

Khmer Angkor Keyboard

Contents

Overview	2
Khmer Angkor Keyboard Layouts	2
Desktop Layout –Windows/macOS	2
Touch Layout –Android/iOS.....	3
Khmer Character Categories and Keystrokes for Desktop.....	5
Consonants	5
Dependent Vowels.....	5
Signs	6
Independent Vowels	6
Diacritics.....	6
Native Punctuations.....	6
Latin Punctuations	7
Currencies	7
Khmer Numerals	7
Divination Lore.....	7
Khmer Lunar Dates.....	7
Discouraged/Obsolete characters	8
Characters used in minority languages.....	8
Joiners and Spaces	9
Typing in Khmer	9
Silent Auto-correction.....	9
Subscript + Vowel.....	9
Subscript + ្ក [U+17D2] ្ខ [U+179A]	9
Subscript + Consonant Shifter + Vowel.....	10
Consonant Shifter + Vowel.....	10
Vowel with Nikahit [U+17C6]	10
Vowel ្ក [U+17C4] and Vowel ្ខ [U+17BE]	10
Visually Identical Subscript	10
Vowel Rotation	10
Beep / Constraints	11
Mistyped Characters	11

Overview

This documentation is adapted from three sources: ‘How to Type Khmer Unicode,’ ‘the Unicode Standard, Version 10.0.0’ and a paper authored by the Keyman team entitled ‘Spoof-Vulnerable Rendering in Khmer Unicode Implementation’ presented at Asia Conference on Information System in 2017. Customizations have been made by the Keyman team. A layer is added to both desktop and touch layout in order to dedicate to the Khmer divination lore and Khmer lunar dates.

Khmer Angkor Keyboard Layouts

Desktop Layout –Windows/macOS

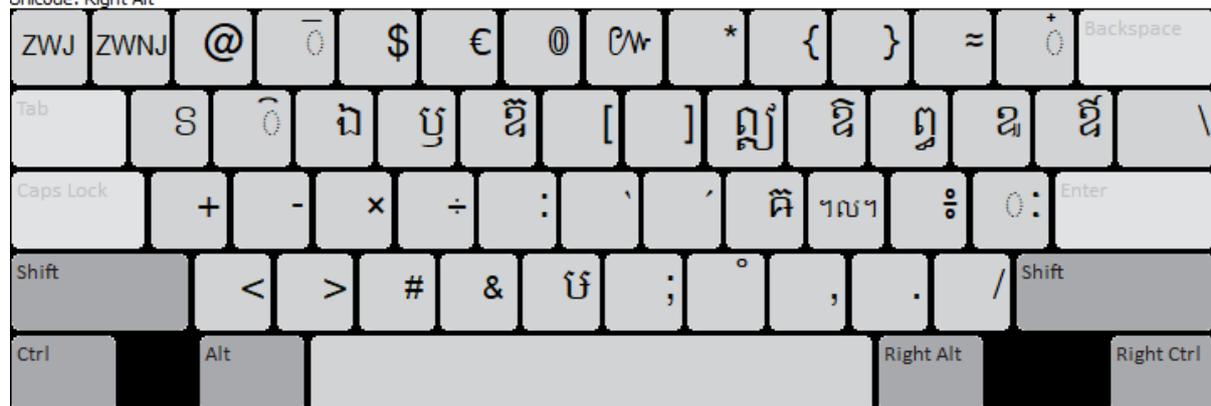
Unicode: Unshifted



Unicode: Shift



Unicode: Right Alt





Touch Layout –Android/iOS

There is no Shift key nor subscript sign in the touch layout to switch between layouts. Strange, right? Yes, maybe, but it is not easy to reach for the “Shift” key one-handedly. More keys are embedded in each key—press and hold for a second or so to see more available keys. For example, to get subscript “្រ”, “្រ” or “្រ”, press and hold on key “រ” for a short while, then more keys would appear, and then you may choose “្រ”. You may notice that on the Khmer Angkor keyboard for desktop, these four characters are on key “R”; which is why they are all packed into this one key hidden in the long press.

