

# L'arte esoterica di scrivere in cirillico con L<sup>A</sup>T<sub>E</sub>X

Enrico Gregorio

## Sommario

La scrittura di parole in caratteri cirillici con L<sup>A</sup>T<sub>E</sub>X è facile una volta che si conosce qualche piccolo trucco del mestiere. Con **babel** è poi facile scrivere documenti con lunghi brani in una lingua che usi l'alfabeto cirillico. Nell'articolo si descrivono anche alcuni difetti di **babel** al riguardo e modi per correggerli.

## Abstract

Writing words in the cyrillic script with L<sup>A</sup>T<sub>E</sub>X is easy once we know some small tricks of the trade. With **babel** it's also easy to write documents with longer parts in a language using the cyrillic script. We describe also some small defects of **babel** in this area and some ways to correct them.

## 1 Introduzione

L'alfabeto cirillico deriva dagli alfabeti greco e glagolitico, che si dice inventato dai fratelli Cirillo e Metodio nel nono secolo nell'ambito dell'evangelizzazione delle popolazioni slave del sud (Bulgaria). Lo stesso alfabeto glagolitico originale ebbe un'evoluzione a parte tra le popolazioni croate dove, a causa della dominazione ungherese, fu soppiantato dall'alfabeto latino. Si veda in figura 1 la forma dei caratteri glagolitici e si consulti l'articolo di D. Žubrinić (ŽUBRINIĆ, 1996) per maggiori notizie su questo alfabeto.

Le forme dei caratteri cirillici ricordano quelle dell'alfabeto greco di Bisanzio per i suoni che hanno un corrispondente; altri sono presi quasi direttamente dal glagolitico. Dalla Bulgaria l'alfabeto si diffuse in molte nazioni slave, diventando la forma comune di scrittura in quelle di religione ortodossa. Con l'espandersi dell'impero russo l'alfabeto fu adattato alle lingue dei paesi sottomessi, soprattutto in Asia. Con la dissoluzione dell'Unione Sovietica, alcuni paesi riadottarono altri alfabeti; è il caso della Moldova, dove la lingua è una variante del romeno e fino al 1991 veniva scritta con l'alfabeto cirillico. Altre lingue, come il lituano, il lettone e l'estone, avevano conservato l'alfabeto latino; in Georgia e in Armenia continuavano a essere usati gli alfabeti tradizionali. Non è il caso di addentrarsi nei meandri di una storia complicata anche da questioni politiche; per esempio, fino alla morte di Tito l'alfabeto usato dai serbi in Jugoslavia era principalmente quello latino, e solo nella repubblica federata di Macedonia era usato l'alfabeto cirillico. Le note vicende prima e

dopo la cruenta dissoluzione della Jugoslavia hanno riportato in auge in Serbia l'alfabeto cirillico.

Nel seguito useremo alcune convenzioni per le traslitterazioni: quella cosiddetta 'scientifica', basata sull'ortografia del ceco, sarà indicata con questo tipo di **carattere**; quella 'internazionale', basata sull'inglese, sarà indicata /così/. Per esempio, il nome russo Горбачёв può essere traslitterato come **Gorbačëv** oppure come /Gorbachev/. Non è necessario usare modi particolari per le parole in alfabeto cirillico che saranno ben distinguibili nel contesto latino, lettere isolate saranno tra virgolette singole. Nel caso del serbo, essendo disponibile una conversione ufficiale da caratteri cirillici a latini, useremo sempre questo **carattere**, ma con quella convenzione.<sup>1</sup>

L'articolo è orientato verso l'utente italiano che abbia bisogno di adoperare caratteri cirillici. Esistono pacchetti sviluppati in Russia per facilitare l'impiego di L<sup>A</sup>T<sub>E</sub>X in ambiente completamente cirillico. Si ringrazia Lorenzo Pantieri per aver stimolato la scrittura di questo articolo (PANTIERI).

## 2 Gli alfabeti cirillici

In realtà non esiste *un* alfabeto cirillico, ma molti (PIŠKA, 1996). Limitandoci alle lingue slave, ciascuna di esse ha distintive particolarità fonetiche e per ciascuna occorrono lettere che ad altre non servono. La faccenda è ancora più complicata da fattori storici; in russo le lettere usate fino al 1918 erano più delle attuali: per motivi etimologici, alcuni suoni erano rappresentati dalla stessa lettera. Un classico esempio è il nome Θεодоръ che dopo la riforma ortografica seguita alla rivoluzione venne scritto Фёдор. Il nome traslitterato è 'Fëdor' che corrisponde a 'Teodoro': il suono simile a quello della lettera greca Θ cambiò pronuncia fino a confondersi con il nostro suono 'F'.<sup>2</sup> La lettera cirillica 'Ө' viene chiamata 'fita'. Le altre lettere cadute sono 'V' (ižica), che discende dalla Υ greca, sostituita con 'И', e 'Ѣ' (ât') sostituita con 'Е' o con 'И'. Fu introdotta anche la lettera 'Ё', che però attualmente è spesso scritta semplicemente come 'Е' (ma pronunciata diversamente, jo). Anche la

1. Esistono altre traslitterazioni, quella scientifica, più propriamente ISO-9:1995, è ufficiale nella Federazione Russa. In una versione precedente le traslitterazioni di И, Ю e Я erano šč, ju e ja.

2. In bulgaro, invece, la pronuncia divenne simile alla nostra Т, ciò che spiega il corrispondente nome Тодор, Todor.

lettera ‘I’ (i *desâteričnoe*<sup>3</sup>) venne eliminata perché inutile; è interessante notare come la lettera ‘И’ sembri la legatura di due ‘I’, ma invece derivi dalla lettera greca ‘eta’.

Nella tabella 2 vediamo l’alfabeto russo attuale. L’alfabeto bulgaro è molto simile, ma mancano le lettere ‘Ė’ e ‘Ѓ’. Nella tabella 3 invece usiamo la latinizzazione ufficiale per l’alfabeto serbo.

Come si vede solo esaminando gli alfabeti, il russo e il serbo hanno differenze fonetiche sostanziali. Si deve notare come il serbo abbia una scrittura quasi completamente fonetica, mentre in russo la stessa lettera può avere pronunce diverse; l’esempio delle desinenze del genitivo degli aggettivi come -oro pronunciate come -ovo è classico, così come quello delle trasformazioni di consonanti sorde in sonore e viceversa. Il serbo ha invece una corrispondenza quasi esatta fra lettere e suoni, dovuta al fatto che l’alfabeto nella forma attuale, sviluppata da Vuk Karadžić (Вук Караџић), fu adottato nel 1868.

La stessa lettera cirillica può indicare suoni diversi in varie lingue: in ucraino la ‘Г’ corrisponde a un’aspirazione meno forte di ‘X’; in bulgaro la lettera ‘Б’ denota una semivocale, presente nel nome България, mentre in russo non ha suono e indica solo una maggiore marcatura della sillaba in cui compare; la lettera ‘Ч’ del russo corrisponde più alla ‘Т’ che alla ‘Ч’ del serbo. La lettera ‘Е’ che in russo è pronunciata ‘ie’ come in ‘ieri’ ma in serbo e ucraino è una ‘e’ come in ‘eri’.

Nella tabella 4 è riportato l’alfabeto ucraino dopo le recenti riforme ortografiche seguite all’indi-

pendenza. Per terminare questa veloce scorsa delle maggiori lingue che usano l’alfabeto cirillico, notiamo che l’alfabeto bielorusso è quasi uguale al russo: la lettera ‘И’ diventa ‘І’; esiste la ‘Ў’ che corrisponde alla ‘u’ semivocalica; mancano poi la ‘ІІ’ e la ‘Б’. L’alfabeto macedone, introdotto ufficialmente solo nel 1944, invece contiene parecchi caratteri diversi sia dal serbo che dal bulgaro (tabella 5).

### 3 Nomi e pronuncia delle lettere

I bambini delle scuole russe imparavano l’alfabeto associando a ciascuna lettera una parola comune o, in alcuni casi, il suono proprio; si veda la tabella 6, dove si usano le lettere soppresse dalla riforma del 1918. Non sono riuscito a trovare l’ordine alfabetico valido prima della riforma, perciò le lettere soppresse sono aggiunte in fondo; altre lettere erano state eliminate nel diciottesimo secolo.

