

Proposed keyboard layout for Swahili in Arabic script

Kevin Donnelly

kevin@dotmon.com

Version 0.1, March 2010

Introduction

Swahili was originally written in Arabic script in its area of origin (the littoral of Kenya and Tanzania), and although standard Swahili is now written in Roman script, it may be useful to have the option of Arabic script (for instance, so that manuscripts can be transcribed and directly converted to digital format as an alternative to, or an extension of, transliteration).

The system set out here is an initial attempt to provide such functionality on the GNU/Linux operating system, the particular version used being Kubuntu 9.10 (Karmic Koala). There still exist some difficulties, and I apologise for the fact that my technical knowledge is too limited to resolve all of them. I hope that others in the community who have more expertise than me will be able to assist in the resolution of these issues.

That said, the system as it stands is fully useable in terms of producing good output in an office suite (OpenOffice.org) and on the Web (provided some conditions are met). It can be typed almost as fast as the normal Roman orthography, and considerably faster than the Roman orthography plus any diacritics for full transliteration.

An example stanza from the *Utendi wa Jaafari* (The Ballad of Ja^cfar) will demonstrate:

عَلَى أَكْئُنْكَ أَكْنَدَ كُوَهْرَكَ
حَتَّى نَيْمِينَ كِفِكَ مَوْنَ فَاطِمَ هُلِيَّيَ

°Alii akainuka akaenda kwaharaka
hata nyumbani kifika Mwana Fatima huliya

(Ali rose up and went with haste
until he reached his home [where] the Lady Fatima was weeping)

Creating a new keyboard layout for Swahili

X11 keyboard layouts are located in the `/usr/share/X11/xkb/symbols` directory in Kubuntu 9.10, and the first step is adding a new layout for Swahili. The proposed keyboard file is attached in *Annex 1*. Run **sudo kate** (or a text editor of your choice), and copy the contents of *Annex 1* into the editor. Save the file as `/usr/share/X11/xkb/symbols/sw`. (Alternatively, download the file from my website, and save it.)

This keyboard layout is based on the Buckwalter layout [1] in the file `/usr/share/X11/xkb/symbols/ara` put together by Martin Vidner. The big advantage of the Buckwalter system is that it allocates Arabic letters so far as possible to keys which have the same sound in English. Although a complete correlation is impossible, the assignment of other letters around this framework means

that typing in Arabic becomes very easy for someone used to a Roman keyboard.

The Swahili layout follows the Buckwalter one fairly closely, but adds letters that are specific to Swahili (eg peh, ain with three dots), and rearranges some other letters to allow for this. The assignments for each key are set out in *Annex 2*. This gives the key in a QWERTY keyboard on the top line, with the Arabic letter assigned to that key in the second line. The third line gives the name used in the `sw` layout file (where they are preceded by *Arabic_*), and the last line gives the Unicode number for that letter (the latter, preceded by *0x100*, are used in the `sw` layout file where there seems to be no assigned X11 name. Names are preferred where possible, since it makes the file easier to understand.

Making X11 aware of the new layout

The information here is based on a very useful blog post by Daniel Paul O'Donnell [2]. Use your editor to open `/usr/share/X11/xkb/rules/evdev.xml`. Go to the end of the `<layoutList>` section (search for `</layoutList>`), and add the following after the last `</layout>` tag and before `</layoutList>`:

```
<layout>
  <configItem>
    <name>sw</name>
    <shortDescription>Swa</shortDescription>
    <description>Swahili in Arabic script</description>
    <languageList>
      <iso639Id>swa</iso639Id>
    </languageList>
  </configItem>
  <variantList/>
</layout>
```

Save the file.

Making KDE aware of the new layout

Go to **K** → **Settings** → **System Settings** → **Regional and Language** → **Keyboard Layout**. Tick **Enable keyboard layouts**. Near the bottom of the list of available layouts you should see an entry for Swahili in Arabic script - click on that, and then click the green right-facing button to move it into the list of active layouts. Click **Apply**, and a keyboard switcher (a little flag) should pop up in the tray section of your panel. Clicking on it will change it into a grey square with **sw** on it - when this is showing, typing on your keyboard will give you the Arabic letters specified in the Swahili layout. The keyboard switch can also be activated by pressing the keys **Ctrl+Alt+K**.

