

Consonant shifts in Mon-Khmer initial stops

André-Georges Haudricourt

▶ **To cite this version:**

André-Georges Haudricourt. Consonant shifts in Mon-Khmer initial stops. 2017. halshs-01631482

HAL Id: halshs-01631482
https://shs.hal.science/halshs-01631482

Preprint submitted on 9 Nov 2017

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.

Non-final version (November 9th, 2017). In preparation for: Haudricourt, André-Georges. Evolution of languages and techniques*. (Ed.) Martine Mazaudon, Boyd Michailovsky & Alexis Michaud. (Trends in Linguistics. Studies and Monographs [TiLSM] 270). Berlin: De Gruyter Mouton.*

Consonant shifts in Mon-Khmer initial stops (1965)

Originally published as: "Les mutations consonantiques des occlusives initiales en môn-khmer". *Bulletin de la Société de Linguistique de Paris* 60 (1): 160–172 (1965). Reprinted in *Problèmes de phonologie diachronique:* 303-316.

translated by Paul Sidwell and Alexis Michaud

Abstract

At least three types of consonant shifts involving initial stops are attested in Mon-Khmer: (i) a "Germanic" type, encountered in Phay, Samre and Khasi, where p, t and k change to p^h, t^h and k^h while b, d and g change to p, t and k; (ii) a specifically "Mon-Khmer" type, where the confusion of p, t and k with b, d and g results in a two-way split of the vowel system (doubling the number of vowel qualities); and (iii) a "Far Eastern" type, where the same confusion results in a two-way split of the tone system.

Until recent years our knowledge of most Mon-Khmer languages was limited to poorly transcribed vocabularies that were barely sufficient as a basis for an initial classification, and definitely inadequate for comparative purposes. Fortunately, some recent publications, together with data newly collected by Michel Ferlus, a teacher in Luang Prabang, allow us to shed some light on the complex problem of correspondences between initial consonants. I will not deal with the Palaung-Wa languages, which have already been discussed by Shafer (1952).


100°E
110°E
N
30°N
20°N
10°N
Yangtze R.
INDIA
CHINA
Mekong R.
KHASI
Chindwin R.
Irrawaddy R.
PALAUNG
Xijiang (Pearl River)
BURMA (MYANMAR)
WA
Red River
RIANG
DANAW
LAMET
KHMU
TAY HAY
T'ENG
MUONG
VIETNAMESE
Salween R.
MRABRI
PHAY
LAOS
TAY POONG
RUC
AREM
Chao Praya R. (Menam R.)
Mekong R.
SÔ
BRU
PACOH
MON
THAILAND
CHAOBUN
(NYAHKUR)
KUY
BOLOVEN
BAHNAR
CHONG
SAMRE
CAMBODIA
KHMER
BIAT
PEAR
STIENG
KÖHO
VIETNAM
State boundary
LAMET Language name
TAY HAY, T'ENG Dialects within the Khmu subgroup
Type of consonant shift:
KHASI "Germanic"
MON "Mon-Khmer"
DANAW "Far Eastern"
RIANG Intermediate between the "Mon-Khmer" and "Far Eastern" types
0 Kilometres
1000

[1. "Germanic" type: devoicing of *voiced, aspiration of *voiceless]

Let us first examine the correspondences between the Phay[1] language (using vocabulary collected by Ferlus in Laos, west of Sainyabuli) and neighbouring languages: to the west, the language of the Mrabri, who are nomads of the western mountains of Nan province (Thailand),[2] and to the east, the language of the Khmu, more specifically the T'eng dialect recorded by Henri Maspero in Nghe-An (Vietnam) near the Laotian border, and the Tay Hay dialect recorded by Catholic missionaries in the province of Sam-Neua, Laos.[3]

[1.1. Phay]

First, we find correspondences between Phay aspirated initials and Khmu unaspirated initials [Table 1].[4]

1. At present, published data on the Phay language consists only of three vocabularies collected in Thailand in the province of Nan: Thin-saḥkaat, Thin-nam-yaa, and Thin-'baan-toey, published alongside the Mrabri vocabulary in the *Journal of the Siam Society* (see Nimmanahaeminda 1963).
2. An initial Mrabri vocabulary, under the name "Yumbri", was collected in 1936 by the photographer-reporter Hugo Bernatzik, working through a series of four interpreters. Understandably, it is not very useful (see Bernatzik 1941). The words cited here were collected by Kraisri Nimmanahaeminda (1963: 184) but there are still some obvious mistakes, *e.g.* when he asked for "dog" he was given "horse"; the two Thai words differ only in tone, and the Mrabri, whose language has no tones, speak Thai without tones.
3. In addition to the sources that I cited in my earlier article (1953: 123n1) [this volume: page xyz, footnote 3], it should be mentioned that Maspero's manuscript (T'eng) has now been published (Maspero 1955). As for W. A. Smalley's Khmu data, only a sketch grammar of Khmu has been published (Smalley 1961). I have also used a manuscript by R. P. Wauthier and John Subra (n.d.).
4. [Michel Ferlus (p.c. 2014) cites slightly different forms for some of these words in the Khmu dialect of the Luang Prabang area: 'breasts' (Table 2) is buʔ, with final glottal stop; 'woman' (Table 3) is cmkɨn (with a short vowel); 'forest, bush' (Table 4) is briʔ (with a short vowel and a final glottal stop); and 'egg' (also in Table 4) is kdoŋ (with a short vowel).]

