SOASMyanmar Burmese and Shan TrueType Font and Keyboard Package

Version 0.94

Copyright © 2004, 2005 SOAS, U of London
For late-breaking info or updates, visit:
http://mercury.soas.ac.uk/wadict/resources.html#fonts
or
http://mercury.soas.ac.uk/private/wadict/misc/working_docs.html

INTRODUCTION

SOASMyanmar is a transitional Unicode font for Burmese and Shan with several accompanying alternate keyboards . The first keyboard developed was a Keyman keyboard based on the "ICMyanmar" font and phonetic keyboard arrangement of Ian Carter. This was in turn an adaptation of the AvaLaser font and keyboard designed by John Okell, to the PC using Microsoft Word macros. But the SOASMyanmar Keyman Burmese keyboard does not have the limitations inherent in Word macros. It may be used in any application running under any version of Windows since Windows 95. The keyboard may be used in conjunction with Keyman 6.0 from Tavultesoft for typing Burmese in a Unicode-compatible, (but not yet fully Unicode-conformant) way. Both the font and the keyboards are meant to be a transitional system for moving toward handling the Burmese language on the computer encoded as actual Burmese characters, not encoded using ASCII a-b-c characters, which are made to look like Burmese by the character pictures in a font. Most characters are encoded with their normal Burmese Unicode encoding. But there are two deviations from "pure" Unicode. First, subscript and other contextual glyphic variations of characters are encoded using codepoints in the Unicode Private User Area; and second, text is "misspelled," in the sense that several characters are out of place in the text stream, especially vowel sign 'e' (U+1031) and subscript 'yá-yiq' (U+101B).

Since version 0.91, ca. September 2004, the SOASMyanmar font also contains draft glyphs needed for typing the Shan (a.k.a. Tai Luang, Tai Long, Tai Yai, etc.) language and these can be typed using the draft "SOASShan" Keyman keyboard. Since Shan is not yet encoded in Unicode, the Unicode Private User Area is used for unique Shan characters.

The font and the keyboards can used separately from each other for displaying or typing Burmese in pure Unicode, but to employ the special transitional features, which enable Burmese to look correct without underlying operating system support, they should be used together.

The font-keyboard package has the following features.

FONT

1) The font retains the glyph outlines of ICMyanmar and AvaLaser as much as possible (although in addition I have increased the size and weight of the glyphs, so that, for example, a 12-point font generated from the "SOASMyanmar" font is closer in size to 12-point for other common fonts, such as the Windows core fonts, Times New Roman, etc.--this is work in progress, with ongoing tuning of the metrics of the font, since Ian's glyphs were designed to mesh with each other at his original sizes; also, I have attempted to increase the weight (stroke thickness) of the font, to make it more suitable as a screen font, so that it will not tend to disappear or turn gray at small point sizes on screen, especially where Windows ClearType technology is in operation.

2) For each character in the Burmese alphabet, the font maps a common or "citation" glyph for that character to a Unicode codepoint in the Unicode Myanmar range U+1000-109F (about 77 different characters in all; all the other 80+ glyphic variants in the font are mapped to codepoints in the Unicode Private User Area (PUA), specifically to codepoints in the range EC00-EC6F, thus making them available for use with the SOASMyanmar keyboard, in a Web browser, or another application, without violating Unicode guidelines, or producing mismatched garbage characters.

3) A few general punctuation characters from ICMyanmar are also retained, and given their correct Unicode values. There are no glyphs in the font for Latin alphanumeric characters. This has the disadvantage that the font name does not show up in applications which only use the font itself to display its name, such as the font menu in certain versions of Microsoft Word.

4) Since version 0.91, ca. September 2004, the SOASMyanmar font package also contains glyphs needed for typing the Shan (a.k.a. Tai Luang) language and these can be typed using the new "SOASShan" Keyman keyboard. Since Shan is not yet encoded in Unicode, the Unicode Private User Area (PUA) is used for special Shan glyphs.

KEYBOARD

4) The SOASMyanmar Keyman keyboard can be used to type and display Burmese in any Windows application which supports Unicode (e.g., MS Office, Notepad, Internet Explorer text boxes, Netscape Composer, etc.), in any version of Windows since Windows 98. The basic keyboard keys are remapped to the Burmese alphabet, retaining virtually completely the phonetic-mnemonic arrangement of Ian Carter's ICMyanmar font and keyboard arrangement (i.e., Burmese MA is mapped to the "m" key, Burmese KA is on the "k" key, etc.. However, unlike ICMyanmar, pressing the "m" key will insert the Burmese letter MA (U+1019), NOT the ASCII letter "m." This transitional version of the keyboard and font does not attempt to encode Burmese fully in accord with Unicode guidelines. So, for example, an invisible ăthaq "killer" (virama) (U+1039) is not inserted between initial and medial consonants. And no attempt is made in the keyboard's internal processing to rearrange the text buffer into a canonical (phonetic) order, rather than visual order.

