

[en.wikipedia.org](https://en.wikipedia.org)

## Khmer script - Wikipedia

34-43 minutes

Khmer	
អក្ខរក្រមខ្មែរ	
Type	<a href="#">Abugida</a>
Languages	<a href="#">Khmer</a> <a href="#">Sanskrit</a> <a href="#">Pali</a>
Time period	About 611 – present <sup>[1]</sup>
Parent systems	<a href="#">Proto-Sinaitic alphabet</a> <ul style="list-style-type: none"> <li><a href="#">Phoenician alphabet</a></li> <li><a href="#">Aramaic alphabet</a></li> <li><a href="#">Brahmi</a></li> <li><a href="#">Tamil-Brahmi</a></li> <li><a href="#">Pallava</a></li> <li>Khmer</li> </ul>
Child systems	<a href="#">Thai</a> <a href="#">Lao</a>
Sister systems	<a href="#">Mon</a> <a href="#">Old Kawi</a>
Direction	Left-to-right
<a href="#">ISO 15924</a>	Khmr , 355

Unicode alias	Khmer
<a href="#">Unicode range</a>	<ul style="list-style-type: none"> <li><a href="#">U+1780–U+17FF</a> Khmer</li> <li><a href="#">U+19E0–U+19FF</a> Khmer Symbols</li> </ul>
<p><b>This article contains <a href="#">IPA</a> phonetic symbols.</b> Without proper <a href="#">rendering support</a>, you may see <a href="#">question marks, boxes, or other symbols</a> instead of <a href="#">Unicode</a> characters. For an introductory guide on IPA symbols, see <a href="#">Help:IPA</a>.</p>	

<a href="#">Brahmic scripts</a>
The Brahmic script and its descendants
Northern Brahmic <a href="#">[show]</a>
Southern Brahmic <a href="#">[show]</a>
<ul style="list-style-type: none"> <li><a href="#">v</a></li> <li><a href="#">t</a></li> <li><a href="#">e</a></li> </ul>

The **Khmer script** (*Khmer*: អក្សរខ្មែរ; IPA: [ʔaʔsaːkʰmae] <sup>[2]</sup> is an [abugida](#) (alphasyllabary) script used to write the [Khmer language](#) (the official language of [Cambodia](#)). It is also used to write [Pali](#) in the Buddhist liturgy of Cambodia and Thailand.

The Khmer script was adapted from the [Pallava script](#), which ultimately descended from the [Brahmi script](#), which was used in southern India and South East Asia during the 5th and 6th centuries AD.<sup>[3]</sup> The oldest dated inscription in Khmer was found at [Angkor Borei District](#) in [Takéo Province](#) south of Phnom Penh and dates from 611.<sup>[4]</sup> The modern Khmer script differs somewhat from precedent forms seen on the inscriptions of the ruins of [Angkor](#). The [Thai](#) and [Lao](#) scripts are descendants of an older form of the Khmer script.





Ancient Khmer script engraved on stone.

Khmer is written from [left to right](#). Words within the same sentence or phrase are generally run together with no [spaces](#) between them. [Consonant clusters](#) within a word are "stacked", with the second (and occasionally third) consonant being written in reduced form under the main consonant. Originally there were 35 consonant characters, but modern Khmer uses only 33. Each character represents a consonant sound together with an [inherent vowel](#), either *â* or *ô*; in many cases, in the absence of another vowel mark, the inherent vowel is to be pronounced after the consonant.

There are some independent [vowel](#) characters, but vowel sounds are more commonly represented as dependent vowels, additional marks accompanying a consonant character, and indicating what vowel sound is to be pronounced after that consonant (or consonant cluster). Most dependent vowels have two different pronunciations, depending in most cases on the inherent vowel of the consonant to which they are added. There are also a number of [diacritics](#) used to indicate further modifications in pronunciation. The script also includes its own [numerals](#) and [punctuation marks](#).

## Consonants[\[edit\]](#)

There are 35 Khmer [consonant](#) symbols, although modern Khmer only uses 33, two having become obsolete. Each consonant has an [inherent vowel](#): *â* /*ɑː*/ or *ô* /*ɔː*/; equivalently, each consonant is said to belong to the *a*-series or *o*-series. A consonant's series determines the pronunciation of the [dependent vowel](#) symbols which may be attached to it, and in some positions the sound of the inherent vowel is itself pronounced. The two series originally represented [voiceless](#) and [voiced](#) consonants respectively (and are still referred to as such in Khmer); [sound changes](#) during the [Middle Khmer](#) period affected vowels following voiceless consonants, and these changes were preserved even though the distinctive voicing was lost (see [phonation in Khmer](#)).

Each consonant, with one exception, also has a subscript form. These may also be called "sub-consonants"; the Khmer phrase is ដើមអក្សរ *cheung âksâr*,

meaning "foot of a letter". Most subscript consonants resemble the corresponding consonant symbol, but in a smaller and possibly simplified form, although in a few cases there is no obvious resemblance. Most subscript consonants are written directly below other consonants, although subscript *r* appears to the left, while a few others have ascending elements which appear to the right. Subscripts are used in writing [consonant clusters](#) (consonants pronounced consecutively in a word with no vowel sound between them). Clusters in Khmer normally consist of two consonants, although occasionally in the middle of a word there will be three. The first consonant in a cluster is written using the main consonant symbol, with the second (and third, if present) attached to it in subscript form. Subscripts were previously also used to write final consonants; in modern Khmer this may be done, optionally, in some words ending *-ng* or *-y*, such as ឲ្យ *aôy* ("give").

The consonants and their subscript forms are listed in the following table. Usual phonetic values are given using the [International Phonetic Alphabet](#) (IPA); variations are described below the table. The sound system is described in detail at [Khmer phonology](#). The spoken [name](#) of each consonant letter is its value together with its inherent vowel. Transliterations are given using the [UNGEGN](#) system;<sup>[5]</sup> for other systems see [Romanization of Khmer](#).

