

# Myanmar: Android Zawgyi and Unicode

 This article is still a stub

This page aims to gather information, experience and best practices about representation of Myanmar characters on Android smartphones. This article is still a stub. Hopefully the ICT4D and CommCare communities of Myanmar will help to flesh it out.

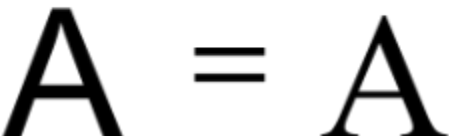
[Charles Flèche](#)

## Unicode vs Zawgyi

Unicode is a **standard** that defines how text should be saved into data and how it is read and written. Almost every script in the world, including Burmese, is defined in the Unicode Standard. Carefully designed by experts of the field around the world, this international standard is supported by virtually all the latest platforms: not only in major operating systems like Windows, Mac, and Linux but also in OSes of mobile phones and many electronic devices.

Operating system	Version since Myanmar UNICODE is supported
Android	4.4
iOS	Unknown
Linux	Any recent distribution supports it
OS X	Unknown
Windows	8.1

Switching the font of a text typed in *Arial* to *Times New Roman* only changes the way the letters are displayed : an *A* would still be an *A*, a *B* would still be a *B*, etc. This is made possible because *Arial* and *Times New Roman* both respect the Unicode standard : internally, in the computer's memory, the different letters of the alphabet are represented the same way in both fonts.

Arial Unicode compliant = Times New Roman Unicode compliant  


Zawgyi is a **font**, invented at a time where Unicode was not a widely used international standard. Zawgyi is currently the most widely used font in Myanmar. However, it is **not Unicode-compliant** and breaks the Unicode standard in many ways.

Padauk Unicode compliant ≠ Zawgyi Not Unicode  


Comparing the Burmese characters Unicode map with the matching Zawgyi one shows the main issues on this font :

1. Lots of Burmese characters do not follow the Unicode standard
2. Zawgyi doesn't support ethnic Myanmar languages

Unicode																	Zawgyi-One																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+100x	က	ခ	ဂ	ဃ	င	စ	ဆ	ဇ	ည	ဋ	ဌ	ဍ	ဎ	တ	ထ	ဒ	က	ခ	ဂ	ဃ	င	စ	ဆ	ဇ	ည	ဋ	ဌ	ဍ	ဎ	တ	ထ	ဒ	
U+101x	ဏ	တ	ထ	ဒ	န	ပ	ဖ	ဗ	ဉ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ	တ	ထ	ဒ	န	ပ	ဖ	ဗ	ဉ	ဏ	ဏ	ဏ	ဏ	ဏ	ဏ		
U+102x	ဋ	ဌ	ဍ	ဎ	တ	ထ	ဒ	န	ပ	ဖ	ဗ	ဉ	ဏ	ဏ	ဏ	ဏ	ဋ	ဌ	ဍ	ဎ	တ	ထ	ဒ	န	ပ	ဖ	ဗ	ဉ	ဏ	ဏ	ဏ		
U+103x	ါ	ေ	ဲ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ါ	ေ	ဲ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+104x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+105x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+106x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+107x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+108x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	
U+109x	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	ံ	

The [Wikipedia Myanmar community](#) page contains a more thorough discussion on Unicode vs Zawgyi and has good links to extra fonts and conversion tools.

# Android support for Myanmar languages

## Before Android 4.4

Before the version 4.4, Android didn't support the Myanmar language at all. Smartphone sold in Myanmar with an earlier Android version usually have Zawgyi installed by default.

## After Android 4.4

From Kit Kat (version 4.4+), Android natively supports the Myanmar languages with a Unicode compliant font. However, **some vendors replaces the standard Unicode compliant font with Zawgyi** for handsets sold in South-East Asia.

## Default Myanmar font for Asian-bought smartphones

Vendor	Model name	Model number	Android version	Myanmar font	Country of buying	Note
Asus	Zenfone	ASUS_T00Q	4.4.2	Unicode	Myanmar	
Huawei		G730-C00	4.1.2	Zawgyi	Myanmar	
Samsung	Galaxy Core Prime	SM-G360H	4.4.4	Zawgyi	Myanmar	The ROM location code is CAM, for Cambodia.