5) The Microsoft Word macros used with the ICMyanmar font to insert the variant glyph characters via Ctrl+Alt+key or Ctrl+Alt+Shift+key extended key combinations are now all generally available with SOASMyanmar in any Windows application. They are implemented in identical fashion (with only the two exceptions noted below, required because the original macros are also "dead keys" as now implemented). But in most cases they generate PUA codepoints, which are meaningless unless rendered with the SOASMyanmar font, or converted to their Burmese character equivalents in pure Unicode with a conversion utility and look-up list (more on this in [7] below).

Changed macros:	was "Ctrl+Alt+b"	is now "Ctrl+Alt+b,b"
	was "Ctrl+Alt+m"	is now "Ctrl+Alt+m,m"

6) The keyboard can be used in either of two ways: (a) to type Burmese in visual order, more or less as it is written by hand, and freely inserting variant glyphs encoded with PUA codes wherever the goal of achieving a correct appearance calls for it. This was the method of ICMyanmar and is also the intended use of this transitional version of the SOASMyanmar keyboard and font package. This produces recognizable, but misspelled Burmese, with mystery characters interspersed.

Or (b) the SOASMyanmar keyboard can be used to type Unicode-conformant Burmese in logical order (generally the same as phonetic or "alphabetical" sort order--the order a dictionary is arranged

in. In this case one types a stream of pure correctly-spelled Unicode text, eschewing any of the variant glyphs in the PUA, and inserting the "killer" by hand by pressing the '\' (backslash) key (order for multiple medial consonants: U+101A 'y' or U+101B 'y' precede U+101D 'w', which in turn precedes U+101F 'h'; order for vowel pairs: U+102F 'u' precedes U+102D 'i' and U+1031 'e' precedes U+102C 'a'). In this case, the text will not look right when rendered by the current SOASMyanmar font, but will look right when rendered by any current fully Unicode-conformant Burmese rendering system (e.g., Unitype Global Writer, Xenotype Burmese Language Kit for OS X), or at the operating system level in the future, in a future version of the Windows and Macintosh operating systems, or even by the next-generation SOASMyanmar OpenType font, with smart font-rendering technology.

7) In any case, Unicode-compatible text can be typed which currently looks right when rendered with the SOASMyanmar font, but which is misspelled (mostly, with the dependent vowel "e" and some medial consonants out of place) and has some mystery PUA characters interspersed in the text stream. But our "Convert Myanmar Encoding" Web page (http://wadict/soas.ac.uk) is able to convert such Unicode-compatible text to completely Unicode-compliant text for (a) exchange with other systems which expect correct Unicode; (b) long-term preservation of data; and (c) our own general use in the future, when pure Unicode Burmese text can be correctly rendered by an OpenType "smart font" using the latest computer technology. This is likely to be apparent when a Microsoft Internet Explorer update begins to render in correct visual order any Burmese text which is encoded in correct logical order (via the UniScribe engine USP10.DLL, which is used for all "difficult" text processing, including Arabic, Hindi, Thai, etc. at the moment).

INSTALLATION

The Keyman 6.0 program itself may be downloaded from: http://www.tavultesoft.com. It has been a free download, but this appears to be changing. However, it is also included as a part of the complete SOASMyanmar keyboard-font package.

First install Keyman on your computer. During installation, note where you install the program (by default it is "C:\Program Files\Tavultesoft").

Then, if you have not yet downloaded the SOASMyanmar keyboard-font package itself, download it now. It is suggested that you place the full package file or the "SOASMyanmar.kmx" keyboard file itself in the Tavultesoft root folder. This is by default "C:\Program Files\Tavultesoft".

Then use the Keyman Configuration utility program to install the "SOASMyanmar" keyboard in Windows. See further details in Keyman Help.

You are encouraged to link the SOASMyanmar keyboard with a Windows (bogus) input-language keyboard, as explained in the Keyman documentation. For example, link Keyman and the SOASMyanmar keyboard with the "Malay" input language. This is simply a trick identification with an otherwise unused and innocuous language ID, since Windows does not allow the end-user to add names for new input languages. The advantage is that the SOASMyanmar Burmese keyboard can now be selected by choosing the Malay keyboard in the Windows Taskbar tray at the lower-right corner of the screen, or on the Language Bar, if enabled.

