

# The Serbian Latin Language in the babel system

Uroš Stefanović\*

September 30, 2022

## The Serbian Latin Language

The file `serbian.dtx` defines all the language definition macros for the Serbian language, typeset in a Latin script.

For this language the character " is made active. In table 1 an overview is given of its purpose. One of the reasons for this is that in the Serbian language some special characters are used.

"c	<code>\v c</code> , also implemented for the lowercase and uppercase s and z.
"d	<code>\dj</code> , also implemented for "D.
"-	An explicit hyphen sign, allowing hyphenation in the rest of the word; inserts a hyphen which is repeated at the beginning of the next line (recommended to use for compound words with hyphen).
"	Disables ligature at this position.
""	Similar to "-" but prints no hyphen sign.
"~	Compound word mark without a breakpoint, prints hyphen prohibiting hyphenation at the point.
"=	A compound word mark with a breakpoint, prints hyphen allowing hyphenation in the composing words.
"“	German opening double quote (looks like „).
"”	German closing double quote (looks like “).
"”	(if the <code>quotes</code> attribute is used) Closing double quote (looks like ”).
"<	French opening double quote (looks like <<).
">	French closing double quote (looks like >>).

Table 1: The extra definitions made by `serbian.ldf`

Macro `\today` prints the date in Serbian. Alternatively, if attribute `datei` is used, `\today` prints the current date, but prints ‘juni’ and ‘juli’ for ‘June’ and ‘July’. If you prefer to use ‘juni’ and ‘juli’ instead of default ‘jun’ and ‘jul’, use the `datei` attribute. Also, the `\today*` macro prints the date without dot after the year (used when after the date is the punctuation mark, such as comma). Alternatively, the commands `\todayRoman` and `\todayRoman*` prints the current date using Roman numerals for months; `\todayGen` and `\todayGen*` prints the current month name in the genitive case, and `\todayArabic` and `\todayArabic*` prints the current month as a number.

The alphabetical enumerations in texts use the Latin alphabet and alphabetic order, but the letters q, w, x and y are omitted by the rules of the Serbian language (22 letters are used). However, if the user wants to use the English alphabet for the enumeration (26 letters), this option is also available. We will also provide the enumeration with the Latin letters but in alphabetic order that matches the

---

\*urostajms@gmail.com

Cyrillic alphabet (30 letters). This of course shouldn't be used when the text is written in the Latin script. However, sometimes the text is written in the Latin script so it can be later converted into Cyrillic script (for example using the OT2 encoding); in such case this alphabet order will be useful. One can manually switch the enumeration alphabet with the commands `\enumCyr`, `\enumLat` and `\enumEng`. This commands can be used after the `\begin{document}` when the serbian language is active. In principle, enumerations are a matter for class and style designers but the same can be said also about things, other than enumerations, such as names of sections and bibliography lists.

Apart from defining shorthands we need to make sure that the first paragraph of each section is indented. Furthermore the following new math operators are defined: `\sh`, `\ch`, `\tg`, `\ctg`, `\arctg`, `\arcctg`, `\th`, `\cth`, `\arsh`, `\arch`, `\arth`, `\arcth`, `\cosec`, `\Prob`, `\Expect`, `\Variance`, `\arcsec`, `\arccosec`, `\sech`, `\cosech`, `\arsech`, `\arcosech`, `\NZD`, `\nzd`, `\NZS`, `\nzs`.

By default, a ekavian spelling is enabled. For ijekavian spelling can be enabled by setting the attribute to `ijekav`. To set an attribute, put the `\languageattribute` macro within a document preamble after `babel`, for example,

```
\usepackage[english,serbian]{babel}
\languageattribute{serbian}{ijekav}
```

Setting the `ijekav` attribute changes the built-in strings (caption names). For example, the part will be entitled as 'Deo' by default and as 'Dio' if the Serbian language attribute is set to `ijekav`. Same result can be achieved using a modifier as follows:

```
\usepackage[english,serbian.ijekav]{babel}
```

Using a modifier in a package option is often better. A modifier is set after the language name, and is prefixed with a dot (only when the language is set as package option — neither global options nor the main key accept them). Also, it's possible to use more than one attribute:

```
\usepackage[english,serbian.ijekav.datei.quotes]{babel}
```

The file `serbian.ldf` is designed to work both with legacy non-unicode (8-bit) and new Unicode encodings of the source document files (input encodings) and of the font files (font encodings).

