

Mon-Khmer Studies

VOLUME 43.1

Papers from the 5th International
Conference on Austroasiatic Linguistics
(Canberra, September 4-5, 2013)

The journal of Austroasiatic
languages and cultures
1964—2014 50 years of MKS

Author: Michel FERLUS
Title: *Arem, a Vietic Language.*
Pages: 1-15

Copyright for this paper vested in the author
Released under Creative Commons Attribution License

Volume 43 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: <http://mksjournal.org>

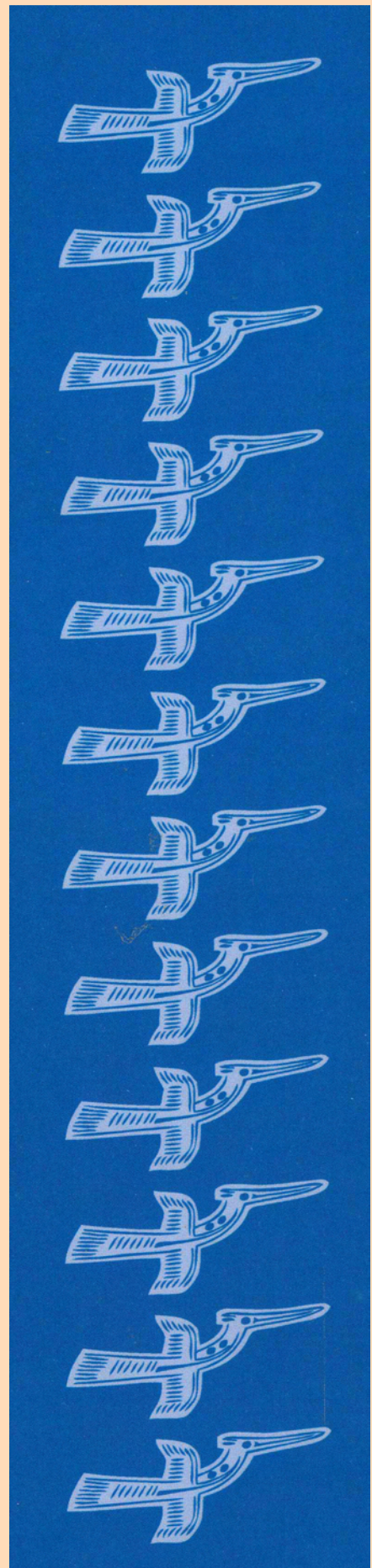
Published by:



Mahidol University (Thailand)



SIL International (USA)



Arem, a Vietic Language

Michel FERLUS

CNRS, France (*retired*)

<jrmferlus@orange.fr>

Abstract

The paper reports on the state of knowledge of Arem, a small endangered language of the Vietic branch of Austroasiatic. Special attention is paid to the synchronic and diachronic phonology, which display many remarkable and highly significant features. The paper is richly illustrated with lexical examples.

Keywords: Vietic languages, lexicon, phonology

ISO 639-3 codes: aem, ptk, vie

1. Introduction

The Arem people are located in the Vietnam's Quảng Bình province, Bồ-trạch district, Tân-trạch commune, Cà Ròng hamlet. Though known of neighboring populations, they were confused with Bru khùa (Katuić) by local authorities: after their presence was reported by the military in 1959, they were eventually identified as a separate ethnic and linguistic group by ethnologists.

Demographics: around 1960, there were 53 people, 30 men and 23 women (Vương Hoàng Tuyên 1963: 69). Recently, there were 102 people (Trần Trí Dõi 1999: 110). In the past, Arem lived in rock shelters or makeshift huts. They exchanged products from hunting and gathering with the neighboring populations, Vietnamese, Bru khùa or Lao. The need to find marital partners outside the group has the consequence that only a minority of Arem, about a quarter, regularly speak their language. The Arem language is in great danger of extinction.

Ethnonym: the term *Arem* (Vietnamese spelling of **are:m**) seems to be an exonym. Those concerned prefer to be named *cmbrâu* [**cmraw**?]. At the end of the 19th century, Captain Rivière reported more than one hundred *Harème* families, sometimes corrupted as *Hémère*, on the Lao side (*Mission Pavie, Géographie et Voyages*, 1902; voc. pp. 285-90). *Harème* has been wrongly identified to Arem (Haudricourt 1966: 136) as recently revealed by Vương Hoàng Tuyên (1963). An examination of the vocabulary of numbers as preserved in Rivière's note clearly shows that *Harème* should be classified in the Maleng subgroup.

	<i>Harème</i>	Maleng brô	Arem		<i>Harème</i>	Maleng brô	Arem
'one'	<i>moitte</i>	<i>mo:c</i>	<i>mù:c</i>	'six'	<i>prao</i>	<i>paraw</i>	<i>paraw</i> [?]
'two'	<i>harr</i>	<i>ha:r</i>	<i>hæ:l</i>	'seven'	<i>paé</i>	<i>paj</i>	<i>pɰh</i>
'three'	<i>pà</i>	<i>pa:</i>	<i>pæ:</i>	'eight'	<i>xâme</i>	<i>sa:m</i>	<i>t^hæ:m</i> [?]
'four'	<i>pône</i>	<i>po:n</i>	<i>puən</i> [?]	'nine'	<i>gîne</i>	<i>ci:n</i>	<i>ci:n</i> [?]
'five'	<i>damm</i>	<i>dâm</i>	<i>dam</i>	'ten'	<i>meû il</i>	<i>mə:j</i>	<i>"cùk</i>

In fact, Arem is an ethnographic term without any clear linguistic relevance: it was apparently used to refer to various semi-nomadic populations moving in the Great Cordillera, between Vietnam and Laos, about the latitude of the Mụ Giạ Pass. Regarding *cmbrâu* [**cmraw**?], the pre-syllable **cm-**, which is attested only in this word, suggests that this ethnonym was borrowed from another language.

Researches on Arem: the data that serves as the basis for the present text were collected during three field trips (June 1991, December 1993, and March 1996) conducted in collaboration with Professor Trần Trí Dõi of Vietnam National University in Hanoi. The Arem language is not homogeneous, the pressure of neighboring languages is exerted on the individuals rather than on the group. The speakers are at least bilingual with Vietnamese, standard as well as dialectal, and some people can also speak Khùa or Lao.

Publications and various data on Arem people and language: the first linguistic data were published by Vương Hoàng Tuyên (1963). They are sufficient to classify Arem in the Viet-Muong

group. More recently, Trần Trí Dõi (1999) cited some vocabulary. It is heartening to mention that Arem, an endangered language, had an academic consecration through the Master of Arts of Atsushi Kasuga (1994), unpublished so far. Consistent ethnographic data should be noticed: Vương Hoàng Tuyên (1963), Đặng Nghiêm Vạn, Chu Thái Sơn, Lưu Hùng (1986), and Trần Trí Dõi (1995, 1999). We chose for our description, the speech of Arem which seemed the most conservative regarding the phonation type register.

