REFLECTIONS ON
THE HISTORY OF THE KHMER PHONEMIC SYSTEM[*]

Heinz-Jürgen PINNOW
Freie Universität, Berlin

The phonetic and phonemic systems of Khmer[1] are many-sided
and complex to an uncommon degree; this is especially true of
the vocalism. The situation is all the more complicated by the
variety of transcriptions employed by different writers, who
disagree in many cases in their interpretation of the articula-
tion of vowels and consonants and of their place in the phone-
mic system. Various dialectal differences and orthographic var-
iations also occur. This being the case, it ought to be useful
to examine critically the systems favored by the principal
authorities and weigh them one against the other. The following
are considered in the presentation hereafter:

1. L. Finot, "Notre transcription du cambodgien," in *BEFEO*, II
    (1902).1: 1-15.

2. P. W. Schmidt, "Grundzüge einer Lautlehre der Mon-Khmer-
    Sprachen," in *Denkschr. d. kais. Akad. d. Wiss. Wien,
    Philos. hist. Kl.*, 51 (1905), Abhandlung III.

3. G. Maspero, *Grammaire de la langue khmère (cambodgien)*
    (Paris: Imprimerie Nationale, 1915).

4. J. Guesdon, *Dictionnaire cambodgien-français*, 2 volumes
    (Paris: Plon, 1930).

5. F. Martini, "Aperçu phonologique du cambodgien," in *BSLP*,
    XLII (1942-1945).1: 112-31.

6. G. Cambefort, *Introduction au cambodgien* (Paris: G.P.

---

[*]Translation of "Sprachgeschichtliche Erwägungen zum Phonem-
system des Khmer," in *Zeitschrift für Phonetik und allgemeine
Sprachwissenschaft*, 10 (1957).4: 378-91.

[1]More accurately, Khmɛr, pronounced [khmaer] or [khmae] with
a distinct [ae] diphthong, as in German *Meyer*.

Maisonneuve, 1950).

7. Eugénie J.A. Henderson, "The Main Features of Cambodian Pronunciation," in *BSOAS*, XIV (1952).1: 149-74.

A careful distinction must be made between *transliterations* on the one hand and, on the other, phonetic and phonemic *transcriptions*. Schmidt gives only a transliteration. Martini uses all three. Henderson uses the Cambodian character together with a phonological transcription, occasionally supplemented by a phonetic transcription. Maspero, Guesdon, Cambefort and Finot all give the Cambodian character along with a romanization which is supposed to be at once a transliteration and a phonetic and phonemic transcription, but which is at variance with the facts. The whole dispute between Schmidt and Finot over romanization is meaningless, since Schmidt uses only transliteration. When Finot asserts that "...la translittération *bhlöñ, *bhnaṃ ne répond à aucune réalité ni moderne, ni ancienne,"[2] he is indeed right as far as pronunciation goes but not as far as written forms are concerned, and it was only the latter that were represented in Schmidt's system.

The Vocalism

In the writing system the following vowels are distinguished: ɔ, ɔ̆, a, ă, ĭ, i, ŭ̄, ɯ, ŭ, u, (r̥̄, r̥, l̥̄, l̥), uo, ɤ, ɯɤ, ie, e, ɛ, ai, o, au, ŭ̄, ɔ̄, ã, (ãŋ), ɔḥ (ɔ̆ḥ, &c.).

In the older though never universally accepted transliteration these were represented a, a´, ā, ā´, i, ī, u, ư, u, ū, (r̥, r̥̄, l̥, l̄), uo, o´, uʼo´, ie, e, ê, ai, o, au, uṃ, aṃ, āṃ, (āṃñ), aḥ, (a´ḥ, &c.). Various other systems have also been used.

The articulation of the vowels is now conditioned by the

---

[2] Finot, op.cit., 4.

nature of the consonants preceding them. These fall into two groups known as *series* or *registers*. To the First Register belong the original voiceless consonants k, kh, c, ch, ṭ, ṭh, t, th, p, ph, s and h, together with n̥ and l̥ and initial vowels. To the Second Register belong the originally voiced consonants g, gh, ɟ, ɟh, ḍ, ḍh, d, dh, b and bh along with the nasals ŋ, ɲ, n and m as well as j [j], r, l and v. Of this Second Register the voiced stops, g and so on, have since lost their voicedness and hence become k and so on, while ṭ and p are now realized as [ɗ] and [ɓ] respectively; see the discussion of the consonants below. To compensate for this shift, words of the Second Register are pronounced in a deeper voice.

> The characteristics of the second register are a deep rather breathy or "sepulchral" voice, pronounced with lowering of the larynx, and frequently accompanied by a certain dilation of the nostrils. Pitch is usually lower than that of the first register in similar contexts.[3]

These relationships as represented by Schmidt, Martini and Henderson are shown in Table 1. The systems used by the others are generally similar but exhibit differences of detail; these are shown in Table 2.

The letters r̃, r̥, l̃ and l̥ are equivalent to combinations of r and l plus w̃ and ɯ. Maspero writes them as r̥ũ, r̥u, l̥ũ, l̥u, Guesdon as rŭ̃, rŭ̃, lŭ̃, lŭ̃.

With Henderson y = [j], and it is to be noted that : appears in her phonological transcription when a long vowel is followed by a final consonant. Her F stands for "final", that is, a following final consonant.

The realization of my /ĭ/, /ŭ/, /ɛ̆/ and /ɔ̆/ is diphthongal: [iə], [uə], [ɛə], [ɔə]. /ŏ/ is realized variously as [o],

---

[3]Henderson, op.cit., 151.

