

Sentiment Analysis

Kevin Chan, Brianna Kincaid, Lauren Vanvalkenburg

April 6, 2018

Comparing Trump's State of the Union Address with Nixon's.

```
library(knitr)
library(tidytext)

## Warning: package 'tidytext' was built under R version 3.4.4

library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(stringr)
library(readtext)
library(wordcloud)

## Warning: package 'wordcloud' was built under R version 3.4.4

## Loading required package: RColorBrewer

library(ggplot2)
library(reshape2)

get_sentiments("bing")

## # A tibble: 6,788 x 2
##   word      sentiment
##   <chr>    <chr>
## 1 2-faced    negative
## 2 2-faces    negative
## 3 a+        positive
## 4 abnormal   negative
## 5 abolish    negative
## 6 abominable negative
## 7 abominably negative
## 8 abominate   negative
## 9 abomination negative
## 10 abort      negative
## # ... with 6,778 more rows

nrc_joy <- get_sentiments("nrc") %>%
  filter(sentiment == "joy")
```

```
nrc_anger <- get_sentiments("nrc") %>%
  filter(sentiment == "anger")

nrc_anticipation <- get_sentiments("nrc") %>%
  filter(sentiment == "anticipation")
```

Trump

Now we look into Trump's State of the Union Address.

```
trump <- readtext("trump first state of the union.txt")
trump <- trump %>% unnest_tokens(word, text)

trump_anger <- trump %>%
  inner_join(nrc_anger) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
trump_anger
```

```
## # A tibble: 64 x 2
##   word      n
##   <chr>    <int>
## 1 deserve      5
## 2 gang         4
## 3 terrible     4
## 4 defense      3
## 5 dictatorship 3
## 6 fighting     3
## 7 aftermath    2
## 8 aggression   2
## 9 bad          2
## 10 broken      2
## # ... with 54 more rows
```

Here, we see some words that are clearly associated with anger, such as aggression and disgrace, but we also see some words that might have been taken out of context and may not actually be used in an angry context like deserve and powerful.

```
trump_anticipation <- trump %>%
  inner_join(nrc_anticipation) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
trump_anticipation
```

```
## # A tibble: 82 x 2
##   word      n
##   <chr>    <int>
## 1 time      9
## 2 finally   8
## 3 proud     6
```

```
## 4 deserve      5
## 5 long         5
## 6 god          4
## 7 plan         4
## 8 young        4
## 9 child        3
## 10 defense     3
## # ... with 72 more rows
```

There are not only more anticipation words, but there are also some anticipation words used more frequently, such as time and plan. This implies that Trump's agenda is filled with

```
trump_joy <- trump %>%
  inner_join(nrc_joy) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
```

```
trump_joy
```

```
## # A tibble: 76 x 2
##   word      n
##   <chr>    <int>
## 1 finally      8
## 2 love         7
## 3 safe         7
## 4 freedom      6
## 5 proud        6
## 6 god          4
## 7 young        4
## 8 achieve      3
## 9 child        3
## 10 confidence  3
## # ... with 66 more rows
```

