A layer of Dongsonian vocabulary in Vietnamese
Michel Ferlus

To cite this version:

HAL Id: halshs-00932218
https://halshs.archives-ouvertes.fr/halshs-00932218v3
Submitted on 8 Sep 2014

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

some corrections : sept. 2014

0 Abstract
The present paper aims at demonstrating by means of linguistic evidence that 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.1

1 Background
The position of the Vietnamese language (or Viet in its shortened form) in Asian phylogeny has varied considerably since the first research on the topic was carried out. After being classified among the Chinese or the Tai-Kadai languages, it was finally integrated to the Mon-Khmer family [for a review, 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 Chỉ 交趾, and Cửu Chân 九真).

Vietnamese and Mường 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 group is presented below.

1- Maleng: Maleng proper, Pakatan, Mǎliěng, Maleng Brô, Kha Phong (or Maleng Kari).
2- Arem: Arem (or Cmrau/Cmbrau).
3- Chứt: Sách (or Chírt, or Salang), Rực.
4- Aheu: Thavung, Phôn Soung, Sô (or Sô Thavung).
5- Pong: Pong (or Phong), Toum, Liha, Đan-lai.
6- Thố: Làng Lờ, Cuội Chắm, Môn.
7- Mường: Mường (or Mọl/Mọn); comprises many dialects, including Mường Đắm, Mường Khóry and Mường Tân Phong and Ngườn.
8- Viet: written standard Vietnamese and its dialects.

1 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
(The current quóc ngữ spelling for the proto phonemes is written in italic)
PVM comprised monosyllables CV(C) and sesquisyllables C-CV(C).

\[
\begin{array}{cccccc}
  p^h & t^h & s & k^h & h \\
p & t & s & k & h \\
\end{array}
\]
\[
\begin{array}{cccccc}
  b & d & c & j & tʃ & k & ? \\
b \sim v & d \sim d & c \sim gi & x \sim gi & c \sim k \sim g \sim gh & # \\
\end{array}
\]
\[
\begin{array}{cccc}
  ɓ & ɗ & ʄ & m & n & nh \\
\end{array}
\]

The aspirated plosives \(p^h\), \(t^h\), \(k^h\) are not frequent and must have evolved from clusters of the type /plos. + h/.

Obstruents \(p\)-, \(t\)-d, \(c\)-, \(s\), \(tʃ\) 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 \(b\sim v\) (\(b\) in monosyllables and \(v\) in ancient sesquisyllables). It must be noticed that, in the 17th century, \(v\) was rendered by \(ɓ/ɗɓ\) in Alexandre de Rhodes’ dictionary [1651].

4 The PVM initial \(tʃ\) and its place in Mon-Khmer
PVM \(tʃ\) (viet \(x\)) is not frequent; however, it is attested in some important words. That proto phoneme is only attested in the northern branch (Viet + Mường). Comparison reveals correspondences between Viet \(x\)- and Khmu \(c\) - [Ferlus 1994]:

Vietnamese | Khmu
---|---
\(xum\) ‘to get together’ | \(cu:m\) ‘classifier for groups’
\(xıəŋ\) ‘bone’ | \(c?əŋ\) ‘bone’
\(xoi\) ‘to dig, to sow, to pierce’ | \(cmɔːl\) ‘to dig, to sow in holes’
\(xe\) ‘to split’ | \(ceh\) ‘to square off’

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

\(\ast s > h\) \hspace{1cm} \(\ast sa:l > ha:l\) ‘to peel’ (Phong Kenieng \(sa:l\))

\(\ast c > s\) \hspace{1cm} \(\ast co? > so?\) ‘dog’ (Viet \(chō\))

\(\ast tʃ > c\) \hspace{1cm} \(\ast tʃu:m > cu:m\) ‘classifier for groups’ (Viet \(xum\))

Apart from those correspondences, Khmu also attests many other examples of words with the initial \(c\)-: \(cit\) ‘grass’, \(cat\) ‘sour’, \(caŋ\) ‘bitter’, \(cu?\) ‘to want, be sick’, \(ca:m\) ‘to weave a piece of thatch’, \(cnɐːm\) ‘a piece of thatch’, …
In Sino-Vietnamese, \textit{x-} rendered the Middle Chinese \textit{*tɕʰ} [Ferlus 1992]. The place of \textit{*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. Our current view is that \textit{*tʃ} is a residual phoneme originating in a North-Austroasiatic substratum partially preserved in Khmu and Vietnamese.

