

GB



[ology@github](mailto:ology@github)

Search this website .

# Inspecting American Inaugural Addresses with Perl and R

2016-10-09 BY GENE

Given all the inaugural addresses of American presidents, what are the readability stats? What is the sentiment over time?

**UPDATE:** Results charted for 2017

As usual I reach for perl to acquire and format the data for exploration with R. The code below (and [available on github](#)) reads and analyzes a collection of text documents. It then prints out the results found in a tab-delimited format for consumption with R.

```
#!/usr/bin/env perl
use strict;
use warnings;

use File::Basename;
use Lingua::EN::Fathom;
use File::Find::Rule;

my $path = shift || die "Usage: perl $0 /path/to/files";
```

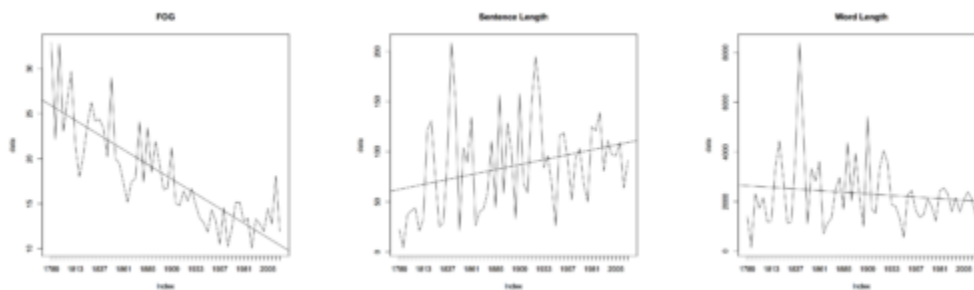


```

abplot( inagural$Chars, 'Raw Length' )
abplot( inagural$Words, 'Word Length' )
abplot( inagural$Sentences, 'Sentence Length' )
abplot( inagural$WpS, 'Inagural Addresses Words per Sentence' )
abplot( inagural$FOG, 'FOG' )
abplot( inagural$Flesch, 'Flesch' )
abplot( inagural$Kincaid, 'Kincaid' )

```

In the word and char plots, we see that the speeches have slowly said less, but that the number of sentences used has increased. Also of note is that the grade-level readability has dropped dramatically. Here are the plots:



So what about the emotional sentiment over time? There is a fabulous R package for this!

```

library(syuzhet)

valence <- function (file) {
  content <- get_text_as_string(file)
  sentences <- get_sentences(content)
  sentiment_vector <- get_sentiment( sentences, method = 'syuzhet' )
  plot( sentiment_vector, type = 'l', main = file, xlab = 'Narrative Time'
  )
}

```

With this function in place, we can process each address into plots.

```

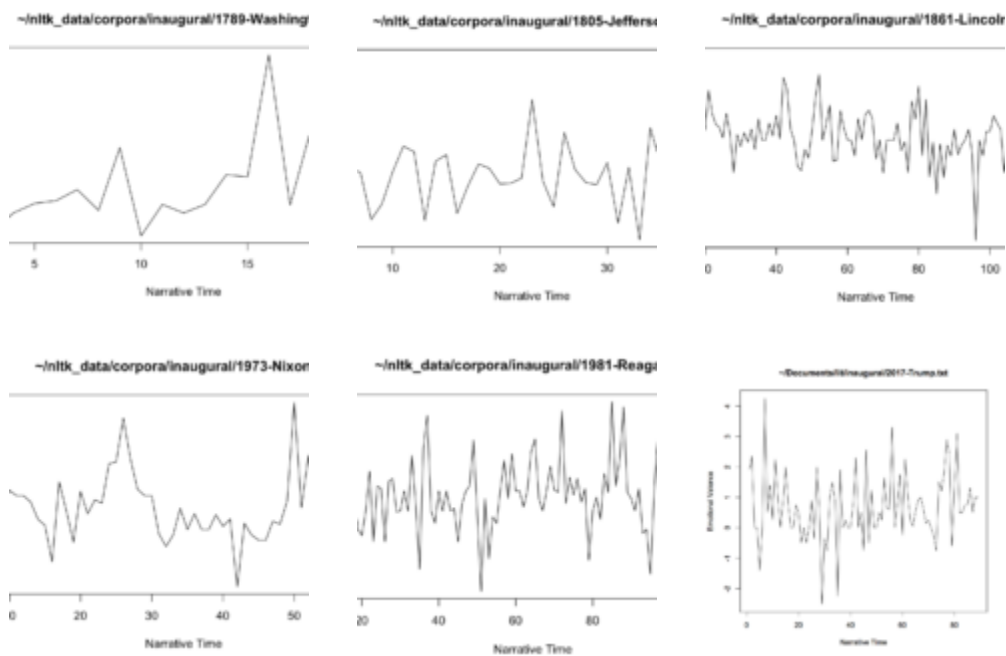
path <- '~/nltk_data/corpora/inaugural/'

for ( f in dir(path) ) {
  valence( paste( path, f, sep = '' ) )
  readline(prompt = "Continue? ")
}

```

}

Here are a few of the “emotional sentiment over time” plots:



Most all

addresses end on an upward trending note.

FILED UNDER: DATA, SOFTWARE

TAGGED WITH: PERL, R, TEXT

Epistemologist-at-large

^ Top