# Myanmar languages support in Android keyboards

Generally, Android phones bought in Myanmar should be able to **display** Burmese characters. However, to **type** in Burmese, a Myanmar language compliant keyboard application must be installed. Depending on font installed on a given device (the standard Unicode compliant font or Zawgyi), different packages should be used.

Font	Keyboard	Note
Unicode compliant	<a href="#">M3 Keyboard (Myanmar)</a>	
Unicode compliant	<a href="#">MUA Keyboard</a>	
Zawgyi	<a href="#">Bagan Keyboard</a>	Advertisement ridden

# Am I using an Unicode compliant font or Zawgyi ?

## Selecting a character to check

The Unicode character MYANMAR SYMBOL AFOREMENTIONED (Unicode character U+104E) allows for a good test as:

- both Unicode and Zawgyi represents this character (it is not an "empty case" in neither options)

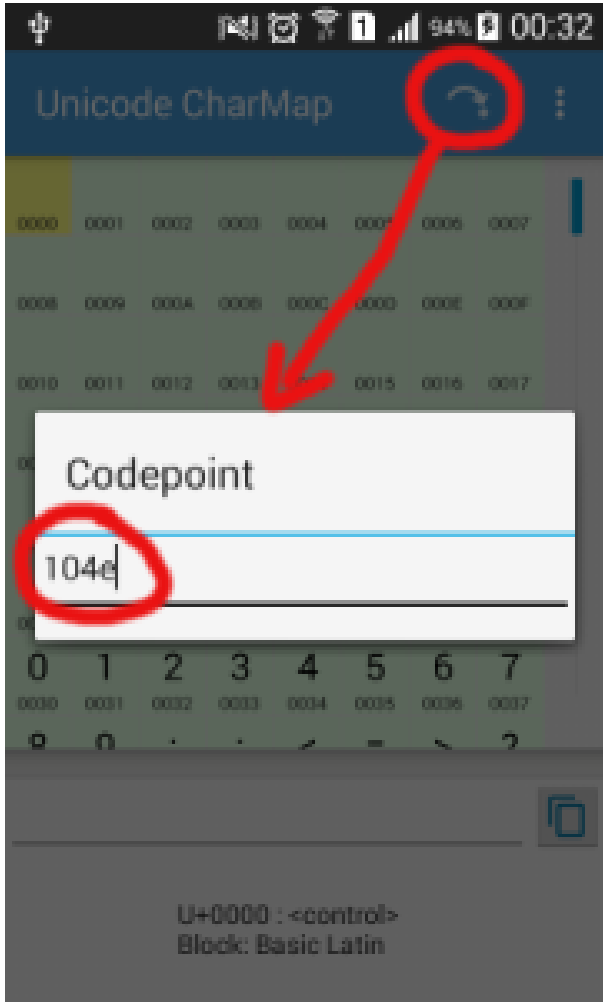
- the two representations are really obviously different for a non-native Burmese reader
- it is a standalone character, not a composition of other glyphs
- it is a common Burmese character and not part of an ethnic language

