Khmer / ខ្មែរ Khmêr

transliteration: KNAB 1994 (modified 2003-01 and 2004-01), romanization: UN 1972

I. Consonant characters

		Trlit	Rom				Trlit	Rom				Trlit	Rom
1	ក	k	*kâ*	12	ឋ	ṭh	*thâ*	24	ភ	bh	*phô*
2	ខ	kh	*khâ*	13	ឌ	ḍ	*dô*	25	ម	m	*mô*
3	គ	g	*kô*	14	ឍ	ḍh	*thô*	26	យ	y	*yô*
4	ឃ	gh	*khô*	15	ណ	ṇ	*nâ*	27	រ	r	*rô*
5	ង	ṅ	*ngô*	16	ត	t	*tâ*	28	ល	l	*lô*
6	ច	c	*châ*	17	ថ	th	*thâ*	29	វ	v	*vô*
7	ឆ	ch	*chhâ*	18	ទ	d	*tô*	30	ស	s	*sâ*
8	ជ	j	*chô*	19	ធ	dh	*thô*	31	ហ	h	*hâ*
9	ឈ	jh	*chhô*	20	ន	n	*nô*	32	ឡ	ḷ	*lâ*
10	ញ	ñ	*nhô*[A]	21	ប	p	*bâ*[B]	33	អ	x	*'â*[C]
11	ដ	ṭ	*dâ*	22	ផ	ph	*phâ*					
				23	ព	b	*pô*					

[A] If used with a subscript character the lower element of the character is omitted: ញ្ច ñja *nhchô*.

[B] If used with a subscript character the romanization of the character will be *p* (see note 3).

[C] See note 4.

II. Subscript consonant characters (ញ stands for any consonant, see also note 2)

Character numbers correspond to those in Table I. In transliteration subscript consonants are not marked; if needed, however, an apostrophe is inserted before the character.

1	ក	k	*k*	12	ឋ	ṭh	*th*	23	ប	b	*p*
2	ខ	kh	*kh*	13	ឌ	ḍ	*d*	24	ភ	bh	*ph*
3	គ	g	*k*	14	ឍ	ḍh	*th*	25	ម	m	*m*
4	ឃ	gh	*kh*	15	ណ	ṇ	*n*	26	យ	y	*y*
5	ង	ṅ	*ng*	16	ត	t	*t*	27	រ	r	*r*
6	ច	c	*ch*	17	ថ	th	*th*	28	ល	l	*l*
7	ឆ	ch	*chh*	18	ទ	d	*t*	29	វ	v	*v*
8	ជ	j	*ch*	19	ធ	dh	*th*	30	ស	s	*s*
9	ឈ	jh	*chh*	20	ន	n	*n*	31	ហ	h	*h*
10	ញ	ñ	*nh*[A]	21	ផ	p	*b*	33	អ	x	*'*
11	ដ	ṭ	*d*	22	ផ	ph	*ph*				

[A] With the same base character, full form is used (ញ្ញ), in all other cases only the lower element is used as subscript, e.g. ខ្ញ khña *khnhâ*.

III. Independent vowel characters

1	ឥ	i̱	*ĕ*	6	ឬ	r̥ű	*ræ*	11	ឱ	o̱	*aô*[B]
2	ឦ	ī̱	*ei*	7	ឭ	l̥ü	*lă*	12	ឳ	a̱u̱	*au*
3	ឧ	u̱	*ŏ, ŭ*[A]	8	ឮ	l̥ű	*læ*				
4	ឩ	û	*âu*	9	ឯ	a̱e̱	*ê*				
5	ឫ	r̥ü	*ră*	10	ឰ	a̱i̱	*ai*				

[A] See note 5. Additional character: ឨ ū̱ [uː].

[B] Variation: ឲ o̱.

IV. Vocalic nuclei (ñ stands for any consonant character)

Where variants in romanization separated by a dash are given the first is to be used in the *â*-series and the second in the *ô*-series.

1	កា ā	*a–éa*[A]	7	ឈ ū	*o–u*	13	កែ ae	*ê*
2	ិ i	*ĕ–ĭ*	8	ឈ uo	*uŏ*	14	កៃ ai	*ai–ey*
3	ី ī	*ei–i*	9	កើ eu	*aeu–eu*	15	កោ o	*aô–oŭ*
4	ឹ ü	*ă*	10	កឿ üa	*œă*	16	កៅ au	*au–ŏu*
5	ឺ ű	*œ*	11	កៀ ie	*iĕ*			
6	ុ u	*ŏ–ŭ*	12	កេ e	*é*			

[A] With the character ប p a ligature ប៉ា pā *ba* is used (see note 3).