				Transcriptions of									
			Schmidt	Martini						Henderson			
Tl.	*phonol.* R¹	R²	*Tl.*	*Tl.*	*Tl.*	*phonet.* R¹	R²	*phonol.* R¹	R²	*phonet.* R¹	R²	*phonol.* R¹	R²
ɔ	ɑ	ɔ	a	a		ɑː	ɔː	ɑ	ɔ	ɔː	ɔ̀ː	ɔ	ɔ̀
ɔ̆	ɑ̆	ŭ, ɔ̞, ɛ̞	a	a′		ɑ, o, ᵊɑ	ʷo	ɑ	uᵒ̆, ɔ̞̊, å̆	ɔ	ùə	ɔF	ẁəF
a	a	iɣ	ā	ā		aː	iːe	a	ie	aː	l̀ə	a	l̀ə
ă	ă	ɔ̆, ɛ̆	ā	ā′		a, ᵊɑ, ɛa	ᵊɑ	ă	ǎ, æ	a	ʷɔ̀ə, ɛ̀ə	aF	ɔ̀əF, ɛ̀əF
ĭ	y̆	ŭ̈, ĭ	i	i		ᵊe, ᵊɛ	ɯ, i	ĕ, ɛ̆	i̞, ĭ̞	ɣ	ù̈	ɣF	ù̈F
i	ɣj	i	ī	ī		ɛi̞	iː	ɛy	i	ɣy	l̀ː	ɣy	l̀
ŭ̈	(y̆)	ŭ̈	ü	i̞		–	ɯ	–	ĭ̞	(ɣ?)	ù̈	(ɣ?)	ù̈F
ɯ	(ɣ)	ɯ	ǖ	ɪ̞		–	ɯː	–	i̞	(ɣ)	ù̈ː	(ɣ)	ù̈
ŭ	ŏ	ŭ	u	u		ɑᵘ	u	ɔ̆	ŭ	ou, oɣ, ʷo	ù	oF	ùF
u	o	u	ū	ū		oː	uː	o	u	oː	ù̀ː	o	ù̀
uo	uɣ	uɣ	uo	w		uːo	uːo	uo	uo	uə	ùə	uə	ùə
ɣ	aɣ	ɣ	ö	œ		aᵊ	əː	ʌ	ə	aɣ	ɣ̀	aɣ	ɣ̀
ɯɣ	ɯɣ	ɯɣ	üö	yœ		ɯːə	ɯːə	i̞ə	i̞ə	ɯə	ù̈ə	ɯə	ù̈ə
ie	ĭ	ĭ	ie	ye		ye	ye	ĭe	ĭe	iə	l̀ə	iə	l̀ə
e	e, ĕ	ɪ, ĭ	é	e		eː	ɪː	e	ɪ	eː, e	è̀ː, l̀	e, eF	l̀F
ɛ	ae, ĕ	ɛ	ē	ê		æː	ɛː	æ	ɛ	ae, e	ɛ̀ː	ae, eF	ɛ̀
ai	aj	ɯj	āi	ai		ai̞	ei̞	ǎy	ĕy	ay	ù̈y	ay	ù̈y
o	ao	ʊ	ō	o		ɑᵒ	oᵘ	å	ʊ	ao	ɔ̀ː	ao	ɔ̀
au	av	ɯv	āu	au		ɑṳ	əṳ	av	əv	av	ù̈ɣ	av	ù̈ɣ
ŭ̃	ɔ̃m	ŭ̃m	ũ	um̊		aᵘm	um	ɔ̃m	ŭ̃m	om	ùm	om	ùm
ɔ̃	ɑ̃m	ɔ̃m	ã	am̊		ɑm	ɔm	ɑ̃m	ɔ̃m	ɔmɛ	ɔ̀əm	ɔm	ɔ̀əm
ã	ẵm	ɛ̃m	ã̃	ām̊		am	ᵊɑm	ẵm	ã̃ᵒm	am	ɛ̀əm	am	ɛ̀əm
ã̃ŋ	ɑ̃ŋ	ɛ̃ŋ	ã̃n̊	ām̊n̊		aŋ	ɛɑŋ	ẵŋ	æ̃ŋ	aŋ	ɛ̀əŋ	aŋ	ɛ̀əŋ
ɔ̣h	ɑ̃h	ɛ̃h	aḥ	aḥ		ah	ɛah	ã̃ḥ	æ̃ḥ	ah	ɛ̀əh	ah	ɛ̀əh

Table 1

Abbreviations: *Tl.* = transliteration; *phonol.* and *phonet.* = phonological and phonetic transcription; R = register.

Tl.	Maspero		Finot		Guesdon		Cambefort	
	R^1	R^2	R^1	R^2	R^1	R^2	R^1	R^2
ɔ	a	ò	a	o	â	o	â	o
ɔ̆	ă	ŏ	ă	ŏ	å̆	ŏ	å̆	ŏ
a	à	ea	à	ã	a	éa	e	ea
ă	ǎ	ǎ	ǎ	ǎ	ǎ	ěa	ǎ	ěa
ɪ̆	ĕ å̂	ĭ	ĕ	ĭ	ĕ	ĭ	ĕ	ĭ
i	ei	i	ĕi	i	ĕi	i	ei	i
ɯ̈	(ɯ̈ʼ)	ɯ̆ʼ	(ɯ̈ʼ)	ɯ̆ʼ	(ɯ̈ʼ)	ɯ̆ʼ	(œ̆)	ɯ̆ʼ
ɯ	(ɯʼ)	ɯʼ	(ɯʼ)	ɯʼ	(ɯʼ)	ɯʼ	(œ)	ɯʼ
ŭ	ǒ	ŭ	ǒ	ŭ	ŏ	ŭ	ŏ	ŭ
u	au	u	ó	u	âu	û	aû	u
uo	uo	uo	uo	ûo	uo	ûo	uo	ûo
ɤ	eu	eu	ɔ̀	ɔʼ	ɔʼ	œu	ɔʼ	ɤʼ
ɯɤ	euo	euo	ɯʼɔʼ	ɯ̂ʼɔʼ	eua	œua	ɯʼɔʼ	ɯʼɔʼ
ie	ie	ie	ie	îe	ie	îe	ie	îe
e	e	e	e	é	é	é	e	é
ɛ	è	è	è	ê	ê	ê	è	ê
ai	ai	ei	ai	ei	ai	ei	ǎi	ĕi
o	ò	ou	ò	ô	ō	ou	ò	ó
au	ao	òu	au	ou	au	ôu	ǎu	ŏu
ŭ̃	ǒm̥	ŭm̥	ǒm̥	ŭm̥	ŏm	ŭm	ǒm	ŭm
ɔ̃	am̥	òm̥	am̥	om̥	å̂m	ŏm	å̂m	ŏm
ã	àm̥	eam̥	àm̥	ām̥	ăm	eå̂m	ǎm	ěam
ãŋ	ǎń	eań	ǎń	āń	ăng	ěang	ǎng	ěang
ɔh̥	ǎ	ǎ	ah̥	áh̥	ǎh	eǎh	ǎh	ěah

Table 2

Abbreviations: *Tl.* = transliteration; R = register.

[ou], [oɤ] and [ᵛo]. e and ε become ĕ in the First Register
before h, s; they become ĭ in the Second Register before h. ă
becomes ε̆ in the Second Register before k, ŋ, h; elsewhere it
is ŏ.