While using the touch keyboard, one has to be aware of the location of each key. This may take a while to get used to, but it would be easier, once passed the learning curve, than pressing “Shift” each time to get addition characters and pressing “្រ” each time to get a subscript.

Phone - Default layout with longpress triggered on រ key



Phone - Numeric layout



Tablet - Default layout



Tablet - Shift layout



Tablet - AltGr layout



Khmer Character Categories and Keystrokes for Desktop

Consonants

Consonant	Keystroke	Consonant	Keystroke
ក [U+1780]	K	ផ [U+1790]	F
ខ [U+1781]	X	ឆ [U+1791]	Shift T
គ [U+1782]	Shift K	ធុ [U+1792]	Shift F
ឃ [U+1783]	Shift X	ន [U+1793]	N
ង [U+1784]	G	ប [U+1794]	B
ច [U+1785]	C	ផ្អ [U+1795]	P
ឆ [U+1786]	Q	ព [U+1796]	Shift B
ជ [U+1787]	Shift C	ភ [U+1797]	Shift P
ឈ [U+1788]	Shift Q	ម [U+1798]	M
ញ [U+1789]	Shift J	យ [U+1799]	Y
ដ [U+178A]	D	រ [U+179A]	R
ប៉ [U+178B]	Z	ល [U+179B]	L
ឌ [U+178C]	Shift D	វ [U+179C]	V
ឍ [U+178D]	Shift Z	ស [U+179F]	S
ណ [U+178E]	Shift N	ហ [U+17A0]	H
ត [U+178F]	T	ឡ [U+17A1]	Shift L
		អ [U+17A2]	Shift G

Dependent Vowels

Dependent Vowel	Keystroke	Dependent Vowel	Keystroke
ា [U+17B6]	A	ៀ [U+17C0]	[
ិ [U+17B7]	I	ៃ [U+17C1]	E
ឹ [U+17B8]	Shift I	ឺ [U+17C2]	Shift E
ី [U+17B9]	W	ឺ [U+17C3]	Shift S
ឺ [U+17BA]	Shift W	ោ [U+17C4]	O
ុ [U+17BB]	U	ៅ [U+17C5]	Shift O
ួ [U+17BC]	Shift U	ុំ [U+17BB] [U+17C6]	,
ូ [U+17BD]	Shift Y	ាំ [U+17B6] [U+17C6]	Shift A
្ល [U+17BE]	;	ុំ៖ [U+17BB] [U+17C7]	Shift ,

ᐃᐃᐃ [U+17BF]	Shift [ᐃᐃᐃᐃ [U+17C1] [U+17C7]	Shift V
		ᐃᐃᐃᐃᐃ [U+17C4] [U+17C7]	Shift ;

Signs

Sign	Keystroke
ᐃᐃ [U+17C6]	Shift M
ᐃᐃᐃ [U+17C7]	Shift H
ᐃᐃᐃᐃ [U+17C8] ¹	AltGr ’

Independent Vowels

Independent Vowel	Keystroke	Independent Vowel	Keystroke
ᐃᐃ [U+17A5]	-	ᐃᐃ [U+17AD]	Shift \
ᐃᐃᐃ [U+17A6]	AltGr I	ᐃᐃ [U+17AE]	\
ᐃᐃ [U+17A7]	Shift]	ᐃᐃ [U+17AF]	AltGr E
ᐃᐃ [U+17A9]	AltGr [ᐃᐃ [U+17B0]	AltGr P
ᐃᐃᐃ [U+17AA]]	ᐃᐃ [U+17B1]	AltGr O
ᐃᐃ [U+17AB]	AltGr R	ᐃᐃ [U+17B2]	=
ᐃᐃ [U+17AC]	Shift R	ᐃᐃ [U+17B3]	AltGr]

Diacritics

Diacritics	Keystroke	Diacritics	Keystroke
ᐃᐃ [U+17C9]	Shift ’	ᐃᐃ [U+17CF]	Shift 8
ᐃᐃ [U+17CA]	/	ᐃᐃ [U+17D0]	Shift 7
ᐃᐃ [U+17CB]	’	ᐃᐃ [U+17CE]	AltGr =
ᐃᐃ [U+17CC]	Shift -	ᐃᐃ [U+17D3]	AltGr M
ᐃᐃ [U+17CD]	Shift 6	ᐃᐃ [U+17D2]	J