Additionally you can add the shortcut "hot" key Ctrl+Shift+5, to harmonize with other handy shortcuts:

Ctrl+Shift=0 English (United Kingdom) or English (United States) keyboard
Ctrl+Shift=9 Chinese (PRC) keyboard
Ctrl+Shift=8 Afrikaans (really Wa-Pinyin keyboard)
-- includes a-e-i-o-u with macron, acute, caron, and grave accents
Ctrl+Shift=7 Faeroese (really "Latvian (Wa-1)" keyboard)
-- includes just a-e-i-o-u with macron
Ctrl+Shift=6 Icelandic (really "U.S. International" keyboard)
Ctrl+Shift=5 Malay (really "SOASMyanmar" keyboard)
Ctrl+Shift=4 X (really "SOASShan" keyboard)

USAGE

Launch Keyman, if you have chosen to launch it manually, rather than have it load automatically whenever Windows is started. Then make the application in which you wish to use the SOASMyanmar keyboard active (e.g., by clicking within its window), then select the SOASMyanmar keyboard, either from the Keyman icon in the Taskbar tray or with the hotkey combination Ctrl+Shift+5 (or whatever you might choose).

For a picture of the SOASMyanmar keyboard arrangement, see the separate file soasmyanmar_kbd.bmp"; or the HTML file soasmyanmar_kbd.html; or the SOASMyanmar documentation PDF file

To display the Keyman on-screen "visual keyboard," right-click on the Keyman icon in the Taskbar tray or on the Language Bar and select "Visual keyboard." You can also set the visual keyboard to display automatically on startup in Keyman Configuration. The Windows on-screen virtual keyboard (Start, All Programs, Accessories, Accessibility, On-Screen Keyboard in Windows XP) does not display Keyman characters.

In order to match a glyphic variant with a Ctrl+Alt key combination, use the chart below, preferably in the MS Word, HTML, or PDF version of this documentation file, which contain the pictures of the glyphs themselves.

BASIC KEYBOARD KEY ASSIGNMENTS

Note that pressing the Space Bar will insert in the text a regular ASCII space. When displayed with the SOASMyanmar font, this space has a deliberately extra-wide width, approximately the same as a Burmese "single-circle" letter, such as "k'a," "da," or "ma." CTRL+ALT+SPACE will insert the Unicode character U+200B ZERO WIDTH SPACE, i.e., an invisible space, which can optionally be inserted in text for use in text processing, such as line-break page layout, spell-checking, dictionary lookup algorithms, etc.

CTRL+ALT KEY ASSIGNMENTS

In general depress and keep holding down the Ctrl and Alt keys, while pressing the other keys indicated, either unshifted, or shifted. Note that in some cases, there are two alternatives: one in which Ctrl and Alt are kept depressed throughout the keystrokes (noted, e.g., "Ctrl+Alt+m,b", the other in which they are released before the second key (noted, e.g., "Ctrl+Alt+m - b")

TABLE OF KEY ASSIGNMENTS

BASIC BURMESE ALPHABET (UNICODE U+1000-U+109F)

Key (combination)	Character	Code	Unicode Name, transcr., and notes
k	က	U+1000	KA /k/
K	ခ	U+1001	KHA /k'/
g	ဂ	U+1002	GA /g/
G	ဃ	U+1003	GHA /g/
c	င	U+1004	NGA /ng/
s	စ	U+1005	CA /s/
S	ဆ	U+1006	CHA /s'/
z	ဇ	U+1007	JA /z/
Z	ဈ	U+1008	JHA /z/
Ctrl+Alt+V,V	ဉ	U+1009	Pali NYA /ny/
v	ည	U+100A	NNYA /ny/
!	ဋ	U+100B	Pali TTA /t/ (handwritten form)
@	ဌ	U+100C	Pali TTHA /t'/
Ctrl+Alt+&,D	ဍ	U+100D	Pali DDA /d/
$	ဎ	U+100E	Pali DDHA /d/
%	ဏ	U+100F	Pali NNA /n/
t	တ	U+1010	TA /t/
T	ထ	U+1011	THA /t'/
d	ဒ	U+1012	DA /d/
D	ဓ	U+1013	DHA /d/
n	န	U+1014	NA /n/
p	ပ	U+1015	PA /p/
P	ဖ	U+1016	PHA /p'/
b	ဗ	U+1017	BA /b/
B	ဘ	U+1018	BHA /b/
m	မ	U+1019	MA /m/
y	ယ	U+101A	YA /y/
r	ရ	U+101B	RA /y/
l	လ	U+101C	LA /l/
w	ဝ	U+101D	WA /w/
q	သ	U+101E	SA /th/
h	ဟ	U+101F	HA /h/
L	ဠ	U+1020	Pali LLA /l/
A	အ	U+1021	ind. vowel A (initial glottal stop)
I	ဣ	U+1023	ind. vowel I /i/
Ctrl+Alt+e	ဤ	U+1024	ind. vowel II /i/
U	ဥ	U+1025	ind. vowel U /u/
Ctrl+Alt+o	ဦ	U+1026	ind. letter UU /u/

