Fonts for Africa: The fc Fonts

Jörg Knappen
Chair of the TWG on African Languages with Latin Writing

Abstract

A font (fc) with 256 characters for the needs of african languages with latin writing is presented. The so-called critical languages with more than 1 mio speakers are supported. It is implemented in METAFONT using Sauter's tools for the parametrisation.

1 Introduction

On the continent of africa, several hundred different languages are spoken, the numbers of speakers of each ranging from more than 100 mio. down to few hundreds. Many languages do not have a written form, other language have some written form, but it is not standardised yet. Or even worse: there exists more than one writing system for the same language (e.g. protestant and catholic orthography). Clearly, it was impossible to check out all writing systems of all possible languages. This work confines itself to the so-called 'critical languages' according to a definition of the US Department of Education in 1986. All these languages have more than 1 mio. speakers.
Three main writing systems are in use in Africa:

- *æthiopian* writing, used for several languages in Æthiopia and Eritrea.
- *arabic* writing, used in northern africa and on the east coast for several languages. It is now losing ground against latin writing; languages like suaheli and hausa are now mainly written in latin.
- *latin* writing is now being introduced or established in the rest of africa.

In this overview, I want to mention also some other writing systems. The Tuareg living in the sahara use the over 2,000 years old *tifinagh* writing system, a right-to-left writing; the Kopts in Ægypt still use *koptic* for religious purposes.

An interesting development was the invention of a somali alphabet by Osman Yusuf, which combined influences from the three major writing systems, but this system did not achieve official status in Somalia.

The introduction of latin writing was first done by missionaries, who invented orthographies quite arbitrarily. But since the first half of this century, there have been attempts to standardise the newly introduced alphabets.

As a result of these attempts, an 'african reference alphabet' emerged, based on the principle of *one sound, one sign* which is also the principle of the international phonetic alphabet. There are many borrowings from the phonetic alphabet in african writing. These efforts are now coordinated by the UNESCO.

2 The fc Fonts

The fc font encoding scheme encodes characters necessary to typeset the major african languages with latin writing. In selecting these languages, I followed the list of so-called critical languages. Unfortunately, it proved impossible to put really all characters into one 256-character font. Therefore, I applied the following selection rules:

- The standard TEX character set is still available.
- Characters with entirely new shapes are included with highest priority.
- Characters which need another accent in some languages have the second highest priority.
- Characters which are difficult to handle by TeX macros are preferred over those which are easier to handle (e.g. characters with diacritic below are preferred to those with diacritics above).
- Characters which occur only in tonal languages and carry a tonal mark, are most likely discarded.

The fc fonts have several other interesting features some of which I list here:

- The lower half of the fc encoding scheme is identical to the Cork (or ec) encoding scheme for european languages.
- If a character occurs both in the fc scheme and in the Cork scheme, it has the same encoding.
- The difference between uppercase and the corresponding lowercase character is a constant.
- It is possible to create virtual fonts to obtain all those characters of tonal languages which needed to be discarded.
- It is possible to rebuild the cm fonts and the ec fonts as virtual fonts from the fc fonts. In the latter case three letters are missing (edh, thorn and Thorn) and cannot be done without larger loss in quality. Of course, there are smaller quality losses in building such a virtual font, e.g. for î, where the ˆ-accent comes out too wide.

The following languages are supported (with the proviso about tones made above): Akan, Bamileke, Basa (Kru), Bemba, Ciokwe, Dinka, Dholuo (Luo), Efik, Ewe-Fon, Fulani (Fulful), Gã, Gbaya, Hausa, Ị̣gbọ, Kanuri, Kikuyu, Kikongo, Kpelle, Krio, Luba, Mandekan (Bambara), Mende, More, Ngala, Nyanja, Oromo, Rundi, Kinya Rwanda, Sango, Serer, Shona, Somali, Songhai, Sotho (two different writing systems), Suaheli, Tiv, Yao, Yoruba, Xhosa, and Zulu.

I decided to support two european languages, which are not covered by the ec-scheme, namely Maltese and Sami[1].

There are some writing systems which are not supported by the fc fonts. First to mention is the writing of Khoi-San languages with their characteristic click sounds. Missing is the letter ≠ which can be constructed as a macro. The obsolete orthographies of Xhosa and Zulu are not supported (they used a cyrillic B (Б) as uppercase of ɓ).

Also not supported is the obsolete orthography of Ịgbọ which contains an o with horizontal bar (ɵ). A proposed alphabet for Tamasheq (a berber language) is not supported because there is no evidence that it ever caught on.

3 Design of the letters

The design of the letters closely follows the *computer modern* fonts by Donald E. Knuth. Much of the METAFONT code made its way unchanged to the fc fonts. However, several modification were necessary in the case of italics. Since Eʋe has both 'f' and 'ƒ', the latter looking like the italic letter f, the italic letter f was redesigned to have a straight, uncurved descender (*f*, *ƒ*). Similarly, the italic letter v was redesigned to have sharp edge (*v*). In consequence of the last change, the italic letters w and y were changed, too (*w*, *y*).

Deliberately, I changed the appearance of the roman letter *scharfes s*. Its new shape exhibits the ligature of long and short s from which it originally derives (ß).

4 How to use the fc fonts

At the present time, I support only LaTeX and plain TeX with the new font selection scheme (NFSS) by R. Schöpf and F. Mittelbach. There are the files `fontdef.fc` and `fclfont.sty` which load the fonts and set the standard TeX commands for accents correctly. But, at the moment there are no standard control sequences to get the newly added characters, except that `\|` shall produce the universal accent (ʼ).

