



**HAL**  
open science

## A layer of Dongsonian vocabulary in Vietnamese

Michel Ferlus

► **To cite this version:**

Michel Ferlus. A layer of Dongsonian vocabulary in Vietnamese. Journal of the Southeast Asian Linguistics Society, 2009, 1, pp.95-108. halshs-00932218v1

**HAL Id: halshs-00932218**

**<https://shs.hal.science/halshs-00932218v1>**

Submitted on 16 Jan 2014 (v1), last revised 8 Sep 2014 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A LAYER OF DONGSONIAN VOCABULARY IN VIETNAMESE

Michel Ferlus  
*Independent Researcher*  
<jrmferlus@orange.fr>

## 0 Abstract

The present paper aims at demonstrating by means of linguistic evidence how the pestle used to husk rice was invented by the Dongsonians, the ancestors of the Vietnamese. That innovation spread in Southeast Asia as far as India, through the Austroasiatic continuum.<sup>1</sup>

## 1 Background

The place of the Vietnamese language (or Viet in its shortened form) in the Asian phylogeny has varied significantly since the first research on the topic was carried out. After being classified among the Chinese or the Tai-Kadai languages, it was finally bound to the Mon-Khmer family [for historical insight see Alves 2006] and more widely to the Austroasiatic family. The discovery (scientifically speaking) of conservative languages related to Vietnamese, made it possible to elaborate a Viet-Muong group (henceforth VM), or Vietic, and to reconstruct a Proto Viet-Muong (henceforth PVM).

Some authors shed light on the close lexical relationship between the VM and the Katuic groups. Historically, it is highly probable that the VM group is the result of an ancient expansion of a form of Katuic coming from Northeast Thailand, which would have covered an Austroasiatic substratum localized in the North Vietnam (corresponding to the ancient *Giao Chi* and *Cru Chân*).

Vietnamese and Mường, its offshoot, include vocabulary and phonetic features which differentiate them from other languages of the same group. The subject covered here relates precisely to Vietnamese vocabulary with the initial *x*- supposed to belong to that particular substratum.

## 2 Languages and dialects of the Viet-Muong (Vietic) group

A simple and practical classification of the VM is presented below.

- 1- *Maleng* : Maleng proper, Malang, Pakatan, Măliêng, Maleng Brô, Kha Phong (or Maleng Kari).
- 2- *Arem* : Arem (or Cmrau/Cmbrau).
- 3- *Chút* : Sách (or Chút, or Salang), Rục.
- 4- *Aheu* : Thavung, Phôn Soung, Sô (or Sô Thavung).
- 5- *Pong* : Pong (or Phong), Toum, Liha, Đan-lai.
- 6- *Thổ* : Làng Lỡ, Cuối Châm, Mọn.
- 7- *Mường* : Mường (or Mọl/Mọn); comprises many dialects of which, Mường Đầm, Mường Khói and Mường Tân Phong and Nguồn.
- 8- *Viet* : written standard Vietnamese and its dialects.

---

<sup>1</sup> I cordially thank Frédéric Pain (Catholic University in Leuven, Belgium), a linguist specialist in Southeast Asia, who read the text over with the greatest attention.

### 3 PVM initial consonants: an outline

PVM comprised monosyllables CV(C) and sesquisyllables C-CV(C). The PVM phonemes (bold) and their modern Vietnamese reflexes (in italic and *quốc ngữ* spelling) are tabled below.

|                                   |                                   |                            |                          |                                   |                      |
|-----------------------------------|-----------------------------------|----------------------------|--------------------------|-----------------------------------|----------------------|
| <b>p<sup>h</sup></b><br><i>ph</i> | <b>t<sup>h</sup></b><br><i>th</i> | <b>s</b><br><i>t~r</i>     |                          | <b>k<sup>h</sup></b><br><i>kh</i> | <b>h</b><br><i>h</i> |
| <b>p b</b><br><i>b~v</i>          | <b>t d</b><br><i>đ~d</i>          | <b>c ʃ</b><br><i>ch~gi</i> | <b>tʃ</b><br><i>x~gi</i> | <b>k g</b><br><i>c/k~g/gh</i>     | <b>ʔ</b><br><i>#</i> |
| <b>ɓ</b><br><i>m</i>              | <b>ɗ</b><br><i>n</i>              | <b>f</b><br><i>nh</i>      |                          |                                   |                      |
| <b>m</b><br><i>m</i>              | <b>n</b><br><i>n</i>              | <b>ɲ</b><br><i>nh</i>      |                          | <b>ŋ</b><br><i>ng/ngh</i>         |                      |
| <b>v</b><br><i>v</i>              |                                   | <b>j</b><br><i>d</i>       |                          |                                   |                      |
|                                   | <b>r</b><br><i>r</i>              | <b>l</b><br><i>l</i>       |                          |                                   |                      |

The aspirated plosives **p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>** are not frequent and must have evolved from clusters of the type /*plos. + h*/.

Obstruents **p-b, t-d, c-ʃ, s, tʃ** and **k-g** underwent two types of phonetic changes, (i) normal changes of initials in monosyllables, (ii) spirantization of medials in sesquisyllables [Ferlus 1982]. For example, the pair of initials **p-b** is on the whole represented now by *b~v* (*b* in monosyllables and *v* in ancient sesquisyllables). It must be noticed that, in the 17th century, *v* was rendered by *ɸ/bɸ* in Alexandre de Rhodes' dictionary [1651].