Al giorno d’oggi le lettere hanno un nome che rispecchia solo il suono; dai nomi tradizionali deriva la parola азбука per indicare l’alfabeto. Nella tabella 7 si trovano i nomi moderni. Si noti che la translitterazione corretta della pronuncia dell’acronimo КГБ (Комитет государственной безопасности), cioè il famigerato KGB, è *kagèbè*.<sup>4</sup>

Non è qui il caso di addentrarci nei problemi della pronuncia delle lingue slave che, fra l’altro, presentano fenomeni molto diversi tra l’una e l’altra. In russo si ha la palatalizzazione delle consonanti davanti a vocali deboli: la ‘Т’ di табак (tabacco) si pronuncia in modo diverso dalla ‘Т’ di тело (corpo), perché la vocale che segue è forte nella prima

3. Il nome significa “i decimale”: nel sistema numerico derivato da quello greco assumeva il valore 10.

4. In ortografia italiana è dunque ‘caghebè’ e non il ridicolo ‘cheghebé’ che si sente spesso.

TABELLA 1: Caratteri dell’alfabeto glagolitico

	‘0	‘1	‘2	‘3	‘4	‘5	‘6	‘7	
‘00x	Ѡ 0	ѡ 1	Ѣ 2	ѣ 3	Ѥ 4	ѥ 5	Ѧ 6	ѧ 7	“0x
‘01x	8	Ѩ 9	10	Ѭ 11	ѭ 12	Ѯ 13	ѯ 14	15	
‘02x	16	Ѱ 17	18	Ѳ 19	ѳ 20	Ѵ 21	ѵ 22	23	“1x
‘03x	24	25	26	27	28	29	30	31	
‘04x	32	! 33	34	35	36	37	38	’ 39	“2x
‘05x	( 40	) 41	• 42	+ 43	, 44	- 45	. 46	47	
‘06x	48	Ѧ 49	ѧ 50	51	52	53	Ѐ 54	55	“3x
‘07x	Ѣ 56	ѣ 57	˙ 58	˚ 59	60	= 61	62	? 63	
‘10x	64	Ѧ 65	ѧ 66	Ѩ 67	ѩ 68	Ѫ 69	ѫ 70	Ѭ 71	“4x
‘11x	ѭ 72	Ѯ 73	ѯ 74	Ѱ 75	ѱ 76	Ѳ 77	ѳ 78	Ѵ 79	
‘12x	ѵ 80	Ѷ 81	ѷ 82	Ѹ 83	ѹ 84	Ѻ 85	ѻ 86	87	“5x
‘13x	— 88	Ѵ 89	ѵ 90	91	92	93	Ѐ 94	95	
‘14x	’ 96	Ѧ 97	ѧ 98	Ѩ 99	ѩ 100	Ѫ 101	ѫ 102	Ѭ 103	“6x
‘15x	ѭ 104	Ѯ 105	ѯ 106	Ѱ 107	ѱ 108	Ѳ 109	ѳ 110	Ѵ 111	
‘16x	ѵ 112	Ѷ 113	ѷ 114	Ѹ 115	ѹ 116	Ѻ 117	ѻ 118	119	“7x
‘17x	— 120	Ѵ 121	ѵ 122	— 123	124	125	126	127	
	“8	“9	“A	“B	“C	“D	“E	“F	

TABELLA 2: Alfabeto russo

Aa	a	Бб	b	Вв	v	Гг	g	Дд	d	Ее	e
Ёё	ë	Жж	ž	Зз	z	Ии	i	Йй	j	Кк	k
Лл	l	Мм	m	Нн	n	Оо	o	Пп	p	Рр	r
Сс	s	Тт	t	Уу	u	Фф	f	Хх	h	Цц	c
Чч	č	Шш	š	Щщ	š	Ъъ	"	Ыы	y	Ьь	'
Ээ	è	Юю	û	Яя	â						

TABELLA 3: Alfabeto serbo

Aa	a	Бб	b	Вв	v	Гг	g	Дд	d	Ђђ	đ
Ee	e	Жж	ž	Зз	z	Ии	i	Јј	j	Кк	k
Лл	l	Љљ	lj	Мм	m	Нн	n	Њњ	nj	Оо	o
Пп	p	Рр	r	Сс	s	Тт	t	Ћћ	ć	Уу	y
Фф	f	Хх	h	Цц	c	Чч	č	Џџ	dž	Шш	š

TABELLA 4: Alfabeto ucraino

Aa	a	Бб	b	Вв	v	Гг	g	Ґґ	ğ	Дд	d
Ee	e	Єє	ê	Жж	ž	Зз	z	Ии	i	Іі	ì
Її	ï	Йй	j	Кк	k	Лл	l	Мм	m	Нн	n
Оо	o	Пп	p	Рр	r	Сс	s	Тт	t	Уу	u
Фф	f	Хх	h	Цц	c	Чч	č	Шш	š	Щщ	š
Ьь	'	Юю	û	Яя	â						

TABELLA 5: Alfabeto macedone

Aa	a	Бб	b	Вв	v	Гг	g	Дд	d	Ѓѓ	ǵ
Ee	e	Жж	ž	Зз	z	Сс	dz	Ии	y	Јј	j
Кк	k	Лл	l	Љљ	lj	Мм	m	Нн	n	Њњ	nj
Оо	o	Пп	p	Рр	r	Сс	s	Тт	t	Ќќ	ќ
Уу	u	Фф	f	Хх	h	Цц	c	Чч	č	Џџ	dž
Шш	š										

parola, debole nella seconda. Il fenomeno è molto meno accentuato in serbo, per esempio, dove invece è più rilevante che in russo la differenza tra vocali lunghe e brevi.

## 4 Scrivere in cirillico con L<sup>A</sup>T<sub>E</sub>X

Da quanto visto nelle pagine precedenti, è evidente che si possono usare caratteri cirillici in L<sup>A</sup>T<sub>E</sub>X. L'American Mathematical Society, nell'ambito del progetto di conversione in T<sub>E</sub>X dei contenuti dei "Mathematical Reviews" commissionò alla University of Washington un font cirillico compatibile con la famiglia Computer Modern disegnata da Knuth.

La particolarità di questo font è che, tramite ben studiate legature, è possibile scrivere parole russe mediante la loro traslitterazione 'americana': per esempio, *Shafarevich* produce, quando si usa il font cirillico, *Шафаревич*. Tuttavia questo sistema di traslitterazione è alquanto rozzo: adatto forse ai giornali, perché i lettori possono ottenere una pronuncia vagamente simile a quella russa leggendo come se fosse inglese. Forse meglio del sistema in uso da noi: il signor Krusciov sarebbe

/Khrushchev/ che, in caratteri cirillici, è *Хрушев*. Il nome corretto sarebbe *Хрущёв*, ma di fatto la distinzione fra 'e' e 'ë' è spesso assente nella stampa russa.

Esiste un modo molto semplice per attivare i caratteri cirillici con il sistema dell'AMS:

```
\usepackage[OT2,OT1]{fontenc}
\makeatletter
\ifundefined{1@nohyphenation}
{\chardef1@nohyphenation\ccclv}{ }
\newcommand{\cyrillic}{%
\language=1@nohyphenation
\fontencoding{OT2}%
\fontfamily{wncyr}%
\selectfont}
\makeatother
\DeclareTextFontCommand
{\textcyrillic}{\cyrillic}
```

Possiamo naturalmente scegliere T1 invece che OT1 come codifica normale per il documento. Scrivendo `\textcyrillic{Shafarevich}` otteniamo *Шафаревич*, e abbiamo a disposizione la dichiarazione `\cyrillic` per testi più lunghi. L'insolito co-

TABELLA 6: Nomi tradizionali delle lettere in russo

Аа азъ	Бб буки	Вв вѣди	Гг глаголь
Дд добро	Ее есть	Жж живѣте	Зз земля
Ии иже	Йй и съ краткой	Кк како	Лл люди
Мм мыслѣте	Нн нашъ	Оо онъ	Пп покой
Рр рцы	Сс слово	Тт твердо	Уу укъ
Фф фертъ	Хх хѣръ	Цц цы	Чч червь
Шш ша	Щщ ща	Ъъ еръ	Ыы еры
Ьь ерь	Ээ э оборотное	Юю ю	Яя я
Өө өита	Іі і десятиричное	Ѳѳ ять	Ѵѵ ижица

TABELLA 7: Nomi delle lettere in russo

Аа а	Бб бэ	Вв вэ
Гг гэ	Дд дэ	Ее е
Ёё ё	Жж жэ	Зз зэ
Ии и	Йй и краткое	Кк ка
Лл эл	Мм эм	Нн эн
Оо о	Пп пэ	Рр эр
Сс эс	Тт тэ	Уу у
Фф эф	Хх ха	Цц це
Чч че	Шш ша	Щщ ща
Ъъ твёрдый знак	Ыы ы (еры)	Ьь мягкий знак
Ээ э	Юю ю	Яя я

dice che riguarda `\language` serve per disabilitare la sillabazione.

In matematica si usa talvolta il simbolo ‘Ш’ per indicare il gruppo di Tate-Shafarevich. Chi ne avesse bisogno può definire il simbolo con

```
\newcommand{\Sha}{\mbox{%
  \usefont{OT2}{wncyr}{m}{n}Sh}}
```

oppure, caricando il pacchetto `amstext`,

```
\newcommand{\Sha}{\text{%
  \usefont{OT2}{wncyr}{m}{n}Sh}}
```

Si noti che `amstext` è caricato automaticamente da `amsmath`; con questa seconda definizione, il simbolo può essere usato anche in pedici ed esponenti.