E	ဧ	U+1027	ind. vowel E /e/
x	ဩ	U+1029	ind. vowel O /aw/
X	ဪ	U+102A	ind. vowel AU /aw/
a	◌ာ	U+102C	dep. vowel AA /a/
i	◌ိ	U+102D	dep. vowel I /i, ei/
^	◌ီ	U+102E	dep. vowel II /i/
u	◌ု	U+102F	dep. vowel U /u/
Ctrl+Alt+u	◌ူ	U+1030	dep. vowel UU /u/
e	◌ေ	U+1031	dep. vowel E /e/
`	◌ဲ	U+1032	dep. vowel AI /eh/
M	◌ံ	U+1036	ANUSVARA ('thèthètin') final nasal
.	◌့	U+1037	DOT BELOW =aukmyit ('auqmyiq')
;	း	U+1038	VISARGA ('shé-gá-bauq')
:	း	U+1038	VISARGA ('shé-gá-bauq')
\	◌္	U+1039	VIRAMA "killer"('ăthaq')U+1039+U+200C
European key102	◌္	U+1039	VIRAMA "killer"('ăthaq')U+1039+U+200C
0	၀	U+1040	MYANMAR DIGIT ZERO
1	၁	U+1041	ONE
2	၂	U+1042	TWO
3	၃	U+1043	THREE
4	၄	U+1044	FOUR
5	၅	U+1045	FIVE
6	၆	U+1046	SIX
7	၇	U+1047	SEVEN
8	၈	U+1048	EIGHT
9	၉	U+1049	NINE
'	၊	U+104A	LITTLE SECTION (danda) ('pouq-ka-le', 'pouq-ma-ngeh')
"	။	U+104B	SECTION (double danda) ('pouq-ma')
Ctrl+Alt+O	၌	U+104C	SYMBOL LOCATIVE ('hnaiq')
j	၍	U+104D	SYMBOL COMPLETED ('ywé')
J (or &c\;)	၎	U+104E	SYMBOL AFOREMENTIONED ('lăgaun')
f	၏	U+104F	SYMBOL GENITIVE (or final part.) (í)
unassigned	၀ etc.	U+1050-U+1059	PALI AND SANSKRIT EXTENSIONS

GENERAL PUNCTUATION AND OTHER SYMBOLS IN FONT

Key	Character	Code	Description or Name
Spacebar		U+0020	ASCII space
unassigned	฿	U+0E3F	THAI CURRENCY SYMBOL BAHT
Ctrl+Alt+Spacebar		U+200B	zero width space (ZWSP)
_	—	U+2014	mdash

<	‘	U+2018	quoteleft
Ctrl+Alt+"	’	U+2019	quoteright
{	“	U+201C	doublequoteleft
}	”	U+201D	doublequoteright
Ctrl+Alt+:	…	U+2026	ellipsis

CONTEXTUAL GLYPHIC VARIATIONS FOR BASIC BURMESE ALPHABET (UNICODE PRIVATE USER AREA)