5 Morphological pairs of words (verb in \textit{x-}, derivative in \textit{ch-})

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

<table>
<thead>
<tr>
<th>Verbal base</th>
<th>Nominal derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td>- xáy ‘dig, hollow, excavate’ / xay ‘grind, husk (rice)’</td>
<td>chày ‘pestle’</td>
</tr>
<tr>
<td>- xeo ‘lift up with a crowbar’</td>
<td>chèo ‘oar’</td>
</tr>
<tr>
<td>‘to propel (a boat) with a long pole’</td>
<td>chêo ‘pitchfork, Trident’</td>
</tr>
<tr>
<td>- xum ‘gather, form groups’ / xúm ‘gather, form groups’</td>
<td>chùm ‘bunch, cluster’</td>
</tr>
<tr>
<td>- xía ‘pick, jab, to put on a stip’</td>
<td>chĩa ‘wax’</td>
</tr>
<tr>
<td>- xỏ ‘sting, pierce’</td>
<td>chỏ ‘pan to cook sticky rice’</td>
</tr>
</tbody>
</table>

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

| Mường | kʰaj² |
| Cuối Chăm | reː¹ |
| Sách | riː¹ |
| Arem | nːriː |

Another example can be found in Nguồn (a Mường dialect whose speakers were resettled in Quang Bình): to the Viet chô ‘pan to cook sticky rice’ corresponds the Nguồn roː⁶.

The change \textit{/tʃ+r/>j} is necessary to understand the relation between \textit{x-} and \textit{ch-} in the morphological pairs. That change is an isolate specific to Vietnamese; in the other VM languages it evolved like the other \textit{/plos.+r/} clusters, some examples of which are given below:

<table>
<thead>
<tr>
<th>PVM</th>
<th>Proto Pong</th>
<th>Rúc</th>
<th>Mường</th>
<th>Viet</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-riː</td>
<td>pʰriː¹</td>
<td>priː¹</td>
<td>kʰaj¹</td>
<td>say ‘be drunk’</td>
</tr>
<tr>
<td>k-ɾoŋ²</td>
<td>kʰɾoŋ¹</td>
<td>kɾoŋ¹</td>
<td>kʰoŋ¹</td>
<td>sǒŋ ‘ridge, back’</td>
</tr>
<tr>
<td>k-ɾaːp</td>
<td>kʰɾaːp⁷</td>
<td>kʰɾaːp⁷</td>
<td>kʰaːp⁷</td>
<td>sáp ‘wax’</td>
</tr>
<tr>
<td>j-ɾuː</td>
<td>kʰɾuː²</td>
<td>cəɾuː¹</td>
<td>kʰuː¹</td>
<td>sâu ‘deep’</td>
</tr>
</tbody>
</table>

5.2 The phonetic history of Lao attests a similar change which supports the change \textit{/tʃ+r/>j} in Viet. Proto Tai possessed the two voiced palatal initials \textit{*j} and \textit{*z} which respectively evolved into \textit{cʰ-} (\textit{ɕ}) and \textit{s-} (\textit{ɕ} or \textit{nɕ}) in Thai, but merged in \textit{s-} (\textit{ɛ}) in Lao [Fang Kuei Li 1977]. A small number of Lao words with the initial \textit{s} (<\textit{*z}) underwent the change \textit{/plos.+r/>z}, the initial of the cluster being a coronal.
seːᴬ² (<*zeː:) រiver’ < Old Khmer *src: ‘ricefield’ (through the semantic change ‘ricefield’ > ‘ricefield + canal’ > ‘canal’ > ‘river’). Not represented in Thai.

saijᴬ² (,<*zaij) ស្ត ‘sand’ < Old Chinese *sCraj [C-raj],  shà 沙 [Baxter 1992: 785]. Thai ทะเล.

saijᴬ² (,<*zaːj) ស氨酸 ‘hog deer (Cervus porcinus)’ < Old Mon drày, Modern Mon ดราย kràj. Thai ทราย.

soːᴬ² (,<*zoi) ស氨酸 ‘two-stringed violin’ < cf. Modern Mon draw krò. Thai ซอ.

sajᴬ² (,<*zaj) ສ氨酸 ‘banyan tree’ < Old Khmer jrai, Modern Khmer jrai crej / Old Mon jrey, jreai. Thai ไทร.