Dalla tabella 8 si può ricavare il modo per inserire lettere che non sono immediatamente disponibili con la traslitterazione americana. Per esempio il nome Пафнутий Львович Чебышёв può essere scritto

```
\textcyrillic{Pafnutii0
  Lp1vovich Chebyshev}
```

usando il comando definito precedentemente. Si tratta di un nome che si trova scritto in varie forme, spesso secondo l’ortografia francese o tedesca; la traslitterazione più appropriata sarebbe Čebyšëv oppure, data l’ormai generale diffusione dell’inglese in questo contesto, /Chebyshev/.

I caratteri della University of Washington sono disponibili in varie forme: diritto, corsivo, nero diritto e diritto senza grazie:

АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

*АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя*

**АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя**

АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

Sarebbe disponibile anche il maiuscoletto, ma nella tradizione tipografica russa non è usato perché la forma di molte lettere minuscole non differisce da quella delle maiuscole. Si noti come il corsivo abbia forme molto diverse dal tondo.

È possibile, con altre legature, inserire anche caratteri propri di altre lingue; la lista completa è nella tabella 9 in appendice. Se si dovessero usare le lettere prerivoluzionarie  $Vv$   $\Theta\theta$   $\mathring{B}\mathring{b}$  è conveniente dare le seguenti definizioni:

```
\chardef\Izhitsa='24
\chardef\izhitsa='34
\chardef\Fita='25
\chardef\fita='35
\chardef\Yat='43
\chardef\yat='53
```

in modo da poter scrivere il nome  $\Theta$ едоръ come `\textcyrillic{\Fita edorp2}` e analogamente

TABELLA 8: Corrispondenze dei caratteri

A	A	K	K	X	Kh	a	a	к	к	х	kh
Б	B	Л	L	П	Ts	б	b	л	l	ц	ts
В	V	М	M	Ч	Ch	в	v	м	m	ч	ch
Г	G	Н	N	Ш	Sh	г	g	н	n	ш	sh
Д	D	О	O	Щ	Shch	д	d	о	o	щ	shch
Е	E	П	P	Ъ	P2	е	e	п	p	ъ	p2
Ё	E0	Р	R	Ы	Y	ё	e0	р	r	ы	y
Ж	Zh	С	S	Ь	P1	ж	zh	с	s	ь	p1
З	Z	Т	T	Э	E1	з	z	т	t	э	e1
И	I	У	U	Ю	Yu	и	i	у	u	ю	yu
Й	I0	Ф	F	Я	Ya	й	i0	ф	f	я	ya

per le altre lettere.<sup>5</sup> Come notato al piede della tabella, è talvolta necessario imporre a TEX di non considerare una legatura, per esempio nella parola Советский; il modo più semplice è di scrivere `\textcyrillic{Sovet\skii0}`. Senza il comando `\/` il risultato sarebbe lo scorretto Совецкий.

Non è necessario, di fatto, richiedere i font della University of Washington; ogni distribuzione sufficientemente ricca di TEX possiede le estensioni cirilliche della famiglia Computer Modern che comprendono i font nella codifica OT2. Perciò la definizione di `\cyrillic` potrebbe diventare

```
\newcommand{\cyrillic}{%
  \language=\l@nohyphenation
  \fontencoding{OT2}\selectfont}
```

che funziona anche con (poche) altre famiglie di font. Si può vedere la differenza:

University of Washington:  
АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

Cyrillic Computer Modern:  
АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

Antikwa Toruńska:  
АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

Iwona:  
АБВГДЕЁЖЗИЙКЛМНОП  
РСТУФХЦЧШЩЪЫЬЭЮЯ  
абвгдеёжзийклмнопрстуфхцчшщъыьэюя

Si noti però che i font Cyrillic Computer Modern nella codifica OT2 sono solo in formato *bitmap*, mentre quelli della University of Washington sono anche in formato vettoriale.

## 5 Unicode

Con l'introduzione di Unicode si è avuta la possibilità di usare caratteri cirillici in modo facile

5. Per la Ъ Ь sarebbero disponibili le abbreviazioni `\#` e `+`, ma sono poco intuitive.

nei documenti LATEX. In passato si era costretti ad acrobazie con codifiche di input strane, mentre ora il formato di input detto UTF-8 permette di usare una sola codifica per documenti che usano caratteri diversi.

Il trucco fondamentale è dichiarare nel proprio documento le codifiche di output per il cirillico *prima* di caricare il pacchetto `inputenc`:

```
\usepackage[T2A,T1]{fontenc}
\usepackage[utf8]{inputenc}
```

Sono disponibili tre codifiche di output per i caratteri cirillici: T2A, T2B e T2C; ciascuna copre un certo spettro di lingue che si scrivono con varie estensioni dell'alfabeto cirillico. Per le lingue slave usate in Europa la codifica T2A è sufficiente. Va ovviamente controllato che la famiglia di caratteri del nostro documento contenga il supporto per i caratteri cirillici. Non è il caso della famiglia Latin Modern, per esempio.

Il pacchetto `babel` fornisce i moduli per il bulgaro, il russo e l'ucraino.<sup>6</sup> Le distribuzioni più recenti di TEX, in particolare TEX Live 2009, rendono disponibili gli schemi di sillabazione per le stesse tre lingue.

Ecco un documento completamente in russo:

```
\documentclass[a4paper]{article}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[russian]{babel}

\begin{document}

Пафнутий Львович Чебышёв

\end{document}
```

Se si usa PDFLATEX e la distribuzione comprende i font CM-Super, il file PDF risultante sarà anche *searchable*, nel senso che i caratteri cirillici sono 'loro stessi'.

Naturalmente è possibile scrivere documenti in più lingue, sfruttando le possibilità offerte da `babel`:

6. Il modulo per il serbo è solo per la scrittura con l'alfabeto latino.

```
\documentclass[a4paper]{article}
\usepackage[T2A,T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[russian,italian]{babel}
\newcommand{\Sha}{\mbox{%
\usefont{T2A}{\rmdefault}{m}{n}III}}
```

```
\begin{document}
```

Bruno De Finetti

```
\begin{otherlanguage*}{russian}
Пафнутий Львович Чебышёв
\end{otherlanguage*}
```

```
Il gruppo $\Sha$ di Tate-%
\foreignlanguage{russian}{Шафаревич}.
```

```
\end{document}
```

Il risultato sarà simile a quello che si vede nella figura 1.

Si può notare come ci siano alcune limitazioni: il passaggio dai caratteri latini a quelli cirillici *deve* essere marcato con uno dei comandi di `babel`, in modo che  $\LaTeX$  si adegui alla richiesta. Non è possibile, quando la codifica principale (quella che appare per ultima nel caricamento di `fontenc`) è T1 scrivere direttamente caratteri cirillici. Viceversa, in un contesto cirillico, cioè quando è attiva la codifica T2A (oppure T2B o T2C), si possono scrivere lettere latine. Infatti le codifiche T2A e T1 differiscono solo per i caratteri con posizione da 128 a 255.

Il comando `\textcyrillic` è comodo, ma potrebbe non essere adatto nel caso in cui il documento sia scritto in più lingue che adoperano caratteri cirillici; infatti questo comando sceglierà le convenzioni, in particolare gli schemi di sillabazione, dell'ultima lingua che compare fra le opzioni a `babel`. Per chiarire, supponiamo che nel preambolo ci sia

```
\usepackage[russian,ukrainian,
italian]{babel}
```

Allora la lingua principale è l'italiano, mentre il comando `\textcyrillic{...}` equivale a

```
\foreignlanguage{ukrainian}{...}
```

Lo stesso discorso vale per `\textcyrillictext`, dichiarazione che nell'esempio attiva una situazione simile a quanto avviene nell'ambiente `otherlanguage*` con argomento `ukrainian`. Le regole precise su questi comandi saranno date più avanti insieme a un modo per avviare all'inghippo.

La situazione è comoda per chi scriva un documento, per esempio, in inglese e russo: basta scegliere la codifica T2A e scrivere senza dare al-

cun comando particolare.<sup>7</sup> La marcatura con la scelta della lingua però è comunque necessaria se si vuole un risultato composto correttamente con le cesure secondo le regole di ciascuna lingua. Se le lingue sono inglese e russo, con

```
\usepackage[russian,english]{babel}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
```

si avrà la corretta sillabazione in inglese, mentre le parole russe non saranno sottoposte alla cesura, dal momento che i caratteri di codice maggiore di 127 non corrispondono a schemi presenti in quelli per l'inglese; parole russe isolate non darebbero problemi di sillabazione scorretta, per brani più ampi si possono adoperare gli ambienti di `babel`. Questa caratteristica non vale se la lingua in caratteri latini non è l'inglese, perché gli schemi di sillabazione potrebbero contenere caratteri con codice maggiore di 127. Esamineremo a fondo la faccenda nelle prossime sezioni.