Out of this considerable disparity between writing and pro-
nunciation the following questions arise:

1. *What was the Old Khmer vowel system like around the*
time when writing was introduced?

2. *What is the nature of the modern system?*

3. *How did the modern system develop out of the old?*

In seeking an answer to the third of these questions we may
either start with the system as it exists today and trace its
development backwards in time, or take the Old Khmer system as
our point of departure. In either case the direction of our en-
quiry must be the former, though the latter is appropriate for
a brief exposition.

Regarding the first question, it must be understood from
the outset that, once adopted, the Indic writing system could
never have been altogether adequate for the representation of
the Khmer vowels.[4] Then too it must be noted that the ortho-
graphy of words has oftentimes varied not only in the course of
historical development but also during given periods.

To begin with, r̥̆, r̥, l̥̄ and l̥ as well as ū̃, õ, ã and ɔḥ can
be dissociated from the rest of the writing system on the
grounds that these represent combinations of consonants and
vowels while nasalized vowels were and still are entirely want-
ing in Khmer; these symbols are purely Indo-Aryan. ai and au
similarly represent combinations of a vowel plus j and v, and
may hence be disregarded.

---

[4]On the Hindī vowel system and its representation in the
devanāgarī vid. my paper "Über die Vokale im Hindī," in *ZfPh*,
7 (1953): 43 sq.

Of the remaining symbols, those corresponding to a, ā, i, ī, u, ū, e and o in the *devanāgarī* and other Indic writing systems represent Khmer ɔ, a, ĭ, i, ŭ, u, e and o. The symbols ṻ, ɯ, γ and ɛ are innovations, γ being compounded of i and e, ɛ being a form of double e.[5] By means of the diacritic ´ (the sɔŋkăt) two additional vowels are distinguished: short ɔ̆ and short ă. Finally, other symbols were invented for three special diphthongs: uo, ie and ɯγ. Some of these symbols were devised in fairly recent times. This applies particularly to ṻ, ɯ and ɯγ, believed to have been introduced by the Brɔḥ Sŭgɔn (Pră sǒkǒn), who died in 1894.[6] In the old inscriptions e was still being used for γ.[7] The relationship of e and ɛ is involved. While e is unquestionably the original symbol for e, i is frequently found in its place in the inscriptions.[8] For this reason Schmidt[9] decided in favor of ɛ as the proper symbol for Khmer e and saw e as having developed from i, which becomes e (Schmidt writes é) when it occurs with a voiced initial.[10] The diphthongs ie and uo, as Schmidt has shown,[11] come from earlier ya [ja] and va [wa] and are hence not to be considered old.[12] The same holds true for ɯγ.

The recent introduction of ɯ and the other new symbols into the writing system does not of course constitute proof that the sounds they represent did not exist previously. Indeed, the latter were expressed formerly, though not accurately. From the

---

[5]Cf. the corresponding symbol in Thai.

[6]Maspero, op.cit., 63.

[7]Ibid., 106.

[8]Ibid., 95.

[9]Schmidt, op.cit., §§165 sq, pp. 159 sq.

[10]Ibid., 160.

[11]Ibid., §§199 sq, pp.180 sq; §§225, pp.198 sq.

[12]Vid. also Maspero, op.cit., 105.

absence of a perfect parallel between Mon, Khmer, Bahnar and
Stieng ɤ, Schmidt infers that ɤ is comparatively recent in
these languages.[13] The same holds true for ɯ, which Mon does
not have. Schmidt advances no opinion on ɯ. It is nevertheless
likely that both sounds existed in Khmer long before they were
represented in writing, possibly even before the adoption of
the writing system.

As to ɔ, ɔ̆ and a, ă (represented by a, a′ and ā, ā′ in Mar-
tini's transliteration), the separation of these two pairs for-
ces itself upon us, as it were. The first vowel of the *devanā-
garī*, namely a, represents a short sound in Sanskrit and kind-
red languages; in Khmer however it is a long vowel which can be
shortened by means of the sɔŋkăt (′). The second vowel of the
writing system, ā, can also be shortened by ′. It follows that
a[14] cannot stand in the same relationship to ā as it does in
Sanskrit; it must have a different articulation. The probable
articulation of Old Khmer a[14] is not hard to define: it must
have been an [ɔ] type vowel. Neither Sanskrit nor the Middle
Indic dialects had an ɔ, and in the New Indic period a was
first pronounced in this manner in Bengali. Hence there were
only two possibilities for the representation of Old Khmer *ɔ:
either by means of the symbol o, or by means of the symbol a.
One reflection of this is the frequent vacillation in the or-
thography between a and o, as in phaŋ[14] ~ phoŋ 'together',
ap ~ op 'to embrace', and ṭaŋ ~ ṭoŋ 'handle'.[15] Vacillation is
also noted between a in one language and o in another; for ex-
ample, Khmer an 'to weaken' : Mon on, Khmer ṭaŋ 'handle' :
Stieng toŋ.[16] The same sort of vacillation is seen within Mon,

---

[13] Schmidt, op.cit., §§253 sq, pp.218 sq.

[14] Still using Martini's transliteration.

[15] Maspero, op.cit., 81 sq.

[16] Id.

of which the name itself is written *mon* or *man*.[17] On the basis
of such comparisons Schmidt[18] concluded that these languages
originally had no independent o at all. In reality, these com-
parisons seem to show that a as well as o and ɔ were in exis-
tence but that words with ɔ had to be written with either a or
o. This situation persisted in Mon, but in Khmer the disadvan-
tages it entailed were overcome by the sɔŋkăt, thanks to which
a could be reserved for ɔ while ā could be reserved for a.
Short syllables then in existence could be marked with ´. That
the latter dispensed with sundry anomalies, aberrations and
mistakes is clear. The existence of an old ɔ sound side by side
with an o sound is supported by the coexistence of ε and e. In
this connection we may well wonder why a new symbol was not
also found for ɔ when one had been found for ε. That one was
not found for ɔ may have been owing to the form (e + ā) of the
symbol o.

It will be clear from the foregoing that transliterating
the sounds in question by a, a´ and ā, ā´ is altogether inap-
propriate and confusing. It is in fact downright absurd, since
ā´ = ǟ or ǎː (!). To be preferred is the transliteration of
these sounds by ɔ, ŏ and a, ă, which fulfill the conditions
stated above and have been used in the foregoing exposition.
The forms cited above should hence not be transliterated phaŋ,
ap, and the like, but phɔŋ, ɔp and so on. Since the "normal"
vowel in Khmer is long, there is an advantage in leaving the
long vowels unmarked (as by ː) and in marking the short vowels
with �‍˘.

