Vietnamese as a Mon-Khmer language

Gérard Diffloth
Cornell University

During roughly the last millennium and a half, the Vietnamese language has had, it seems, a more dramatic history than other languages in Southeast Asia. Current scholarship shows that the language has changed from having the typical sesquisyllabic word-structure of an ordinary Mon-Khmer language with a great variety of rimes, no tones, and complex initials possibly containing affixes, to being strictly monosyllabic, having few final consonants, six (or five) tones and hardly any initial clusters.

Notions about the history of the language also have had their own, equally spectacular history. British scholars like Logan and Forbes were the first to upset the traditional view, echoed by Taberd (1838), that Vietnamese was simply an offshoot of Chinese. They proposed instead that VN was related to Mon, to Khmer and to several other languages of the Indochinese peninsula. Their idea of a Mon-Anam family (Logan 1856, Forbes 1878, 1881) was looked at with a good deal of suspicion, partly because Vietnamese had tones, and tone languages were considered to be radically different from others. It took another century, and Haudricourt's insistence, for the traditional view to be overturned and the original insight of Logan and Forbes' to gain credibility.

The history of the emergence of tone thus plays, and will surely continue to play, a prominent role in studies of VN historical phonology; I will offer some
views on the problem in another paper. But now looms the larger and more difficult question of finding a place for Vietnamese in the Austroasiatic family, or more precisely for the Vietic branch\(^1\) away from, or among, the various known branches of the Mon-Khmer family.

Haudricourt was elusive on the matter. But Ferlus (1991) suggested "de solides affinités" between Katuic, Bahnaric and Việt-mương\(^2\) without providing any evidence. He was presumably thinking about the lexicon, and I would share that impression, especially with regards to Vietic and Katuic.

Respect for strict methodology would lead us to reject the question as being undecidable for the moment. The current classification of Mon-Khmer into eleven branches (Thomas and Headly 1970, Diffloth 1974), though still valid, rests to some extent on now generally discredited lexicostatistical techniques, and no one has yet proposed a set of innovations distinguishing each branch of Mon-Khmer from the rest. The higher grouping into three major divisions: North MK, East MK, South MK, (Diffloth 1979) also remains little more than a guess. But fresh linguistic data\(^3\) collected on the lesser-known languages of the Vietic branch may allow us to break new ground.

The first phonological correspondence I will now document serves to identify all and only the languages which belong to the Vietic branch; it probably represents a Proto-Vietic innovation. The other correspondence sets apart all Vietic and all Katuic languages together, and no others in the family. Depending on reconstructions, this could indicate that Vietnamese, far from being a remote offshoot of the Austroasiatic phylum, as is sometimes suggested, belongs not simply to the Mon-Khmer family, but in fact to a subdivision within it. We could say that Vietnamese, and with it the Vietic branch, is a member of a Northeastern subdivision of Mon-Khmer consisting of two branches,
Vietic and Katuic; this in turn would be a part of the larger Eastern Mon-Khmer division which also includes Khmeric, Bahnaric and perhaps Pamic, the other two major divisions of the family being Northern and Southern Mon-Khmer.

This proposal entails, I realize, some drastic revisions of commonly held views on Vietnamese proto-history, cultural links, and early movements, but it is based on the most rigorous arguments available at present, which unfortunately is not to say very much. And other scenarios remain possible.

The Proto-Vietic *k-/ʔ- correspondence.

There are several Mon-Khmer, or even Austroasiatic, etyma which begin with a glottal stop in all branches of the family, but are found, surprisingly, with an initial *k- in every known Vietic language. Only six or seven examples have been found so far, but several of them are often recorded and well documented words. Cognates from all branches of Mon-Khmer are quoted here and give an idea of how Vietic fits, or does not fit, in the rest of the family.

"fire, firewood"

Vietic:
  - VN: cúi, ; Mường Khoi: kuy4, Mường Khên: kuy45
  - Thavung⁴: kuy⁴
  - Maieng: kuy⁴, Ruc: kurh, Arem (Ferlus 1991): kuh
  - Tum: ku: c31

The breathy register of Thavung does not necessarily indicate a proto-voiced initial (contra Ferlus, 1974) but is conditioned in part by the quality of the proto-vowel. The Vietic forms above indicate a Proto-Vietic *kuː:s, agreeing with Ferlus (1991).

In the rest of MK, the etymon has been identified for a long time (Schmidt 1905:22, Skeat and Blagden 1936:...
F124). It is absent from Munda and the whole Northern-MK division except for Western Khmuic: Mal ?ο:ŋ, Prai: ?ο:t "fire"; otherwise, it is almost omnipresent in Eastern MK:

Khmer: ?ơh: "firewood"


Bahnaric has a cognate in only one language: Sre: ?ơ:s "fire", otherwise a similar sounding *ʔο:n has replaced this etymon.

It is omnipresent in the Southern division:

Monic: Nyah Kur: ʔyaŋ, Proto-Mon *ʔơh (Difflloth 1984:128) "firewood"

Aslian: Proto-Semai *ʔơ:s, Temiar: ʔơ:s, Jah Hut: ʔơh, Semelai: ʔus "fire"

Nicobarese: Central: ʔơh, Teresa-Bompaka: ʔsh "fire"

Note that all non-Vietic items have initial *ʔ-, and that the semantic extension from "fire" to "firewood", also found in Vietic, is older than the split between East and South Mon-Khmer.

"(part of the body) to swell"

Vietic:

- Thavung: kɛh
- Maleng: kəl, Ruc: kɛ:i, kɛ:ɛ
- Tum: kɛ:ɛ

The word is not attested in Việt or Mường, but the correspondence: Thavung -h, Tum -ɛ indicates a final *-s, and a probable Proto-Vietic form *kɛ:s.

In the rest of MK, his etymon, absent from Khmeric and Monic, has received little attention so far, but goes back to Proto-MK. It is attested in the Eastern MK division:

Pearic: Chong: ʔat, Song: ʔah

Bahnaric: Tarieng, Sre: ʔas "to swell", Brao: ʔɛh "abcess"