All Women Short lists Methodology

# All Women Short lists Methodology

* [TO DO:](#to-do)
* [Methodology](#methodology)
  + [Descriptive Statistics](#descriptive-statistics)
  + [Women vs Men](#women-vs-men)
  + [Short lists vs Non-Short lists](#short-lists-vs-non-short-lists)
  + [Conservatives vs Labour](#conservatives-vs-labour)
  + [All MPs Gender Differences](#all-mps-gender-differences)
  + [POS Analysis](#pos-analysis)
  + [Tokenising / Keyness](#tokenising-keyness)
    - [Men vs Women](#men-vs-women)
    - [Short lists vs Non-Short lists](#short-lists-vs-non-short-lists-1)
    - [Labour vs Conservative](#labour-vs-conservative)
  + [Bigrams](#bigrams)
  + [Naive Bayes classification](#naive-bayes-classification)
  + [Topic Models](#topic-models)
    - [Short lists vs Non-Short lists](#short-lists-vs-non-short-lists-2)
  + [Discussion](#discussion)
* [Appendix](#appendix)
  + [Full topic model summary](#full-topic-model-summary)
* [References](#references)

## TO DO:

1. Male MP topic models matching

# Methodology

Previous research on gender differences in political speech patterns has focused on differences between male and female politicians (Yu, [2014](#ref-yu2014)) or on variations in Hilary Clinton’s speech patterns (Bligh, Merolla, Schroedel, & Gonzalez, [2010](#ref-bligh2010); Jones, [2016](#ref-jones2016)). This paper focuses on differences in speech patterns between female Labour MPs nominated through All Women Shortlists (AWS) and female Labour MPs nominated through open short lists. We examined differences in speaking styles using the Linguistic Inquiry and Word Count 2015 (LIWC) dictionary (Pennebaker, Boyd, Jordan, & Blackburn, [2015](#ref-pennebaker2015)) and the spaCy (Honnibal & Montani, [2017](#ref-honnibal2017)) Parts-of-Speech (POS) tagger. We examined differences in the topics discussed by AWS and non-AWS MPs, using \({\chi}^2\) tests for individual words and for bigrams. We experimented with unsupervised Latent Dirichlet Allocation (Blei, Ng, & Jordan, [2003](#ref-blei2003)), and trained a Naive Bayes classifier to distinguish AWS and non-AWS speeches.

To account for the possible effects of age, parliamentary experience and cohort, and in order to compare women selected through all women short lists to women who were not (but who theoretically had the opportunity to contest all-women short lists), our analysis is been restricted only to Labour MPs first elected to the House of Commons in the 1997 General Election, up to but excluding the 2017 General Election. Comparisons between MPs of different parties are also restricted to MPs first elected in the 1997 General Election, and before the 2017 General Election. Speeches made by the Speaker, including Deputy Speakers, were also excluded. Words contained in parentheses were removed, as they are added by Hansard to provide additional information not actually spoken by the MP.[1](#fn1) Speeches and data on MPs’ gender and party affiliation are from a previously assembled dataset (Odell, [2018](#ref-odell2018)). Information on candidates selected through all women short lists is from the House of Commons Library (Kelly, [2016](#ref-kelly2016)). Unsuccessful General Election candidates selected through all women short lists who were subsequently elected in a byelection are classified as having been selected on an all women short list.

Word classification used the Linguistic Inquiry and Word Count 2015 (LIWC) dictionary (Pennebaker et al., [2015](#ref-pennebaker2015)) and tokenising tools from the Quanteda R package (Benoit, [2018](#ref-benoit2018)). Word counts and words-per-sentence were calculated using stringi (Gagolewski, [2018](#ref-gagolewski2018)), a wrapper to the ICU regex library.

Following Yu ([2014](#ref-yu2014)) drawing on (Newman, Groom, Handelman, & Pennebaker, [2008](#ref-newman2008)) we used the following LIWC categories:

* All Pronouns (pronoun)
* First person singular pronouns (i)
* First person plural pronouns (we)
* Verbs (verb)
* Auxiliary verbs (auxverb)
* Social processes (social)
* Positive emotions (posemo)
* Negative emotions (negemo)
* Tentative words (tentat)
* Articles (article)
* Prepositions (preps)
* Anger words (anger)
* Swear words (swear)
* Cognitive processes (cogproc)
* Words longer than six letters (Sixltr)

We also included mean words-per-sentence (WPS), total speach word count (WC) and Flesch–Kincaid grade level (FK) (Kincaid, Fishburne, Rogers, & Chissom, [1975](#ref-kincaid1975)), calculated using Quanteda (Benoit, [2018](#ref-benoit2018)) and stringi (Gagolewski, [2018](#ref-gagolewski2018)).

### Descriptive Statistics

Labour MPs and Intakes

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| General Election | Total MPs | Labour MPs | Female Labour MPs | Labour MPs Intake | Intake Women | Intake Short list | Nominated Short list |
| 1997 | 659 | 418 | 101 (24%) | 177 | 64 (36%) | 35 | 38 |
| 2001 | 659 | 412 | 95 (23%) | 38 | 4 (11%) | 0 | 0 |
| 2005 | 646 | 355 | 98 (28%) | 40 | 26 (65%) | 23 | 30 |
| 2010 | 650 | 258 | 81 (31%) | 64 | 32 (50%) | 28 | 63 |
| 2015 | 650 | 232 | 99 (43%) | 49 | 31 (63%) | 31 | 77 |

Data in this table is from House of Commons library reports (Audickas, Hawkins, & Cracknell, [2017](#ref-audickas2017); Kelly, [2016](#ref-kelly2016)). All women short lists were not used by Labour during the 2001 General Election.

Number of Speeches and Words in Dataset

Gender

Speeches

Words

All

656412

111180398

Female

148702

26231034

Male

507710

84949364

**Conservatives**

All

285291

44800169

Female

48768

7363031

Male

236523

37437138

**Labour**

All

261942

46494850

Female

84569

15897929

Non-All Women Shortlists

28695

5422776

All Women Shortlists

55874

10475153

Male

177373

30596921

**Liberal Democrat**

All

72716

13485902

Female

7552

1503459

Male

65164

11982443

**Other**

All

36463

6399477

Female

7813

1466615

Male

28650

4932862

## Women vs Men

Effect Sizes for Male and Female Labour MPs

Women

Men

Effect Size

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Pronouns

10.07

4.60

10.15

4.99

0.02

negligible

First person singular pronouns

1.89

2.41

2.02

2.55

0.05

negligible

First person plural pronouns

0.97

1.42

0.99

1.51

0.01

negligible

Verbs

12.82

5.00

12.67

5.36

-0.03

negligible

Auxiliary verbs

7.91

3.45

7.93

3.69

0.01

negligible

Social processes

8.47

4.82

8.18

5.11

-0.06

negligible

Positive emotions

2.73

2.49

2.57

2.54

-0.06

negligible

Negative emotions

1.15

1.68

1.07

1.77

-0.05

negligible

Tentative words

1.48

1.74

1.58

1.90

0.05

negligible

More than six letters

10.58

3.68

10.22

3.92

-0.10

negligible

Articles

7.65

3.30

7.96

3.55

0.09

negligible

Prepositions

12.58

4.42

12.14

4.73

-0.10

negligible

Anger words

0.23

0.81

0.24

0.77

0.01

negligible

Swear words

0.00

0.06

0.00

0.09

0.01

negligible

Cognitive processes

8.68

4.83

8.82

5.15

0.03

negligible

Words per Sentence

43.63

19.68

41.15

20.04

-0.13

negligible

Total Word Count

402.79

691.27

370.18

647.36

-0.05

negligible

Flesh-Kincaid Grade Level

10.81

7.68

9.78

7.87

-0.13

negligible

There are no categories where gender differences meet the effect size threshold of \(|0.2|\) suggested by Cohen ([1988](#ref-cohen1988), pp. 25–26) to indicate a small effect. 4 categories – words with more than six letters, prepositions, words-per-sentence and Flesh-Kincaid grade level – exceeded the \(|0.1|\) threshold suggested by Newman et al ([2008](#ref-newman2008)).