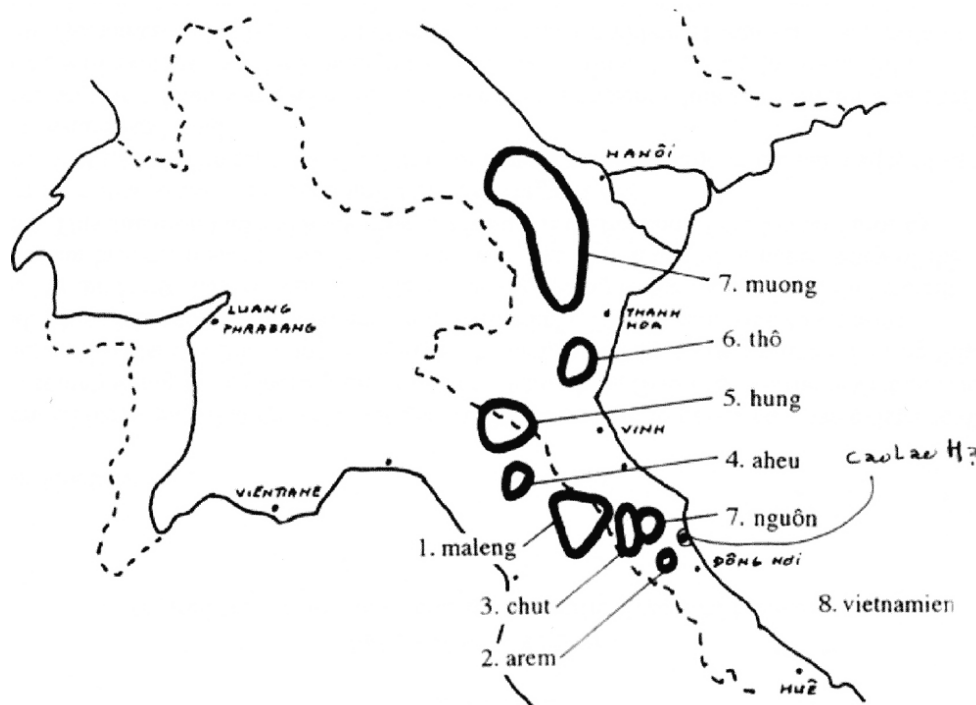


Figure 1: Map of Vietic languages showing approximate locations.

2. Phonetic system

2.1 Syllabic structure

The vocabulary of Arem has monosyllables and sesqui-syllables.

monosyllable: C V (C)
 sesqui-syllable: C-C V (C)

A sesqui-syllable is composed of a main syllable C V (C) preceded by a pre-syllable C-. The main syllable is identical to a monosyllable. The pre-syllable is short and unstressed, without distinctive vowel, and its consonantal system is reduced. A pre-syllable can not exist alone. Arem is the language of the Viet-Muong group that has preserved the highest percentage of sesqui-syllable (prenasalized being considered as initials sesqui-syllables). The percentage of retention in subgroups increases from north to south.

Vietnamese	monosyllabic
Muong/mường	monosyllabic
Thổ	monosyllabic
Toum-Liha	monosyllabic
Pong	10% of sesqui-syllabic
Thavung	35-40%
Maleng Brô	35-40%
Sách-Rục	35-40%
Arem	55-60%

2.2 Voice-type register system of Arem:

- Arem has two sets of vowels: Series 1 with a high-clear register phonation; Series 2, with a low-breathy register phonation (indicated by the grave accent).

- In each series, we must distinguish the syllables ending in voiced finals (# w j l m n ɲ ɳ), and the syllables ending in voiceless finals (-p -t -c -k) and (-h).

- In the syllables with voiced finals, the series 1 is indicated by /a - a^ʔ/, and corresponds to tones *ngang-sắc* in Vietnamese. The series 2 is indicated by /à - à^ʔ/ and corresponds to *huyền-nặng*.

- Final -h is preserved in the both series; /ah/ corresponds to tone *hỏi*, and /àh/ to *ngã*.

- Syllables in /aC/ and /àC/ are checked by voiceless finals (-p -t -c -k); they respectively correspond to tones *sắc* and *nặng*.

- Syllables in voiced finals show a contrast between a modal tone and a glottalized tone: in series 1 /a~a^ʔ/, and in series 2 /à~à^ʔ/, and in series 2 /à~à^ʔ/. It should be noted that [ʔ] indicates a glottal constriction of the vowel, also named creakiness, but by no means a creaky voice. I think that this phonetic feature is not a voice type register, just a weak glottal stop.

- The combination of the contrast /clear~breathy/ and /unchecked~glottal/ forms a four tones system:

tone *a*: high, clear, unmarked (corresponds to tone *ngang*)

tone *a^ʔ*: high, glottalized, slightly raising (corresponds to tone *sắc*)

tone *à*: low breathy (corresponds to tone *huyền*)

tone *à^ʔ*: low breathy glottalized (corresponds au to tone *nặng*)

Arem		voiced finals (# w j l m n ɲ ɳ)		unvoiced finals (h) / (p t c k)	
series 1 (hight-clear)		a	a ^ʔ	ah	aC
series 2 (low-breathy)		à	à ^ʔ	àh	àC

Vietnamese		voiced finals		unvoiced finals	
series 1 (tones <i>a á ả</i>)		<i>ngang</i>	<i>sắc</i>	<i>hỏi</i>	<i>sắc</i>
series 2 (tones <i>à ạ ã</i>)		<i>huyền</i>	<i>nặng</i>	<i>ngã</i>	<i>nặng</i>

Because of the phenomenon of vowel splitting in two series, it is difficult to find a same vowel represented with the four tones. Only rhymes with vowel *iə* provide enough possibilities of combinations.

tone *a*: *at^hiəl* ‘écaille / fish scale’

tone *a^ʔ*: *akiəl^ʔ* ‘sur le côté / on side’

tone *à*: *ciəl* ‘filet-épervier / cast net’

tone *à^ʔ*: *patiəl^ʔ* ‘milieu / middle’

2.3 Initial consonants

Simple initials

p^h	t^h			k^h	
p	t	$tʃ$	c	k	$ʔ$
b	d		$ʃ$		
m	n		$ɲ$	$ŋ$	
	s	$ʃ$			h
v	(z)	$(ʒ/ʒ)$	j		
	l	r			

Clusters with r / l

pr	tr		cr	kr
br				
pl	tl			kl
	t^hl			

Pre-nasalized initials

	nth			nk	nh
mp	nt	ntʃ	nc	nk	nʔ
mb	nd		nʃ		
nv	nl	$^nr/^ndʒ$	nj		

2.4 Finals

				h
p	t		c	k
$m(?)$	$n(?)$		$ɲ(?)$	$ŋ(?)$
$w(?)$		$l(?)$	$j(?)$	$ʃ(?)$

Final liquid $-l$ [l] is vocalized in [ɛ] by some informants.