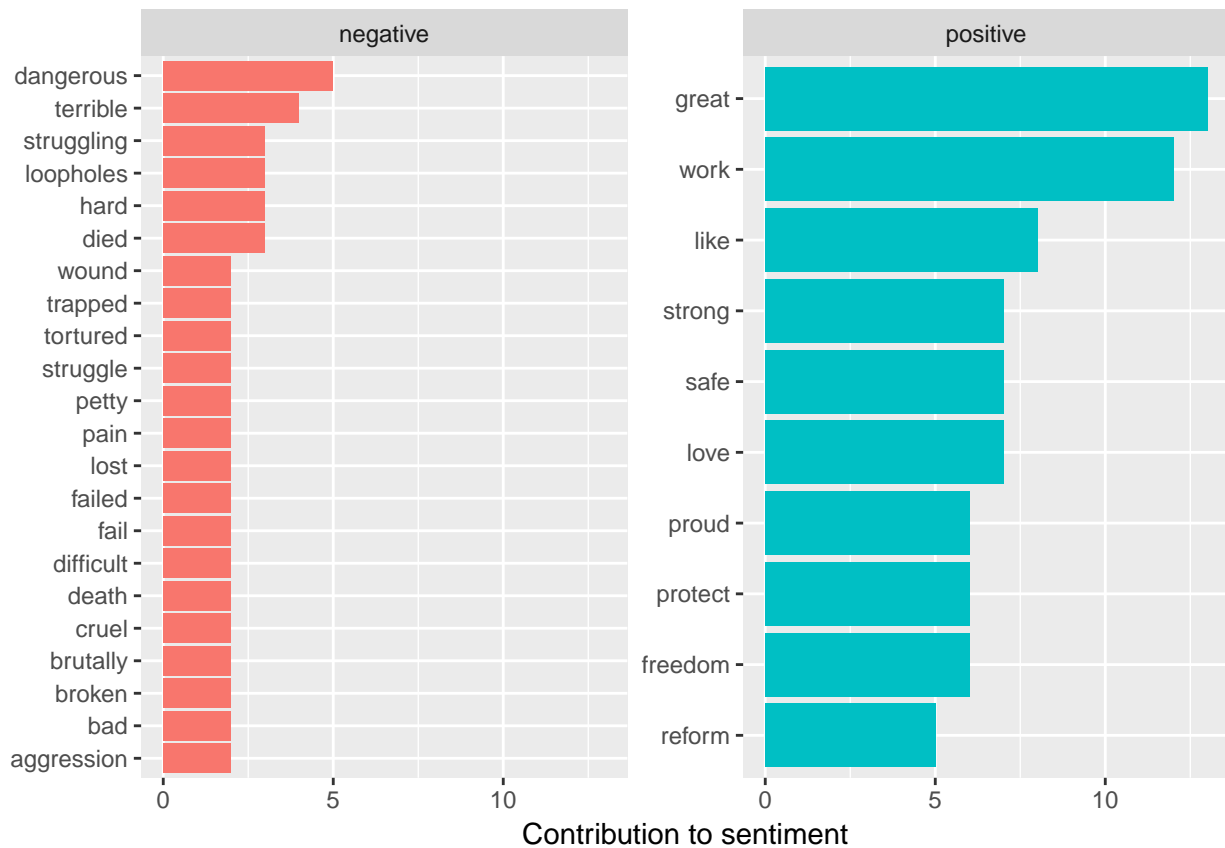
Despite what the media has said about Trump so far during his presidency, his state of the union address was has more words associated with joy than words associated with anger. However, there are plenty of words here that may have been taken out of context, such as finally, create, and yearning.

```
trump_bing <- trump %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()
```

```
## Joining, by = "word"
```

```
trump_bing %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y = "Contribution to sentiment",
       x = NULL) +
  coord_flip()
```

```
## Selecting by n
```

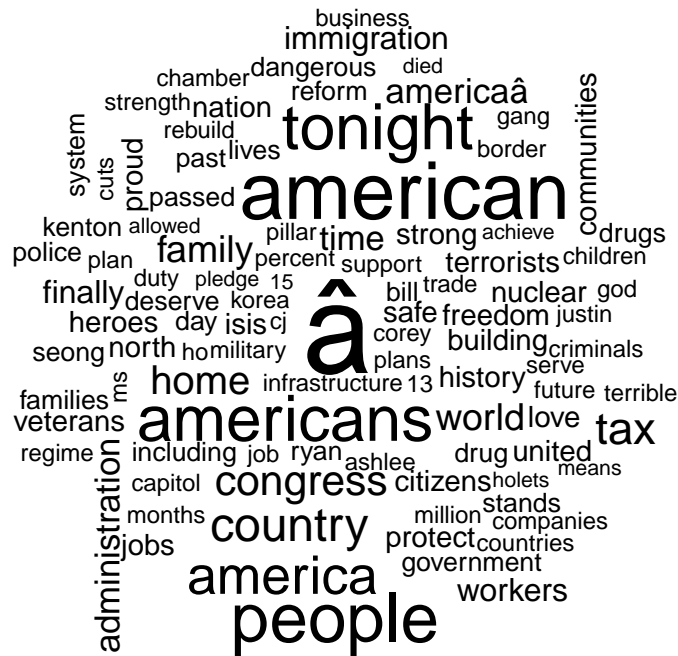


Here, we see that there are many more negative words than positive, but the most frequently used positive word, great, is used more than the most frequently used negative word, which is dangerous. This implies that Trump sees the current condition of the US in a bad light and is not very optimistic about the US' potential future. Like in the previous sentiment analyses, some words may have been taken out of context like reform.