There is another style option (`fcuse.sty`) which makes the special characters accesible via active characters. Three characters needed to be activated,

Name	Shape	Input
Hooktop b	ɓ, Ɓ	`+b, +B`
Hooktop c	ƈ, Ƈ	`+c, +C`
Hooktop d	ɗ, Ɗ	`+d, +D`
Open e	ɛ, Ɛ	`+e, +E`
Long f	ƒ, Ƒ	`+f, +F`
Ipa gamma	ɣ, Ɣ	`+g, +G`
Latin iota	ɩ, Ɩ	`+i, +I`
Enj	ɲ, Ɲ	`+j, +J`
Hooktop k	ƙ, Ƙ	`+k, +K`
Eng	ŋ, Ŋ	`+n, +N`
Open o	ɔ, Ɔ	`+o, +O`
Hooktop p	ƥ, Ƥ	`+p, +P`
Esh	ʃ, Ʃ	`+s, +S`
Hooktop t	ƭ, Ƭ	`+t, +T`
Variant u	ʋ, Ʋ	`+u, +U`
Ezh	ʒ, Ʒ	`+z, +Z`
Crossed d	đ, Đ	`/d, /D`
Crossed h	ħ, Ħ	`/h, /H`
Crossed t	ŧ, Ŧ	`/t, /T`
Tailed d	ɖ, Ɖ	`=d, =D`
Inverted e	ə, Ǝ	`=e, =E`
Long t	ʈ, Ʈ	`=t, =T`

Table 1: Overview of the special letters in the fc fonts and their access via `fcuse.sty`

since the letters d and t carry the load of three new variants (ɗ, ɖ, đ; ƭ, ʈ, ŧ). An overview of the assignments is given in Table 1.

5 Implementation

The implementation follows closely to the one of the cm fonts. The distinction between parameter files, driver files and programme files is kept. The parameters are computed from the design size of the fonts with the help of the Sauter tools. This allows an easy nonlinear scaling. I added only one new parameter, `univ_acc_breath`, which governs the shape of the universal accent. To adjust the height of accented capitals the parameter `comma_depth` was employed. The result is excellent with serif fonts, but not that good with sans serif ones. I made up one new parameter file, creating a `sans serif typewriter` fontshape. There were only minor modifications needed in the programmes of the letters i, I, and l to make it work.

There is not much to say about the driver files, except that there are additional kerns for kV, kW,

[1] The Cork scheme includes the letter eng (ŋ) especially for Sami, but it is missing the letter t with bar (ŧ). However, with the letter eng several african languages can be written using the Cork scheme as well.

	x0	x1	x2	x3	x4	x5	x6	x7	x8	x9	xA	xB	xC	xD	xE	xF
0y	`	´	ˆ	˜	¨	˝	˚	ˇ	˘	¯	˙	¸	˛	,	‹	›
1y	“	”	„	«	»	–	—		0	ı	ȷ	ff	fi	fl	ffi	ffl
2y	␣	!	"	#	$	%	&	’	(	)	*	+	,	-	.	/
3y	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4y	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5y	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	ˆ	_
6y	‘	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7y	p	q	r	s	t	u	v	w	x	y	z	{	|	}	˜	-
8y	Ɓ	Ɗ	Ɛ	Ǝ	Ƒ	Ě	Ɣ	Ħ	Ƙ	Ɲ	Ɔ	Ń	Ʃ	Ŋ	Ʊ	Ƴ
9y	Ƈ	Ƥ	Š	Ṅ	Ṉ	Ṣ	Ʒ	Ŧ	Ė	Ẹ	Ƭ	Ʈ	ʧ	ʤ	đ	‟
Ay	ɓ	ɗ	ɛ	ə	ƒ	ě	ɣ	ħ	ƙ	ɲ	ɔ	ń	ʃ	ŋ	ʊ	ƴ
By	ƈ	ƥ	š	ṅ	ṉ	ṣ	ʒ	ŧ	ė	ẹ	ƭ	ʈ	"	¡	¿	'
Cy	Ɩ	Ị	Ɛ̃	Ã	Ḿ	Ɔ̃	Æ	Ç	È	É	Ê	Ë	E̱	Ē	Ẽ	Ĩ
Dy	Đ	Ñ	Ò	Ȯ	Ô	Õ	Ö	Œ	Ø	Ọ	O̱	Ō	Ǒ	Ụ	Ũ	‐
Ey	ɩ	ị	ɛ̃	ã	ḿ	ɔ̃	æ	ç	è	é	ê	ë	e̱	ē	ẽ	ĩ
Fy	ɖ	ñ	ò	ȯ	ô	õ	ö	œ	ø	ọ	o̱	ō	ǒ	ụ	ũ	ß

Table 2: The font fcr10

mV, mW, and eV. The ligatures from the Cork scheme to get german and french quotes are included, too. The programme files are completely reorganised. The huge files are split into smaller ones each containing only one to three letters of the alphabet. The shape of the letter is saved as a picture and only the diacritic is calculated to get an accented letter. This is done to save cpu time and to make the maintenance of the files easier. The accented letters are generated only on demand, i.e. if their code is known. The codes are given in the file `coding.fc`. It is possible to use the same source to create other real fonts by replacing the coding file. The so-created fonts may not be called fc, since this abbreviation is reserved to fonts which are fc-encoded.

6 Outlook

The fonts are done now, but there is lot of work left in creating suitable styles for african languages and virtual fonts, if needed. For this work, the speakers and users of the african languages are qualified best, and I hope that there will be some results of this work soon.

◇ Jörg Knappen
Chair of the TWG on African Languages with Latin Writing
Institut für Kernphysik
D 55 099 Mainz
R. F. A.
`knappen@vkpmzd.kph.uni-mainz.de`