2.5 Pre-syllables

Note: infrequent pre-syllables are placed in the brackets:

$a-$ $i-$ $u-$

$pa-$ ($pi-$) ($ba-$) ($ma-$ / $mu-$)

$ta-$ (tam) N- (first segment of pre-nasalized)

$ci-$ (cm)

$ka-$ $ku-$ ($kə-$)

$la-$ ($le-$) ($ra-$ / $rə-$)

Pre-syllables in decreasing order, with the number of occurrences in the lexicon:

$a-$ (240), N- (141), $ka-$ (83), $u-$ (79), $ci-$ (54), $pa-$ (44), $ta-$ (36), $i-$ (30), $la-$ (30), $ku-$ (14).

2.6 Vowel system

breathy vowels					clear vowels						
<i>ì</i>	<i>ì:</i>	<i>î</i>	<i>î:</i>	<i>ù</i>	<i>ù:</i>	<i>i</i>	<i>(i:)</i>	<i>í</i>	<i>í:</i>	<i>u</i>	<i>u:</i>
						<i>ɪ</i>	<i>ɪ:</i>	#	#:	<i>ʊ</i>	<i>ʊ:</i>
	<i>ie</i>		<i>iẽ</i>		<i>ùo</i>	<i>ie</i>		<i>iẽ</i>		<i>uo</i>	
<i>è</i>	<i>(è:)</i>	<i>(ə̃)</i>	<i>ə̃:</i>	<i>ò</i>	<i>ò:</i>	<i>(e)</i>	<i>e:</i>	<i>ə</i>	<i>ə:</i>	<i>o</i>	<i>o:</i>
		<i>è̃</i>	<i>è̃:</i>			<i>ɛ</i>	<i>ɛ:</i>	(ɐ)	ɐ:	<i>ɔ</i>	<i>ɔ:</i>
						<i>æ</i>	<i>æ:</i>	ʌ	ʌ:	<i>a</i>	<i>a:</i>
								<i>a</i>	<i>a:</i>		

Units placed between brackets are poorly represented. The vowel **æ:** is realized [ɛa]. The difference between **(ɐ)/** and **ʌ/ʌ:** has no etymological justification. The series of high vowels /**i i: í í:** **u u:**/ occurs only in Vietnamese words or in Arem words whose pronunciation is influenced by Vietnamese pronunciation. This series is not in contrast with the regular series /**ɪ ɪ: ʊ ʊ:**/.

3. Diachrony: identification of some pairs of vowels contrasting by voice type

A pair of vowels contrasting by voice type, in short a registral pair, is composed of two vowels (one in each series) that derived from one single proto vowel. It is particularly useful to identify these pairs for the reconstruction of the vowel system of Arem that existed before the formation of series. We will proceed by comparing Arem to Vietnamese and Proto Viet-Muong. In some cases one will have to distinguish between primary and secondary correspondences.

3.1. PVM **a:**

PVM	Arem	Viet
		<i>a - ủa/ớ</i>

primary (regular) correspondences

<i>r.ka:</i>	<i>lakæ:</i>		<i>gà</i>	‘poulet / chicken’
<i>s.la:?</i>	<i>ulæ:?</i>		<i>lá</i>	‘feuille / leaf’
<i>ha:r</i>	<i>hæ:l</i>		<i>hai</i>	‘deux / two’
<i>sa:m?</i>	<i>t^hæ:m?</i>		<i>tám</i>	‘huit / eighth’
<i>də:k</i>	<i>dæ:k</i>		<i>nước</i>	‘eau / water’
<i>k.ma:</i>		<i>mìə</i>	<i>mưa</i>	‘pluie / rain’
<i>ʃa:l</i>		<i>ciəl</i>	<i>chài</i>	‘filet-épervier / cast net’
<i>la:s</i>		<i>liəh</i>	<i>lưỡi</i>	‘langue / tongue’
<i>tʃa:ŋ/ʃ.ʔa:ŋ</i>		<i>siəŋ</i>	<i>xương</i>	‘os / bone’

secondary correspondences

<i>la:c</i>		<i>lè:t</i>	<i>lạt</i>	‘lien / bond’
<i>(la:ŋ)</i>		<i>lè:ŋ</i>	<i>làng</i>	‘village / village’
<i>gə:n?</i>		<i>kè:n?</i>	<i>cạn</i>	‘peu profond / shallow’
<i>kr.ha:k</i>		<i>kah̃:k</i>	<i>khác</i>	‘cracher / to spit’

PVM a: (before palatals -c -ɲ -j)

PVM	Arem		Viet	
<i>a:</i>	<i>a:</i>	<i>è:</i>	<i>a - ươ</i>	
<i>t.ka:c</i>	<i>taka:c</i>		<i>cát</i>	‘sable / sand’
<i>t.sa:c</i>	<i>at^ha:c</i>		<i>tát</i>	‘écoper (de l'eau) / to bail out’
<i>ta:ɲ</i>	<i>ta:ɲ</i>		<i>đan</i>	‘tisser / to weave’
<i>ma:ɲ[?]</i>	<i>ma:ɲ[?]</i>		<i>mượn</i>	‘emprunter / to borrow’
<i>C.la:j[?]</i>	<i>la:j[?]</i>		<i>lưới</i>	‘filet de pêche / fishing net’
<i>s.ɲa:j[?]</i>	<i>ɲa:j[?]</i>		<i>ngái</i>	‘loin de / far from’
<i>sa:j</i>	<i>t^ha:j</i>		<i>tai</i>	‘oreille / ear’
<i>la:j[?]</i>		<i>lè:j[?]</i>	<i>lại</i>	‘revenir / to return’

PVM a: (before -w)

PVM	Arem		Viet	
<i>a:</i>	<i>ɛ:</i>	<i>è:</i>	<i>a</i>	
<i>ca:w[?]</i>	<i>ce:w[?]</i>		<i>cháo</i>	‘brouet de riz / rice gruel’
<i>k.ta:w</i>	<i>te:w</i>		<i>dao</i>	‘sabre / sword’
<i>?a:w[?]</i>	<i>?e:w[?]</i>		<i>áo</i>	‘veste / upper garment’
<i>(ga:w[?])</i>		<i>kè:w[?]</i>	<i>cạo</i>	‘se raser / to shave’
<i>(la:w)</i>		<i>lè:w</i>	<i>lào</i>	‘Lao (peuple) / Lao (people)’
<i>(C.la:w)</i>		<i>tè:w</i>	<i>trào</i>	‘déborder / to overflow’