So the old vowel system as it existed prior to the devoic-
ing of the originally voiced stops was substantially as fol-
lows. First, the normal (long) vowels:

---

[17]Cf. Maspero, op.cit., 96.

[18]Schmidt, op.cit., §230, p.195.

i [iː], written i

e [eː], written e

ɛ [ɛː], written ɛ

ɯ [ɯː], written i (ɯ)

ɤ [ɤː], written e (ɤ)

a [aː], written a

u [uː], written u

o [oː], written o

ɔ [ɔː], written ɔ

In addition to the above, the diphthongs ie, ɯɤ, uo developed from jɔ (ja) and wɔ (wa), written yɔ (ya), vɔ (va). The ortho-graphic diphthongs ai and au are combinations of vowel plus j or w.

Once the devoicing of the originally voiced stops had set in, this change affected the vowels. These split up into the two so-called registers, differentiated by pitch and openness. The First Register, here marked ´, had greater openness and higher pitch, while the Second Register, marked `, had less openness and lower pitch. Thus the total number of vowels was doubled, resulting in a system which, with the diphthongs in-cluded, was structured as follows:

ì	ɯ̀	ù
í[19]	ɯ́	ú
ìe	ɯ̀ɤ	ùo
íe	ɯ́ɤ	úo
è	ɤ̀	ò
é	ɤ́	ó
ɛ̀	à	ɔ̀
ɛ́	á	ɔ́

[19]That i after voiceless consonants (First Register) is rare

In the course of time these sounds underwent further chan-
ges of various kinds, culminating in the following inventory in
which diphthongization plays an important part. The changes are
best shown in tabular form:

old → modern		old → modern		old → modern	
M[20]	H	M	H	M	H
ì → i	ì	ɯ̀ → i̯	ɯ̀	ù → u	ù
í → ɛj	ɤj	(ɯ́)		ú	
ìe → ĭe	ìə	ɯ̀ɤ → i̯ə	ɯ̀ɤ	ùo → uo	ùə
íe → ĭe	iə	ɯ́ɤ → i̯ə	ɯɤ	úo → uo	uə
è → ɪ	è	ɤ̀ → ə	ɤ̀	ò → u	ò
é → e	e	ɤ́	(ɤ)	ó	o	o
ɛ̀ → ɛ	ɛ̀	à		ɔ̀ → ɔ	ɔ̀
ɛ́ → æ	ae	á	ʌ ay	ɔ́ → å	ao
a	a			ɑ	ɔ

According to Martini the old à became ie, which hence differs
from the ĭe from old ìe.

The reader will recognize the clear parallel between the
central and the back vowels. í develops into a diphthong or,
more accurately, into a vowel plus j, thus deviating from the

does not indicate that Mon-Khmer originally had no i in this
environment, as Schmidt supposed (op.cit., §178, p.168). What
it does suggest is that even before the adoption of writing
Khmer and Mon i after voiceless consonants had become diph-
thongized to ai [aj] (cf. Schmidt, op.cit., §242, p.211). Cf.
Khmer ṭai, Mon tai (toa) 'hand' with Bahnar ti, Khaɽia ti?,
Santali ti; Khmer cai, Mon cai 'louse' with Bahnar ʃi, Santali
se, Munḍari si-ku.

[20]M = Martini; H = Henderson.

pattern. For all practical purposes ìe and íe and the other
pairs are identical, the sole difference between them being
their register. Like à, old έ, ύ, ó are diphthongized. For the
ə in Henderson's iə and uə it is preferable to write ɣ, since
no new ə phoneme is involved. In the same way, it is better to
use ɯ and ɣ for Martini's i̧ and ə.

With Martini the modern vowel system (long vowels) consists
of the following:

i	ɯ	u
ie	ɯɣ	uo
ɪ		ʊ
e	ɣ	o
ɛ		ɔ
æ	ʌ	å
a		ɑ

For his ǐe (iĕ) see the short vowels.

But if with Henderson we take register rather than openness
as the relevant characteristic, we obtain a vowel system which
is quite different from Martini's, namely:

monophthongs:	i	ɯ	u
	e	ɣ	o
	ɛ	a	ɔ

diphthongs:	iɣ	ɯɣ	uɣ
	ae	aɣ	ao

To the above the symbol ` must be added to mark the Second Reg-
ister. As the First Register is not abnormal in any way, it
need not be marked.

Thus the curious result is that the modern system matches
the old one tolerably well, but that register — which, how-
ever, does not belong directly to the vowel system — now be-

comes operative in place of the earlier voicing. There was a return, in a sense, to the old system after the vowel changes consequent upon the old devoicing had been carried out.
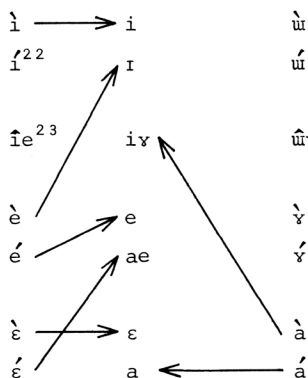
Yet the system proposed by Henderson suffers in one respect. Out of a total of nine vowels (monophthongs) register is relevant in only three cases: e : è, o : ò, ɔ : ɔ̀. The six other vowels always take either the Second Register (ì, ǜ, ù, ỳ, ɛ̀) or the First Register (a). Among the diphthongs, ae, aɣ, ao are invariably of the First Register. The three remaining diphthongs occur with both registers, but the phonetic distinction between them (iɣ : ìɣ, uɣ : ǜɣ, uɣ : ùɣ) is virtually non-existent and carries no weight in favor of including register within the phonemic system.[21] Since the relative openness of the sounds in question — e : è, o : ò, ɔ : ɔ̀ — is what is perceptible, we shall give preference to Martini's interpretation and with him set up ɪ for è, ʊ for ò, and ɑ for ɔ, while ɔ will be used for Henderson's ɔ̀. This ɪ and ʊ do not represent somewhat centralized sounds, but degrees of openness midway between i and e on the one hand and between u and o on the other.