Native Punctuations

Native Punct.	Keystroke	Native Punct.	Keystroke
ᐃ [U+17D4]	.	ᐃ [U+17D9]	AltGr 6
ᐃᐃ [U+17D5]	Shift .	ᐃᐃ [U+17DA]	AltGr 7
ᐃᐃ [U+17D6]	AltGr ;	ᐃᐃ [U+17D7]	Shift 2

¹ It is actually not a vowel, but it behaves like one. To some, this is a diacritics.

Latin Punctuations

Latin Punct.	Keystroke	Latin Punct.	Keystroke
! [U+0021]	Shift 1	« [U+00AB]	`
? [U+003F]	Shift ?	» [U+00BB]	Shift `
“ [U+0022]	Shift 3	- [U+002D]	AltGr S
. [U+002E]	AltGr .	: [U+003A]	AltGr G
, [U+002C]	AltGr ,	‘ [U+2018]	AltGr H
; [U+003B]	AltGr N	’ [U+2019]	AltGr J
# [U+0023]	AltGr C	& [U+0026]	AltGr V

Currencies

₪ [U+17DB]	Shift 4
------------	---------

Khmer Numerals

Khmer Numeral	Keystroke	Khmer Numeral	Keystroke
១ [U+17E1]	1	៦ [U+17E6]	6
២ [U+17E2]	2	៧ [U+17E7]	7
៣ [U+17E3]	3	៨ [U+17E8]	8
៤ [U+17E4]	4	៩ [U+17E9]	9
៥ [U+17E5]	5	០ [U+17E0]	0

Divination Lore

Divination Lore	Keystroke	Divination Lore	Keystroke
^ [U+17F1]	AltGr Shift 1	˘ [U+17F6]	AltGr Shift 6
ˆ [U+17F2]	AltGr Shift 2	˜ [U+17F7]	AltGr Shift 7
˘ [U+17F3]	AltGr Shift 3	/ [U+17F8]	AltGr Shift 8
˙ [U+17F4]	AltGr Shift 4	˚ [U+17F9]	AltGr Shift 9
˚ [U+17F5]	AltGr Shift 5	◦ [U+17F0]	AltGr Shift 0

Khmer Lunar Dates

Khmer Lunar Date	Keystroke	Khmer Lunar Date	Keystroke
០ [U+19E0]	AltGr Shift Q	៤ [U+19F0]	AltGr Shift G
១ [U+19E1]	AltGr Shift W	៥ [U+19F1]	AltGr Shift H
២ [U+19E2]	AltGr Shift E	៦ [U+19F2]	AltGr Shift J

ၵ [U+19E3]	AltGr Shift R	ၵ [U+19F3]	AltGr Shift K
ၶ [U+19E4]	AltGr Shift T	ၶ [U+19F4]	AltGr Shift L
ၷ [U+19E5]	AltGr Shift Y	ၷ [U+19F5]	AltGr Shift ;
ၸ [U+19E6]	AltGr Shift U	ၸ [U+19F6]	AltGr Shift '
ၹ [U+19E7]	AltGr Shift I	ၹ [U+19F7]	AltGr Shift Z
ၺ [U+19E8]	AltGr Shift O	ၺ [U+19F8]	AltGr Shift X
ၻ [U+19E9]	AltGr Shift P	ၻ [U+19F9]	AltGr Shift C
ၼ [U+19EA]	AltGr Shift [ၼ [U+19FA]	AltGr Shift V
ၽ [U+19EB]	AltGr Shift]	ၽ [U+19FB]	AltGr Shift B
ၾ [U+19EC]	AltGr Shift A	ၾ [U+19FC]	AltGr Shift N
ၿ [U+19ED]	AltGr Shift S	ၿ [U+19FD]	AltGr Shift M
ၠ [U+19EE]	AltGr Shift D	ၠ [U+19FE]	AltGr Shift ,
ၡ [U+19EF]	AltGr Shift F	ၡ [U+19FF]	AltGr Shift .