Vediamo ora come, disponendo della possibilità di inserire caratteri Unicode, si può risolvere il problema di scrivere parole isolate in una lingua che usa il cirillico senza necessariamente adoperare il macchinario di `babel`. Come prima, apriremo il preambolo con

```
\documentclass[a4paper]{article}
\usepackage[T2A,T1]{fontenc}
\usepackage[utf8]{inputenc}
```

e daremo qualche definizione, per esempio:

```
\newcommand{\cyrbox}[1]{%
\mbox{\cyrfont#1}}
\newcommand{\cyrfont}{%
\fontencoding{T2A}\selectfont}
```

Una breve tabella di corrispondenza di parole può essere inserita allora con

```
\begin{tabular}{>\cyrfont}ll}
\toprule
russo & italiano \\
\midrule
палец & pollice \\
сад & giardino \\
успех & successo \\
\bottomrule
\end{tabular}
```

che, composta,<sup>8</sup> darebbe

russo	italiano
город	città
палец	pollice
сад	giardino
успех	successo

7. Non ci sono problemi a usare un numero qualsiasi di lingue purché solo una usi i caratteri cirillici.

8. Sono necessari i pacchetti `booktabs` e `array`.

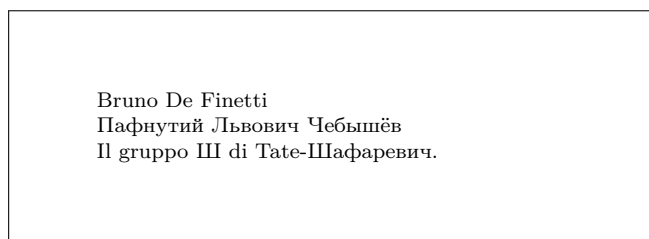


FIGURA 1: Esempio di documento in più lingue

Si noti che non occorre prendere precauzioni per l'intestazione della prima colonna, in questo caso, perché la codifica T2A prevede che i caratteri della prima metà coincidano con quelli della codifica T1. Ci sono anche gli accenti più comuni. Con `\multicolumn{1}{1}{...}` si possono risolvere eventuali problemi. Nel caso volessimo scrivere una parola isolata, potremmo usare `\cyrbox`:

Il nome `'\cyrbox{Иван}'` è  
'Giovanni'.

per scrivere che il nome 'Иван' è 'Giovanni'.

## 6 Documenti in una sola lingua

La parte più facile è appunto la scrittura di un documento in una sola lingua; ci accontenteremo di quelle per le quali è disponibile un modulo di `babel`, cioè bulgaro, russo e ucraino. Alla fine parleremo del serbo.

La prima cosa da ricordare è di impostare il proprio programma di scrittura affinché usi la codifica UTF-8. In alcuni programmi è possibile farlo su un singolo documento, in altri occorre impostare una preferenza globale. Per esempio, con TEXShop o TEXWorks basta che tra le prime righe del documento ci sia

```
%!TEX encoding = UTF-8 Unicode
```

Per Emacs (in qualsiasi incarnazione), si può impostare la codifica UTF-8 cominciando il documento con la riga

```
% -*- coding: utf-8 -*-
```

Si consulti il manuale del proprio programma se non è fra questi. Nel seguito fingeremo che il documento sia scritto con TEXWorks e che la lingua del documento sia il bulgaro; useremo la classe `article`. Nella figura 2 troviamo il sorgente e il risultato di un documento minimale.

Il testo è l'inizio della voce 'Sofia' nella pagina bulgara di Wikipedia; il nome dell'autore è inventato, si tratta di un nome piuttosto comune in Bulgaria. Le modifiche per il russo o l'ucraino sono ovvie.

Per quanto riguarda il serbo, non esiste un modulo ufficiale per la scrittura in caratteri cirillici; tuttavia caricando

```
\usepackage[...serbian,...]{babel}
\input{cyrserbian}
```

dove il contenuto del file `cyrserbian.tex` sia come in figura 4 in appendice, si dovrebbero sistemare almeno gli aspetti principali della scrittura serba, anche se la correttezza della sillabazione non è garantita.

## 7 Documenti in più lingue, come correggere un difetto di babel

Non ci sono grandi difficoltà per un documento in italiano con inserti in ucraino:

```
%!TEX encoding = UTF-8 Unicode
\documentclass[a4paper]{article}
\usepackage[T2A,T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ukrainian,italian]{babel}
```

In questo caso il comando `\textcyrillic` punterà all'ucraino. Diventa disponibile anche `\textlatin` per inserti in caratteri latini in contesto cirillico; si tenga però presente che l'argomento di questo comando continuerà a seguire le regole di sillabazione dell'ucraino e non saranno disponibili i comandi di `babel` specifici per l'italiano. Dunque `\textlatin` ha un'utilità molto relativa.

Il comando `\Ukrainian` è una dichiarazione che può tornare molto utile per definire un ambiente di citazioni in ucraino, per esempio

```
\newenvironment{quotationUA}
{\begin{quotation}\Ukrainian}
{\end{quotation}}
```

Analogamente potrebbe essere indicato, nel caso la lingua principale sia l'ucraino, con inserti in italiano, definire un comando e una dichiarazione appropriati:

```
\makeatletter
\DeclareRobustCommand{\Italian}{%
\fontencoding\latinencoding
\selectfont
\let\encodingdefault\latinencoding
\foreign@language{italian}}
\makeatother
```

Purtroppo non esiste (ancora) un'interfaccia più comoda per definire questi comandi.

```
%!TEX encoding = UTF-8 Unicode
\documentclass[a4paper]{article}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[bulgarian]{babel}

\begin{document}
\author{Димитри Димитров}
\title{София}
\maketitle

София е столицата и най-големият град на България.
Разположена е в северозападната част на Южна България,
в полите на планината Витоша и е основен административен,
индустриален, транспортен, културен и университетски център
на страната. София носи името на раннохристиянския храм и
бивша катедрала на града „Света София“. Към 2009 година
регистрираните в София постоянни жители наброяват
1\,267\,098 души.
\end{document}
```

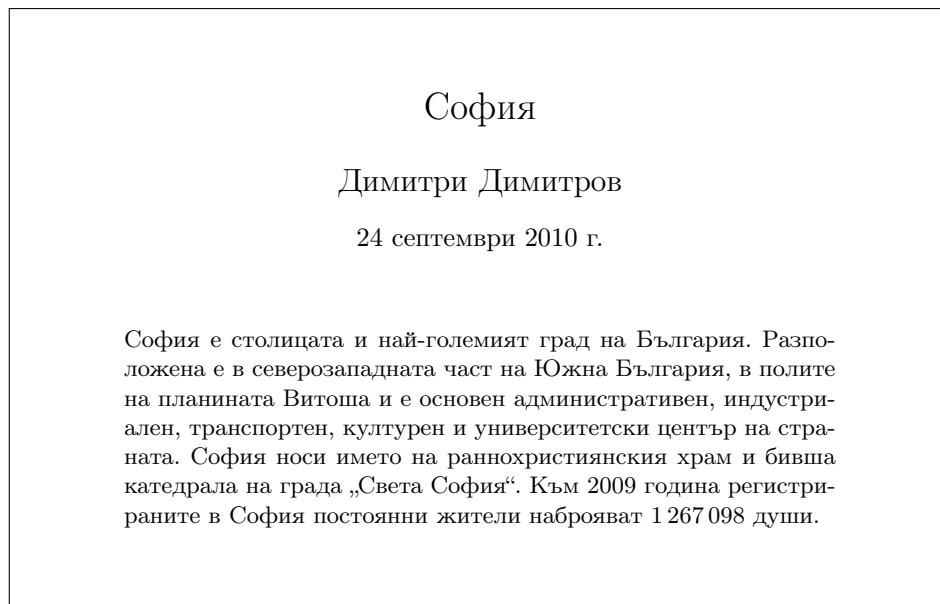


FIGURA 2: Esempio di documento in bulgaro



Supponiamo però che il nostro documento sia in russo, con inserti in italiano, francese e ucraino. Il preambolo comincerà con

```
%!TEX encoding = UTF-8 Unicode
\documentclass[a4paper]{article}
\usepackage[T1,T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[french,italian,
    ukrainian,russian]{babel}
```

Ricordiamo, per chi ancora non l'avesse capito, che la lingua principale del documento è l'ultima specificata nelle opzioni a `babel`; analogamente, l'ultima opzione a `fontenc` stabilisce la codifica di output principale, quella cioè che vale all'inizio del documento.