Consonant	Subscript form	Full value (with inherent vowel)		Consonant value	
		IPA	UN	IPA	UN
ក	ក្រ	[kɑː]	<i>kâ</i>	[k]	<i>k</i>
ខ	ខ្រ	[kʰɑː]	<i>khâ</i>	[kʰ]	<i>kh</i>
គ	គ្រ	[kɔː]	<i>kô</i>	[k]	<i>k</i>
ឃ	ឃ្រ	[kʰɔː]	<i>khô</i>	[kʰ]	<i>kh</i>
ង	ង្រ	[ŋɔː]	<i>ngô</i>	[ŋ]	<i>ng</i>
ច	ច្រ	[cɑː]	<i>châ</i>	[c]	<i>ch</i>
ឆ	ឆ្រ	[cʰɑː]	<i>chhâ</i>	[cʰ]	<i>chh</i>
ជ	ជ្រ	[cɔː]	<i>chô</i>	[c]	<i>ch</i>

ឈ	្រ័	[cʰɔː]	<i>chhô</i>	[cʰ]	<i>chh</i>
ញ	្រ័	[ɲɔː]	<i>nhô</i>	[ɲ]	<i>nh</i>
ដ	្រ័	[dɔː]	<i>dâ</i>	[d]	<i>d</i>
ថ	្រ័	[tʰɔː]	<i>thâ</i>	[tʰ]	<i>th</i>
ឌ	្រ័	[dɔː]	<i>dô</i>	[d]	<i>d</i>
ឆ	្រ័	[tʰɔː]	<i>thô</i>	[tʰ]	<i>th</i>
ណ	្រ័	[nɔː]	<i>nâ</i>	[n]	<i>n</i>
ត	្រ័	[tɔː]	<i>tâ</i>	[t]	<i>t</i>
ថ	្រ័	[tʰɔː]	<i>thâ</i>	[tʰ]	<i>th</i>
ទ	្រ័	[tɔː]	<i>tô</i>	[t]	<i>t</i>
ឆ	្រ័	[tʰɔː]	<i>thô</i>	[tʰ]	<i>th</i>
ន	្រ័	[nɔː]	<i>nô</i>	[n]	<i>n</i>
ប	្រ័	[bɔː]	<i>bâ</i>	[b], [p]	<i>b, p</i>
ផ	្រ័	[pʰɔː]	<i>phâ</i>	[pʰ]	<i>ph</i>
ព	្រ័	[pɔː]	<i>pô</i>	[p]	<i>p</i>
ភ	្រ័	[pʰɔː]	<i>phô</i>	[pʰ]	<i>ph</i>
ម	្រ័	[mɔː]	<i>mô</i>	[m]	<i>m</i>
យ	្រ័	[jɔː]	<i>yô</i>	[j]	<i>y</i>
រ	្រ័	[rɔː]	<i>rô</i>	[r]	<i>r</i>
ល	្រ័	[lɔː]	<i>lô</i>	[l]	<i>l</i>
វ	្រ័	[vɔː]	<i>vô</i>	[v]	<i>v</i>
គ	្រ័	Obsolete; historically used for <a href="#">palatal s</a> Used only for Pali/Sanskrit transliteration <sup>[6]</sup>			
ថ	្រ័	Obsolete; historically used for <a href="#">retroflex s</a>			

		Used only for Pali/Sanskrit transliteration <sup>[6]</sup>			
ស	្រ	[sɑː]	<i>sâ</i>	[s]	<i>s</i>
ហ	្រ	[hɑː]	<i>hâ</i>	[h]	<i>h</i>
ឡ	none <sup>[7]</sup>	[lɑː]	<i>lâ</i>	[l]	<i>l</i>
អ	្រ	[ʔɑː]	<i>'â</i>	[ʔ]	<i>'</i>

The letter *bâ* appears in somewhat modified form (e.g. ឃ) when combined with certain dependent vowels (see [Ligatures](#)).

The letter ញ *nhô* is written without the lower curve when a subscript is added. When it is subscripted to itself, the subscript is a smaller form of the entire letter: ញ្ច -*nhnh*-.

Note that ដ *dâ* and ត *tâ* have the same subscript form. In initial clusters this subscript is always pronounced [d], but in medial positions it is [d] in some words and [t] in others.

The series ដ *dâ*, ថ *thâ*, ឌ *dô*, ឍ *thô*, ណ *nâ* originally represented [retroflex consonants](#) in the Indic parent scripts. The second, third and fourth of these are rare, and occur only for etymological reasons in a few Pali and Sanskrit loanwords. Because the sound /n/ is common, and often grammatically productive, in Mon-Khmer languages, the fifth of this group, ណ, was adapted as an a-series counterpart of ន *nô* for convenience (all other nasal consonants are o-series).

### Variation in pronunciation<sup>[edit]</sup>

The aspirated consonant letters (*kh*-, *chh*-, *th*-, *ph*-) are pronounced with aspiration only before a vowel. There is also slight aspiration with *k*, *ch*, *t* and *p* sounds before [certain consonants](#), but this is regardless of whether they are spelt with a letter that indicates aspiration.

A Khmer word cannot end with more than one consonant sound, so subscript consonants at the end of words (which appear for etymological reasons) are not pronounced, although they may come to be pronounced when the same word begins a compound.

In some words, a single medial consonant symbol represents both the final consonant of one syllable and the initial consonant of the next.

The letter ប *bâ* represents [b] only before a vowel. When final or followed by a subscript consonant, it is pronounced [p] (and in the case where it is followed by a subscript consonant, it is also romanized as *p* in the UN system). For modification to *p* by means of a diacritic, see [Supplementary consonants](#). The letter, which represented /p/ in Indic scripts, also often maintains the [p] sound in certain words borrowed from Sanskrit and Pali.