[Table 1. Aspirated stop in Phay vs. unaspirated in Khmuic languages.]

gloss	Phay	Khmuic languages[5]			
		Mrabri	T'eng	Tay hay	Khmu
three	pʰɛʔ	pɛh			
four	pʰoːn	poːn			
fruit	pʰlɛʔ		pleʔ	pleʔ	
bush leech	pʰlɔːm			pluəm	
eight	tʰiʔ	ti			
six	tʰuol	tuon			
to do/make	tʰɛːŋ		tɛːŋ	tɛːŋ	
to sell	tʰəːc			teːt	
hand	ntʰiː	tiː	tiː	tiʔ	
bat (animal)	ntʰəːl		tɯər	ntɯəl	
tongue	ntʰaːk			ntaːk	
fish	kʰaː	kaː	kaː	kaʔ	kaʔ
arrow	kʰam		kam		
child	kʰuon	kɔːn	kɔːn		
thunder	kʰəːl	kɯl	kur		
tuber (edible)	kʰwaːj		kwaï		

There is also a correspondence between unaspirated voiceless stops and voiced stops [Table 2].

[Table 2. Voiceless unaspirated stop in Phay vs. voiced stop in Khmuic languages.]

gloss	Phay	Khmu dialects			
		Mrabri	T'eng	Tay hay	Khmu
bear	piek	bɛːk			
two	pie	bɛːr	baːr	kbaːr	
ash	puo		bɔʔ	bɔh	
alcohol	puc		buic	buːc	
breasts	poʔ	boʔ	buː	buː	
to cut	pok		bɔk		
to eat	pəŋ	baŋ			
coconut tree	plaːw		blaːw		

5. [In the original article, this column header reads "Khmu dialects", but the preceding text indicates that Mrabri and possibly Tay Hay are separate languages; this is consistent with current classifications, which subgroup both (and Phay) as "Khmuic", not "Khmu". See Sidwell (2009).]

leg	pluː			bluː	
gaur *(Bibos gaurus)*	tiːŋ				kdiːŋ
egg	tɔŋ		kdoːŋ	kdoːŋ	
to go down	tjuːl	djuːl	jur	djuːr	
mortar	kuːl		guol		
to winnow	kuːm		guːm		
stone	klieŋ		glaːŋ	glaːŋ	

This is clearly a shift of the Germanic type, whereby:

p > p^h	at the same time as	b > p
t > t^h		d > t
k > k^h		g > k

[1.2. Samre]

The same shift is found in the language which will be referred to here as Samre,[6] spoken between Tonle Sap (the great lake of Cambodia) and the Gulf of Siam by people known as the Čɔng or Chawng of Thailand, the Pear of the Cardamom mountains, and the Suoi of Kampong Speu. The shift did not take place in the languages spoken across the Mekong, namely Köho, Biat and Bahnar,[7] all of which are as conservative as Khmu in this respect. Examples are shown in Tables 3 and 4.

[Table 3. Correspondence between Samre aspirated stops and unaspirated stops in more conservative languages.]

gloss	Samre dialects	conservative languages
three	p^hai, p^hei, p^heːk	pɛː (Köho, Biat, Boloven)
four	p^hoːn, p^houn	puon (Köho, Biat, Boloven)
bush leech	p^hliːm, p^hleːm	pləːm (Köho)
fruit	p^hliː	plaj (Köho, Boloven)
sky	p^hliŋ	pleŋ (Bahnar)

6. The data are from Baradat's manuscript (1941). For other sources, see Haudricourt (1953: 124n3) [this volume: page xyz, footnote 7] and the entries "Pear", "Chong" and "Samreh" in Shorto, Jacob et al. 1963.
7. For bibliographical references, see Haudricourt (1953: 124n2) [this volume: page xyz, footnote 6] and the entries "Köho group", "Pnong group" and "Jəru group" in Shorto, Jacob et al. 1963.

squirrel	pʰruːk, pʰroːk	prɔːk (Bahnar), prɔːʔ (Köho)
tree trunk	tʰiːm, tʰeːm	təːm (Köho, Bahnar)
to plait	tʰaːɲ	taːɲ (Köho, Bahnar)
crab	tʰaːm	taːm (Köho), ktaːm (Bahnar)
earth	tʰai, tʰei, tʰeːk	tieh (Köho), tɛːh (Bahnar)
woman	cəmkʰɯːn	cəmkɯːn (T'eng, Tay hay)
arrow	kʰam	kam (Köho, T'eng)
to harvest	kʰac	kac (Köho)

[Table 4. Correspondence between Samre unaspirated initials and voiced initials in more conservative languages.]

gloss	Samre dialects	conservative languages
two	paːr	baːr (Köho, Biat), bəːr (Boloven)
ash	pɔːh	buːh (Köho)
cotton yarn	praːj	braːj (Köho, Boloven)
thigh	pluː	bluː (Biat, Tay hay)
forest, bush	priː	briː (Köho, Biat, Bahnar, T'eng, Tay hay)
water	taːk	daːk (Biat, Boloven), daːʔ (Köho)
duck	taː	adaː (Köho), daː (Boloven)
egg	tuŋ	kdoːŋ (T'eng, Tay hay)
house	tɔŋ	doŋ (Kuy)
to winnow	kuːm	guːm (T'eng)
lean	rkiː, rkiːk	rgaj (Köho), rgiː (Bahnar)

[1.3. Khasi]

A third language, Khasi, has the same correspondences [Table 5]. Khasi is currently spoken in India, in the mountains south of Assam; it probably originated in Burma.

