Khmer Angkor Keyboard

Contents
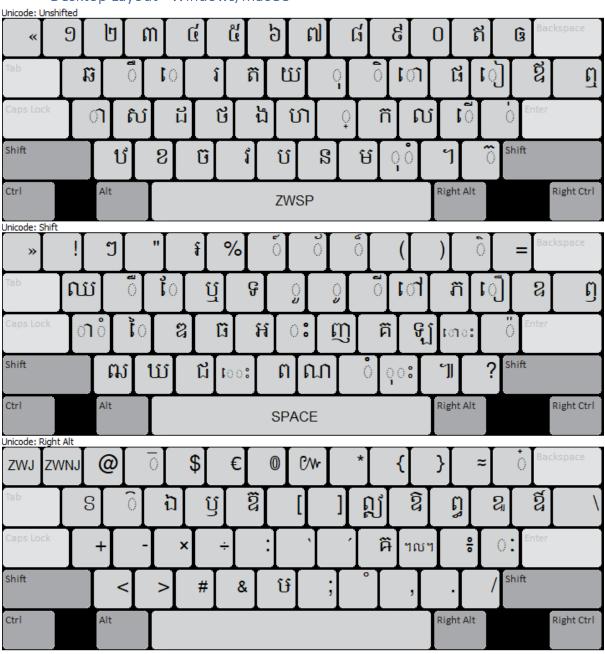
Overview

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

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", "ឫ" or "ឬ", press and hold on key "ិ". 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 ្ each time to get a subscript.

Default layout

Numeric layout

Divination Lore and Khmer Lunar Dates layout

Khmer Character Categories and Keystrokes for Desktop

Consonants

Consonant	Keystroke	Consonant	Keystroke
ក [U+1780]	K	ថ [U+1790]	F
ខ [U+1781]	X	ទ [U+1791]	Shift T
គ [U+1782]	Shift K	ធ [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	ផ [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
ឋ [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]	L
		អ [U+17A2]	Shift G

Dependent Vowels

Dependent Vowel	Keystroke
ា [U+17B6]	A
ិ [U+17B7]	I
ី [U+17B8]	Shift I
ឹ [U+17B9]	W
ឺ [U+17BA]	Shift W
ុ [U+17BB]	O
ូ [U+17BC]	Shift O
ួ [U+17BD]	Shift Y
ើ [U+17BE]	;
ឿ [U+17BF]	Shift [
ៀ [U+17C0]	[
េ [U+17C1]	E

Dependent Vowel	Keystroke
ែ [U+17C2]	Shift E
ៃ [U+17C3]	Shift S
ោ [U+17C4]	O
ៅ [U+17C5]	Shift O
ុំ [U+17BB][U+17C6]	,
ំ [U+17C6]	Shift M
ាំ [U+17B6][U+17C6]	Shift A
ះ [U+17C7]	Shift H
ុះ [U+17BB][U+17C7]	Shift ,
េះ [U+17C1][U+17C7]	Shift V
ោះ [U+17C4][U+17C7]	Shift ;
ៈ [U+17C8][1]	AltGr '

Independent Vowels

Independent Vowel	Keystroke
ឥ [U+17A5]	-
ឦ [U+17A6]	AltGr I
ឧ [U+17A7]	Shift]
ឩ [U+17A9]	AltGr [
ឪ [U+17AA]	]
ឫ [U+17AB]	AltGr R
ឬ [U+17AC]	Shift R

Independent Vowel	Keystroke
ឭ [U+17AD]	Shift \
ឮ [U+17AE]	\
ឯ [U+17AF]	AltGr E
ឰ [U+17B0]	AltGr P
ឱ [U+17B1]	AltGr O
ឲ [U+17B2]	=
ឳ [U+17B3]	AltGr]

Diacritics

Diacritics	Keystroke
៉ [U+17C9]	Shift '
៊ [U+17CA]	/
់ [U+17CB]	'
៌ [U+17CC]	Shift -
៍ [U+17CD]	Shift 6

Diacritics	Keystroke
៏ [U+17CF]	Shift 8
័ [U+17D0]	Shift 7
៎ [U+17CE]	AltGr =
៓ [U+17D3]	AltGr M
្ [U+17D2]	J

---

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

Native Punctuations

Native Punct.	Keystroke	Native Punct.	Keystroke
�។ [U+17D4]	.	៕៙៕ [U+17D8]	AltGr L
៕ [U+17D5]	Shift .	៙ [U+17D9]	AltGr 6
៖ [U+17D6]	AltGr ;	៚ [U+17DA]	AltGr 7
ៗ [U+17D7]	Shift 2		

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
៱ [U+17F1]	AltGr Shift 1	៶ [U+17F6]	AltGr Shift 6
៲ [U+17F2]	AltGr Shift 2	៷ [U+17F7]	AltGr Shift 7
៳ [U+17F3]	AltGr Shift 3	៸ [U+17F8]	AltGr Shift 8
៴ [U+17F4]	AltGr Shift 4	៹ [U+17F9]	AltGr Shift 9
៵ [U+17F5]	AltGr Shift 5	៰ [U+17F0]	AltGr Shift 0