Key combination	Glyph	PUA Code	Unicode character equivalent(s)
Ctrl+Alt+k	ဃ်	U+EC00	U+1000 subscript var.
Ctrl+Alt+K	ဈ်	U+EC01	U+1001 subscript var.
Ctrl+Alt+g	ဉ်	U+EC02	U+1002 subscript var.
Ctrl+Alt+G	ဋ်	U+EC03	U+1003 subscript var.
Ctrl+Alt+s	ဏ်	U+EC04	U+1005 subscript var.
Ctrl+Alt+S	ဘ်	U+EC05	U+1006 subscript var.
Ctrl+Alt+z	မ်	U+EC06	U+1007 subscript var.
Ctrl+Alt+Z	ဠ်	U+EC07	U+1008 subscript var.
Ctrl+Alt+V,A	ဩ	U+EC08	U+1009 + U+102C conjunct
Ctrl+Alt+v	ဥ	U+EC09	U+1009 var.
unassigned	ဣ္ဍ်	U+EC0A	U+1000 +100A+1015+1039+200C 'kyats'
Ctrl+Alt+!,_	ဍ	U+EC0B	U+100B var. (printed form)
Ctrl+Alt+!,!	ဍ္ဍ	U+EC0C	U+100B + subscript U+100B conjunct
Ctrl+Alt+@ (Shft-2)	ဍ္ဎ	U+EC0D	U+100B + subscript U+100C conjunct
#	ဎ္ဍ	U+EC0E	U+100D + subscript U+100D conjunct
Ctrl+Alt+$,D	ဳ	U+EC0F	U+100E + U+100D conjunct
Ctrl+Alt+%,%	ဿ	U+EC10	U+100F + subscript U+100F conjunct
Ctrl+Alt+%,D	ဿ	U+EC11	U+100F + subscript U+100D conjunct
Ctrl+Alt+%,@	ဿ	U+EC12	U+100F + subscript U+100C conjunct
Ctrl+Alt+%,$	ဿ	U+EC13	U+100F + subscript U+100E conjunct
Ctrl+Alt+%,!	ဿ	U+EC14	U+100F + subscript U+100B conjunct
Ctrl+Alt+t	ဓ်	U+EC15	U+1010 subscript var.
Ctrl+Alt+T	ဝ်	U+EC16	U+1011 subscript var.
Ctrl+Alt+d	ဟ်	U+EC17	U+1012 subscript var.
Ctrl+Alt+D	ဠ်	U+EC18	U+1013 subscript var.
N	ဓ	U+EC19	U+1014 var.
unassigned	ဃ	U+EC1A	U+1000 + subscript U+1000
unassigned	ဈ	U+EC1B	U+1002 + subscript U+1002
unassigned	ဋ	U+EC1C	U+1002 + subscript U+1003
unassigned	ဠ	U+EC1D	U+1007 + subscript U+1008
unassigned	ဿ	U+EC1E	U+1014 + subs. U+1010 + subs. U+101B
unassigned	ဓ	U+EC1F	U+1014 + subscript U+1014

~		U+EC20	U+1014 subscript var. (alt. for next)
Ctrl+Alt+N,N		U+EC20	U+1014 subscript var (alt. for previous)
Ctrl+Alt+n,t		U+EC21	U+1014 + subscript U+1010 conjunct
Ctrl+Alt+N,T		U+EC22	U+1014 + subscript U+1011 conjunct
Ctrl+Alt+p		U+EC23	U+1015 subscript var.
Ctrl+Alt+P		U+EC24	U+1016 subscript var.
Ctrl+Alt+b,b		U+EC25	U+1017 subscript var.
Ctrl+Alt+b - b		U+EC25	U+1017 subscript var. (alt. to above)
Ctrl+Alt+b,B		U+EC26	U+1017 + subscript U+1018 conjunct
Ctrl+Alt+B		U+EC27	U+1018 subscript var.
Ctrl+Alt+m,m		U+EC28	U+1019 subscript var.
Ctrl+Alt+m,b		U+EC29	U+1019 + subscript U+1018 conjunct
Ctrl+Alt+m - b		U+EC29	U+1019 + subscript U+1018 conjunct (alt.)
R		U+EC2A	U+101B var.
Ctrl+Alt+l ("el")		U+EC2B	U+101C var.
W		U+EC2C	U+101D var.
Ctrl+Alt+h,w		U+EC2D	U+101D (round) + U+101F high) conjunct
Ctrl+Alt+h - w		U+EC2D	U+101D + U+101F conjunct (alt. to above)
Ctrl+Alt+w,y		U+EC2E	U+101A + U+101D conjunct
Ctrl+Alt+w - y		U+EC2E	U+101A + U+101D conjunct (alt. to above)
Ctrl+Alt+W,Y		U+EC2F	U+101A + U+101D + U+101F conjunct
Q		U+EC54	U+101E subscript var.
Ctrl+Alt+Q		U+EC30	U+101E + U+101E conjunct
Ctrl+Alt+a		U+EC31	U+102C var.
Ctrl+Alt+A		U+EC32	U+102C + U+1039+ U+200C conjunct
Ctrl+Alt+I,M		U+EC33	U+102D + U+1036 conjunct
Ctrl+Alt+# (shft-3)		U+EC34	U+101F + U+1030 conjunct
Ctrl+Alt+U		U+EC5A	U+1030 var.
o		U+EC58	U+102F var.
Ctrl+Alt+H		U+EC35	U+101F var.
H		U+EC36	U+101F var.
Ctrl+Alt+y		U+EC37	U+101A var.
Y		U+EC38	U+101A + U+101F conjunct
Ctrl+Alt+/		U+EC39	U+101A var.
Ctrl+Alt+9,9		U+EC3A	U+101B var. ('yá yiq')
Ctrl+Alt+9 - 9		U+EC3A	U+101B var. (alt. to above)
Ctrl+Alt+[,[		U+EC3B	U+101B var.
Ctrl+Alt+[- [		U+EC3B	U+101B var. (alt. to above)
Ctrl+Alt+(,(		U+EC3C	U+101B var.
Ctrl+Alt+(- (		U+EC3C	U+101B var. (alt. to above)
Ctrl+Alt+{,{		U+EC3D	U+101B var.
Ctrl+Alt+9,u		U+EC3E	U+101B + U+102F conjunct var.

