55. Numeral Classifiers

David Gil

1. Introduction

In English and in other languages, nouns may vary with respect to the property of **countability**. Nouns of high countability such as *woman*, *dog* and *pencil* denote objects which are conceptualized in terms of highly individuated units typically associated with a characteristic shape. In contrast, nouns of low countability such as *water*, *sand* and *smoke* denote objects which are conceptualized in terms of masses without unitary structure or characteristic shape.

Countability bears a number of grammatical consequences, one of which is the way in which nouns may occur in construction with numerals. Nouns of high countability generally occur in direct construction with numerals, for example *one woman, two dogs, three pencils*. In contrast, nouns of low countability typically do not occur in direct construction with numerals; instead, an additional item must be present, for example *one **glass** of water*, *two **pounds** of sand*, *three **plumes** of smoke*. Such additional items are sometimes referred to as **mensural numeral classifiers**, since they provide nouns of low countability with a unit of measure by means of which they may then be counted.

However, in many other languages, nouns in construction with numerals may occur with an additional grammatical element even when such nouns are of high countability:

(1)   Minangkabau (Sundic, Austronesian; Indonesia; own
         knowledge)
         a.   *sa-**urang**   padusi*
                one-CLF   woman
                'one woman'

b. *duo* **ikue** *anjiang*
two   CLF   dog
'two dogs'

c. *tigo* **batang** *pituluik*
three   CLF   pencil
'three pencils'

Such elements are typically referred to as **sortal numeral classifiers**. This term makes reference to one of the most salient functions of such forms, which is to divide the inventory of count nouns into semantic classes, each of which is associated with a different classifier. Thus, for example, in Minangkabau, as illustrated in (1) above, human nouns take the classifier *urang*, nonhuman animate nouns the classifier *ikue*, and elongated objects the classifier *batang*. In languages with sortal classifiers, the number of such classifiers may vary considerably, ranging from a mere handful to several dozen.

This map displays the worldwide distribution of sortal numeral classifiers. The map is not concerned with mensural numeral classifiers, which most or all languages have. Accordingly, sortal numeral classifiers are referred to here as *numeral classifiers* or simply as *classifiers*, in contexts where the meaning is clear.

2. Feature values

The map divides the world's languages up into three types, in accordance with whether, in the numeral-plus-noun construction, numeral classifiers are absent, optional or obligatory.

@	1.	Numeral classifiers are absent		260
@	2.	Numeral classifiers are optional		62
@	3.	Numeral classifiers are obligatory		78
			total	400

In languages of the first type, there are no numeral classifiers; a numeral always occurs in direct construction with a noun without the additional presence of a classifier. One example of such a language is English. Another is Tagalog, in which the numeral-plus-noun construction is marked with the further presence of a ligature of morphophonemically variable form, in the following example the enclitic =*ng*:

(2)   Tagalog (own knowledge)
        dalawa=ng aso
        two=LIG   dog
        'two dogs'

Other examples of languages without numeral classifiers include Yoruba (Welmers 1973: 289), Hebrew (own knowledge), Tiwi (Osborne 1974), Maricopa (own field work) and Shipibo-Konibo (Pilar Valenzuela p.c.).

In languages of the second type, numeral classifiers are present, but their occurrence in the numeral-plus-noun construction is optional. One example of such a language is Minangkabau, in which classifiers such as those illustrated in (1) are sometimes present, but in other cases absent:

(3)   Minangkabau (own knowledge)
        duo   (ikue)   anjiang
        two   (CLF)   dog
        'two dogs'

Other examples of languages with optional numeral classifiers include Hungarian (Beckwith 1992), Chantyal (Noonan 2000), Hatam (Reesink 1999: 57–58), Tongan (own field work), and Haida (Swanton 1911: 277–278).

In languages of the third type, numeral classifiers are obligatory in the numeral-plus-noun construction: a numeral cannot quantify a noun without such a classifier being present. One example of such a language is Vietnamese:

(4)   Vietnamese (own fieldwork)
       hai   **con**   *chó*
       two   CLF   dog
       'two dogs'

Other examples of languages with obligatory numeral classifiers include Boko (Jones 1998: 128), Nivkh (Daniel Abondolo p.c.), Nyelâyu (Françoise Ozanne-Rivierre, Isabelle Bril p.c.), Coast Tsimshian (Boas 1911: 396–398) and Warekena (Aikhenvald 1998: 298).

