

# The Coptic Standard Character Code (CSCC)

## Preview

This is a project of standardization of the Coptic Fonts that work on windows – format: \*.ttf – so, as to make a unified Character coding. It is done by Remenkimi (Coptic language e-mail discussion group, where most of correspondence take place in Coptic language).

## Summary

The first Coptic Fonts were created for ancient printers dating to 19<sup>th</sup> century then it developed to typewriters then to Electric Typewriters roughly in 1950. Computer fonts for DOS were as early as 1980s and then for windows. But their initiators had different needs like church text writings, manuscript re-typing, new compositions...etc. The fonts that were done were almost independent to each other separate in approach & arrangement due to lack of communications of different initiators (sometimes on opposite sides of the globe). All that caused that a total of 16 Coptic fonts are now on the market each distributed differently on the keyboard thus making a font change a disaster that would surely destroy the text on converting from one font to another. In addition, one has to have the font exactly in order to be capable of reading, that's unlike any font encoding system where morphology of the font is the only variable amongst various fonts of the same language.

Therefore, Remenkimi (Coptic Community) on the Internet (130 members) has undertaken this project of Font Code Standardization. A technical coordinator was elected and a Computer Engineer was the executor. A study was conducted to analyze how previous different fonts have been built and the standardization actions have been set towards a revival of Coptic in its different dialects in all fields of life.

The keyboard distribution was set within a participatory approach including polls and voice discussions for a period of 30 days before a trial version was issued, and 60 days to establish the final version. Comments and feedback from members enabled the computer engineer to finalize the layout.

We ask all new initiators of fonts to cooperate by accepting this new distribution to be capable to unify & standardize one font system. More details follow below:

### Definition:

This is a construction of a standard complete layout of the Coptic language on the standard keyboard.

### Aim:

Construct a unified system for placement of letters, djinkims, abbreviations, overlines, underlines, punctuation marks etc. on the keyboard. In the context, of coping with Coptic language as contemporary language, a competent keyboard to withstand all what's needed in day to-day work is demanded which is done here.

### Problem and reasons:

The main problem rose from the lack of a unified keyboard due to

1. Different goals on making keyboards, some had the intention of using it in ecclesiastical purposes, others to write manuscripts, others to scan manuscripts
2. Lack of communication among font makers, and lack of observation of each font.

Result:

1. There is no unified keyboard
2. Moving a document from a font to another font requires enormous effort in working with macros etc.
3. To be capable to read a document one has to have the font itself because not any other font can compensate the lack of original font
4. A typist has to learn a new keystroke system with every new font he installs.
5. Incapability of reading destruction of documents, sometimes is inevitable
6. No professional typists in the field.
7. To preserve a document one has to make it a picture file (\*.jpeg) which is impractical on production to scale, and on handling.
8. Lack of distribution of its software, and fonts render thinking about Coptic as a PC hostile fonts.
9. Lack of unity results in more separation, leading to disastrous incomprehensiveness font development

### Principles and strategies behind this standardization system:

1. Arrangement of letters on the keyboard on a pronunciation based system. This required sacrificing the morphologically based system & the standard system of placing keys that's by amount repetition. The pronunciation based system advantages are that it is almost standardized in the context that all other measurements are relative e.g. the morphologically based in one font may place theta on zero other may place it on q, tau may be on t ones opinion other see it closer to the j. As regards, the amount of repetition it was not used because, we do not have statistics for it; also, we do not have ready made keyboards with stickers for Coptic letters on keyboard. Moreover, most of typists are not of good knowledge of Coptic so it would be much easier for a typist to have a system that is close to what he knows in English. For letters that have no equivalent in English approximate pronunciation was applied e.g. shy on s (an s in a language is spelled as sh in the other, f is approximated to v , Soou is added on shift 6 because it means 6 etc.)
2. Adding djinkims & overlines of many types. They were all added to the shift first row due to frequency of repetition. Bohairic djinkim( ` ) key, was made free from m & n to allow ecclesiastical & educational typists to use it freely. While, classically it should be on m & n only.
3. The row of numbers is kept almost as it is because, many people believe in the application of Arabic numeral system (decimal system 0123456789) to Coptic in the context of revival, others see that it is a must to keep all mathematical marks in place for programs as MS Excel. This does not harm the font because the overlines made for typing classical Coptic numerals are available. Also, marks as these ( % \* !) are considered by some Copts & Copt linguists as neutral marks that should be applied to Coptic language, on handling the subject as a competent day-to-day language. This does not harm a classical typist because he can simply ignore using the line row and other keys would satisfy the needs for overlines, abbreviations, dots & djinkims of various types.