5.3 The instrumental infix -rn- 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 -n- (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

| xáy  ‘dig, hollow, excavate’ / xay ‘grind, husk (rice)’ | chày ‘pestle’ |

6.1 PVM presents two basic verbs from which chày ‘pestle’ can have derived: (i) PVM tʃeʔ (xáy) ‘dig, hollow, excavate’ and (ii) PVM tʃeː (xay) ‘grind, husk (rice)’. The root tʃeː, which has a specialized meaning, probably derives from tʃeʔ, which has a general meaning. Let’s now try to explain the phonetic change which led tʃeʔ (xáy) ‘dig, hollow, excavate’ to tʃeː (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 tʃeʔ > tʃeʔ-tʃeʔ. Thereafter, the reduplicate form was reduced to tʃ-tʃeʔ, which is nothing else than a structural adaptation to a sesquisyllabic constraint.

6.2 Before going further in the explanation of phonetic changes from 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 both initials articulations in each syllable. By contrast, the feature of laxness developed in monosyllables. Consequently to monosyllabization, the tense–lax contrast (henceforth T–L) became distinctive, creating two types of syllables which most sinologists name A and B.
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, thus creating open syllables. Let us point out some examples illustrating those changes:

<table>
<thead>
<tr>
<th>Early PVM (Khmu)</th>
<th>Late PVM</th>
<th>Rúc</th>
<th>Viet</th>
</tr>
</thead>
<tbody>
<tr>
<td>*k-maʔ</td>
<td>(kmaʔ)</td>
<td>*k-ma: k̂mə́ʔ</td>
<td>mua</td>
</tr>
<tr>
<td>*c-ruʔ</td>
<td>(jruʔ)</td>
<td>*c-ru: ĉə́ʔ</td>
<td>sā́u</td>
</tr>
</tbody>
</table>

Concerning the vocabulary which interests us here:

*tʃeʔ > tʃ-tʃeʔ  ---  *tʃ-tʃe:  ---  xay  ‘to husk (rice)’
*tʃ-reʔ  (cn⁴reʔ)  *tʃ-re: nri:²  chày  ‘pestle’

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

*əcoʔ | (səʔ) | *coʔ | aco:³ | chó | ‘dog’
*okaʔ | (kaʔ) | *kaʔ | aka:³ | cá | ‘fish’
*tʃeʔ  ---  *tʃeʔ  ---  xáý  ‘dig, excavate’

6.4 To summarize:

*tʃeʔ (xáý) ‘dig, hollow, excavate’.
*tʃeʔ > (reduplication) tʃeʔ-tʃeʔ > (sesquisyllabization) tʃe-tʃeʔ > (tenseness and loss of final -ʔ) tʃ-tʃe: (xay) ‘to husk (rice)’.
*tʃeʔ + infix -r- > tʃ-reʔ > (tenseness and loss of final -ʔ) tʃ-re: > tʃre: > (reduction) je: (chày) ‘pestle’.

To sum up, xay ‘to husk (rice)’ is the result of an old process of reduplication of xáý ‘dig, hollow, excavate’, while chày derive from xáý by the infixation of -r-. All changes involved in the demonstrations are in keeping with regular phonetic laws.
The morphological pair ‘to husk (rice) - pestle’ in Austroasiatic

The vocabulary analyzed here comes from personal collected materials [Ferlus, Marie Martin] and from linguists’ publications [Sidwell, Zide, Diffloth, ...] as well as of non linguists’ ones [Baradat, Skeat & Blagden]. For the sake of convenience, the various linguistic reconstructions proposed in the literature are not reviewed here.

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.