Ctrl+Alt+[,u		U+EC3F	U+101B + U+102F conjunct var.
Ctrl+Alt+(,U		U+EC40	U+101B + U+102F conjunct var.
Ctrl+Alt+{,U		U+EC41	U+101B + U+102F conjunct var.
Ctrl+Alt+9,w		U+EC42	U+101B + U+101D conjunct var.
Ctrl+Alt+9 - w		U+EC42	U+101B + U+101D conjunct var. (alt.)
Ctrl+Alt+[,w		U+EC43	U+101B + U+101D conjunct var.
Ctrl+Alt+[- w		U+EC43	U+101B + U+101D conjunct var. (alt.)
Ctrl+Alt+(,W		U+EC44	U+101B + U+101D conjunct var.
Ctrl+Alt+(- W		U+EC44	U+101B + U+101D conjunct var. (alt.)
Ctrl+Alt+{,W		U+EC45	U+101B + U+101D conjunct var.
Ctrl+Alt+{ - W		U+EC45	U+101B + U+101D conjunct var. (alt.)
Ctrl+Alt+9,0 (zero)		U+EC46	U+101B + U+101D+ U+101F conjunct var.
Ctrl+Alt+9 - 0 (zero)		U+EC46	U+101B + U+101D+ U+101F conjunct (alt.)
Ctrl+Alt+[,0 (zero)		U+EC47	U+101B + U+101D+ U+101F conjunct var.
Ctrl+Alt+[- 0 (zero)		U+EC47	U+101B + U+101D+ U+101F conjunct(alt.)
Ctrl+Alt+(,)		U+EC48	U+101B + U+101D+ U+101F conjunct var.
Ctrl+Alt+(-)		U+EC48	U+101B + U+101D+ U+101F conjunct (alt.)
Ctrl+Alt+{,}		U+EC49	U+101B + U+101D+ U+101F conjunct var.
Ctrl+Alt+{ - }		U+EC49	U+101B + U+101D+ U+101F conjunct (alt.)
Ctrl+Alt+0,0 (zero)		U+EC4A	U+101B var.
Ctrl+Alt+0 - 0 (zero)		U+EC4A	U+101B var. (alt. to above)
Ctrl+Alt+],]		U+EC4B	U+101B var.
Ctrl+Alt+} - }		U+EC4B	U+101B var.
Ctrl+Alt+),)		U+EC4C	U+101B var.
Ctrl+Alt+] -]		U+EC4C	U+101B var. (alt. to above)
Ctrl+Alt+},)		U+EC4D	U+101B var.
Ctrl+Alt+.		U+EC4E	U+1037 var.
>		U+EC4F	U+1037 var.
Ctrl+Alt+\,\		U+EC50	U+1004 var. (kinzi)
Ctrl+Alt+\ - \		U+EC50	U+1004 var. (kinzi) (alt. to above)
C		U+EC51	U+1004 var. (kinzi) + U+1036 conjunct
Ctrl+Alt+\,i		U+EC52	U+1004 var. (kinzi) + U+102D conjunct
Ctrl+Alt+\ - i		U+EC52	U+1004 var. (kinzi) + U+102D conj. (alt.)
Ctrl+Alt+|,^		U+EC53	U+1004 var. (kinzi) + U+102E conjunct
Q		U+EC54	U+101E subscript var.
V		U+EC55	U+100A var.
O		U+EC56	U+101F + U+102F conjunct
&		U+EC57	U+104E partial glyph
o		U+EC58	U+102F var.
unassigned		U+EC59	U+1020 + subscript U+101F var.
Ctrl+Alt+U		U+EC5A	U+1030 var.
Ctrl+Alt+$,$		U+EC5B	U+100E subscript var.