Discouraged/Obsolete characters

Character	Keystroke
ၵ [U+17D8]	AltGr L
ၶ [U+17A8]	AltGr T
ၷ [U+17DC]	AltGr Q
ၸ ² [U+17D1]	AltGr 3
ၹ [U+17D3]	AltGr M

Characters used in minority languages

Other Character	Keystroke
ၵ [U+179D]	AltGr K
ၶ [U+179E]	AltGr B
ၷ [U+17DD]	AltGr W

² Note the different between this character and ၸ [U+17CA] .

Joiners and Spaces

Key name	Keystroke
Zero-width non-joiner [U+200C]	AltGr + 1
Zero-width joiner [U+200D]	AltGr + ‘
Zero-width space [U+200B]	Spacebar
White space [U+0020]	Shift + Spacebar or Press spacebar twice
No break space [U+00A0]	AltGr + Spacebar

Typing in Khmer

Khmer Angkor keyboard adopted the key arrangement from that of Khmer NiDA keyboard. More functionalities have been added to it to assist in typing (i.e. reduce inconsistency in character sequences). Regardless of how the user may type, a well-formed word will be produced. Please note that only issues identified are covered. For how to install the Khmer Angkor keyboard, please click [here](#) for the instructions.

You can simply use Khmer Angkor as you would for any NiDA based keyboard. You may type a word in an order of how it is spelled, not how it appears to be, especially when the vowel is to the left of the consonant. For example, to type the word “តែ” which means ‘tea’, one should type the consonant first and then the vowel (i.e. press key “T” and then “Shift E”).

The order of characters is:

Consonant + Subscript(s) + Consonant Shifter + Vowel + Diacritic

Silent Auto-correction

There are eight cases of invisible typos that Khmer Angkor keyboard automatically corrects. They could not be detected with the naked eye; and the rendering engines may render them all the same on the screen. The auto-correction involves reordering and/or transducing of certain in correct characters to the correct ones in a split second.

Subscript + Vowel

When a subscript and a vowel occur together, the vowel has to come after the subscript. For instance, “ខ្មែរ” and “ខ្មែរ” look the same, but they are encoded differently.

- ខ្មែរ = ខ ្ម ែ រ (correct)
- ខ្មែរ = ខ ែ ្ម រ (incorrect)

Thanks to the context dependent rules, Khmer Angkor keyboard is able to output the second typing sequence the same as that of the first one.

Subscript + ្រ [U+17D2] ្រ [U+179A]

According to the spelling rule, in a sequence of two subscript and one of which is ្រ [U+17D2] [U+179A], the ្រ [U+17D2] [U+179A] should come after another subscript.

- កញ្ជីល = ក ញ ្រ ្រ ី ល (correct)
- កញ្ជីល = ក ញ ្រ ្រ ី ល (incorrect)

As you can see the two-character sequences yield the same visual output on the screen. Once again, Khmer Angkor will be able to automatically swap the sequence of the second to the first; so that the output would be both looked and encoded correctly and consistently.

Subscript + Consonant Shifter + Vowel

This is a case whereby the character sequence is vague in the Unicode Standard. Unicode Standard 4.0 and later have the consonant shifter come between the consonant and the subscript while Unicode Standard 3.0 has it after the subscript. Khmer Angkor conforms to Unicode 3.0 and it is taken as the standard sequence. If one typed a consonant shifter before a subscript, the order would be switched silently.

ម ័ ្យ ា ង ⇒ ម ្យ ័ ា ង

Consonant Shifter + Vowel

Consonant shifters should be typed before a vowel. (See the same in the section above)

Vowel with Nikahit [U+17C6]

Nikahit is considered as a sign in the Unicode Standard. However its function is more like that of a vowel. It is usually combined with another vowel, i.e. ្ក [U+17BB] ា [U+17B6], and it should always be typed after these vowels.

ក ័ ្ក ⇒ ក ្ក ័
ក ័ ា ⇒ ក ា ័

Vowel ៃា [U+17C4] and Vowel ើ [U+17BE]

These two vowels could be mistakenly typed by adding ា [U+17B6] and ៃ [U+17C1] respectively to ៃ [U+17C1]. Given this sequence, Khmer Angkor would transform them to ៃា [U+17C4] and ើ [U+17BE], the correct vowel encodings.