<table>
<thead>
<tr>
<th>Group / Language</th>
<th>‘to husk’</th>
<th>‘to pound’</th>
<th>‘pestle’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VIETIC</strong> [Ferlus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTO VIET-MUONG</td>
<td>(tʃe? &gt;) tʃe:</td>
<td>(tʃre? &gt;) tʃe:</td>
<td></td>
</tr>
<tr>
<td>Viet</td>
<td>(xày &gt;) xày</td>
<td>chày</td>
<td></td>
</tr>
<tr>
<td>Mường [Nguyễn VK 2002]</td>
<td>saj1 (xày)</td>
<td>kʰaj2 (khày)</td>
<td></td>
</tr>
<tr>
<td>Cuội Chấm</td>
<td>saj¹</td>
<td>re:¹</td>
<td></td>
</tr>
<tr>
<td>Làng Lộ</td>
<td>saj¹</td>
<td>tʃe:¹</td>
<td></td>
</tr>
<tr>
<td>PROTO PONG</td>
<td></td>
<td>tʃp⁸</td>
<td>≈sərə:¹</td>
</tr>
<tr>
<td>Thavung</td>
<td>muːl¹</td>
<td>aʰə:¹</td>
<td></td>
</tr>
<tr>
<td>Sách</td>
<td>cuk⁷</td>
<td>oɾi:¹</td>
<td></td>
</tr>
<tr>
<td>Arem</td>
<td>tluh</td>
<td>ʷəɾi:</td>
<td></td>
</tr>
<tr>
<td>Maleng Kari</td>
<td>kəluː⁵⁶</td>
<td>soɾe:¹</td>
<td></td>
</tr>
<tr>
<td><strong>KATUIC</strong> [Ferlus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suei</td>
<td></td>
<td>n⁴ɾə:</td>
<td></td>
</tr>
<tr>
<td>Ong</td>
<td>kloh</td>
<td>n²ɾaj</td>
<td></td>
</tr>
<tr>
<td>Kantou</td>
<td>kloh</td>
<td>ɾnɾe:</td>
<td></td>
</tr>
<tr>
<td>Sò</td>
<td>cikloŋ</td>
<td>ɾnʈi:</td>
<td></td>
</tr>
<tr>
<td><strong>KATUIC</strong> [Sidwell]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTO KATUIC [2005]</td>
<td>kloh</td>
<td>?n⁴ɾee</td>
<td></td>
</tr>
<tr>
<td>Souei</td>
<td>ntap</td>
<td>ɾnɾe:</td>
<td></td>
</tr>
<tr>
<td>Sò/Bru</td>
<td>kloh</td>
<td>ɾnɾi</td>
<td></td>
</tr>
<tr>
<td><strong>BAHNARIC</strong> [Sidwell]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTO BAH. [1998]</td>
<td>pəh</td>
<td>?əɾɾaj</td>
<td></td>
</tr>
<tr>
<td>NORTH BAHNARIC [Sidwell]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTO NORTH BAH. [2002]</td>
<td>pəh</td>
<td>?əɾɾaj</td>
<td></td>
</tr>
<tr>
<td>Jeh</td>
<td>pəh¹</td>
<td>?əɾɾaj¹</td>
<td></td>
</tr>
<tr>
<td>Halang</td>
<td>pəh¹</td>
<td>həɾɾaj</td>
<td></td>
</tr>
<tr>
<td>Rengao</td>
<td>pəh¹</td>
<td>həɾɾi[L]</td>
<td></td>
</tr>
<tr>
<td>Sedang</td>
<td>pəj</td>
<td>draj[T]</td>
<td></td>
</tr>
<tr>
<td>Bahnar</td>
<td>pəh</td>
<td>hdroj</td>
<td></td>
</tr>
<tr>
<td>SOUTH BAHNARIC [Sidwell]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTO SOUTH BAH. [2000]</td>
<td>pəh</td>
<td>r-n-aj</td>
<td></td>
</tr>
<tr>
<td>Mnong</td>
<td>pəh</td>
<td>nə</td>
<td></td>
</tr>
<tr>
<td>Stieng</td>
<td>pəh</td>
<td>ranaj</td>
<td></td>
</tr>
<tr>
<td>Chrau</td>
<td>pəh</td>
<td>rənaj</td>
<td></td>
</tr>
</tbody>
</table>
### WEST BAHNARIC [Ferlus]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laven</td>
<td>tpeh</td>
<td>?rej</td>
</tr>
<tr>
<td>Nhaheun</td>
<td>ja?</td>
<td>?re:</td>
</tr>
<tr>
<td>Brao</td>
<td>tveh</td>
<td>raj</td>
</tr>
<tr>
<td>Sapouan</td>
<td>ja?</td>
<td>araj</td>
</tr>
<tr>
<td>Lave</td>
<td>tveh</td>
<td>araj</td>
</tr>
<tr>
<td>Cheng</td>
<td>tveh</td>
<td>raj</td>
</tr>
</tbody>
</table>