Henderson's three diphthongs ae, aɣ, ao are preferable to Martini's interpretation which posits /æ/, /ʌ/, /å/ as phonemic monophthongs realized phonetically as diphthongs, inasmuch as no useful purpose is served in treating real diphthongs as monophthongs. What is more, ae, aɣ, ao occupy a place in the system which is analogous to that of iɣ, uɣ, uɣ.
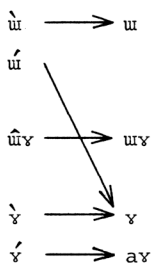
In this way the following scheme of changes results:

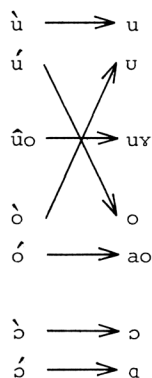[21]But cf. Henderson, op.cit., 154 sq.

old → modern	old → modern	old → modern

```
ì ──────▸ i          ɯ̀ ──────▸ ɯ         ù ──────▸ u
í²²      ↗ɪ           ɯ́                   ú       ↗ʊ
                                            ╲   ╱
îe²³   ╱   iɣ◂        ɯ̂ɣ ─────▸ ɯɣ        ûo ──╳──▸ uɣ
      ╱        ╲                          ╱   ╲
è   ╱    ↗e          ɣ̀ ───────▸ ɣ        ò ╱      ╲ o
é ╱    ↗ae           ɣ́ ───────▸ aɣ       ó ───────▸ ao

ɛ̀  ┼───▸ ɛ           à                    ɔ̀ ──────▸ ɔ
ɛ́ ╱      a ◂───── á                      ɔ́ ──────▸ ɑ
```

For the long vowels, therefore, the modern phonological system is as follows:

monophthongs:

i	ɯ	u
ɪ		ʊ
e	ɣ	o
ɛ		ɔ
a		ɑ

diphthongs:

iɣ	ɯɣ	uɣ
ae	aɣ	ao

This system is valid only for the long vowels, the only ones that can be deemed normal. Short vowels occur only before a final consonant and are hence marked with ˘. They must be considered next.

---

[22] *í becomes ɣj.*

[23] The symbol ^ represents both registers, i.e. ´ and `. îe becomes ĭ [iə], which is shorter than iɣ. Vid. the discussion of short vowels.

The old system regularly distinguished ĭ, ŭ, ă and ɔ̆. The long phonemes ɯ, ɤ and o as well as e and ε appear to have had no short counterparts in early times; hence graphemes for *ɤ̆, *ŏ, *ĕ and *ε̆ are not found. The modern writing system has a symbol for ŭ̈, while the sɔŋkăt (ʹ) is also used at times to distinguish short ĕ, as in t̪ĕc 'to break off'. The old system of short vowels was therefore:

> ĭ [i], written ĭ
>
> ă [a], written ă
>
> ŭ [u], written ŭ
>
> ɔ̆ [ɔ], written ɔ̆

With the advent of register, consequent upon the loss of voicing, these developed into:

> ĭ̌   ǔ
>
> ĭ̷   ǔ̷
>
> ǎ   ɔ̌
>
> ă̷   ɔ̷

With the lapse of time these sounds, like their long counterparts, then underwent various modifications, culminating in the following inventory:

> ĭ̌ ⟶ ŭ̈̌   ǔ ⟶ ǔ
>
> ĭ̷   ǔ̷
>
>   ⟍ ɤ̆̌   ⟍ ŏ
>
> ǎ   ɔ̌
>
> ă̷ ⟶ ā   ɔ̷ ⟶ ɔ̆
>
>   ⟍   ⟍ ŭ̈ə
>
>   ⟍ ε̆ə   ⟍ ɔ̆ə

The front vowels ĭ̌ and ĭ̷ were thus altered to nonlabial back

vowels (ŭ̈, y̆); ĭ and ŭ were lowered to (*ĕ >) y̆ and ŏ; while ă and ɔ̆ were diphthongized, the former to ɛ̆ə or ɔ̆ə, the latter usually to ɛ̆ə before h, to ɔ̆ə before m, and to ŭə[24] elsewhere. ŏ is phonetically [o], [ou], [oɣ] or [ᵁo]. The resulting system was therefore:
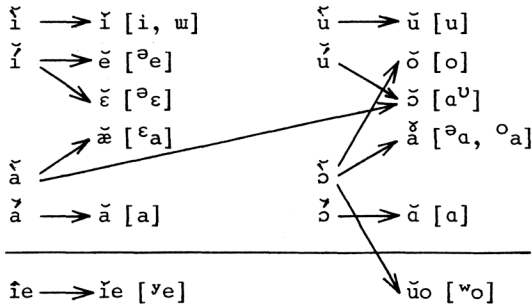
monophthongs: ŭ̈   ŭ

  y̆   ŏ

  ă   ɔ̆

  ŭɣ

diphthongs: ɛ̆ɣ

  ɔ̆ɣ

y̆, ŏ, ă, ɔ̆ belong to the First Register, the others to the Second Register.

There was an urge to adjust this new system of short vowels to that of the long vowels, inasmuch as front *ĭ, *ĕ, *ɛ̆ were lacking. Old ĭ did not invariably become ŭ̈, but in some cases remained as ĭ. Also, the old long è before h was shortened and raised to ĭ, so that a definite contrast arose between ĭ : ŭ̈ : ŭ. The old é and ɛ́ before h and s were similarly shortened, resulting in the contrast ĕ : y̆ : ŏ as a secondary development. The old shortened ɛ́, however, developed into ĕ, with the result that there was no *ɛ̆ to complete the ă : ɔ̆ contrast.

The system presented here accords with Henderson's interpretation in all essentials. Martini's analysis of the short vowels is quite different. According to him, the following developmental scheme obtains:

---

[24]These three diphthongs are properly ɛ̆ɣ, ɔ̆ɣ, ŭɣ.

ĭ ⟶ ĭ [i, ɯ]   ŭ ⟶ ŭ [u]
ĭ ⟶ ĕ [ᵊe]   ŭ ⟶ ŏ [o]
   ⟶ ɛ̆ [ᵊɛ]   ⟶ ɔ̆ [aᵛ]
      ⟶ æ̆ [ᶜa]   ⟶ ǎ [ᵊɑ, ᵒa]
ǎ   ɔ̌
ǎ ⟶ ă [a]   ɔ̆ ⟶ ā [ɑ]
_____

îe ⟶ ĭe [ʸe]   ŭo [ʷo]

According to Martini, the old ĭ became ɛ̆ before k, ŋ, h but ĕ elsewhere. The old ǎ had two replacements while ɔ̌ had three. The realization of ĭ is [i] or, more often, [ɯ] (i.e., i̥). Martini maintains that the diphthong ie is short and contrasts with i:e from old ǎ. His system hence looks as follows:

          ĭ   ŭ
          ĭe   ŭo
          ĕ   ŏ
          ɛ̆   ɔ̆
          æ̆   ǎ
          ă   ā

This corresponds to Henderson's system as below:

Old   →   Modern
                M   H[25]

ĭ   ĭ   ĭ
               ɯ̆   ɯ̆

ĭ   ĕ   ɤ̆
               ɛ̆

ĕ   –   ĭ

ĕ   –   ĕ
ɛ̆

_____

[25]Henderson's F is replaced by ˘.