Unicode																Zawgyi-One																	
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+100x	၀	၁	၂	၃	၄	၅	၆	၇	၈	၉	၁၀	၁၁	၁၂	၁၃	၁၄	၁၅	၀	၁	၂	၃	၄	၅	၆	၇	၈	၉	၁၀	၁၁	၁၂	၁၃	၁၄	၁၅	
U+101x	၁၆	၁၇	၁၈	၁၉	၂၀	၂၁	၂၂	၂၃	၂၄	၂၅	၂၆	၂၇	၂၈	၂၉	၃၀	၃၁	၁၆	၁၇	၁၈	၁၉	၂၀	၂၁	၂၂	၂၃	၂၄	၂၅	၂၆	၂၇	၂၈	၂၉	၃၀	၃၁	
U+102x	၃၂	၃၃	၃၄	၃၅	၃၆	၃၇	၃၈	၃၉	၄၀	၄၁	၄၂	၄၃	၄၄	၄၅	၄၆	၄၇	၃၂	၃၃	၃၄	၃၅	၃၆	၃၇	၃၈	၃၉	၄၀	၄၁	၄၂	၄၃	၄၄	၄၅	၄၆	၄၇	
U+103x	၄၈	၄၉	၅၀	၅၁	၅၂	၅၃	၅၄	၅၅	၅၆	၅၇	၅၈	၅၉	၆၀	၆၁	၆၂	၆၃	၄၈	၄၉	၅၀	၅၁	၅၂	၅၃	၅၄	၅၅	၅၆	၅၇	၅၈	၅၉	၆၀	၆၁	၆၂	၆၃	
U+104x	၆၄	၆၅	၆၆	၆၇	၆၈	၆၉	၇၀	၇၁	၇၂	၇၃	၇၄	၇၅	၇၆	၇၇	၇၈	၇၉	၆၄	၆၅	၆၆	၆၇	၆၈	၆၉	၇၀	၇၁	၇၂	၇၃	၇၄	၇၅	၇၆	၇၇	၇၈	၇၉	
U+105x	၈၀	၈၁	၈၂	၈၃	၈၄	၈၅	၈၆	၈၇	၈၈	၈၉	၉၀	၉၁	၉၂	၉၃	၉၄	၉၅	၈၀	၈၁	၈၂	၈၃	၈၄	၈၅	၈၆	၈၇	၈၈	၈၉	၉၀	၉၁	၉၂	၉၃	၉၄	၉၅	
U+106x	၉၆	၉၇	၉၈	၉၉	၁၀၀	၁၀၁	၁၀၂	၁၀၃	၁၀၄	၁၀၅	၁၀၆	၁၀၇	၁၀၈	၁၀၉	၁၁၀	၁၁၁	၉၆	၉၇	၉၈	၉၉	၁၀၀	၁၀၁	၁၀၂	၁၀၃	၁၀၄	၁၀၅	၁၀၆	၁၀၇	၁၀၈	၁၀၉	၁၁၀	၁၁၁	
U+107x	၁၁၂	၁၁၃	၁၁၄	၁၁၅	၁၁၆	၁၁၇	၁၁၈	၁၁၉	၁၂၀	၁၂၁	၁၂၂	၁၂၃	၁၂၄	၁၂၅	၁၂၆	၁၂၇	၁၁၂	၁၁၃	၁၁၄	၁၁၅	၁၁၆	၁၁၇	၁၁၈	၁၁၉	၁၂၀	၁၂၁	၁၂၂	၁၂၃	၁၂၄	၁၂၅	၁၂၆	၁၂၇	
U+108x	၁၂၈	၁၂၉	၁၃၀	၁၃၁	၁၃၂	၁၃၃	၁၃၄	၁၃၅	၁၃၆	၁၃၇	၁၃၈	၁၃၉	၁၄၀	၁၄၁	၁၄၂	၁၄၃	၁၂၈	၁၂၉	၁၃၀	၁၃၁	၁၃၂	၁၃၃	၁၃၄	၁၃၅	၁၃၆	၁၃၇	၁၃၈	၁၃၉	၁၄၀	၁၄၁	၁၄၂	၁၄၃	
U+109x	၁၄၄	၁၄၅	၁၄၆	၁၄၇	၁၄၈	၁၄၉	၁၅၀	၁၅၁	၁၅၂	၁၅၃	၁၅၄	၁၅၅	၁၅၆	၁၅၇	၁၅၈	၁၅၉	၁၄၄	၁၄၅	၁၄၆	၁၄၇	၁၄၈	၁၄၉	၁၅၀	၁၅၁	၁၅၂	၁၅၃	၁၅၄	၁၅၅	၁၅၆	၁၅၇	၁၅၈	၁၅၉	

### Checking the character map

A character map is an organized, visual representation of all the characters and the associated values used internally by a device for a given font. Using a character map application, it is possible to check the representation of the Unicode Character U+104E. The following illustrations have been taken using [Unicode CharMap](#).

### Finding the U+104E character



## Unicode representation of the U+104E character



The screenshot shows the Unicode CharMap application interface. At the top, the title bar reads "Unicode CharMap" and the time is 12:39. Below the title bar is a grid of Myanmar characters. The grid is organized into rows and columns, with each cell containing a character and its corresponding Unicode code point. The character U+104E is highlighted in yellow. Below the grid, there is a search bar and a status bar at the bottom.

