

A Babel language definition file for Icelandic

Einar Árnason
einar@lif.hi.is

1 The Icelandic language

1.1 Overview

The file `iceland.dtx`¹ defines all the language definition macros for the Icelandic language

Customization for the Icelandic language was made following several official and semiofficial publications [2, 3, 1, 6, 5]. These publications do not always agree and we indicate those instances.

For this language the character " is made active. In table 1 an overview is given of its purpose. The shorthands in table 1 can also be typeset by using the commands in table 2.

References

- [1] Alþingi. *Reglur um frágang þingskjala og prentun umræðna*, 1988.
- [2] Auglýsing um greinarmerkjasetningu. Stj.tíð B, nr. 133/1974, 1974.
- [3] Auglýsing um breyting auglýsingu nr. 132/1974 um íslenska stafsetningu. Stj.tíð B, nr. 261/1977, 1977.
- [4] Einar Haugen, editor. *First Grammatical Treatise*. Longman, London, 2 edition, 1972.
- [5] Staðlaráð Íslands og Fagráð í upplýsingatækni, Reykjavík. *Forstaðall FS 130:1997*, 1997.
- [6] STRÍ Staðlaráð Íslands. *SI - kerfið*, 2 edition, 1994.

1.2 TeXnical details

When this file was read through the option `icelandic` we make it behave as if `icelandic` was specified.

```
1 \def\bb1@tempa{icelandic}
2 \ifx\CurrentOption\bb1@tempa
3   \def\CurrentOption{icelandic}
4 \fi
```

¹The file described in this section has version number ? and was last revised on ?.

"	disable ligature at this position.
"-	an explicit hyphen sign, allowing hyphenation in the rest of the word.
""	like "-", but producing no hyphen sign (for compound words with hyphen, e.g. x-"y).
"~	for a compound word mark without a breakpoint.
"=	for a compound word mark with a breakpoint, allowing hyphenation in the composing words.
"‘	for Icelandic left double quotes (looks like „).
"’	for Icelandic right double quotes.
">	for Icelandic ‘french’ left double quotes (similar to >>).
"<	for Icelandic ‘french’ right double quotes (similar to <<).
"o	for old Icelandic o
"O	for old Icelandic O
"ó	for old Icelandic ó
"Ó	for old Icelandic Ó
"e	for old Icelandic e
"E	for old Icelandic E
"é	for old Icelandic é
"É	for old Icelandic É
\tala	for typesetting numbers
\grada	for the ‘degree’ symbol
\gradur	for ‘degrees’, e.g. 5 °C
\upp	for textsuperscript

Table 1: The shorthands and extra definitions made by `icelandic.ldf`

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

```
5 ⟨*code⟩
6 \LdfInit\CurrentOption{captions\CurrentOption}
```

When this file is read as an option, i.e., by the `\usepackage` command, `icelandic` will be an ‘unknown’ language, so we have to make it known. So we check for the existence of `\l@icelandic` to see whether we have to do something here.

```
7 \ifx\l@icelandic\@undefined
8 \nopatterns{Icelandic}
9 \adddialect\l@icelandic0
10 \fi
```

`\if@Two@E` We will need a new ‘if’: `\if@Two@E` is true if and only if L^AT_EX 2_ε is running *not* in compatibility mode. It is used in the definitions of the command `\tala` and `\upp`.

<code>\ilqq</code>	for Icelandic left double quotes (looks like „).
<code>\irqq</code>	for Icelandic right double quotes (looks like “).
<code>\ilq</code>	for Icelandic left single quotes (looks like ,).
<code>\irq</code>	for Icelandic right single quotes (looks like ‘).
<code>\iflqq</code>	for Icelandic ‘french’ left double quotes (similar to >>).
<code>\ifrqq</code>	for Icelandic ‘french’ right double quotes (similar to <<).
<code>\ifrq</code>	for Icelandic ‘french’ right single quotes (similar to <).
<code>\iflq</code>	for Icelandic ‘french’ left single quotes (similar to >).
<code>\dq</code>	the original quotes character (“).
<code>\oob</code>	for old Icelandic o
<code>\Oob</code>	for old Icelandic O
<code>\oob</code>	for old Icelandic ó
<code>\OOb</code>	for old Icelandic Ó
<code>\eob</code>	for old Icelandic e
<code>\Eob</code>	for old Icelandic E
<code>\eeob</code>	for old Icelandic é
<code>\EEob</code>	for old Icelandic É