### 4 The PVM initial tʃ and its place in Mon-Khmer

PVM **tʃ** (Viet. *x*) while not frequent is attested in significant vocabulary. That proto phoneme is only attested in the northern branch (Viet + Mường). The comparison shows some correspondences between Viet *x*- and Khmu **c-** [Ferlus 1994]:

|  |  |
|--|--|
| Vietnamese                             | Khmu                                   |
| <i>xum</i> 'to get together'           | <b>cu:m</b> 'classifier for groups'    |
| <i>xuong</i> 'bone'                    | <b>cʔaŋ</b> 'bone'                     |
| <i>xoi</i> 'to dig, to sow, to pierce' | <b>cmɔ:l</b> 'to dig, to sow in holes' |
| <i>xẻ</i> 'to split'                   | <b>ceh</b> 'to square off'             |

To support the correspondences put forward above, it should be added that Khmu underwent the following chain of phonetic changes:

|                      |                      |  |
|----------------------|----------------------|--|
| *s > <b>h</b>        | *sa:l > <b>ha:l</b>  | 'to peel' (Phong Kenieng <b>sa:l</b> )     |
| -- *c > <b>s</b>     | *cɔʔ > <b>sɔʔ</b>    | 'dog' (Viet <i>chó</i> )                   |
| -- -- *tʃ > <b>c</b> | *tʃu:m > <b>cu:m</b> | 'classifier for groups' (Viet <i>xum</i> ) |

Apart from those correspondences, Khmu also attests many other examples of words with the initial **c-**: **cit** 'grass', **cat** 'sour', **caŋ** 'bitter', **cuʔ** 'to want, be sick', **cam** 'to weave a piece of thatch', **crna:m** 'a piece of thatch', ...

In Sino-Vietnamese, *x-* rendered the Middle Chinese \*tɕʰ [Ferlus 1992].

The place of \*tʃ in Viet and Khmu raises some problems. That proto phoneme is poorly represented if compared to the major units in the system, but, nevertheless, it exists in basic vocabulary. As far as we are concerned, \*tʃ is a residual phoneme originating in a North-Austroasiatic substratum partially preserved in Khmu and Vietnamese.

## 5 Morphological pairs of words (verb in *x-*, derivative in *ch-*)

5.1 One of the most remarkable characteristics of the Vietnamese lexicon is to possess a short list of five morphological pairs made up of a verbal base in *x-* associated with a derivative in *ch-* with an instrumental meaning.

| Verbal base  | Nominal derivative   |
|--|--|
| - <i>xáy</i> ‘dig, hollow, excavate’ /<br><i>xay</i> ‘grind, husk (rice)’      | <i>chày</i> ‘pestle’   |
| - <i>xeo</i> ‘lift up with a crowbar’<br>‘to propel (a boat) with a long pole’ | <i>chèo</i> ‘oar’  |
| - <i>xum</i> ‘gather, form groups’ /<br><i>xúm</i> ‘gather, form groups’       | <i>chùm</i> ‘bunch, cluster’<br><i>chụm</i> ‘assemble, gather’ |
| - <i>xĩa</i> ‘pick, jab, to put on a stip’                                     | <i>chĩa</i> ‘pitchfork, trident’                               |
| - <i>xỏ</i> ‘sting, pierce’  | <i>chõ</i> ‘pan to cook sticky rice’                           |

How could a nominal derivative in *ch-* (PVM ʃ), with a low serie tone, derive from a verbal base in *x-* (PVM tʃ), with a high serie tone? Correspondences between the attestation of ‘pestle’ among the VM languages suggest an old *-r-* infix:

|           |                                |
|-----------|--------------------------------|
| Mường     | k <sup>h</sup> aj <sup>2</sup> |
| Cuối Chặm | re: <sup>1</sup>               |
| Sách      | ri: <sup>1</sup>               |
| Arem      | <sup>n</sup> ri:               |

Another example can be found in Nguồn (a Mường dialect whose speakers were resettled in Quảng Bình): to the Viet *chõ* ‘pan to cook sticky rice’ corresponds the Nguồn rɔ:<sup>6</sup>.

The change /tʃ+ r/> ʃ is necessary to understand the relation between *x-* and *ch-* in the morphological pairs. That change is an isolate specific to Vietnamese ; in the other VM languages it evolved like current clusters /plos.+ r/ whose some examples are given below:

| PVM                 | Proto Pong                       | Rục                             | Mường                           | Viet        |               |
|---------------------|----------------------------------|---------------------------------|---------------------------------|-------------|---------------|
| p-ri:               | p <sup>h</sup> ri: <sup>1</sup>  | pri: <sup>1</sup>               | k <sup>h</sup> aj <sup>1</sup>  | <i>say</i>  | ‘be drunk’    |
| k-ro:ŋ <sup>2</sup> | k <sup>h</sup> ro:ŋ <sup>3</sup> | kro:ŋ <sup>3</sup>              | k <sup>h</sup> o:ŋ <sup>3</sup> | <i>sống</i> | ‘ridge, back’ |
| k-rap               | k <sup>h</sup> ra:p <sup>7</sup> | k <sup>h</sup> rap <sup>7</sup> | k <sup>h</sup> a:p <sup>7</sup> | <i>sáp</i>  | ‘wax’         |
| ʃ-ru:               | k <sup>h</sup> ru: <sup>2</sup>  | cəru: <sup>1</sup>              | k <sup>h</sup> u: <sup>1</sup>  | <i>sâu</i>  | ‘deep’        |

5.2 The phonetic history of Lao attests a similar change which supports the change /tʃ+ r/> ʃ in Viet. Proto Tai possessed the two voiced palatal initials \*ʃ and \*z which respectively evolved into c<sup>h</sup>- (ʃ) and s- (ʃ or ɲ) in Thai, but merged in s- (ʃ) in Lao [Fang Kuei Li 1977]. A short list of Lao words with the initial s (<\*z) underwent the change /plos.+ r/> z, the initial of the cluster being a coronal.

**se:**<sup>A2</sup> (<\*ze:) ទេ ‘river’ < Old Khmer \*sre: ‘ricefield’ (through the semantic change ‘ricefield’ > ‘ricefield + canal’ > ‘canal’ > ‘river’). Not represented in Thai.

**saj:**<sup>A2</sup> (<\*zaj) ខ្សាម ‘sand’ < Old Chinese \*sCraj [C-raj], *shā* 沙 [Baxter 1992: 785]. Thai ทราบ.

**saj:**<sup>A2</sup> (<\*zaj) ខ្សាម ‘hog deer (*Cervus porcinus*)’ < Old Mon *drāy*, Modern Mon *drāy* **kràj**. Thai ทราบ.

**so:**<sup>A2</sup> (<\*zo:) ខ្សែ ‘two-stringed violin’ < cf. Modern Mon *draw* **krò**. Thai ฆอ.

**saj:**<sup>A2</sup> (<\*zaj) ខ្សែ ‘banyan tree’ < Old Khmer *jrai*, Modern Khmer *jrai* **crej** / Old Mon *jrey*, *jreai*. Thai ไม้ท้อ.