Wordcloud for Trump's State of the Union Address

```
trump %>%
  anti_join(stop_words) %>%
  count(word) %>%
  with(wordcloud(word, n, max.words = 100))
```

```
## Joining, by = "word"
```



Wordcloud comparing positive and negative words in Trump's State of the Union Address

Nixon

```
trump %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  acast(word ~ sentiment, value.var = "n", fill = 0) %>%
  comparison.cloud(colors = c("gray20", "gray80"),
                   max.words = 100)

## Joining, by = "word"

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): loyalty could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): powerful could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): progress could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): promise could not be fit on page. It will not be plotted.
```

```

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): prosperity could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): supported could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): wonderful could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): menace could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): mistakes could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): needy could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): pains could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): randomly could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): reckless could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): savage could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): slashing could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): sorrow could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): threaten could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): tragedy could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): tragically could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unaccustomed could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): undermine could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): undermined could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unfortunately could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unlawful could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): weakness could not be fit on page. It will not be plotted.

```

```
## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): weep could not be fit on page. It will not be plotted.
```



Nixon's State of the Union Address

```
nixon <- readtext("nixon first state of the union.txt")
nixon <- nixon %>% unnest_tokens(word, text)
```

```
nixon_anger <- nixon %>%
  inner_join(nrc_anger) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
```

```
nixon_anger
```

```
## # A tibble: 28 x 2
##   word      n
##   <chr>    <int>
## 1 crime      7
## 2 blame      4
## 3 bad        3
## 4 challenge  3
## 5 defense    3
## 6 money      3
## 7 fight      2
```

```
## 8 involvement      2
## 9 moral            2
## 10 argue           1
## # ... with 18 more rows
```

Similar to Trump's state of the union address, there are not too many of Nixon's words associated with anger. Moreover, Nixon had way less words associated with anger than Trump in his State of the Union, which infers that Nixon had a more favorable outlook and attitude toward the US as well as leading the US to become a better nation.

```
nixon_anticipation <- nixon %>%
  inner_join(nrc_anticipation) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
```

```
nixon_anticipation
```

```
## # A tibble: 58 x 2
##   word      n
##   <chr>    <int>
## 1 time      17
## 2 peace     16
## 3 progress   6
## 4 defense    3
## 5 hope       3
## 6 money       3
## 7 opportunity 3
## 8 plan       3
## 9 vision     3
## 10 abundance  2
## # ... with 48 more rows
```

Like in Trump's address, Nixon's address has the word 'time' as the most frequently used word related to anticipation, but Nixon uses it even more than Trump, seventeen times to be exact, compared to nine times by Trump.

```
nixon_joy <- nixon %>%
  inner_join(nrc_joy) %>%
  count(word, sort = TRUE)
```

```
## Joining, by = "word"
```

```
nixon_joy
```

```
## # A tibble: 57 x 2
##   word      n
##   <chr>    <int>
## 1 peace     16
## 2 clean      9
## 3 progress   6
## 4 freedom    4
## 5 fulfill    4
## 6 hope       3
## 7 money       3
## 8 abundance  2
## 9 achieve    2
## 10 income    2
```



```
## # ... with 47 more rows
```