The letters ដ *dâ* and ឌ *dô* are pronounced [t] when final. The letter ត *tâ* is pronounced [d] in initial position in a weak syllable ending with a nasal.

In final position, letters representing a [k] sound (*k-*, *kh-*) are pronounced as a glottal stop [ʔ] after the vowels [ɑː], [aː], [iə], [iə̃], [uə], [ɑ], [a], [ěə], [ũə]. The letter រ *rô* is silent when final (in most dialects; see [Northern Khmer](#)). The letter ស *sâ* when final is pronounced /h/ (which in this position approaches [ç]).

### Supplementary consonants[\[edit\]](#)

The Khmer writing system includes supplementary consonants, used in certain [loanwords](#), particularly from [French](#) and [Thai](#). These mostly represent sounds which do not occur in native words, or for which the native letters are restricted to one of the two vowel series. Most of them are [digraphs](#), formed by stacking a subscript under the letter ហ *hâ*, with an additional *treisâpt* [diacritic](#) if required to change the inherent vowel to *ô*. The character for *pâ*, however, is formed by placing the *musĕkâtônd* ("mouse teeth") diacritic over the character ប *bâ*.

Supplementary consonant	Description	Full value (with inherent vowel)		Consonant value		Notes
		IPA	UN <a href="#">[citation needed]</a>	IPA	UN <a href="#">[citation needed]</a>	
ហ្គ	<i>hâ + kô</i>	[gɑː]	<i>gâ</i>	[g]	<i>g</i>	Example: ហ្គាស, [gas] ( <i>'gas'</i> )
ហ្គ័	<i>hâ + kô + diacritic</i>	[gɔː]	<i>gô</i>	[g]	<i>g</i>	

ហ្ល	<i>hâ + nô</i>	[nɑː]	<i>nâ</i>	[n]	<i>n</i>	Example: ហ្លាំង or ហ្លាំង, [nan] ( <i>'shadow play'</i> from Thai: หนัง)
ប៉	<i>bâ +</i> diacritic	[pɑː]	<i>pâ</i>	[p]	<i>p</i>	Example: ប៉ាក់, [pak] (to 'embroider'), ប៉ង, [pan] ( <i>'bread'</i> )
ហ្គ	<i>hâ + mô</i>	[maː]	<i>mâ</i>	[m]	<i>m</i>	Example: គ្រូហ្គ, [kruː maː] ( <i>'shaman'</i> , from Thai: หมอ)
ហ្ល	<i>hâ + lô</i>	[lɑː]	<i>lâ</i>	[l]	<i>l</i>	Example: ហ្លង, [luəŋ] ( <i>'king'</i> , from Thai: หลวง)
ហ្វ	<i>hâ + vô</i>	[fɑː], [vɑː]	<i>fâ, vâ</i>	[f], [v]	<i>f, v</i>	Pronounced [v] in ហ្វង, [vɑŋ] ( <i>'clear'</i> ) and [f] in កាហ្វេ, [kaafeɛ] ( <i>'coffee'</i> )
ហ្វ៊	<i>hâ + vô +</i> diacritic	[fɔː], [vɔː]	<i>fô, vô</i>	[f], [v]	<i>f, v</i>	Example: ហ្វីល, [fiːl] ( <i>'film'</i> )
ហ្ស	<i>hâ + sâ</i>	[ʒɑː], [zɑː]	<i>žâ, zâ</i>	[ʒ], [z]	<i>ž, z</i>	Example: ហ្សាស, [ʒas]



						('jazz')
ហ្ស៊	<i>hâ + sâ +</i> diacritic	[ʒoː], [zoː]	žô, zô	[ʒ], [z]	ž, z	Example: ហ្ស៊ីប, [ʒiːp] ('jeep')

## Dependent vowels[\[edit\]](#)

Most Khmer vowel sounds are written using dependent, or [diacritical](#), vowel symbols, known in Khmer as ស្រៈនិស្ស័យ *srăk nissăy* or ស្រៈផ្សំ *srăk phsâm* ("connecting vowel"). These can only be written in combination with a consonant (or consonant cluster). The vowel is pronounced after the consonant (or cluster), even though some of the symbols have graphical elements which appear above, below or to the left of the consonant character. Most of the vowel symbols have two possible pronunciations, depending on the inherent vowel of the consonant to which it is added. Their pronunciations may also be different in [weak syllables](#), and when they are shortened (e.g. by means of a diacritic). [Absence of a dependent vowel](#) (or diacritic) often implies that a syllable-initial consonant is followed by the sound of its inherent vowel.

In determining the inherent vowel of a consonant cluster (i.e. how a following dependent vowel will be pronounced), [stops](#) and [fricatives](#) are dominant over [sonorants](#). For any consonant cluster including a combination of these sounds, a following dependent vowel is pronounced according to the dominant consonant, regardless of its position in the cluster. When both members of a cluster are dominant, the subscript consonant determines the pronunciation of a following dependent vowel. A non-dominant consonant (and in some words also ហ្ម *hâ*) will also have its inherent vowel changed by a preceding dominant consonant in the same word, even when there is a vowel between them, although some words (especially among those with more than two syllables) do not obey this rule.

The dependent vowels are listed below, in conventional form with a dotted circle as a dummy consonant symbol, and in combination with the a-series letter អ 'â. The IPA values given are representative of dialects from the northwest and central plains regions, specifically from the [Battambang](#) area, upon which [Standard](#) Khmer is based. Vowel pronunciation varies widely in other dialects such as [Northern Khmer](#), where diphthongs are leveled, and

Western Khmer, in which breathy voice and modal voice phonations are still contrastive.