၀	၁	၂	၃	၄	၅	၆	၇
1040	1041	1042	1043	1044	1045	1046	1047
၈	၉	၁၀	၁၁	၁၂	၁၃	၁၄	၁၅
1048	1049	104A	104B	104C	104D	104E	104F
၁၆	၁၇	၁၈	၁၉	၂၀	၂၁	၂၂	၂၃
1050	1051	1052	1053	1054	1055	1056	1057

U+104E: MYANMAR SYMBOL AFOREMENTIONED  
Block: Myanmar

## Zawgyi representation of the U+104E character

Unicode CharMap

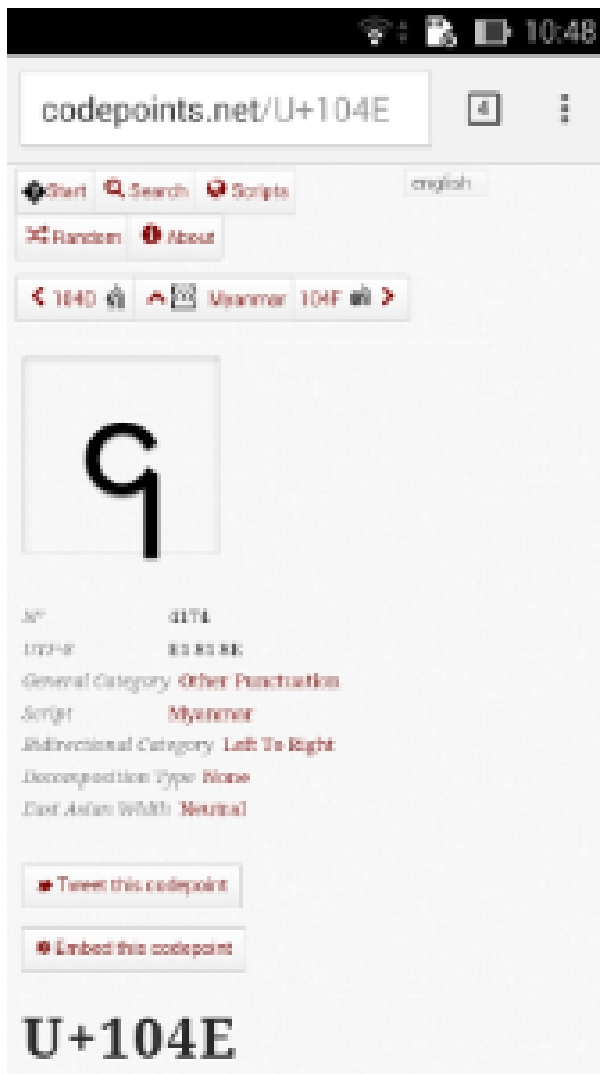
၉	အ		အိ	အိ	ဥ	ဥ	ဇ
1020	1021	1022	1023	1024	1025	1026	1027
	ဩ	ဩ	ါ	ာ	ာ	ာ	ါ
1028	1029	102A	102B	102C	102D	102E	102F
။	ေ	်	ါ	။		့	့
1030	1031	1032	1033	1034	1035	1036	1037
း	့	ါ	ါ	့	ါ		
1038	1039	103A	103B	103C	103D	103E	103F
့	ာ	ါ	ါ	ါ	ါ	ါ	ါ
1040	1041	1042	1043	1044	1045	1046	1047
့	ါ	ါ	။	ါ	ါ	၎င်း	ါ
1048	1049	104A	104B	104C	104D	104E	104F
1050	1051	1052	1053	1054	1055	1056	1057
		ါ					

U+104E : MYANMAR SYMBOL AFOREMENTIONED  
Block: Myanmar

In the web browser

Web browsers not only rely on system fonts, they also embed their own. The representation of the U+104E character can be checked on [Codepoints](#).

Web browser Unicode representation of the U+104E character



Web browser Zawgyi representation of the U+104E character

codepoints.net/U+104E

Start Search Scripts english

Random About

104D Myanmar 104F



N<sup>o</sup> 4174  
UTF-8 E1 81 8E  
General Category **Other Punctuation**  
Script **Myanmar**  
Bidirectional Category **Left To Right**  
Decomposition Type **None**  
East Asian Width **Neutral**