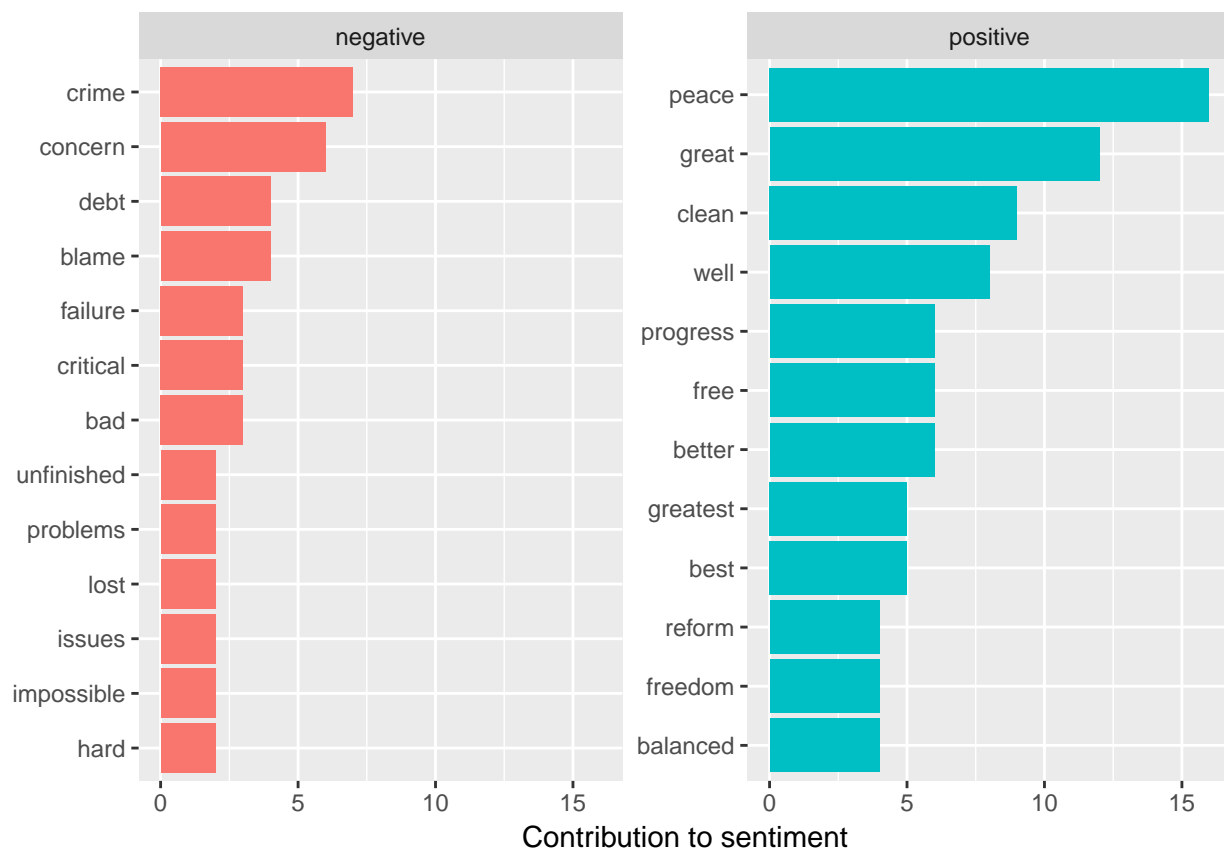
Although Trump may have used more words relating to joy, Nixon's most frequent joy word, namely peace, is used sixteen times in his speech, as opposed to Trump's most frequent word, finally, which is used eight times and may not have necessarily been used in such a positive context. Moreover, there are less words associated with joy that are used only once in Nixon's speech than in Trump's speech, further implying that Nixon approached leading the US in a more optimistic way.

```
nixon_bing <- nixon %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()
```

```
## Joining, by = "word"
```

```
nixon_bing %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  ungroup() %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y = "Contribution to sentiment",
       x = NULL) +
  coord_flip()
```

```
## Selecting by n
```



Unlike in Trump's State of the Union Address, where there were clearly more negative words used, Nixon's State of the Union Address has a relatively equal balance between positive and negative words. Therefore, we can infer that Nixon's agenda approached resolving the nation's issues with a more optimistic outlook, as opposed to Trump, who seems to really focus on the negative side of the nation's condition and has a little more pessimistic view on resolving the problems.

Wordcloud for Nixon's State of the Union Address

```
nixon %>%
  anti_join(stop_words) %>%
  count(word) %>%
  with(wordcloud(word, n, max.words = 100))

## Joining, by = "word"

## Warning in wordcloud(word, n, max.words = 100): people could not be fit on
## page. It will not be plotted.

## Warning in wordcloud(word, n, max.words = 100): responsibility could not be
## fit on page. It will not be plotted.

## Warning in wordcloud(word, n, max.words = 100): world could not be fit on
## page. It will not be plotted.

## Warning in wordcloud(word, n, max.words = 100): peace could not be fit on
## page. It will not be plotted.

## Warning in wordcloud(word, n, max.words = 100): congress could not be fit
## on page. It will not be plotted.
```



Wordcloud comparing positive and negative words in Nixon's State of the Union Address

```
nixon %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort = TRUE) %>%
  acast(word ~ sentiment, value.var = "n", fill = 0) %>%
  comparison.cloud(colors = c("gray20", "gray80"),
                   max.words = 100)

## Joining, by = "word"

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): overwhelming could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): overwhelmingly could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): polluter could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): rejecting could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): relentlessly could not be fit on page. It will not be plotted.
```

```

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): scarce could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): spite could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): stark could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): suffered could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): surrender could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): tragic could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unable could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unbearable could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): underestimate could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): unrest could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): urgent could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): violent could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): warning could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): waste could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): weak could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): weaken could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): weakness could not be fit on page. It will not be plotted.

## Warning in comparison.cloud(., colors = c("gray20", "gray80"), max.words =
## 100): worst could not be fit on page. It will not be plotted.

```