[Table 5. Khasi aspirates vs. conservative Mon-Khmer unaspirated voiceless initials.]

gloss	Khasi	conservative languages
you	pʰi	pe (Boloven)
to weave	tʰaːiŋ	taːŋ (Köho, Bahnar)
land crab	tʰaːm	taːm (Köho), ktaːm (Bahnar)
fish	kʰaː	kaː (Köho, Bahnar)
child	kʰuːn	kɔːn (Köho, Bahnar), kuon (Boloven)
arrow	kʰnam	kam (Köho, T'eng)
tiger	kʰlaː	klaː (Bahnar), kluə (Boloven)

Unfortunately, there is no convincing example of the expected correspondence between Khasi unaspirated voiceless stops and voiced stops in the conservative languages. However, there are traces of a third series of stops in Khasi. These stops are currently realized as voiced, but the fact that this series lacks a velar, **g** (a sound which is only found in recent borrowings), suggests that it originates in a former preglottalized series; a good example is the word for 'paddy', **kba**, which is **ɓa** in Mon and Bahnar, with a preglottalized **ɓ**.[8]

[2. "Mon-Khmer" type consonant shift]

[2.1. Mon]

The second type of consonant shift (the Mon-Khmer type) occurs in Mon, and we have all the elements at hand to study it thanks to the excellent dictionary by Shorto (1962). Recall that this language is written in an alphabet of Indian origin which uses the Indian symbols for voiceless stops, (**p, t, c** and **k**) and for voiced stops (**b, d, j** [IPA: **ɟ**] and **g**). Mon preglottalized **ɗ** is written with the Indian letter for retroflex *d* [IPA: **ɖ**] (transliterated *ḍ*) and preglottalized **ɓ** by the Indian *v* with an added dot (transliterated *ḅ*).

8. [Throughout the article, Modern Mon forms from Shorto's *Dictionary of Modern Spoken Mon* (1962) have been added after Written Mon forms. In these modern forms, Shorto's notations for the preglottalized stops have been modified from ɓ and ɗ to the IPA symbols, ɓ and ɗ.]

The first issue to be addressed is whether the confusion of voiced and voiceless stops is completed in Mon or if it is only an ongoing change. In modern Mon there is an opposition between two syllable prosodies: "head register", pronounced in a clear voice, and "chest register", pronounced in a muffled, breathy voice. The latter is characterized by a general relaxation of the speech organs, including the larynx, resulting in "breathy" voice and centralization of the vowel.

Recall that in Mon, words are monosyllabic or quasi-monosyllabic: the main syllable may be preceded by a reduced syllable with the same "voice register". Thus we find, in "head" register:

'to obtain' **ket** (written *ket*)
'side' **həkoa**[9] (written *skay*)

and in "chest" register:

'to turn (tr.)' **kèt**[10] (written *get*)
'cat' **həkòa** (written *bagay*).

Shorto (1962: xi) tells us that in "chest-register" words **k**, **c**, **t**, **p** are partly voiced in initial position and fully voiced in medial position. In other words, **kèt** 'to turn' is pronounced **[kgèt]**, and **həkòa** 'cat' is pronounced **[həgòa]**. Table 6 provides a summary.

[Table 6. Mon data illustrating allophonic variation of voiceless stops depending on their initial or medial position. 'Monosyll.' = monosyllable.]

	register	gloss	Written Mon	Modern Mon	
				phonemic	phonetic
monosyll.	"head"	to obtain	ket	ket	ket
	"chest"	to turn tr.	get	kèt	^{k}gèt
quasi-monosyll.	"head"	side	skay	həkoa	həkoa
	"chest"	cat	bagay	həkòa	həgòa

Should we conclude that the voiceless/voiced opposition is still [phonemically] distinctive and that the "voice register" is merely its phonetic consequence? No, because for nasals we have examples of

9. [There is a typo in the original, where the word is written as kəkoa.]
10. [In the original transcription, the low (breathy) register was not indicated; it was clear from the heading. In the translation, this indication was added for all the Mon data in the article, using the standard transcription: a grave accent for low (breathy) register, and the absence of accent for high register.]

dissociation between the voicing of the initial and the "head" or "chest" register of the syllable. Thus, we have, in head register:

'widow' **mai** (written *kmay*) vs 'omens' **hnoa**[11] (written *snay*)

and in chest register:

'indigo' **mài** (written *may*) vs 'comb (of cock)' **hnòa** (written *dnay*).

[Since these words are identical in terms of initial consonants, the contrast between them has to be recognized as consisting in the voice register.] This influence of the "register", which, in synchrony, causes a difference in the tenseness or laxness of the vowel, is also observed in diachrony. Comparing quasi-homographs which, in writing, differ only in their initials, it can be observed that their actual pronunciation differs not only in register [i.e. phonation type and pitch] but also in vowel quality [Table 7].

[Table 7. Modern Mon pronunciation of identically written vowels in "head" and "chest" registers.]

"Head register" [i.e. modal voice]			"Chest register" [i.e. breathy voice]		
gloss	written as	pronounced	gloss	written as	pronounced
fish	*ka*	kaʔ	cotton blanket	*yāt ga*[12]	jàt kɛ̀ʔ
to split	*pāk*	paik	to send away	*bak*[13]	pɛ̀ak
four	*pan*	pɔn	to kiss	*ban*	pòn
earth	*ti*	tɔeʔ	to strike	*di*	tìʔ
to rub	*tuh*	taoh	fault	*duh*	tùh

[2.2. Sô]

Ferlus has recently discovered a similar phenomenon in the Sô language, spoken in the vicinity of the city of Thakhek in Laos: words borrowed from Thai have retained their original vowel quality when the initial was voiceless, but their vowel quality has changed when the initial was voiced [Table 8].