**5.3** The instrumental infix **-r-** can only be reconstructed after the PVM initial **tʃ**. That infix has only been detected in the North-Austroasiatic substratum of Vietnamese. In the Mon-Khmer languages of Southeast Asia, the most commonly attested infix is **-rn-** (in its full form) or **-n-** (in its reduced form). The origin of the infix **-r-** and its place in Austroasiatic morphological system are a new subject of research which will not be dealt here.

## 6 The morphological pair ‘to husk (rice) - pestle’ in PVM

|  |
|--|
| <p><i>xáy</i> ‘dig, hollow, excavate’ /<br/> <i>xay</i> ‘grind, husk (rice)’ &gt; <i>chày</i> ‘pestle’</p> |
|--|

**6.1** PVM presents two basic verbs from which *chày* ‘pestle’ can have derived: (i) PVM **tʃe?** (*xáy*) ‘dig, hollow, excavate’ and (ii) PVM **tʃe:** (*xay*) ‘grind, husk (rice)’. The root **tʃe:**, which has a specialized meaning, must probably derive from **tʃe?**, which has a general meaning. Let’s now try to explain the phonetic change which led **tʃe?** (*xáy*) ‘dig, hollow, excavate’ to **tʃe:** (*xay*) ‘grind, husk (rice)’.

It is a well known fact in general linguistics that a repetitive action is generally expressed by a reduplication of the basic verb indicating the simple motion. We can consequently supposed the following change **tʃe?** > **tʃe?-tʃe?**. Thereafter, the reduplicate form was reduced to **tʃ-tʃe?**, which is nothing else than a structural adaptation to a sesquisyllabic constraint.

**6.2** Before going further in the explanation of phonetic changes which brought PVM to Vietnamese, it is necessary to point out some phonetic changes that affected Chinese and which occurred between the stage of Old Chinese and Middle Chinese. The formation of the Vietnamese language since its origin has been strongly influenced by some phonetic changes that affected the Chinese language. One could even say that the phonetic changes in Vietnamese are aftereffects of the phonetic changes that affected the Chinese language.

Between the final stage of Old Chinese (2nd-1st BC) and that of Middle Chinese (7th AD), a phonetic feature of tenseness developed in sesquisyllables as a consequence of the coalescence of primary tenseness of initials in each syllable. Both separate tenseness merged into one stronger tenseness. By contrast, the feature of laxness developed in monosyllables. Consequently to monosyllabization, the *tense-lax* contrast (henceforth T~L) became relevant in creating two types of syllables which most sinologists name A and B.

|         |   |         |                      |   |
|---------|---|---------|----------------------|---|
| C-CV(C) | > | CV(C)/T | ( <i>tenseness</i> ) | A |
| CV(C)   | > | CV(C)/L | ( <i>laxness</i> )   | B |

Thereafter, the T and L features modified the apertures of the vocalic onsets, lowering in A, raising and associated with breathiness in B. That theory was developed in our two communications at the *31st and 39th International Conference on Sino-Tibetan Languages and Linguistics* [Ferlus 1998, 2006]. It should be mentioned, however, that our theory is far from being accepted in the sinologists' world.

**6.3** By the Han time, the T~L contrast in the Chinese syllables was transferred to PVM in the same context: sesquisyllables developed a *tenseness* feature, while monosyllables developed a *laxness* feature. T~L contrast on PVM, however, acted differently than on Chinese. Those rather complex changes brought us to view two stages for PVM: an Early PVM and a Late PVM (the traditional PVM). That theory was presented at the *11th Annual Meeting of the Southeast Asian Linguistic Society*, Mahidol University at Salaya, 2001 [Ferlus 2004].

In Early PVM, the tenseness on sesquisyllables caused the final -ʔ loss, creating so open syllables. Let us point out some examples illustrating those changes:

|           |        |          |                    |            |        |
|-----------|--------|----------|--------------------|------------|--------|
| Early PVM | (Khmu) | Late PVM | Rục                | Viet       |        |
| *k-maʔ    | (kmaʔ) | *k-ma:   | kəmæa <sup>2</sup> | <i>mua</i> | 'rain' |
| *c-ruʔ    | (ɽruʔ) | *c-ru:   | cəru: <sup>1</sup> | <i>sâu</i> | 'deep' |

Concerning the vocabulary which interests us here:

|               |                       |          |                   |             |                 |
|---------------|-----------------------|----------|-------------------|-------------|-----------------|
| *tʃeʔ>tʃ-tʃeʔ | ---                   | *tʃ-tʃe: | ---               | <i>xáy</i>  | 'to husk(rice)' |
| *tʃ-reʔ       | (cn <sup>d</sup> reʔ) | *tʃ-re:  | nri: <sup>2</sup> | <i>chày</i> | 'pestle'        |

In monosyllables, on the other hand, the final glottal stop was preserved (the presyllabic vowel was not taken into account as a presyllable):

|       |       |       |                   |            |                 |
|-------|-------|-------|-------------------|------------|-----------------|
| *əcɔʔ | (sɔʔ) | *cɔʔ  | acɔ: <sup>3</sup> | <i>chó</i> | 'dog'           |
| *əkaʔ | (kaʔ) | *kaʔ  | aka: <sup>3</sup> | <i>cá</i>  | 'fish'          |
| *tʃeʔ | ---   | *tʃeʔ | ---               | <i>xáy</i> | 'dig, excavate' |

#### 6.4 To summarize:

- \*tʃeʔ (*xáy*) 'dig, hollow, excavate'.
- \*tʃeʔ > (reduplication) tʃeʔ-tʃeʔ > (sesquisyllabization) tʃ-tʃeʔ > (tenseness and loss of final -ʔ) tʃ-tʃe: > (monosyllabization) tʃe: (*xáy*) 'to husk (rice)'.
- \*tʃeʔ + infix -r- > tʃ-reʔ > (tenseness and loss of final -ʔ) tʃ-re: > tʃre: > (reduction) ʃe: (*chày*) 'pestle'.

It is clear that *xáy* 'to husk (rice)' is the result of an old process of reduplication of *xáy* 'dig, hollow, excavate', while *chày* derive from *xáy* by the infixation of -r-. All changes involved in the demonstrations are in conformity with the regularity of the phonetic laws.