### WEST BAHNARIC [Sidwell, Jacques]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laven/Jru’</td>
<td>tōpeh</td>
<td></td>
</tr>
<tr>
<td>Nhaheun</td>
<td>ja?</td>
<td>?ree</td>
</tr>
<tr>
<td>Sapuan</td>
<td>ja?</td>
<td>?raj</td>
</tr>
</tbody>
</table>

### BOLYU [Edmondson 1995]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tōn</td>
<td>xwuŋk</td>
</tr>
</tbody>
</table>

### MANG

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tō:</td>
<td>tuŋ</td>
</tr>
</tbody>
</table>

### KHIMUIC [Ferlus]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khmu</td>
<td>hic</td>
<td>cnʷre?</td>
</tr>
<tr>
<td>Phay</td>
<td>kʰə:t</td>
<td>ɡgle?</td>
</tr>
<tr>
<td>Thin</td>
<td>kʰə:t</td>
<td>ɡgre?</td>
</tr>
<tr>
<td>Pray</td>
<td>kʰə:t</td>
<td>ɡgia?</td>
</tr>
<tr>
<td>Lamet</td>
<td>peh</td>
<td>ntoro:</td>
</tr>
<tr>
<td>Keneng</td>
<td>kal</td>
<td>kanreː</td>
</tr>
<tr>
<td>Hat</td>
<td>suʔ</td>
<td>ndraː</td>
</tr>
<tr>
<td>Khang</td>
<td>tepeː</td>
<td>heʔ</td>
</tr>
<tr>
<td>Kесing Mul</td>
<td>bǒk</td>
<td>hagēː</td>
</tr>
</tbody>
</table>

### PALAUNGIC [Ferlus]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>daʔak</td>
<td>aduh</td>
<td>ŋkrej</td>
</tr>
<tr>
<td>taʔaŋ</td>
<td>doh</td>
<td>greː</td>
</tr>
<tr>
<td>raʔaŋ</td>
<td>dìh</td>
<td>ɡloŋ achoŋ</td>
</tr>
</tbody>
</table>

### WAIC [Ferlus]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>pəzaʔk</td>
<td>tah</td>
<td>griʔ</td>
</tr>
<tr>
<td>vaʔ</td>
<td>kujh</td>
<td>ŋiʔ</td>
</tr>
<tr>
<td>Sem</td>
<td>taoh</td>
<td>glliʔ</td>
</tr>
<tr>
<td>Phalɔk</td>
<td>dəh</td>
<td>ŋiʔ</td>
</tr>
<tr>
<td>Samtao</td>
<td>tih</td>
<td>nreʔ</td>
</tr>
<tr>
<td>laviaʔ</td>
<td>blouh</td>
<td>kʰouʔ</td>
</tr>
<tr>
<td>La-oop</td>
<td>toh</td>
<td>greiʔ</td>
</tr>
<tr>
<td>Lawa</td>
<td>pouh</td>
<td>kʰo: toh</td>
</tr>
<tr>
<td>PROTO WA [Diffloth 1980]</td>
<td>toh</td>
<td>ŋriʔ</td>
</tr>
</tbody>
</table>

### RIAANG [Luce 1965]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rèʔ</td>
<td></td>
</tr>
</tbody>
</table>

### DANAW [Luce 1965]

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rèʔ</td>
<td></td>
</tr>
</tbody>
</table>

### MONIC

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nyah Kur [Theraphan 1984]</td>
<td>jáːk</td>
<td>ŋrìːʔ</td>
</tr>
</tbody>
</table>
KHMER

Khmer

\begin{align*}
\text{bok} & \quad \text{kin} \\
puk & \quad \text{kin}
\end{align*}

\begin{align*}
?\text{ŋre:} & \quad 7\text{ʔnræ}
\end{align*}

PEARIC [Baradat 1941]

Pear, Kpg Speu

\begin{align*}
\text{chhák} & \quad \text{ken} \\
\text{bok} & \quad \text{ken}
\end{align*}

\begin{align*}
\text{rōhi-i} & \\
\text{rê}
\end{align*}

Pear, Kpg Thom

\begin{align*}
\text{chhák} & \quad \text{ken}
\end{align*}

\begin{align*}
\text{rōhi-i}
\end{align*}

Pear, west

\begin{align*}
\text{chhák}
\end{align*}

\begin{align*}
\text{rōhik}
\end{align*}

Pear, east

\begin{align*}
\text{chhák}
\end{align*}

\begin{align*}
\text{rōhik}
\end{align*}

PEARIC [Martin]

