun in enumerated noun constructions, and

(3) the existence of the quantifier construction with the order: thematic noun followed by the quantifier expression.

The development into a full-fledged classifier system must be thought of as having spread simultaneously along the latitudinal and longitudinal axes. Once such a nominal classifying principle has been set into motion, it quickly encompasses all countable

objects. The domain of classifiers with regard to semantically
structured noun classes will vary from dialect to dialect de-
pending on the strength with which individual classifiers engulf
itemized objects. The classifying particles which in the begin-
ning carried their full meaning content with them lose it later
on, leaving behind the empty nutshell of sound and their 'new'
meaning constituted by the common semantic denominator of the
class of nouns characterized by the individual classifier.
There are two factors which are responsible for the further
latitudinal and longitudinal spread of classifiers. They affect
both the syntactic and the semantic component of the diachronic
development.

(1) The first component is system *pressure* which I have
already mentioned. System pressure is created by the fact that
the construction of measure nouns has three compulsory construc-
tion slots. It causes padding for the construction of countable
nouns with its two compulsory slots (originally).

(2) As far as the semantic component is concerned, the
introduction of classificatory particles added to the bare noun
is subject to the ever ongoing processes of sound attrition and
sound loss of the syllable expressing the noun. The reduction
and loss of sounds causes the notorious problem of homonymy.
In the phonetically most advanced variety of Chinese dialects,
viz., Mandarin, the number of different syllables has dwindled
to a mere 400, discarding tonal differentiation, or 1200, in-
cluding tonal differentiation. The other extreme of maximum
preservation of different syllables in Chinese is represented
by the Yue dialect group, such as the Cantonese dialect, in
which the number of different syllables reaches over 700,
discarding tonal differentiation, or 1800 including tonal
differentiation. There are two possibilities of avoiding
homonymy of nouns: either by directly putting together two
monosyllabic stems so that a nominal compound will result, or

by leaving the noun monosyllabic but adding the right classifier within numerical or certain other contexts. The communicational problem caused by homonymy of monosyllables is obvious. The Chinese language solved this problem by leaving the most commonly used nouns monosyllabic thus serving the purpose of easy communication. Only if absolutely required, such as in numerical expressions, does a further syllable indicating the classifier combine with the noun in order to give it more sound substance or sound redundancy.

Leaving the Chinese language area, we can now proceed to a study of the factors that have to be made responsible for the spread of classifiers in East and Southeast Asia. On the basis of my research on Southeast Asian languages like Malay, Burmese, Vietnamese, and numerous languages of Assam, I have been able to demonstrate that the emergence of classifier systems in Southeast Asia is either due to direct Chinese contact itself or to contact with neighbouring languages which ultimately received their classifier system through Chinese contact. A very convincing example may suffice to illustrate this aspect of classifier diffusion. Old Malay which has been preserved by a few inscriptions found in Sumatra and dated back to the 7th century A.D. does not show any trace of classifiers. About 1000 years later, at the stage called Classical Malay, due to intense Chinese commercial interpenetration along the coasts of Malakka, Sumatra and Java, a fully developed classifier system arose. Since the sound structure of Malay is of a polysyllabic nature, it never had to struggle against homonymy. I suspect that the lack of homonmy is one of the reasons for the simplification of the classifier system that has occurred within the last 2000 years. Nowadays only a handful of classifiers have remained over.