Dependent vowel	Example	IPA <sup>[2]</sup>		UN		Notes
		a-series	o-series	a-series	o-series	
(none)	អ៊	[ɑ:]	[ɔ:]	â	ô	See <a href="#">Modification by diacritics</a> and <a href="#">Consonants with no dependent vowel</a> .
ា	អ៊ា	[a:]	[i:ə] <sup>[8]</sup>	a	éa	See <a href="#">Modification by diacritics</a> . អ៊ា, the o-series of ា, is slightly distinct from េ. (អ៊ា ~ "air" vs េ ~ "ear")
ិ	អ៊ិ	[ə], [e]	[ɨ], [i]	ě	ĩ	Pronounced [e]/[i] in syllables with no written final consonant (a glottal stop is then added if the syllable is

						stressed; however in some words the vowel is silent when final, and in some words in which it is not word- final it is pronounced [əj]). In the o-series, combines with final អ yô to sound [i:]. (See also <a href="#">Modification by diacritics.</a> )
័	អ័	[əj]	[i:]	ei	i	
័	អ័	[ə]	[i]	œ		
័	អ័	[əi]	[i:]	œ		
័	អ័	[o]	[u]	ō	ū	See <a href="#">Modification by diacritics.</a> In a stressed syllable with no written final consonant, the vowel is followed by

						a glottal stop [ʔ], or by [k] in the word តុ <i>tōk</i> ("table") (but the vowel is silent when final in certain words).
្ដ	អ្ដ	[ou]	[u:]	<i>o</i>	<i>u</i>	Becomes [əw]/[iɰ] before a final វ <i>vô</i> .
្ដ	អ្ដ	[uə]	<i>uǎ</i>			
្ដ	អ្ដ	[aə]	[ə:]	<i>aeu</i>	<i>eu</i>	See <a href="#">Modification by diacritics</a> .
្ដ	អ្ដ	[iə]	<i>œă</i>			
្ដ	អ្ដ	[iə]	<i>iě</i>			
្ដ	អ្ដ	[ei]	[e:]	<i>é</i>		Becomes [ə]/[i] before palatals (or in the a-series, [a] before [c] in some words). Pronounced [ae]/[ɛ:] in some

						words. See also <a href="#">Modification by diacritics</a> .
ខែ	ខែ	[ae]	[ɛ:]	ê		See <a href="#">Modification by diacritics</a> .
ខែ	ខែ	[aj]	[i:]	ai	ey	
ខែ	ខែ	[ao]	[o:]	aô	oŭ	See <a href="#">Modification by diacritics</a> .
ខែ	ខែ	[aw]	[i:w]	au	ôu	

The spoken name of each dependent vowel consists of the word ស្រៈ *srǎk* [sraʔ]("vowel") followed by the vowel's a-series value preceded by a glottal stop (and also followed by a glottal stop in the case of short vowels).

### Modification by diacritics[\[edit\]](#)

The addition of some of the [Khmer diacritics](#) can modify the length and value of inherent or dependent vowels.

The following table shows combinations with the *nĭkkôhĕt* and *reăhmŭkh* diacritics, representing final [m] and [h]. They are shown with the a-series consonant អ 'ă.


Combination	IPA		UN		Notes
	a-series	o-series	a-series	o-series	
អំ	[om]	[um]	<i>om</i>	<i>ŭm</i>	
អំ	[am]	[um]	<i>âm</i>	<i>um</i>	The word ធំ "big" is pronounced

					[tʰom] (but [tʰum] in some dialects).
អំ	[am]	[ǽm]	<i>ǎm</i>	<i>ǽm</i>	When followed by ង <i>ngô</i> , becomes [aŋ]/[eəŋ] <i>ǎng/eǎng</i> .
អ៊	[aʰ]	[ǽəʰ]	<i>ǎh</i>	<i>eǎh</i>	
អិ្គ	[eʰ]	[iʰ]	<i>ěh</i>	<i>ĩh</i>	
អុ្គ	[oʰ]	[uʰ]	<i>ǫh</i>	<i>ũh</i>	
អេ្គ	[eʰ]	[iʰ]	<i>éh</i>		
អែ្គ	[ɑʰ]	[ũəʰ]	<i>aôh</i>	<i>ǫǎh</i>	The word នោះ "that" is pronounced [nuʰ].

The first four configurations listed here are treated as dependent vowels in their own right, and have names constructed in the same way as for the other dependent vowels (described in the previous section).

Other rarer configurations with the *reǎhmŭkh* are អើ្គ (or អិ្គ), pronounced [əh], and អែ្គ, pronounced [eh]. The word ចាំ៖ "yes" (used by women) is pronounced [ca:h].

The *bântăk* (a small vertical line written over the final consonant of a syllable) has the following effects:

- in a syllable with inherent *â*, the vowel is shortened to [ɑ], UN transcription *á*
- in a syllable with inherent *ô*, the vowel is modified to [u] before a final [labial](#), otherwise usually to [ǫə]; UN transcription *ó*
- in a syllable with the *a* dependent vowel symbol (  ) in the *a*-series, the vowel is shortened to [a], UN transcription *ǎ*
- in a syllable with that vowel symbol in the *o*-series, the vowel is modified to [ǫə], UN transcription *oǎ*, or to [ǽə] *eǎ* before *k*, *ng*, *h*

The *sanhyōk sannha* is equivalent to the *a* dependent vowel with the *bântāk*. However, its o-series pronunciation becomes [ɨ] before final *y*, and [ɔə] before final (silent) *r*.

The *yŭkôleăkpĭntŭ* (pair of dots) represents [a] (a-series) or [ĕə] (o-series), followed by a glottal stop.

### Consonants with no dependent vowel[\[edit\]](#)

There are three environments where a consonant may appear without a dependent vowel. The rules governing the inherent vowel differ for all three environments. Consonants may be written with no dependent vowel as an initial consonant of a [weak syllable](#), an initial consonant of a strong syllable or as the final letter of a written word.