## Short lists vs Non-Short lists

The following plots show changes in the occurences of selected LIWC terms, words-per-sentence, total word count and Flesch–Kincaid grade level, over the course of an MP’s career. There do not appear to be any notable changes in speaking style over the course of female Labour MPs’ careers.

Effect Sizes for Female Labour MPs by selection process

All Women Short lists

Open Shorlists

Effect Size

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Pronouns

10.01

4.67

10.19

4.48

-0.04

negligible

First person singular pronouns

1.86

2.41

1.95

2.42

-0.04

negligible

First person plural pronouns

0.88

1.36

1.16

1.51

-0.19

negligible

Verbs

12.88

5.10

12.69

4.80

0.04

negligible

Auxiliary verbs

7.94

3.49

7.86

3.38

0.02

negligible

Social processes

8.48

4.94

8.46

4.59

0.00

negligible

Positive emotions

2.69

2.52

2.81

2.42

-0.05

negligible

Negative emotions

1.16

1.69

1.13

1.67

0.02

negligible

Tentative words

1.48

1.75

1.49

1.73

0.00

negligible

More than six letters

10.52

3.73

10.70

3.58

-0.05

negligible

Articles

7.69

3.38

7.55

3.15

0.04

negligible

Prepositions

12.55

4.54

12.63

4.15

-0.02

negligible

Anger words

0.23

0.78

0.24

0.88

-0.01

negligible

Swear words

0.00

0.06

0.00

0.05

0.01

negligible

Cognitive processes

8.59

4.90

8.86

4.69

-0.06

negligible

Words per Sentence

44.02

20.45

42.85

18.05

0.06

negligible

Total Word Count

401.77

704.40

404.78

664.97

0.00

negligible

Flesh-Kincaid Grade Level

10.97

7.97

10.48

7.06

0.07

negligible

There are no categories among female Labour MPs by selection process meeting the \(|0.2|\) threshold. Only one category – first person plural pronouns, *d*=0.19 – exceeded \(|0.1|\).

## Conservatives vs Labour

Effect Sizes for All Labour and Conservative MPs

Labour

Conservatives

Effect Size

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Pronouns

10.12

4.87

10.62

4.84

0.10

negligible

First person singular pronouns

1.98

2.51

2.15

2.56

0.07

negligible

First person plural pronouns

0.98

1.48

1.22

1.70

0.15

negligible

Verbs

12.72

5.24

12.93

5.13

0.04

negligible

Auxiliary verbs

7.92

3.61

8.17

3.58

0.07

negligible

Social processes

8.28

5.02

8.13

4.80

-0.03

negligible

Positive emotions

2.62

2.53

2.85

2.66

0.09

negligible

Negative emotions

1.10

1.74

1.05

1.78

-0.03

negligible

Tentative words

1.55

1.85

1.57

1.88

0.01

negligible

More than six letters

10.34

3.85

10.28

3.76

-0.02

negligible

Articles

7.86

3.48

7.82

3.45

-0.01

negligible

Prepositions

12.28

4.64

12.38

4.49

0.02

negligible

Anger words

0.24

0.78

0.24

0.82

0.01

negligible

Swear words

0.00

0.08

0.00

0.10

0.00

negligible

Cognitive processes

8.77

5.05

8.86

5.06

0.02

negligible

Words per Sentence

41.95

19.96

42.76

20.16

0.04

negligible

Total Word Count

380.71

662.03

335.54

592.41

-0.07

negligible

Flesh-Kincaid Grade Level

10.12

7.82

10.41

7.91

0.04

negligible

There are no categories with effect sizes exceeding \(|0.2|\) between Labour and Conservative MPs, like inter-Labour differences.

## All MPs Gender Differences

Effect Sizes for Male and Female MPs, All Parties

Women

Men

Effect Size

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Pronouns

10.31

4.65

10.26

4.90

-0.01

negligible

First person singular pronouns

1.99

2.45

2.00

2.52

0.00

negligible

First person plural pronouns

1.11

1.57

1.08

1.59

-0.02

negligible

Verbs

12.88

4.97

12.80

5.26

-0.02

negligible

Auxiliary verbs

8.00

3.45

8.08

3.64

0.02

negligible

Social processes

8.45

4.77

8.00

4.93

-0.09

negligible

Positive emotions

2.84

2.53

2.69

2.58

-0.06

negligible

Negative emotions

1.10

1.65

1.08

1.78

-0.01

negligible

Tentative words

1.47

1.73

1.61

1.91

0.08

negligible

More than six letters

19.73

6.94

19.25

7.18

-0.07

negligible

Articles

7.62

3.31

8.00

3.51

0.11

negligible

Prepositions

12.58

4.36

12.22

4.62

-0.08

negligible

Anger words

0.23

0.78

0.25

0.82

0.02

negligible

Swear words

0.00

0.05

0.00

0.10

0.01

negligible

Cognitive processes

8.67

4.79

8.93

5.12

0.05

negligible

Words per Sentence

43.25

19.45

42.06

20.12

-0.06

negligible

Total Word Count

377.31

648.92

358.13

623.49

-0.03

negligible

Flesh-Kincaid Grade Level

10.63

7.61

10.16

7.89

-0.06

negligible

## POS Analysis

Part-of-speech (POS) tagging was done using spaCy (Honnibal & Montani, [2017](#ref-honnibal2017)) and the spacyr package (Benoit & Matsuo, [2018](#ref-benoit2018a)).

Part-of-Speech Effect Sizes for Male and Female Labour MPs

Women

Men

Effect Size

Word Type

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Nouns

22.18

9.60

21.66

10.96

-0.04

negligible

Plural Nouns

5.85

3.72

5.03

3.79

-0.16

negligible

Singular Nouns

15.62

9.84

16.01

11.19

0.02

negligible

Adjectives

9.58

4.78

9.28

5.29

-0.02

negligible

Adverbs

4.91

4.26

5.07

4.91

0.03

negligible

Verbs

20.94

9.52

20.78

10.28

-0.02

negligible

Part-of-Speech Effect Sizes for AWS and non-AWS Labour MPs

All Women Short lists

Open Shorlists

Effect Size

Word Type

Mean

SD

Mean

SD

Cohen’s D

Magnitude

All Nouns

22.16

8.78

22.18

10.00

-0.04

negligible

Plural Nouns

6.03

3.60

5.76

3.77

-0.16

negligible

Singular Nouns

15.51

8.97

15.67

10.26

0.02

negligible

Adjectives

9.83

4.59

9.45

4.86

-0.02

negligible

Adverbs

4.95

3.78

4.89

4.49

0.03

negligible

Verbs

20.88

9.04

20.97

9.76

-0.02

negligible

## Tokenising / Keyness

The most commonly used words by both men and women would be protocol decorum expressions, so we calculate the keyness of words to identify gender differences in the choices of topics raised by men and women, and by short-list and non-short list women.