Ctrl+Alt+$ - $	ᦲ	U+EC5B	U+100E subscript var. (alt. to above)
unassigned	ᦳ	U+EC5C	U+1021 subscript var.
unassigned	ᦴ	U+EC5D	U+1007 subs. (EC06) + U+101A var. (EC39) conjunct
unassigned	ᦵ	U+EC5E	U+101B initial var.
unassigned	ᦶ	U+EC5F	U+1004 ('halbard tip' kinzi) + U+102C var.
unassigned	ᦷ	U+EC60	U+101A + U+102D + U+102F conjunct
unassigned	ᦸ	U+EC61	U+1004 (halberd-tip kinzi) + U+1002 + subscript U+101B + U+102F + U+102D
unassigned	ᦹ	U+EC62	U+1009 NYA + U+1005 conjunct
unassigned	ᦺ	U+EC63	U+1009 NYA + U+1007 conjunct
unassigned	ᦻ	U+EC64	U+101F subscript var.
unassigned	ᦼ	U+EC65	U+1000 + subscript U+1008
unassigned	ᦽ	U+EC66	U+1010 subs. + U+101D subs. conjunct
unassigned	ᦾ	U+EC67	U+101C + subscript U+101C conjunct
unassigned	ᦿ	U+EC68	U+101A subscript var.
unassigned	ᧀ	U+EC69	U+100C subscript var.
unassigned	ᧁ	U+EC6A	U+100B subscript var.
unassigned	ᧂ	U+EC6B	U+1018 subs. + U+1037 subs. conjunct
unassigned	ᧃ	U+EC6C	U+100F subscript var.
unassigned	ᧄ	U+EC6D	U+101B + U+102F + U+102D conjunct
unassigned	ᧅ	U+EC6E	U+101B + U+101D (triangle var.) conjunct
unassigned	ᧆ	U+EC6F	U+101B + U+101D (triangle var.) conjunct
unassigned	ᧇ	U+EC70	U+101D (triangle var.) + U+101F conjunct
unassigned	ᧈ	U+EC71	U+101A + U+101D (triangle var.) + U+101F conjunct
unassigned	ᧉ	U+EC72	U+101A + U+101D (triangle var.) conjunct
unassigned	᧊	U+EC73	U+1009 + subscript U+1005 conjunct
unassigned	᧋	U+EC74	U+1009 + subscript U+1007 conjunct
unassigned	᧌	U+EC75	U+101D (triangle var.) + U+101F conjunct
unassigned	᧍	U+EC76	U+101D round var. + U+101F low tail conjunct
unassigned	᧎	U+EC77	U+101B + U+102F + U+102D conjunct

SHAN ALPHABET (INCLUDING GLYPHS FOR BOTH NEW AND OLD SHAN SCRIPT IN UNICODE PRIVATE USER AREA)

(Use SOASShan Keyboard)

Key (combination)	Glyph	Code	Unicode character equivalent(s) (tentative)
k	ᧁ	U+EC80	Shan KA /k/
K	ᧂ	U+EC81	Shan KHA /k'/
c	ᧃ	U+1004	NGA /ng/
s	ᧄ	U+EC85	Shan CA /ts/
S	ᧅ	U+EC86	Shan CHA /s/ (glyph same as U+101E)

