

# Obama's speech analysis

*Evelyn Sanchez*

5/28/2018

## Text analysis of two speeches

In the following exercise, I analyzed two speeches of the former President Obama given in remarkably different contexts in order to identify the change in rhetoric. The speeches that I analyzed are: Obama's Presidential Farewell Address, delivered in Chicago on January 10th, 2017 and Address on Drones and Terrorism at the National Defense University, delivered in Washington, D.C. on May 23rd, 2013, obtained as PDF documents at the American Rhetoric (<https://www.americanrhetoric.com/>) website.

To present this information, I first used a wordcloud visualization which identifies the most used words during the speech. This visualization is easy to understand for any kind of public, since the larger the size of the word represent the words that appeared in the speech more frequently. Additionally, this visualization includes colors association with the frequency, so even when the public cannot distinguish the size the color difference can help to understand the difference in frequency.

To make the difference in the use of words clearer, I later select the most popular words, compare their relation in a hierarchy cluster, that help to identify the combination of the words and their frequency.

# Obama's Presidential Farewell Address

## Obama's Presidential Farewell Address

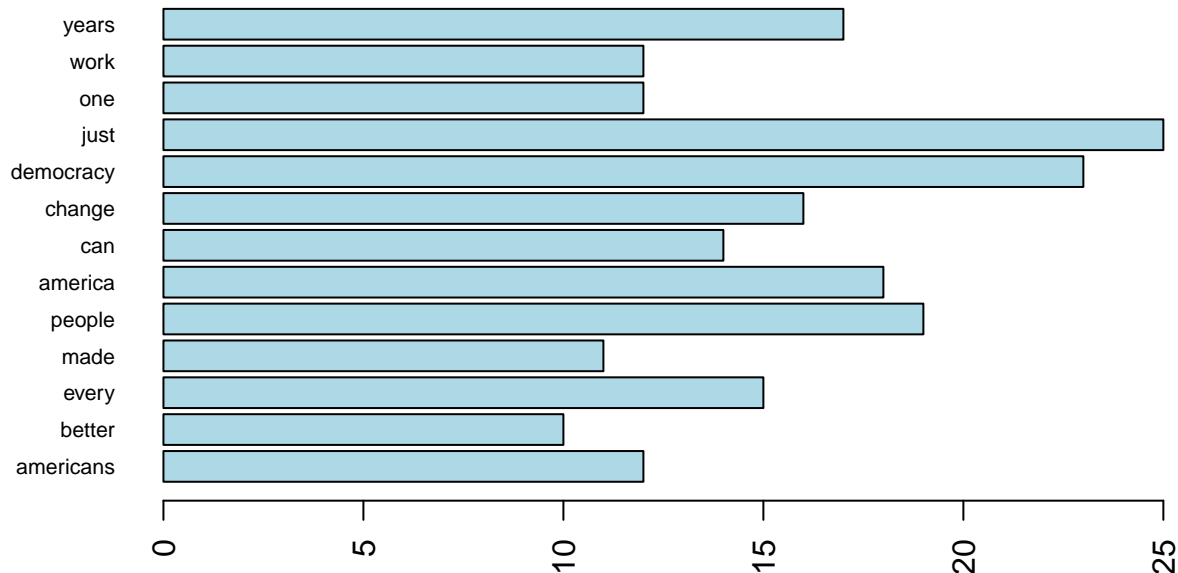


In the wordcloud we are able to see the 70 most popular word of the speech. On it, we can observe that the words used the most are, just and 's, following by democracy and then people, years and America. We can appreciate the tone of the speech is trying to give a review of what had been done in his period by looking at the smaller words, such as change, can, better, good, work, made, economic. In the same way, we encounter some other emotional words, such as faith, believe and spirit.

```
## <<TermDocumentMatrix (terms: 1299, documents: 13)>>
## Non-/sparse entries: 1972/14915
## Sparsity          : 88%
## Maximal term length: 16
## Weighting          : term frequency (tf)

## [1] "americans" "better"    "every"    "made"     "people"
## [6] "'s"         "'ve"       "america"  "can"      "change"
## [11] "democracy" "just"      "one"      "work"     "years"
```

### Most used words



Number of repetitions  
Obama's Presidential Farewell Address

Even when with the wordcloud it is possible to appreciate the most popular words, for a deeper analysis it is helpful to reduce the words to a smaller sample of word that are repeated at least 10 times. This filter give a sub sample of 13 words ('s and 've were omitted). The bar chart above helps to illustrate the difference in use of this top speech words, in which we can see that just reach 25 repetitions, followed by democracy and years.

