MEMORANDUM
To: All BA members
From: Yuchen Liu
Date: Nov 16, 2016
Subject: Trump's Linguistic Effectiveness Report

Introduction

In this report, we are going to select four speeches to analyze Trump's speech style. Our object is to find the insights on Trump's linguistic effectiveness.

Methods

1. Create a Corpus for the Speeches

a. Create Corpus

```
newscorpus2 <- corpus(DT$Full_Text,
docnames=DT$Speech_Topic, docvar=NULL)
#clean corpus
newcorpus2<- toLower(newscorpus2, keepAcronyms = FALSE)
cleancorpus2 <- tokenize(newscorpus2,
                         removeNumbers=TRUE,
                         removePunct = TRUE,
                         removeSeparators=TRUE,
                         removeTwitter=FALSE,
                         verbose=TRUE)
```

b. Explore the Corpus:

```
> summary(newscorpus2)  #summary of corpus
Corpus consisting of 4 documents.

    Text Types Tokens Sentences
 Speech1   536   1910       170
 Speech2  1469   6426       453
 Speech3  1129   8613       639
 Speech4   959   2783       138

Source:  /Users/apple/* on x86_64 by apple
Created: Mon Nov 14 22:01:04 2016
```

2. Complete a frequency analysis of word usage

a. Generate (DFM):

to create a custom dictionary list of stop words:

```
swlist2 = c("thank", "much", "can","will","just","trump")

dfm2<- dfm(cleancorpus2,
           toLower = TRUE,
           ignoredFeatures = c(swlist2, stopwords("english")),
           verbose=TRUE,
           stem=TRUE)
```

Reviewing top features:

```
view(dfm2)
```

	everybody	sorry	keep	waiting	complicated	business	received	call	secretary	clinton	congratulated	us	victory	family	hard-fought	campaign	mean	fought	hard	hillary
Speech1	1	1	1	1	2	2	1	4	1	1	2	7	2	5	1	5	2	1	3	1
Speech2	3	0	4	1	1	1	0	4	2	14	0	6	0	0	0	0	3	0	5	15
Speech3	10	1	2	0	0	1	0	5	1	1	0	27	1	7	0	1	1	0	1	0
Speech4	0	0	0	0	0	0	1	1	3	13	0	5	0	0	0	2	0	0	0	11

```
topfeatures(dfm2, 100)        # displays 100 features
```

```
> topfeatures(dfm2, 100)     # displays 100 features
      people         going         great          know            re       country           now           one 
         121           102            73            72            72            70            54            50 
        want           don           say            us          said           get   immigration         right 
          50            50            46            45            42            40            36            35 
        like          back       clinton          take          even        states          make        number 
          33            32            29            29            29            29            29            29 
        need          many        really           new            ll       hillary         world       illegal 
          29            28            28            28            28            27            27            27 
        jobs          time        united          also       billion      military      american           big 
          27            26            26            26            25            25            24            24 
       years         china     president       believe           got         never          good          work 
          24            24            23            23            22            22            22            21 
      happen          tell         think          love           way        mexico        border         money 
          21            21            21            20            20            20            20            20 
    audience          come administration    immigrants       defense       america           law        member 
          20            19            19            19            19            18            18            18 
     million            ve          plan            go           ever        office         bring         build 
          18            18            17            17            17            17            17            17 
      things         first     countries        system         obama           day           put       nothing 
          16            16            16            16            16            16            15            15 
  thousands          iraq     everybody          call           job          look           lot         place 
          15            15            14            14            14            14            14            14 
     needs       criminal          nice       support        folks           end           two         state 
          14            14            14            13            13            13            13            13 
     today     politicians       leaders        family 
          13            13            13            12 
```

Analysis of DFM:

First, because the texts are speeches, so I choose to create new dictionary to screen out some common words in speeches, such as “can”, “thank”, “will”, “much” and “just”. In the DFM, we can see that Mr. Trump’s most frequent words are “people”, “going”, “great”, “know”, “re-”,” country”, “now”, “one”, “want”. These words are all simple monosyllables. In addition, besides the top 10 frequent words, we can also see some common topics, such as “immigration”, “military”, “jobs”, “Mexico”, “China”, “criminal”. Furthermore, he also mentioned his competitors Hillary a lot in his speeches.

b. Find Root words:

```
dfm.stem2 <- dfm(cleancorpus2, toLower = TRUE,
                 ignoredFeatures = c(swlist2, stopwords("english")),
                 verbose=TRUE,
                 stem=TRUE)
topfeatures.stem2 <-topfeatures(dfm.stem2, n=50) #fifty common words
topfeatures.stem2
```

```
> topfeatures.stem2
   peopl       go  countri     know    great       re     want   immigr      now      say      one      don      get 
     121      119       87       75       74       72       59       58       54       54       54       50       48 
    need       us     said    state      job     take     make    right     work     like     year  clinton american 
      46       45       42       42       41       38       38       35       34       34       34       32       32 
    back    illeg   number     come     even    build   happen      law     mani   realli      new       ll  hillari 
      32       31       31       30       30       30       29       29       28       28       28       28       27 
    time    world     unit    thing     also     tell   border    think  billion  militari     call 
      27       27       26       26       26       26       26       26       25       25       24 
```