11. [The transcriptions *hn* (etc.) represent unvoiced nasals (IPA **n̥**, etc.)]
12. [Transliterated by Haudricourt as *yat ga*.]
13. [Transliterated by Haudricourt as *bāk*.]

[Table 8. Vowels in breathy-register Sô borrowings of Thai origin, and in the corresponding Thai etyma (with voiced initials in Old Thai).]

gloss	Lao pronunciation	Thai[14]	Sô pronunciation	[Thai script]
big knife	pʰaː	*brā 2*	prɛa	พร้า
harrow	kʰaːt	*grāt*	krɛat	คราด
slowly	saː	*jā 2*	cɛa	ช้า
time	veːlaː	*velā*	veːlɛa	เวลา
price	lakʰaː	*lahgā 1*	lakɛa	ราคา
one million	laːn	*lān*	lɛa	ล้าน
copper	thɔːŋ	*dåṅ*	tɔaŋ	ทองแดง
gong	kʰɔːŋ	*gåṅ*	kɔaŋ	ฆ้อง
gently	kʰɔːj	*gåy*	kɔaj	ค่อย
bar of wood	thɔːn	*dån*	tɔan	ท่อน
expensive	pʰɛːŋ	*beːṅ*	peːŋ	แพง
to be	mɛːn	*meːn 1* [15]	meːn	แม่น
waning (moon)	hɛːm	*reːm*	zeːm	แรม

Words of Mon-Khmer origin show the same evolution [Table 9].

[Table 9. Vowels in breathy-register Sô words of Mon-Khmer origin, and in cognates (with voiced initials) in conservative Mon-Khmer languages.]

gloss	Other Mon-Khmer	Sô pronunciation
duck	ada (Köho), daː (Boloven)	tɛa
violin	*draw* (Mon)[16]	trɔa
banana	*brāt* (Mon)	priat
salt	boh (Köho)	pɔh
belly	*buṅ* (Mon)	puŋ

On the other hand, former voiced preglottalized initials did not lose their voicing; they are reflected by plain voiced stops [Table 10].

14. [This column reproduces Haudricourt's transliteration of Thai.]
15. [The meaning in contemporary Thai is 'accurately, precisely, truly'.]
16. [Transliterated by Haudricourt as *dråˀ*.]

[Table 10. Sô voiced initials corresponding to preglottalized initials in conservative Mon-Khmer languages.]

gloss	Other Mon-Khmer[17]	Sô pronunciation
two	ḅā [ɓa] (Mon), ɓaːr (Bahnar)	baːr
bamboo shoot	tḅaṅ [ɓɛŋ] (Mon), tɓaŋ (Bahnar)	abaŋ
water	ḍāk [ɗaik] (Mon), ɗaːk (Bahnar)	dəːʔ
house	ḍuṅ [ɗɜŋ] (Mon: 'town, country')	duŋ

The Sô language is spoken as far east as the hills bordering the Gulf of Tonkin by the Van Kieu, the Khua, the Bru, the Pacoh, and the Katu. In these dialects, too, the consonant shift seems to have reached completion, with the confusion of two series of initials, voiced and voiceless, resulting in a split in the vowel system – a system that was already crowded, as it distinguished four degrees of aperture as well as length. Thus, according to Watson (1964: 136-139) there are at present 30 vowel phonemes in Pacoh, and the Millers and R. Phillips[18] found 41 vowel phonemes in Bru.

Concerning high vowels, Ferlus observed that for the Sô of Thakhek, there is only a difference in register, while for Bru in Vietnam, in the transcriptions of J. and C. Miller[19] we find diphthongs after former unvoiced initials [Table 11].

[Table 11. Sô and Bru data.]

Following formerly voiceless series:	Sô (Ferlus p.c.)	Bru (J. and C. Miller n.d.)
hand	atiː	atei
to take, to obtain	ʔiːt	ʔeit
dead	kuciːt	kuceit
five	sɯːŋ	səɯŋ
to see	huːm	həɯm
four	puːn	poun
knife	acuː	acou
name of ethnic group	bruː	brou

17. [In Haudricourt's notations, consonant preglottalization is indicated with an apostrophe, hence Mon *'ba*, *t'bang*, *'dāk*, *'duṅ* and Bahnar *'baar*, *t'bang*, *'daak*.]
18. [Possibly a pre-publication version of Phillips et al. 1976.]
19. [Haudricourt probably had access to J. Miller (n.d.) and J. and C. Miller (n.d. [c. 1963]). See references.]

Following formerly voiced series:		
pig	ali:k	ali:k
village	vi:l	vi:l
brother-in-law	anji:	anji: (uncle)
pair	ku:	ku: (Lao: $k^{h}u:$ < proto-Tai *gu:)

[2.3. Kuy]

Smalley (1964) provides some details on the Kuy dialects, spoken across the Mekong in Thailand and Cambodia, based on the unpublished notes of Richard Johnston. A total of 24 vowel phonemes are distinguished by quality and length (11 short and 13 long), but each can also be "breathy" or "non-breathy", which could yield a total of as many as 48 vowel phonemes. But in the present state of the language, the voiced stops and the voiceless aspirates do not occur before breathy vowels. The voiced stops are former preglottalized stops that behave like voiceless stops [in that they yield a "non-breathy" register], and the aspirates are loans from Tai, and proto-Tai had only one series of aspirated consonants, which were voiceless. In his discussion of the rendering of Kuy by means of the Thai alphabet, Smalley explains that breathy voice quality tends to cause a preceding stop to be perceived as aspirated (there is no phonological opposition with a true aspirate in that context), and that speakers of the language, who are bilingual, transcribe stop consonants followed by breathy vowels[20] using the symbols for aspirated consonants, adding the diacritic for low tone.