3.2. *PVM i:*

PVM	Arem		Viet	
<i>i:</i>	<i>i: / i:</i>	<i>ì:</i>	<i>ay - i</i>	
<i>si:</i>	<i>t^hi:</i>		<i>tay</i>	‘main / hand’
<i>p.ri:</i>	<i>pari:</i>		<i>say</i>	‘ivre / be drunk’
<i>ci:m</i>	<i>ci:m</i>		<i>chim</i>	‘oiseau / bird’
<i>ⁿsi:t</i>		<i>t^hi:t</i>	<i>thít</i>	‘viande / flesh’
<i>di: / ti:</i>		<i>tì:</i>	<i>đi</i>	‘aller / to go’
<i>ʰi:[?]</i>		<i>ci:[?]</i>	<i>chí</i>	‘sœur aînée / elder sister’
<i>ci:n[?]</i>	<i>ci:n[?]</i>		<i>chín</i>	‘neuf (9) / nine’
<i>ci:n[?]</i>	<i>ci:n[?]</i>		<i>chín</i>	‘cooked, ripe / cuit, mûr’

3.3. *PVM e:*

PVM	Arem		Viet	
<i>e:</i>	<i>i: / iə</i>	<i>iə</i>	<i>ê - ay</i>	
<i>de:p</i>	<i>di:p</i>		<i>nếp</i>	‘(riz) glutineux / sticky (rice)’
<i>tʰ.re:</i>	<i>ⁿri:</i>		<i>chày</i>	‘pilon / pestle’
<i>k.cet</i>	<i>ci:t</i>		<i>chết</i>	‘mourir / to dead’
<i>kp.cet</i>	<i>kacit</i>		<i>giết</i>	‘tuer / to kill’
<i>-se:l</i>	<i>at^hiəl</i>		<i>tê (tê)</i>	‘écaille, pangolin / scale, anteater’

In the tense series, we see that PVM **e:** is treated as **i:**

3.4. PVM *ɛ:*

PVM	Arem		Viet	
<i>ɛ:</i>	<i>iə</i>	<i>iə</i>	<i>e</i>	
<i>t.kɛ:t</i>	<i>kiət</i>		<i>ghét</i>	‘détester, hair / to hate’
<i>dɛ:n</i>		<i>tiən</i>	<i>đen</i>	‘chandelle / candle’

Data on PVM *ɛ:* are insufficient.

3.5. PVM *u:*

PVM	Arem		Viet	
<i>u:</i>	<i>aw / u:</i>	<i>ù:</i>	<i>au/u - ô/âu</i>	
<i>p.ru:?</i>	<i>paraw?</i>		<i>sáu</i>	‘six / six’
<i>cu:?</i>	<i>caw?</i>		<i>cháu</i>	‘petit-fils / grand son’
<i>t.mu:?</i>	<i>maw?</i>		<i>máu</i>	‘résine, sang /
<i>t.pu:j</i>	<i>pɔ:j</i>		<i>vui</i>	‘joyeux / happy’
<i>C.hu:l</i>	<i>pahɔ:l</i>		<i>hôi</i>	‘sueur / sweat’
<i>k.pu:r</i>	<i>apɔ:l</i>		<i>vôi</i>	‘chaux / lime’
<i>p.tu:c</i>	<i>utu:t</i>		<i>đốt</i>	‘nœud / node’
<i>tu:c</i>	<i>tɔ:c</i>		<i>đốt</i>	‘allumer (feu) / to lit (fire)’
<i>C.bu:c</i>	<i>cibɔ:c</i>		<i>mút</i>	‘sucrer / to suck’
<i>t.lu:m?</i>		<i>lù:m?</i>	<i>trộm</i>	‘voler / to steal’
<i>C.ru:</i>		<i>cirù:</i>	<i>sâu</i>	‘profond / deep’
<i>bu:?</i> / <i>pu:?</i>		<i>pù:?</i>	<i>bú</i>	‘téter (au sein) / to suck (breast)
<i>mu:c/mo:c</i>		<i>mù:c</i>	<i>một</i>	‘un / one’
<i>gu:r?</i> / <i>ku:r?</i>		<i>kù:l?</i>	<i>cúi</i>	‘porc / pig’
<i>C.gu:r?</i> / <i>C.ku:r?</i>		<i>cikù:?</i>	<i>gấu</i>	‘ours / bear’

3.6. PVM *o:*

PVM	Arem		Viet	
<i>o:</i>	<i>uə</i>	<i>ùə</i>	<i>ô</i>	
<i>po:n?</i>	<i>puən?</i>		<i>bốn</i>	‘quatre / four’
<i>k.lo:n?</i>	<i>tluən?</i>		<i>trốn</i>	‘se cacher / to hide’
<i>k.mo:l?</i>	<i>umùəl?</i>		<i>mối</i>	‘termite / white ant’
<i>po:ŋ</i>		<i>ṃpùəŋ</i>	<i>bông</i>	‘fleur, floraison / blossom’
<i>C.ko:l?</i>	<i>akuəl?</i>		<i>gối</i>	‘oreiller, appui-tête / pillow’
<i>k.no:l?</i>	<i>tanuəl?</i>		<i>gối</i>	‘oreiller, appui-tête / pillow’

In these two examples PVM *u:* is treated as *o:*

<i>u:</i>	<i>uə</i>	<i>ùə</i>	<i>ô</i>	
<i>t.ku:l?</i>	<i>lakuəl?</i>		<i>gối</i>	‘genou / knee’
<i>ŋu:j</i>		<i>ŋùəj</i>	<i>ngôi</i>	‘s’asseoir / to sit down’

3.7. PVM ɔ:

PVM	Arem		Viet	
	<i>a:</i>	<i>ùə</i>	<i>o - ua/uô</i>	
<i>ʔa.cə:ʔ</i>	<i>aca:ʔ</i>		<i>chó</i>	‘chien / dog’
<i>ʔa.lə:ʔ</i>	<i>ala:ʔ</i>		<i>lúa</i>	‘paddy / unhusked rice’
<i>k.ləjʔ</i>	<i>tlajʔ</i>		<i>trói</i>	‘ligoter / to tie up’
<i>kən</i>	<i>ka:n</i>		<i>con</i>	‘fils, fille / son, daughter’
<i>t.kə:c</i>	<i>ka:c</i>		<i>gót</i>	‘tailler, ciseler / to carve, chisel’
<i>p.rək</i>	<i>para:k</i>		<i>sóc</i>	‘écureuil / squirrel’
<i>m.rəj</i>		<i>urùəj</i>	<i>ruôj</i>	‘mouche / a fly’
<i>C.pət</i>		<i>apùət</i>	<i>vuót</i>	‘lisser / to smooth’
<i>rə:c</i>		<i>rùəc</i>	<i>ruôt</i>	‘intestins / intestines’
<i>rəŋʔ</i>		<i>rùəŋʔ</i>	<i>ruông</i>	‘rizière / ricefield’