Table 2: Commands which produce quotes and old Icelandic diacritics, defined by `icelandic.ldf`

The definition is somewhat complicated, due to the fact that `\if@compatibility` is not recognized as a `\if` in L^AT_EX-2.09 based formats.

```

11 \newif\if@Two@E \@Two@Etrue
12 \def\@FI@{\fi}
13 \ifx\@compatibilitytrue\@undefined
14 \@Two@Efalse \def\@FI@\relax}
15 \else
16 \if@compatibility \@Two@Efalse \fi
17 \@FI@

```

`\extrasicelandic` The macro `\extrasicelandic` will perform all the extra definitions needed for the Icelandic language. The macro `\noextrasicelandic` is used to cancel the actions of `\extrasicelandic`.

For Icelandic the " character is made active. This is done once, later on its definition may vary.

```

18 \initiate@active@char{"}
19 \@namedef{extras\CurrentOption}{%
20 \languageshorthands{icelandic}}
21 \expandafter\addto\csname extras\CurrentOption\endcsname{%
22 \bbl@activate{"}}

```

Don't forget to turn the shorthands off again.

```
23 \addto\noextrasicelandic{\bbl@deactivate{}}
```

The icelandic hyphenation patterns can be used with `\lefthyphenmin` and `\righthyphenmin` set to 2.

```
24 \providehyphenmins{\CurrentOption}{\tw@}{\tw@}
```

The code above is necessary because we need an extra active character. This character is then used as indicated in table 2.

To be able to define the function of ", we first define a couple of 'support' macros.

1.3 Captionnames and date

The next step consists of defining the Icelandic equivalents for the L^AT_EX captionnames.

`\captionsicelandic` The macro `\captionsicelandic` will define all strings used used in the four standard document classes provided with L^AT_EX.

```
25 \@namedef{captions\CurrentOption}{%
26   \def\prefacename{Form\'}{a}li}%
27   \def\refname{Heimildir}%
28   \def\abstractname{\'}{U}tdr\'}{a}ttur}%
29   \def\bibname{Heimildir}%
30   \def\chaptername{Kafli}%
31   \def\appendixname{Vi\'}{dh}auki}%
32   \def\contentsname{Efnisyfirlit}%
33   \def\listfigurename{Myndaskr\'}{a}}%
34   \def\listtablename{T\'}{o}fluskr\'}{a}}%
35   \def\indexname{Atri\'}{dh}isor\'}{dh}askr\'}{a}}%
36   \def\figurename{Mynd}%
37   \def\tablename{Tafla}%
38   \def\partname{Hluti}%
39   \def\enclname{Hj\'}{a}lagt}%
40   \def\ccname{Samrit}%
41   \def\headtoname{Til:}% in letter
42   \def\pagename{Bla\'}{dh}s\'}{i}{\'}{dh}a}%
43   \def\seenname{Sj\'}{a}}%
44   \def\alsoname{Sj\'}{a} einnig}%
45   \def\proofname{S\'}{o}nnun}%
46   \def\glossaryname{Or\'}{dh}alisti}%
47 }
```

`\dateicelandic` The macro `\dateicelandic` redefines the command `\today` to produce Icelandic dates.

```
48 \def\dateicelandic{%
49   \def\today{\number\day.~\ifcase\month\or
50     jan\'}{u}ar\or febr\'}{u}ar\or mars\or apr\'}{i}1\or ma\'}{i}\or
51     j\'}{u}n\'}{i}\or j\'}{u}1\'}{i}\or \'}{a}g\'}{u}st\or september\or
52     okt\'}{o}ber\or n\'}{o}vember\or desember\fi
53     \space\number\year}}
```

1.4 Icelandic quotation marks

`\dq` We save the original double quote character in `\dq` to keep it available, the math accent `\"` can now be typed as `"`.

```
54 \begingroup \catcode'\ "12
55 \def\x{\endgroup
56 \def@SS{\mathchar"7019 }
57 \def\dq{"}}
58 \x
```