### 7 The morphological pair ‘to husk (rice) - pestle’ in Austroasiatic

The vocabulary analyzed here comes from personal collected materials [Ferlus, Martin] and from linguists’ publications [Sidwell, Zide, Diffloth, ...] as well as of non linguists’ ones [Baradat, Skeat & Blagden]. The authors implied in works of linguistic reconstruction were conveniently not quoted.

It was quite difficult to collect the two words for ‘to husk (rice)’ and ‘pestle’, particularly when they were scattered in general studies or lexicons in which target language is placed in input. There are often ambiguities between ‘to husk’ and ‘to pound’ ; the Western authors being sometimes not accurate on those technical actions, while are so fundamental in the concerned societies.

| Group/Language                  | ‘to husk’              | ‘to pound’        | ‘pestle’                              |
|---------------------------------|------------------------|-------------------|---------------------------------------|
| <i>VIETIC</i> [Ferlus]          |                        |                   |                                       |
| PROTO VIET-MUONG                | (tʃeʔ >) tʃe:          |                   | (tʃreʔ >) ʃe:                         |
| Viet                            | (xáy >) xay            |                   | chày                                  |
| Mường [Nguyễn VK 2002]          | saj <sup>1</sup> (xay) |                   | k <sup>h</sup> aj <sup>2</sup> (khày) |
| Cuối Chàm                       | saj <sup>1</sup>       |                   | re: <sup>1</sup>                      |
| Làng Lỡ                         | saj <sup>1</sup>       |                   | ʃe: <sup>1</sup>                      |
| PROTO PONG                      |                        | təp <sup>8</sup>  | re: <sup>1</sup>                      |
| Thavung                         | mu: <sup>1</sup>       |                   | ahə: <sup>1</sup>                     |
| Sách                            | cuk <sup>7</sup>       | tu:ɲ <sup>2</sup> | əri: <sup>1</sup>                     |
| Arem                            | tluh                   | tù:ɲ              | <sup>h</sup> ri:                      |
| Maleng Kari                     | kəlu: <sup>56</sup>    |                   | səre: <sup>1</sup>                    |
| <i>KATUIC</i> [Ferlus]          |                        |                   |                                       |
| Suei                            |                        | ntap              | n <sup>d</sup> rè:                    |
| Ong                             | kloh                   |                   | ndraj                                 |
| Kantou                          | kloh                   |                   | ntre:                                 |
| Sô                              | cikloh                 |                   | ntɹi:                                 |
| <i>KATUIC</i> [Sidwell]         |                        |                   |                                       |
| PROTO KATUIC [2005]             | kloh                   | tap               | ʔn <sup>d</sup> ree                   |
| Souei                           |                        | ntap              | ntɹe:                                 |
| Sô/Bru                          | kloh                   |                   | ntɹii                                 |
| <i>BAHNARIC</i> [Sidwell]       |                        |                   |                                       |
| PROTO BAH. [1998]               | pəh                    |                   | ʔənrəj/r(ən)aj                        |
| <i>NORTH BAHNARIC</i> [Sidwell] |                        |                   |                                       |
| PROTO NORTH BAH. [2002]         | pəh                    |                   | ʔəraj                                 |
| Jeh                             | pəh <sup>T</sup>       |                   | ʔədraj <sup>T</sup>                   |
| Halang                          | pəh <sup>T</sup>       |                   | hədraj                                |
| Rengao                          | pih <sup>T</sup>       |                   | hədrii <sup>L</sup>                   |
| Sedang                          | pej                    |                   | draj <sup>T</sup>                     |
| Bahnar                          | pəh                    |                   | hdəj                                  |
| <i>SOUTH BAHNARIC</i> [Sidwell] |                        |                   |                                       |
| PROTO SOUTH BAH. [2000]         | pəh                    |                   | r-n-aj                                |
| Mnong                           | pəh                    |                   | ne                                    |
| Stieng                          | pəh                    |                   | rənaj                                 |
| Chrau                           | pəh                    |                   | rənaj                                 |

|   |                   |                                |
|---|-------------------|--------------------------------|
| <i>WEST BAHNARIC</i> [Ferlus]           |                   |                                |
| Laven                                   | tpeh              | ʔrej                           |
| Nhaheun                                 |                   | ʔre:                           |
| Brao                                    | tveh              | raj                            |
| Sapouan                                 |                   | araj                           |
| Lave                                    | tveh              | araj                           |
| Cheng                                   | tveh              | raj                            |
| <i>WEST BAHNARIC</i> [Sidwell, Jacques] |                   |                                |
| PROTO WEST BAH. [2000]                  | təpeh             | ʔraj                           |
| PROTO WEST BAH. [2003]                  | tʔpeh             | ʔraj                           |
| Laven/Jru'                              | təpeh             | ʔraj                           |
| Nyaheun                                 |                   | ʔree                           |
| Sapuan                                  |                   | ʔraj                           |
| <i>BOLYU</i> [Edmondson 1995]           |                   | tən <sup>53</sup>              |
| <i>MANG</i>                             | tə:               | tuj                            |
| <i>KHMUIC</i> [Ferlus]                  |                   |                                |
| Khmu                                    | hic               | cn <sup>d</sup> reʔ            |
| Phay                                    | k <sup>h</sup> ət | ŋgleʔ                          |
| Thin                                    | k <sup>h</sup> ət | ŋgreʔ                          |
| Pray                                    | k <sup>h</sup> ət | ŋgiaʔ                          |
| Lamet                                   | peh               | ntro:                          |
| Keneng                                  | kal               | kanre:                         |
| Hat                                     | suʔ               | ndra:                          |
| Khang                                   | təpe:             | he <sup>ʔ</sup>                |
| Kesing Mul                              | bək               | hagè:                          |
| <i>PALAUNGIC</i> [Ferlus]               |                   |                                |
| daʔa:k                                  | aduh              | ŋkrej                          |
| taʔa:ŋ                                  | dəh               | gre:                           |
| raʔa:ŋ                                  | dih               | gləŋ achom                     |
| <i>WAIC</i> [Ferlus]                    |                   |                                |
| pəzaək                                  | tah               | griʔ                           |
| va <sup>ʔ</sup>                         | kujh              | ŋiʔ                            |
| Sem                                     | taoh              | gliʔ                           |
| Phalək                                  | dəh               | ŋiʔ                            |
| Samtao                                  | tih               | nreʔ                           |
| lavia <sup>ʔ</sup>                      | bləuh             | k <sup>h</sup> ou <sup>ʔ</sup> |
| La-ooop                                 | toh               | grei <sup>ʔ</sup>              |
| Lawa                                    | pəuh              | k <sup>h</sup> o: toh          |
| PROTO WA [Diffloth 1980]                | toh               | ŋriʔ                           |
| <i>RIANG</i> [Luce 1965]                |                   | rèʔ                            |
| <i>DANAW</i> [Luce 1965]                |                   | réʔ                            |
| <i>MONIC</i>                            |                   |                                |
| Môn [Shorto 1962]                       | yàik [jàc]        | ri <sup>ʔ</sup>                |
|   | yāk               | ri                             |
| Nyah Kur [Theraphan 1984]               | jà:k              | ŋri: <sup>ʔ</sup>              |