3.8. PVM a

PVM	Arem		Viet	
	<i>a</i>	<i>è</i>	<i>ă / a(y)</i>	
<i>k.carʔ</i>	<i>calʔ</i>		<i>cháy</i>	‘incendie / fire’
<i>C.pat</i>	<i>pat</i>		<i>vất</i>	‘presser / to squeeze’
<i>k.lam</i>	<i>tlam</i>		<i>trăm</i>	‘cent / hundred’
<i>s.ranʔ</i>	<i>ranʔ</i>		<i>sắn</i>	‘manioc / cassava’
<i>t.baŋ</i>	<i>abaŋ</i>		<i>măng</i>	‘pousses (bambou) / bamboo shoots’
<i>C.kamʔ</i>		<i>kəmʔ</i>	<i>gặm</i>	‘mordre / to bite’
<i>t.ŋarʔ</i>		<i>aŋəlʔ</i>	<i>ngáy</i>	‘ronfler / to snore’
<i>mat</i>		<i>mèt</i>	<i>mắt</i>	‘œil / eye’
<i>k.məŋʔ</i>		<i>mènʔ</i>	<i>mặn</i>	‘salé / salted’

In these two examples PVM ə is treated as a

<i>a / ə</i>	<i>a</i>	<i>è</i>	
<i>p.sat/-sət</i>	<i>at^hat</i>		‘éteindre / to extinguish’
<i>pər</i>	<i>pal</i>		‘voler / to fly’

3.9. PVM ə

PVM	Arem		Viet	
	<i>i / I / ɨ</i>	<i>ì / ì</i>	<i>ã / â</i>	
<i>p.səŋʔ</i>	<i>ut^hiŋʔ</i>		<i>rắn</i>	‘serpent / snake’
<i>pəŋʔ</i>	<i>pìŋʔ</i>		<i>bắn</i>	‘tirer (fusil, arbalète) / to shoot’
<i>pəc</i>	<i>pìc</i>		<i>bắt, bắt</i>	‘cueillir / to pick, pluck’
<i>təŋʔ</i>	<i>latəŋʔ</i>		<i>đững</i>	‘cloison / partition’
<i>bət</i>	<i>bət</i>		<i>mất</i>	‘perdre (objet) / to lose (object)’
<i>m.rəŋʔ</i>		<i>riŋʔ</i>	<i>rân</i>	‘pou / louse’
<i>məc</i>		<i>mít</i>	<i>mật</i>	‘fiel, bile / gall’
<i>t.kəŋ</i>		<i>kìŋ</i>	<i>gần</i>	‘près de / near of’
<i>k.vət</i>		<i>vít</i>	<i>quát</i>	‘fouetter / to whip’
<i>k.rəmʔ</i>		<i>karìmʔ</i>	<i>sấm</i>	‘tonnerre / thunder’
<i>jərʔ</i>		<i>jìlʔ</i>	<i>dậy</i>	‘se lever / to get up’

3.10. PVM u

PVM	Arem		Viet	
<i>u</i>	<i>u</i>		<i>o / ô</i>	
<i>-suk</i>	<i>ut^hok</i>		<i>tóc</i>	‘cheveux / hair’
<i>s.ʔuh</i>	<i>nuh</i>		<i>ô, tô</i>	‘nid / nest’

Not that data for PVM **u** are insufficient.

4. The influence of a Vietnamese dialect on Arem

One can note the unexpected correspondence between pairs with tones *hỏi-ngã* in Viet, and pairs of type *a²-â²* in Arem. If these correspondences were regular, Arem words should have the final **-h**.

PVM	Arem		Viet	
<i>-h</i>	tone <i>a²</i>	tone <i>â²</i>	<i>hỏi-ngã</i>	
<i>bəs</i>	<i>baj²</i>		<i>bây</i>	‘piège, tonilière / a trap’
<i>ka:s</i>	<i>kə:l²</i>		<i>cái</i>	‘moutarde / mustard’
...	<i>t^haŋ²</i>		<i>thẳng</i>	‘droit / straight’
...	<i>ròŋ²</i>		<i>rông</i>	‘vide / empty’
<i>kuh</i>	<i>ku²</i>		<i>củ</i>	‘tubercule / tuber’
<i>fòh</i>	<i>fo:²</i>		<i>nhô</i>	‘arracher, désherber / to weed’
<i>k.cəh</i>	<i>kace:²</i>		<i>giẻ</i>	‘torchon / duster’
<i>k.pa:s</i>	<i>kupə:l²</i>		<i>vải</i>	‘éttoffe / cloth’
<i>cis</i>	<i>cin²</i>		<i>chí</i>	‘fil / thread’
<i>k.cìh</i>		<i>cì:²</i>	<i>giữ</i>	‘surveiller / to watch’
<i>k.cus</i>		<i>cù:²</i>	<i>chổi</i>	‘balai / broom’

Arem words that show these unexpected correspondences are probably borrowings from a local Vietnamese dialect, in which the expected tones corresponding regularly to tones *hỏi-ngã* have merged to tones *sắc-nặng* according to two modalities (see tables below). In type A, the expected tone *hỏi* merged in the tone *sắc*, while the expected tone *ngã* merged in tone *nặng*. This is the dialect to which Arem borrowed a part of his vocabulary. Note in passing, the existence of another dialect, the type B, in which the expected tones *hỏi-ngã* merged in the tone *nặng*.

type A		type B	
<i>ngang</i>	<i>sắc (hỏi)</i>	<i>ngang</i>	<i>sắc (hỏi)</i>
<i>huyền</i>	<i>nặng (ngã)</i>	<i>huyền</i>	<i>nặng (ngã)</i>

The existence of type A was revealed to me by the linguist Võ Xuân Trang in 1991. Type B is the dialect of Cao Lao Hạ (Ferlus 1995).

5. Correspondences of initial consonants

In order to present the comparative evolutions of initials in Arem and Vietnamese, we have classified PVM initials into five categories:

- Obstruents ***/p-b t-d c-j k-g/** and ***/s/** which have undergone a double processing in Vietnamese according to their position within the word when the change occurs (Ferlus 1982, 1992), as initial of monosyllable, or as medial in sesqui-syllable (see table 5.1.). Remember that in the 17th century, *v* was *ɸ / ɸě*, and *d* was *d / ɸ* (de Rhodes 1651, 1991).