Qui però accade il pasticcio. La struttura dei moduli di `babel` per le tre lingue che si scrivono in caratteri cirillici non permette l'uso contemporaneo di schemi di sillabazione di due di esse. Lo si può provare aggiungendo al preambolo di prima le righe

```
\begin{document}
The language is \number\language;
it should be
\number\csname l@russian\endcsname

\selectlanguage{ukrainian}
Now the language is \number\language;
it should be
\number\csname l@ukrainian\endcsname
\end{document}
```

Una breve spiegazione preliminare a quella dell'inghippo. Il comando `l@russian`, secondo i canoni di `babel` contiene il numero assegnato alla lingua russa durante la compilazione dei formati; analogamente per `l@ukrainian`. La costruzione

```
\number\csname l@russian\endcsname
```

permette di stamparne il valore. Il parametro interno `\language` contiene il numero associato alla lingua in uso (precisamente agli schemi di sillabazione attivi). Sulla macchina dove questo documento è stato composto, il russo ha il numero 56, mentre l'ucraino ha il numero 66.<sup>9</sup> Con `\number\language` possiamo stampare il numero del linguaggio attivo: compilando l'esempio (ammesso che gli schemi di sillabazione di russo e ucraino siano stati letti durante la compilazione dei formati), ci si accorge che nonostante l'istruzione

```
\selectlanguage{ukrainian}
```

il numero stampato da `\number\language` è lo stesso! In figura 3 è possibile vedere il risultato, insieme al prologo che LATEX stampa sul terminale a ogni compilazione; basta contare (da zero) le lingue e ci si rende conto che i numeri sono quelli giusti.

9. Su ciascun sistema questi numeri possono essere diversi; si suppone, ovviamente, che le lingue siano state abilitate al momento della creazione del formato LATEX.

Dov'è dunque l'inghippo? Ogni modulo delle tre lingue (bulgaro, russo e ucraino) definisce il comando `\cyrillictext`; nei tre moduli la definizione termina con

```
\language\l@bulgarian
\language\l@russian
\language\l@ukrainian
```

rispettivamente. Successivamente il comando è inserito fra quelli eseguiti al momento del cambio di lingua, cioè in `\extrasbulgarian`, `\extrarussian` e `\extrasukrainian`. Nel nostro esempio viene letto per primo il modulo per l'ucraino e poi quello per il russo, dove la definizione di `\cyrillictext` sovrascrive la precedente. Ed ecco che quando si dà `\selectlanguage{ukrainian}` fra le istruzioni eseguite c'è `\language\l@russian` che imposta gli schemi di sillabazione al russo! Infatti l'analoga istruzione `\language\l@ukrainian` viene eseguita da `\selectlanguage{ukrainian}` ma *prima* che venga eseguito il comando `\extrasukrainian`.

Come risolvere la faccenda? Ridefinendo i comandi incautamente introdotti da `babel`. Ecco un modo che dovrebbe funzionare; c'è una differenza sostanziale rispetto ai comandi originali (sbagliati): le dichiarazioni `\Russian` e simili non modificano le parole chiave e vanno date in un gruppo. Infatti, se si desidera cambiare completamente la lingua, il comando `\selectlanguage{russian}` è più chiaro ed espressivo. Se si volesse mantenere il comportamento originale, si sostituisca `\foreign@language` con `\selectlanguage`.

```
\makeatletter
\newcommand{@cyrillictext}{%
    \fontencoding\cyrillicencoding
    \selectfont
    \let\encodingdefault\cyrillicencoding
}
\DeclareRobustCommand{\cyrillictext}{%
    \@cyrillictext
}
\let\cyr\cyrillictext
\DeclareRobustCommand{\Bulgarian}{%
    \@cyrillictext
    \foreign@language{bulgarian}}
\let\Bul\Bulgarian \let\Bg\Bulgarian
\DeclareRobustCommand{\Russian}{%
    \@cyrillictext
    \foreign@language{russian}}
\let\Rus\Russian
\DeclareRobustCommand{\Ukrainian}{%
    \@cyrillictext
    \foreign@language{ukrainian}}
\let\Ukr\Ukrainian
% \newcommand*{\defaultcyrlanguage}{%
%     russian}
\AtBeginDocument{
    \ifundefined{defaultcyrlanguage}
    {\let\defaultcyrlanguage
        \bbl@main@language}
```

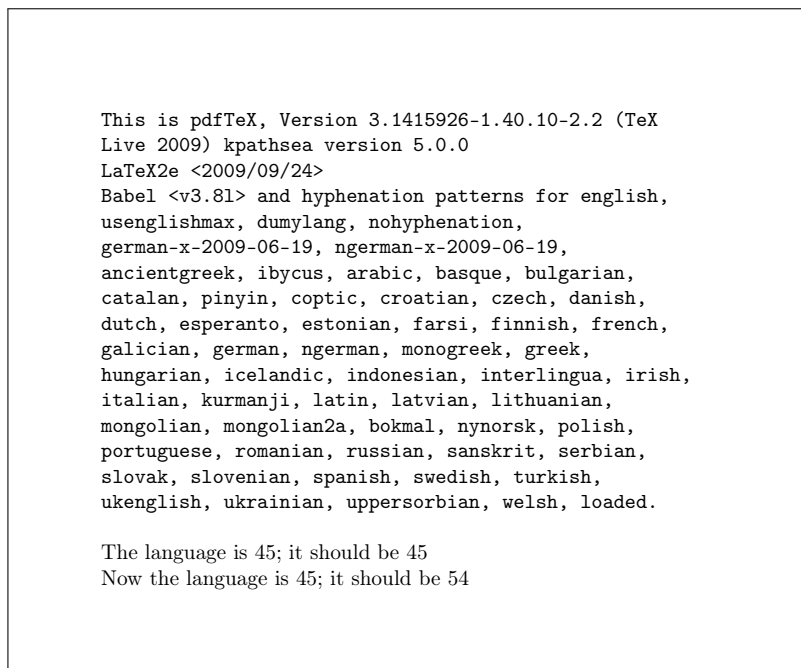


FIGURA 3: Esempio di errata impostazione degli schemi di sillabazione

```
{%
}
\renewcommand{\textcyrillic}[2]{%
  [\defaultcyrlanguage]{%
    \expandafter\foreignlanguage
    \expandafter{#1}{#2}}
\makeatother
```

Naturalmente non è necessario scrivere le definizioni che riguardano una lingua non scelta fra le opzioni a `babel`, nel nostro esempio quelle riguardanti il bulgaro. Vengono ridefiniti tutti i comandi ‘pericolosi’, comprese le abbreviazioni definite dai moduli di `babel`. In più ridefiniamo `\textcyrillic` con un argomento opzionale come equivalente di `\foreignlanguage`; il valore di default dell’argomento opzionale è la lingua principale del documento, che `babel` memorizza in `\bbl@main@language`. La definizione di `\defaultcyrlanguage` è perciò da dare nel caso in cui una delle tre lingue (bulgaro, russo o ucraino) sia la lingua principale del documento. Perciò, in un contesto italiano, ma con lingua principale il russo, potremo dare

```
\textcyrillic{образования}
\textcyrillic[ukrainian]{агломерацію}
```

invece di

```
\foreignlanguage{russian}{образования}
\foreignlanguage{ukrainian}
{агломерацію}
```

È chiaro che se la lingua principale non si scrive in caratteri cirillici, il comando `\textcyrillic` dovrà ricevere il corretto argomento opzionale oppure bisognerà definire, come indicato nell’esempio, il comando `\defaultcyrilliclanguage`, togliendo il segno di commento e modificando la lingua.

## 8 La codifica X2

Oltre alle codifiche T2A, T2B e T2C che servono a contenere i numerosi caratteri degli alfabeti cirillici in modo che ciascuna lingua possa usare uno di essi ed evitare problemi con l’interpretazione dei caratteri di codice minore di 128 necessari per scrivere i comandi  $\LaTeX$ , esiste anche un contenitore generale dei caratteri cirillici, chiamato “codifica X2”. In questa codifica trovano posto, per esempio, le lettere prerivoluzionarie  $\Theta\epsilon\text{I}\text{V}\text{v}\text{B}\text{f}$ .

