
narmer Documentation

Release 0.1.2

Christopher C. Little

Sep 30, 2018

Contents:

1 narmer package	1
1.1 Submodules	1
1.1.1 narmer.phonetic module	1
1.1.2 narmer.stats module	3
2 narmer	5
3 Indices and tables	7
3.1 Bibliography	7
Python Module Index	9

narmer.

Narmer NLP/IR library by Christopher C. Little

This library contains code I'm using for research, in particular dissertation research & experimentation.

Further documentation to come. . .

1.1 Submodules

1.1.1 narmer.phonetic module

narmer.phonetic.

The phonetic module implements phonetic algorithms including:

- `german_ipa`

`narmer.phonetic.enhg_ipa` (*word*)

Convert Early New High German to IPA.

This is based on TODO

Parameters `word` (*str*) – the ENHG word to transcribe to IPA

Returns the ENHG word's approximate IPA equivalent

Return type `str`

`narmer.phonetic.german_ipa` (*word*, *period='u'nhg'*)

Convert German to IPA.

Wrapper for other, more specific functions to convert German of various periods to IPA.

Parameters

- **word** (*str*) – the German word to transcribe to IPA

- **period** (*str*) – a period of German from the set:
 - nhg (default) – New High German
 - enhg – Early New High German
 - mhg – Middle High German
 - ohg – Old High German

Returns the German word’s approximate IPA equivalent

Return type str

```
>>> german_ipa('Ehre')
'ere'
>>> german_ipa('Kohl')
'kol'
>>> german_ipa('Schiffahrt')
'iffart'
>>> german_ipa('Schiller')
'iller'
>>> german_ipa('Tschechien')
'teçin'
```

`narmer.phonetic.mhg_ipa` (*word*)

Convert Middle High German to IPA.

This is based on <http://users.clas.ufl.edu/hasty/resources/CHAPTER1.HTM>

Parameters **word** (*str*) – the ENHG word to transcribe to IPA

Returns the ENHG word’s approximate IPA equivalent

Return type str

`narmer.phonetic.nhg_ipa` (*word*)

Convert New High German to IPA.

This is based largely on the orthographic mapping described at: https://en.wikipedia.org/wiki/German_orthography

No significant attempt is made to accommodate loanwords.

Parameters **word** (*str*) – the NHG word to transcribe to IPA

Returns the NHG word’s approximate IPA equivalent

Return type str

```
>>> nhg_ipa('Ehre')
'ere'
>>> nhg_ipa('Kohl')
'kol'
>>> nhg_ipa('Schiffahrt')
'iffart'
>>> nhg_ipa('Schiller')
'iller'
>>> nhg_ipa('Tschechien')
'teçin'
```

`narmer.phonetic.ohg_ipa` (*word*)

Convert Old High German to IPA.

This is based on TODO

Parameters `word` (*str*) – the ENHG word to transcribe to IPA

Returns the ENHG word’s approximate IPA equivalent

Return type `str`

1.1.2 narmer.stats module

`narmer.stats`.

The stats module defines functions for calculating various statistical data about linguistic objects, including:

- Weissman score calculation

`narmer.stats.weissman` (*r_tar*, *t_tar*, *r_src*, *t_src*, *alpha=1.0*)

Calculate Weissman score based on entered statistics.

The score is: $W =$

$c \cdot$

$\frac{r_{tar} r_{src}}{c \cdot}$

$c \cdot$

$\frac{\log t_{src}}{\log t_{tar}}$

In practice, the score can be used to rate time-intensive tasks on the basis of other metrics, also, e.g. F_1 score.

Sources: <http://spectrum.ieee.org/view-from-the-valley/computing/software/a-madefortv-compression-metric-moves-to-the-real-world>

Parameters

- **r_tar** (*float*) – the target algorithm’s compression ratio
- **t_tar** (*float*) – the target algorithm’s compression time
- **r_src** (*float*) – a standard algorithm’s compression ratio
- **t_src** (*float*) – a standard algorithm’s compression time
- **alpha** (*float*) – a scaling constant (1.0 by default)

Returns the Weissman score

Return type `float`

```
>>> weissman(1, 1, 1, 1)
1.0
>>> weissman(1, 1, 1, 5)
7248263982714164.0
>>> weissman(1.2, 1.6, 4.8, 5)
0.8560773855177113
>>> weissman(1, 1, 1, 1, alpha=2)
2.0
>>> weissman(1.2, 1.6, 4.8, 5, alpha=2)
1.7121547710354226
```


CHAPTER 2

narmer

- `genindex`
- `modindex`
- `search`

3.1 Bibliography

References

n

`narmer`, 1
`narmer.phonetic`, 1
`narmer.stats`, 3

E

enhg_ipa() (in module narmer.phonetic), 1

G

german_ipa() (in module narmer.phonetic), 1

M

mhg_ipa() (in module narmer.phonetic), 2

N

narmer (module), 1

narmer.phonetic (module), 1

narmer.stats (module), 3

nhg_ipa() (in module narmer.phonetic), 2

O

ohg_ipa() (in module narmer.phonetic), 2

W

weissman() (in module narmer.stats), 3