Analysis of Root Words:

The root words can be used to justify the conclusion showed in the "Analysis of DFM", which indicate the most frequent root words and the most likely topics in Mr. Trump's speeches.

c. Analyzing DFM with bigrams:

```
cleancorpus_bi <- tokenize(newscorpus2,
                           removeNumbers=TRUE,
                           removePunct = TRUE,
                           removeSeparators=TRUE,
                           removeTwitter=FALSE,
                           ngrams=2, verbose=TRUE)

dfm.bigram2 <- dfm(cleancorpus_bi, toLower = TRUE,
                   ignoredFeatures = c(swlist2, stopwords("english")),
                   verbose=TRUE,
                   stem=FALSE)
topfeatures.bigram2 <-topfeatures(dfm.bigram2, n=50)
topfeatures.bigram2
```

```
> topfeatures.bigram2
      united_states     hillary_clinton      audience_member   illegal_immigrants         middle_east
                 25                  24                  17                   13                  11
          re_going     law_enforcement      criminal_aliens     president_obama          bring_back
                 11                   9                   9                    7                   7
           ll_say           net_worth      missile_defense          take_care           right_now
                  6                   6                   6                    5                   5
        great_job   immigration_system    special_interests        open_borders          number_one
                  5                   5                   5                    5                   5
      right_people    islamic_terrorism             web_site        make_america          member_yes
                  5                   5                   5                    5                   5
      saudi_arabia         nice_person      foreign_policy         great_people     american_people
                  5                   5                   5                    4                   4
         many_many     sanctuary_cities           air_force      billion_dollars           one_thing
                  4                   4                   4                    4                   4
           day_one           even_know        common_sense        need_somebody           re_gonna
                  4                   4                   4                    4                   4
       common_core     radical_islamic tremendous_potential        inner_cities         truly_great
                  4                   4                   3                    3                   3
   talented_people            new_york         four_years         white_house illegal_immigration
                  3                   3                   3                    3                   3
```

Analysis of bigrams:

The bigrams provide more detailed information. Mr. Trump's speeches focus on making United States a great country, his most frequent topic would be "illegal immigrants" and "jobs".

3. Complete a sentiment analysis

```
mydict2<- dictionary(list(negative = c("detriment*", "bad*", "awful*", "terrib*",
"horribl*","stupid","weak", "loser","tough","dangerous","zeor","hate","worse"),

postive = c("fantastic","classy","good", "great", "super*", "excellent",
"yay","win","smart","amazing","terrific")))      ###create your own dictionary
```

```
dfm.sentiment2 <- dfm(cleancorpus2, dictionary = mydict2)
topfeatures(dfm.sentiment2)
View(dfm.sentiment2)
```

	negative	postive
Speech1	6	34
Speech2	20	28
Speech3	24	53
Speech4	2	3

Sentiment analysis:

	negative	positive
Speech1	15%	85%
Speech2	41.67%	58.33%
Speech3	31.17%	68.83%
Speech4	40%	60%

The outcomes indicate that speech 1 (victory speech) used more positive words; speech 2 used more positive words but the difference between negative words and positive words is small; speech 3 use more positive words; speech 4 is more neutral because there are few emotional words.

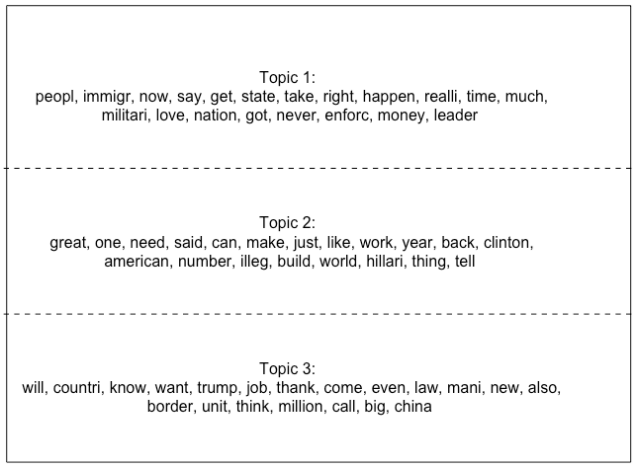
4. Common topics in the corpus

```
prevfit2 <-stm(docs , vocab ,
               K=3,
               verbose=TRUE,
               data=meta,
               max.em.its=10)

topics <-labelTopics(prevfit2 , topics=c(1:3))
topics      #shows topics with highest probability words

#explore the topics in context.    Provides an example of the text
help("findThoughts")
findThoughts(prevfit2, texts = DT,    topics = 1,    n = 1)

help("plot.STM")
plot.STM(prevfit2, type="summary")
plot.STM(prevfit2, type="labels", topics=c(1,2,3))
plot.STM(prevfit2, type="perspectives", topics = c(1,2))
```



