

## XV. MECHANICAL TRANSLATION\*

V. H. Yngve  
Carol M. Bosche  
Elinor K. Charney  
J. L. Darlington  
J. M. Dolan

R. S. Fabry  
Elfriede Goody  
T. R. Hofmann  
L. Jenkins  
Muriel S. Kannel  
E. S. Klima

Elizabeth Landers  
W. A. Lea  
Ellie Reed Lewis  
B. K. T'sou  
R. J. Zatorski†

### A. THE PLACE OF CLASSIFIERS IN A GENERATIVE GRAMMAR OF CHINESE

The use of numerals with nouns introduces a degree of measure and calls for the use of 'numeral measure words' or 'classifiers' in most languages. Such usage of classifiers is particularly important in the "isolating" languages of Asia,<sup>1</sup> especially in Chinese. In general, it is extremely difficult to determine the rules that would match the nouns and their proper classifiers. The inventory and usage of classifiers vary in the dialects of Chinese, and traditionally nouns are learned together with the classifiers they use. Very often a person's proficiency in a particular dialect may be noted on the basis of his usage of classifiers. Thus a speaker of Fukienese or Hakka, in speaking Mandarin or Cantonese, may inadvertently use 'chih' in 'san chih jen' (three-classifier-men) instead of 'ko' which is the proper general classifier for nouns. Similarly, a speaker of Mandarin or Shanghainese, in speaking Cantonese, may use 'liang' or 'pu' for motorcars instead of 'chia', which is the proper classifier for motorcar in Cantonese.

In the traditional treatment of Chinese Grammar, the place of classifiers in the grammar stands separate from the nouns, and in recent generative grammars of Chinese,<sup>2</sup> the classifiers still remain separate from the nouns in the lexicon, although they are related by syntax. We propose here that the classifiers be listed together with the nouns with which they occur,<sup>3</sup> and that the usage of proper classifiers be determined by the information supplied by the syntactic rules. Since the number of classifiers in Chinese exceeds 100, this treatment will decrease the number of rules in the grammar. Furthermore, the listing need not be random. There can be, at most, four such listings.

We shall begin by observing that in English one important syntactic distinction of nouns is in separating count nouns from noncount nouns. In the case of count-nouns we can immediately precede any one of them with a numeral, e. g., 'three men' and 'four horses', but for the others we have to employ count nouns as counters, e. g., 'three pounds of meat' and 'two kinds of hopes', where 'pounds' and 'kinds' are counters and their function is analogous to that of the classifiers. There is a further semantic distinction in English concerning exact and inexact measures, e. g., 'Three cows' is an exact measure of 'cows', whereas 'three herds of buffalos' is an inexact measure of 'buffalos'. We can perhaps illustrate this scheme by considering the different measures that we can apply to 'newspaper'.

---

\*This work was supported in part by the National Science Foundation (Grant GN-244).

†Visiting from the University of Melbourne, Australia.

(XV. MECHANICAL TRANSLATION)

Newspaper		Exact	
		+	-
Entity	+	sheet	(bundle)
	-	pound	kind

Here, 'sheet' is one of the rare classifiers in English and it goes with 'paper', and perhaps 'blanket'. For regular count nouns we have no classifiers for exact and entity measure. For example, for the noun 'lady' we can have three kinds of measures but only two 'classifiers':

Lady		Exact	
		+	-
Entity	+	φ	bevy
	-		type

For a mass noun such as 'beef' we have only three measures and three 'classifiers':

Beef		Exact	
		+	-
Entity	+		(order)
	-	pound	kind

By using our scheme and by making observations on other languages, we can see that classifiers exist in all languages. It is only in isolating languages such as Chinese, Thai, and Vietnamese that we have a particular abundance<sup>4</sup> of classifiers for exact and entity measures, and that the absence or wrong usage of classifiers in measure phrases yields ungrammatical sentences. In 'inflectional' languages such as English and German there are generally no classifiers for count nouns in exact and entity measures because the notion of entity is usually included in the nouns.

For Chinese we can define 'classifier' as having the following position in a noun phrase:

$$\left[ [X \ [Q \ Cl]_{mp} \ (Adjm)]_{am} \ N \right]_{NP}$$

Q = quantifier, containing numerals and interrogative pronouns

Cl = classifier

(XV. MECHANICAL TRANSLATION)

mp = measure phrase

Adjm = adjectival modifier construction<sup>5</sup>

am = adjectival modification

N = noun

NP = noun phrase

Note that this is not the only position for classifiers in a noun phrase. In sentences involving the enumeration of objects the measure phrase may be placed after the noun in the noun phrase:

$$\left[ (\text{Adjm}) \text{ N } [\text{Q Cl}]_{\text{mp}} \right]_{\text{NP}}$$

This is a rather restricted type of syntactic pattern in Chinese, and it is likely that it is carried over from classical Chinese. In any event, this does not affect our proposal here.

Applying our scheme to some examples of Chinese we have:

Exact

niu-ju (beef)		+	-
		+	p'an
Entity	-	chin	chung

Exact

jen (man)		+	-
		+	chün
Entity	-		chung

and

Exact

chi (chicken)		+	-
		+	chün
Entity	-	chin	chung

(XV. MECHANICAL TRANSLATION)

In the case of 'chi' (chicken) which normally belongs to the class of count nouns, such as 'niu' (cow), 'jen' (man), etc., it has acquired one feature of mass nouns such as 'hiu-ju' (beef), in that we can say both 'san chin chi' (three catties of chicken) as well as 'san chin chi-ju' (three catties of chicken meat).