### Men vs Women

Keyness – a linguistic measure of the frequency of different words in two groups of texts – reveals clear gender differences in the most disproportionately common words used by female and male Labour MPs. Unsurprisingly, despite male MPs saying almost twice as many words (30601887 vs 15898845) as their female colleagues, female Labour MPs were more than two-and-a-half (2.61) times as likely to say “women”. They were also much more likely to refer to “women’s” and “woman”. Female Labour MPs also appear much more likely to discuss “children”, “people”, “care”, “families”, “home”, “parents”, “work” and social policy areas such as “services”, “disabled [people]” and “housing” than their male colleagues. Male MPs were more likely to refer to military topics (“Iraq”, “nuclear”), and to parliamentary process and protocol – “question”, “political”, “conservative”, “electoral”, “house”, “party”, “argument” “liberal” and “point” are far more common in speeches by male Labour MPs than by female ones. This could suggest that male MPs are more comfortable using the traditional language of House of Commons debate.

![](data:image/png;base64;base64,)

### Short lists vs Non-Short lists

Keyness differences by selection process are not as obviously stereotypical. Nonetheless, the most common words amongst AWS MPs included “carers”, “disabled”, “bedroom” and “sen”[2](#fn2). Also of note is AWS MPs making more references to their “constituency” and its “constituents”, suggesting that AWS MPs may draw on the fact they were elected by their constituents as a source political legitimacy, at least more than non-AWS MPs.

![](data:image/png;base64;base64,)

### Labour vs Conservative

The keyness differences between Labour and Conservative MPs are much greater than gender differences within Labour. The very high use of “Lady” by Conservative MPs is reflective of the greater proportion of female MPs in other parties, as it is often used to refer to comments by other members of the house. It may also represent a greater use of traditional hosue decorum by Conservative MPs.

![](data:image/png;base64;base64,)

## Bigrams

We created bigrams of all first person plural and singular pronouns for female Labour MPs. As above, AWS MPs are far more likely to make references to their constituency or their constituents.

![](data:image/png;base64;base64,)

## Naive Bayes classification

We trained a Naive Bayes classifier with document-frequency priors and a multinomial distribution to predict the gender of speakers when given speeches by all Labour MPs in our dataset, and the selection process when only given female Labour MPs. The accuracy of both models were roughly equivalent, 70.55% accuracy when predicting gender and 70.84% when predicting short lists. By contrast, the classifier could distinguish between Labour and Conservative speeches with 74.23% accuracy.

## Topic Models

Using topic models to classify text is widely used in social sciences (Grimmer & Stewart, [2013](#ref-grimmer2013)), as, when combined with the large volume of plain text data available, it allows for a rapid and consistent method of analysis . Topic modelling and other statistic methods of textual analysis are not a substitute for reading the texts themselves, but can augment other analysis or – as in this case – analyse and classify larger amounts of text than would be feasible using human coders (Grimmer & Stewart, [2013](#ref-grimmer2013)). Topic models classify a series of documents (in this case individual speeches) into one of a given number of topics, identifying terms that are common in some documents but rare in others. When developing topic models, there is a trade-off between high precision in the classification of each document with broader topics when using smaller numbers of topics, or lower precision in individual speech classification with more finely-grained topics when using larger numbers of topics. Grimmer & Stewart ([2013](#ref-grimmer2013)) also highlight the importance of validating unsurpervised topic models when applied to new sets of texts.

The R package stm (Roberts, Stewart, & Tingley, [2018](#ref-roberts2018)) implements a structured topic model (STM) (Arora et al., [2013](#ref-arora2013); Roberts, Stewart, & Airoldi, [2016](#ref-roberts2016)). An STM incorporates data about the writer or speaker into the topic classification algorithm. We incorporated the AWS status of speakers into our topic model. We used an algorithm developed by Lee and Mimno ([2014](#ref-lee2014c)), implemented in the stm package (Roberts et al., [2018](#ref-roberts2018)), to estimate the number of topics across all speeches made by female Labour MPs, using the “spectral” method developed by Arora et al. ([2013](#ref-arora2013)). The resulting topic model has 69 topics, across 81,607 documents and a dictionary of 115,477 words.

### Short lists vs Non-Short lists

We created a Fruchterman-Reingold (Fruchterman & Reingold, [1991](#ref-fruchterman1991)) diagram of all 69 topic models. Larger vertices indicate more common topics, and the plot implements a colour scale to indicate the proportion of speeches classed in that topic made by AWS and non-AWS female Labour MPs. Vertices

![](data:image/png;base64;base64,)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Topic Number | AWS Speeches | Percent of AWS Speeches | Non-AWS Speeches | Percent of non-AWS Speeches |
| Topic 1 | 1,272 | 2.37% | 353 | 1.27% |
| Topic 2 | 334 | 0.62% | 127 | 0.46% |
| Topic 3 | 241 | 0.45% | 71 | 0.25% |
| Topic 4 | 550 | 1.02% | 133 | 0.48% |
| Topic 5 | 826 | 1.54% | 206 | 0.74% |
| Topic 6 | 978 | 1.82% | 915 | 3.28% |
| Topic 7 | 648 | 1.21% | 236 | 0.85% |
| Topic 8 | 70 | 0.13% | 25 | 0.09% |
| Topic 9 | 265 | 0.49% | 309 | 1.11% |
| Topic 10 | 1,024 | 1.91% | 513 | 1.84% |
| Topic 11 | 940 | 1.75% | 580 | 2.08% |
| Topic 12 | 313 | 0.58% | 319 | 1.14% |
| Topic 13 | 325 | 0.61% | 146 | 0.52% |
| Topic 14 | 1,596 | 2.97% | 461 | 1.65% |
| Topic 15 | 1,386 | 2.58% | 642 | 2.30% |
| Topic 16 | 1,407 | 2.62% | 525 | 1.88% |
| Topic 17 | 3,690 | 6.87% | 1,459 | 5.23% |
| Topic 18 | 1,026 | 1.91% | 847 | 3.04% |
| Topic 19 | 640 | 1.19% | 423 | 1.52% |
| Topic 20 | 872 | 1.62% | 216 | 0.77% |
| Topic 21 | 658 | 1.23% | 363 | 1.30% |
| Topic 22 | 818 | 1.52% | 439 | 1.57% |
| Topic 23 | 795 | 1.48% | 518 | 1.86% |
| Topic 24 | 385 | 0.72% | 199 | 0.71% |
| Topic 25 | 240 | 0.45% | 74 | 0.27% |
| Topic 26 | 788 | 1.47% | 200 | 0.72% |
| Topic 27 | 266 | 0.50% | 120 | 0.43% |
| Topic 28 | 847 | 1.58% | 350 | 1.25% |
| Topic 29 | 1,110 | 2.07% | 327 | 1.17% |
| Topic 30 | 1,132 | 2.11% | 462 | 1.66% |
| Topic 31 | 996 | 1.85% | 975 | 3.49% |
| Topic 32 | 76 | 0.14% | 64 | 0.23% |
| Topic 33 | 1,238 | 2.31% | 985 | 3.53% |
| Topic 34 | 1,124 | 2.09% | 521 | 1.87% |
| Topic 35 | 650 | 1.21% | 657 | 2.35% |
| Topic 36 | 601 | 1.12% | 154 | 0.55% |
| Topic 37 | 455 | 0.85% | 194 | 0.70% |
| Topic 38 | 1,246 | 2.32% | 991 | 3.55% |
| Topic 39 | 1,917 | 3.57% | 936 | 3.35% |
| Topic 40 | 848 | 1.58% | 290 | 1.04% |
| Topic 41 | 63 | 0.12% | 40 | 0.14% |
| Topic 42 | 853 | 1.59% | 590 | 2.11% |
| Topic 43 | 1,344 | 2.50% | 604 | 2.16% |
| Topic 44 | 814 | 1.52% | 288 | 1.03% |
| Topic 45 | 602 | 1.12% | 474 | 1.70% |
| Topic 46 | 709 | 1.32% | 150 | 0.54% |
| Topic 47 | 664 | 1.24% | 245 | 0.88% |
| Topic 48 | 940 | 1.75% | 901 | 3.23% |
| Topic 49 | 835 | 1.55% | 563 | 2.02% |
| Topic 50 | 1,328 | 2.47% | 1,219 | 4.37% |
| Topic 51 | 1,076 | 2.00% | 323 | 1.16% |
| Topic 52 | 196 | 0.36% | 85 | 0.30% |
| Topic 53 | 590 | 1.10% | 293 | 1.05% |
| Topic 54 | 1,057 | 1.97% | 824 | 2.95% |
| Topic 55 | 302 | 0.56% | 157 | 0.56% |
| Topic 56 | 535 | 1.00% | 398 | 1.43% |
| Topic 57 | 656 | 1.22% | 314 | 1.13% |
| Topic 58 | 468 | 0.87% | 182 | 0.65% |
| Topic 59 | 426 | 0.79% | 183 | 0.66% |
| Topic 60 | 562 | 1.05% | 297 | 1.06% |
| Topic 61 | 86 | 0.16% | 28 | 0.10% |
| Topic 62 | 550 | 1.02% | 343 | 1.23% |
| Topic 63 | 690 | 1.28% | 252 | 0.90% |
| Topic 64 | 594 | 1.11% | 244 | 0.87% |
| Topic 65 | 662 | 1.23% | 457 | 1.64% |
| Topic 66 | 1,493 | 2.78% | 527 | 1.89% |
| Topic 67 | 737 | 1.37% | 451 | 1.62% |
| Topic 68 | 279 | 0.52% | 145 | 0.52% |
| Topic 69 | 1 | 0.00% | NA | NA% |