V. Shortened syllables and vocalic nuclei with *anuswara* or *visarga*

Where variants in romanization separated by a dash are given the one before the dash is to be used in the *â*-series and the one(s) after the dash in the *ô*-series. (ñ stands for any consonant character.)

1	ញ៎ំ _a_`	*á–ó*	7	ញះ _aḥ	*ăh–eăh*
2	កា៎ំ _ā_`	*ă–oă,eă*[A]	8	ុះ _uḥ	*ŏh–ŭh*
3	័ញ _ă_	*ă–oă,eă*[A]	9	កេះ _eḥ	*éh*
4	ំ _aṁ	*âm–um*	10	កោះ _oḥ	*aôh-ŏăh*
5	ុំ _uṁ	*om–ŭm*	11	ាំង _āṁṅ	*ăng-eăng*
6	ាំ _āṁ	*ăm–ŏâm*			

[A] *eă* if followed by *k, ng, h*; otherwise *oă*.

VI. Diacritical signs and other symbols (ក stands for any consonant character)

Sign			Name	Explanation
1	កក់	`	បន្តក់ pantak` *bânták*	Marks a short vowel: បត់ pat` *bát*, ខ្បស់ khbas` *khpós*, ចាក់ cāk` *chăk*, របាក់ rabāk` *rôpeăk*, មាត់ māt` *moăt*. See Table V 1, 2.
2	ក័ក	ă	សំយោគសញ្ញា saṁyog·saññā *sângyoŭksânhnha*	Marks a short vowel between two consonant characters: ច័ក căk *chăk*, វ័ង្ក văṅk *veăngk*, ភ័ព្វ bhăbv *phoăpv*. See Table V 3.
3	កំ	ṁ	និគ្គហិត niggahit *nĭkkôhĕt* (អនុស្វារ xanusvāra *ânŭsvarâ*)	Marks a short-vowel-syllable ending in *m*, see Table V 4–6, 11.
4	ៈ	ḥ	វិសជ្ជនីយ visajjanīy *vĭsârchôniy* (រៈមុខ raḥmukh *reăh mŭk*)	Marks a short-vowel-syllable ending in *h*, see Table V 7–10.
5	៉	"	ធ្មេញកណ្ដុរ dhmeñ·kaṇṭur *thménh kândŏr* (មូសិកទន្ត mūsikadant *musĕkâtôn*)	1. Transfers the main character from the *ô*-series into the *â*-series or *vice versa*: ម៉ា m"ā *mâ*, ញ៉ាង ñ"aṅ *nhâng*. 2. Makes ប p unvoiced (note 3).
6	៊	^	ត្រីស័ព្ទ trīsăpd *trei săbt*	Transfers the main character from the *â*-series into the *ô*-series: ហ៊ា h^a *hô*, ហ៊ាង h^āṅ *héang*. If ៉ or ៊ are in conflict with a vowel sign, they are substituted by ុ or ៈ: ម៉ឺ m"u *mœ*, ហ៊ី h^īu *hi*, ដំរី ṭaṁrīu *dâmrei*. (In Unicode the original codes are preserved.) In everyday texts these signs are often omitted, esp. in Indic words.
7	ៈ	à	យុគលពិន្ទុ yugalabindu *yŭkôlôpĭntŭ*	Marks an open short-vowel-syllable in Pali and Sanskrit words: សិល្បៈ silpà *sĕlpă* [selɒpaʔ].
8	៌	ŕ	រប៉ាទ rapād *rô bat*	Marks a pre-consonant *r*: ព័ណ៌ băṇŕ *poărn* [pɔɒ] (*r* is not pronounced if the main character