- Pre-glottalized */**ʔ** **ɗ** **f**/
- Unproblematic units */**ʔ** **h** **m** **n** **ɲ** **v** **j** **r** **l**/
- Clusters with **-r-**
- Clusters with **-l-**

5.1. Table of the evolutions of PVM */**p-b t-d c-ʃ k-g**/ and */**s**/

PVM	Arem	Vietnamese		
		*monosyll.	*sesqui-syllable	
			modern	1651
<i>p - b</i>	<i>p</i>	<i>b</i> [b]	<i>v</i> [v]	< <i>ɸ/ɸě</i> [β]
<i>t - d</i>	<i>t</i>	<i>đ</i> [d]	<i>d</i> [z/j]	< <i>d/dě</i> [ð]
<i>c - ʃ</i>	<i>c</i>	<i>ch</i> [c]	<i>gi</i> [z/j]	< <i>gi</i> [j]
<i>k - g</i>	<i>k</i>	<i>c/k</i> [k]	<i>g/gh</i> [ɣ]	< <i>g/gh</i> [ɣ]
<i>s</i>	<i>t^h</i>	<i>t</i> [t]	<i>r</i> [r/z/j]	< <i>r</i> [r]

Labials **p-b**

PVM	Arem	Vietnamese		
<i>p-b</i>	<i>p</i>	<i>b</i>	<i>v</i>	
<i>pa:</i>	<i>pæ:</i>	<i>ba</i>		‘trois / three’
<i>pəp[?]</i>	<i>pjɲ[?]</i>	<i>bắn</i>		‘tirer (fusil, arbalète) / to shoot’
<i>po:n[?]</i>	<i>puən[?]</i>	<i>bốn</i>		‘quatre / four’
<i>k.pa:s</i>	<i>kupɛ:l[?]</i>		<i>vải</i>	‘éttoffe / fabric’
<i>k.pu:r</i>	<i>apɯ:l</i>		<i>vôi</i>	‘chaux / lime’
<i>k.pa:[?]</i>	<i>pæ:[?]</i>		<i>vá</i>	‘rapiécer, réparer / to mend, repair’

Alveolars **t-d**

<i>t-d</i>	<i>t</i>	<i>đ</i>	<i>d</i>	
<i>tɯ:c</i>	<i>tɯ:c</i>	<i>đốt</i>		‘allumer / to light’
<i>ta:ɲ</i>	<i>ta:ɲ</i>	<i>đan</i>		‘tisser / to weave’
<i>l.ta:[?]</i>	<i>atæ:[?]</i>	<i>đá</i>		‘pierre / stone’
<i>k.ta:l[?]</i>	<i>atæ:l[?]</i>		<i>dái</i>	‘scrotum / scrotum’
<i>təɲ[?]</i>	<i>latəɲ[?]</i>		<i>đứng</i>	‘cloison / partition’
<i>k.ta:w</i>	<i>tɛ:w</i>		<i>dao</i>	‘sabre, couteau / sword, knife’

Palatals **c-ʃ**

<i>c-ʃ</i>	<i>c</i>	<i>ch</i>	<i>gi</i>	
<i>-ci:m</i>	<i>ici:m</i>	<i>chim</i>		‘oiseau / bird’
<i>k.car[?]</i>	<i>cal[?]</i>	<i>cháy</i>		‘incendie / fire’
<i>ʃa:l</i>	<i>ciəl</i>	<i>chài</i>		‘filet-épervier / cast net’
<i>k.ʃə:ɲ</i>	<i>kacɛ:ɲ</i>		<i>giường</i>	‘lit / bed’
<i>C.ca:m</i>	<i>cæ:m</i>		<i>giàm</i>	‘emprisonner / to imprison’
<i>k.caj[?]</i>	<i>kacaj[?]</i>		<i>giấy</i>	‘papier / paper’

Velars **k-g**

<i>k-g</i>	<i>k</i>	<i>c / k</i>	<i>g / gh</i>	
<i>t.kɑ:c</i>	<i>takɑ:c</i>	<i>cát</i>		‘sable / sand’
<i>kɑ:n</i>	<i>kɑ:n</i>	<i>con</i>		‘fils, fille / son, daughter’
<i>ke:ʔ</i>	<i>ke:ʔ</i>	<i>cái</i>		‘féminin / féminin’
<i>r.kɑ:</i>	<i>lakæ:</i>		<i>gà</i>	‘poulet / chicken’
<i>c.kɑ:ŋ</i>	<i>cikæ:ŋ</i>		<i>gang</i>	‘empan / handspan’
<i>c.gi:ʔ</i>	<i>cikèjʔ</i>		<i>gây</i>	‘bâton, canne / walking stick’

Fricative alveolar **s**

<i>s</i>	<i>t^h</i>	<i>t</i>	<i>r</i>	
<i>sɑ:mʔ</i>	<i>t^hæ:mʔ</i>	<i>tám</i>		‘huit / eight’
<i>sa:j</i>	<i>t^ha:j</i>	<i>tai</i>		‘oreille / hear’
<i>si:</i>	<i>t^hi:</i>	<i>tay</i>		‘main / hand’
<i>p.səŋʔ</i>	<i>ut^hŋʔ</i>		<i>rǎn</i>	‘serpent / snake’
<i>p.se:ʔ</i>	<i>ut^he:ʔ</i>		<i>rái</i>	‘loutre / otter’
<i>k.saŋ</i>	<i>at^haŋ</i>		<i>rǎng</i>	‘dent / teeth’

5.2. Pre-glottalized **ʈ ɖ ʃ**

PVM	Arem	Viet	
<i>ʈ</i>	<i>b</i>	<i>m</i>	
<i>c.ɓu:c</i>	<i>cibuc</i>	<i>mút</i>	‘sucrer (sucre) / to suck (sugar)’
<i>ɓah</i>	<i>abah</i>	<i>mũa</i>	‘vomir / to vomit’
<i>ɓɑ:jʔ</i>	<i>ba:jʔ</i>	<i>muối</i>	‘sel / salt’
<i>ɓu:k</i>	<i>buk</i>	<i>múc</i>	‘puiser / to draw (water)’
<i>t.ɓaŋ</i>	<i>abaŋ</i>	<i>mãng</i>	‘pousses (bambou) / bamboo shoot’