## The code

The macro `\LdfInit` takes care of preventing that this file is loaded more than once, checking the category code of the `@` sign, etc.

```
1 (*code)
2 \LdfInit{serbian}{captionsserbian}
```

First, we check if `LuaLATEX` or `XeLATEX` is running. If so, we set boolean key `\if@srb@uni@ode` to true.

```
3 \ifdefined\if@srb@uni@ode
4   \PackageError{babel}{\if@srb@uni@ode already defined.}
5   \relax
6 \fi
7 \newif\if@srb@uni@ode
8 \ifdefined\luatexversion \@srb@uni@odetrue \else
9 \ifdefined\XeTeXrevision \@srb@uni@odetrue \fi\fi
```

Check if hyphenation patterns for the Serbian language have been loaded in `language.dat`. Namely, we check for the existence of `\l@serbian`. If it is not defined, we declare Serbian as dialect for the default language number 0 which almost for sure is English.

```
10 \ifx\l@serbian\@undefined
11   \@nopatterns{Serbian}
12   \adddialect\l@serbian0
13 \fi
```

For Serbian the " character is made active. This is done once, later on its definition may vary. Other languages in the same document may also use the " character for shorthands; we specify that the Serbian group of shorthands should be used. We save the original double quote character in the `\dq` macro to keep it available. The shorthand "- should be used in places where a word contains an explicit hyphenation character. According to the rules of the Serbian language, when a word break occurs at an explicit hyphen it must appear both at the end of the first line and at the beginning of the second line.

```

14 \initiate@active@char{"}
15 \begingroup \catcode'\ "12
16 \def\reserved@a{\endgroup
17 \def\@SS{\mathchar"7019 }
18 \def\dq{"}}
19 \reserved@a
20 \declare@shorthand{serbian}{c}{\textormath{\v c}{\check c}}
21 \declare@shorthand{serbian}{d}{\textormath{\dj}{\textnormal{\dj}}}
22 \declare@shorthand{serbian}{s}{\textormath{\v s}{\check s}}
23 \declare@shorthand{serbian}{z}{\textormath{\v z}{\check z}}
24 \declare@shorthand{serbian}{C}{\textormath{\v C}{\check C}}
25 \declare@shorthand{serbian}{D}{\textormath{\DJ}{\textnormal{\DJ}}}
26 \declare@shorthand{serbian}{S}{\textormath{\v S}{\check S}}
27 \declare@shorthand{serbian}{Z}{\textormath{\v Z}{\check Z}}
28 \declare@shorthand{serbian}{"'}{\quotedblbase}
29 \declare@shorthand{serbian}{"'}{\textquotedblleft}
30 \declare@shorthand{serbian}{"<}{\guillemotleft}
31 \declare@shorthand{serbian}{">}{\guillemotright}
32 \declare@shorthand{serbian}{""}{\hskip\z@skip}
33 \declare@shorthand{serbian}{"~}{\textormath{\leavevmode\hbox{-}}{-}}
34 \declare@shorthand{serbian}{"=}{\nobreak-\hskip\z@skip}
35 \declare@shorthand{serbian}{"|}{\textormath{\nobreak\discretionary{-}{-}{\kern.03em}\allowhyphens}{}}
36 \declare@shorthand{serbian}{"-}{\nobreak\discretionary{-}{-}{-}\bbl\allowhyphens}