Visually Identical Subscript

Two subscripts look exactly the same, but they are encoded with different Unicode code points, ្ក [U+17D2] [U+178A] and ្ខ [U+17D2] [U+178F]. According to the spelling rule, it governs that the first should come after ឃ [U+178E] and the latter should come after ៩ [U+1793].

Vowel Rotation

Since no two-vowel sequence is allowed, once one is typed after another, the vowel would switch to the one typed last. For example, if the current character is vowel ា [U+17B6], and then ៃ [U+17C2] is typed, the final output would be ៃ [U+17C2], not ា [U+17B6].

NOTE: The vowel rotation would not be triggered in the following cases:

- If the first vowel is ៃ [U+17C1] and the second is ើ [U+17B8] or vice versa, then the switch would not happen, but the output would be ើ [U+17BE] instead.
- If the first vowel is ៃ [U+17C1] and the second is ៃា [U+17B8] or vice versa, then the switch would not happen, but the output would be ៃា [U+17C4] instead.

- If the first vowel is on the above vowels³ and the second is ្ក [U+17BB] or vice versa, then the switch would not happen, but the output would be the corresponding consonant shifter instead. For examples: a sequence of ស ី ្ក or ស ្ក ី would be automated as សី (ស ិ ី); ម ី ្ក or ម ្ក ី would be automated as មី (ម ិ ី).
- When ្ក [U+17BB] is typed after ្ក, it will be transformed to a consonant shifter corresponding with the consonant before ្ក before which the shifter will be placed. For example,
 - សាំ + ្ក ⇒ សាំ (ស ិ ្ក) and
 - មាំ + ្ក ⇒ មាំ (ម ិ ្ក).
- When ្ក or ្ក is typed after ្ក, the order will be reversed. For example,
 - ក + ្ក + ្ក ⇒ ក្ក (ក ្ក ្ក) and
 - ក + ្ក + ្ក ⇒ ក្ក (ក ្ក ្ក).

Beep / Constraints

- No two consecutive vowels in a syllable.
- No Bantoc is allowed after a vowel, a pseudo vowel⁴, a consonant shifter or a subscript.
- No two consecutive subscript signs (្ក) are allowed.
- No two consecutive consonant shifters are allowed.

Mistyped Characters

No	Mistyped as ...	Transformed to ...
1	ឡ (ឡ ្ក)	ឡ
2	ប្ក (ប ្ក)	ប្ក
3	ប្ក (ប ្ក ្ក)	ប្ក
4	ញ (ញ ្ក)	ញ
5	ញ (ញ ្ក)	ញ
6	ញ (ញ ្ក)	ញ
7	ញ (ញ ្ក ្ក)	ញ
8	ញ (ញ ្ក)	ញ
9	ឌី (ឌ ិ ី)	ឌី
10	ឌី (ឌ ិ ី)	ឌី
11	ឌី (ឌ ិ ី)	ឌី
12	ដ្ក (ដ ្ក)	ដ្ក
13	ឡ (ឡ ្ក)	ឡ

³ ិ [U+17B7], ី [U+17B8], ឹ [U+17B9], ឺ [U+17BA], ើ [U+17BE], and ្ក [U+17B6] [U+17C6] are considered as superscript vowels or above vowels.

⁴ Three characters play a role similar to that of vowels, thus they are called 'pseudo vowel'. They are: ្ក [U+17C6], ្ក [U+17C7], and ្ក [U+17C8].

14	ពន្ធ (ព ន ្ន)	ពន្ធ
15	ព័ន្ធ (ព ័ ន ្ន)	ព័ន្ធ
16	ឡី (ឡ ី)	ឡី
17	ឡី (ឡ ី)	ឡី
18	ញ (ញ ូ)	ញ (ឧទា. សព្វានុវ)
19	ច្រើន	ច្រើន
20	ច្រើន	ច្រើន
21	ច្រើន	ច្រើន