In order to see how a classifier system may originate due

to the influence of a neighbouring language other than Chinese,
I shall now present the case of modern Assamese, the Indo-Aryan
language of Assam. Assamese is as noteworthy as Malay because,
having basically a polysyllabic sound structure with no danger
of homonymy, it has produced a classifier system due to areal
contact, but later on reduced the number of classifiers. The
reduction in the number of classifiers creates an imbalanced
structural system as is observed today in which latitudinal
development of the leftover classifiers may increase on the
one hand (that is, lead to the development of one or two
"general" classifiers); but a decrease in the use of classifiers
is caused within the longitudinal development by making them
either optional or dropping them altogether for large classes
of countable nouns. The language with the donor classifier
system within the Assam territory was Ahom, a Northwestern
Thai language. The people speaking this language had entered
Assam in the 12th century A.D. and established a royal dynasty
that ruled over the plains of Assam for 700 years up to the
beginning of the 19th century. Early documents written in the
Ahom language show that a classifier system was still in the
optional stage of the longitudinal development; in later
centuries it presumably became fully established. All languages
with close contact to Ahom adopted, in more or less elaborated
form, the Ahom system of nominal counting by means of classi-
fiers. In general it can be said that the closer the linguistic
contact existed the more elaborate the classifier system became,
provided that a basically polysyllabic sound structure did not
create too much of an obstacle for its development. The factors
which have to be taken into account for the development of a
classifier system within a basically polysyllabic Indo-Aryan
language are both external and internal ones. The external
factors refer to the strength of stimulus diffusion created
through political, economical and cultural contacts. A par-
ticularly reinforcing factor of stimulus diffusion is

bilingualism, when the recipient people speak both the culturally
superior official donor language as well as their own native
tongue fluently. Bilingualism must have indeed been the de-
cisive factor for interethnic social communication in Assam
as is amply demonstrated by the present-day language situation
of Assam itself as well as by the intense acculturation process
the Ahoms underwent during their reign in the Assam plains until
their ultimate absorption within the Hinduistic cultural main-
stream. The proneness or readiness of a language to incorporate
both phonetic and grammatical features of a donor language
within a contact area usually leads to what has been termed
areal features. Investigating the exact amount of linguistic
transfer from one language to another is one of the thorny
problems of linguistic analysis. In general, the spread of
classifier systems in Southeast Asia can more easily be shown
if there are enough historical records to account for the
social interaction required to explain the spread. We luckily
possess enough records written during the Ahom reign to provide
us with an approximate timetable for the spread of classifiers
into individual recipient languages after either political
subjection, interethnic marriage, interethnic trade and com-
merce, or a combination of these, had taken place.

The internal factors necessary to explain the spread of
classifier systems refer to the sound structure and grammatical
structure of the recipient language. The following example
written in Assamese language shows the structural identity of
the mass noun construction with the enumerated noun construc-
tion. Just like the case of Chinese presented before, structural
identity of the parts of speech within a similar construction
type appears as the main internal prerequisite for the accept-
ance of the third redundant slot within the enumerated noun
construction. The following list of measured and counted
items is taken from an Assamese chronicle about 1700 A.D. It

contains the amount of foodstuffs supplied to a royal messenger for an 8 days' journey:

Assamese text:

cāul 7 purā	husked rice 7-baskets
cikan cāul 7 kalah	fine rice 7-pitchers
māṭimāh 3 purā	lentils 3-baskets
mug 2 purā	? 2-baskets
gur 4 kalah	molasses 4-pitchers
ceni 2 ser	sugar 2-kilos
ghiu 7 ser	ghee 7-kilos
jāluk 2 ser	black pepper 2-kilos
loṇ 8 ser	salt 8-kilos
tel 4 kalah	oil 4-pitchers
khāhi 2-ṭā	castrated goat 2-clf
kukurā 5-ṭā	dog 5-clf
māch 7-ṭā	fish 7-clf

Such enumerations show the typical parallelism of the parts of speech under the influence of the quantifier construction.

This excursion into the linguistic geography of East and Southeast Asia brings us back to the classifier system of Newari. The internal factors responsible for the spread of a classifier system are comparable to those of Archaic Chinese mentioned before. First, Newari basically has a monosyllabic sound structure, with a large percentage of basic noun vocabulary expressed by a single syllable. Similar to the phonological history of Chinese, sound attrition and sound reduction is noticeable when comparing Newari lexemes with Tibeto-Burman cognages. The process of sound reduction usually affects the

syllable-final part where consonants easily drop off. There is also a remarkable degree of syllable homonymy, which may affect communication if there is no other means of built-in redundancy. Secondly, the quantifier construction with its normally ordered sequence N + Num + Quant creates pressure from within the sytagmatic arrangement of the parts of speech. If compared with the quantifier expression, the lack of a third structural slot within enumerated noun expressions would easily be perceived, e.g,

jākhi cha-khwalā		a cupful of rice
lā cha-kuu	meat 1-piece	a piece of meat
manuu cha-kwaṭhā		a roomful of people