Chi dovesse trascrivere qualche lungo tratto di testo in russo con queste lettere, oppure in lingue per le quali la codifica T2A non è sufficiente, può usare i caratteri UTF-8 e definire comandi appositi come

```
\newcommand{\cyrx}[1]{%
  \language\l@nohyphenation
  \fontencoding{X2}\selectfont}
\DeclareTextFontCommand{\textcyrx}
{\cyrx}
```

dopo che sia stata dichiarata la codifica X2 fra le opzioni a `fontenc`. Così il nome scritto in grafia prerivoluzionaria `\textcyrx{\Theta\epsilon\text{I}\text{V}\text{v}\text{B}\text{f}}` è stampato correttamente come  $\Theta\epsilon\text{I}\text{V}\text{v}\text{B}\text{f}$  e

```
{\cyrx
азъ, буки, вѣди, глаголь, добро, есть,
живѣте, земля, иже, и съ краткой, како,
люди, мыслѣте, нашъ, онъ, покой, рцы,
слово, твердо, укъ, фертъ, хѣръ, цы,
червь, ша, ща, еръ, еры, ерь,
э оборотное, ю, я\par}
```

viene correttamente composto come

азъ, буки, вѣди, глаголь, добро, есть, живѣте, земля, иже, и съ краткой, како, люди, мыслѣте, нашъ, онъ, покой, рцы, слово, твердо, укъ, фертъ, хѣръ, цы, червь, ша, ща, ерь, еры, ерь, э оборотное, ю, я

Altro esempio è l'inizio della dichiarazione dei diritti dell'uomo in lingua /sakha/, parlata nella Саха Республика (Якутия), repubblica autonoma della Federazione Russa che comprende gran parte della Siberia orientale.<sup>10</sup> La lingua appartiene allo stesso ceppo del turco.

Дьон барыта бэйэ суолтатыгар уонна  
быраабыгар тэн буолан төрүүлэр.  
Кинилэр бары өркөн өйдөөх, суобастаах  
буолан төрүүлэр, уонна бэйэ  
бэйэлэригэр тылга кииринигэс быдыылара  
доҕордоһуу тыыннаах буолуохтаах.

che risulta nel testo seguente (composto a bandiera perché non conosco le regole di sillabazione della lingua).

Дьон барыта бэйэ суолтатыгар уонна  
быраабыгар тэн буолан төрүүлэр.  
Кинилэр бары өркөн өйдөөх, суобастаах  
буолан төрүүлэр, уонна бэйэ бэйэлэригэр  
тылга кииринигэс быдыылара доҕордоһуу  
тыыннаах буолуохтаах.

Si noti che in questa lingua il carattere 'Ө' indica un suono simile a quello del danese 'ø'. Lo stesso testo si potrebbe comporre usando la codifica T2B (per questa lingua la codifica T2A non è invece adatta):

Дьон барыта бэйэ суолтатыгар уонна  
быраабыгар тэн буолан төрүүлэр.  
Кинилэр бары өркөн өйдөөх, суобастаах  
буолан төрүүлэр, уонна бэйэ бэйэлэригэр  
тылга кииринигэс быдыылара доҕордоһуу  
тыыннаах буолуохтаах.

Per la lingua abcaso è necessaria la codifica T2C. Ecco l'alfabeto, prima in carattere dattilografico, poi reso con la codifica T2C e quindi con la codifica X2.

А а Б б В в Г г Гь гь Гэ гэ Ѓ ґ Ѓь ґь  
Ђа ґа Д д Дэ дэ Е е Ж ж Жь жь Жа жа З з  
Зэ зэ И и К к Кь кь Кэ кэ Қ қ  
Қь қь Қэ қэ Л л М м  
Н н О о П п Ѓ ґ Р р С с Т т Тэ тэ  
Т т Тэ тэ У у Ф ф Х х Хь хь Хэ хэ  
Х х Хэ хэ Ц ц Цэ цэ Цҫ цҫ Ч ч  
Ч ч Ё ё Ё ё Ш ш Шь шь Шэ шэ Ы ы  
Ѡ ѡ Ц ц Ць ць Ь ь Ѣ ѣ

10. La superficie della repubblica è di 3 103 200 km<sup>2</sup>, dieci volte l'Italia; la popolazione è di circa un milione di abitanti. Corrisponde, più o meno, alla Jacuzzi del RisiKo!<sup>TM</sup>.

T2C:

А а Б б В в Г г Гь гь Гэ гэ Ѓ ґ Ѓь ґь Ѓа ґа  
Д д Дэ дэ Е е Ж ж Жь жь Жа жа З з Зз  
Зэ зэ И и К к Кь кь Кэ кэ Қ қ Қь қь Қэ қэ  
К к Кь кь Кэ кэ Л л М м Н н О о П п Ѓ ґ  
Р р С с Т т Тэ тэ Т т Тэ тэ У у Ф ф Х х  
Хь хь Хэ хэ Х х Хэ хэ Ц ц Цэ цэ Цҫ цҫ  
Ч ч Ч ч Ё ё Ё ё Ш ш Шь шь Шэ шэ Ы ы  
Ѡ ѡ Ц ц Ць ць Ь ь Ѣ ѣ

X2:

А а Б б В в Г г Гь гь Гэ гэ Ѓ ґ Ѓь ґь Ѓа ґа  
Д д Дэ дэ Е е Ж ж Жь жь Жа жа З з Зз  
Зэ зэ И и К к Кь кь Кэ кэ Қ қ Қь қь Қэ қэ  
К к Кь кь Кэ кэ Л л М м Н н О о П п Ѓ ґ  
Р р С с Т т Тэ тэ Т т Тэ тэ У у Ф ф Х х  
Хь хь Хэ хэ Х х Хэ хэ Ц ц Цэ цэ Цҫ цҫ  
Ч ч Ч ч Ё ё Ё ё Ш ш Шь шь Шэ шэ Ы ы  
Ѡ ѡ Ц ц Ць ць Ь ь Ѣ ѣ

Si noti che la più recente ortografia dell'abcaso richiede la lettera Гҫ al posto di Ѓ ґ anche nei digrammi Гь гь e Гэ гэ; questa lettera non compare nella codifica T2C, ma compare nella X2. Un'analoga modifica è necessaria per la lettera Ѓ ґ, ma il carattere con il discendente invece dell'uncino non è disponibile con i font Computer Modern. L'alfabeto comprende ben 64 lettere, fra cui 24 digrammi!

Come si vede, il risultato è identico. Dunque la codifica X2 è comoda nel caso si debbano usare caratteri cirillici che appartengono alle tre codifiche senza dover caricare font per tutte e tre. Tuttavia solo la famiglia Computer Modern Cyrillic, nella distribuzione T<sub>E</sub>X Live, contiene font adatti per quella codifica.

## Riferimenti bibliografici

PANTIERI, L. «L'arte di scrivere in cirillico con L<sup>A</sup>T<sub>E</sub>X». [http://www.lorenzopantieri.net/LaTeX\\_files/Cirillico.pdf](http://www.lorenzopantieri.net/LaTeX_files/Cirillico.pdf).

PÍŠKA, K. (1996). «Cyrillic alphabets». *TUGboat*, **17** (2), pp. 92–98. <http://www.tug.org/TUGboat/Articles/tb17-2/tb51pisk.pdf>.

ŽUBRINIĆ, D. (1996). «Croatian fonts». *TUGboat*, **17** (1), pp. 29–33. <http://www.tug.org/TUGboat/Articles/tb17-1/tb50zubr.pdf>.

▷ Enrico Gregorio  
Dipartimento di Informatica – Università di Verona  
Enrico dot Gregorio at univr dot it

TABELLA 9: Lista completa dei caratteri nei font della University of Washington

Output	Input	Output	Input	Output	Input	Output	Input
A	A	Б	B	Ц	C	Д	D
E	E	Ф	F	Г	G	X	H
И	I	J	J	К	K	Л	L
M	M	H	N	O	O	П	P
Ч	Q	P	R	С	S	Т	T
У	U	В	V	Щ	W	Ш	X
Ы	Y	З	Z				
a	a	б	b	ц	c	д	d
e	e	ф	f	г	g	х	h
и	i	ј	j	к	k	л	l
м	m	н	n	о	o	п	p
ч	q	р	r	с	s	т	t
у	u	в	v	щ	w	ш	x
ы	y	з	z				
Ѓ	C1	Ч	Ch	Ђ	D1	Ї	D2
S	D3	Ђ	Dj	Ё	E0	Э	E1
Є	E2	Ї	I0	І	I1	Я	J1
Ю	J2	X	Kh	Љ	L1	Ў	Lj
№	N0	Њ	N1	Ћ	Nj	Ь	P1
Љ	P2	Ш	Sh	Щ	Shch	Ц	Ts
Я	Ya	Ю	Yu	Ж	Z1	Ж	Zh
ћ	c1	ч	ch	ђ	d1	џ	d2
s	d3	ђ	dj	ё	e0	э	e1
є	e2	ї	i0	і	i1	я	j1
ю	j2	х	kh	љ	l1	ў	lj
		њ	n1	ћ	nj	ь	p1
џ	p2	ш	sh	щ	shch	ц	ts
я	ya	ю	yu	ж	z1	ж	zh
Ѐ	\'C	Ѐ	\'G	Ѐ	\'K	Ѐ	\u{U}
ђ	\'c	ђ	\'g	ђ	\'k	ђ	\u{u}
«	[[	»	]]	«	<	»	>
Ђ	\#	V	\char'24	Ѐ	\char'25		
ђ	+	v	\char'34	Ѐ	\char'35		