|  |  |
| --- | --- |
| Topic Number | Top Ten Words |
| Topic 1 | secretary, state, tell, ministers, given, today, department, can, confirm, said |
| Topic 2 | safety, register, registration, indicated, registered, electoral, risk, risks, number, individual |
| Topic 3 | make, sure, statement, progress, difference, northern, ireland, towards, representations, responsibilities |
| Topic 4 | debt, water, credit, charges, pay, loan, loans, people, financial, cost |
| Topic 5 | house, committee, parliament, leader, select, motion, parliamentary, debate, scrutiny, business |
| Topic 6 | new, development, work, need, investment, strategy, must, programme, working, also |
| Topic 7 | road, petition, residents, car, vehicles, petitioners, dogs, roads, site, house |
| Topic 8 | important, agree, welcome, country, making, particularly, thank, part, makes, good |
| Topic 9 | companies, market, company, competition, energy, consumers, prices, price, consumer, customers |
| Topic 10 | women, men, equality, women’s, discrimination, rights, gender, equal, woman, marriage |
| Topic 11 | energy, climate, fuel, change, green, carbon, emissions, gas, environmental, industry |
| Topic 12 | office, post, offices, royal, service, closure, mail, services, network, christmas |
| Topic 13 | mr, north, south, east, west, spoke, friends, birmingham, talked, central |
| Topic 14 | pension, scheme, benefit, pensions, benefits, pensioners, system, credit, allowance, income |
| Topic 15 | economy, jobs, economic, growth, unemployment, country, investment, chancellor, budget, crisis |
| Topic 16 | schools, school, education, children, teachers, parents, pupils, educational, special, primary |
| Topic 17 | want, say, one, think, know, need, us, get, go, see |
| Topic 18 | review, report, commission, independent, process, recommendations, inquiry, also, system, standards |
| Topic 19 | business, businesses, small, financial, bank, banks, insurance, rates, industry, enterprise |
| Topic 20 | wales, industry, welsh, north-east, england, assembly, constituency, jobs, manufacturing, uk |
| Topic 21 | care, services, social, mental, need, health, home, provision, service, older |
| Topic 22 | pay, work, workers, employment, working, wage, minimum, employers, paid, national |
| Topic 23 | amendment, clause, amendments, new, 1, lords, section, 2, act, clauses |
| Topic 24 | report, last, since, said, received, published, year, following, official, end |
| Topic 25 | made, clear, impact, decision, changes, recent, assessment, decisions, effect, proposed |
| Topic 26 | funding, cuts, fund, cut, budget, grant, spending, bbc, review, flood |
| Topic 27 | money, spent, extra, spend, liberal, cost, spending, value, opposition, tory |
| Topic 28 | constituency, great, community, proud, many, sport, one, also, world, new |
| Topic 29 | families, child, poverty, children, parents, work, credit, working, family, living |
| Topic 30 | party, conservative, vote, parliament, political, election, labour, parties, scottish, elected |
| Topic 31 | point, can, may, issue, take, however, whether, matter, understand, consider |
| Topic 32 | member, said, lady, mentioned, raised, comments, speech, referred, points, remarks |
| Topic 33 | european, uk, eu, countries, united, union, europe, states, british, trade |
| Topic 34 | education, skills, young, training, students, university, college, higher, science, apprenticeships |
| Topic 35 | local, authorities, authority, planning, community, communities, councils, area, guidance, system |
| Topic 36 | disabled, carers, disability, support, disabilities, needs, caring, autism, learning, can |
| Topic 37 | environment, marine, fishing, sea, industry, natural, fish, countryside, rural, fisheries |
| Topic 38 | justice, court, violence, victims, cases, criminal, domestic, courts, prison, offence |
| Topic 39 | international, foreign, rights, human, peace, un, conflict, world, aid, war |
| Topic 40 | day, family, never, told, families, life, happened, constituent, man, went |
| Topic 41 | proposals, future, forward, consultation, plans, meet, paper, current, discuss, bring |
| Topic 42 | behaviour, crime, antisocial, alcohol, young, drugs, drug, problem, use, tackle |
| Topic 43 | housing, homes, social, affordable, private, home, accommodation, rent, need, properties |
| Topic 44 | question, order, mr, put, asked, answer, questions, ask, speaker, time |
| Topic 45 | research, cancer, treatment, medical, condition, screening, disease, can, patients, use |
| Topic 46 | online, internet, farmers, animals, digital, animal, broadband, sites, tickets, technology |
| Topic 47 | defence, forces, armed, plymouth, personnel, service, military, army, nuclear, royal |
| Topic 48 | information, home, security, data, immigration, control, orders, system, terrorism, appeal |
| Topic 49 | police, officers, crime, policing, home, force, service, forces, officer, chief |
| Topic 50 | nhs, hospital, patients, health, services, hospitals, care, service, trust, trusts |
| Topic 51 | tax, budget, cut, chancellor, cuts, rate, income, vat, benefit, hit |
| Topic 52 | years, now, two, time, first, three, past, one, months, ago |
| Topic 53 | staff, doctors, emergency, medical, service, training, nurses, royal, junior, ambulance |
| Topic 54 | bill, legislation, act, law, rights, provisions, powers, regulations, place, believe |
| Topic 55 | public, sector, private, organisations, service, voluntary, services, society, community, organisation |
| Topic 56 | health, national, inequalities, programme, suicide, disease, department, prevention, among, risk |
| Topic 57 | council, london, areas, city, area, constituency, centre, rural, county, liverpool |
| Topic 58 | advice, legal, cases, civil, hull, aid, case, compensation, claims, service |
| Topic 59 | people, work, many, young, get, people’s, can, help, lives, job |
| Topic 60 | tax, revenue, relief, duty, uk, avoidance, hmrc, charities, companies, taxation |
| Topic 61 | government, government’s, policy, labour, previous, scotland, scottish, commitment, policies, coalition |
| Topic 62 | trafficking, home, uk, asylum, refugees, immigration, country, human, migration, britain |
| Topic 63 | food, products, industry, smoking, advertising, tobacco, ban, product, standards, shops |
| Topic 64 | members, debate, many, issues, also, today, heard, opportunity, hope, issue |
| Topic 65 | children, child, parents, young, children’s, family, contact, vulnerable, adoption, abuse |
| Topic 66 | transport, rail, bus, services, line, travel, train, network, passengers, london |
| Topic 67 | year, million, number, increase, figures, increased, billion, 1, average, cost |
| Topic 68 | support, ensure, can, help, aware, taking, take, provide, action, continue |
| Topic 69 | deal, recently, new, can, lack, great, concern, done, move, given |