In careful speech, initial consonants without a dependent vowel in weak initial syllables are pronounced with their inherent vowel shortened as if modified by the *bantak* diacritic (see previous section). For example the first-series letter "ច" in "ច្បុះ" ("torch") is pronounced with the short vowel /a/. The second-series letter "ត" in "ត្បឿង" ("light") is pronounced with the short diphthong /öə/. In casual speech, these are most often reduced to /ə/ for both series.

Initial consonants in strong syllables without written vowels are pronounced with their inherent vowels. The word ចង ("to tie") is pronounced /cɑːŋ/, ជក់ ("weak", "to sink") is pronounced /cɔːt/. In some words, however, the inherent vowel is pronounced in its reduced form, as if modified by a *bântāk* diacritic, even though the diacritic is not written (e.g. សព [sqp] "corpse"). Such reduction regularly takes place in words ending with a consonant with a silent subscript (such as សព្វ [sqp] "every"), although in most such words it is the *bântāk*-reduced form of the vowel *a* that is heard, as in សព្វ [sap] "noise". The word អ្នក "you, person" has the highly irregular pronunciation [nĕəʔ].

Consonants written as the final letter of a word usually represent a word-final sound and are pronounced without any following vowel and, in the case of stops, with [no audible release](#) as in the examples above. However, in some words adopted from [Pali](#) and [Sanskrit](#), what would appear to be a final consonant under normal rules can actually be the initial consonant of a following syllable and pronounced with a short vowel as if followed by ាត់. For example, according to rules for native Khmer words, សុត្តិ ("good",

"clean", "beautiful") would appear to be a single syllable, but, being derived from Pali *subha*, it is pronounced /soʔ pʰěəʔ/.

## Ligatures[\[edit\]](#)

Most consonants, including a few of the subscripts, form [ligatures](#) with the vowel *a* (ា) and with all other dependent vowels that contain the same cane-like symbol. Most of these ligatures are easily recognizable, but a few may not be, particularly those involving the letter ប *bâ*. This combines with the *a* vowel in the form បា, created to differentiate it from the consonant symbol ហ *hâ* and also from the ligature for ច *châ* with *a* (ចា).

Some more examples of ligatured symbols follow:

បៅ *bau* /baw/ Another example with ប *bâ*, forming a similar ligature to that described above. Here the vowel is not *a* itself, but another vowel (*au*) which contains the cane-like stroke of that vowel as a graphical element.

លា *léa* /liə/ An example of the vowel *a* forming a connection with the [serif](#) of a consonant.

ផ្កា *phba* /pʰbaː/ Subscript consonants with ascending strokes above the baseline also form ligatures with the *a* vowel symbol.

ម្សៅ *msau* /msaw/ Another example of a subscript consonant forming a ligature, this time with the vowel *au*.

ត្រា *tra* /traː/ The subscript for រ *rô* is written to the left of the main consonant, in this case ត *tâ*, which here forms a ligature with *a*.

## Independent vowels[\[edit\]](#)

Independent vowels are non-diacritical vowel characters that stand alone (i.e. without being attached to a consonant symbol). In Khmer they are called ស្រះពេញត្ថ *srăk pénhtuŏ*, which means "complete vowels". They are used in some words to represent certain combinations of a vowel with an initial [glottal stop](#) or [liquid](#). The independent vowels are used in a small number of words, mostly of Indic origin, and consequently there is some inconsistency in their use and pronunciations.<sup>[2]</sup> However, a few words in which they occur are used quite frequently; these include: ឥឡូវ [ʔəjləw] "now", ឪពុក [ʔəwpuk]



Independent vowel	IPA	UN
ඊ	[ʔə], [ʔɪ], [ʔe]	ě
ඹ	[ʔəj]	ei
උ	[ʔo], [ʔu], [ʔao]	ǫ, ŭ
ඉ	Obsolete (equivalent to the sequence ඉඟ්) <sup>[9]</sup>	
ඊ	[ʔou], [ʔu:]	not given ( <i>ou</i> in GD system)
ඊ	[ʔəw]	âu
ඊ	[ra~ru]	rǽ
ඊ	[ra:~ru:]	rœ
ඊ	[la~lu]	lǽ
ඊ	[la:~lu:]	lœ
ඊ	[ʔae], [ʔɛ:], [ʔe:]	ê
ඊ	[ʔaj]	ai
ඊ, ඊ	[ʔao]	aô
ඊ	[ʔaw]	au

Diacritics[[edit](#)]

16/03/2020 à 18:54

Diacritic	Khmer name	Function
◌ <sup>◌</sup>	និក្ខហិត <i>nīkkôhēt</i>	The Pali <i>niggahīta</i> , related to the <a href="#">anusvara</a> . A small circle written over a consonant or a following dependent vowel, it <a href="#">nasalizes</a> the inherent or dependent vowel, with the addition of [m]; long vowels are also shortened. For details see <a href="#">Modification by diacritics</a> .
◌ <sup>◌◌</sup>	រះមុខ <i>reăhmŭkh</i> "shining face"	Related to the <a href="#">visarga</a> . A pair of small circles written after a consonant or a following dependent vowel, it modifies and adds final <a href="#">aspiration</a> /h/ to the inherent or dependent vowel. For details see <a href="#">Modification by diacritics</a> .
◌ <sup>◌◌</sup>	យុគលពិន្ទុ <i>yŭkôleăkpîntŭ</i>	A "pair of dots", a fairly recently introduced diacritic, written after a consonant to indicate that it is to be followed by a short vowel and a glottal stop. See <a href="#">Modification by diacritics</a> .
◌ <sup>៊៊</sup>	មូសិកទន្ត <i>musĕkâtônd</i> "mouse teeth"	Two short vertical lines, written above a consonant, used to convert some o-series consonants (ង ញ ម យ រ វ) to a-series. It is also used with ប <i>bâ</i> to convert it to a <i>p</i> sound (see <a href="#">Supplementary consonants</a> ).
◌ <sup>〰</sup>	ត្រីសព្ទ <i>treisăpt</i>	A wavy line, written above a consonant, used to convert some a-series consonants (ស ហ ប ផ) to o-series.
◌ <sub>៊</sub>	ក្បៀសក្រោម <i>kbiĕh kraôm</i>	Also known as បុកជើង <i>bŏkcheung</i> ("collision foot"); a vertical line written under a consonant, used in place of the diacritics <i>treisăpt</i> and <i>musĕkâtônd</i> when they would be impeded by superscript vowels.
◌ <sup>៊</sup>	បន្តក់ <i>bântăk</i>	A small vertical line written over the last consonant of a syllable, indicating shortening