Khmer Lunar Dates

Khmer Lunar Date	Keystroke	Khmer Lunar Date	Keystroke
᧠ [U+19E0]	AltGr Shift Q	᧰ [U+19F0]	AltGr Shift G
᧡ [U+19E1]	AltGr Shift W	᧱ [U+19F1]	AltGr Shift H
᧢ [U+19E2]	AltGr Shift E	᧲ [U+19F2]	AltGr Shift J
᧣ [U+19E3]	AltGr Shift R	᧳ [U+19F3]	AltGr Shift K
᧤ [U+19E4]	AltGr Shift T	᧴ [U+19F4]	AltGr Shift L
᧥ [U+19E5]	AltGr Shift Y	᧵ [U+19F5]	AltGr Shift ;
᧦ [U+19E6]	AltGr Shift U	᧶ [U+19F6]	AltGr Shift '
᧧ [U+19E7]	AltGr Shift I	᧷ [U+19F7]	AltGr Shift Z
᧨ [U+19E8]	AltGr Shift O	᧸ [U+19F8]	AltGr Shift X
᧩ [U+19E9]	AltGr Shift P	᧹ [U+19F9]	AltGr Shift C
᧪ [U+19EA]	AltGr Shift [	᧺ [U+19FA]	AltGr Shift V
᧫ [U+19EB]	AltGr Shift]	᧻ [U+19FB]	AltGr Shift B
᧬ [U+19EC]	AltGr Shift A	᧼ [U+19FC]	AltGr Shift N
᧭ [U+19ED]	AltGr Shift S	᧽ [U+19FD]	AltGr Shift M
᧮ [U+19EE]	AltGr Shift D	᧾ [U+19FE]	AltGr Shift ,
᧯ [U+19EF]	AltGr Shift F	᧿ [U+19FF]	AltGr Shift .

Deprecated characters

Deprecated Character	Keystroke
ឝ [U+179D]	AltGr K
ឞ [U+179E]	AltGr B
ឨ [U+17A8]	AltGr T
ᯬ [U+17DC]	AltGr Q
៝ [U+17DD]	AltGr W
៑² [U+17D1]	AltGr 3
៓ [U+17D3]	AltGr M

---

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

Joiners and Spaces

Key name	Keystroke
Zero-width non-joiner [U+200C]	Ctrl + Alt + 1
Zero-width joiner [U+200D]	Ctrl + Alt + '
Zero-width space [U+200B]	Spacebar
White space [U+0020]	Shift + Spacebar
No break space [U+00A0]	Ctrl + Alt + 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.

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 characters.

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] [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.

- កញ្ច្រឹល = ក ញ ្ច ្រ ឹ ល   (correct)
- កញ្ច្រឹល = ក ញ ្រ ្ច ឹ 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 agrees with the later 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. ុ [U+17BB] ា [U+17B6], and it should always be typed after these vowels.

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, ្ដ [U+17D2][U+178A] and ្ដ [U+17D2][U+178F]. Since a spelling rule governs that the first should come after ណ [U+178E] and the later should come after ន [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], 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 vowels[3] and the second is ◌ុ [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 ើ ុ or ស ុ ើ would be automated as ស៊ើ (ស ៊ ើ); ម ើ ុ or ម ុ ើ would be automated as ម៉ើ (ម ៉ ើ).

Beep / Constraints

- No two consecutive vowels in a syllable.
- No Bantoc is allowed after a vowel, a pseudo vowel[4], a consonant shifter or a subscript.
- No two subscript signs (◌្) are allowed.

Mistyped Characters

No	Mistyped as …	Transformed to …
1	ទ្យ (ទ ្យ)	ទ្យ
2	ប្ុ (ប ្ុ)	ប៊ុ
3	ប្ុ (ប ្ុ ុ)	ប៊ុ
4	ញ (ញ ា)	ញ
5	ញ (ញ ា)	ញ
6	ព្ (ព ្ុ)	ព្
7	ព្ (ព ្ុ ុ)	ព្
8	ព្ (ព ្ុ)	ព្
9	ឱិ (ឱ ិ)	ឱ
10	ឱ៊ី (ឱ ៊ី)	ឱ
11	ឱ៎ (ឱ ៎)	ឱ
12	ៗល៕ (៕ ល ៕)	ៗល៕
13	ដ្ឋ (ដ ្ឋ)	ដ្ឋ
14	ទ្ធ (ទ ្ធ)	ទ្ធ

---

[3] ◌ិ [U+17B7], ◌ី [U+17B8], ◌ឹ [U+17B9], ◌ឺ [U+17BA], ើ◌ [U+17BE], and ឿ◌ [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: ◌ំ [U+17C6], ◌ះ [U+17C7], and ◌ៈ [U+17C8].