

CLASSIFIERS AND STANDARDISATION:
SOME SOUTH AND SOUTH-EAST ASIAN COMPARISONS

R.K. Barz and A.V.N. Diller

INTRODUCTION

Areal and typological treatments of Asian classifier systems have led to proposals about syntactic development of classifier constructions within specific languages and about how they may spread among languages or language families. Proposals to date have mainly focused on syntactic issues, such as word order relative to head nouns. We suggest that for a more detailed understanding of classifier evolution and spread, sociolinguistic and stylistic issues need to be considered. Among classifier systems across southern Asia stylistic norms and attitudinal factors exert pressures both for and against classifier use on given speech levels. There is a general areal pattern: toward the west classifiers are normatively devalued, traditionally occur only in vernacular speech and even there are little proliferated; toward the east they are normatively valued, occur in standard languages as well as in spoken vernaculars and are typically proliferated.¹

Among Asian languages, areal patterning of numeral classifier systems has been recognised for some time. Emeneau (1956) placed importance on classifier distribution in his initial treatment of India as a 'linguistic area' and later (1965) he went on to document classifier constructions in a host of Asian languages to the east. Heston (1980) has recently extended them westward into Iranian languages. Jones (1970) has described South-East Asian classifiers in particular detail, tracing not only their constructions with numerals but their interactions with deictics and other modifiers.

Classifier constructions have been linked to other linguistic features, particularly to obligatory singular-plural marking. Sanchez (1973), Greenberg (1972, 1975) and others have established this (inverse) relationship and noticed the similarity between classifiers and units of measure as used with mass nouns in languages which otherwise have obligatory plural marking. Similarities between classifiers and generic nouns can also be drawn along syntactic lines (Krupa 1978).

Although syntactic approaches to understanding areal classifier phenomena are warranted, we suggest below that semantic and sociolinguistic analysis may contribute more explanatory detail as well, particularly when we turn to investigate diachronic issues and how they relate to language standardisation.

Below we document the use of classifiers in languages of the Hindi group in some detail. We find that Standard Hindi makes no use of classifiers at all (excluding measuring terms), however as one moves eastwards classifier use not only increases, but social attitudes toward classifiers shift as well. In Standard Bengali several classifiers are now normatively acceptable, but in an

David Bradley, ed. *Papers in South-East Asian linguistics* No.9:
Language policy, language planning and sociolinguistics in South-East Asia, 155-184. *Pacific Linguistics*, A-67, 1985.

archaic Sanskritised style of Bengali they do not occur. Assamese uses upwards of a dozen classifiers with no stylistic devaluation. This general areal shift in social attitude continues into South-East Asia where, e.g. in Thai and Lao, stylistic norms actually favour classifiers and their proliferation, the opposite of the situation within the Hindi group. We propose then that a gradient areal shift occurs across southern Asia with respect to normative attitudes taken towards classifiers, and that such attitudes are significant in tracing syntactic change.

CLASSIFIER USE IN HINDI VARIETIES

Rather than falling into a single cluster of closely-knit forms of speech, as do for example Bengali and Gujarati dialects, the so-called Hindi dialects constitute a linguistic continuum or language group consisting of five definite subgroups. Although the genetic interrelationship of these subgroups, which are known under the geographical designations of Rajasthani, Western Hindi, Eastern Hindi, Bihari, and Pahari, cannot be denied, the lack of a single clear-cut line of linguistic descent has led to a state of fluidity that makes the drawing of language and dialect borders difficult. It is this situation which has caused a blurring along the western edge of the Hindi group where its Rajasthani subgroup merges with Gujarati and on the eastern side where some of its Bihari languages and their dialects tend to slip toward the Bengali linguistic orbit. Such conditions have meant that culture and politics have played as great a role as have purely linguistic factors in determining the limits of the Hindi group of languages.

The most important of these non-linguistic influences has been the political and cultural hegemony over the entire Hindi-speaking area for the past 800 years at least of the western section of Uttar Pradesh in general and of Delhi in particular. It was around the city of Mathura just to the south of Delhi from which Brajbhāṣā spread all over northern India as the premier vehicle for written vernacular poetry on Hindu religious themes from the 15th century or earlier to the end of the 19th. The same period saw a parallel diffusion of Urdu, which is based on the dialect of Delhi itself, as the dominant language of Islamic and secular literature and polite urban culture and interregional commerce. These trends have culminated in the present century with the rise of Urdu and its sister Standard Hindi to the status of world languages. Since Brajbhāṣā, Urdu, and Standard Hindi are all three Western Hindi languages and share the same or very similar grammatical features, the prestige of that subgroup has become so great that linguistic forms which are characteristic of it have come to be the criteria for polished, sophisticated speech against which all other members of the Hindi language group are measured. By the same token, those grammatical traits which are not present in Western Hindi but are found in one or more of the other subgroups of the Hindi family tend to seem inelegant and rustic - even to those who use them in their mother tongue.

One such grammatical element is the numeral classifier, which occurs in the Eastern Hindi, Bihari, and Pahari subgroups, but not in Western Hindi or Rajasthani. Of all the Hindi family subgroups it is in the Bihari that the numeral classifier reaches the height of its vigour and shows its greatest variety of forms. As we see below, this confirms the basic areal pattern of eastward classifier acceptance, with neighbouring Bengali admitting classifiers in both its literary and colloquial forms. Three languages, Maithilī, Magahī,

and Bhojpurī, have long been accepted as independent members of the Bihari subgroup. In addition to these, on the basis of linguistic, cultural and even political grounds, Nāgpurī (also called Sadānī and Sadrī) has recently begun to be seen as an independent language rather than as a variety of Bhojpurī (Tivārī 1970:88-118). The following brief survey of the present condition of these languages will begin with Maithilī in the north-eastern quarter of Bihar and move through the other three on a course to the south and then to the west. Population figures for members of the Hindi group are notoriously difficult to estimate (Sinha 1973:123-124). Statistics below, based on 1971 population of the districts in the heartland of each language area (unless otherwise noted, M.I.B. 1978:409, 421, 438-439), should be taken as approximate.