It is sometimes the case that in languages with numeral classifiers, their occurrence in the numeral-plus-noun construction is dependent on the choice of the numeral or of the noun. For example, with regard to numerals, in some languages numeral bases such as 'ten', 'hundred', 'thousand' and the like are themselves treated as numeral classifiers, and therefore take the place of another numeral classifier which might otherwise have been present. For the purposes of this map, classifiers are treated as obligatory if there exists a major class of numerals, such as for example two through nine, for which the numeral classifier is obligatory, even if for other numerals the classifier may be optional or absent. Similarly, in the absence of a major class of numerals for which a numeral classifier is obligatory, classifiers are treated as optional if there exist's a major class of numerals for which the numeral classifier is optional, even if for other numerals it may be absent. With regard to nouns, in some

languages different noun classes may differ with respect to the requirement of a numeral classifier in the numeral-plus-noun construction. For example, in Toqabaqita (Oceanic, Austronesian; Solomon Islands) only a small and semantically heterogeneous set of nouns may occur with classifiers; most other nouns do not (Frank Lichtenberk p.c.). For the purposes of this map, classifiers are treated as obligatory if for all major classes of nouns, the classifier is obligatory. Alternatively, if this condition is not met, classifiers are treated as optional if there exists at least one class of nouns for which they are either optional or obligatory, as is the case for Toqabaqita.

The distinction between optional and obligatory numeral classifiers is further complicated by the existence of specific contexts and special styles of speech in which the distribution of numeral classifiers diverges from the norm. For example, grammar books of Vietnamese usually claim that numeral classifiers are obligatory in the numeral-plus-noun environment, and speakers of Vietnamese will typically judge the absence of a numeral classifier in such constructions to be ungrammatical. Accordingly, as noted above, Vietnamese is characterized as a language of the third type, with obligatory numeral classifiers. Nevertheless, there is at least one common context in which numeral classifiers are frequently absent, that associated with food stalls and restaurants. In a simple eating establishment, there might be a choice between two kinds of noodle soup, containing either chicken or beef. A group of, say, five people placing their order might ask for *ba gà hai bò*, or 'three chicken two beef', without any classifiers; and the person taking the order will shout back *ba gà hai bò* to the cook, again without any classifiers. For the purposes of the present map, such specific contexts and specialized styles of speech are not taken into consideration; the usage of numeral classifiers is described as it occurs in what might be characterized as "ordinary citation" language.

3.   Geographical distribution

Numeral classifiers exhibit striking worldwide distribution at the global level. The main concentration of numeral classifiers is in a single zone centered in East and Southeast Asia, but reaching out both westwards and eastwards. To the west, numeral classifiers peter out as one proceeds across the South Asian subcontinent; thus, in this particular region, the occurrence of numeral classifiers cross-cuts what has otherwise been characterized as one of the classical examples of a linguistic area, namely, South Asia. However, numeral classifiers pick up again, albeit in optional usage, in parts of western Asia centering on Iran and Turkey; it is not clear whether this should be considered as a continuation of the same large though interrupted isogloss, or as a separate one. To the east, numeral classifiers extend out through the Indonesian archipelago, and then into the Pacific in a grand arc through Micronesia and then down to the southeast, tapering out in New Caledonia and western Polynesia. Interestingly, whereas in the western parts of the Indonesian archipelago numeral classifiers are often optional, in the eastern parts of the archipelago and in Micronesia numeral classifiers tend once more, as in mainland East and Southeast Asia, to be obligatory.

Outside this single large zone, numeral classifiers are almost exclusively restricted to a number of smaller hotbeds, in West Africa, the Pacific Northwest, Mesoamerica, and the Amazon basin. In large parts of the world, numeral classifiers are completely absent.

4.   Theoretical Issues

Numeral classifiers have captured the imagination of linguists and others interested in language for a variety of reasons. Most obviously, linguists, psychologists, anthropologists and other scholars have been interested in the classificatory function of

