# Khmer

## Consonants

| Full Form | Subscript | Romanization | Full Form | Subscript | Romanization |
|-----------|-----------|--------------|-----------|-----------|--------------|
| ñ | ្ | k | ឌ | ្ | d |
| ខ | ្ | kh | ឍ | ្ | dh |
| គ | ្ | g | ឞ | ្ | n |
| ឃ | ្ | gh | ប | ្ | p |
| ង | ្ | ṅ | ផ | ្ | ph |
| ច | ្ | c | ព | ្ | b |
| ឆ | ្ | ch | ភ | ្ | bh |
| ជ | ្ | j | ម | ្ | m |
| ឈ | ្ | jh | យ | ្ | y |
| ញ | ្ or ្ | ñ | រ | ្ | r |
| ដ | ្ | ṭ | ល | ្ | l |
| ឋ | ្ | ṭh | វ | ្ | v |
| ឌ | ្ | ḍ | ឝ | ្ | ś |
| ឍ | ្ | ḍh | ឞ | ្ | ṣ |
| ណ | ្ | ṇ | ស | ្ | s |
| ត | ្ | t | ហ | ្ | h |
| ថ | ្ | th | ឡ | - | ḷ |
|   |   |   | អ | ្ | q |

## Vowels

| Independent | Romanization | Independent | Romanization |
|-------------|--------------|-------------|--------------|
| ឥ | i | ឱ or ឱ | o |
| ឦ | ī | ឱ or ឱ | au |
| ឧ | u | ឫ | ṛ̥ |
| ឨ or ឩ | ū | ឬ | r̥̄ |
| ឯ | e | ឭ | ḷ̥ |
| ឰ | ai | ឮ | l̥̄ |

**Vowels**

| Dependent | Romanization | Dependent | Romanization |
|---|---|---|---|
| □ | □a | ◌ | □ua |
| □□ | □a□ | ເ◌ | □oe |
| ◌ | □á□ | ເ◌ | □ẏa |
| ◌ | □ă□ | ເ◌ | □ia |
| ◌ | □ā | ເ□ | □e |
| ◌ | □â□ | ◌ | □ae |
| ◌ | □i | ເ□ | □ai |
| ◌ | □ĭ | ເ◌ | □o |
| ◌ | □ẏ | ເ◌ | □au |
| ◌ | □ȳ | ◌ | □aṃ |
| ◌ | □u | □ះ | □aḥ |
| ◌ | □ū | □ៈ | □à |

**Diacritical Marks**

| Vernacular | Alternative | Romanization |
|---|---|---|
| ◌ | ◌ | ″ (hard sign) |
| ◌ | ◌ | ′ (soft sign (prime)) |
| ◌ | | r□ |
| ◌ | | □° (circle above) |
| ◌ | | □' (alif) |
| ◌ | | □' (ayn) |
| ◌ | | □˙ (dot above) |

**Notes**

1. In the consonant portion of this romanization table, the special character □ shows the position of a Khmer script character below which a subscript character is written. A subscript character is always romanized after a full form character, without an intervening vowel, as in ក្រខ្វាក់ (*krakhvák*).

2. When ញ (*ñ*) occurs with a subscript character, the lower element is omitted, as in ញ្ (*ñj*). When ញ occurs as its own subscript, it takes the full form ញ, as in កញ្ញា (*kaññā*). Otherwise, the subscript has the form of the lower element alone, as in ខ្ញ (*khñ*).

3. The consonant ប (*p*), followed by the vowel ា (*ā*), takes the special form បា.

4.   In the vowel columns, □ shows the position of the consonant relative to the vowel. This applies to both the Khmer vernacular and to the romanization columns. It should be noted that □ in the Khmer vernacular column can also represent a final consonant with no vowel following, in which case it is romanized simply as □ , as in ដាប់ (*dăb*).

5.   The consonants ° (*m*), and ៈ (*ḥ*) are always preceded by a vowel, but, being finals, never themselves bear a vowel. Vowels other than *a* may precede them, as in ទុំ (*ṭuṃ*), សេះ (*seḥ*).

6.   The diacritics ៉ and ៓ are romanized by // and ′ respectively, immediately following the consonant they modify. They have the alternative form when they co-occur with one of the superscript vowels ◌ី, ◌ឹ, ◌ឺ, and ◌ឿ. When ◌ុ co-occurs with one of the superscript vowels and with one of the consonants គ, ម, យ, ឈ, ឞ, រ, or ប, it is romanized as // , as in ពី̃ (*p″ī*). When ◌ុ co-occurs with one of the superscript vowels and with one of the consonants ន, ឩ, or ហ, it is romanized ′ , as in ស̃ី (*s′ī*). Otherwise, ◌ុ represents the vowel *u*, as in មុន (*mun*).

7.   The diacritics □°, □′, □‘, and □˙ in the romanization column are placed after the last letter of the word in which they occur, as in ក្សត្រិយ៍ (*ksatriy°*); ចាះ (*cāḥ′*); ត្ (*ṭa′*); អាត្មន៍ (*qātman˙*).

8.   Conventional signs are: ៗ, romanized by repeating the preceding word or phrase; ៚ romanized as *.l.* ; ៕ , romanized as *.p.* ; ◌ , romanized by means of a hyphen ( - ); ៖ , romanized by means of a colon ( : ), and ។ and ៕ , romanized by means of a period ( . ). The signs ៙ and ៘ are omitted in romanization.

9.   The numerals are: ០ (0), ១ (1), ២ (2), ៣ (3), ៤ (4), ៥ (5), ៦ (6), ៧ (7), ៨ (8), and ៩ (9).

10.   Khmer words are not written separately and spacing occurs only after longer phrases. When romanizing, the shortest written form which can stand alone as a word is treated as such. This applies also to Pali and Sanskrit loan-words. Other loan-words are divided as the original language.

## SPECIAL CHARACTERS AND CHARACTER MODIFIERS IN ROMANIZATION

| Special Characters | Name | USMARC hexadecimal code |
|---|---|---|
| ′ | soft sign (prime) | A7 |
| ᾽ | alif | AE |
| ‘ | ayn | B0 |
| // | hard sign (double prime) | B7 |

| Character Modifiers | Name | USMARC hexadecimal code |
|---|---|---|
| ◌̀ | grave | E1 |
| ◌́ | acute | E2 |
| ◌̂ | circumflex | E3 |
| ◌̃ | tilde | E4 |
| ◌̄ | macron | E5 |
| ◌̆ | breve | E6 |
| ◌̇ | dot above | E7 |
| ◌̊ | circle above | EA |
| ◌̣ | dot below | F2 |
| ◌̥ | circle below | F4 |