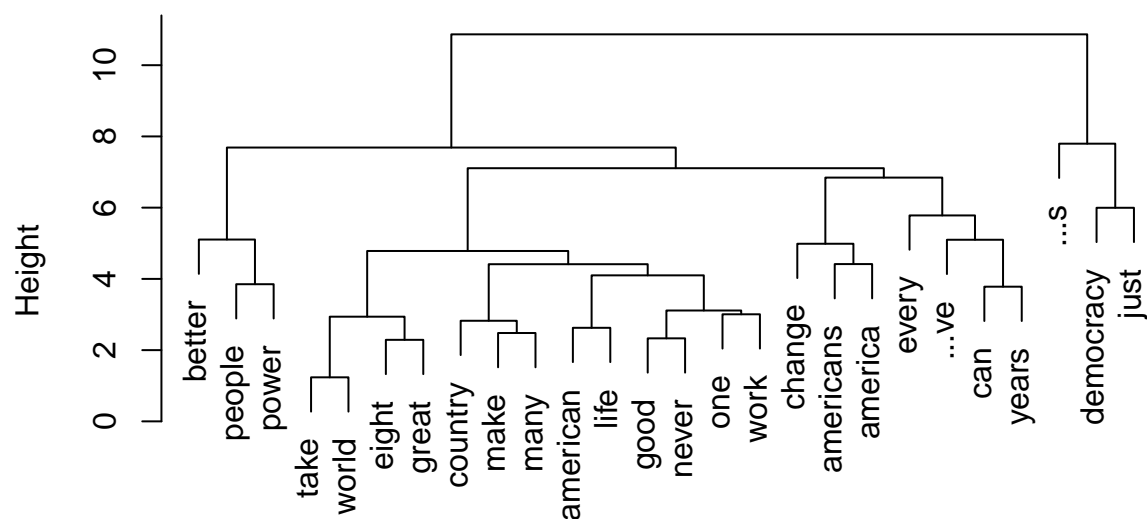
```
## $democracy
## democratic eventually low protect sense
## 0.88 0.88 0.88 0.88 0.88
## terrorism accounts affiliation anyone argued
## 0.88 0.84 0.84 0.84 0.84
## argument badly basic begin beginning
## 0.84 0.84 0.84 0.84 0.84
```

##	century	compromised	cost	costs	covers
##	0.84	0.84	0.84	0.84	0.84
##	credit	decency	demographic	demonstrably	determine
##	0.84	0.84	0.84	0.84	0.84
##	educate	everyone	expected	expense	fairer
##	0.84	0.84	0.84	0.84	0.84
##	fall	falling	focus	forces	growing
##	0.84	0.84	0.84	0.84	0.84
##	haven't	improvements	inequality	knew	ladders
##	0.84	0.84	0.84	0.84	0.84
##	less	lower	market	moments	near
##	0.84	0.84	0.84	0.84	0.84
##	news	particular	paying	plan	potential
##	0.84	0.84	0.84	0.84	0.84
##	poverty	prosper	publicly	quarreled	rate
##	0.84	0.84	0.84	0.84	0.84
##	realized	records	reflects	restore	retirement
##	0.84	0.84	0.84	0.84	0.84
##	rising	score	security	shatters	shrinking
##	0.84	0.84	0.84	0.84	0.84
##	slowest	solidarity	specter	stark	state
##	0.84	0.84	0.84	0.84	0.84
##	stock	tenyear	tested	testing	throughout
##	0.84	0.84	0.84	0.84	0.84
##	unemployment	uniformity	uninsured	want	words
##	0.84	0.84	0.84	0.84	0.84
##	works				
##	0.84				