## Discussion

There do not appear to be substantial or meaningful differences in the speaking styles of female Labour MPs selected through all women short lists when compared to their female colleagues selected through open short lists using LIWC.

There is more gender distinction in some selected terms and topics. AWS MPs are far more likely to make reference to their constituency and their constituents.

# Appendix

## Full topic model summary

## A topic model with 69 topics, 81607 documents and a 115477 word dictionary.

## Topic 1 Top Words:  
## Highest Prob: secretary, state, tell, ministers, given, today, department   
## FREX: secretary, state, confirm, tell, ministers, state's, minister's   
## Lift: dhar, lowell, qatada's, #nationalistsconfused, 1135, 2.5bn, 36,500   
## Score: secretary, state, confirm, state's, tell, ministers, department   
## Topic 2 Top Words:  
## Highest Prob: safety, register, registration, indicated, registered, electoral, risk   
## FREX: registration, indicated, hse, canvass, register, gurkhas, safety   
## Lift: aps, hse, 10-litre, 13a, 14,940, 1760, 1867   
## Score: safety, registration, register, electoral, indicated, registered, hse   
## Topic 3 Top Words:  
## Highest Prob: make, sure, statement, progress, difference, northern, ireland   
## FREX: statement, make, sure, progress, ireland, representations, difference   
## Lift: 101269, 101548, 102938, 103414, 106583, 107305, 109413   
## Score: make, statement, progress, sure, ireland, northern, milton   
## Topic 4 Top Words:  
## Highest Prob: debt, water, credit, charges, pay, loan, loans   
## FREX: payday, loan, lenders, debts, loans, debt, charges   
## Lift: 1,021, 1,025, 1,106, 1,189, 1,273, 1,385, 1,413   
## Score: debt, water, payday, loan, loans, lenders, credit   
## Topic 5 Top Words:  
## Highest Prob: house, committee, parliament, leader, select, motion, parliamentary   
## FREX: select, leader, house, motion, committee, backbench, scrutiny   
## Lift: e-petitions, praying-against, sherlock, wednesdays, sittings, thursdays, 10,000-signature   
## Score: committee, house, leader, select, scrutiny, parliament, motion   
## Topic 6 Top Words:  
## Highest Prob: new, development, work, need, investment, strategy, must   
## FREX: development, strategy, develop, project, regional, projects, partnership   
## Lift: #1,150, 1.245, 1.875, 128406, 131,000, 18-the, 2002-around   
## Score: development, regional, investment, strategy, infrastructure, projects, work   
## Topic 7 Top Words:  
## Highest Prob: road, petition, residents, car, vehicles, petitioners, dogs   
## FREX: petitioners, lap-dancing, petition, dogs, dog, pedestrians, cycling   
## Lift: 0.037, 0.044, 0fficial, 1,042, 1,072, 1,108, 1,122   
## Score: petitioners, petition, dogs, road, residents, dog, declares   
## Topic 8 Top Words:  
## Highest Prob: important, agree, welcome, country, making, particularly, thank   
## FREX: agree, welcome, important, absolutely, makes, making, friend's   
## Lift: adequacies, and-importantly-much, ayse, ballantine, bobtail-and, bostrom, bucketfuls   
## Score: agree, important, thank, welcome, friend's, absolutely, country   
## Topic 9 Top Words:  
## Highest Prob: companies, market, company, competition, energy, consumers, prices   
## FREX: competition, companies, market, wholesale, suppliers, company, regulator   
## Lift: 1,105, 1,345, ashington, boakye, nord, over-charging, price-fixing   
## Score: companies, consumers, energy, market, company, prices, competition   
## Topic 10 Top Words:  
## Highest Prob: women, men, equality, women's, discrimination, rights, gender   
## FREX: gender, bishops, transgender, women's, women, abortion, same-sex   
## Lift: balmforth, bpas, celebrants, cohabitants, jessy, msafiri, natal   
## Score: women, women's, equality, men, gender, discrimination, marriage   
## Topic 11 Top Words:  
## Highest Prob: energy, climate, fuel, change, green, carbon, emissions   
## FREX: renewables, solar, insulation, feed-in, biofuels, greenhouse, dioxide   
## Lift: fossil, #2,500, #solar, 1-are, 1,129, 1,214, 1,343   
## Score: energy, fuel, carbon, emissions, climate, renewable, renewables   
## Topic 12 Top Words:  
## Highest Prob: office, post, offices, royal, service, closure, mail   
## FREX: offices, mail, sub-post, post, sub-postmasters, closures, consignia   
## Lift: #1.8, #210, #450, 1,001, 1,352, 1,639, 1,827   
## Score: post, offices, office, mail, closure, postal, sub-post   
## Topic 13 Top Words:  
## Highest Prob: mr, north, south, east, west, spoke, friends   
## FREX: ealing, spoke, dorset, lothian, ayrshire, glasgow, chris   
## Lift: argos's, blackford, cairns's, clemency, marxist, no2id, 0.66   
## Score: mr, east, north, south, west, spoke, birmingham   
## Topic 14 Top Words:  
## Highest Prob: pension, scheme, benefit, pensions, benefits, pensioners, system   
## FREX: pension, esa, pensions, claimants, retirement, pip, pensioners   
## Lift: means-testing, #20,000, #400, 0º, 1,052, 1,366, 1,482   
## Score: pension, pensions, pensioners, allowance, scheme, retirement, credit   
## Topic 15 Top Words:  
## Highest Prob: economy, jobs, economic, growth, unemployment, country, investment   
## FREX: unemployment, recession, growth, economy, obr, deficit, inflation   
## Lift: 0.76, 0.83, 1,196, 1,319, 1.04, 1.32, 10-about   
## Score: economy, jobs, unemployment, growth, economic, recession, chancellor   
## Topic 16 Top Words:  
## Highest Prob: schools, school, education, children, teachers, parents, pupils   
## FREX: academies, pupil, grammar, schools, pupils, teachers, ofsted   
## Lift: sen2, 11-plus, leas, pupil, 000-to, 1-regardless, 1,000-pupil   
## Score: schools, school, teachers, pupils, children, education, parents   
## Topic 17 Top Words:  
## Highest Prob: want, say, one, think, know, need, us   
## FREX: think, say, things, want, something, saying, going   
## Lift: about-part, arguing-i, beneficial-we, career-many, cause-the, clam, compatriot   
## Score: think, want, get, say, things, going, us   
## Topic 18 Top Words:  
## Highest Prob: review, report, commission, independent, process, recommendations, inquiry   
## FREX: recommendations, inquiry, panel, audit, independent, recommendation, reviews   
## Lift: bourn, cag's, clarke's, clowes, dg, emag, equitable's   
## Score: fsa, inquiry, review, commission, recommendations, report, independent   
## Topic 19 Top Words:  
## Highest Prob: business, businesses, small, financial, bank, banks, insurance   
## FREX: smes, medium-sized, businesses, bank, enterprises, enterprise, banking   
## Lift: 0.21, 0.84, 1,034, 1,130, 10-fold, 1130, 12.19   
## Score: businesses, business, bank, banks, banking, insurance, small   
## Topic 20 Top Words:  
## Highest Prob: wales, industry, welsh, north-east, england, assembly, constituency   
## FREX: welsh, wales, steel, cardiff, north-east, assembly, visteon   
## Lift: a55, angharad, bethesda, co-investment, dewhirst, dogger, gorge   
## Score: wales, welsh, assembly, manufacturing, steel, north-east, yorkshire   
## Topic 21 Top Words:  
## Highest Prob: care, services, social, mental, need, health, home   
## FREX: mental, care, social, elderly, older, advocacy, services   
## Lift: #900,000, 1,051, 1,312, 10,758, 104962, 1089, 13,198   
## Score: care, mental, services, social, health, older, homes   
## Topic 22 Top Words:  
## Highest Prob: pay, work, workers, employment, working, wage, minimum   
## FREX: wage, workers, zero-hours, employees, paternity, employer, minimum   
## Lift: 6.70, 8.20, 85p, awb, e-balloting, hannett, increments   
