PALAUNGIC VOWELS
IN MON-KHMER PERSPECTIVE

G. Diffloth

Thirteen years ago, H.L. Shorto pointed to the vowel system of Proto-Mon-Khmer as being the 'crux' in the historical phonology of this family (Shorto 1976). His assessment remains as valid today as it was then, even though some advances in reconstruction have been made; our data base has considerably expanded and improved, but the mirage of Proto-Mon-Khmer vowels continues to recede, even as we penetrate further into the past.

The solution proposed then: vowel variation in the proto-language, is consistent with certain facts which can be observed in several Mon-Khmer languages spoken today. In Bahnar, Sre, Khmu and Semai, to select but a few, whole families of Expressives (Diffloth 1979; in press) are often built on vowel permutations, and such Expressives occasionally find their way into the prosaic (Non-Expressive) vocabulary; conversely, prosaic words often serve as a starting point for building families of Expressives which differ, for example, only by their major vowel. This has surely contributed to the formation of word-families such as those identified in Shorto (1973). This explanation, however, has its limits: presumably, these processes would have affected a word here and a word there, at different times, but it is difficult to see how it could have pervaded the thousands of items which form the non-expressive lexicon of one language, not to speak of an entire family. Other factors are needed in order to account for the numerous vowel correspondences which have been detected so far.

For example, it may well be that the Proto-Mon-Khmer vowel system reconstructed until now, although sizable, is not rich enough for the purpose, and that we need to expand it with some additional phonological dimensions.

Tone has been practically ruled out for Proto-Mon-Khmer since the simple tone systems of Bùlăng and Riang, and the tone-cum-register system of Nyah Kur (Diffloth 1980, 1984) can all be explained as innovations; but the newly recorded Angkuic languages U and Man Met (see below) have four-tone and six-tone systems respectively, the origins of which remain partly unknown for the moment. Then again, Haudricourt's account of Vietnamese (VN) tonogenesis has generally been accepted, but it leaves out, as tonally irregular or unexplained, a large number of words which do belong to the indigenous Mon-Khmer stratum of the language. And the recently discovered Palyu language, called Lai in Chinese, also has six tones which may, or may not, turn out to be

recalcitrant—Palyu is apparently Mon-Khmer (Liang 1986, Benedict, in press), but its position in the family is still undecided.

Register is a better candidate. This is, typologically, a well-established feature of Mon-Khmer languages (Huffman 1976). The general consensus is that Register is a relatively recent phenomenon, and Shorto, accordingly, does not reconstruct it for Proto-Mon-Khmer. Ferlus (1979) desribed all Mon-Khmer register systems found until then as being the result of one type of evolution: devoicing of initial consonants. This explanation has long been accepted in the case of Spoken Mon (Blagden 1910) and of Modern Khmer (which ironically has now lost phonation-type distinctions), and it does account for the registers of several other newly recorded languages (e.g. Kuy, Bruu, Phalok). But it is inadequate in certain other languages: the Pacoh register system, admittedly an innovation, has nothing to do with the process of devoicing which has independently taken place in this language. It is also inapplicable to the North Bahnaric languages where no devoicing has taken place, except in Sedang. In Pacoh, the genesis of register is due to changes in vowel quality, namely, the fronting or backing of certain proto-central vowels (Diffloth 1982), and a similar innovation has apparently also taken place in North Bahnaric (Diffloth 1983). So, we do not have yet a case of reconstructing register as being ancient in Mon-Khmer.[1]

However, the Pearic branch might force us to do that: recently, Huffman (1985) has shown that Chong, a language of the Pearic branch, had a Clear vs. Breathy distinction, criss-crossing a Plain vs. Glottalised contrast, giving rise in effect to a four-register system. Theraphan (this volume) describes in detail the complex bundle of phonetic features these four registers contain. This phenomenon has no historical explanation, and Headley himself (1985) has abandoned to the sagacity of future historical linguists any attempt in this regard.

Gage (1985) has pointed out that certain unexpected occurrences of the *sắc* tone in Vietnamese seem to find an echo in the register system of Pearic. The tonogenesis of Vietnamese requires that the *sắc* tone occur with final proto-stops, and indeed cannot explain the tones of many VN words which have excellent Mon-Khmer etymologies, such as: *bốn* 'four', *chín* 'cooked', *gió* 'wind', or: *ngái* 'far'. Cognates to all four of these words happen to have glottalisation in Pearic. In Chong, as I have recorded it, the first three have the 'tight' register:[2] /phɔ̰ɔn/ 'four', /chḭin/ 'cooked', /kəya̰ay/ 'wind'; the fourth word has a 'breathy-creaky' register[3]: /ŋa̤ay/ 'far'. Other examples can be found, e.g. VN: *cám*, Chong /kəŋa̰am/ 'rice-husk', but there are counter-examples as well, e.g.: VN: *chim*, Chong /chḭim/ 'bird'. Since the Pearic and the Việt-Mường branches are only distantly related, the implications of this fact could go back directly to Proto-Mon-Khmer.

---

1. Smith's opinion to this effect (Smith 1972) was not based on the establishment of sound correspondences, but on statistical tendencies within a very small set of possible Mon-Khmer cognates, which a more thorough comparison does not confirm.
2. Clear voice plus glottalisation in Huffman's (1985) analysis.
3. Huffman's (1985) breathy voice plus glottalisation.

