EARLY MY/TB LOAN RELATIONSHIPS
Paul K. Benedict

The Miao-Yao (MY) languages appear to have split off from the mainland bloc of Austro-Tal (AT) languages at an early period, moving to the north and west and there coming into contact with the Sino-Tibetan (ST) languages. To make use of an analogy from geomorphology, they came to be positioned at the 'grinding edges' of the vast AT and ST tectonic plates, along a line of maximum anticipated activity. In this case the activity involved the transformation of MY into a monosyllabic, fully tonal language family, as earlier described by the writer (1975 - hereafter ATL, with many of the details yet to be explored. As regards the monosyllabism, this is a feature of both Tibeto-Burman (TB) and Sinitic (Chinese, Bai) and must be set up for the parental proto-language (PST), hence no clues as to the precise source of the influences are provided. The tones are another matter, however; at an early period (2nd/1st mill. B. C.) Chinese converted the two-tone (or two-accent) *A ~ *B system of PST into a three-tone (or three-accent) system with the addition of a third, sandhi tone (or accent) *C and a system of precisely this kind must be reconstructed for the parent MY language (PMY), with one-to-one tonal correspondences in the early loanwords for such cultural items as 'horse' and 'charcoal'. This is an example of 'direct' diffusion, in contrast to the 'stimulus' diffusion found in Mon-Khmer (e.g. Riang, Khamu) and even in Austronesian (Huihui, the Chamic language spoken on Hainan); the tonal system itself was borrowed, not simply the 'idea of tones', with the loanwords serving as the bearers of tone. The process undoubtedly began in a highly selective manner, with later spread throughout the language; cf. the situation in Tin, a Mon-Khmer language now on its way to becoming fully tonal, which has borrowed Thai (Siamese) numerals along with the tones (see Benedict 1984: 67).

This sort of linguistic happening cannot occur at a distance, inasmuch as it requires a virtual cultural 'flooding', with extensive diglossia on the part of the 'submerged' population (as in the case of Tin). The early (Archaic level) Chinese/MY loans, as currently analyzed (Benedict 1986), point to exactly this kind of prolonged, intimate relationship, thus dovetailing with the tonal evidence. The picture is complicated, however, by the fact that the bulk of Chinese loans into MY reflect early dialect(s) that are distinct from the 'standard' Archaic, notably in the retention of PST final *-a, where the latter shows a regular shift to -o. These loans also commonly show evidence of Chinese 'processing', reflecting manifold prefix + initial shifts of the type described in an earlier paper (1976) by the writer. Thus, the picture is exceedingly complex, even under the best of circumstances, yet all would make good sense historically were it not for the following: a TB rather than Chinese source must be recognized for the basic cardinal numerals of MY as well as for the 'core' (basic) lexical items: 'sun/day' ~ 'moon/month', along with a number of other items rather less basic in nature. This paper presents the relevant linguistic data, attempts to delineate (if not identify) the donor source or sources (DMY or DMY's) and,
finally, offers some preliminary speculation as to how this strange linguistic state of affairs is likely to have come about.

The numerals of MY, like those of the Kadai languages and even Austronesian (Huihui), nicely exemplify selective lexical retention (SLR - see Benedict 1983b), with a primary 'rule' that cardinals are better retained than ordinals and a secondary 'rule' that the lowest numerals are better retained than the higher (it should be noted here that the Swadesh basic 100-word list includes only 'one' and 'two'). In the case of the MY numeral system, only one loanword managed to break into the native 'one' through 'three' alignment and the intrusion was relatively late: Middle Chinese (MC) ʔ1 ʔet 'one' > Yao (Mien only) *yet1. The MY numerals above 'ten' are also of Chinese origin but those from 'four' through 'ten' appear to have been borrowed as a set from a TB source (or sources). Both Shafer (1964) and Downer (1971) took special note of the resemblances shown by these numerals while the writer (1976) has analyzed them at some length, along with the early Chinese loans to Tai. Contra the view expressed there, it now seems evident (see fn. 1) that 'four' also belongs in this set of early loans from TB, with the further strong likelihood that 'five' must be included as well, i.e. these numerals were borrowed as a set. As shown in the following listing, the corresponding Chinese numerals developed in different ways (see the above-cited paper), precluding the possibility of their having served as sources for the MY numerals (see Benedict 1972 - hereafter STC - for the TB numerals):

'four': PTB *b-lɔy; PMY *p leiA; MC ʔ4 s\. PMY regularly has -ei for earlier (PAT-level) final -i, with -i as an alternative reconstruction, hence the indicated DMY form is *pliA, with -*i for final *-ɔy (as in most TB languages) along with unvoicing of the prefix.