(1) Nelle legature il simbolo 0 è la cifra zero

(2) Per spezzare tc si usi t\s oppure t\cydot s

(3) Tutte le legature come 'Yu' o 'Ts' funzionano anche se la seconda lettera è maiuscola

TABELLA 10: Tabella della codifica T2A

	0	1	2	3	4	5	6	7	
'00x	` 0	´ 1	^ 2	~ 3	¨ 4	˜ 5	° 6	ˇ 7	"0x
'01x	˘ 8	¯ 9	· 10	¸ 11	€ 12	ℐ 13	⟨ 14	⟩ 15	
'02x	“ 16	” 17	^ 18	˘ 19	ˆ 20	– 21	— 22	23	"1x
'03x	0 24	1 25	2 26	ff 27	fi 28	fl 29	ffi 30	ffl 31	
'04x	˘ 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	"2x
'05x	( 40	) 41	* 42	+ 43	, 44	- 45	. 46	/ 47	
'06x	0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	"3x
'07x	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63	
'10x	@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	"4x
'11x	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79	
'12x	P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	"5x
'13x	X 88	Y 89	Z 90	[ 91	\ 92	] 93	^ 94	_ 95	
'14x	‘ 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	"6x
'15x	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111	
'16x	p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	"7x
'17x	x 120	y 121	z 122	{ 123	124	} 125	~ 126	- 127	
'20x	Г 128	Г 129	Ђ 130	Ћ 131	ћ 132	Ж 133	З 134	Љ 135	"8x
'21x	Ї 136	Ќ 137	Њ 138	Ћ 139	Æ 140	Ҁ 141	҂ 142	Ѕ 143	
'22x	Ө 144	Ќ 145	Ў 146	Ҁ 147	҂ 148	Ҁ 149	҂ 150	҂ 151	"9x
'23x	҂ 152	€ 153	Ә 154	҂ 155	Ё 156	№ 157	☒ 158	§ 159	
'24x	г 160	Г 161	ђ 162	ћ 163	ћ 164	ж 165	з 166	љ 167	"Ax
'25x	ї 168	ќ 169	њ 170	ќ 171	æ 172	Ҁ 173	҂ 174	ѕ 175	
'26x	ө 176	ѕ 177	ў 178	Ҁ 179	҂ 180	Ҁ 181	҂ 182	҂ 183	"Bx
'27x	҂ 184	€ 185	Ә 186	҂ 187	ё 188	,, 189	« 190	» 191	
'30x	А 192	Б 193	В 194	Г 195	Д 196	Е 197	Ж 198	З 199	"Cx
'31x	И 200	Й 201	К 202	Л 203	М 204	Н 205	О 206	П 207	
'32x	Р 208	С 209	Т 210	У 211	Ф 212	Х 213	Ц 214	Ч 215	"Dx
'33x	Ш 216	Щ 217	Ъ 218	Ы 219	Ь 220	Э 221	Ю 222	Я 223	
'34x	а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	"Ex
'35x	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239	
'36x	р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	"Fx
'37x	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255	
	"8	"9	"A	"B	"C	"D	"E	"F	

TABELLA 11: Tabella della codifica T2B

	0	1	2	3	4	5	6	7	
'00x	` 0	´ 1	^ 2	~ 3	¨ 4	˜ 5	° 6	˘ 7	"0x
'01x	˘ 8	¯ 9	· 10	¸ 11	ε 12	ı 13	⟨ 14	⟩ 15	
'02x	“ 16	” 17	^ 18	˘ 19	ˆ 20	— 21	— 22	23	"1x
'03x	o 24	l 25	j 26	ff 27	fi 28	fl 29	ffi 30	ffl 31	
'04x	„ 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	"2x
'05x	( 40	) 41	* 42	+ 43	, 44	- 45	. 46	/ 47	
'06x	0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	"3x
'07x	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63	
'10x	@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	"4x
'11x	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79	
'12x	P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	"5x
'13x	X 88	Y 89	Z 90	[ 91	\ 92	] 93	^ 94	_ 95	
'14x	‘ 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	"6x
'15x	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111	
'16x	p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	"7x
'17x	x 120	y 121	z 122	{ 123	124	} 125	~ 126	- 127	
'20x	ƒ 128	Ƒ 129	Ɠ 130	Ɣ 131	ħ 132	℥ 133	đ 134	Ʒ 135	"8x
'21x	Љ 136	Њ 137	Ћ 138	Ќ 139	Ј 140	Ў 141	Ѓ 142	Ѕ 143	
'22x	Θ 144	Ɔ 145	Ÿ 146	Υ 147	Χ 148	Χ 149	Χ 150	Ϣ 151	"9x
'23x	Ϣ 152	Ѓ 153	Θ 154	Ε 155	Ё 156	№ 157	Ϣ 158	§ 159	
'24x	ƒ 160	Ƒ 161	Ɠ 162	Ɣ 163	ħ 164	℥ 165	đ 166	Ʒ 167	"Ax
'25x	Љ 168	Њ 169	Ћ 170	Ќ 171	Ј 172	Ў 173	Ѓ 174	Ѕ 175	
'26x	Θ 176	Ɔ 177	Ÿ 178	Υ 179	Χ 180	Χ 181	Χ 182	Ϣ 183	"Bx
'27x	Ϣ 184	Ѓ 185	Θ 186	Ε 187	Ё 188	„ 189	« 190	» 191	
'30x	A 192	Б 193	В 194	Г 195	Д 196	Е 197	Ж 198	З 199	"Cx
'31x	И 200	Й 201	К 202	Л 203	М 204	Н 205	О 206	П 207	
'32x	Р 208	С 209	Т 210	У 211	Ф 212	Х 213	Ц 214	Ч 215	"Dx
'33x	Ш 216	Щ 217	Ъ 218	Ы 219	Ь 220	Э 221	Ю 222	Я 223	
'34x	a 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	"Ex
'35x	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239	
'36x	р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	"Fx
'37x	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255	
	"8	"9	"A	"B	"C	"D	"E	"F	

TABELLA 12: Tabella della codifica T2C

	0	1	2	3	4	5	6	7	
'00x	` 0	´ 1	^ 2	~ 3	¨ 4	˜ 5	° 6	ˇ 7	"0x
'01x	˘ 8	¯ 9	· 10	¸ 11	€ 12	ı 13	‹ 14	› 15	
'02x	“ 16	” 17	^ 18	˘ 19	˙ 20	– 21	— 22	23	"1x
'03x	0 24	1 25	2 26	ff 27	fi 28	fl 29	ffi 30	ffl 31	
'04x	˘ 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	"2x
'05x	( 40	) 41	* 42	+ 43	, 44	- 45	. 46	/ 47	
'06x	0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	"3x
'07x	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63	
'10x	@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	"4x
'11x	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79	
'12x	P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	"5x
'13x	X 88	Y 89	Z 90	[ 91	\ 92	] 93	^ 94	_ 95	
'14x	‘ 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	"6x
'15x	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111	
'16x	p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	"7x
'17x	x 120	y 121	z 122	{ 123	124	} 125	~ 126	- 127	
'20x	Ѐ 128	Ё 129	Ђ 130	Ѓ 131	ђ 132	Ѕ 133	Ї 134	Њ 135	"8x
'21x	Ќ 136	Ѕ 137	Ї 138	Њ 139	Ћ 140	Ќ 141	Ќ 142	Ќ 143	
'22x	Ќ 144	Ќ 145	Ќ 146	Ќ 147	Ќ 148	Ќ 149	Ќ 150	Ќ 151	"9x
'23x	Ќ 152	Ќ 153	Ќ 154	Ќ 155	Ќ 156	Ќ 157	Ќ 158	Ќ 159	
'24x	Ќ 160	Ќ 161	Ќ 162	Ќ 163	Ќ 164	Ќ 165	Ќ 166	Ќ 167	"Ax
'25x	Ќ 168	Ќ 169	Ќ 170	Ќ 171	Ќ 172	Ќ 173	Ќ 174	Ќ 175	
'26x	Ќ 176	Ќ 177	Ќ 178	Ќ 179	Ќ 180	Ќ 181	Ќ 182	Ќ 183	"Bx
'27x	Ќ 184	Ќ 185	Ќ 186	Ќ 187	Ќ 188	Ќ 189	Ќ 190	Ќ 191	
'30x	А 192	Б 193	В 194	Г 195	Д 196	Е 197	Ж 198	З 199	"Cx
'31x	И 200	Й 201	К 202	Л 203	М 204	Н 205	О 206	П 207	
'32x	Р 208	С 209	Т 210	У 211	Ф 212	Х 213	Ц 214	Ч 215	"Dx
'33x	Ш 216	Щ 217	Ъ 218	Ы 219	Ь 220	Э 221	Ю 222	Я 223	
'34x	а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	"Ex
'35x	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239	
'36x	р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	"Fx
'37x	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255	
	"8	"9	"A	"B	"C	"D	"E	"F	