Now we can define the icelandic and icelandic ‘french’ quotes. The icelandic ‘french’ guillemets are the reverse of french guillemets. We define single icelandic ‘french’ quotes for compatibility. Shorthands are provided for a number of different quotation marks, which make them useable both outside and inside mathmode.

```
59 \let\ilq\grq
60 \let\irq\grq
61 \let\iflq\frq
62 \let\ifrqq\flq
63 \let\ilqq\glqq
64 \let\irqq\grqq
65 \let\iflqq\frqq
66 \let\ifrqq\flqq

67 \declare@shorthand{icelandic}{"'}{\glqq}
68 \declare@shorthand{icelandic}{"'}{\grqq}
69 \declare@shorthand{icelandic}{">}{\frqq}
70 \declare@shorthand{icelandic}{"<}{\flqq}
```

and some additional commands:

```
71 \declare@shorthand{icelandic}{"-}{\nobreak\-\bbl@allowhyphens}
72 \declare@shorthand{icelandic}{"}{\%}
73 \textormath{\nobreak\discretionary{-}{-}{\kern.03em}%
74 \bbl@allowhyphens}{-}}
75 \declare@shorthand{icelandic}{""}{\hskip\z@skip}
76 \declare@shorthand{icelandic}{""~}{\textormath{\leavevmode\hbox{-}}{-}}
77 \declare@shorthand{icelandic}{"=}{\nobreak-\hskip\z@skip}
```

1.5 Old Icelandic

In old Icelandic some letters have special diacritical marks, described for example in *First Grammatical Treatise* [4, 5]. We provide these in the T1 encoding with the ‘ogonek’. The ogonek is placed with the letters ‘o’, and ‘O’, ‘ó’ and ‘Ó’, ‘e’ and ‘E’, and ‘é’ and ‘É’. Shorthands are provided for these as well.

The following code by Leszek Holenderski lifted from `polish.dtx` is designed to position the diacritics correctly for every font in every size. These macros need a few extra dimension variables.

```
78 \newdimen\pl@left
79 \newdimen\pl@down
80 \newdimen\pl@right
```

```

81 \newdimen\pl@temp

\sob The macro \sob is used to put the ‘ogonek’ in the right place.

82 \def\sob#1#2#3#4#5{%parameters: letter and fractions hl,ho,vl,vo
83 \setbox0\hbox{#1}\setbox1\hbox{\k{}}\setbox2\hbox{p}%
84 \pl@right=#2\wd0 \advance\pl@right by-#3\wd1
85 \pl@down=#5\ht1 \advance\pl@down by-#4\ht0
86 \pl@left=\pl@right \advance\pl@left by\wd1
87 \pl@temp=-\pl@down \advance\pl@temp by\dp2 \dp1=\pl@temp
88 \leavevmode
89 \kern\pl@right\lower\pl@down\box1\kern-\pl@left #1}

\sob
\Ob 90 \DeclareTextCommand{\sob}{T1}{\sob {o}{.85}{0}{.04}{0}}
\oob 91 \DeclareTextCommand{\Ob}{T1}{\sob {0}{.7}{0}{0}{0}}
\OOob 92 \DeclareTextCommand{\ooob}{T1}{\sob {ó}{.85}{0}{.04}{0}}
\eob 93 \DeclareTextCommand{\OOob}{T1}{\sob {Ó}{.7}{0}{0}{0}}
\Eob 94 \DeclareTextCommand{\eob}{T1}{\sob {e}{1}{0}{.04}{0}}
\eeob 95 \DeclareTextCommand{\Eob}{T1}{\sob {E}{1}{0}{.04}{0}}
\EEob 96 \DeclareTextCommand{\eeob}{T1}{\sob {é}{1}{0}{.04}{0}}
97 \DeclareTextCommand{\EEob}{T1}{\sob {É}{1}{0}{.04}{0}}

98 \declare@shorthand{icelandic}{"o}{\sob}
99 \declare@shorthand{icelandic}{"0}{\Ob}
100 \declare@shorthand{icelandic}{"ó}{\oob}
101 \declare@shorthand{icelandic}{"Ó}{\OOob}
102 \declare@shorthand{icelandic}{"e}{\eob}
103 \declare@shorthand{icelandic}{"E}{\Eob}
104 \declare@shorthand{icelandic}{"é}{\eeob}
105 \declare@shorthand{icelandic}{"É}{\EEob}

```

1.6 Formatting numbers

This section is lifted from `frenchb.dtx` by D. Flipo. In English the decimal part starts with a point and thousands should be separated by a comma: an approximation of 1000π should be inputed as $\$3{,}141.592{,}653\$$ in math-mode and as 3,141.592,653 in text.