'five': PTB *r-ŋa ~ (secondary) *b-ŋa; PMY *praA; MC ʔŋ uo. PTB prefixed *r- is maintained in Old Kuki but replaced in Written Tibetan (WT) by ʔ-, apparently through influence from PTB *lak 'arm/hand'; general replacement by prefixed *b- through influence from the preceding numeral (see above), with occasional preemption of initial *ŋ-. The indicated DMY form is *praA, with the secondary *b- (> p-) preceding the *r- rather than replacing it, followed by preemption of the initial and unvoicing of the prefix, as in 'four'.

'six': PTB *d-ruk ~ *k-ruk; PMY *tru?; MC ʔ liuk. The PMY initial *tr- has been reconstructed on the basis of the Na-e evidence (see Benedict: forthcoming); the final *?- for earlier *-k is a MY feature, hence the indicated DMY form is *truk, with unvoicing of the prefix as in 'four' and 'five'.

'seven': PTB *snis; PMY *zal[a]C; MC ʔ tsi ꞉ t. Contra STC, the PTB *s- stands for the first part of the *s- cluster rather than for prefixed *s-, with Stau exhibiting a distinction between /zal/ 'seven' (secondary voicing and palatalization with loss of final *-s) and /sfi/ 'day', from *s-nəyA (see below). The nasal element of the MY root is maintained only in Yao: Mun (ǎi) but it yielded typical secondary nasalization of the final in Miao, with some Western dialects reflecting secondary unvoicing (*zǎiaC > *siəjaC). The initial of the DMY form can be reconstructed as *z[n,ŋ]-, with secondary voicing as in Stau; the ambiguity results from the fact that
secondary palatalization is characteristic of MY, tying in with the same feature in Chinese, probably reflecting an ancient Sprachbund. The final of the DMY form could hardly have been *i, with loss of *s (as in Stau), since this would have yielded PMY *ei (cf. 'four'); PMY does have medial *ia- for medial *i-, however (see 'ten' and 'year', below), hence the indicated reconstruction of the DMY final is *is, with the vocalic shift preceding the eventual loss of *s (lacking in PMY).

'eight': PTB *-ryåt; PMY *hayt; MC /pøåt/. The earliest ST prefixation pattern, reflected in WT (bryad < *b-g-ryat) and Jingpho (møtsåt) as well as in Chinese (apparently unvoiced by an original prefixed *s-) can be set up as *b-g-; but within TB the Kuki-Naga forms reflect prefixed *d- (< *d-råt) while those of Burmese-Lolo reflect prefixed *s- > Written Burmese (WB) hraç (< *hrayt), with /hyat/ appearing in the inscriptions, exactly matching the PMY root. The indicated DMY form is *h[ry,y]at, with *ry- a possibility in view of the fact that PMY lacked this cluster and may well have substituted *y-, paralleling the Burmese development.

'nine': PTB *d-køw ~ *d-gøw; PMY *g[you]A; MC /kïu/. The initial *g- form, represented notably in WT (dgú), shows secondary voicing after the prefix, which was replaced by *s- in some languages (Jingpho, Bodo-Garo). The PMY palatalized initial is generally reflected by palatais or dentals (but Na-e kò); the reconstruction of the final is provisional, with the Yao forms apparently influenced by the similar ordinal forms (< Chinese). PMY regularly has final *ou for earlier (PAT-level) *u, paralleling *ei for earlier *i (see 'four', above), hence the indicated DMY form here is *[prefix+][g,gy]u, with ambiguity arising from the palatalization (see the discussion under 'seven', above). The initial voicing points to earlier prefixed *d-, as in WT, with later shift to *t- (cf. 'six', above).

'ten': PTB *gilp; PMY *gylap; MC /ziop/. The PMY initial palatalization appears to have been variable, with Yao forms perhaps influenced by competing forms of Chinese origin. As in 'seven', the PMY medial *-ia- is to be considered secondary, with *gilp indicated for the DMY form.

Reconstructed tones (*A or *B) have not been cited for the three relevant PTB numeral roots ('four', 'five' and 'nine') in view of the widespread variation shown by the modern forms, including tonal 'form classes'. e.g. WB all < tone *B; Trung (Nungish) all < tone *A, for these three numerals as well as for 'three' (PTB *g-sum). PMY tone *A for the trio indicates that DMY had a 'form class' like that of Trung. The fact that PMY has tone *C for 'seven' supports the final *s reconstruction since this tone has a sandhi origin, reflecting syllable-final features.