TABELLA 13: Tabella della codifica X2

	0	1	2	3	4	5	6	7	
'00x	` 0	´ 1	^ 2	~ 3	¨ 4	˜ 5	° 6	˘ 7	"0x
'01x	˘ 8	¯ 9	· 10	¸ 11	€ 12	ı 13	‹ 14	› 15	
'02x	“ 16	” 17	^ 18	˘ 19	ˆ 20	— 21	— 22	23	"1x
'03x	o 24	˘ 25	ˆ 26	- 27	đ 28	đ 29	ђ 30	ђ 31	
'04x	˘ 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	"2x
'05x	( 40	) 41	* 42	+ 43	, 44	- 45	. 46	/ 47	
'06x	0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	"3x
'07x	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63	
'10x	@ 64	Æ 65	Ƨ 66	Ƨ 67	Ƨ 68	Ƨ 69	Ƨ 70	Ƨ 71	"4x
'11x	Ƨ 72	Ƨ 73	Ƨ 74	Ƨ 75	Ƨ 76	Ƨ 77	Ƨ 78	Ƨ 79	
'12x	Ƨ 80	Ƨ 81	Ƨ 82	Ƨ 83	Ƨ 84	Ƨ 85	Ƨ 86	Ƨ 87	"5x
'13x	Ƨ 88	Ƨ 89	Ƨ 90	Ƨ 91	Ƨ 92	Ƨ 93	Ƨ 94	Ƨ 95	
'14x	‘ 96	æ 97	ħ 98	ħ 99	€ 100	€ 101	Ƨ 102	Ƨ 103	"6x
'15x	Ƨ 104	ı 105	ı 106	Ƨ 107	Ƨ 108	Ƨ 109	Ƨ 110	Ƨ 111	
'16x	Ƨ 112	Ƨ 113	Ƨ 114	Ƨ 115	Ƨ 116	Ƨ 117	Ƨ 118	Ƨ 119	"7x
'17x	Ƨ 120	Ƨ 121	Ƨ 122	{ 123	124	} 125	˘ 126	- 127	
'20x	Ƨ 128	Ƨ 129	Ƨ 130	Ƨ 131	Ƨ 132	Ƨ 133	Ƨ 134	Ƨ 135	"8x
'21x	Ƨ 136	Ƨ 137	Ƨ 138	Ƨ 139	Ƨ 140	Ƨ 141	Ƨ 142	Ƨ 143	
'22x	Ƨ 144	Ƨ 145	Ƨ 146	Ƨ 147	Ƨ 148	Ƨ 149	Ƨ 150	Ƨ 151	"9x
'23x	Ƨ 152	Ƨ 153	Ƨ 154	Ƨ 155	Ƨ 156	Ƨ 157	Ƨ 158	Ƨ 159	
'24x	Ƨ 160	Ƨ 161	Ƨ 162	Ƨ 163	Ƨ 164	Ƨ 165	Ƨ 166	Ƨ 167	"Ax
'25x	Ƨ 168	Ƨ 169	Ƨ 170	Ƨ 171	Ƨ 172	Ƨ 173	Ƨ 174	Ƨ 175	
'26x	Ƨ 176	Ƨ 177	Ƨ 178	Ƨ 179	Ƨ 180	Ƨ 181	Ƨ 182	Ƨ 183	"Bx
'27x	Ƨ 184	Ƨ 185	Ƨ 186	Ƨ 187	Ƨ 188	„ 189	« 190	» 191	
'30x	А 192	Б 193	В 194	Г 195	Д 196	Е 197	Ж 198	З 199	"Cx
'31x	И 200	Й 201	К 202	Л 203	М 204	Н 205	О 206	П 207	
'32x	Р 208	С 209	Т 210	У 211	Ф 212	Х 213	Ц 214	Ч 215	"Dx
'33x	Ш 216	Щ 217	Ъ 218	Ы 219	Ь 220	Э 221	Ю 222	Я 223	
'34x	а 224	б 225	в 226	г 227	д 228	е 229	ж 230	з 231	"Ex
'35x	и 232	й 233	к 234	л 235	м 236	н 237	о 238	п 239	
'36x	р 240	с 241	т 242	у 243	ф 244	х 245	ц 246	ч 247	"Fx
'37x	ш 248	щ 249	ъ 250	ы 251	ь 252	э 253	ю 254	я 255	
	"8	"9	"A	"B	"C	"D	"E	"F	



```

\ProvidesFile{cyrserbian}[2009/12/01 %
  Elementary support for cyrillic serbian]
\providecommand\cyrtext{\fontencoding{T2A}\selectfont}
\providecommand\CYRJ{J}\providecommand\cyrj{j}
\def\captionsserbian{%
  \def\prefacename{\cyrtext\CYRP\cyrr\cyre\cyrd\cyrg\cyro
    \cyrv\cyro\cyrr}}%
  \def\refname{\cyrtext\CYRL\cyri\cyrt\cyre\cyrr\cyra\cyrt
    \cyru\cyrr\cyra}}%
  \def\abstractname{\cyrtext\CYRS\cyra\cyrrh\cyre\cyrt\cyra\cyrk}}%
  \def\bibname{\cyrtext\CYRE\cyri\cyrb\cyrl\cyri\cyro\cyrg\cyrr\cyra
    \cyrf\cyri\cyrj\cyra}}%
  \def\chaptername{\cyrtext\CYRG\cyrl\cyra\cyrv\cyra}}%
  \def\appendixname{\cyrtext\CYRD\cyro\cyrd\cyra\cyrt\cyra\cyrk}}%
  \def\contentsname{\cyrtext\CYRS\cyra\cyrd\cyrr\cyrrh\cyra\cyrj}}%
  \def\listfigurename{\cyrtext\CYRS\cyrl\cyri\cyrk\cyre}}%
  \def\listtablename{\cyrtext\CYRT\cyra\cyrb\cyre\cyrl\cyre}}%
  \def\indexname{\cyrtext\CYRI\cyrn\cyrd\cyre\cyrk\cyrs}}%
  \def\figurename{\cyrtext\CYRS\cyrl\cyri\cyrk\cyra}}%
  \def\tablename{\cyrtext\CYRT\cyra\cyrb\cyre\cyrl\cyra}}%
  \def\partname{\cyrtext\CYRD\cyre\cyro}}%
  \def\enclname{\cyrtext\CYRP\cyrr\cyri\cyrl\cyro\cyrz\cyri}}%
  \def\ccname{\cyrtext\CYRK\cyro\cyrp\cyri\cyrj\cyre}}%
  \def\headtoname{\cyrtext\CYRP\cyrr\cyri\cyrm\cyra}}%
  \def\pagename{\cyrtext\CYRS\cyrt\cyrr\cyra\cyrn\cyra}}%
  \def\seenname{\cyrtext\CYRV\cyri\cyrd\cyri}}%
  \def\alsiname{\cyrtext\CYRV\cyri\cyrd\cyri\ \cyrt\cyra\cyrk
    \cyro\cyrdje\cyre}}%
  \def\proofname{\cyrtext\CYRD\cyro\cyrk\cyra\cyrz}}%
  \def\glossaryname{\cyrtext\CYRG\cyrl\cyro\cyrs\cyra\cyrr}}%
}%
\def\dateserbian{%
  \def\today{\number\day .~\ifcase\month\or
    {\cyrtext\cyrj\cyra\cyrn\cyru\cyra\cyrr}\or
    {\cyrtext\cyrf\cyre\cyrb\cyrr\cyru\cyra\cyrr}\or
    {\cyrtext\cyrm\cyra\cyrr\cyrt}\or
    {\cyrtext\cyra\cyrp\cyrr\cyri\cyrl}\or
    {\cyrtext\cyrm\cyra\cyrj}\or
    {\cyrtext\cyrj\cyru\cyrn\cyri}\or
    {\cyrtext\cyrj\cyru\cyrl\cyri}\or
    {\cyrtext\cyra\cyrv\cyrg\cyru\cyrs\cyrt}\or
    {\cyrtext\cyrs\cyre\cyrp\cyrt\cyre\cyrm\cyrb\cyra\cyrr}\or
    {\cyrtext\cyro\cyrk\cyrt\cyro\cyrb\cyra\cyrr}\or
    {\cyrtext\cyrn\cyro\cyrv\cyre\cyrm\cyrb\cyra\cyrr}\or
    {\cyrtext\cyrd\cyre\cyrc\cyre\cyrm\cyrb\cyra\cyrr}\fi
    \space \number\year}}

```

FIGURA 4: Come sistemare le parole fisse per il serbo in cirillico