Package: nsyllable (via r-universe)

October 16, 2024

Version 1.0.1

Title Count Syllables in Character Vectors

Description Counts syllables in character vectors for English words.
Imputes syllables as the number of vowel sequences for words
not found.

License GPL-3

Depends R (>= 2.10)

Suggests knitr, testthat, spelling

URL <https://github.com/quanteda/nsyllable>

Encoding UTF-8

BugReports <https://github.com/quanteda/nsyllable/issues>

LazyData TRUE

Language en-GB

RoxygenNote 7.1.2

Roxygen list(markdown = TRUE)

Repository https://quanteda.r-universe.dev

RemoteUrl https://github.com/quanteda/nsyllable

RemoteRef HEAD

RemoteSha 124d2cf48c11387f98b3b19935e925213f9bd523

Contents

---

nsyllable-package   *nsyllable: Count syllables in character vectors*

Description

Counts syllables from character vector inputs. Imputes syllables as the number of vowel sequences for words not found.

Author(s)

Kenneth Benoit

See Also

Useful links:

- <https://github.com/quanteda/nsyllable>
- Report bugs at <https://github.com/quanteda/nsyllable/issues>

---

data_syllables_en   *Syllable counts of English words*

Description

A named integer vector of syllable counts for English words. Based on a pronunciation dictionary for North American English that contains over 134,000 words and their pronunciations, from the [Carnegie Mellon University Pronouncing Dictionary](#).

Usage

```
data_syllables_en
```

Format

An object of class `integer` of length 125698.

Note

`data_syllables_en` is a data object consisting of a named numeric vector of syllable counts for the words used as names. This is the default object used to count English syllables. For words with multiple pronunciation variants, we use the first entry.

This object that can be accessed directly, but we strongly encourage you to access it only through the [nsyllable()](#) wrapper function.

Source

Version 0.7b of the CMU Pronouncing Dictionary. See <https://github.com/cmusphinx/cmudict>.

---

nsyllable   *Count syllables in a text*

Description

Returns a count of the number of syllables in texts. For English words, the syllable count is exact and looked up from the CMU pronunciation dictionary, from the default syllable dictionary `data_int_syllables`. For any word not in the dictionary, the syllable count is estimated by counting vowel clusters.

Usage

```
nsyllable(x, language = "en", syllable_dictionary = NULL, use.names = FALSE)
```

Arguments

x	character vector whose syllables will be counted. This will count all syllables in a character vector without regard to separating tokens, so it is recommended that x be individual terms.
language	specify the language for syllable counts by ISO 639-1 code. The default is English, using the data object `data_syllables_en`, an English pronunciation dictionary from CMU.
syllable_dictionary	
	optional named integer vector of syllable counts where the names are lower case tokens. This can be used to override the language setting, when set to `NULL` (the default). If a syllable dictionary is supplied, this will override the `language` argument.
use.names	logical; if `TRUE`, assign the tokens as the names of the syllable count vector

Value

an integer vector of the counts of the syllables in each element, named with the element if `use.names = TRUE`

Examples

```
# character
nsyllable(c("cat", "syllable", "supercalifragilisticexpialidocious",
            "Brexit", "Administration"), use.names = TRUE)
```

Index