[2.4. Chaobun (Nyahkur)]

This explains the state of affairs found in the Chaobun [Nyahkur] dialects of Thailand, of which we only have transcriptions in the Thai alphabet:[21] [Table 12a].

20. [The original reads "followed by breathy consonants" instead of "followed by breathy vowels".]
21. See Shorto et al. (1963: 14), "Niakuol" entry, and Chintana (1958: 34–45).

[Table 12a. Aspirated initials in Chaobun vs. voiced initials in Written Mon.

gloss	Chaobun dialects[22]	Mon[23]
belly	phung [pʰùŋ]	buṅ [pɜ̀ŋ]
wife	phau [pʰràw]	brau [prɛ̀a]
rain	phəːy, phləːy [pʰɛ̀j]~[pʰə̀j]	bray [pròa]
to bind	thak [tʰàk]	dak [tɛ̀ak]
duck	thaː, thia [tʰàː]	ada [ətɛ̀a]
water turtle	thi [tʰwìːˀ]	gwi (dwi) [kwìˀ]
rope	choːk [sòːk]~[cʰòːk]	juk [cɜ̀k]
liver	khi [kʰrìh]	gruih [krɜ̀h]
forest	kheːp, khrəːp [kʰrə̀p]	gruip [krɜ̀p]

Thus former voiced initials are represented by aspirates, while former voiceless initials are unchanged [Table 12b].

[Table 12b. Voiceless initials in Chaobun vs. voiceless initials in Written Mon.]

gloss	Chaobun dialects	Mon
ground, earth	tiʔ [tiːˀ]	ti [tɔeˀ]
hand	təy [tɛj]~[təj]	tay [toa]
tongue	taːk [ntaːk]	latāk [kətaik]
cooked rice	pong [poːŋ]	puṅ [pɜŋ]
three	pi [piːˀ]	pi [pɔeˀ]
four	pan [pan]	pan [pɔn]
buffalo	pieng, plieng [priɛŋ]	preṅ [preaŋ]
fish	kaʔ [kaːˀ]	ka [kaˀ]
child	kuon [kuan]	kon [kon]

The former glottalized series retained its voiced character [and thus remained distinct from the other series of consonants: see Table 12c].

22. [Nyahkur data from Ferlus's unpublished field notes are provided in square brackets in the *Chaobun* column; the grave accent indicates breathy voice.]
23. [Modern Mon, from Shorto 1962, has been added after Written Mon forms.]

[Table 12c. Voiced initials in Chaobun vs. preglottalized initials in Written Mon.]

gloss	Chaobun dialects	Mon
two	baː	ḅā [ɓa]
water	daːk	ḍāk [ɗac]
village	dong	ḍuṅ [ɗɜŋ]

[2.5. The Khmer shift]

The type of consonant shift which we have observed in Mon and Sô is reminiscent of that found in Khmer, presented implicitly by Martini (1946: 112–131) when he pointed out the relationship between the traditional orthography and the pronunciation.

The essential difference between the type of onset consonant shift found in Khmer and that found in Mon and Sô is that in Khmer, as also in the languages to the west (Chong and Samre) and east (Stieng and Köho), there is no trace of the existence of a former preglottalized series: the initial consonants in the two diagnostic examples 'two' and 'water' evolved like ordinary voiced stops [i.e. they became voiceless]. The consequence of this gap was that nothing prevented the former voiceless stops from becoming more tense, more glottalized, until they finally became voiced [and preglottalized] when followed by a vowel. Examples are shown in Table 13a.

[Table 13a. Voiceless stops in Written Khmer and conservative languages vs. preglottalized stops in present-day Khmer.]

gloss	Khmer orthography	pronunciation	other Mon-Khmer	
			Köho	Stieng
three	*pi*	ɓei	pɛ	pai
four	*puon*	ɓuon	pɔn	puon
earth	*ti*	ɗei	tiah	teh
hand	*tai*	ɗai	ti	ti

The consonants of the former voiced series are now voiceless, as shown in Table 13b.

[Table 13b. Voiced stops in Written Khmer and conservative languages and voiceless stops in present-day Khmer.]

gloss	Khmer orthography	pronunciation	other Mon-Khmer
two	*bīr*	piː	baːr (Köho)
forest	*brai*	prei	bri (Köho)
thigh	*bhlau*	p^{h}ləw	bluː (Biat)
water	*dḭḵ*[24]	tɨk	daːʔ (Köho)
duck	*da*	tea	da (Stieng), ada (Köho)
to think	*git*	kit ~ kɨt	git (Köho ‘to know’)

According to Martini (1946) there are 32 vowel phonemes in present-day Khmer, placing this language between Pacoh (30 vowels) and Bru (41 vowels) in terms of richness of the vocalic inventory. According to Henderson (1952), the language has two vowel registers, as in Mon.

[3. “Far-Eastern” type: initial merger and tonal split]

There is obviously a close relationship between the Mon-Khmer type of consonant shift, where confusion of initials leads to two voice registers that affect the quality of vowels, and the shifts observed in tone languages of the Far East (see the discussion of the two-way split of tonal systems in Haudricourt 1961 [translated in this volume]). One wonders if these two voice registers are interpreted as pitch differences only in languages that already have tones, in which case the observations of G. H. Luce on Riang and of K. G. Izikowitz on Lamet should rather be interpreted in another way. Luce and Izikowitz were conscientious observers and were familiar with tone languages, but were not professional phoneticians. It is not unlikely that Riang and Lamet are actually voice register languages rather than tone languages.[25] Ferlus tells me there is a Lamet dialect with “diphthongization” and hence probably with voice register.