Note that should you make changes to the keyboard layout file in `/usr/share/X11/xkb/symbols/sw`, you must reload the keyboard switcher. The easiest way of doing this is to right-click on it and select **Configure**. Tick **Disable keyboard layouts**, and next tick **Enable keyboard layouts**, and then **Apply**, and the switcher will reload.

Installing Arabic language support

Run the following commands to install packages giving support for Arabic:

```
sudo apt-get install language-support-fonts-ar
sudo apt-get install language-support-input-ar
sudo apt-get install language-support-writing-ar
```

Alternatively, use Synaptic or Aptitude to do the same thing.

(Note that I am uncertain as to whether all of the items in these packages are required, but this is what I did.)

Choosing a font

Unfortunately, the default Liberation fonts on Kubuntu do not display Arabic letters properly - some medials and finals are not joined, and some letters, (eg *peh*) appear as boxes. So you need to choose a font that avoids these issues.

Kubuntu includes a variety of attractive fonts for Arabic, and the one I am using is *SIL Scheherazade*. To install it, run **apt-get install ttf-sil-scheherazade**.

Typing some text

Open *OpenOffice.org Writer*, and select Scheherezade as the font. I also change the font size to 20 to make it easier to read. Press **Ctrl+Alt+K** to switch the keyboard to Swahili, and start typing. I have found that for some reason DejaVu Sans gets used as the font, even though I have specified Scheherazade, so I press **Ctrl+A** and select Scheherezade again, and then it works properly.

Remaining issues

In OpenOffice.org, the text is left-justified, so you have to click the right-justification button. This may be a function of my Swahili layout, because the original Arabic layout does not have this problem. I did in fact originally try to add the Swahili layout as a variant of the Arabic layout, but I couldn't get X11 to see it, hence my decision to set up a separate layout.

The problems with the default Liberation fonts are vexing, and it would be nice to adjust the glyph specification so that all the letters show properly.

This also has an impact on the use of Swahili in Arabic script on the Web. I have found [3] that if the CSS properties are set to Scheherezade on the website, with a fallback of monospace, and the browser is running on a machine that has the Scherezade font installed, the text will display perfectly on Firefox when running GNU/Linux, but with some degradation on Konqueror. However, the display on other operating systems is less attractive. It would be nice to resolve this, but I confess I am not entirely sure where to start!

Conclusion

This howto has put forward a suggested configuration that allows you to use Swahili written in Arabic script. If anyone reading it has suggestions for further improvement, or corrections or amendments to offer, I would be very glad to receive them.