## Score: wage, workers, employers, employment, pay, employees, minimum   
## Topic 23 Top Words:  
## Highest Prob: amendment, clause, amendments, new, 1, lords, section   
## FREX: amendment, nos, insert, subsection, clause, amendments, clauses   
## Lift: 153a, 22a, 287, 50b, 51b, counter-notice, insured's   
## Score: clause, amendment, amendments, lords, nos, insert, subsection   
## Topic 24 Top Words:  
## Highest Prob: report, last, since, said, received, published, year   
## FREX: march, vol, official, january, july, november, published   
## Lift: 1-2ws, 1,033, 1,099, 1,124,818, 1,337, 1,368,186, 1,595   
## Score: report, official, vol, published, march, april, november   
## Topic 25 Top Words:  
## Highest Prob: made, clear, impact, decision, changes, recent, assessment   
## FREX: made, decision, assessment, clear, decisions, impact, implications   
## Lift: 104963, 107312, 116,400, 125214, 125828, 125830, 126370   
## Score: made, assessment, impact, changes, decision, decisions, clear   
## Topic 26 Top Words:  
## Highest Prob: funding, cuts, fund, cut, budget, grant, spending   
## FREX: flood, funding, bbc, formula, grant, flooding, floods   
## Lift: #10.89, #12, #3.3, #bbcdiversity, 1,027, 1,536-will, 1,546   
## Score: funding, cuts, flood, bbc, budget, spending, flooding   
## Topic 27 Top Words:  
## Highest Prob: money, spent, extra, spend, liberal, cost, spending   
## FREX: money, spent, liberal, spend, democrats, tories, tory   
## Lift: 1-but, 10,309.63, 1228, 158.8, 1763, 18-that, 1979-80   
## Score: money, liberal, tory, democrats, conservatives, tories, spending   
## Topic 28 Top Words:  
## Highest Prob: constituency, great, community, proud, many, sport, one   
## FREX: maiden, arts, football, museum, museums, sport, olympic   
## Lift: 0.27, 0.51, 1,084, 1,126, 1,468, 1,580-plus, 1,983   
## Score: arts, sport, museum, maiden, heritage, football, constituency   
## Topic 29 Top Words:  
## Highest Prob: families, child, poverty, children, parents, work, credit   
## FREX: lone, poverty, childcare, families, low-income, child, nursery   
## Lift: 1,000-discriminates, 1,000-not, 1,080, 1,142,600, 1,170, 1,390, 1,664   
## Score: poverty, child, families, children, parents, credit, lone   
## Topic 30 Top Words:  
## Highest Prob: party, conservative, vote, parliament, political, election, labour   
## FREX: party, vote, voting, conservative, party's, voters, election   
## Lift: alphabetical, gentry, olga, one-party, randomisation, 1,166, 1,294   
## Score: party, conservative, vote, scottish, election, elections, political   
## Topic 31 Top Words:  
## Highest Prob: point, can, may, issue, take, however, whether   
## FREX: matter, point, understand, consider, certainly, accept, possible   
## Lift: 450,000-in, advised-by, backslid, bill-albeit, bizarre-but, can-enable, cases-roughly   
## Score: point, matter, issue, gentleman's, consider, shall, whether   
## Topic 32 Top Words:  
## Highest Prob: member, said, lady, mentioned, raised, comments, speech   
## FREX: member, lady, comments, remarks, bromley, interesting, chislehurst   
## Lift: 12.20, 130b, 1991-forcing, 2,784, 54,000-worth, achieved-is, arguments-and   
## Score: member, lady, comments, said, speech, raised, points   
## Topic 33 Top Words:  
## Highest Prob: european, uk, eu, countries, united, union, europe   
## FREX: accession, enlargement, wto, lisbon, treaty, eu, doha   
## Lift: 13652, 13653, 13654, 1707, balkan, barnier, blackmailing   
## Score: eu, european, countries, union, treaty, europe, trade   
## Topic 34 Top Words:  
## Highest Prob: education, skills, young, training, students, university, college   
## FREX: ema, fe, students, apprenticeship, universities, qualifications, apprenticeships   
## Lift: ema, #ne, 1,188, 1,308, 1,555-what, 1,740, 1,803   
## Score: students, education, young, skills, apprenticeships, training, universities   
## Topic 35 Top Words:  
## Highest Prob: local, authorities, authority, planning, community, communities, councils   
## FREX: authorities, local, authority, planning, councils, councillors, locally   
## Lift: achcew, central-local, laa, lsp, maas, observances, place-shaping   
## Score: local, authorities, authority, councils, planning, communities, community   
## Topic 36 Top Words:  
## Highest Prob: disabled, carers, disability, support, disabilities, needs, caring   
## FREX: carers, autism, autistic, disabled, disabilities, disability, dementia   
## Lift: autism, rnib, #185, #85, #hellomynameis, 1,400-one, 10-person   
## Score: carers, disabled, disability, autism, disabilities, caring, dementia   
## Topic 37 Top Words:  
## Highest Prob: environment, marine, fishing, sea, industry, natural, fish   
## FREX: fishermen, cod, forestry, biodiversity, habitats, mmo, fishing   
## Lift: aquaculture, arable, bee-friendly, biodiversity, birdlife, bycatch, caterpillar   
## Score: marine, fishing, fishermen, fish, fisheries, wildlife, conservation   
## Topic 38 Top Words:  
## Highest Prob: justice, court, violence, victims, cases, criminal, domestic   
## FREX: attorney-general, defendants, defendant, prison, prosecutors, solicitor-general, stalking   
## Lift: #9, 0.08, 0.48, 1,046, 1,237, 10,544, 10.15   
## Score: violence, prison, court, offence, criminal, rape, victims   
## Topic 39 Top Words:  
## Highest Prob: international, foreign, rights, human, peace, un, conflict   
## FREX: israel, palestinian, israeli, gaza, sri, zimbabwe, iran   
## Lift: lankan, saddam, #aleppo, 1,010, 1,476, 1,591, 1224   
## Score: un, israel, syria, humanitarian, palestinian, israeli, iraq   
## Topic 40 Top Words:  
## Highest Prob: day, family, never, told, families, life, happened   
## FREX: man, died, son, story, stories, hillsborough, tragedy   
## Lift: 10,000-seat, 12-inch, 1234, 1519-20, 1635, 1710, 174,995   
## Score: families, holocaust, family, constituent, man, died, mother   
## Topic 41 Top Words:  
## Highest Prob: proposals, future, forward, consultation, plans, meet, paper   
## FREX: proposals, consultation, paper, plans, forward, discuss, white   
## Lift: 10.42, 107910, 109648, 114061, 119621, 141605, 141607   
## Score: proposals, consultation, plans, future, forward, paper, white   
## Topic 42 Top Words:  
## Highest Prob: behaviour, crime, antisocial, alcohol, young, drugs, drug   
## FREX: antisocial, asbos, alcohol, alcohol-related, binge, psychoactive, drinking   
## Lift: acquisitive, addaction, auto, bailes, crawlers, ghb, gilpin   
## Score: antisocial, crime, behaviour, alcohol, drug, drugs, cannabis   
## Topic 43 Top Words:  
## Highest Prob: housing, homes, social, affordable, private, home, accommodation   
## FREX: housing, tenants, rented, tenancies, homelessness, leasehold, landlords   
## Lift: one-for-one, one-bedroom, rented, right-to-buy, #19, #21.5, #28.5   
## Score: housing, homes, tenants, rented, rent, landlords, affordable   
## Topic 44 Top Words:  
## Highest Prob: question, order, mr, put, asked, answer, questions   
## FREX: question, answer, questions, speaker, asked, deputy, answers   
## Lift: 11.00, 11.57, 12.26, 1223, 1232, 1412, 1555-56   
## Score: question, speaker, mr, answer, deputy, order, questions   
## Topic 45 Top Words:  
## Highest Prob: research, cancer, treatment, medical, condition, screening, disease   
## FREX: embryos, prostate, cervical, hepatitis, cloning, transplant, embryo   
## Lift: abnormalities, cystic, embryo, fertilisation, marrow, @cfaware, #500   
## Score: cancer, patients, embryos, screening, treatment, tissue, breast   
## Topic 46 Top Words:  
## Highest