24. [In the original article, the transliterations of Written Khmer appear as bi for ‘two’ and dik for ‘water’; the pronunciation indicated for ‘water’ is tik.]
25. [This conjecture was confirmed since. On Riang, see Shorto and Luce (2013) and Shintani (2014). The synchronic facts for Lamet are more complex (Lindell et al. 1978) because some dialects have become predominantly tonal (Ferlus, p.c. 2016).]

On the other hand, Danaw, spoken in Burma east of Inle Lake, which was not available to Shafer (1952), but was recently rediscovered by Luce (1965), is indeed a tone language. Comparing Danaw with Riang, Luce observed that the low tone of Riang corresponds to the high falling tone of Danaw, but that the high tone of Riang corresponds to two distinct Danaw tones: low falling and high level, the latter for words which in Riang retain a final glottal closure that has disappeared in Danaw. Final glottalization is precisely what gave rise to the *sắc-nặng* tone in Vietnamese [Haudricourt 1954]. Examples of the correspondence between Danaw High level and Vietnamese *sắc-nặng* tones are shown in Table 14.

[Table 14. Correspondence of Danaw high level tone with the tones *sắc* and *nặng* (B1 and B2) of Vietnamese.]

gloss	Danaw high level tone	Vietnamese[26]	Vietnamese orthography
			sắc tone (B1)
dog	tso	[cɔ ^{B1}]	*chó*
louse	tsi	[cɤ̆j ^{B1}]	*chấy*
leaf	la	[la ^{B1}]	*lá*
fruit	ple	**[**ʈaj ^{B1}**]**	*trái*
to open (mouth)	ha	**[**ha ^{B1}**]**	*há*
			nặng tone (B2)
husked rice	ko	**[**ɣaw ^{B2}**]**	*gạo*
thread	lu	**[**luə ^{B2}**]**	*lụa*
pot	lo	**[**lɔ ^{B2}**]**	*lọ*

Finally, Danaw has also preserved the former preglottalized stops as a [distinct] series, now realized as voiced stops, *e.g.* 'paddy' *ba*, 'bamboo shoot' *tubong*.

In Vietnamese and Muong, the merger of initials occurred after the development of tones.[27] The voicing contrast is perhaps maintained in the

26. [Haudricourt's phonetic rendering of the Vietnamese examples is: ćɔ, ćei, la, ʈaːy, ha, ɣaːw, luo, lɔ.]
27. See Haudricourt (1961), pp. 172–173, and p. 165, note 3. To the sources cited in Haudricourt (1953: 123, note 2) must now be added the word lists for Poong, Arem, Ruc and May published by Vương Hoàng-Tuyên (1963: 196).

mountainous dialects, such as Tay Poong,[28] Khong-Kheng, Hung etc. [Table 15].

[Table 15. Possible preservation of an old opposition of voice in Tay Poong as compared with Muong, 17th-century Vietnamese (VN), and modern VN dialects.]

gloss	Tay Poong	Muong	17th-c. VN	southern VN	northern VN
sky	bləj	tləj	*blời*	*trời*	*giời*
moon	plɔŋ	tlaŋ	*blăng*	*trăng*	*giăng*

In Muong, the stops resulting from the merger of voiced and voiceless stops have remained voiceless, and preglottalized stops have remained unchanged.

In Vietnamese, preglottalized stops have become nasals, with the same high series tones as the former voiceless nasals. Thus, Muong **ɗam** 'five' corresponds to Vietnamese *năm*, which is homophonous with *năm* 'year' [illustrating the merger of the former preglottalized stops with the former voiceless nasals]. (That the word for 'year' used to have a voiceless nasal is shown by the Mon *cnam* > **hnam**.)

On the other hand, the single Muong series **k, c, t, p** corresponds to two series in Vietnamese: (i) a series of stops, some of which are preglottalized: **k-, c-, ɗ-,** and **ɓ-**, written as *c, ch, đ* and *b*; (ii) a series of voiced spirants: **ɣ, ʒ**, ð,[29] and **ʋ**, written as *g, gi, d* and *v*. This second series could be explained

28. [In spite of its name, Tay Poong is not related to Tay (Tai-Kadai family): it is a language of the Vietic group of the Mon-Khmer family. It is also known as Poong.]

29. [Haudricourt's notation is *ɣ, ʒ, ź* and *ʋ*, i.e. suggesting ʑ or z̬ as the pronunciation of Vietnamese *d*. Comparative evidence, and Vietnamese borrowings into neighbouring languages, support the view that at the time when the Vietnamese orthography was designed, the phoneme written as *d* was a dental spirant, written as δ by Ferlus 1982 (see also Ferlus 2001). The closest equivalent to this sound in the International Phonetic Alphabet is ð. On *spirants*, see the Glossary. On the Vietnamese alphabet, see Haudricourt (1949).]

by the existence of an **r-** prefix found in certain dialects, such as Ruc and Arem [Table 16].[30]

[Table 16. Correspondences between Vietnamese, Muong and other Mon-Khmer languages, suggesting the presence of an r-prefix in these words. "Viet."=Vietnamese, "Viet. spell."=Vietnamese spelling.]

	Viet.	Viet. spell.	Muong	[Conservative Mon-Khmer languages]
chicken	ɣa	*gà*	ka	rəka (Ruc), ləkɛ (Arem)
husked rice	ɣaːw	*gạo*	kaːw	rəko (Ruc, May) rəkaw (Palaung)
cotton	ʋaːj	*vải*	paːj	kupal (Ruc), kəpai (May), *karpās* (Khmer)

[4. Conclusion]

From a general point of view, we can classify the consonant shifts observed above as follows.