Kevin Donnelly (kevin@dotmon.com)
kevindonnelly.org.uk

[1] http://en.wikipedia.org/wiki/Buckwalter_transliteration

[2] <http://people.uleth.ca/~daniel.odonnell/Blog/custom-keyboard-in-linuxx11>

[3] <http://kevindonnelly.org.uk/2010/03/swahili-in-arabic-script-on-wordpress/>

Annex 1

Proposed X11 keyboard layout for Swahili in Arabic script

```
// $XKeyboardConfig$

// Keyboard layout for Swahili in Arabic script.
// Based on Martin Vidner's Buckwalter transliteration variant of the Arabic keyboard
// Please notify any corrections or omissions to
// Kevin Donnelly (kevin@dotmon.com)

partial alphanumeric_keys
xkb_symbols "swahili" {
    name[Group1]= "Swahili";

    key <TLDE> { [ 0x1000670, VoidSymbol ] };

    key <AE01> { [ 0x1000661, VoidSymbol ] };
    key <AE02> { [ 0x1000662, VoidSymbol ] };
    key <AE03> { [ 0x1000663, VoidSymbol ] };
    key <AE04> { [ 0x1000664, VoidSymbol ] };
    key <AE05> { [ 0x1000665, Arabic_percent ] };
    key <AE06> { [ 0x1000666, VoidSymbol ] };
    key <AE07> { [ 0x1000667, VoidSymbol ] };
    key <AE08> { [ 0x1000668, VoidSymbol ] };
    key <AE09> { [ 0x1000669, 0x100fd3e ] };
    key <AE10> { [ 0x1000660, 0x100fd3f ] };
    key <AE11> { [ minus, Arabic_tatweel ] };
    key <AE12> { [ equal, plus ] };



    key <AD01> { [ Arabic_qaf, Arabic_gaf ] };
    key <AD02> { [ Arabic_waw, Arabic_hamzaonwaw ] };
    key <AD03> { [ 0x1000656, Arabic_ain ] };
    key <AD04> { [ Arabic_ra, Arabic_tteh ] };
    key <AD05> { [ Arabic_teh, Arabic_tah ] };
    key <AD06> { [ Arabic_yeh, Arabic_alefmaksura ] };
    key <AD07> { [ Arabic_damma, Arabic_dammatan ] };
    key <AD08> { [ Arabic_kasra, Arabic_hamzaonyeh ] };
    key <AD09> { [ Arabic_sukun, 0x1000657 ] };
    key <AD10> { [ Arabic_peh, Arabic_tehmarbuta ] };
    key <AD12> { [ VoidSymbol, VoidSymbol ] };

    key <AC01> { [ Arabic_fatha, Arabic_alef ] };
    key <AC02> { [ Arabic_seen, Arabic_sad ] };
    key <AC03> { [ Arabic_dal, Arabic_dad ] };
    key <AC04> { [ Arabic_feh, Arabic_fathatan ] };
    key <AC05> { [ 0x10006A0, Arabic_ghain ] };
    key <AC06> { [ Arabic_ha, Arabic_hah ] };
    key <AC07> { [ Arabic_jeem, VoidSymbol ] };
    key <AC08> { [ Arabic_kaf, Arabic_kasratan ] };
    key <AC09> { [ Arabic_lam, Arabic_shadda ] };
    key <AC10> { [ Arabic_semicolon, VoidSymbol ] };
    key <AC11> { [ Arabic_hamza, VoidSymbol ] };

    key <AB01> { [ Arabic_zain, Arabic_zah ] };
    key <AB02> { [ Arabic_khah, VoidSymbol ] };
    key <AB03> { [ Arabic_tcheh, Arabic_sheen ] };
    key <AB04> { [ 0x10006A8, Arabic_theh ] };
    key <AB05> { [ Arabic_beh, Arabic_thal ] };
    key <AB06> { [ Arabic_noon, VoidSymbol ] };
    key <AB07> { [ Arabic_meem, VoidSymbol ] };
    key <AB08> { [ Arabic_comma, Arabic_hamzaunderalef ] };
    key <AB09> { [ 0x10006d4, Arabic_hamzaonalef ] };
    key <AB10> { [ VoidSymbol, Arabic_question_mark ] };
};
```

Annex 2:

Key assignments for the proposed X11 keyboard layout for Swahili in Arabic script

`	1	2	3	4	5	6	7	8	9	0	-
'	١	٢	٣	٤	٥	٦	٧	٨	٩	.	-
0670	0661	0662	0663	0664	0665	0666	0667	0668	0669	0660	
					%						-
					percent						tatweel
					066A				FD3E	FD3F	

q	w	e	r	t	y	u	i	o	p	[
ق	و	ا	ر	ت	ي	ء	ِ	◌◌◌	پ	
qaf	waw		ra	tch	yeh	damma	kasra	sukun	peh	
0642	0648	0656	0631	062A	064A	064F	0650	0652	067E	
گ	ؤ	ع		ط	ى	ة	ئ	‘	ة	أ
gaf	hamzaonwaw	ain		tah	alefmaksura	dammatan	hamzaonyeh		tehmarbuta	
06AF	0624	0639		0637	0649	064C	0626	0657	0629	0671

a	s	d	f	g	h	j	k	l	;	'	#
َ	س	د	ف	غ	ه	ج	ك	ل	؛	ء	
fatha	seen	dal	feh		ha	jeem	kaf	lam	semicolon	hamza	
064E	0633	062F	0641	06A0	0647	062C	0643	0644	061B	0621	
ا	ص	ض	ّ	غ	ح	چ	؁	س			آ
alef	sad	dad	fathatan	ghain	hah	tchch	kasratan	shadda			maddaonalef
0627	0635	0636	064B	063A	062D	0686	064D	0651			

Z	X	c	v	b	n	m	,	.	/
ز	خ	چ	ق	ب	ن	م	،	-	
zain	khah	tchch		bch	noon	meem	comma		
0632	062E		06A8	0628	0646	0645	060C	06D4	
ظ		ش	ث	ذ			إ	أ	؟
zah		sheen	thch	thal			hamzaunderalef	hamzaonalef	questionmark
0638			062B	0630			0625	0623	061F