Compare this with the normal order of the classifier construction:

chę cha-khā	house 1-clf	one house
pwāā cha-pwāā	hole 1-hole	one hole
khwāā cha-pāā	face 1-clf	one face

As far as the external factors of areal diffusion are concerned, I suspect that the strength of classifier stipulation in Newari may ultimately be due to Chinese influence itself. Such an influence must of course be validated on the basis of historical records; their study may probably shed light on Central Asian trade links in the past (cf. also the interesting remarks and observations by Mary Slusser 1982). The immediate Tibeto-Burman neighbours of Newari to the North, viz., Tibetan with all its numerous dialects, have not developed, to my knowledge, any classifier system to date. (The only Himalayan language I have heard of possessing a classifier system is Thakali, a tribe well known for its trading activities). External factors for such a structural divergence in Tibetan may be found by referring both to the geographic isolation caused by the Himalayas and to the religious forces which have influenced Tibetan culture and language ever since the spread

of Buddhism in Tibet, thus perhaps counterweighing socio-linguistically low level pressures exerted by the communicational necessities of trade and commerce. However, such assumptions are purely speculative for the time being. No such inhibiting forces are noticeable for Newari, so the origin and development of a fully developed classifier system must be mainly attributed to the sociolects of trade and commerce.

Finally, the fact that the Newars have always been famous for their trade skills shall be demonstrated by quoting a passage from the work of Brian Houghton Hodgson, Resident of the British Empire to the Kingdom of Nepal in the first half of the 19th century and one of the first Europeans who brought expert knowledge about Nepal to the Western world. The following passage mentioning the Newars is contained in Hodgson's *Essays on the Languages, Literature and Religion of Nepal and Tibet*, and was probably written about 1830:

> "When we consider how much intelligent activity the native inhabitants of Calcutta have, of late years, been manifesting, we cannot help wondering that none of the mercantile class among them should have yet turned their attention to the commerce of Nepal. Do they not know that the Newars, or aborigines of the great Valley of Nepal, have, from the earliest times, maintained an extensive commercial intercourse between the plains of India on the hand and those of Tibet on the other. ..."

REFERENCES

Burling, Robbins. 1970. "Classifier Constructions in Southeast Asia." *Journal of the American Oriental Society* 90, 1-12.

Dewall, Magdalene von. 1964. *Pferd und Wagen im frühen China.* (Saarbrücker Beiträge zur Altertumskunde, Bd. 1.) Bonn: Habelt.

Dobson, W.A.C.H. 1959. *Late Archaic Chinese - a Grammatical Study.* Toronto: University Press.

————. 1962. *Early Archaic Chinese.* Toronto: University Press.

————. 1968. *The Language of the Book of Songs.* Toronto: Univ. Press.

Erkes, Eduard. 1956. *Chinesische Grammatik. Nachtrag zur Chinesischen Grammatik von G.v.d. Gabelentz.* Leipzig: VEB Deutscher Verlag der Wissenschaften.

Greenberg, Joseph. 1972. "Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type." *Working Papers on Language Universals* 9, 1-37.

Hale, Austin, and Iswaranand Shresthacharya. 1973. "Is Newari a Classifier Language?" *Contributions to Nepalese Studies,* vol. 1, no. 1, 1-21.

Hashimoto, Mantaro J. 1977. "The genealogy and the role of the classifier in Sino-Tibetan." *Computational Analyses of Asian and African Languages,* no. 7, 69-78.

Hobson-Jobson. 1903. ([2]1968) *A Glossary of Anglo-Indian Colloquial Words and Phrases and of Kindred Terms.*

Slusser, Mary Shepherd. 1982. *Nepal Mandala. A cultural study of the Kathmandu Valley.* Princeton.