Maithilī

Maithilī as the only member of the Bihari subgroup to have developed a literary tradition enjoys considerable prestige both in its own region and among scholars of Indian literature. Its position is considerably enhanced by the fact that the famous 15th century poet Vidyāpati chose it for the composition of some of his poetry (Mishra 1976:93-95). Like some of the other members of the Hindi group it has been revived as a vehicle for minor literary use in the present century. Nevertheless, it has no political status and remains primarily a village language. Maithilī is spoken mainly in the districts of Muzaffarpur, Sitamarhi, Vaishali, Darbhanga, Madhubani, Samastipur, and Saharsa in the Indian state of Bihar and in the neighbouring areas of southern Nepal. The number of Maithilī speakers at the present day could be reasonably put at about 20 million. On the north Maithilī is bordered by Nepali, on the east and south by Bengali, and on the south and west by Magahī and Bhojpurī. While as many as 12 different numeral classifiers have been listed for Maithilī (Jha 1958:353-354), the most common in written use today seems to be *goṭ* which is suffixed to numerical adjectives as with *ek* (*one*) in the following sentence (Deshmukh 1976:16):

Nāmdev ekgoṭ sāmāsik shabd, jakar arth bhel nāmhi dev achi.
Nāmdev one-CLF compound word, which-of meaning is name-INTENSIVE god is
Nāmdev is a compound word the meaning of which is 'the name itself is god'.

Contrary to modern usage, old texts like the poems of Vidyāpati do not contain the classifier and, even as late as the last decades of the 19th century, many learned Maithilī writers preferred to suppress its use. For example, between 1883 and 1887 George Grierson produced a set of grammatical sketches of Maithilī, Magahī - which he termed 'Magadhī' - and Bhojpurī. In order to obtain illustrative and comparative material he asked an educated speaker of one or more dialects of each language to translate the same set of fables into his particular form of speech. The results are informative. Not a single classifier is employed by the translators of the northern Maithilī (Grierson 1883:30-38), mixed southern Maithilī-Bengali (Grierson 1887:82-89), and mixed southern Maithilī-Magadhī (Grierson 1886:88-94) versions while classifiers are used normally in the mixed Maithilī-Bengali (Grierson 1887:80-86), southern Maithilī (Grierson 1885:94-101), and mixed Maithilī-Bhojpurī (Grierson 1884:92-98) translations. Since it is unlikely that classifiers were not used at that time in speech uniformly all over the Maithilī area, the explanation seems to be that some translators were writing in a traditional literary style in which classifiers were omitted.

Magahī

Despite its being the speech of the area at the political centre of the state of Bihar - the districts of Patna, Gaya, Nalanda, Aurangabad, Nawadah, Monghyr, Hazaribagh, Girdih, and Dhandad - and the home language of some 16 million people, Magahī is the least developed of the four Bihari languages. It has a fine tradition of folklore, but possesses neither an old nor a modern written literature.

As in Bhojpurī, with which it shares a long linguistic border on the west, the usual numeral classifiers in Magahī seem to be *ṭho* and *go* placed between the numerical adjective and the noun it modifies. Again as in Bhojpurī a composite form, *ego*, results when *go* is used with *ek*. The following examples are taken from folk tales (Aryānī 1965, pt.1:2,13):

ek ṭho rājā halā.
one CLF king was
There was a king.

ego jangal mē ego bāgh raha halai.
one-CLF wilderness in one-CLF tiger used to live
A tiger used to live in a wilderness.

Since Magahī has never been a literary language it is not strange that Grierson's texts show no evidence of a suppression of the classifier, as for example (Grierson 1883:84):

ego cṭlh apanā ṭhor me ēk ghōghā lēlē halai.
one-CLF kite his-own mouth in one cockle had taken
A kite had taken a cockle in his mouth.

Nāgpurī

To the south of Magahī and Bhojpurī, to the east of the Chattīsgarhī language of the Eastern Hindi subgroup, and to the north of the Oriya-speaking area, lies the region of the third Bihari language, Nāgpurī. Unlike most other members of the Hindi group Nāgpurī does not have a large population of native speakers. This is due to the fact that Nāgpurī functions as a lingua franca or bazaar language for intercommunication among speakers of the various non-Indo-Aryan languages of southern Bihar, among which Mundari and Kurukh (Oraon) are the most important, and between those people and immigrants from other parts of Bihar. Thus, though Nāgpurī is spoken as a second language by a great many people over the districts of Palamau (Palāmū), Ranchi, and Singhbhum, in the 1961 census only 459,143 people listed Nāgpurī as their mother tongue (Jordan-Horstmann 1969: 8-11). While Nāgpurī until the present century was never thought worthy of literary development, in the past couple of decades a small but dedicated group of writers has begun to use it for short stories, plays, poetry, and - especially among Christians - religious literature. Nāgpurī has several numeral classifiers. *Go*, *goṭ*, *goṛ* and *ṭho* are all used freely while *mūr* is restricted to use in counting cattle and *khār* for pieces of cloth (Nowrangi 1956:32-33). This sentence illustrates the use of the numeral classifier in Nāgpurī (Navrangī 1965:51):

mor dui go beṭāman ahai.
my two CLF sons are
I have two sons.