		(and corresponding change in quality) of certain vowels. See <a href="#">Modification by diacritics</a> .
◌̣	រតាទ <i>rôbat</i> រេផៈ <i>répheăk</i>	This superscript diacritic occurs in Sanskrit loanwords and corresponds to the <a href="#">Devanagari</a> diacritic <i>repha</i> . It originally represented an <i>r</i> sound (and is romanized as <i>r</i> in the UN system). Now, in most cases, the consonant above which it appears, and the diacritic itself, are unpronounced. Examples: ធម៌ /tʰɔː/ ("dharma"), កាណី /kaː/ (from <i>karṇa</i> ), ស្នេហា /suəkie ~ suəkie/ (" <a href="#">Svarga</a> ").
◌̤	ទណ្ឌាយាដ <i>tôndâkhéat</i>	Written over a final consonant to indicate that it is unpronounced. (Such unpronounced letters are still romanized in the UN system.)
◌̥	កាកបាទ <i>kakâbat</i>	Also known as a "crow's foot", used in writing to indicate the rising intonation of an exclamation or <a href="#">interjection</a> ; often placed on <a href="#">particles</a> such as /na/, /naː/, /nɛː/, /vəːj/, and on បាទ /caːh/, a word for "yes" used by females.
◌̦	អស្តា <i>âsda</i> "number eight"	Used in a few words to show that a <a href="#">consonant with no dependent vowel</a> is to be pronounced with its inherent vowel, rather than as a final consonant.
◌̧	សំយោគសញ្ញា <i>sanhyoŭk</i> <i>sannha</i>	Used in some Sanskrit and Pali loanwords (although alternative spellings usually exist); it is written above a consonant to indicate that the syllable contains a particular short vowel; see <a href="#">Modification by diacritics</a> .
◌̨	វិរាម <i>vîréam</i>	A mostly obsolete diacritic, corresponding to the <a href="#">virama</a> , which suppresses a consonant's inherent vowel.

## Dictionary order[\[edit\]](#)

For the purpose of [dictionary ordering](#)<sup>[11]</sup> of words, main consonants, subscript consonants and dependent vowels are all significant; and when they appear in combination, they are considered in the order in which they would be spoken (main consonant, subscript, vowel). The order of the [consonants](#) and of the [dependent vowels](#) is the order in which they appear in the above tables. A syllable written without any dependent vowel is treated as if it contained a vowel character that precedes all the visible dependent vowels.

As mentioned above, the four [configurations with diacritics](#) exemplified in the syllables អំ អំ អំ អំ are treated as dependent vowels in their own right, and come in that order at the end of the list of dependent vowels. Other configurations with the *reăhmŭkh* [diacritic](#) are ordered as if that diacritic were a final consonant coming after all other consonants. Words with the *bântăk* and *sanhyoŭk sannha* diacritics are ordered directly after identically spelled words without the diacritics.

Vowels precede consonants in the ordering, so a combination of main and subscript consonants comes after any instance in which the same main consonant appears unsubscripted before a vowel.

Words spelled with an [independent vowel](#) whose sound begins with a glottal stop follow after words spelled with the equivalent combination of អ ʼâ plus dependent vowel. Words spelled with an independent vowel whose sound begins [r] or [l] follow after all words beginning with the consonants រ rô and ល lô respectively.

Words spelled with a consonant modified by a diacritic follow words spelled with the same consonant and dependent vowel symbol but without the diacritic.<sup>[*dubious* – *discuss*][*citation needed*]</sup> However, words spelled with ប̌ (a *bâ* converted to a *p* sound by a diacritic) follow all words with unmodified ប *bâ* (without diacritic and without subscript).<sup>[*dubious* – *discuss*][*citation needed*]</sup> Sometimes words in which ប̌ is pronounced *p* are ordered as if the letter were written ប̌..

## Numerals<sup>[[edit](#)]</sup>

The numerals of the Khmer script, similar to that used by other civilizations in Southeast Asia, are also derived from the southern Indian script. Western-style [Arabic numerals](#) are also used, but to a lesser extent.

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

In large numbers, [groups of three digits](#) are delimited with Western-style [periods](#). The [decimal point](#) is represented by a comma. The Cambodian currency, the [riel](#), is abbreviated using the symbol ៛ or simply the letter ៛ *rô*.

## Spacing and punctuation[[edit](#)]

[Spaces](#) are not used between all words in written Khmer. Spaces are used within sentences in roughly the same places as [commas](#) might be in English, although they may also serve to set off certain items such as numbers and proper names.

Western-style [punctuation marks](#) are quite commonly used in modern Khmer writing, including French-style [guillemets](#) for [quotation marks](#). However, traditional Khmer punctuation marks are also used; some of these are described in the following table.