```

The macro `\captionsserbian` defines all strings used in the four standard documentclasses provided with L<sup>A</sup>T<sub>E</sub>X.

```

37 \addto\captionsserbian{%
38 \def\prefacename{Predgovor}%
39 \def\refname{Literatura}%
40 \def\abstractname{Sa\v zetak}%
41 \def\bibName{Bibliografija}%
42 \def\chaptername{Glava}%
43 \def\appendixname{Dodatak}%
44 \def\contentsname{Sadr\v zaj}%
45 \def\listfigurename{Spisak slika}%
46 \def\listtablename{Spisak tabela}%
47 \def\indexname{Indeks}%
48 \def\figurename{Slika}%
49 \def\tablename{Tabela}%
50 \def\partname{Deo}%
51 \def\enclname{Prilozi}%
52 \def\ccname{Kopije}%
53 \def\headtoname{Prima}%
54 \def\pagename{strana}%
55 \def\seename{vidi}%
56 \def\alsosome{vidi tako\dj e}%
57 \def\proofname{Dokaz}%
58 \def\glossaryname{Re\v cnik}%
59 }%
60 \if@srb@uni@ode
61 \addto\captionsserbian{%
62 \def\abstractname{Sažetak}%

```

```

63   \def\contentsname{Sadržaj}%
64   \def\alsoname{vidi takođe}%
65   \def\glossaryname{Rečnik}%
66   }%
67 \fi
68 \let\captionsserbian@ijekav=\captionsserbian
69 \addto\captionsserbian@ijekav{%
70   \def\partname{Dio}%
71   \def\glossaryname{Rje\v cnik}%
72 }
73 \if@srb@uni@ode
74   \addto\captionsserbian@ijekav{%
75     \def\glossaryname{Rječnik}%
76   }
77 \fi

```

The macro `\dateserbian` redefines the commands `\today`, `\today*`, `\todayRoman` and `\todayRoman*` to produce Serbian dates. Also, the commands `\todayGen`, `\todayGen*`, `\todayArabic` and `\todayArabic*` are provided.

```

78   \def\dateserbian{%
79     \def\month@serbian{\ifcase\month\or
80       januar\or
81       februar\or
82       mart\or
83       april\or
84       maj\or
85       jun\or
86       jul\or
87       avgust\or
88       septembar\or
89       oktobar\or
90       novembar\or
91       decembar\fi}%
92     \def\month@serbian@gen{\ifcase\month\or
93       januara\or
94       februara\or
95       marta\or
96       aprila\or
97       maja\or
98       juna\or
99       jula\or
100      avgusta\or
101      septembra\or
102      oktobra\or
103      novembra\or
104      decembra\fi}%
105     \def\today{\number\day.\~\month@serbian\ \number\year\@ifstar{}{.}}%
106     \def\todayRoman{\number\day.\~\@Roman\month\~\number\year\@ifstar{}{.}}%
107     \def\todayGen{\number\day.\~\month@serbian@gen\ \number\year\@ifstar{}{.}}%
108     \def\todayArabic{\number\day.\~\number\month.\~\number\year\@ifstar{}{.}}
109   \let\dateserbian@datei=\dateserbian
110   \addto\dateserbian@datei{
111     \def\month@serbian@datei{\ifnum\month=6 juni%
112       \else\ifnum\month=7 juli\else\month@serbian\fi\fi}%
113     \def\today{\number\day.\~\month@serbian@datei\ \number\year\@ifstar{}{.}}
114 }

```

The Serbian hyphenation patterns can be used with `\lefthyphenmin` and `\righthyphenmin` set to 2. (Actually, the “official” definition allows even one character for `\lefthyphen`, but it is recommended to use value two for better results.)

```

115 \providehyphenmins{\CurrentOption}{\tw@\tw@}
116 \providehyphenmins{serbian}{\tw@\tw@}

```

We specify that the Serbian group of shorthands should be used.

```

117 \addto\extrasserbian{\languageshorthands{serbian}}
118 \addto\extrasserbian{\bbl@activate{}}
119 \addto\noextrasserbian{\bbl@deactivate{}}

```

Serbian typesetting requires frenchspacing. So, we add commands to `\extrasserbian` and `\noextrasserbian` to turn it on and off, respectively.

```

120 \addto\extrasserbian{\bbl@frenchspacing}
121 \addto\noextrasserbian{\bbl@nonfrenchspacing}