The correlation that contrasts two series of stops (*fortis* voiceless vs. *lenis* voiced) can evolve in either of two ways.

The first type of evolution originates in a reinforcement of the *fortis* series: the contraction of the laryngeal-oral muscles increases the air pressure inside the mouth during the interval when the stop is maintained; the air escapes with plosion, turning the voiceless stops into aspirated stops. The voiced stops may then lose their voicing without danger of lexical confusions. This is the "Germanic" type of consonant shift.

The second originates in a weakening of the *lenis* series: the relative laxness of the laryngeal-oral muscles is prolonged from the articulation of the initial consonant into that of the following vowel. The relaxation of the larynx lets breathy voice come through and lowers the pitch of the voice, while the relaxation of the muscles of the mouth results in a "lax" vowel quality.

When this process takes place in languages without tones, it is the difference in vowel quality that eventually becomes distinctive. A correlation of consonants disappears as a correlation of vowels appears, decreasing the number of consonants, while increasing the number of

30. [This hypothesis was confirmed by later comparative research about Vietnamese: see the study of the spirantization of medial consonants by Ferlus (1982).]

vowels. On the other hand, when this happens in a tone language, it is the change of register that becomes distinctive: the correlation between consonants disappears and a tonal correlation appears, decreasing the number of consonants while causing a split of the tone system.

These two subtypes can be superimposed, as in the tonal southern Tai[31] languages (Thai, Lao, and Saek). These languages were already tonal when the onset consonant shift took place, but the shift was carried out on a Mon-Khmer substratum and adstratum: the breathy character of the vowel that followed the former voiced stops became aspiration, such that **b, d, ɟ** and **g** became **pʰ, tʰ, cʰ** and **kʰ**.[32] On the surface, this change seems to be exactly the inverse of the "Germanic" shift with which we started. I believe I have shown that it is fundamentally the same phenomenon, and that the differences between the phonological systems of the languages at issue and their historical contact situations suffice to explain the different outcomes.

Comments

This article builds on "Two-way and three-way splitting of tonal systems in some Far Eastern languages" (this volume), which had been published four years earlier. In Vietnamese, the voicing opposition on initial consonants transphonologized to new tonal oppositions, resulting in an increase in the

31. [Haudricourt's "southern Tai" (original French: "langues Thai méridionales") corresponds to Li Fang-kuei's (1977) Southwestern Tai. Haudricourt uses the term "Thai" to refer to a set of languages that by and large corresponds to Li's Southwestern Tai and Central Tai, exclusive of Northern Tai (which Haudricourt calls "Dioi", or "Zhuang", after the Chinese ethnonym *Zhuàng* 壯). The more modern term "Tai" is used in the present translation; "Thai" refers to the national language of Thailand, also known as "Siamese". For more details, see "On the reconstruction of initial consonants in monosyllabic languages: the case of proto-Tai" and "What geography, ecology, and semantics can tell us about the origin of the Tai peoples", in this volume.]
32. A necessary requirement for this change is the existence, prior to the change, of a series of voiceless aspirates. This condition was met in Tai as well as in Chinese. In Mon-Khmer languages, on the other hand, the sounds transcribed with the series of letters used for voiceless aspirated stops in Indic languages are in fact clusters formed of a voiceless stop followed by h [so the condition for the reinterpretation of breathiness as aspiration was not met in Mon-Khmer].

number of lexical tones. The present article brings out other evolutionary paths followed by this opposition in different languages.

Later work has confirmed the links between the type of change which Haudricourt calls "Far Eastern" (split of the tone system) and that which he calls "specifically Mon-Khmer" (split of the vowel system) (Ferlus 1979). These two types share a common initial stage from the phonetic point of view: the relaxation of the larynx. At that stage, the phonetic cues to this opposition include, in addition to phonation type proper, some differences in pitch, as well as differences in vowel articulation, as was already noted for a conservative variety of Khmer by Henderson 1952. At a later stage, one or the other of the cues becomes dominant: this is where the evolution branches into the "Mon-Khmer" type (where vowel quality stabilizes as the new distinctive property) and the "Far Eastern type" (where the distinctions become tonal).

An extension of the general theory of the changes associated with "registrogenesis" has been proposed in a study of vowel changes from Old Chinese to Middle Chinese (Ferlus (2009: 192).

References

Baradat, Raymond. 1941. Les dialectes des tribus sâmrê [The dialects of Sâmrê tribes]. Unpublished manuscript of the Ecole Française d'Extrême-Orient, Paris. 267 pp. Available online from http://sealang.net/monkhmer/database (Accessed in 2015).

Bernatzik, Hugo. 1941. *Die Geister der Gelben Blätter* [The spirits of the yellow leaves]. Leipzig: Koehler & Voigtländer.

Chintana, But. 1958. Lawa Chao Bon [The Lawa Chaobon language]. *Silpakorn* (Bangkok) 2(1/2). 52–62/34–45.

Haudricourt, André-Georges. 1953. La place du vietnamien dans les langues austroasiatiques [The place of Vietnamese in Austroasiatic]. *Bulletin de la Société de Linguistique de Paris* 49(1). 122–128.

Haudricourt, André-Georges. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient [Two-way and three-way splitting of tonal systems in some Far Eastern languages]. *Bulletin de la Société de Linguistique de Paris* 56(1). 163–180.

Henderson, Eugénie J.A. 1952. The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies* 14(1). 149–174.