N	၎	U+EC8A	Shan NNYA /ny/ or final /[u]y/, /[ɯ,ɤ]y/, /n/?
t	တ	U+1010	TA /t/
T	ထ	U+1011	THA /t'/
d	ဓ	U+1013	DHA stressed /t/ in Old Shan religious texts
n	ၼ	U+EC94	Shan NA /n/
p	ပ	U+1015	PA /p/
P	ၽ	U+EC96	Shan PHA /p'/
b	ဗ	U+1017	BA stressed /p/ in Old Shan religious texts
m	မ	U+1019	MA /m/
f	ၾ	U+ECF5	Shan eastern dialect FA /f/
y	ယ	U+101A	YA /y/
Y	◌ၢ	U+EC37	U+101A medial var.
r	ရ	U+101B	RA /r/
R	◌ၣ	U+EC3A	U+101B medial single-circle var.
L	◌ၤ	U+EC3C	U+101B medial double-circle var.
l	လ	U+101C	LA /l/
w	ဝ	U+101D	WA /w/
W	◌ွ	U+ECE8	U+101D medial (hook) var. /w/ ("Shan medial VA")
o	◌ၜ	U+ECF4	U+101D WA medial (triangular) var. /w/ (Shan vowel sign? /ɔ/ in CVC syllable)
O	◌ဝ	U+EC2C	U+101D WA medial (round) var. /w/
q	ဆ	U+EC9E	Shan SA /th/ (glyph similar to U+1006)
h	ၶ	U+EC9F	Shan HA /h/
A	ဣ	U+ECA1	Shan ind. vowel A
z	◌ါ	U+ECEF	Shan dep. vowel sign AA /aa/
a	◌ာ	U+ECF8	Shan dep. vowel sign U+ECEF AA var. in [C]VC syllables /aa/
i	◌ိ	U+102D	dep. vowel sign I /i/ in [C]VC syllables
I (or ^)	◌ီ	U+102E	dep. vowel sign II /i/ (written w/ U+102D in [C]VC syllables)
u	◌ု	U+102F	dep. vowel sign U /u/ in [C]VC syllables
U (or Ctrl+Alt+u)	◌ူ	U+1030	dep. vowel sign UU /u/ (written w/ U+102F in [C]VC syllables)
e	ေ◌	U+1031	dep. vowel E /e/ (written w/ U+ECEB in Shan [C]VC syllables)
E	◌ဵ	U+ECEB	Shan U+1031 var. in [C]VC syllables /e/ ("Shan vowel sign superscript A")
g	◌ဲ	U+ECF9	Shan dep. vowel sign AE?AI?AY? /ɛ/ (written w/ U+ECEC in [C]VC syllables)
G	◌ဳ	U+ECEC	U+ECF9 var. in [C]VC syllables /ɛ/ ("Shan vowel sign superscript AY")
j	◌ဴ	U+ECED	"Shan vowel sign superscript AI" /ai/
M	◌ံ	U+1036	ANUSVARA final /m/ abbreviation

:	း	U+1038	VISARGA (Shan 4th tone mark [high level tone] (or use colon char?)
\	◌်	U+1039	VIRAMA "killer" U+1039+U+200C
European key102	◌်	U+1039	VIRAMA "killer" U+1039+U+200C
C	◌ံ	U+ECF6	U+1039 + U+1036 conjunct (Old Shan /eo/ /eow/ + ANUSVARA /m/ (e.g. Cushing)
0	၀	U+1040	MYANMAR DIGIT ZERO
1	၁	U+1041	ONE
2	၂	U+1042	TWO
3	၃	U+1043	THREE
4	၄	U+1044	FOUR
5	၅	U+1045	FIVE
6	၆	U+1046	SIX
7	၇	U+1047	SEVEN
8	၈	U+1048	EIGHT
9	၉	U+1049	NINE
'	၊	U+104A	LITTLE SECTION (danda)
"	။	U+104B	SECTION (double danda)
!	၍	U+104D	SYMBOL COMPLETED (Shan abbr. for U+ECA1+102F+102D+EC8A+1039+1038 'Hey')
,	◌ၽ	U+ECE7	Shan 2nd tone mark [low level tone] (or use comma?)
;	◌ၷ	U+ECE6	Shan 3rd tone mark [mid level tone] (or use semicolon?)
.	◌ၺ	U+ECF2	Shan 5th tone mark [falling tone] (or use period/full stop?) ("Shan tone mark BELOW AI" ?)
~	◌ၻ	U+ECF7	Shan 6th (emphatic) tone mark (cf. U+0E30)
unassigned	◌ၵ	U+ECE9	("Shan medial PHA" ?)
unassigned	◌ၢ	U+ECEA	("Shan vowel sign AOU" ?)
unassigned	◌ံ	U+ECEE	("Shan vowel sign superscript AIM") cf. ECF6
unassigned	◌ၰ	U+ECF0	("Shan tone mark AI" ?)
unassigned	◌ၱ	U+ECF1	("Shan tone mark AII" ?)
unassigned	◌ၳ	U+ECF3	("Shan tone mark BELOW AII" ?)
unassigned	◌ၸ	U+ECFA	("Shan Pa[-]Oh tone mark MAI PAK NGA"?)
unassigned	◌ၹ	U+ECFB	("Shan Pa[-]Oh tone mark MAI NGA" ?)

SOASMyanmar Keyboard Pictures

Unicode: Unshifted

Unicode: Shift

Unicode: Ctrl+Alt

Unicode: Shift+Ctrl+Alt

SOASShan Keyboard Pictures

Unicode: Unshifted

Unicode: Shift

Unicode: Ctrl+Alt

Unicode: Shift+Ctrl+Alt

Rev. 2005.05.26 rak