<i>ɖ</i>	<i>d</i>	<i>n</i>	
<i>ɖɑ:k</i>	<i>dæ:k</i>	<i>nước</i>	‘eau / water’
<i>ɖam</i>	<i>dam</i>	<i>năm</i>	‘cinq / five’
<i>ɖi:p</i>	<i>di:p</i>	<i>nếp</i>	‘(riz) glutineux / sticky (rice)’
<i>k.ɖe:</i>	<i>kade:</i>	<i>nai</i>	‘cerf sambar / stag’
<i>ɖɑ:</i>	<i>da:</i>	<i>no</i>	‘rassasié / satiated’

<i>f</i>	<i>t</i>	<i>nh</i>	
<i>fuk</i>	<i>tuk</i>	<i>nhúc</i>	‘arbre à amadou / touchwood’
<i>faw</i>	<i>taw</i>	<i>nhau</i>	‘placenta / placenta, afterbirth’
<i>foh</i>	<i>to:ʔ</i>	<i>nhỏ</i>	‘arracher / to snatch’

5.3. Evolutions of others initials

PVM	Arem	Viet
ʔ	ʔ	# [ʔ]
h	h	h [h]
m	m	m [m]
n	n	n [n]
ɲ	ɲ	nh [ɲ]
ŋ	ŋ	ng(h) [ŋ]
v	v	v [v]
ʃ	ʃ	d [z/j]
r	r	r [r/z/j]
l	l	l [l]

The Arem initials **ʃ** and **ʃ** are only interpretations of Vietnamese phonemes.

5.4. Clusters with -r-

PVM	Arem	Viet	
<i>C.r-</i>	<i>Cr-/r-</i>	<i>s-</i>	
<i>g.rəjʔ</i>	<i>cirɛjʔ</i>	<i>sợi</i>	‘fil / thread’
<i>p.rək</i>	<i>para:k</i>	<i>sóc</i>	‘écureuil / squirrel’
<i>k.rəjʔ</i>	<i>kara:lʔ</i>	<i>sói</i>	‘loup / wolf’
<i>s.ranʔ</i>	<i>ranʔ</i>	<i>sắn</i>	‘manioc / cassava’
<i>(C.rəŋ)</i>	<i>kəra:ŋ</i>	<i>song</i>	‘rotin / rattan’
<i>c.ru:</i>	<i>cirù:</i>	<i>sâu</i>	‘profond / deep’
<i>p.ru:ʔ</i>	<i>parawʔ</i>	<i>sáu</i>	‘six / six’
<i>p.ri:</i>	<i>pari:</i>	<i>say</i>	‘ivre / drunk’
<i>c.ra:ŋ</i>	<i>ʳra:ŋ</i>	<i>sườn</i>	‘côtes / ribs’

5.5. Clusters with -l-

PVM	Arem	Viet	
<i>C.l-</i>	<i>l-/tl-/Cl-</i>	<i>tr-</i>	
...	<i>lù:mʔ</i>	<i>trộm</i>	‘voler, dérober / to steal’
<i>k.lən</i>	<i>ulan</i>	<i>trăn</i>	‘python / python’
<i>p.le:ʔ</i>	<i>ule:ʔ</i>	<i>trái</i>	‘fruit / fruit’
<i>k.la:</i>	<i>cilə</i>	<i>trưa</i>	‘midi / midday’
<i>p.ləŋ</i>	<i>ʳləŋ</i>	<i>tranh</i>	‘herbe à pailote / thatch-grass’
<i>b.ləj</i>	<i>tləj</i>	<i>trời</i>	‘ciel / sky’
<i>k.ləjʔ</i>	<i>tləjʔ</i>	<i>trói</i>	‘ligoter / to tie up’
<i>k.lam</i>	<i>tlam</i>	<i>trăm</i>	‘cent / hundred’
<i>k.lo:nʔ</i>	<i>tluənʔ</i>	<i>trốn</i>	‘s’ enfuir / to run away’

6. Finals consonants

6.1. Table of the evolution of final consonants

PVM	Arem	Viet
<i>p</i>	<i>p</i>	<i>p</i>
<i>t</i>	<i>t</i>	<i>t</i>
<i>c</i>	<i>c</i>	<i>t</i>
<i>k</i>	<i>k</i>	<i>c</i>
<i>ʔ</i>	<i>-ʔ</i>	<i>(sắc năng)</i>
<i>h</i>	<i>h</i>	<i>(hỏi ngã)</i>
<i>s</i>	<i>h</i>	<i>i / y (hỏi ngã)</i>
<i>m</i>	<i>m</i>	<i>m</i>
<i>n</i>	<i>n</i>	<i>n</i>
<i>ɲ</i>	<i>ɲ</i>	<i>n</i>
<i>ŋ</i>	<i>ŋ</i>	<i>ng(h)</i>
<i>r</i>	<i>l</i>	<i>i / y</i>
<i>l</i>	<i>l</i>	<i>i / y</i>

6.2. Final palatals PVM *c ɲ*

PVM	Arem	Viet	
<i>t.kɑ:c</i>	<i>taka:c</i>	<i>cát</i>	‘sable / sand’
<i>rɔ:c</i>	<i>rùɔc</i>	<i>ruột</i>	‘intestins / intestines’
<i>c.bu:c</i>	<i>cibũ:c</i>	<i>mút</i>	‘sucrer (sucre) / to suck (sugar)’
<i>p.səɲ</i>	<i>ut^hiɲ^ʔ</i>	<i>rắn</i>	‘serpent / snake’
<i>ta:ɲ</i>	<i>ta:ɲ</i>	<i>đan</i>	‘tisser / to plait’
<i>m.rəɲ^ʔ</i>	<i>ⁿriɲ^ʔ</i>	<i>rận</i>	‘pou / louse’

6.3. Final fricatives PVM *h s*

PVM	Arem	Viet	
<i>t.pah</i>	<i>^mpah</i>	<i>vả</i>	‘giffler / to slap’
<i>cəh</i>	<i>acəh</i>	<i>chẻ</i>	‘fendre (du bois) / to split (wood)’
<i>-bah</i>	<i>abah</i>	<i>mửa</i>	‘vomir / to vomit’
<i>la:s</i>	<i>liəh</i>	<i>lưỡi</i>	‘langue / tongue’
<i>mu:s</i>	<i>mùh</i>	<i>mũi</i>	‘nez / nose’
<i>C.kas</i>	<i>akæ:h</i>	<i>gãi</i>	‘gratter (de l’ongle) / to scrape (of nail)’

6.4. Final approximants PVM *r l*

The PVM *r-l* distinction is preserved only in some dialects in the Maleng group. Comparison of Arem and Vietnamese can not reconstruct this distinction.