### Association with “democracy”

Excluding “just”, democracy is the most popular word, so to see how this word relates with other words in the speech, I check the number of words that are correlated by 80 %. We found an extensive list of words, that even when they are not used with high frequency are highly correlated, which can indicate a harmony in the tone of the speech in relation with the democracy word

## Hierarchy cluster of words



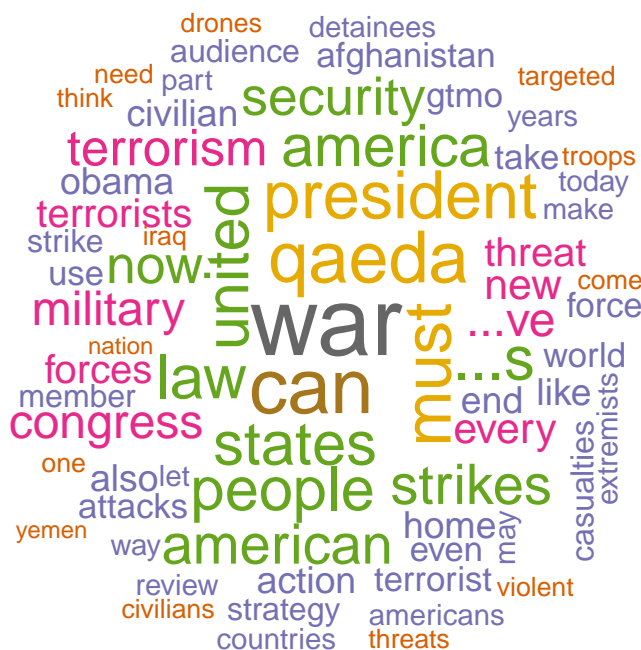
distMatrix\_f

## Obama's Presidential Farewell Address

Finally, coming back to the sub sample of the most used words, I analyze the relation between the most popular 15 words ('s and 've included) through an hierarchy cluster of words, that cluster words in groups depending on their “distance”, which can give us an idea how these words were used together in a sentence or in paragraphs.

## Address on Drones and Terrorism at the National Defense University

## Obama's Address on Drones and Terrorism at the National Defense University



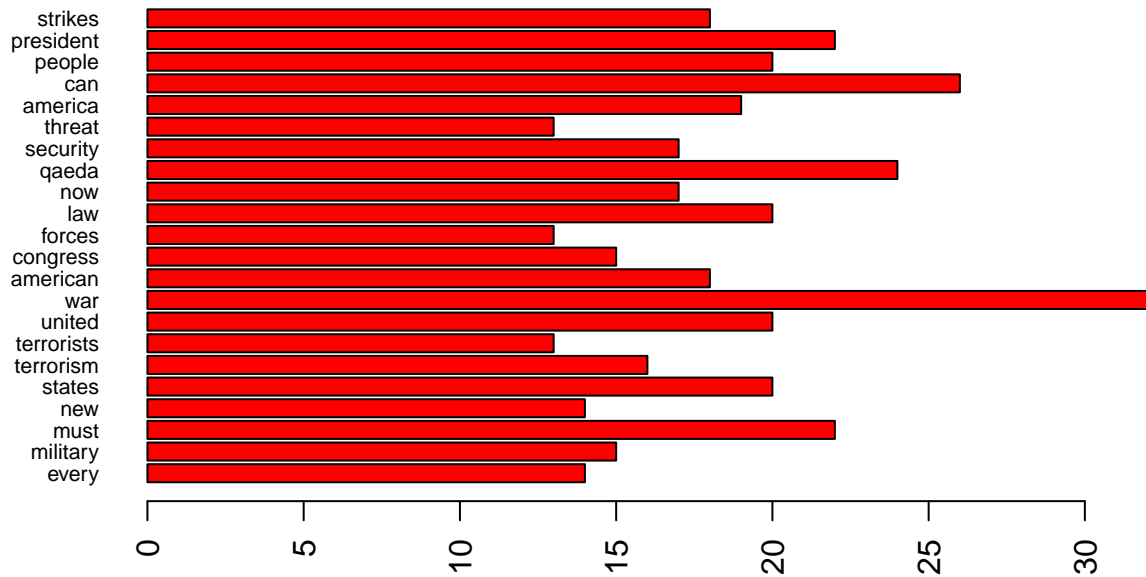
For this second speech I found a very different set of words, with “war” as the most popular followed by “can”, and later for a long list of words, such as law, states, America, united. This might indicate that the speech covers a broader set of primary topics compared with the other speech.