```

In Serbian the first paragraph of each section should be indented.

```

122 \let\@aifORI\@afterindentfalse
123 \def\bbl@serbianindent{\let\@afterindentfalse\@afterindenttrue\@afterindenttrue}
124 \def\bbl@nonserbianindent{\let\@afterindentfalse\@aifORI\@afterindentfalse}
125 \addto\extrasserbian{\bbl@serbianindent}
126 \addto\noextrasserbian{\bbl@nonserbianindent}

```

We redefine the macro `\Alph`, which now produces (uppercase) Latin letters without the letters q, w, x and y when Serbian is switched on, but we will keep the English alphabet if the user wants to use it. Also we will define Latin alphabet in order that matches Cyrillic alphabet. The user can choose which alphabet to use through the commands `\enumCyr`, `\enumLat` and `\enumEng` (or even to switch from one enumeration to another).

```

127 \newcount\srbl@lettering \srbl@lettering=\z@
128 \addto\extrasserbian{\babel@save\@Alph \let\@Alph\srbl@Alph}
129 \def\srbl@Alph#1{%
130 \ifcase\srbl@lettering
131   \ifcase#1\or A\or B\or C\or D\or E\or F\or G\or H\or I\or
132   J\or K\or L\or M\or N\or O\or P\or R\or S\or T\or U\or V\or
133   Z\else\@ctrerr\fi
134 \or
135   \if@srb@uni@ode
136     \ifcase#1\or A\or B\or V\or G\or D\or Đ\or E\or Ž\or Z\or
137     I\or J\or K\or L\or Lj\or M\or N\or Nj\or O\or
138     P\or R\or S\or T\or Ć\or U\or F\or H\or C\or
139     Ć\or Dž\or Š\else\@ctrerr\fi
140   \else
141     \ifcase#1\or A\or B\or V\or G\or D\or DJ\or E\or v Z\or Z\or
142     I\or J\or K\or L\or Lj\or M\or N\or Nj\or O\or
143     P\or R\or S\or T\or 'C\or U\or F\or H\or C\or
144     v C\or D\or v z\or v S\else\@ctrerr\fi
145   \fi
146 \or
147   \ifcase#1\or A\or B\or C\or D\or E\or F\or G\or H\or I\or
148   J\or K\or L\or M\or N\or O\or P\or Q\or R\or S\or T\or U\or V\or
149   W\or X\or Y\or Z\else\@ctrerr\fi
150 \fi}%

```

The same thing will be done with the macro `\alph`.

```

151 \addto\extrasserbian{\babel@save\@alph \let\@alph\srbl@alph}
152 \def\srbl@alph#1{%
153 \ifcase\srbl@lettering
154   \ifcase#1\or a\or b\or c\or d\or e\or f\or g\or h\or i\or
155   j\or k\or l\or m\or n\or o\or p\or r\or s\or t\or u\or v\or
156   z\else\@ctrerr\fi
157 \or
158   \if@srb@uni@ode
159     \ifcase#1\or a\or b\or v\or g\or d\or đ\or e\or ž\or z\or

```

```

160     i\or j\or k\or l\or lj\or m\or n\or nj\or o\or
161     p\or r\or s\or t\or ć\or u\or f\or h\or c\or
162     č\or dž\or š\else\@ctrerr\fi
163 \else
164     \ifcase#1\or a\or b\or v\or g\or d\or dj\or e\or\ v z\or z\or
165     i\or j\or k\or l\or lj\or m\or n\or nj\or o\or
166     p\or r\or s\or t\or\ 'c\or u\or f\or h\or c\or
167     \v c\or d\ v z\or\ v s\else\@ctrerr\fi
168 \fi
169 \or
170 \ifcase#1\or a\or b\or c\or d\or e\or f\or g\or h\or i\or
171     j\or k\or l\or m\or n\or o\or p\or q\or r\or s\or t\or u\or v\or
172     w\or x\or y\or z\else\@ctrerr\fi
173 \fi}%
174 \addto\extrasserbian{%
175 \babel@save\enumEng\def\enumEng{\srbl@lettering=\tw@}%
176 \babel@save\enumLat\def\enumLat{\srbl@lettering=\z@}%
177 \babel@save\enumCyr\def\enumCyr{\srbl@lettering=\one}%
178 }%