Samray

\begin{align*}
\text{chuuk} & \quad \text{ken}
\end{align*}

\begin{align*}
(\text{rōhi-i})
\end{align*}

Samree

\begin{align*}
\text{chook} & \quad \text{kun}
\end{align*}

\begin{align*}
(\text{rōhik})
\end{align*}

PEARIC (various)

Pear [Headley 1978]

\begin{align*}
\text{čha:k}
\end{align*}

\begin{align*}
\text{rōhi:}
\end{align*}

Saaoch

\begin{align*}
\text{tʰa:k}
\end{align*}

\begin{align*}
\text{ʁi}
\end{align*}

Chong [Siriphen 2001]

\begin{align*}
\text{cʰɔːk}[^{R1}]
\end{align*}

\begin{align*}
\text{kəhi}[^{R1}][\text{kaˈhi:}]
\end{align*}

KHASI [Singh 1920]

\begin{align*}
\text{synrei}
\end{align*}

ASLIAN

Jahai [Burenhult 2001]

\begin{align*}
\text{sntip/tiʔ/sih/patim/tił}
\end{align*}

\begin{align*}
\text{gul}
\end{align*}

Tembi [Skeat & B. 1906]

\begin{align*}
\text{rentik}
\end{align*}

Serau [Skeat & B. 1906]

\begin{align*}
\text{kənəh, kənuʔ(?)}
\end{align*}

NICOBAR

- 

NORTH MUNDA [Zide 1976]

Korku

\begin{align*}
\text{rum-}
\end{align*}

\begin{align*}
\text{toko} / \text{tuki}
\end{align*}

Ho

\begin{align*}
\text{ruŋ-}
\end{align*}

- 

Santali

\begin{align*}
\text{ruŋŋ-}
\end{align*}

- 

Santali [Macphail 1954]

\begin{align*}
\text{huŋŋ}
\end{align*}

\begin{align*}
\text{sok’}
\end{align*}

\begin{align*}
\text{tok}
\end{align*}

SOUTH MUNDA [Zide 1976]

Kharia

- 

\begin{align*}
\text{e(n)ɾi} / \text{eɾndi}
\end{align*}

Remo

- 

\begin{align*}
\text{tɾeʔ}
\end{align*}

Gtaʔ

- 

\begin{align*}
\text{tɔŋkæ}
\end{align*}

Gorum

- 

\begin{align*}
\text{tanɾləd}
\end{align*}

\begin{align*}
\text{in(d)ɾi}
\end{align*}

Sora

- 

\begin{align*}
\text{tanɾləd}
\end{align*}

\begin{align*}
\text{ŋɾij}
\end{align*}

PROTO MON-KHMER

[Shorto 2006]

\begin{align*}
\text{pis/pəs}
\end{align*}

\begin{align*}
[k]bok
\end{align*}

\begin{align*}
\text{nrəʔ?} / \text{nrəʔy}
\end{align*}

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 (Cuôii Chám, Lãng Lở), 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 arōšh, Teressa aros) is genuinely Portuguese [de Röepstorff 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.

PVM tʃeːw (xeo) and tʃ-r-eːw > jeːw (chèo) must be reconstructed.

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

The word chèo, 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 crɔvɑː cravā. In Lao we find sɛːwʰ (<*jeːw) ແຊວ ‘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:

tʃeːw > (infixation) tʃ-ɾ-n-eːw > (monosyllabization) nɛːw neo ‘anchor’.

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

- sɛk - srɛɨk ‘to comb - a comb’
- tajə - traŋjə ‘to light with a steel lighter - a lighter’
- kɔ̃ - kɾnpə ‘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-:

- dăn - nan ‘to plait - bamboo split’
- đút - nut ‘to cork (a bottle) - a cork’
- chọc - pọc ‘to shake down (with a long pole) - a long pole’
- xếp - nɛp ‘to fold - a fold’

9

xum ‘gather, form group’ > chūm ‘bunch, cluster’

xùm ‘gather, form groups’ > chūm ‘assemble, gather’

The place of xum in dictionaries needs some further remarks. Xum is not attested in the modern Vietnamese dictionaries, while in others, xum and xùm 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 *xím ‘se réunir, réunir, rassembler, convoquer, grouper’. This distinction can be interpreted as *xum ‘to meet, to get together’, with an intransitive meaning, and *xím ‘to gather, to collect, to call together’ with a causative aspect.