Old	→	Modern	
		M	H
ǎ̯		æ̆	ɛ̆ə
		ɔ̆	ɔ̆ə
ǎ̮		ă	ă
ǔ̯		ŭ	ŭ
ǔ̮		ɔ̆	ŏ
		ŏ	ɔ̆ə
ɔ̯̌		å̆	ɛ̆ə
		ŭo	ŭə
ɔ̮̌		ɑ̆	ɔ̆

Old ǐ, ǔ and ǎ̮ are in correspondence, while the interpreta-
tions of ɔ̮̌, ǎ̯, ě, ě̮ and ě̯ differ in ways which are relatively
unimportant, as also does ǐe. Markedly different in interpreta-
tion are the correspondences for ǐ̮, ǔ̮ and ɔ̌: here the realiza-
tions which Martini advances say more in favor of Henderson.

Henderson's system is accepted, then, but it is preferable
to follow Martini to the extent of taking ɑ̆ instead of Hender-
son's ɔ̆.[26] Thus we obtain the following system of short vow-
els:

monophthongs: ĭ   ᵾ̆   ŭ

                 ě   ɤ̆   ŏ

                 ă   –   ɑ̆

diphthongs: ŭɤ

                   ɛ̆ɤ

                     ɔ̆ɤ

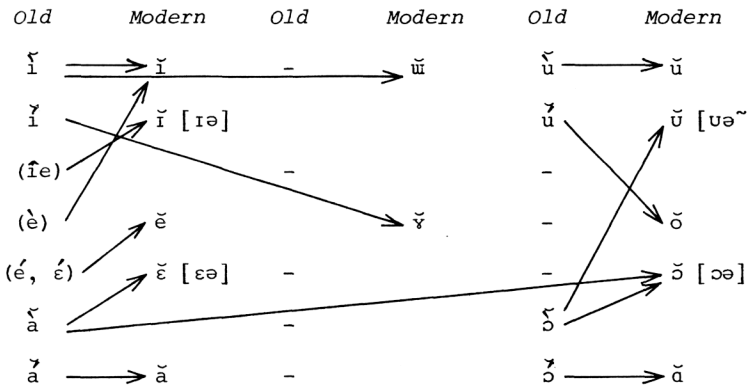Instead of ŭɤ it might be better to posit ᵾ̆ɤ, since the first

---

[26]Cf. also the long vowels!

member of this short diphthong is somewhat open.

Diphthongization may be said to be a basic drift with the short vowels, as is evidenced by the realization of ŏ and by Martini's setting up such phonetic values as [ᵊe], [ᵊɛ], [ᵋa] and [ᵊɑ], as above. For this reason it would not be amiss to set up the phonemes /ɛ̆/ [ɛɤ], /ɔ̆/ [ɔɤ] and /ŭ/ [ʊɤ] rather than Henderson's short diphthongs. To be sure, these are realized as diphthongs, but their relevant features are openness and advancement. In addition, Martini's statement that his ĭe is a short diphthong — hence ĭɤ or, more accurately, ĭɤ — can be accepted, and there is nothing to prevent our setting it up as phoneme /ĭ/. It is true, however, that this /ĭ/ differs from the other short vowels in that it can stand in final position.

Thus the realizations of ĭ, ŭ, ɛ̆, ɔ̆, all four of which exhibit phonetic lowering, are always diphthongal, with a short [ə] or [ɤ] as their second member.

This yields the following developmental scheme:

Old	Modern	Old	Modern	Old	Modern
ĩ	ĭ	-	ũ	ũ	ŭ
ĭ	ĭ [ɪə]			ŭ	ŭ [ʊə̃]
(ĩe)		-			-
(è)	ĕ		ɤ̆	-	ŏ
(é, ɛ́)	ɛ̆ [ɛə]	-		-	ɔ̆ [ɔə]
ã		-		ɔ̃	
ã	ã	-		ɔ̃	ɑ̆

Thus the modern short-vowel system matches that of the long vowels perfectly except for the fact that phonemic diphthongs are lacking here.

The end result of all these complex changes is an essen-

tially simple system:

normal (long)			short			diphthongs		
i	ɯ	u	ĭ	ɯ̆	ŭ	iɣ	ɯɣ	uɣ
ɪ		ʊ	ɪ̆		ʊ̆			
e	ɣ	o	ĕ	ɣ̆	ŏ	ae	aɣ	ao
ɛ		ɔ	ɛ̆		ɔ̆			
a		ɑ	ă		ɑ̆			

Thus the vowel system, comprising thirty phonemes, can be expressed without ambiguity with thirteen symbols: a, ɑ, e, ɛ, i, ɪ, o, ɔ, u, ʊ, ɣ, ɯ and ˘.

The Consonantism

In the writing system the following consonants are distinguished:

p	b	t	d	ṭ	(ḍ)	c	ɟ	k	g	ʔ
ph	bh	th	dh	(ṭh)	(ḍh)	ch	ɟh	kh	gh	
m		n		ṇ		ɲ		ŋ		
		l		ḷ						
		r								
	s		(ṣ)		(ʃ)				ḥ h	
v					ẏ					

To these are to be added the consonant symbols taking the diacritic ¨ (the sɔ̃ḷãp):

ṁ   (n̈)   j̇   n̈ (ï)   ṙ   v̈   ẏ   s̈   ḧ   ṗ   (b̈)

Also to be noted are the diacritic ˜ and the symbols r̆, r̥, ḷ̆ and ḷ̥ (vid. the vowels). ˜ is equivalent to -m. The symbols for the sonorant r and l of Sanskrit are graphic only, being

realized as /r, l/ plus /ŭ/ or /ɯ/. For all practical purposes
ḍ, ṭh, ḍh, ṣ and ʃ are obsolete; ñ, ï and ƀ are listed in the
Abhidhānaśabda but are equivalent to ṇ, ḷ and p̈ — though nï is
usually written for ṇḷ. ʔ is the vowel support, also used in
ligatures.