		has no vowel signs), ធម៌ dhamŕ *thôrm*. NB! Transliterated (and encoded in Unicode) after the main character.
9	៌	ř ...	Marks a pre-consonant *r* with a short vowel: ការ៌ gārř *koărr*. (NB! Information on the sign in sources is dubious.)
10	់	å	លេខអស្ដា lekh·xasṭā *lék âsda*	Marks vocalization of the main character (with *â* or *ô*): ស់ så *sâ*, ស់សស såsas *sâsâs*, តំណ៉ taṁṇå *tâmnâ*, ពម៉ bamå *pômô*.
11	៍	'	ទណ្ឌឃាដ daṇḍaghāṭ *tôndô khéat*	Marks the absence of vocalization: វិវាទន៍ vivād·n' *vĭvéatn* [vivɵət], បុន្យ៍ puṇy' *bŏny*, ពោធិ៍ bodhi' *poŭthĭ*, ភូមិ៍ bhūmi' *phumĭ*.
12	៉	i	កាកបាទ kākapād *kakâ bat*	Indicates some modal words: អ្ហ៊ឬ xh¡ū *'ho* [oː], ចាៗទ c¡āde *chaté*.
13	ៗ	~	លេខទោ lekh·do *lék toŭ*	Indicates duplication of the word: ផ្សេង ៗ phseṅ~ *phséng-phséng*.
14	។	·	ខណ្ឌ khaṇḍ *khând* (ល្បះ lpaḥ *lôbăh*)	Phrase ending, full stop.
15	៕	|	បរិយាន pariyān *bârĕyan* (ល្បះចប់ lpaḥcap' *lôbăh cháp*)	Paragraph ending, full stop.

Notes

1. Khmer consonants are divided into 2 series, the *â*-series and the *ô*-series, as indicated in the consonant table. With vocalic nuclei these consonants may produce different romanizations, as shown in the abbreviated vowel table: ក ka *kâ*, ក្រ kra *krâ*, គ ga *kô*, គ្រ gra *krô*. A Khmer consonant in syllable-final position, not accompanied by a vowel marker or by ់, should generally be romanized without a vowel letter following: កក kak *kâk*, អង្ក xaṅg *'ângk* (exception: ពង្រ baṅra *pôngrô*, also written ពង្រ៌ baṅrar *pôngrôr* and ពង្រហ baṅrah *pôngrôh*).

2. The second consonant of a Khmer graphic cluster is generally written below the base consonant in the special form called a "foot": ខ្ង khnaṅ *khnâng*. There is no foot for the character ឡ ḷ. The "feet" ្ត and ្ធ usually represent the characters ដ ṭ and ឍ dh, respectively, rather than ត t and ថ ṭh: ក្ដី kṭī *kdei*, កន្ធាយ kandhāy *kânthéay*, but កន្ត្រប់ kantrap` *kântráb*. A "foot" determines the series of the following vocalic nucleus unless it is a nasal (ង ṅ, ញ ñ, ណ ṇ, ន n, ម m) or យ y, រ r, ល l, វ v, ស s. In that case, the base consonant determines the vocalic series. E.g. ខ្ពង់ khbaṅ *khpông*, ល្ងa lxa *l'â*, ថ្ម thma *thmâ*, ស្វាយ svāy *svay*. Syllable-final យ y and ង ṅ sometimes appear as "feet": ស្វាយ svāy or ស្វ្យ svā'y *svay*, ទាំង dāṁṅ or ទាំ dāṁ'ṅ *teăng*. This practice appears to be optional and such irregular Khmer spellings are not reflected as such in romanization.

3. The combination ប p plus ា ā is written បា pā *ba*. The latter character is a graphic device designed to prevent confusion with ហ h. The characters ប p and បា pā with the diacritical mark ៉ are romanized *p* in the *â*-series, rather than as *b* in the *ô*-series: ប៉ង់ p"aṅ *pâng*, ប៉ាតៅ p"ātau *patau*. The diacritical mark ុ or ៀ is substituted where a conflict with another symbol written above a character would occur: ប៉ុ pīu *pei*. The characters ប p and បា pā when accompanied by a "foot" are also romanized as *p* in the *â*-series, although the Khmer diacritical mark is generally omitted: ផ្លែង plaeṅ *plêng*, ផ្ល pxa *p'â*, ប្រាប់ prāp` *prăb*.

4. The *â*-series consonant អ x is romanized by means of an apostrophe ('): ក្អែក kxaek *k'êk*, ចង្អៀត caṅxiet *châng'iet*, រអិល raxil *rô'el*, អ្វី xvī *'vei*, អាង xāṅ *'ang*. In word-initial position before a vowel, ' may be omitted: អាង xāṅ *ang*.

5. The independent character ឱ្ u^ is romanized either *ŏ* or *ŭ*. A reference source should be consulted where doubt arises. The character ឧ u is usually romanized *ŭ*: ឧស្សា ussā *ŭssa*, ឧត្តម uṭṭam *ŭtdâm* [utdɒm].