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

*tfjúm > (prefixation) *p-*tfjúm > (tenseness and glottalization) *p-tfjúm > (monosyllabization) *tfjúm (xím). On the circumstances of the occurrences of glottalization in sesquisyllables, see Ferlus [2004].

Formation of derivatives with the infix -r-: *tfj-r-úm > *júm (chúm ‘bunch, cluster’) and (p-)*tfj-r-úm > *júm (chum ‘assemble, gather’).

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

The prefixed form passed in Khmer, pracum prajum, then in Thai pracbúm^A2 ประชุม and in Lao, pasum^A2 ประชุม.

10  | xia ‘pick, jab, to put on a step’ > chăa ‘pitchfork, trident’

PVM *tʃeh (xia) and *tʃ-r-eh > *jeh (châa) must be reconstructed.

Derivative formed with -rn- infix: *tʃeh > (infixation) *tʃ-rn-eh > (monosyllabization)

neh nā ‘fork’.

These words remain confined in the Vietnamese area.

11  | xō ‘sting, pierce’ > chô ‘pan to cook sticky rice’

PVM *tʃɔh (xō) and *tʃ-r-ɔh > *jɔh (chô) must be reconstructed.

These words remain confined in the Vietnamese area.

12  Conclusions

The PVM proto phoneme *tʃ is specific to the Vietnamese language and to some very close VM languages. Words opening with the initial *tʃ (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 - chây, (2) xẹo - chẹo, (3) xum/xím - chúm/chùm, (4) xia - chăa and (5) xō - chô. 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 - chây. It was explained how from PVM *tʃe? (xáy) ‘to dig, excavate’ was formed the derivative *tʃe: (xay) ‘to husk (rice)’ with a more specialized meaning, and also was formed *tʃre? > je: (chây) ‘pestle’. 
It was also noted that, in the primordial PVM pair tʃeʔ-tʃreʔ, the reflexes of the basic verb (tʃeʔ→ tʃe: ‘to husk (rice)’ remained restricted to Vietnamese, while the reflexes of the derivative *tʃreʔ ‘pestle’ spread to most Austroasiatic languages. Bolyu, Māng, Aslian, Nicobarese, North Munda and some languages of South Munda did not receive that derivative. We are facing 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 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, which was technically superior to all earlier methods for 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 was presumably used before. The other advantage is that utensils made of wood are easier to make than those made 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 common actions, the nominal derivatives of which are utensils or concepts useful in 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 Đōng 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 Dongsonian’ language.

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

References
Ferlus, Michel. *Unpublished materials on several Mon-Khmer languages, specially Viet-Muong (Vietic) languages, collected in Laos, Thailand, Burma and Vietnam*.


Ferlus, Michel. 2006. What were the four Divisions (đèng 等) of the Middle Chinese. *The 39th International Conference on Sino-Tibetan Languages and Linguistics*, University of Washington at Seattle, September 14-17.


Martin, Marie A. *Unpublished materials on Pearic languages*.


A rice-husking scene engraved on a Dongsonian bronze drum [Parmentier 1918: *Pl. IV, fig. l*]. Museum of History in Hanoi.
### Summarized chart: ‘to husk (rice) - pestle’ in Austroasiatic