Using this scheme of categorization we may, in the lexicon of the grammar, list each noun with the different kinds of classifiers it may have. Each listing, then, may contain more than one classifier, which functions as one class of syntactic elements. Thus, for example, 'liang' and 'pu' are syntactically equivalent classifiers for 'motorcar', so are 'chin' (catty) and 'liang' (ounce) for 'niu-ju' (beef). In generating sentences we need not write rules that will match the nouns with their particular classifiers. We consider the classifiers to be parts of the nouns, which may be brought out by simple rules.

This may systematically explain why we do not have \* san ko niu-ju not because we have left out classifiers of exact entity measure for mass nouns in our matching rules but because there simply does not exist such a measure for the nouns concerned as shown in the lexicon. Also, for a verb such as 'fa-ming' (to invent) we may have:

ta fa-ming-le san-chung wu-hsiang niu-ju.

(He has invented three kinds of spiced beef.)

but not: \*ta fa-ming le san  $\left\{ \begin{array}{l} \text{chin} \\ \text{pan} \end{array} \right\}$  wu-hsiang niu-ju.

(He has invented three  $\left\{ \begin{array}{l} \text{catties} \\ \text{plates} \end{array} \right\}$  of spiced beef.)

This is because 'fa-ming' does not take non entity-exact measure nor entity-exact measure in the object noun phrase. If we had listed the classifiers separately from the nouns, we would have a great deal of difficulty in doing the proper matching in the verb phrase. But if this scheme is followed, we would merely have to include in the verb the information that it does not take non entity-exact measure nor entity-exact measure.

Returning now to our comparison of English with Chinese, we have said that classifiers for exact entity measure are most important in the measure phrase of Chinese, whereas in languages such as English there is generally no need for them. The reason offered was that the notion of entity was usually included in the nouns. This indicates a difference in the nature of nouns of both types of languages. One might raise the query whether all Chinese nouns are of the type of collective nouns in English. This is because collective nouns in English, such as 'cattle' and 'paper' (in one sense of the word), each takes a classifier in exact entity measures, that is, we do not have two cattle<sup>6</sup> or two papers<sup>6</sup> (in the sense of 'two sheets or pieces of paper') but rather 'two head(s) of cattle' and 'two sheets of paper'. Furthermore, we do not use the plural morpheme[S] for

'cattle' or 'paper' in the examples above, whereas we use it with the other nouns such as in 'two cats' and 'two boys'. The fact that Chinese has no plural morpheme (except in the pronouns) may indicate that there is a similarity between Chinese substantives and English collective nouns, in that the obligatory classifiers in Chinese function analogously to those of English collective nouns. Also, the function of the classifier in exact entity measure for Chinese partially parallels that of the plural morpheme for English. This is because the latent notion of entity included in the English count noun is manifested in the determiner system composed of the numeral plus the plural morpheme. On the other hand, in the collective noun this entity notion is brought out and externally expressed by the use of the classifier without the plural morpheme. An isolated instance from German may further substantiate this reasoning in that we use 'glas' without the plural morpheme in 'zwei glas Bier'; that is, we do not have 'zwei gläser Bier'. Here it seems that 'glas' functions exactly as a Chinese classifier.<sup>6</sup>

The English collective noun, however, may generally utilize its classifiers in nominalization processes, for example, 'cattle herds', 'savage hordes', 'mankind', etc. On the other hand, most Chinese nouns may not undergo this; that is, we do not have \*jen-ko, \*chi-chih, \*chè-pu, etc., although we may have 'jen-chung', 'ma-p'i', etc. This indicates that both our positive and negative thoughts about this categorical similarity are worth entertaining. Further discussion or resolution of our doubts is quite beyond the scope of this paper.

B. K. T'sou

#### References

1. See, for example, M. B. Emeneau, Studies in Vietnamese (Annamese) Grammar (University of California Press, Berkeley, 1951), pp. 84-114; Mary R. Haas, "The Use of Numeral Classifiers in Thai," Language 18, 201-205 (1942).
2. A. Yue Hashimoto, "A Transformational Outline of Cantonese Grammar," M. A. Thesis, University of Texas, 1963 (unpublished).
3. Professor Mary Haas, op. cit., has suggested that "the most useful type of dictionary for the language (Thai) would be one which provided each noun with a parenthetical indication of the classifiers required to be used with it." It is proposed here that a general definition of classifiers yields four kinds and that while it is extremely difficult to manipulate one of these four kinds, we should consider all four together.
4. There are more than 80 classifiers in Thai, as cited from Frankfurter by Prof. Haas, op. cit., and there are more than 100 classifiers in Vietnamese recorded by Professor Emeneau, op. cit. For Chinese, the number also exceeds 100. See, for example, Kao Ming-K'ai, Han-yü yü-fa-lun (On Chinese Grammar) (K'ai Ming, 1948).
5. More detailed treatment on this may be found in B. T'sou, "Chinese Grammar 1: Adjectival Modifier Construction," Internal Memorandum, Mechanical Translation Group, Research Laboratory of Electronics, M. I. T., Cambridge, Mass., 1963 (unpublished).

(XV. MECHANICAL TRANSLATION)

6. Similarly, and curiously enough, Webster's New Collegiate Dictionary records the identical form for the plural of 'head' when it is used as a classifier for 'cattle'. As for 'herd' it records an optional plural form 'herds'. 'Sheet', however, takes the plural form when used in plural measure. This tends to show that the English (and perhaps German) collective noun is developing in the direction of the Chinese substantive.