6. The main character ហ h in combinations has also the following functions:

   a. ហ្វ hv indicates *f*-sound: អាប្រ្វ៊ីក xāhvr^ik *Ahvrĭk* [a:frɯk], កាហ្វេ kāhve *kahvé*;

b. in some traditional words transfers the syllable into the *â*-series; ពេទ្យហ្ម bedy·hma *péthmâ*, ហ្មត់ចត់ hmat`cat` *hmátchát* [mɒtcɒt];

c. indicates some foreign sounds: ហ្ស៊ីណែវ hsŭnaev *Hsœnêv* [ʒəːnaew]; ហ្គោ hgo *hkoŭ* (transcription of the syllable *go*).

7. Affixed words usually preserve the same series than the word stem: សើម seum *saeum* – សន្សើម sanseum *sânsaeum* [sɒnsaəm]. In polysyllabic non-compound words the series of non-initial syllables depends on the series of the first syllable: សំរាប់ saṁrāp` *sâmrăp*, ទាហាន dāhān *téahéan* (not *téahan*).

8. When transliterating Khmer script, the logical pronunciation order of the characters is generally followed: first the main character, then the subscript(s), then the diacritical signs $\tilde{\mathrm{o}}$, $\ddot{\mathrm{o}}$ or $\overset{+}{\mathrm{o}}$, then vowel and other signs (starting with the upper ones) and finally $\acute{\mathrm{o}}$ or $\overset{\circ}{\mathrm{o}}$ ṁ. If the main character has an inherent vowel (*â* or *ô*), in transliteration a will be inserted: កម្ពុជា kambujā *Kâmpŭchéa*.

9. In transliteration, two successive consonants (except if the first is ṁ or ḥ) would mean that the second is a subscript. If there are two successive main characters, a middle dot (·) is inserted: អណ្ដូងពេជ្រ xaṇṭūṅ·bejr *Ândong Péch*.

Numbers

០ 0   ១ 1   ២ 2   ៣ 3   ៤ 4   ៥ 5   ៦ 6   ៧ 7   ៨ 8   ៩ 9

Pronunciation of romanization

a	[aː]	chh	[cʰ]	éh	[ɛh,aeh]–[ih]^D
aeu	[aə]	d	[d], [t]^B	énh	[eɲ]–[ẹɲ]^D
aô	[ao]	eă	[ɛə]	hs	[ʒ]
au	[əw]	ei	[ɛj]	hv	[f]
ă	[a]	eu	[ə̣ː]	i	[iː], [ɛj]
â	[ɒː], [ɒ]^A	ey	[ej]	iĕ	[iə]
âm	[ɒm]	ĕ	[e], [ə]^C	ĭ	[i], [ɯ]^C
âu	[əw]	ê	[ae]–[ɛː]^D	kh	[kʰ], [k]^B
á	[ɒ]	é	[eː]–[ẹː]^D	lœ	[lɯː]
b	[b], [p]^B	éa	[ẹə]	lŏ	[lɯ]
ch	[c]	éch	[ec]–[ẹc]^D	ng	[ŋ]

nh	[ɲ]	ô	[ɔː], [ɔ]^A	th	[tʰ], [t]^B
o	[oː]	ó	[ɔ]	u	[uː], [u]^A
oă	[ɔʊ]	œ	[əː]–[ɯː]^D	um	[ɔm]
om	[om]	œă	[ɯə]	uŏ	[uə]
oŭ	[ou]	œ̆	[ə]–[ɯ]^D	uv	[ə̯w]
ov	[əw]	ph	[pʰ], [p]^B	ŭ	[u]
ŏ	[o]	r	[r], [-]^B	v	[v], [w]^B
ŏăh	[ɔh]	rœ	[rɯː]	y	[j]
ŏăm	[ɔʊm]	rœ̆	[rɯ]	'	[ʔ]
ŏu	[ə̯w]	s	[s], [h]^B		

[A] Usually in an unstressed syllable.
[B] Syllable-finally.
[C] Usually in a closed syllable.
[D] The first in the *â*-series, the second in the *ô*-series.

Other reading rules (mainly in Pali and Sanskrit words)

1. Only the first character in syllable-final combinations is pronounced: រាជាណាចក្រ rājāṇācakr *réachéanachâkr* [rɛ̯əcɛ̯əna:cak], រាស្ត្រ rāstr *réastr* [rɛ̯əh].

2. Word-final i and u are often not pronounced: ជាតិ jāti *chéatĭ* [cɛ̯ət], ភូមិ bhūmi *phumĭ* [pʰuːm] (aga vrd មតិ mati *môtĕ* [mɛəteʔ]).