Common topic analysis:

According to the two charts shown above, the common topics are:

1. Immigration Issue
2. How to make America to be great
3. Trump will bring jobs to the country

Context analysis (the screen shots below are just two sample words):

kwic(cleancorpus2, "believe", 5)

kwic(cleancorpus2, "great", window = 3)

```
                                  contextPre keyword                      contextPost
[Speech1, 1163]          A very special person who [ believe ] me I read reports that
[Speech2, 1791]               are we doing Hard to [ believe ] Hard to believe Now that
[Speech2, 1794]           Hard to believe Hard to [ believe ] Now that you've heard about
[Speech2, 2089]             work with us I really [ believe ] it Mexico will work with
[Speech2, 2098]          work with us I absolutely [ believe ] it And especially after meeting
[Speech2, 2112] wonderful president today I really [ believe ] they want to solve this
[Speech2, 2258]              so great It's hard to [ believe ] people don't even talk about
[Speech2, 2847]              And they will go face [ believe ] me They're going to go
[Speech2, 4080]          the right people doing it [ believe ] me very very few will
[Speech2, 4362]             take them back Hard to [ believe ] with the power we have
[Speech2, 4370]              power we have Hard to [ believe ] We're like the big bully
[Speech2, 4735]         If people around the world [ believe ] they can just come on
 [Speech3, 242]            They are not our friend [ believe ] me But they re killing
 [Speech3, 457]             hotel in Syria Can you [ believe ] this They built a hotel
 [Speech3, 794]              from to percent Don t [ believe ] the Don t believe it
 [Speech3, 798]                t believe the Don t [ believe ] it That s right A
[Speech3, 1138]            They will not bring us [ believe ] me to the promised land
[Speech3, 1274]            level that you wouldn t [ believe ] It makes it impossible for
[Speech3, 3190]          my opinion the new China [ believe ] it or not in terms
[Speech3, 3460]          them one for each country [ Believe ] me folks We will do
[Speech3, 4236]            there except for us And [ believe ] me you look at the
[Speech3, 5671]                   I don t have to [ believe ] it or not I m
[Speech3, 5894]        builds walls better than me [ believe ] me and I ll build
```

	contextPre	keyword	contextPost
[Speech1, 196]	and unify our [	great]	country As I've
[Speech1, 214]	an incredible and [	great]	movement made up
[Speech1, 421]	care of our [	great]	veterans who have
[Speech1, 505]	We have a [	great]	economic plan We
[Speech1, 546]	We will have [	great]	relationships We expect
[Speech1, 552]	expect to have [	great]	great relationships No
[Speech1, 553]	to have great [	great]	relationships No dream
[Speech1, 564]	challenge is too [	great]	Nothing we want
[Speech1, 709]	me right now [	Great]	people I've learned
[Speech1, 724]	every regard Truly [	great]	parents I also
[Speech1, 757]	brother Robert my [	great]	friend Where is
[Speech1, 779]	that's okay They're [	great]	And also my
[Speech1, 786]	late brother Fred [	great]	guy Fantastic guy
[Speech1, 796]	was very lucky [	Great]	brothers sisters great
[Speech1, 799]	Great brothers sisters [	great]	unbelievable parents To
[Speech1, 876]	much What a [	great]	group You've all
[Speech1, 1034]	is Jeff A [	great]	man Another great
[Speech1, 1037]	great man Another [	great]	man very tough
[Speech1, 1432]	to do a [	great]	job and I
[Speech1, 1449]	will do a [	great]	job We will
[Speech1, 1455]	will do a [	great]	job I look
[Speech2, 209]	also discussed the [	great]	contributions of Mexican-American
[Speech2, 679]	many of the [	great]	parents who lost
[Speech2, 998]	Force veteran a [	great]	woman according to
[Speech2, 1536]	our country with [	great]	dignity So important
[Speech2, 1911]	Immigration offices very [	great]	people Among the
[Speech2, 1976]	will build a [	great]	wall along the
[Speech2, 2006]	it And they're [	great]	people and great
[Speech2, 2009]	great people and [	great]	leaders but they're

The outcomes indicate that Mr. Trump applied catchphrases, which are the speech styles that salesmen use, such as “believe me”, “many people are saying” and “great”.

Summary

Mr. Trump usually applies simple monosyllables in his speeches; he basically just uses casual speech in a public setting. In addition, Mr. Trump’s speeches are filled with sentiments when the speech topics are relating immigration and jobs, which indicates that he is adept in connecting audiences on an emotional level. He often uses catchphrases, which are actually versions of speech mechanisms that salesmen use. Furthermore, the topics he often targets in his speeches are “immigration” and “jobs”.

I will be glad to discuss these conclusions and follow through on how to compute statistical significance between Trump’s and other presidents’ speech styles.

Thank you.