|                                |                                      |                          |   |
|--------------------------------|--------------------------------------|--------------------------|---|
| <i>KHMER</i>                   |                                      |                          |   |
| Khmer                          | <b>bok</b><br><i>puk</i>             | <b>kɤn</b><br><i>kin</i> | <b>ʔəŋre:</b><br><i>'aŋræ</i>                 |
| <i>PEARIC</i> [Baradat 1941]   |                                      |                          |   |
| Pear, Kpg Speu                 | <i>chhâk</i>                         | <i>ken</i>               | <i>rôhi-i</i>                                 |
| Pear, Kpg Thom                 | <i>bok</i>                           | <i>ken</i>               | <i>ré</i>                                     |
| Pear, west                     | <i>chhûk</i>                         |                          | <i>rôhi-i</i>                                 |
| Pear, east                     | <i>chhâk</i>                         |                          | <i>rôhik</i>                                  |
| <i>PEARIC</i> [Martin]         |                                      |                          |   |
| Samray                         | <b>chuuk</b>                         | <b>ken</b>               | ( <i>rôhi-i</i> )                             |
| Səmree                         | <b>chœk</b>                          | <b>kum</b>               | ( <i>rôhik</i> )                              |
| <i>PEARIC</i> (various)        |                                      |                          |   |
| Pear [Headley 1978]            | <b>čha:k</b>                         |                          | <b>rəhi:</b>                                  |
| Saoch                          | <b>t<sup>h</sup>a:k</b>              |                          | <b>ri</b>                                     |
| Chong [Siriphen 2001]          | <b>c<sup>h</sup>œ:k<sup>R1</sup></b> | <b>bot</b>               | <b>kəhi:<sup>R1</sup> [kə<sup>h</sup>hi:]</b> |
| <i>KHASI</i> [Singh 1920]      |                                      |                          |   |
|                                |                                      |                          | <i>synrei</i>                                 |
| <i>ASLIAN</i>                  |                                      |                          |   |
| Jahai [Burenhult 2001]         | <b>sntip/tiʔ/sih/patim/til</b>       |                          | <b>gul</b>                                    |
| Tembi [Skeat & B. 1906]        |                                      |                          | <i>rentik</i>                                 |
| Serau [Skeat & B. 1906]        |                                      |                          | <i>kěnöh, kěnu<sup>?</sup>(?)</i>             |
| <i>NICOBAR</i>                 |                                      |                          |   |
|                                | -                                    |                          | -   |
| <i>NORTH MUNDA</i> [Zide 1976] |                                      |                          |   |
| Korku                          | <b>rum-</b>                          |                          | <b>toko / tuki</b>                            |
| Ho                             | <b>ruuŋ-</b>                         |                          | -   |
| Santali                        | <b>ruŋŋ-</b>                         |                          | -   |
| Santali [Macphail 1954]        | <i>huruf</i>                         | <i>sok'</i>              | <i>tok</i>                                    |
| <i>SOUTH MUNDA</i> [Zide 1976] |                                      |                          |   |
| Kharia                         | -                                    |                          | <b>ě(n)ri / enđi</b>                          |
| Remo                           | -                                    |                          | <b>tiŋeʔ</b>                                  |
| Gtaʔ                           | -                                    |                          | <b>toŋkæ</b>                                  |
| Gorum                          | -                                    | <b>taŋlad</b>            | <b>in(d)ri</b>                                |
| Sora                           | -                                    | <b>taŋlad</b>            | <b>əŋrij</b>                                  |
| <i>PROTO MON-KHMER</i>         |                                      |                          |   |
| [Shorto 2006]                  | <b>pis/pøs</b>                       | <b>[k]bok</b>            | <b>nrøyʔ / nrœy</b>                           |

General remarks: (see *Summarized chart* and map at the end of article)

A remarkable fact arises from the reading of the table: the verbal base ‘to husk (rice)’ and the nominal derivative ‘pestle’ form a morphological pair only in the subgroups of Vietnamese, Mường and Thổ (Cuối Chặm, Làng Lỗ), i.e. in the most septentrional languages of the VM group. On the other hand, the same derivative ‘pestle’, recognizable by the presence of *r* in its various forms, is attested in the other VM languages and in most groups of the Austroasiatic family.

The languages or groups of languages which attest other roots for ‘pestle’ are Bolyu (Guangxi - Zhuang Autonomous Region), Mãng (Lai Châu, Vietnam), the Aslian group (Peninsular Malaysia) and North Munda (India). As far as Nicobarese is concerned, it does

not seem to have proper vocabulary for rice and its culture ; the word for ‘rice’ (Nancowry *arōsh*, Teressa *aros*) is genuinely Portuguese [de Röpstorff 1875].

It is obvious that the derivation which produced the word ‘pestle’ took place in a northern VM language, direct ancestor of Vietnamese. From there, the object and its name spread through most Austroasiatic languages, as far as in India.

In current classifications, Munda forms a clearly characterized branch within the Austroasiatic family. However, it seems surprising that the word for ‘pestle’ reached South Munda and missed North Munda. The Munda branch might be the result of a symbiosis of several waves of Austroasiatic languages coming from the Austroasiatic *Urheimat*, somewhere in the heart of China.