Mark	Khmer name	Function
។	ខណ្ឌ <i>khăṇ</i>	Used as a <a href="#">period</a> (the sign resembles an <a href="#">eighth rest</a> in music writing). However, consecutive sentences on the same theme are often separated only by spaces.
។ល។	ល្អៈ <i>lăk</i>	Equivalent to <a href="#">etc.</a>
២	លេខទោ <i>lékhtoŭ</i> ("figure two")	Duplication sign (similar in form to the <a href="#">Khmer numeral</a> for 2). It indicates that the preceding word or phrase is to be repeated ( <a href="#">duplicated</a> ), a common feature in Khmer syntax.
។្រ	បរិយោសាន <i>bâriyaôsan</i>	A period used to end an entire text or a chapter.
្រ	គោម្មត្រ <i>koŭmot</i> ("cow urine")	A period used at the end of poetic or religious texts.
្រ	ភ្នែកម៉ាន់ <i>phnêkmoăṅ</i>	A symbol (said to represent the elephant trunk of <a href="#">Ganesha</a> ) used at the start of poetic or religious

	("cock's eye")	texts.
៖	ចំណុចពីរគូស <i>châmnöch pi kus</i> "two dots (and a) line"	Used similarly to a <a href="#">colon</a> . (The middle line distinguishes this sign from a <a href="#">diacritic</a> .)

A [hyphen](#) (Khmer name សហសញ្ញា *sâhâ sânhnha*) is commonly used between components of personal names, and also as in English when a word is divided between lines of text. It can also be used between numbers to denote ranges or dates. Particular uses of Western-style periods include grouping of digits in large numbers (see [Numerals](#) hereinbefore) and denotation of [abbreviations](#).

## Styles[\[edit\]](#)

Several styles of Khmer writing are used for varying purposes. The two main styles are *âksâr chriêng* (literally "slanted script") and *âksâr mul* ("round script").



*Âksâr khâm* (អក្សរខម, *Aksar Khom*), an antique style of the Khmer script as written in Uttaradit, Thailand. In this picture, although it was written with Khmer script, all texts in this manuscript are in Thai languages.

- ***Âksâr chriêng*** (អក្សរជ្រៀង) refers to [oblique](#) letters. Entire bodies of text such as novels and other publications may be produced in *âksâr chriêng*. Unlike in [written English](#), oblique lettering does not represent any grammatical differences such as [emphasis](#) or quotation. Handwritten Khmer is often written in the oblique style.
- ***Âksâr chhôr*** (អក្សរឈរ) or ***Âksâr trâng*** (អក្សរត្រង់) refers to upright or 'standing' letters, as opposed to oblique letters. Most modern Khmer [typefaces](#) are designed in this manner instead of being oblique, as text can

be italicized by way of word processor commands and other computer applications to represent the oblique manner of *âksâr chriêng*.

- **Âksâr khâm** (អក្សរខ្មែរ) is a style used in Pali [palm-leaf manuscripts](#). It is characterized by sharper serifs and angles and retainment of some antique characteristics, notably in the consonant *kâ* (ក្ក). This style is also for [yantra tattoos](#) and [yantras](#) on cloth, paper, or engravings on brass plates in Cambodia as well as in Thailand.<sup>[12][13][14][15]</sup>
- **Âksâr mul** (អក្សរមូល) is [calligraphical](#) style similar to *âksâr khâm* as it also retains some characters reminiscent of antique Khmer script. Its name in Khmer means literally 'round script' and it refers to the bold and thick lettering style. It is used for titles and headings in Cambodian documents, on books, banknotes, shop signs and banners. It is sometimes used to emphasize royal names or other important names.

## Unicode[\[edit\]](#)

The basic [Khmer block](#) was added to the [Unicode](#) Standard in version 3.0, released in September 1999. It then contained 103 defined code points; this was extended to 114 in version 4.0, released in April 2003. Version 4.0 also introduced an additional block, called [Khmer Symbols](#), containing 32 signs used for writing [lunar dates](#).

The Unicode block for basic Khmer characters is U+1780–U+17FF:

<a href="#">Khmer</a> <sup>[1][2][3]</sup> <a href="#">Official Unicode Consortium code chart</a> (PDF)																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+178x	ក	ខ	គ	ឃ	ង	ច	ឆ	ជ	ឈ	ញ	ដ	ប	ឧ	ឝ	ណ	ត
U+179x	ថ	ទ	ធ	ន	ប	ផ	ព	ភ	ម	យ	រ	ល	វ	គ	ម	ស
U+17Ax	ហ	ឡ	អ	អ	អ	ក	ឡ	ខ	ខ	ខ	ខ	ប	ប	ព	ព	ឯ
U+17Bx	ញ	ឌ	ឍ	ឌ	KIV AQ	KIV AA	្ក	្ខ	្គ	្ឃ	្ង	្ច	្ឆ	្ជ	្ឈ	្ញ
U+17Cx	្ដ	្ឋ	្ឌ	្ឍ	្ណ	្ត	្ថ	្ទ	្ធ	្ន	្ប	្ផ	្ព	្ភ	្ម	្យ
U+17Dx	្រ	្ល	្វ	្ឝ	្ឞ	្ស	្ហ	្ឡ	្អ	្ឣ	្ឤ	្ឥ	្ឦ	្ឧ	្ឨ	្ឩ

U+17Ex	០	១	២	៣	៤	៥	៦	៧	៨	៩						
U+17Fx	០	១	២	៣	៤	៥	៦	៧	៨	៩						

**Notes**

1. ^ As of Unicode version 13.0
2. ^ Grey areas indicate non-assigned code points
3. ^ U+17A3 and U+17A4 are deprecated as of Unicode versions 4.0 and 5.2 respectively

The first 35 characters are the [consonant letters](#) (including two obsolete). The symbols at U+17A3 and U+17A4 are deprecated (they were intended for use in Pali and Sanskrit transliteration, but are identical in appearance to the consonant 𑀓, written alone or with the *a* vowel). These are followed by the 15 [independent vowels](#) (including one obsolete and one variant form). The code points U+17B4 and U+17B5 are invisible combining marks for inherent vowels, intended for use only in special applications. Next come the 16 [dependent vowel signs](#) and the 12 [diacritics](#) (excluding the *kbiəh kraôm*, which is identical in form to the *ö* dependent vowel); these are represented together with a dotted circle, but should be displayed appropriately in combination with a preceding Khmer letter.