Prob: online, internet, farmers, animals, digital, animal, broadband   
## FREX: cull, badgers, badger, fur, bovine, mink, culling   
## Lift: culling, @daisydumble, @donna\_smiley, @jimspin, @leamingtonsbc, @maggieannehayes, @nhconsortium   
## Score: farmers, animals, internet, cull, animal, online, badgers   
## Topic 47 Top Words:  
## Highest Prob: defence, forces, armed, plymouth, personnel, service, military   
## FREX: mod, naval, hms, submarines, dockyard, veterans, armed   
## Lift: hms, submarine, submarines, 1,000-people, 1,625, 1,705, 10-3   
## Score: defence, armed, forces, plymouth, military, personnel, mod   
## Topic 48 Top Words:  
## Highest Prob: information, home, security, data, immigration, control, orders   
## FREX: extradition, tpims, sia, warrant, detention, checks, tpim   
## Lift: carlile's, carlile, sia, 10-month, 10,410, 10,500-for, 10.45   
## Score: immigration, terrorism, detention, terrorist, tpims, home, security   
## Topic 49 Top Words:  
## Highest Prob: police, officers, crime, policing, home, force, service   
## FREX: constable, constables, officers, policing, police, soca, ipcc   
## Lift: 2003-morecambe, 9,650, a19s, ashleys, bigg, bounties, ckp   
## Score: police, officers, policing, crime, forces, constable, neighbourhood   
## Topic 50 Top Words:  
## Highest Prob: nhs, hospital, patients, health, services, hospitals, care   
## FREX: dentists, dentistry, pharmacies, pct, nhs, hospitals, hospital   
## Lift: acos, bequest, bernstein, bodmin, catto, cayton, chailey   
## Score: nhs, patients, hospital, health, patient, hospitals, care   
## Topic 51 Top Words:  
## Highest Prob: tax, budget, cut, chancellor, cuts, rate, income   
## FREX: 50p, vat, millionaires, hit, tax, allowances, credits   
## Lift: #840, 0.76p, 1,003, 1,009, 1,226, 1,275, 1,296   
## Score: tax, vat, budget, credits, chancellor, cuts, income   
## Topic 52 Top Words:  
## Highest Prob: years, now, two, time, first, three, past   
## FREX: years, three, months, ago, two, past, weeks   
## Lift: 10-week-old, 10,616, 11-month, 11-point, 1758, 196b, 63,500   
## Score: years, months, two, ago, three, past, weeks   
## Topic 53 Top Words:  
## Highest Prob: staff, doctors, emergency, medical, service, training, nurses   
## FREX: ambulance, junior, staffing, doctors, halifax, posts, nurses   
## Lift: #i'm, 03, 1-who, 1,454, 1,631, 10,000-strong, 10,000-with   
## Score: staff, doctors, ambulance, nurses, medical, emergency, junior   
## Topic 54 Top Words:  
## Highest Prob: bill, legislation, act, law, rights, provisions, powers   
## FREX: bill, legislation, bill's, provisions, passage, regulations, legislative   
## Lift: 1865, 1990s-when, 1998-it, 19may2000, 2003-largely, 2005-have, 29-year   
## Score: bill, legislation, provisions, rights, law, powers, regulations   
## Topic 55 Top Words:  
## Highest Prob: public, sector, private, organisations, service, voluntary, services   
## FREX: public, voluntary, organisations, sector, private, co-operative, volunteering   
## Lift: 1844, af, carpetbaggers, nebulous, puk, 1075, 170-year   
## Score: public, sector, private, voluntary, organisations, service, services   
## Topic 56 Top Words:  
## Highest Prob: health, national, inequalities, programme, suicide, disease, department   
## FREX: flu, hiv, pandemic, inequalities, infections, suicide, mortality   
## Lift: kirkley, lowestoft, nihr, acupuncture, influenza, #148, #3.6   
## Score: health, vaccine, flu, inequalities, hiv, infection, suicide   
## Topic 57 Top Words:  
## Highest Prob: council, london, areas, city, area, constituency, centre   
## FREX: county, mayor, borough, cities, liverpool, city, regeneration   
## Lift: #12,000, #14.4, #356, #38, #5,000, #500,000, #66.6   
## Score: london, council, city, regeneration, county, rural, borough   
## Topic 58 Top Words:  
## Highest Prob: advice, legal, cases, civil, hull, aid, case   
## FREX: hull, tribunal, legal, compensation, solicitors, advice, concentrix   
## Lift: 0300, 1,997, 1.148, 1.4m, 112.8, 1147, 128,687   
## Score: legal, advice, hull, aid, compensation, civil, tribunal   
## Topic 59 Top Words:  
## Highest Prob: people, work, many, young, get, people's, can   
## FREX: people, people's, get, getting, work, young, jobcentre   
## Lift: 2,425, 294,488, 3,699, 37,290, 5,320, 50-to-64, 75589   
## Score: people, young, work, get, youth, many, people's   
## Topic 60 Top Words:  
## Highest Prob: tax, revenue, relief, duty, uk, avoidance, hmrc   
## FREX: evasion, hmrc, gaar, avoidance, inland, stamp, revenue   
## Lift: 1,643, 3.12, 32.2, 44a, 80g, aaronson, aat   
## Score: tax, hmrc, avoidance, revenue, relief, evasion, territories   
## Topic 61 Top Words:  
## Highest Prob: government, government's, policy, labour, previous, scotland, scottish   
## FREX: government, previous, policy, government's, scotland, coalition, scottish   
## Lift: 2005-perhaps, 80994, actually-that, agency-when, agenda-access, alloway, aware-in   
## Score: government, scotland, scottish, labour, policy, government's, previous   
## Topic 62 Top Words:  
## Highest Prob: trafficking, home, uk, asylum, refugees, immigration, country   
## FREX: trafficking, slavery, trafficked, sierra, leone, slave, dubs   
## Lift: #7, 0.025, 1-yes, 1,060, 1,483, 1,746, 1.123   
## Score: trafficking, refugees, asylum, slavery, trafficked, immigration, sierra   
## Topic 63 Top Words:  
## Highest Prob: food, products, industry, smoking, advertising, tobacco, ban   
## FREX: gambling, betting, sunbed, tobacco, cocoa, meat, supermarkets   
## Lift: 0.7p, 00, 0157, 1,000-almost, 1,032, 1,200-i, 1,666   
## Score: food, smoking, products, tobacco, advertising, gambling, industry   
## Topic 64 Top Words:  
## Highest Prob: members, debate, many, issues, also, today, heard   
## FREX: members, debate, heard, speak, sides, issues, hear   
## Lift: noakes, 6.47pm, accept-or, analysis-but, aspect-listening-is, bunfight, called-making   
## Score: members, debate, issues, many, opposition, heard, constituents   
## Topic 65 Top Words:  
## Highest Prob: children, child, parents, young, children's, family, contact   
## FREX: csa, adopters, adoption, child's, cafcass, looked-after, children's   
## Lift: csa, @mandatenow, 10-month-old, 10-went, 12j, 150765, 16-only   
## Score: children, child, parents, young, children's, adoption, child's   
## Topic 66 Top Words:  
## Highest Prob: transport, rail, bus, services, line, travel, train   
## FREX: rail, passengers, passenger, heathrow, hs2, freight, high-speed   
## Lift: 12-car, 15.15, 50.1, adtranz, anti-icing, bahn, chilterns   
## Score: rail, transport, bus, passengers, fares, trains, hs2   
## Topic 67 Top Words:  
## Highest Prob: year, million, number, increase, figures, increased, billion   
## FREX: million, figures, figure, increased, increase, compared, year   
## Lift: #112, #3,850, #84.3, #87.2, 1,249, 102269, 122.9   
## Score: million, year, billion, increase, figures, average, increased   
## Topic 68 Top Words:  
## Highest Prob: support, ensure, can, help, aware, taking, take   
## FREX: aware, ensure, support, taking, steps, continue, help   
## Lift: 103684, 103965, 107320, 111532, 112584, 113698, 117890   
## Score: support, ensure, steps, aware, help, taking, department   
## Topic 69 Top Words:  
## Highest Prob: deal, recently, new, can, lack, great, concern   
## FREX: deal, recently, lack, elsewhere, concern, great, improved   
## Lift: 2004-a, 721,000, added-gva-per, age-of, centres-jacs-which, employment-can, index-the   
## Score: deal, recently, new, worktrack, lack, can, great