numeral classifiers and what insights it might provide into how speakers view the world (Haas 1942; Burling 1965; Craig 1986). Whereas in some languages each noun is associated with exactly one classifier, in other languages there is a degree of flexibility with regard to the choice of classifier, permitting the differentiation of subtle shades of meaning. For example, in Mandarin, when *sān* 'three' occurs in construction with *huā* 'flower', *sān zhī huā,* with classifier *zhī* for elongated objects, conjures up an image of three flowers on their stalks, whereas *sān duǒ huā*, with classifier *duǒ* for round objects, focuses in on the actual florescences, to the exclusion of the stalks.
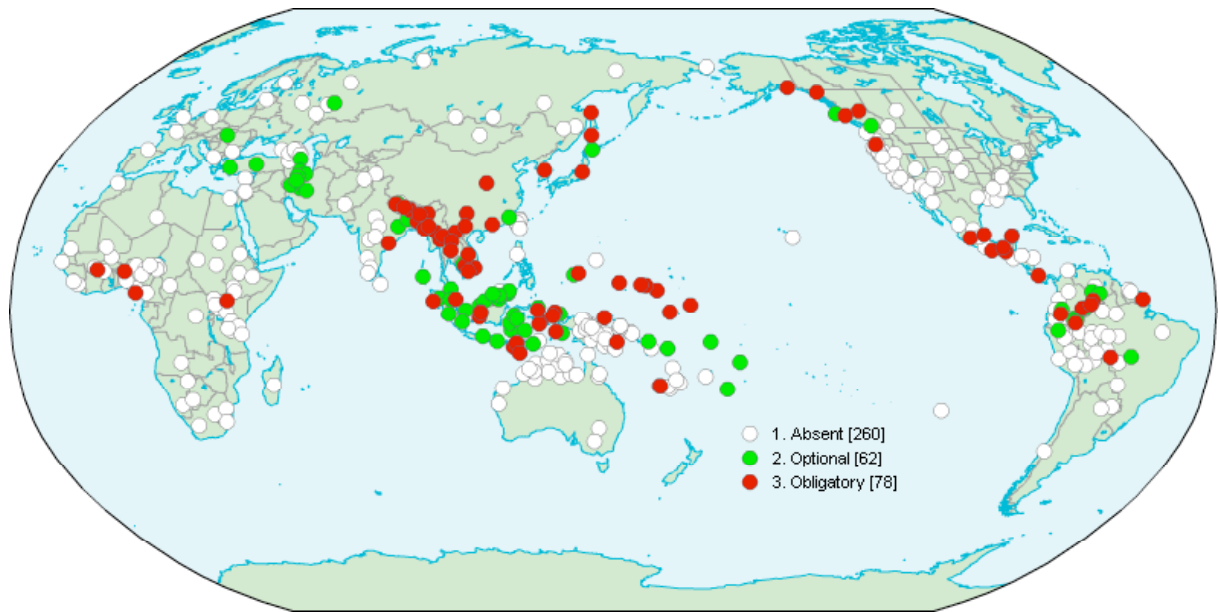
Other interesting issues pertain to the grammar of numeral classifiers. One question is that of constituency: does the classifier form a constituent with the numeral or with the noun? There is at least some reason to believe that the answer may vary across languages, and also across different kinds of classifiers. A second question is that of headedness: given a particular constituency, which of the elements is the head and which is the modifier? Again, different scholars have proposed different answers to this question, which may be due to differences between languages, differences between theoretical approaches, or perhaps other reasons.

Whereas in some languages, for example Minangkabau, numeral classifiers occur only with numerals, in other languages they may occur in a variety of other constructions. For example, in Mandarin they may also occur with several other quantifiers as well as with demonstratives, as in *zhèi ge píngguǒ* 'this CLF apple', while in Cantonese, in addition to the above, they may occur also with possessor noun phrases, as in *a³³-faay⁵⁵ lap5 piŋ¹¹-guo³⁵* 'Ah Fai's apple'. In addition, numeral classifiers sometimes have a nominalizing function of the kind described in chapter 61. For example, in Vietnamese, they may be used to nominalize demonstratives, adjectives and relative clauses, as in *con Quân dã mua* 'one that Quan bought'. Finally, in some languages, numeral classifiers may occur in construction with a

noun without any other attributive expression being present; in such cases, the classifier functions rather like an article, though varying from language to language with respect to definiteness.

Why do languages have numeral classifiers? Most commonly, classifier languages are ones in which an NP consisting entirely of a bare noun may be understood as either mass or count, and as either singular or plural (Gil 1987). For example, in Mandarin, the noun-phrase *píngguǒ* 'apple' may be understood as 'apple' (mass), 'an apple' (singular), or 'some apples' (plural). This cross-linguistic generalization has led to a widespread belief that in such languages, the classifier is needed in order to individuate the noun and provide the necessary units to facilitate quantification (Thompson 1965; Quine 1969; Stein 1981; Link 1991). Thus, according to this view, the Mandarin *sān píngguǒ* 'three apple(s)' is semantically ill-formed for the same reason that the English *three water(s)* is: just as English *water* requires an explicit mensural classifier before it can be quantified, as in *three glasses/ounces/drops of water,* so Mandarin *píngguǒ* requires a sortal classifier before it can be successfully enumerated, as in *sān* **gè** *píngguǒ*. However, this explanation for the occurrence of classifiers is belied by a variety of additional facts (Gil 1996). Unlike English *water,* the Mandarin noun *píngguǒ* does indeed come with a conventional unit of enumeration, namely the individual round fruit that one can hold in one's hand, and therefore does not need a classifier in order to be counted. Evidence for this is provided by size and shape adjectives: whereas in English, constructions such as *big water* are bizarre, in Mandarin, constructions such as *dà píngguǒ* 'big apple' are syntactically well-formed, and understood in the same way as their English counterparts. A further argument against this explanation is provided by the existence of languages in which, as in Mandarin and other classifier languages, an NP consisting entirely of a bare noun may be understood as either mass or count, and as either singular or plural, but in which there are no numeral classifiers.

For example, in Tagalog, the noun phrase *mansana* 'apple' may be understood as 'apple' (mass), 'an apple' (singular), or 'some apples' (plural); however, it does not require a classifier in order to occur with a numeral, as in *tatlong mansana* 'three apples'. Thus, the occurrence of classifiers in classifier languages has no straightforward semantic explanation; in the absence of any better-motivated account, it must accordingly be viewed as just another arbitrary conventionalized fact about the grammars of individual languages.

Numeral Classifiers