

# The luavlna package

Michal Hoftich (michal.h21@gmail.com)

Miro Hrončok (miro@hroncok.cz)

September 22, 2016

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Usage</b>	<b>2</b>
<b>3</b>	<b>Commands</b>	<b>3</b>
3.1	Turning off language switching . . . . .	4
3.2	Turning off processing . . . . .	4
3.3	Debugging commands . . . . .	4
<b>4</b>	<b>Lua module langno.lua</b>	<b>4</b>
4.1	Recognized languages . . . . .	5
4.1.1	luaTeX and lua <del>La</del> TeX . . . . .	5
4.1.2	CSplain . . . . .	5

## 1 Introduction

This is a small package for plain luaTeX and lua~~La~~TeX. In some languages, like Czech or Polish, there should be no single letter words at the line end, according to the typographical norms. There exists some external commands (like vlna) or packages (encxvlna for encTeX, xevlna for XeTeX, impnattypo for lua~~La~~TeX).

Other feature of this package is including of non-breakable space after initials, like in personal names, after or before academic degrees and between numbers and units (SI and others).

The code is modified version of Patrick Gundlach's answer on TeX.sx<sup>1</sup>. The difference is that it is possible to specify which single letters should be taken into account for different languages. The support for degrees and units was added as well.

## 2 Usage

The usage is simple:

```
\input ucode
\uselanguage{czech}
% in the case of luacsplain, use instead:
% \chyph
\input luavlna
\preventsingleddebugon
\input luaotfload.sty
\font\hello={name:Linux Libertine 0:+rlig;+clig;+liga;+tlig} at 12pt
\hsize=3in
\hello
Příliš žluťoučký kůň úpěl ďábelské ódy.
Text s krátkými souhláskami a samohláskami i dalšími jevy
z nabídky možností (v textu možnými).

I začátek odstavce je třeba řešit, i když výskyt zalomení není pravděpodobný.

Co třeba í znaky š diakritikou?

Různé možnosti [v závorkách <i jiných znacích

Podpora iniciál a titulů: M. J. Hegel, Ing. Běháková, Ph.D., Ž. Zíbrt.

Podpora jednotek: 100,5 MN\cdot{}s, 100.5 kJ, 200 μA, $-1$ dag, 12 MiB, 1 m$^3$/s.

\preventsingleddebugoff
\bye
```

---

<sup>1</sup><http://tex.stackexchange.com/a/28128/2891>

Příliš *žlutoučký kůň* úpěl ďábelské ódy. Text s krátkými souhláskami a samohláskami i dalšími jevy z nabídky možností (v textu možnými).

I začátek odstavce je třeba řešit, i když výskyt zalomení není pravděpodobný.

Co třeba í znaky š diakritikou?

Různé možnosti [v závorkách < i jiných znacích

Podpora iniciál a titulů: M. J. Hegel, Ing. Běháková, Ph.D., Ž. Zíbrt.

Podpora jednotek: 100,5 MN·s, 100.5 kJ, 200 μA, −1 dag, 12 MiB, 1 m<sup>3</sup>/s.

It is also possible to use the package with  $\text{lua}\text{\LaTeX}$ , just use

```
\usepackage{luavlna}
```

in the preamble.

### 3 Commands

```
\singlechars{<language name>}{<letters>}
```

Enable this feature for certain letters in given language.

Default values:

```
% only Czech and Slovak are supported out of the box
```

```
\singlechars{czech}{AIiVvOoUuSsZzKk}
```

```
\singlechars{slovak}{AIiVvOoUuSsZzKk}
```

```
\compoundinitials{<language name>}{<compounds>}
```

Declare compound letters for given language. Second argument should be comma separated list of compound letters, in exact form in which they can appear.

Default values:

```
\compoundinitials{czech}{Ch,CH}
```

### 3.1 Turning off language switching

By default, language of the nodes is taken into account. If you want to use settings for one language for a whole document, you can use following command:

```
\preventsingl-lang{<language name>}
```

### 3.2 Turning off processing

If you want to stop processing of the spaces in the document you can use command

```
\preventsingleoff
```

To resume processing, use

```
\preventsingleon
```

You can also disable units and degrees processing:

```
\nounits
```

```
\noprededegrees
```

```
\nosufdegrees
```

### 3.3 Debugging commands

```
\preventsingledbugon
```

```
\preventsingledbugoff
```

Insert debugging marks on/off. Default off.

## 4 Lua module langno.lua

When we process glyph nodes with lua<sub>TeX</sub> callbacks, there are lang fields available. These are numerical codes of languages, but no information about language names easily accesible from lua side is available.<sup>2</sup> These numbers are format dependent, majority of formats like lua<sub>TeX</sub> use

---

<sup>2</sup>Language names are stored in <sub>TeX</sub>macros like \l@langname, but different formats use different naming of these macros

language.dat file provided by babel for assign numbers to languages, but for example csplain use its own system.

To allow easy setting of language dependent parameters, langno module was created. It's purpose is to translate language code to language name and the other way. lua $\LaTeX$ , lua $\TeX$  and CSplain formats are supported at the moment.

## 4.1 Recognized languages

### 4.1.1 lua $\TeX$ and lua $\LaTeX$

File language.dat is processed to load language names, aliases and assigned numbers. These language names are the same as supported by babel package.

UKenglish, USenglish, afrikaans, american, amharic, ancientgreek, arabic, armenian, assamese, basque, bengali, bokmal, british, bulgarian, catalan, churchslavonic, classiclatin, coptic, croatian, czech, danish, dumylang, dutch, english, espanol, esperanto, estonian, ethiopic, farsi, finnish, francais, french, friulan, galician, geez, georgian, german, greek, gujarati, hindi, hungarian, ibycus, icelandic, indonesian, interlingua, irish, italian, kannada, kurmanji, latin, latvian, lithuanian, liturgicallatin, malayalam, marathi, mongolian, mongolianlmc, monogreek, ngerman, nohyphenation, norsk, norwegian, nynorsk, occitan, oriya, panjabi, patois, persian, piedmontese, pinyin, polish, polygreek, portuges, portuguese, romanian, romansh, russian, sanskrit, serbian, serbianc, slovak, slovene, slovenian, spanish, swedish, swissgerman, tamil, telugu, thai, turkish, turkmen, ukenglish, ukrainian, uppersorbian, usenglish, usenglishmax, welsh

### 4.1.2 CSplain

Different method is used. File hyphen.1an is included in CSplain, where language numbers are assigned to ISO-639-1 or ISO-639-2 language codes. These language codes were then normalized to names used by babel, or standard English language names.

afrikaans, armenian, assamese, basque, bokmål, catalan, chinese, coptic, croatian, czech, danish, dutch, english, estonian, finnish,

french, galician, german, greek, hindi, icelandic, indonesian, inter-  
lingua , irish, italian, kannada, kurdish, lao, latin, latvian, lithua-  
nian, magyar, malayalam, marathi, mongolian, norsk, oriya, pan-  
jabi, polish, portuges, romanian, russian, sanskrit, slovak, slove-  
nian, spanish, tamil, telugu, turkish, turkmen, ukrainian, upper-  
sorbian, welsh