# References

Arora, S., Ge, R., Halpern, Y., Mimno, D., Moitra, A., Sontag, D., … Zhu, M. (2013). A Practical Algorithm for Topic Modeling with Provable Guarantees. In S. Dasgupta & D. McAllester (Eds.), *Proceedings of the 30th International Conference on Machine Learning* (Vol. 28, pp. 280–288). Atlanta, Georgia, USA: PMLR. Retrieved from <http://proceedings.mlr.press/v28/arora13.pdf>

Audickas, L., Hawkins, O., & Cracknell, R. (2017). *UK Election Statistics: 1918-2017* (Briefing Paper No. CBP7529) (p. 89). London: House of Commons Library. Retrieved from <http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7529>

Benoit, K. (2018). *Quanteda: Quantitative Analysis of Textual Data*. <https://doi.org/10.5281/zenodo.1004683>

Benoit, K., & Matsuo, A. (2018). *Spacyr: Wrapper to the ’spaCy’ ’NLP’ Library*. Retrieved from <http://github.com/quanteda/spacyr>

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bligh, M., Merolla, J., Schroedel, J. R., & Gonzalez, R. (2010). Finding Her Voice: Hillary Clinton’s Rhetoric in the 2008 Presidential Campaign. *Women’s Studies*, *39*(8), 823–850. <https://doi.org/10.1080/00497878.2010.513316>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. <https://doi.org/10.1002/spe.4380211102>

Gagolewski, M. (2018). R package stringi: Character string processing facilities. <https://doi.org/10.5281/zenodo.1292492>

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*(03), 267–297. <https://doi.org/10.1093/pan/mps028>

Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. *To Appear*. Retrieved from <https://spacy.io>

Jones, J. J. (2016). Talk "Like a Man": The Linguistic Styles of Hillary Clinton, 1992-2013. *Perspectives on Politics*, *14*(03), 625–642. <https://doi.org/10.1017/S1537592716001092>

Kelly, R. (2016). *All-women shortlists* (Briefing Paper No. 5057) (p. 34). London: House of Commons Library. Retrieved from <https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN05057>

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel:* Fort Belvoir, VA: Defense Technical Information Center. <https://doi.org/10.21236/ADA006655>

Lee, M., & Mimno, D. (2014). Low-dimensional Embeddings for Interpretable Anchor-based Topic Inference. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1319–1328). Doha, Qatar: Association for Computational Linguistics. <https://doi.org/10.3115/v1/D14-1138>

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*(3), 211–236. <https://doi.org/10.1080/01638530802073712>

Odell, E. (2018). Hansard Speeches and Sentiment V2.5.1 [dataset]. <https://doi.org/10.5281/zenodo.1306964>

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The Development and Psychometric Properties of LIWC2015, 26. Retrieved from <https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf>

Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A Model of Text for Experimentation in the Social Sciences. *Journal of the American Statistical Association*, *111*(515), 988–1003. <https://doi.org/10.1080/01621459.2016.1141684>

Roberts, M. E., Stewart, B. M., & Tingley, D. (2018). *Stm: R Package for Structural Topic Models*. Retrieved from <http://www.structuraltopicmodel.com>

Yu, B. (2014). Language and gender in Congressional speech. *Literary and Linguistic Computing*, *29*(1), 118–132. <https://doi.org/10.1093/llc/fqs073>

* e.g. a reference to “the member for Bethnal Green and Bow” in keeping with Parliamentary convention of identifying MPs by their seat rather than their name would be followed by “(Rushnara Ali)”.[↩](#fnref1)
* Special Educational Needs[↩](#fnref2)