

Package ‘koRpus.lang.nl’

July 11, 2016

Type Package

Title Language support for koRpus: Dutch

Author m.eik michalke [aut, cre], Christophe Bechet [ctb]

Maintainer m.eik michalke <meik.michalke@hhu.de>

Depends R (>= 2.10.0),koRpus (>= 0.07-1),
methods

Description Adds support for the Dutch language to the koRpus package. Due to some restrictions on CRAN, the full package sources are only available from the project homepage. To ask for help, report bugs, suggest feature improvements, or discuss the global development of the package, please subscribe to the koRpus-dev mailing list: https://ml06.ispgateway.de/mailman/listinfo/korpus-dev_r.reaktanz.de

License GPL (>= 3)

Encoding UTF-8

LazyLoad yes

URL <http://reaktanz.de/?c=hacking&s=koRpus>

Version 0.01-3

Date 2016-07-11

RoxygenNote 5.0.1

R topics documented:

koRpus.lang.nl-package	2
hyph.nl	2
lang.support.nl	3

Index	5
--------------	----------

koRpus.lang.nl-package

The koRpus.lang.nl Package

Description

Language support for koRpus: Dutch.

Details

Package: koRpus.lang.nl
Type: Package
Version: 0.01-3
Date: 2016-07-11
Depends: R (>= 2.10.0),koRpus (>= 0.07-1), methods
Encoding: UTF-8
License: GPL (>= 3)
LazyLoad: yes
URL: <http://reaktanz.de/?c=hacking&s=koRpus>

Adds support for the Dutch language to the koRpus package. Due to some restrictions on CRAN, the full package sources are only available from the project homepage. To ask for help, report bugs, suggest feature improvements, or discuss the global development of the package, please subscribe to the koRpus-dev mailing list: https://ml06.ispgateway.de/mailman/listinfo/korpus-dev_r.reaktanz.de

Author(s)

m.eik michalke, with contributions from Christophe Bechet

hyph.nl

Hyphenation patterns for Dutch

Description

Hyphenation patterns for Dutch to be used by [hyphen](#). These data objects are not really intended to be used directly, but rather to be consulted by the `hyphen()` function without further user interaction.

Usage

hyph.nl

Format

The pattern slot of each hyphenation pattern object has three columns:

orig The original pattern in patgen style format.

char Only the character elements of the pattern which can be matched to parts of an actual word.

nums A code of digits defining the possibility to split syllables at respective places in this pattern.

Source

The patterns (as they are present in the "orig" column described above) were originally provided by the LaTeX developers[1], under the terms of the LaTeX Project Public License[2]. Refer to Liang (1983) for a detailed explanation. From these original patterns the values in the remaining columns were created using [read.hyph.pat](#).

In case any changes to the patterns were necessary to be used in this package, they are fully documented in the file "ChangeLog_hyph_patterns.txt" in the sources for this package. The unchanged original patterns can be found under [1].

References

Liang, F.M. (1983). *Word Hy-phen-a-tion by Com-put-er*. Dissertation, Stanford University, Dept. of Computer Science.

[1] <http://tug.ctan.org/tex-archive/language/hyph-utf8/tex/generic/hyph-utf8/patterns/>

[2] <http://www.ctan.org/tex-archive/macros/latex/base/lppl.txt>

See Also

[read.hyph.pat](#), [manage.hyph.pat](#)

lang.support.nl

Language support for Dutch

Description

This function adds support for Dutch to the koRpus package. You should not need to call it manually, as that is done automatically when this package is loaded.

Usage

lang.support.nl()

Details

In particular, this function adds the following:

- `lang`: The additional language "nl" to be used with `koRpus`
- `hyphen`: An additional set of hyphenation patterns (see [hyph.nl](#))
- `treetag`: The additional preset "nl-utf8", implemented according to the respective `TreeTagger[1]` script
- `POS_tags`: An additional set of tags, implemented using the documentation for the corresponding `TreeTagger` parameter set[2]

References

[1] <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

[2] <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/dutch-tagset.txt>

Index

*Topic **datasets**

hyph.nl, [2](#)

*Topic **package**

koRpus.lang.nl-package, [2](#)

hyph.nl, [2](#), [4](#)

hyphen, [2](#)

koRpus.lang.nl-package, [2](#)

lang.support.nl, [3](#)

manage.hyph.pat, [3](#)

read.hyph.pat, [3](#)