```

An `ijekav` attribute changes default behavior and activates an alternative set of captions suitable for typesetting in ijekavian dialect. The `quotes` attribute changes the `"` shorthand to produce `"` as closing quote, instead the traditional `“` quote of Serbian language. Also, the `datei` attribute will produce `'juni'` and `'juli'` instead `'jun'` and `'jul'` for date.

```

179 \bbl@declare@ttribute{serbian}{ijekav}{%
180 \PackageInfo{babel}{Serbian attribute set to ijekav}%
181 \let\captionsserbian=\captionsserbian@ijekav }
182 \@onlypreamble\captionsserbian@ijekav
183 \bbl@declare@ttribute{serbian}{quotes}{%
184 \PackageInfo{babel}{Serbian attribute set to quotes}%
185 \declare@shorthand{serbian}{"'}{\textquotedblright} }
186 \bbl@declare@ttribute{serbian}{datei}{%
187 \PackageInfo{babel}{Serbian attribute set to datei}%
188 \let\dateserbian=\dateserbian@datei }
189 \@onlypreamble\dateserbian@datei

```

Some math functions in Serbian math books have other names: e.g. `sinh` in Serbian is written as `sh` etc. So we define a number of new math operators.

```

190 \def\sh{\mathop{\operator@font sh}\nolimits}
191 \def\ch{\mathop{\operator@font ch}\nolimits}
192 \def\tg{\mathop{\operator@font tg}\nolimits}
193 \def\ctg{\mathop{\operator@font ctg}\nolimits}
194 \def\arctg{\mathop{\operator@font arctg}\nolimits}
195 \def\arcctg{\mathop{\operator@font arcctg}\nolimits}
196 \addto\extrasserbian{%
197 \babel@save{\th}%
198 \let\ltx@th\th
199 \def\th{\textormath{\ltx@th}}%
200     {\mathop{\operator@font th}\nolimits}}%
201 }
202 \def\cth{\mathop{\operator@font cth}\nolimits}
203 \def\arsh{\mathop{\operator@font arsh}\nolimits}
204 \def\arch{\mathop{\operator@font arch}\nolimits}
205 \def\arth{\mathop{\operator@font arth}\nolimits}
206 \def\arcth{\mathop{\operator@font arcth}\nolimits}
207 \def\cosec{\mathop{\operator@font cosec}\nolimits}
208 \def\arcsec{\mathop{\operator@font arcsec}\nolimits}
209 \def\arccosec{\mathop{\operator@font arccosec}\nolimits}
210 \def\sech{\mathop{\operator@font sech}\nolimits}
211 \def\cosech{\mathop{\operator@font cosech}\nolimits}

```

```

212 \def\arsech{\mathop{\operator@font arsech}\nolimits}
213 \def\arcosech{\mathop{\operator@font arcosech}\nolimits}
214 \def\Prob{\mathop{\kern\z@\mathsf{P}}\nolimits}
215 \def\Expect{\mathop{\kern\z@\mathsf{E}}\nolimits}
216 \def\Variance{\mathop{\kern\z@\mathsf{D}}\nolimits}
217 \def\nzs{\mathop{\operator@font nzs}\nolimits}
218 \def\nzd{\mathop{\operator@font nzd}\nolimits}
219 \def\NZS{\mathop{\operator@font NZS}\nolimits}
220 \def\NZD{\mathop{\operator@font NZD}\nolimits}

```

The macro `\ldf@finish` takes care of looking for a configuration file, setting the main language to be switched on at `\begin{document}` and resetting the category code of `@` to its original value.

```

221 \ldf@finish{serbian}
222 \code

```