Izikowitz, Karl Gustav. 1951. *Lamet. Hill Peasants in French Indochina*. Göteborg, Etnografiska Museet (Etnologiska Studier 17).

Luce, Gordon. 1965. Danaw, a dying Austroasiatic language. *Lingua* 14. 98–129.

Martini, François. 1946. Aperçu phonologique du cambodgien [A phonological outline of Cambodian (Khmer)]. *Bulletin de la Société de Linguistique de Paris* 42. 112–131.

Maspero, Henri. 1955. Matériaux pour l'étude de la langue T'èng [Materials for the study of the T'eng language]. *Bulletin de l'École Française d'Extrême-Orient* 47. 457–507.

Miller, John. n.d. A tentative statement of Bru phonology. s.l., s.n. 7 p. [catalogued in the David Thomas Library, Bangkok].

Miller, John & Carolyn Miller. n.d. [c. 1963]. *Bru Dictionary* s.l., s.n. 190 p. [catalogued in the David Thomas Library, Bangkok. Later publications by these authors on Bru are listed in the library catalog; several are available on line. Website: www.msea-ling.info/dt-library.]

Nimmanahaeminda, Kraisri. 1963. The Mrabri language. *Journal of the Siam Society* 51(2). 179–184.

Phillips, Richard L., John Miller & Carolyn Miller. 1976. The Brũ vowel system: alternate analyses. *Mon-Khmer Studies* 5. 203–217. [An editor's note (p. 216n1) mentions that the paper was written "well over a decade" before publication.]

Shafer, Robert. 1952. Etudes sur l'austroasien [Austroasian studies]. *Bulletin de la Société de Linguistique de Paris* 48(1). 111–158.

Shintani, Tadahiko. 2014. *The Riang language*. (Linguistic Survey of Tay Cultural Area 101). Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Shorto, Harry Leonard. 1962. *A dictionary of modern spoken Mon.* London: Oxford University Press.

Shorto, Harry Leonard, Judith M. Jacob & E.H.S. Simmonds. 1963. *Bibliographies of Mon-Khmer and Tai linguistics.* Oxford & London: Oxford University Press.

Smalley, William A. 1961. *Outline of Khmuʔ structure.* New Haven: American Oriental Society.

Smalley, William A. (ed.). 1964. *Orthography studies: Articles on new writing systems*. London: United Bible Society.

Vương Hoàng-Tuyên. 1963. *Các dân tộc nguồn gốc Nam Á ở miền bắc Việt-Nam* [Ethnic groups of South Asian origin in Northern Vietnam]. Hanoi: Nhà xuất bản Giáo dục.

Watson, Richard. 1964. Pacoh phonemes. *Mon-Khmer Studies* 1. 135–148.

Wauthier, R.P. [Jean] & Jean Subra. n.d.. Syllabaire thay-hay. *ms. (not located).*

Additional references

Ferlus, Michel. 1979. Formation des registres et mutations consonantiques dans les langues mon-khmer [Development of phonation-type registers and consonant shifts in Mon-Khmer languages]. *Mon-Khmer Studies* 8. 1–76.

Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien [The spirantization of medial obstruents in Vietnamese, and the development of the Vietnamese consonant system]. *Cahiers de Linguistique - Asie Orientale* 11(1). 83–106.

Ferlus, Michel. 2001. *Les hypercorrections dans le thổ de Làng Lỡ (Nghệ An, Vietnam) ou les pièges du comparatisme* [Hypercorrections in Làng Lỡ Thổ (Nghệ An, Vietnam): the pitfalls of historical-comparative phonology]. Quinzièmes Journées de Linguistique de l'Asie Orientale, École des Hautes Études en Sciences Sociales, Paris.

Ferlus, Michel. 2009. What were the four divisions of Middle Chinese? *Diachronica* 26 (2). 184–213.

Ferlus, Michel. n.d.. Nyahkur (Chaobon). Unpublished field notes.

Haudricourt, André-Georges. 1949. L'origine des particularités de l'alphabet vietnamien. *Dân Việt-Nam* 3. 61–68. Translated as: The origin of the

peculiarities of the Vietnamese alphabet. *Mon-Khmer Studies* 39. 89–104 (2010).

Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien [The origin of tones in Vietnamese]. *Journal Asiatique* 242. 69–82.

Li, Fang-Kuei. 1977. *A Handbook of Comparative Tai*. Oceanic Linguistics Special Publications no. 15. Honolulu: University of Hawai'i Press.

Lindell, Kristina, Jan-Olof Svantesson & Damrong Tayanin. 1978. Two dialects of the Rəmeet (Lamet) language. *Cahiers de linguistique – Asie orientale* 4(1). 5–22.

Martinet, André. 1956. *La Description phonologique avec application au parler franco-provençal d'Hauteville (Savoie)* [Phonological description, with an application to the Franco-Provençal dialect of Hauteville, Savoie]. Genève: Droz.

Martinet, André. 1981. Fricatives and spirants. In Bhakti Prasad Mallik (ed.), *Suniti Kumar Chatterji Commemoration Volume,* 145–151. Burdwan, West Bengal, India: Burdwan University Press.

Martinet, André. 1985. Two proposals. *The Study of Sounds (Onsei no Kenkyuu), Commemorative volume for the 50th anniversary of the Phonetic Society of Japan* 21. 67–72.

Shorto, Harry L. & Gordon Hannington Luce. 2013. *Riang-Lang vocabulary: Compiled from the materials collected by G.H. Luce*. Canberra: Asia-Pacific Linguistics.

Sidwell, Paul. 2009. *Classifying the Austroasiatic languages: History and state of the art.* Munich: Lincom.