PVM	Arem	Viet	
<i>pər</i>	<i>pal</i>	<i>bay</i>	‘voler / to fly’
<i>har</i>	<i>hæ:l</i>	<i>hai</i>	‘deux / two’
<i>t.kar^ʔ</i>	<i>ikal^ʔ</i>	<i>gáy</i>	‘chanter (coq) / to crow (rooster)’
<i>ʃa:l</i>	<i>ciəl</i>	<i>chài</i>	‘filet-épervier / cast net’
<i>k.ta:l^ʔ</i>	<i>atæ:l^ʔ</i>	<i>đái</i>	‘scrotum / scrotum’
<i>k.mo:l^ʔ</i>	<i>umùəl^ʔ</i>	<i>mối</i>	‘termite / white ant’

7. Brief conclusions

Despite its great interest for comparatism and reconstruction of Viet-Muong, the Arem has the disadvantage of being not a more homogeneous language. In addition to its own lexical funds, it contains layers of borrowings from standard Vietnamese and from a local dialect yet to be described. One can also detect borrowings from Katuic, Lao and Cham.

Notwithstanding these problems and challenges, it is important to present a substantial study to the community of specialists, but we must recognize that there remains important lexical work to do.

References

- Alves, Mark. 2006. Linguistic Research on the Origin of the Vietnamese Language: An Overview. *Journal of Vietnamese Studies* 1(1-2): 104-130.
- Barker, Muriel A. & Milton E. Barker. 1970. Proto-Vietnamuong (Annamuong) final consonants and vowels. *Lingua* 24: 268-85.
- Đặng Nghiêm Vạn, Chu Thái Sơn & Lưu Hùng. 1986. *Les ethnies minoritaires du Vietnam*. Editions en langues étrangères, Hanoi. [Arem are included in the ethnic group Chứt, pp. 38-42].
- Diffloth, Gérard. 1989. Proto-Austroasiatic Creaky Voice. *Mon-Khmer Studies* 15: 139-154.
- Enfield, N.J. & Gérard Diffloth. 2009. Phonology and sketch grammar of Kri, a Vietic language of Laos. *Cahiers de Linguistique – Asie Orientale* 38(1): 3-69.
- Ferlus, Michel. 1975. Vietnamien et proto viet-muong", *ASEMI* 6(4): 21-55.
- Ferlus, Michel. 1979. Lexique thavung-français, *Cahiers de Linguistique, Asie Orientale* 5: 71-94.
- Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. *Cahiers de Linguistique – Asie Orientale* 11(1): 83-106. [traduit en vietnamien: Sự biến hóa của các âm tắc giữa (obstruantes médiales) trong tiếng việt, *Ngôn ngữ* 1981(2): 1-22.].
- Ferlus, Michel. 1992. Histoire abrégée de l'évolution des consonnes initiales du Vietnamien et du Sino-Vietnamien. *MKS* XX: 111-125.
- Ferlus, Michel. 1996. Langues et peuples viet-muong", *Mon-Khmer Studies* 26 (Special Volume Dedicated to Dr. David D. Thomas on The Occasion of his Sixty-Fifth Birthday): 7-28.
- Ferlus, Michel. 1997. Problèmes de la formation du système vocalique du vietnamien. *Cahiers de linguistique – Asie Orientale* 26(1): 37-51.
- Ferlus, Michel. 1997. Le maleng brô et le vietnamien", *Mon-Khmer Studies* 27 (Hommages à Paul K. Benedict): 55-66.
- Ferlus, Michel. 1998. Les systèmes de tons dans les langues Viet-Muong. *Diachronica* XV: 11-27.
- Ferlus, Michel. 1999. Les disharmonies tonales en Viet-Muong et leurs implications historiques. *Cahiers de linguistique – Asie Orientale* 28(1): 83-100.
- 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.
- Gage, William. 1985. Vietnamese in Mon-Khmer perspective. In: *Southeast Asian Linguistic Studies presented to André-G. Haudricourt*. Edited by Surya Ratanakul, David D. Thomas & Suwilai Premsirat: 493-524. Bangkok: Mahidol University.
- Gregerson, Kenneth J. 1969. A Study of Middle Vietnamese Phonology. *Bulletin de la Société des Etudes Indochinoises* 44(2): 131-193.
- Haudricourt, André-Georges. 1953. La place du vietnamien dans les langues austroasiatiques. *Bulletin de la Société de Linguistique de Paris* 49(1): 122-8.
- Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242: 69-82.
- Kasuga, Atsushi. 1994. *A Study on Arem Phonology*. Master's thesis submitted to Tokyo University.

- Kasuga, Atsushi. 2008. *Arem Vocabulary*. Manuscript, organized in order Vietnamese-Arem-English.
- Maspero, Henri. 1912. Etudes sur la phonétique historique de la langue annamite: Les initiales. *Bulletin de l'Ecole Française d'Extrême-Orient* 12(1): 1-127.
- Nguyễn Tài Cẩn. 1995. *Giáo trình lịch sử ngữ âm tiếng Việt* [Course in historical phonetics of the Vietnamese language]. Hà Nội: Nhà xuất giáo.
- Nguyễn Văn Khang, Bùi Chi & 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.
- 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.
- Suwilai Premrirat. 2000. *So (Thavung) preliminary Dictionary*. Institute of Language and Culture for Rural Development, Mahidol University, Thailand, Department of Linguistics and Applied Linguistics, University of Melbourne, Australia.
- Taylor, Kate W. 2001. On being Muonged. *Asian Ethnicity* 2(1): 24-34.
- Thompson, Laurence C. 1973. Proto-Viet-Muong Phonology. *Austroasiatic Studies*, Oceanic Linguistics Special Publication n° 13, Edited by Philip N. Jenner, Laurence C. Thompson and Stanly Starosta, part II: 1113-1203.
- Trần Trí Dõi. 1995. *Thực trạng kinh tế và văn hóa của ba nhóm tộc người đang có nguy cơ bị biến mất* [Actual state of the economy and culture of three endangered ethnic groups]. Nhà xuất bản văn hóa dân tộc. [Arem: population pp. 71-76; no vocabulary]
- Trần Trí Dõi. 1999. *Nghiên cứu ngôn ngữ các dân tộc thiểu số Việt Nam* [Study of languages of ethnic minorities in Vietnam]. Nhà xuất bản đại học quốc gia Hà-Nội. [Arem: demography p. 110; vocabulary p. 16 and 37]
- Vương Hoàng Tuyên. 1963. *Các dân tộc nguồn gốc Nam-Á ở miền bắc Việt-Nam* [Ethnic groups of Austro-Asiatic origin in North Vietnam]. Nhà xuất bản Giáo Dục, Hà-Nội. [Arem: vocabulary p. 71; Compare list of a hundred words in Viet-Muong languages, including Arem; see fold-out page V-VIII, end of the book]