- |   |
|---|
| <p>8 <i>xeo</i> ‘lift up with a crowbar,<br/>to propel (a boat) with a long pole’ &gt; (<i>cái</i>) <i>chèò</i> ‘paddle, oar’</p> |
|---|

PVM **tʃɛ:w** (*xeo*) and **tʃ-r-ɛ:w** > **ʃɛ:w** (*chèò*) must be reconstructed.

*Chèo* must have originally named the long pole used to propel boats ; today, it means ‘to paddle, to row’, while *cái chèo* means ‘paddle, oar’.

The word *chèò*, verb or noun, is quite common among the VM languages and many languages of Vietnam and neighbouring countries. It is represented in Khmer by **caew cæv** ‘to paddle, to row, paddle’, while ‘oar’ is **crəva: cravā**. In Lao we find **sɛ:w<sup>A2</sup>** (<\***ʃɛ:w**) ສຳຮັດ ‘to row’.

To the same word family we must add *neo* ‘anchor’, formed by the insertion of an old **-rn-** infix with an instrumental meaning:

**tʃɛ:w** > (infixation) **tʃ-rn-ɛ:w** > (monosyllabization) **nɛ:w** *neo* ‘anchor’.

Note: (i) The infix **-rn-** has been preserved in some Maleng dialects of the VM group. For exemple, in Maleng Brô [Ferlus 1997]:

**sɛk - srnɛk** ‘to comb - a comb’  
**taj<sup>2</sup> - trnaj<sup>2</sup>** ‘to light with a steel lighter - a lighter’  
**ko<sup>2</sup> - krno<sup>2</sup>** ‘to dwell, to stay at - a house’

(ii) The Vietnamese vocabulary attests many examples of the type *xeo-neo* which reinforce the reconstruction of an infix **-rn-**:

*đan - nan* ‘to plait - bamboo split’  
*đút - nút* ‘to cork (a bottle) - a cork’  
*chọc - nọc* ‘to shake down (with a long pole) - a long pole’  
*xếp - nếp* ‘to fold - a fold’

- |  |
|--|
| <p>9 <i>xum</i> ‘gather, form group’ &gt; <i>chùm</i> ‘bunch, cluster’<br/> <i>xúm</i> ‘gather, form groups’ &gt; <i>chụm</i> ‘assemble, gather’</p> |
|--|

The place of *xum* in dictionaries needs some further remarks. *Xum* is not attested in the modern Vietnamese dictionaries, while in others, *xum* and *xúm* are presented as synonyms.

Father E. Gouin [1957] was the only one to establish a clear distinction between (in French) *xum* ‘se réunir, rassembler’ and *xúm* ‘se réunir, réunir, rassembler, convoquer, grouper’. This distinction can be interpreted as *xum* ‘to meet, to get together’, with an intransitive meaning, and *xúm* ‘to gather, to collect, to call together’ with a causative aspect.

We can then reconstruct PVM **tfum** (*xum*) as the basic root with the meaning ‘to meet, to get together’ and suppose a causative derivation, **p-tfum** with the following chain of changes:

**tfum** > (prefixation) **p-tfum** > (tenseness and glottalization) **p-tfum<sup>2</sup>** > (monosyllabization) **tfum<sup>2</sup>** (*xúm*). On the circumstances of the occurrences of glottalization in sesquisyllables, see Ferlus [2004].

Formation of derivatives with the infix **-r-**: **tf-r-um** > **jum** (*chùm* ‘bunch, cluster’) and **(p-)tf-r-um<sup>2</sup>** > **jum<sup>2</sup>** (*chum* ‘assemble, gather’).

The prefixed form **p-tfum** gave *giùm* ‘give help, help’ by spirantization of **tf** in medial position: **p-tfum** > (spirantization) **p-ju:m** > (monosyllabization) **jum** (*giùm*). Old dictionaries also attest *gium* ‘help’, *giùm* ‘to help each other’ and *giùm* ‘to put together’.

The prefixed form passed in Khmer, **prəcum** *prajum*, then in Thai **prac<sup>h</sup>um<sup>A2</sup>** ปรๅจๅม and in Lao, **pasum<sup>A2</sup>** ປๅຊຸມ.

10  *xia* ‘pick, jab, to put on a stip’ > *chĩa* ‘pitchfork, trident’

PVM **tfɛh** (*xia*) and **tf-r-ɛh** > **jeɛh** (*chĩa*) must be reconstructed.

Derivative formed with **-rn-** infix: **tfɛh** > (infixation) **tf-rn-ɛh** > (monosyllabization) **neh** *nĩa* ‘fork’.

These words remain confined in the Vietnamese area.

11  *xỏ* ‘sting, pierce’ > *chỗ* ‘pan to cook sticky rice’

PVM **tfɔh** (*xỏ*) and **tf-r-ɔh** > **jeɔh** (*chỗ*) must be reconstructed.

These words remain confined in the Vietnamese area.

## 12 Conclusions

The PVM proto phoneme **tf** is specific to the Vietnamese language and to some very close VM languages. Words opening with the initial **\*tf** (*x-*) are very few but belong to the significant vocabulary of everyday life. Correspondences with Khmu have been noticed.

In Vietnamese, there are five morphological pairs of words associating a verb in *x-* with a nominal derivative in *ch-*. These five pairs are: (1) *xáy/xay* - *chày*, (2) *xeo* - *chèo*, (3) *xum/xúm* - *chùm/chum*, (4) *xia* - *chĩa* and (5) *xỏ* - *chỗ*. The verb expresses a basic action, while the derivative indicates an object or a concept related to the exercise of the action. Correspondences in VM make it possible to highlight an old nominalizing **-r-** infix with an instrumental meaning.

Among these morphological pairs, the most striking is *xáy/xay* - *chày*. It was explained how from PVM **tfɛ?** (*xáy*) ‘to dig, excavate’ was formed the derivative **tfɛ:** (*xay*) ‘to husk (rice)’ with a more specialized meaning, and also was formed **tfre?** > **je:** (*chày*) ‘pestle’.