In Icelandic the decimal part starts with a comma and thousands should be separated by a space [1] or by a period [5]; we have the space. The above approximation of 1000π should be inputed as $\$3\ ;141{,}592\ ;653\$$ in math-mode and as something like $3\sim 141,592\sim 653$ in text. Braces are mandatory around the comma in math-mode, the reason is mentioned in the `TEXbook` p. 134: the comma is of type `\mathpunct` (thus normally followed by a space) while the point is of type `\mathord` (no space added).

Thierry Bouche suggested that a second type of comma, of type `\mathord` would be useful in math-mode, and proposed to introduce a command (named `\decimalsep` in this package), the expansion of which would depend on the current language.

Vincent Jalby suggested a command `\nombre` to conveniently typeset numbers: inputting `\nombre{3141,592653}` either in text or in math-mode will format this number properly according to the current language (Icelandic or non-Icelandic). We use `\nombre` to define command `\tala` in Icelandic.

`\tala` accepts an optional argument which happens to be useful with the extension ‘dcolumn’, it specifies the decimal separator used in the *source code*:

```
\newcolumnmtype{d}{D{,}{\decimalsep}{-1}}
\begin{tabular}{d}\hline
  3,14 \\\
  \tala[,]{123,4567} \\\
  \tala[,]{9876,543}\\\hline
\end{tabular}
```

will print a column of numbers aligned on the decimal point (comma or point depending on the current language), each slice of 3 digits being separated by a space or a comma according to the current language.

`\decimalsep` We need a internal definition, valid in both text and math-mode, for the comma
`\thousandsep` (`\@comma@`) and another one for the unbreakable fixed length space (no glue) used in Icelandic (`\f@thousandsep`).

The commands `\decimalsep` and `\thousandsep` get default definitions (for the English language) when `icelandic` is loaded; these definitions will be updated when the current language is switched to or from Icelandic.

```
106 \mathchardef\m@comma="013B \def\@comma@{\ifmmode\m@comma\else,\fi}
107 \def\f@thousandsep{\ifmmode\mskip5.5mu\else\penalty\M\kern.3em\fi}
108 \newcommand{\decimalsep}{.} \newcommand{\thousandsep}{\@comma@}
109 \expandafter\addto\csname extras\CurrentOption\endcsname{%
110     \def\decimalsep{\@comma@}%
111     \def\thousandsep{\f@thousandsep}}
112 \expandafter\addto\csname noextras\CurrentOption\endcsname{%
113     \def\decimalsep{.}%
114     \def\thousandsep{\@comma@}}
```

`\tala` The decimal separator used when *inputting* a number with `\tala` has to be a *comma*. `\tala` splits the inputted number into two parts: what comes before the first comma will be formatted by `\@integerpart` while the rest (if not empty) will be formatted by `\@decimalpart`. Both parts, once formatted separately will be merged together with between them, either the decimal separator `\decimalsep` or (in \LaTeX 2_ϵ *only*) the optional argument of `\tala`.

```
115 \if@Two@E
116 \newcommand{\tala}[2][\decimalsep]{%
117     \def\@decimalsep{#1}\@tala#2\@empty,\@empty,\@nil}
118 \else
119 \newcommand{\tala}[1]{%
120     \def\@decimalsep{\decimalsep}\@tala#1\@empty,\@empty,\@nil}
121 \fi
122 \def\@tala#1,#2,#3\@nil{%
123     \ifx\@empty#2%
```

```

124     \@integerpart{#1}%
125     \else
126     \@integerpart{#1}\@decimalsep\@decimalpart{#2}%
127     \fi}

```

The easiest bit is the decimal part: We attempt to read the first four digits of the decimal part, if it has less than 4 digits, we just have to print them, otherwise `\thousandsep` has to be appended after the third digit, and the algorithm is applied recursively to the rest of the decimal part.

```

128 \def\@decimalpart#1{\@decimalpart#1\@empty\@empty\@empty}
129 \def\@decimalpart#1#2#3#4{#1#2#3%
130   \ifx\@empty#4%
131   \else
132     \thousandsep\expandafter\@decimalpart\expandafter#4%
133   \fi}