The code point U+17D2, called ជើង *ceung*, meaning "foot", is used to indicate that a following consonant is to be written in subscript form. It is not normally visibly rendered as a character. U+17D3 was originally intended for use in writing lunar dates, but its use is now discouraged (see the Khmer Symbols block hereafter). The next seven characters are the [punctuation marks](#) listed hereinbefore; these are followed by the [riel](#) currency symbol, a rare sign corresponding to the Sanskrit [avagraha](#), and a mostly obsolete version of the *vīréam* diacritic. The U+17Ex series contains the [Khmer numerals](#), and the U+17Fx series contains variants of the numerals used in [divination](#) lore.

The block with additional lunar date symbols is U+19E0–U+19FF:

<a href="#">Khmer Symbols<sup>[1]</sup></a> <a href="#">Official Unicode Consortium code chart</a> (PDF)															
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	F



U+19Ex	០	១	២	៣	៤	៥	៦	៧	៨	៩	១០	១១	១២	១៣	១៤	១៥
U+19Fx	៤	៥	៦	៧	៨	៩	១០	១១	១២	១៣	១៤	១៥	១៦	១៧	១៨	១៩
<b>Notes</b>  1. <sup>^</sup> As of Unicode version 13.0																

The symbols at U+19E0 and U+19F0 represent the first and second "eighth month" in a lunar year containing a leap-month (see [Khmer calendar](#)). The remaining symbols in this block denote the days of a lunar month: those in the U+19Ex series for waxing days, and those in the U+19Fx series for waning days.

## See also<sup>[[edit](#)]</sup>

- [Khmer Braille](#)
- [Romanization of Khmer](#)

## Notes<sup>[[edit](#)]</sup>

- <sup>^</sup> [Herbert, Patricia; Anthony Crothers Milner \(1989\). South-East Asia: languages and literatures : a select guide. University of Hawaii Press. pp. 51–52. \[ISBN 0-8248-1267-0\]\(#\).](#)
- <sup>^</sup> [Jump up to: <sup>a</sup> <sup>b</sup> <sup>c</sup>](#) Huffman, Franklin. 1970. *Cambodian System of Writing and Beginning Reader*. Yale University Press. [ISBN 0-300-01314-0](#).
- <sup>^</sup> Punnee Soonthornpocht: *From Freedom to Hell: A History of Foreign Interventions in Cambodian Politics And Wars*. Page 29. Vantage Press.
- <sup>^</sup> Russell R. Ross: *Cambodia: A Country Study*. Page 112. Library of Congress, USA, Federal Research Division, 1990.
- <sup>^</sup> [Report on the Current Status of United Nations Romanization Systems for Geographical Names – Khmer](#), UNGEGN Working Group on Romanization Systems, September 2013 (linked from [WGRS website](#)).
- <sup>^</sup> [Jump up to: <sup>a</sup> <sup>b</sup>](#) *"Unicode 12.1 Character Code Charts - Khmer" (PDF)*.
- <sup>^</sup> The letter ្ក្រ *lâ* has no subscript form in standard orthography, but some fonts include one, as a form to be rendered if the character appears after the Khmer subscripting character (see under [Unicode](#)).

8. <sup>^</sup> [Jacob, Judith M. \(1968\). \*Introduction to Cambodian\*. Internet Archive. London ; Bombay \[etc.\] : Oxford University Press. pp. \[19\]\(#\), 29–30.](#)
9. <sup>^</sup> [Official Unicode Consortium code chart for Khmer](#) (PDF)
10. <sup>^</sup> Huffman (1970), p. 29.
11. <sup>^</sup> Different dictionaries use slightly different orderings; the system presented here is that used in the official *Cambodian Dictionary*, as described by Huffman (1970), p. 305.
12. <sup>^</sup> May, Angela Marie. (2014). *Sak Yant: The Transition from Indic Yantras to Thai Magical Buddhist Tattoos (Master's thesis)* (p. 6). The University of Alabama at Birmingham.
13. <sup>^</sup> Igunma, Jana. (2013). Aksoon Khoom: Khmer Heritage in Thai and Lao Manuscript Cultures. *Tai Culture, 23: Route of the Roots: Tai-Asiatic Cultural Interaction*.
14. <sup>^</sup> Tsumura, Fumihiko. (2009). Magical Use of Traditional Scripts in Northeastern Thai Villages. *Senri Ethnological Studies*, 74, 63-77.
15. <sup>^</sup> This particular style of Khmer shall not be confused with another script with the same name, described by [Paul Sidwell](#) (see [Khom script](#)).

## References<sup>[edit](#)</sup>

- *Dictionnaire Cambodgien*, Vol I & II, 1967, L'institut Bouddhique (Khmer Language)
- Jacob, Judith. 1974. *A Concise Cambodian-English Dictionary*. London, Oxford University Press.

## External links<sup>[edit](#)</sup>

- [FAQ and Resources on Khmer in Unicode](#)
- [Enabling Khmer Unicode](#)
- [Khmer Unicode in some mobile phones](#)
- [Khmer Consonant Chart with Audio](#)
- [Khmer Vowel Chart with Audio](#)
- [How to Install Khmer Unicode on your Windows 7 Computer](#)

- [How to Install Khmer Unicode on your Windows XP Computer](#)
- [Omniglot entry on Khmer](#)
- [Khmer Romanization Table](#) (PDF)
- [Evolution of the Khmer script](#)
- [Authentic Khmer Online](#) (common phrases in Khmer script with audio file examples)
- [Khmer wordlist sortet frequenzy](#)
- [CBC radio documentary referring to development of keyboard for Khmer script](#)
- [A small Primer on the Khmer Language](#)
- [A Khmer Language Primer](#)
- <http://unicode-table.com/en/sections/khmer/>
- <http://unicode-table.com/en/sections/khmer-symbols/>
- <http://www.angkorone.com/installkhmerfont.aspx>