It was also noted that, in the primordial PVM pair **tʃe?** - **tʃre?**, the reflexes of the basic verb (**tʃe?**>) **tʃe:** ‘to husk (rice)’ remained restricted to Vietnamese, while the reflexes of the derivative **\*tʃre?** ‘pestle’ spread in the most Austroasiatic languages. Bolyu, Mãng, Aslian, Nicobarese, North Munda and some languages of South Munda did not receive that derivative. We are faced to a rather exceptional case, considering the antiquity of the phenomenon, where a word created in a limited area invaded the quasi totality of a linguistic family.

This phenomenon is not only of linguistic nature, it is also necessary to take into account also the technological component and more generally the level of civilization in the area of origin. It is obvious that the word for ‘pestle’ spread with the object itself. Such an expansion does not have any equivalent in the old times. It is the object itself more than the carrying languages, that spread through the Austroasiatic family. That means that the pestle was an innovating invention, the technical superiority of which was higher than all that preceded in the manner of husking rice. The complex ‘pestle - mortar’ (in French ‘pilon - mortier’), made possible a better husking of the grain than the complex ‘saddle quern - rubber stone’ (in French ‘meule dormante - molette mobile’) which might exist before. The other advantage is that utensils out of wooden are easier to make than those out of stone.

The continuity of the morphological pairs in a layer of the Vietnamese vocabulary (the layer of PVM **tʃ**) can only be explained if one population went on speaking the same language in the same place. Moreover, the verbs of the morphological pairs imply current actions, the nominal derivatives of which are utensils or concepts useful in the everyday life: : ‘pestle’, ‘oar’, ‘group’, ‘trident’ and ‘pan to cook sticky rice’. The speakers of that language belonged to a culture which encouraged them to innovate.

As the Đông Sơn culture (c. 7th BC to 1st AD), famous for its bronze drums [Parmentier 1918: *Pl. IV, fig. 1*], was precisely located in the North of Vietnam, at the same place as the area of origin of our morphological pairs, one can conclude from it that this layer comes from the Dongsonians’s language.

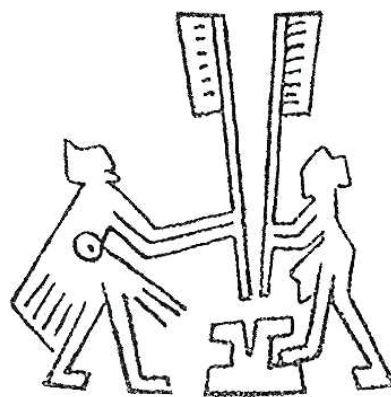
In conclusion: the Vietnamese language preserved a part of the Dongsonians’ language, and the Vietnamese are the most direct heirs of the Dongsonian culture.

## References

- Alves, Mark. 2006. Linguistic Research on the Origin of the Vietnamese Language: An Overview. *Journal of Vietnamese Studies* 1(1-2): 104-130.
- Baradat, R. 1941. *Les dialectes des tribus sâmrê*. Manuscript, École Française d’Extrême-Orient. Paris.
- Burenhult, Niclas. 2005. *A Grammar of Jahai*. Pacific Linguistics 566. Canberra, The Australian National University.
- Diffloth, Gérard. 1980. The Wa Languages. *Linguistics of the Tibeto-Burman Area* 5(2).
- Edmondson, Jerold. 1995. English-Bolyu Glossary. *Mon-Khmer Studies* 24: 133-159.
- Fang Kuei Li. 1977. *A Handbook of Comparative Tai*. The University Press of Hawaii.
- Ferlus, Michel. Unpublished materials on several Mon-Khmer languages, specially Viet-Muong (Vietic) languages, collected in Laos, Thailand, Burma and Vietnam.
- Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. *Cahiers de linguistique Asie Orientale* 11(1): 83-106.

- Ferlus, Michel. 1992. Histoire abrégée de l'évolution des consonnes initiales du vietnamien et du sino-vietnamien. *Mon-Khmer Studies* 20: 111-125.
- Ferlus, Michel. 1994. Contacts anciens entre viet-muong et austroasiatique-nord. *Kristina Lindell Symposium on Southeast Asia*. University of Lund. May 16, 1994.
- Ferlus, Michel. 1997. Le maleng brô et le vietnamien. *Mon-Khmer Studies* 27: 55-66.
- Ferlus, Michel. 1998. Du chinois archaïque au chinois ancien: monosyllabisation et formation des syllabes *tendu/lâche* (Nouvelle théorie sur la phonétique historique du chinois). *The 31st International Conference on Sino-Tibetan Languages and Linguistics*. University of Lund, Sept. 30 - Oct. 4, 1998.
- Ferlus, Michel. 2004. The Origin of Tones in Viet-Muong. *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistic Society 2001*. Edited by Somsong Burusphat. Arizona State University: 297-313.
- Ferlus, Michel. 2006. What were the four Divisions (*děng* 等) of the Middle Chinese. *The 39th International Conference on Sino-Tibetan Languages and Linguistics*, University of Washington at Seattle, September 14-17.
- Gouin, Eugène. 1957. *Dictionnaire vietnamien chinois français*. Saigon, Imprimerie d'Extrême-Orient.
- Headley, Robert K. 1978. An English-Pearic Vocabulary. *Mon-Khmer Studies* VII.
- Jacq, Pascale & Paul Sidwell. 2000. *A Comparative West Bahnaric Dictionary*. Lincom Europa.
- Luce, Gordon H. 1965. Danaw, A Dying Austroasiatic Language. *Lingua* 14: 98-129.
- Macphail, R.M. (edited by). 1954. *Campbell's English-Santali Dictionary*. Santal Mission Press, Benageria, India.
- Man, Edward Horace. 1889. *Dictionary of the Central Nicobarese Language (English-Nicobarese and Nicobarese-English)*, ... London, W.H. Allen and Co. Reprint 1975, Delhi.
- Martin, Marie A. Unpublished materials on Pearic languages.
- Nguyễn Văn Khang, Bùi Chí and Hoàng Văn Hành. 2002. *Từ Điển Mường-Việt* [Mường-Việt Dictionary] . Hà Nội, Nhà xuất bản văn hóa dân tộc.
- Nguyễn Văn Lợi. 1993. *Tiếng Rục* [The Rục language]. Hà Nội, Nhà xuất bản khoa học xã hội.
- Parmentier, Henri. 1918. Anciens tambours de bronze. *Bulletin de l'Ecole Française d'Extrême-Orient* 18(1): 1-30 + planches.
- Peiros, Ilija. 1996. *Katuic Comparative Dictionary*. Pacific Linguistics C-132. Canberra, The Australian National University.
- Rhodes, (Père) Alexandre de. 1651. *Dictionarium annamiticum, lusitanum, et latinum*. Rome. Reprinted with a translation into Modern Vietnamese: Viện Khoa Học Xã Hội tại T.P. Hồ Chí Minh, *Từ Điển Annam-Lusitan-Latinh*, 1991, Nhà xuất bản khoa học xã hội.
- Röepstorff, Frederick. Ad. de. 1875. *Vocabulary of Dialects Spoken in the Nicobar and Andaman Isles*. Calcutta, Superintendent Government Printing. Second edition: 1987, New Delhi, Asian Educational Services.