Later, I analyze the most repeated words. This time I used a threshold of over 13 repetitions, since the number of words used more than 10 times was too large. The bar chart below shows how the word “war” is used over 30 times, followed with can (27), Qaeda (24) and must (23). It is possible to observe a higher frequency of repetition than the other speech, as well the tone of the speech is more encouraging. One similarities between speeches is that the words “can”, “people”, “Americans” are one of most used words.

```
## <<TermDocumentMatrix (terms: 1678, documents: 16)>>
## Non-/sparse entries: 2927/23921
## Sparsity           : 89%
## Maximal term length: 16
## Weighting          : term frequency (tf)

## [1] "every"      "military"   "must"       "new"        "states"
## [6] "terrorism"  "terrorists" "united"     "war"        "american"
## [11] "congress"   "forces"     "law"        "now"        "qaeda"
## [16] "security"   "threat"     "america"    "can"        "people"
## [21] "president"  "'s"         "'ve"        "strikes"
```

## Most used words



## Number of repetitions

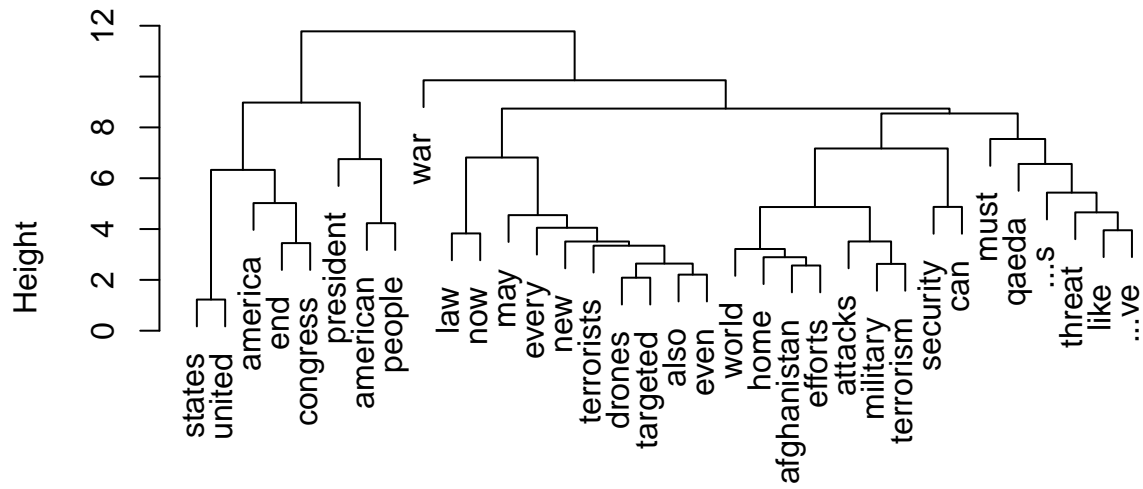
### Obama's Address on Drones and Terrorism at the National Defense University

##	\$democracy				
##	must	group	uniform	absence	education
##	0.71	0.65	0.65	0.65	0.65
##	engaging	perpetual	rebuke	success	ways
##	0.65	0.65	0.65	0.65	0.65
##	working	organizations	challenges	forward	protect
##	0.65	0.65	0.62	0.62	0.62
##	aumf	keep			
##	0.62	0.62			

### Association with “democracy”

I test the correlation of the words in this speech with the word democracy, to have a comparative between both speeches. Here I reduce the correlation limit to 60% since there are not words associated at the 80% level. We can see that the set of words related with democracy are very different, and must is the most correlated one.

## Hierarchy cluster of words



distMatrix\_t

## Obama's Address on Drones and Terrorism at the National Defense University

Finally, I tested the association of the most popular words, in order to see how they are jointly used in the speech, that can give a clearer way of how the speech used this words. In the first cluster, we can See how states and united are in the same clustered words, since they were used to make reference to the name of the country.