3. Consonant characters without vowel signs are often pronounced with a short vowel [ɒ]~[a] in the *â*-series, and [ɛə]~[ɔɒ] in the *ô*-series, in the latter case the choice of the diphthong depends on the following consonant (see V 2 and 3; open syllables end with a glottal stop and therefore are pronounced [ɛə]): សព្វ sabv *sâbv* [sɒp], សត្វ satv *sâtv* [sat], ភរិយា bhariyā *phôrĭyéa* [pʰɛərijɛ̯ə].

4. Vowel sign ា ā is often pronounced [a:] in the *ô*-series: ឃោសនា ghosanā *khŏŭsânéa* [kʰŏusana:], សុរា surā *sŏréa* [sora:], ចលនា calanā *châlônéa* [calana:].

5. Vowel sign េ e is often used instead of ែ ae: កេណ្ឌ keṇḍ *kénd* [kaen], គ្រូពេទ grūbed *krupét* [kru:pɛ:t].

6. Consonants ត t and ប p may also denote syllables [dɒ:] and [pa]: ពតិមាន batŕamān *pôrtâméan* [pɔ:dɒ:mɛ̯ən], បរមាណូ paramānū *bârâmano* [parama:no:], បណ្ណាល័យ paṇṇālăy *bânnalăy* [panna:laj].

Khmer script styles

Khmer script is used mainly in two styles that are known អក្សរឈរ *âksâr chhor* ('standing script') and **អក្សរមូល** *âksâr mul* ('round script'). The characters sometimes differ considerably.

Standing	Round	Translit.	Subscr. (st.)	Subscr. (rd.)		Standing	Round	Translit.	Subscr. (st.)	Subscr. (rd.)		Standing	Round	Translit.	Form (st.)	Form (rd.)
ក	ក	k	◌	◌		ធ	ធ	dh	◌	◌				ü	◌	◌
ខ	ខ	kh	◌	◌		ន	ន	n	◌	◌				ű	◌	◌
គ	គ	g	◌	◌		ប	ប	p	◌	◌		ឧ	ឧ	u	◌ុ	◌ុ
ឃ	ឃ	gh	◌	◌		ផ	ផ	ph	◌	◌		ឩ	ឩ	ū	◌ូ	◌ូ
ង	ង	ṅ	◌	◌		ព	ព	b	◌	◌		ឪ	ឪ	û		
ច	ច	c	◌	◌		ភ	ភ	bh	◌	◌				uo	◌ួ	◌ួ
ឆ	ឆ	ch	◌	◌		ម	ម	m	◌	◌		ឫ	ឫ	rü		
ជ	ជ	j	◌	◌		យ	យ	y	◌	◌		ឬ	ឬ	rű		
ឈ	ឈ	jh	◌	◌		រ	រ	r	◌	◌		ឭ	ឭ	lü		
ញ	ញ	ñ	◌	◌		ល	ល	l	◌	◌		ឮ	ឮ	lű		
ដ	ដ	ṭ	◌	◌		វ	វ	v	◌	◌				eu	េ◌ើ	េ◌ើ
ឋ	ឋ	ṭh	◌	◌		ស	ស	s	◌	◌				üa	េ◌ឿ	េ◌ឿ
ឌ	ឌ	ḍ	◌	◌		ហ	ហ	h	◌	◌				ie	◌ៀ	◌ៀ
ឍ	ឍ	ḍh	◌	◌		ឡ	ឡ	!						e	េ◌	េ◌
ណ	ណ	ṇ	◌	◌		អ	អ	x	◌	◌		ឯ	ឯ	ae	ែ◌	ែ◌
ត	ត	t	◌	◌				ā	◌ា	◌ា		ឰ	ឰ	ai	ៃ◌	ៃ◌
ថ	ថ	th	◌	◌				i	◌ិ	◌ិ		ឱ	ឱ	o	ោ◌	ោ◌
ទ	ទ	d	◌	◌				ī	◌ី	◌ី		ឲ	ឲ	ò		

		au	ເ◌	ເ◌		ḥ	◌ះ	◌ៈ		ŕ	◌̀	◌̃
ឨ	ឩ	`	◌់	◌៉		à	◌៉ះ	◌៉ះ		å	◌៌	◌៌
		ă	◌៊	◌៊		"	◌̈	◌̈		'	◌៍	◌៍
		ṁ	◌̊	◌̊		^	◌̃	◌̃	ៗ	�	~	៕	៕