As regards articulation, original b, d, ɟ, g, bh, dh, ɟh,
gh have become voiceless, merging with p, t, c, k, ph, th, ch,
kh. The retroflex symbols are purely graphic and have no coun-
terpart in pronunciation, where they are realized as the cor-
responding dentals: ṭ [ɗ], ḍ [ɗ], th [th], ḍh [th], ṇ [n], ḷ
[l], ṣ [s]. Palatal ʃ = [s].

According to Martini c is [c] or [ʈ], being close to Rus-
sian ТЬ, and is hence often romanized *ti* or *tch*. Again accord-
ing to Martini, ɟ is [c̃] while ch and ɟh are both [tʃ]. Hender-
son, however, indicates that c and ɟ are both [c] and that ch
and ɟh are [ch]. Dialect differences may be involved here. Ac-
cording to Martini again, ṭ (as well as ḍ) and p are [d] and
[b] respectively; according to Henderson they are voiced implo-
sive [ɗ] and [ɓ], represented hereafter as preglottalized 'd
and 'b. Graphic ḥ is [h], while ~ is [m]; v is either bilabial
[w] or labiodental [v] generally without friction: [ʋ]. The
ligature hv is [f]. p̈ is [p]; m̈, ñ, ŋ and so on are realized
as the corresponding symbols without the sɔ̃ḷãp.

In final position only the following phonemes appear:

p	t	c	k
m	n	ɲ	ŋ
l			
(r)			
(s)			h
w		j	

Final r and s occur in only a few dialects. Elsewhere final orthographic r is silent and final s is replaced by /h/. Final /w, j/, phonologically consonants, form diphthongs.

The consonants fall into two groups, the registers, according to their effect on the vowels following them.[27] The sɔ̆ḻăp (¨) serves to change the register while preserving the consonant's value. Thus m̈, ji̇, ÿ, r̈, v̈ and ÿ belong with the First Register, s̈ and ḧ with the Second. The sole exception is p̈, which remains with the First Register; in this case the sɔ̆ḻăp indicates that it is to be pronounced as voiceless [p] rather than as implosive [ɓ] ('b). This ¨ was introduced in recent times,[28] and most of the forms with which it occurs are loans or derivatives in which an earlier consonant cluster is embedded; should the cluster be lost through the dropping of a prefix or addition of an infix, the second member of the cluster presupposes the same vowel as before. To mark this, ¨ is used, as in thvay /twaj/ 'to offer' (First Register) → tɔŋv̈ay /'daŋwaj/ 'offering' (First Register).[29] Otherwise v (Second Register) + a must be read /iɤ/. Compare also prɤ 'to order' (First Register) → pɔ̃r̈ɤ 'servant' (First Register),[30] m̈ɔp = hmɔp 'folded over, doubled'.

It is possible that in m̈, r̈, v̈ and the like we have the vestiges of former voiceless consonants. Yet, since all these take the First Register, it is not impossible that m̈, ji̇, ÿ and ṇ as well as r̈, ḻ, v̈ and ÿ were intended to represent former preglottalized imposives ('m, 'ɲ, 'ŋ, 'n, 'r, 'l, 'w, 'j) still

---

[27]Vid. the vowels.

[28]Maspero, op.cit., 130.

[29]Ibid., 158.

[30]Ibid., 178.

actually found in other Mon-Khmer languages such as Bahnar.[31]
In Mon and in Khmer alike it is the rule that the implosives
(except Mon 'b), whatever their origin, belong to the First
Register, which is otherwise the domain of the voiceless con-
sonants. Khmer 'b and 'd, which developed from former voiceless
consonants, are now voiced but still belong to the First Regis-
ter. Mon 'b and 'd, which developed from former voiced conso-
nants, are treated in the same way. Compare:

- Khmer ṭi, Mon tɪ 'land', Munḍari ɔte 'field';
  Khmer ṭ /'d/ < *t.
- Khmer puon, Mon pɔn, Munḍari upun 'four';
  Khmer p /'b/ < *p.

- Mon ḍaʔ, Khmer dũk, Munḍari daʔ 'water';
  Mon ḍ /'d/ < *d.
- Mon 'ba, Khmer bir, Munḍari bar 'two';
  Mon 'b /'b/ < *b.

All of these Mon and Khmer forms belong with the First Regis-
ter. So the principle is not violated, as Schmidt mistakenly
supposed,[32] but is as follows:

    implosives: *First Register*

    nonimplosives:

        formerly voiceless: *First Register*

        formerly voiced: *Second Register*

Thus voice plays no part in determining the register of the im-
plosives.

    The number of forms with ¨ is not inconsiderable. Exam-
ples:[33]

---

[31]Cf. André-G. Haudricourt, "Les consonnes préglottalisées
en Indochine," in *BSLP*, 46 (1950): 172-182.

[32]Schmidt, op.cit., §143, p.140.

[33]My examples are taken from Guesdon, whose romanizations
are given in parentheses. Cf. also Maspero, 69, 72, 130 sq.

ŋ̊   ŋ̊ak (*ngak*) 'onomatopoeia for dull sounds or cries';

             ŋ̊ai (*ngai*) 'day', for the more usual thŋai;

             phlĭt ŋ̊oy (*phlĕt ngōy*) 'monk's fan'.

ɲ̊   ɲ̊ɔɲ (*nhânh*) 'elder sister';

             thwɤ ɲ̊ɔɲ̊ɔ̆k (*thvœ nhonhẵk*) 'to toady';

             ɲɔɲ̊ieŋ (*nhonhêng*) 'to look sour, sullen';

             ɲ̊aɲ (*nhanh*) 'grimace'.

n̥ = n̈   n̥ɔj (*nây*) 'coat';

             n̥a (*na*) 'where, who';

             n̥oh̥ (*nōh*) 'that';

             n̥ɔŋ (*nōng*) 'firm, solid'.

m̈   m̈ɔ (*mâ*) 'medicine';

             m̈uv (*mâuv*) 'furrow';

             m̈au (*mau*) 'to be angry'.

ÿ   ÿɔŋ (*yâng*) 'species of plant';

             ÿav (*yav*) 'slow';

             paÿap (*bayap*) 'idol'.

r̈   r̈ɔp (*râp*) 'to prepare';

             r̈ɔs (*râs*) 'alive';

             r̈aŋ (*rang*) 'hole, pit';

             r̈uv (*râuv*) 'right, good'.