- Shorto, Harry L. 1962. *A Dictionary of Modern Spoken Mon*. London, Oxford University Press.
- Shorto, Harry L. 2006. *A Mon-Khmer Comparative Dictionary*. Edited by Paul Sidwell, Doug Cooper and Christian Bauer. Pacific Linguistics 579. The Australian National University.
- Sidwell, Paul. 1998. *A Reconstruction of Proto-Bahnaric*. Thesis. University of Melbourne.
- Sidwell, Paul. 2000. *Proto South Bahnaric, A Reconstruction of a Mon-Khmer language of Indo-China*. Pacific Linguistics, The Australian National University.
- Sidwell, Paul. 2005. *The Katuic Languages, Classification, Reconstruction and Comparative Lexicon*. Lincom Europa.
- Sidwell, Paul & Pascale Jacq. 2003. *A Handbook of Comparative Bahnaric, Volume 1: West Bahnaric*. Pacific Linguistics, The Australian National University.
- Singh, U Nissor. 1920. *English-Khasi Dictionary*. Assam. [1993. Reprint Mittal Publications. India]
- Siriphen Ungsitipoonporn. 2001. *A Phonological Comparision between Khlong Phlu Chong and Wangkraphræ Chong*. MA thesis. Institute of Language and Culture for Rural Development, Mahidol University at Salaya.
- Skeat, Walter W. and Blagden, Charles Otto. 1906. *Pagan Races of the Malay Peninsula*. Two volumes. London, Frank Cass. Reprint 1966.
- Suwilai Premsrirat. 2002. *Thesaurus of Khmu Dialects in Southeast Asia*. Salaya (Nakhon Pathom, Thailand), Mahidol University, Institute of Language and Culture for Rural Development.
- Theraphan L. Thongkum. 1984. *Nyah Kur (Chao Bon)-Thai-English Dictionary*. Monic Language Studies II. Bangkok, Chulalongkorn University Printing House.
- Zide, Arlene R. K. & Norman H. Zide. 1976. Proto-Munda Cultural Vocabulary: Evidence for Early Agriculture. *Austroasiatic Studies*, part II. Edited by Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta: 1295-1334.



A husking rice scene engraved on a Dongsonian bronze drum [Parmentier 1918: Pl. IV, fig. 1]. Museum of History in Hanoi.

## Summarized chart: 'to husk (rice) - pestle' in Austroasiatic

| <i>Groups/Languages</i>   | <i>to husk (rice)</i>   | <i>to pound</i>          | <i>pestle</i>  |
|---|---|--------------------------|--|
| PROTO VIET-MUONG<br>Viet<br>Mường Bì  | (tʃeʔ >) tʃe:<br>(xáy >) xay<br>saj <sup>1</sup> (xay)                |                          | (tʃreʔ >) ʃe:<br>chày<br>k <sup>h</sup> aj <sup>2</sup> (khày)                             |
| Sách<br>Arem  | cuk <sup>7</sup><br>tloh  | tuɲ <sup>2</sup><br>tù:ɲ | əri: <sup>1</sup><br><sup>n</sup> ri:  |
| PROTO KATUIC  | kloh  | tap                      | ʔn <sup>d</sup> ree  |
| PROTO BAHNARIC<br>PROTO NORTH BAH.<br>Rengao<br>Bahnar<br>PROTO SOUTH BAH.<br>Stieng<br>PROTO WEST BAH.<br>Laven/Jru' | pəh<br>peh<br>pih <sup>T</sup><br>peh<br>pəh<br>peh<br>tʔpeh<br>təpeh | jaʔ                      | ʔənɾəj/r(ən)aj<br>ʔəraj<br>hədrii <sup>L</sup><br>hdɾəj<br>r-n-aj<br>rənaj<br>ʔraj<br>ʔraj |
| BOLYU   |   | tən <sup>53</sup>        | xuək <sup>31</sup>   |
| MANG  | tə:   |                          | tuɲ  |
| KHMUIC<br>Khmu<br>Thin<br>Keneng  | hic<br>k <sup>h</sup> ət<br>kal                                       |                          | cn <sup>d</sup> reʔ<br>ŋgreʔ<br>kanre:   |
| PALAUNGIC taʔa:ŋ<br>PROTO WAIC  | doh<br>toh  |                          | gre:<br>ŋriʔ   |
| RIANG   | -   | -                        | rèʔ  |
| MÔN   | yàik [jàc]  |                          | ri <sup>ʔ</sup>  |
| KHMER   | bok   | kɤn                      | ʔoŋre:   |
| PEARIC<br>Saoch<br>Chong  | t <sup>h</sup> a:k<br>c <sup>h</sup> ɔ:k <sup>R1</sup>                | bət                      | ri<br>kəhi: <sup>R1</sup> [kə'hi:]   |
| KHASI   |   |                          | synrei   |
| ASLIAN<br>Jahai<br>Tembi  | sntip/tiʔ/sih/...   |                          | gul<br>rentik  |
| NICOBAR   | -   |                          | -  |
| NORTH MUNDA<br>Korku<br>Santali   | rum-<br>huɾuŋ   | sok'                     | toko / tuki<br>tok   |
| SOUTH MUNDA<br>Kharia<br>Sora   | -<br>-  | taŋlad                   | ē(n)ri/əndi<br>oŋrij   |
| PROTO MON-KHMER   | pis/pəs   | [k]bok                   | nrəyʔ/nrəəy  |