```

Formatting the integer part is more difficult because the slices of 3 digits start from the *bottom* while the number is read from the top! This (tricky) code is borrowed from David Carlisle's `comma.sty`.

```

134 \def\@integerpart#1{\@integerpart{ }#1\@empty\@empty\@empty}
135 \def\@integerpart#1#2#3#4{%
136   \ifx\@empty#2%
137     \@addthousandsep#1\relax
138   \else
139     \ifx\@empty#3%
140       \@addthousandsep\@empty\@empty#1#2\relax
141     \else
142       \ifx\@empty#4%
143         \@addthousandsep\@empty#1#2#3\relax
144       \else
145         \@integerpartafterfi{#1#2#3#4}%
146       \fi
147     \fi
148   \fi}
149 \def\@integerpartafterfi#1\fi\fi\fi{\fi\fi\fi\@integerpart{#1}}
150 \def\@addthousandsep#1#2#3#4{#1#2#3%
151   \if#4\relax
152   \else
153     \thousandsep\expandafter\@addthousandsep\expandafter#4%
154   \fi}

```

1.7 Extra utilities

We now provide the Icelandic user with some extra utilities.

`\upp` `\upp` is for typesetting superscripts. `\upp` relies on `\upp@size` The internal macro `\upp@size` holds the size at which the superscript will be typeset. The reason for this is that we have to specify it differently for different formats.


```

155 \ifx\sevenrm\@undefined
156 \ifx\@ptsize\@undefined
157   \let\upp@size\small
158 \else
159   \ifx\selectfont\@undefined

```

In this case the format is the original L^AT_EX-2.09:

```

160   \ifcase\@ptsize
161     \let\upp@size\ixpt\or
162     \let\upp@size\xpt\or
163     \let\upp@size\xipt
164   \fi

```

When `\selectfont` is defined we probably have NFSS available:

```

165   \else
166     \ifcase\@ptsize
167       \def\upp@size{\fontsize\@ixpt{10pt}\selectfont}\or
168       \def\upp@size{\fontsize\@xpt{11pt}\selectfont}\or
169       \def\upp@size{\fontsize\@xipt{12pt}\selectfont}
170     \fi
171   \fi
172 \fi
173 \else

```

If we end up here it must be a plain based T_EX format, so:

```

174   \let\upp@size\sevenrm
175 \fi

```

Now we can define `\upp`. When L^AT_EX 2_ε runs in compatibility mode (L^AT_EX-2.09 emulation), `\textsuperscript` is also defined, but does no good job, so we give two different definitions for `\upp` using `\if@Two@E`.

```

176 \if@Two@E
177   \DeclareRobustCommand*\upp}[1]{\textsuperscript{#1}}
178 \else
179   \DeclareRobustCommand*\upp}[1]{%
180     \leavevmode\raise1ex\hbox{\upp@size#1}}
181 \fi

```

Some definitions for special characters. `\grada` needs a special treatment: it is `\char6` in T1-encoding and `\char23` in OT1-encoding.

```

182 \ifx\fmtname\LaTeXeFmtName
183   \DeclareTextSymbol{\grada}{T1}{6}
184   \DeclareTextSymbol{\grada}{OT1}{23}
185 \else
186   \def\T@one{T1}
187   \ifx\fontencoding\T@one
188     \newcommand{\grada}{\char6}
189   \else
190     \newcommand{\grada}{\char23}
191   \fi
192 \fi

```

`\gradur` Macro for typesetting the abbreviation for ‘degrees’ (as in ‘degrees Celsius’). As the bounding box of the character ‘degree’ has *very* different widths in CMR/DC and PostScript fonts, we fix the width of the bounding box of `\gradur` to 0.3em, this lets the symbol ‘degree’ stick to the preceding (e.g., `45\gradur`) or following character (e.g., `20~\gradur C`).

```
193 \DeclareRobustCommand*\gradur}{%  
194         \leavevmode\hbox to 0.3em{\hss\grada\hss}}
```

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.

```
195 \ldf@finish\CurrentOption  
196 \code)
```