ï   pɔnïa (*bânla*) 'thorn'.

l̥ = ï   l̥ɔk (*lâk*) 'to scratch, wound';

             l̥ay (*lay*) 'all';

             l̥ek (*lêk*) 'numeral', Sanskrit *lekha*; the l̥ preserves
                the /e/.

v̈   v̈aŋ (*vang*) 'to prowl around';

             v̈ĭk (*vĕk*) 'to break in, train';

             lɔv̈a (*lova*) 'species of banana'.

It goes without saying that the question of former implo-
sives does not arise in the case of s̈ and ḧ.

š   šɔk (*sok*) 'to insert';

     ši (*si*) 'to eat'.

ḧ   ḧaŋ (*ḧeang*) = haŋ (*hang*) 'shop, store';

     ḧan (*ḧean*) 'to dare; bold';

     ḧam (*ḧeam*) 'to flow, gush'.

p̈ is voiceless /p/.

p̈   p̈ɔŋ (*p̈âng*) 'to wish, desire';

     p̈ɔdhŭc (*p̈âthŏch*) = prɔdhŭc (*prâthŏch*) 'to risk, be in
     danger';

     p̈ɔthɔm (*p̈âthâm*) = prɔthɔ̃ (*prâthãm*) 'first', Sanskrit
     prathama.

Consonant clusters are important in Khmer. To them belong
the aspirated stops, which are combinations to be interpreted
not as unit phonemes but as stops + /h/, inasmuch as they may
be dissociated, as with dhɔ̃ 'big' → dɔ̃hɔ̃ 'size'. Clusters com-
prise only two consonants, though between these a phonetic [h]
may intervene. As prior member r, y, v, ŋ, ɲ, n, 'b may not oc-
cur (in terms of pronunciation, not of orthography), while the
latter member may not be one of the aspirated series. The tran-
sition between the two members may be simple — without a pho-
netic bridge — or it may be effected by junctural [h] or [ə].
Henderson represents this purely phonetic interpolation by un-
derscored h̲ and ə̲, both devoid of phonological relevance. Pos-
sible combinations, in phonemic form rather than translitera-
tion, are the following:

   1. *Simple:* kh, ch, th, ph; kr, cr, tr, pr; sk, st,
s'd, sp, s'b, sŋ, sɲ, sn, sm, sj, sw, sr, sl.

   2. *With junctural h:* ck, tk, pk; kc, kj, tj, pc, pj;
kt, ks, pt, ps; kp, kw, cp, cw, tp, tw; kl, cl, tl, pl. These
are realized [kʰc], [kʰj], &c., and written chk, chg, thk, thg,
phk, phg, khc, khj, phc, phj, &c.

3. *With junctural* ə: k'd, k'b, c'd, c'b, t'b, p'd,
'd'b. These are realized as [kə'd], [kə'b], &c., and written
kt̬, kp, ct̬, cp, tp, pt̬, t̬p.

Clusters with nasals call for separate treatment. With
these [h] or [ə] is again interpolated, but usage fluctuates
considerably.

1. *With h:* cŋ, tŋ, pŋ; kɲ, pɲ; kn, cn, tn, pn; km,
cm, tm.

2. *With ə:* kŋ, lŋ; kn, tn, pn; km, tm, lm.

Thus bnau [pənəu] or bhnau [pʰnəu] *'Aegle marmelos,
Malabar orange'* and gman [kʰmiːən, kəmiːən] 'to have not'.[34]

3. *Clusters with m* as the prior member show junctural
ə: mr, mc, m'd, mn, realized [mər], [məc], &c. Combinations
of m + aspirated stops (e.g., mkh) also occur.

Clusters with l as the prior member include lv and lk, for
example lvɛŋ /lwɛŋ/ [lɔwɛːŋ] 'far, distant'. This is according
to Martini. Henderson[35] reports the optional insertion of junc-
tural ə or simple clustering, as in lva [lə̯vi̯ə, lvi̯ə] 'fig-
tree'.

Clusters with ʔ as the second member are articulated with
an intercalated vowel according to Martini but without any vow-
el according to Henderson. Thus Martini writes kʔ, cʔ, tʔ, pʔ,
sʔ, rʔ, lʔ as [kɑʔ], [tʃɑʔ], [tɑʔ], [pɑʔ], [sɑʔ], [rɔʔ], [lɔʔ].
For example, lʔɔ /lʔɑ/ 'good, fine' is [lɔʔɑː] for Martini,
[lʔɔ] for Henderson.

Henderson's analysis disagrees with Martini's in a few
other details. For example, she treats r as a possible cluster

---

[34]The pronunciation is as given by Martini, op.cit., 128.
[35]Henderson, op.cit., 167.

initial, as in rɔʈuv [rədɤv] 'season', and so forth.[36]

In the light of the above, the consonant system of Khmer may be said to comprise the following:

p	t	c	k	ʔ
m	n	ɲ	ŋ	
b	d			
	l			
	r			
(f)	s			h
w		y		

Plain b and d may be used for 'b and 'd inasmuch as no confusion can arise. As in the case of the vowels, then, the true consonantism is different from what the writing system would suggest.

The older consonant system may be set up as follows:

p	b	t	d	c	ɟ	k	g	ʔ
	m		n		ɲ		ŋ	
	'b		'd					
	'm		'n		'ɲ		'ŋ	
			l					
			'l					
			r					
			'r					
		s						h
	v				y			
	'v				'y			

---

[36]For details see Henderson, op.cit.

Admittedly, the positing of the implosives other than 'b and 'd is problematic. ? could be taken as earlier *'g.

The changes which culminated in the modern language may be recapitulated in the following terms. The old voiced stops became voiceless after they, together with the explosive nasals and l, r, w and y, had brought the Second Register into being. Meanwhile, 'm, 'n and the other implosives except 'b and 'd coalesced with m, n and the other explosives but kept the First Register.

Objection to this theory on the grounds that the introduction of the sɔ̆ḷăp is of recent date scarcely affects its validity, inasmuch as we are concerned not with the symbol but with register, which can be older than the diacritic. It is likely that register was marked only in later times and for the sake of precision and the avoidance of ambiguities. In this light, the assumption that words which now take the sɔ̆ḷap may formerly have had the other register would be mistaken. For the theory to be demonstrated, however, extensive etymological work would have to be done to throw more light on these complex questions. Also to be carried out is more reliable research in the phonology of related languages.