## Justifying placements on keyboard:

Djinkim on first key as its position is that high to letters. Keys of numbers are same to facilitate typing numbers. q = qhy, usually the letter qhy is written as khy some times khy. As the k is both morphologically and phonetically identical to the kappa so there is no place else for qhy except for q. w, e, r & t are omega, ei ro & tau phonetically, y is rendered to eeta for phonetic basis, u, i, o, p, are upsilon, iota, o, py for phonetics, chima [, ti ] because, they are after each other on alphabet as well this placement is borrowed from other older versions of fonts to which no objections were raised, in addition to that the chima should be closer to a typist because of high incidence of repetition in Sahidic a = alpha, s = shy (sima was not placed because it is both phonetically & morphologically closer to C, while in some Languages as German they spell stern, strauss as shtern, shtrauss, that is close to the shy. f, g, h, j, k, l, are fy, gamma, hoary, djendja, kappa, laula, Theta on ; while psy on ' as theta is used in Coptic & Greek words while psy is used in Greek words mainly that makes theta more recurrent. z, x, c, for zita, ksy, cima, v had a phey on it because it is phonetically close to it (not identical by any mean) also, classical Ecclesiastical pronunciation offered by Arian (3arian) effendy Guirguis Moftah, indicated that phey is v. Though proved to be wrong later on, but gave a hint about a letter that is close to ph ( in Coptic societies one day) b , n , m, for beta nei & mei (,) for chey, and (.) for so as chey is more repetitive.

## Conclusion

Following this arrangement one can type all Coptic Letters (including letters in Achmimic Dialect) in almost a totally pronunciation based system that facilitates both typing & education. Also, all types of djinkims Bohairic, Sahidic, Achmimic could be written. Arabic numerals, simple mathematical marks & modern punctuations are present. Ecclesiastical common abbreviations are present on Alt keys.

First row lower case

`	1	2	3	4	5	6	7	8	9	0	-	=	\	backspace
`	1	2	3	4	5	6	7	8	9	0	-	overline	Ⲅ	backspace

First row upper case

~	!	@	#	\$	%	^	&	*	(	)	_	+		backspace
underdot	!	:	#	\$	%	Ⲛ	;	*	(	)	=	+	ⲉ	backspace

Second row lower case

tab	q	w	e	r	t	y	u	i	o	p	[	]	Enter
tab	Ⲛ	ⲟ	ⲉ	ⲣ	ⲧ	Ⲏ	ⲩ	ⲓ	ⲟ	ⲡ	Ⲅ	ⲧ	Enter

Second row upper case

tab	Q	W	E	R	T	Y	U	I	O	P	{	}	Enter
-----	---	---	---	---	---	---	---	---	---	---	---	---	-------

tab	<b>ⲃ</b>	<b>Ⲙ</b>	<b>Ⲏ</b>	<b>Ⲡ</b>	<b>Ⲫ</b>	<b>Ⲭ</b>	<b>Ⲯ</b>	<b>Ⲱ</b>	<b>Ⲳ</b>	<b>Ⲵ</b>	<b>Ⲷ</b>	<b>Ⲹ</b>	Enter
-----	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-------

Third row lower case

caps lock	a	s	d	f	g	h	j	k	l	;	'	Enter
caps lock	<b>ⲁ</b>	<b>ⲥ</b>	<b>ⲉ</b>	<b>ⲋ</b>	<b>ⲍ</b>	<b>ⲏ</b>	<b>ⲓ</b>	<b>ⲑ</b>	<b>ⲇ</b>	<b>ⲕ</b>	<b>ⲓ</b>	Enter

Third row upper case

caps lock	A	S	D	F	G	H	J	K	L	:	"	Enter
caps lock	<b>Ⲁ</b>	<b>Ⲥ</b>	<b>Ⲅ</b>	<b>Ⲇ</b>	<b>ⲇ</b>	<b>Ⲉ</b>	<b>Ⲋ</b>	<b>ⲋ</b>	<b>Ⲍ</b>	<b>Ⲏ</b>	<b>ⲏ</b>	Enter

Fourth row lower case

Shift	z	x	c	v	b	n	m	,	.	/	Shift
Shift	<b>ⲗ</b>	<b>ⲙ</b>	<b>Ⲙ</b>	<b>ⲛ</b>	<b>ⲏ</b>	<b>ⲓ</b>	<b>ⲑ</b>	<b>ⲕ</b>	<b>ⲉ</b> period	<b>ⲇ</b>	Shift

Fourth row Upper case

Shift	Z	X	C	V	B	N	M	<	>	?	Shift
Shift	<b>Ⲙ</b>	<b>ⲙ</b>	<b>Ⲙ</b>	<b>ⲛ</b>	<b>ⲏ</b>	<b>ⲓ</b>	<b>ⲑ</b>	<b>ⲕ</b>	<b>ⲉ</b> comma	<b>ⲇ</b>	Shift

0128	0129
	<b>ⲛ</b>



	<	[	]	®					
0180	0181	0182	0183	0184	0185	0186	0187	0188	0189
		Ⲛ			//		>	†	Ⲛ
0190	0191	0192	0193	0194	0195	0196	0197	0198	0199
ⲛ	+								\$
0200	0201	0202	0203	0204	0205	0206	0207	0208	0209
/	.	,	=	=	-	@	\$	^	&
0210	0211	0212	0213	0214	0215	0216	0217	0218	0219