<table>
<thead>
<tr>
<th>Groups/Languages</th>
<th>to husk (rice)</th>
<th>to pound</th>
<th>pestle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PROTO VIET-MUONG</strong></td>
<td>(tʃe? &gt;) tʃe: (xay &gt;) xay</td>
<td>tuːp²</td>
<td>(tʃe? &gt;) je: chày kʰaj² (khày)</td>
</tr>
<tr>
<td>Viet Mường Bi</td>
<td>saj¹ (xay)</td>
<td>tuːp</td>
<td>əri:¹</td>
</tr>
<tr>
<td>Sách</td>
<td>cuk⁷ tluh</td>
<td>tuːp²</td>
<td>əri:¹</td>
</tr>
<tr>
<td>Arem</td>
<td></td>
<td>tuːp²</td>
<td>əri:¹</td>
</tr>
<tr>
<td><strong>PROTO KATUIC</strong></td>
<td>kloh</td>
<td></td>
<td>?n⁴ree</td>
</tr>
<tr>
<td><strong>PROTO BAHNARIC</strong></td>
<td>pəh</td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td><strong>PROTO NORTH BAH.</strong></td>
<td></td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td>Rengao</td>
<td>pəh</td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td>Bahnar</td>
<td>pəh</td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td><strong>PROTO SOUTH BAH.</strong></td>
<td>pəh</td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td>Stieng</td>
<td></td>
<td></td>
<td>?onəj/r(ə)naj</td>
</tr>
<tr>
<td><strong>PROTO WEST BAH.</strong></td>
<td>tʔpeh</td>
<td>jaʔ</td>
<td>?raj</td>
</tr>
<tr>
<td>Laven/Jru’</td>
<td>tʔpeh</td>
<td>jaʔ</td>
<td>?raj</td>
</tr>
<tr>
<td><strong>BOLYU</strong></td>
<td>ton⁵³</td>
<td>xuɔk³¹</td>
<td></td>
</tr>
<tr>
<td><strong>MANG</strong></td>
<td>toː</td>
<td></td>
<td>tuŋ</td>
</tr>
<tr>
<td><strong>KHMUIC</strong></td>
<td>hic</td>
<td></td>
<td>?n⁴ree</td>
</tr>
<tr>
<td>Khmu</td>
<td>kʰə:t</td>
<td></td>
<td>?n⁴ree</td>
</tr>
<tr>
<td>Thin</td>
<td>kal</td>
<td></td>
<td>?n⁴ree</td>
</tr>
<tr>
<td>Keneng</td>
<td></td>
<td></td>
<td>?n⁴ree</td>
</tr>
<tr>
<td><strong>PALAUNGIC təʔanj</strong></td>
<td>dɔh</td>
<td></td>
<td>gre:</td>
</tr>
<tr>
<td><strong>PROTO WAIC</strong></td>
<td>toh</td>
<td></td>
<td>ɲrǐ</td>
</tr>
<tr>
<td><strong>RIANG</strong></td>
<td>-</td>
<td>-</td>
<td>rɛʔ</td>
</tr>
<tr>
<td><strong>MÓN</strong></td>
<td>yàik [jəc]</td>
<td></td>
<td>r̥ʔ</td>
</tr>
<tr>
<td><strong>KHMER</strong></td>
<td>bok</td>
<td>kvn</td>
<td>?ŋreː</td>
</tr>
<tr>
<td><strong>PEARIC</strong></td>
<td>tʰaː:k</td>
<td></td>
<td>ri kɔhiːR¹[kə’hiː]</td>
</tr>
<tr>
<td>Sa och</td>
<td>cʰɔ:kR¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chong</td>
<td>bɔt</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KHASI</strong></td>
<td></td>
<td></td>
<td>synrei</td>
</tr>
<tr>
<td><strong>ASLIAN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jahai</td>
<td>sntip/tiʔ/sih/…</td>
<td></td>
<td>gul</td>
</tr>
<tr>
<td>Tembi</td>
<td></td>
<td></td>
<td>rentik</td>
</tr>
<tr>
<td><strong>NICOBAR</strong></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>NORTH MUNDA</strong></td>
<td>rum-</td>
<td>sɔk’</td>
<td>toko / tuki</td>
</tr>
<tr>
<td>Korku</td>
<td>hurug</td>
<td></td>
<td>tok</td>
</tr>
<tr>
<td>Santali</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SOUTH MUNDA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td>-</td>
<td>taŋlad</td>
<td>ə(n)ɾi/əndi</td>
</tr>
<tr>
<td>Sora</td>
<td>-</td>
<td>taŋlad</td>
<td>əɾri</td>
</tr>
<tr>
<td><strong>PROTO MON-KHMER</strong></td>
<td>pis’pɔs</td>
<td>[k]bok</td>
<td>nɾəʔ/nɾəŋ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Map showing the expansion of Proto Viet-Muong *tʃre? 'pestle'

- Cradle of *tʃre?
- Expansion of *tʃre?
- Area of Việt chọi
- Others