There are also less exotic vowel features which have not been fully used in Mon-Khmer reconstruction. Diphthong systems can be much richer than the simple *iə and *uə usually proposed; I have reconstructed Proto-Katuic with five proto-diphthongs (Diffloth 1982), and Nancowry Nicobar (Radakrishnan 1981:25) is described even today as also having five diphthongs: /iá/, /ía/, uá/, /úa/ and /úa/, which seem to correspond with what we can reconstruct for Proto-Aslian.

Some of these phonological features may have to be reconstructed back to Proto-Mon-Khmer, and could well explain a number of Shorto's variations as being regular outcomes of a much richer proto-vowel system. But then, the number of proposed Proto-Mon-Khmer etyma becomes a relevant issue, and what has been published so far can be said to represent only a sample.

In this paper, I will not explore these possibilities, but only prepare the comparative ground to do so; I will try to clarify some points in the history of vowel systems in the Palaungic branch, where recently recorded material allows us to make systematic reconstructions. This may seem at first to be somewhat irrelevant: if Waic and Palaung are notable for one thing, it is precisely the poverty of their vowel systems. The old vowel-length contrast was already lost in Proto-Waic, and the best source of information on Milne's Palaung (=Ta-ang)[4] appears not to have a phonemic contrast of this kind.

But the Palaungic branch has an important role to play in reconstruction: it belongs to a distinct division of the family, the Northern Division, and it provides us with an independent testimony for the reconstruction of Proto-Mon-Khmer vowels. Besides, as I will try to show, Palaungic vowel systems are not as poor as they first seem to be.

1. *Proto-Waic*

The term 'Waic' covers (1) several Wa languages, e.g. Paraok, Avüa', La (Zhōu & Yán 1984) and their dialects; (2) the Phalok language,[5] formerly referred to as Khalo or Mae Rim Lawa (Flatz 1970); (3) Lawa and its dialects (Mitani 1972); and (4) the Bulang-Phang complex with its many dialects[6] (Diffloth 1980). Certain Waic languages, Lawa and Paraok in particular, currently have rich and complicated vowel systems[7] but this is

---

15

due, in part, to the influx of Tai borrowings[8] and in part to recent processes of vowel warp, conditioned by Registers and final consonants. Only nine proto-vowels are needed at the Proto-Waic stage[9]

Proto-Waic Vowel system

i	ɨ	
e	ɤ	o
ɛ		ɔ
a	ɒ	

This maximum system is found with most final consonants, but there are certain distributional gaps; for example, with final *-ʔ, only eight proto-vowels are found (all the above except *ɒ); with final *-h, only seven are found (*ɤ and *ɒ are excluded); and there are no open final syllables in Proto-Waic.

2. *Proto-Palaung-Rumai*

The term 'Palaung-Rumai' also covers several languages, the best known of which is Ta-ang, i.e. the Palaung of Nam Hsan described by Milne (1931). The Rumai language and its dialects, also belongs here[10], as well as the dialects of Riang;[11] it also includes another distinct group sometimes called 'Palê', which contains at least Da-ang and Na-ang; and several other languages, like Ka-ang and Ra-ang[12]. Other Palaung-Rumai languages surely await description in Burma or Yunnan, and they may or may not belong to one of the seven groups mentioned here.

There is considerable diversity within Palaung-Rumai, but this is not the place to present all the phonological innovations which can now be documented. Mitani (1977, 1979) has already reconstructed the Palaung-

---

8. Lawa has borrowed from Northern Thai and more recently from Standard Thai; Paraok has borrowed from Shan (referred to in China as 'Déhóng Dǎi'), Bùlǎng has borrowed from Lü (referred to in China as 'Xīshūangbǎnnà Dǎi', or Xī Dǎi for short).
9. This reconstruction was presented in (Diffloth 1980); since then, a Paraok-Chinese dictionary has been published (Yán *et al.* 1981), and, in 1984, I was fortunate to collect Waic linguistic material in China, with the help of Zhōu, Z.-Z., under another NSF grant entitled 'Comparison of the Mon-Khmer languages of China with other languages of the Mon-Khmer family'; with this new material, the number of reconstructed Proto-Waic words has now more than doubled, and the reconstruction of a few etyma given in Diffloth (1980) has been modified; this new information confirmed the nature of the Proto-Waic system I had reconstructed in 1980 with the help of Y. Mitani.
10. All the Rumai examples quoted here were collected in May 1981, with the help of a family of Rumai emigrants living in Chiang Mai at the time.
11. See note 4 above.
12. The information on Na-ang included here was kindly given to me by Yán, Q.-X. as part of a research programme in China (see n.9). She is the author of a valuable sketch on the Běnglóng language(s) (Yán 1983). I collected myself the Ka-ang data from a native speaker in Kūnmíng, in the course of the same research programme. The Da-ang and Ra-ang data were collected in 1981 (see n.5), and 1984, during my stays in Thailand. It was not possible for me to determine with precision the geographic spread of these languages, especially for those spoken in the Shan States (Burma) where there seems to be a lot of small-scale migrations. For China, Svantesson *et al.* (1981), Yán (1983), Zhōu & Yán (1983, 1984) provide geographic and demographic information.

16