To sum up, the DMY numerals were probably as follows:

<table>
<thead>
<tr>
<th></th>
<th>DMY</th>
<th></th>
<th>DMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>four</td>
<td>*plíaA</td>
<td>eight</td>
<td>*hryat</td>
</tr>
<tr>
<td>five</td>
<td>*praA</td>
<td>nine</td>
<td>*t-guA</td>
</tr>
<tr>
<td>six</td>
<td>*truk</td>
<td>ten</td>
<td>*gilp</td>
</tr>
<tr>
<td>seven</td>
<td>*znis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The DMY phonology is featured by the unvoicing of stop prefixes (the *t- of 'nine' only by inference), along with the *s-r- > *hr-shift, contrasting
with secondary voicing of the *sn- cluster; also the parallel final *-ey > *-i and *-aw > *-u shifts. Although only one numeral root (*gip 'ten') remained without change, none of the DMY forms appears very unusual for TB, with the exception of that for 'five', and even here there are parallels in the modern languages to the initial-preemption involved.

Three other lexical items, all at least partially of calendrical type, have long attracted the attention of comparativists because of the obvious MY/TB resemblances. As reported in the above-cited article (Benedict 1976), there is a competing AT etymology in the case of one of these items ('moon/month') but this must now be discarded, especially in view of the additional evidence available here (see below). As in the case of the numerals, the Chinese cognates exhibit different lines of development, effectively precluding them as possible loan sources:

'sun/day': PTB *nayA 'sun' ~ *s-nayA 'day'; PMY *hnɔiA (often prefixed) 'sun/day'; MC ʰnʒʰet 'id.'. The 'sun' vs. 'day' distinction can be established at the PTB level; cf. Stau sāi 'day' (above); Jingpho ʃəŋf 'id.'; WB ne (< *nayA) 'sun' ~ ne' (< *s-nayA) 'day', but has been lost in some TB languages, e.g. Lushai has /nɪ/ for both 'sun' and 'day'. Chinese also has lost the distinction: ʰnʒʰet < *s-nʒʰet (Min evidence for the prefixation), with secondary final -t after the prefix (paralleling the WB glottalization; see Benedict 1983a). In MY the 'sun' vs. 'day' distinction is often made through prefixation but basically the PMY root shows a loss here, of the Chinese rather than Lushai type, i.e. it is the *s- prefixed rather than the unprefixed form that has prevailed. On the basis of this MY loan evidence one can posit a similar loss of distinction in the donor language, although it is possible that MY innovated here. The indicated *hn- for DMY, from an earlier (PTB) *s-n-, contrasts with *zn- from a *sn- cluster in 'seven', precisely as in Stau (see above). The final presents a problem, however, inasmuch as in 'four' PTB final *-ey is represented by PMY *-ei, regularly from an earlier *-i, whereas in this root it is represented by PMY *-o, very likely standing for *-ey itself. At least three possible solutions present themselves: (a) the final reflex was conditioned by the initial (*pl- vs. *hn-) (b) PMY lacked the specific *hn*i combination (Mien, at least, appears to lack it) and substituted *hnɔi (3) a somewhat different DMY, perhaps a dialect of the 'standard' DMY, which retained final *-ey, was the source in the case of this item. Of these three possible solutions (a) seems the least likely and (c) the most likely.

'moon/month': PTB *(s-)(g-){l}aA: PMY *hlaC; MC ʰsək [GSR 769a] 'new moon', from *s-glāk (see Benedict 1976:189 for the phonetic element here), with prefixed *s- yielding secondary -k after the low vowel, in contrast to the final -t after high front vowel that appears in the preceding root (see Benedict 1983a). The Garo cognate: ja (< *sglāA) is also glossed as 'spirit/apparition/phantom'; WT has the doublet zla (< *sglāA) 'moon' ~ hla (< *sla) 'the gods'; Jingpho has only šətə (< *sglāA). Chang-Tangs (STC: Konyak group) only *glāA and WB only la' (< *s-laA), all in the meaning 'moon' (~ 'month'), but Lushai parallels WT in showing a doublet, in this case the product of regular tonal changes (see Benedict 1983a): thlə (< *s-glāA) 'moon' ~ thlə (< *sglāA) 'spirit/soul/one's double'. This all points to an