

X-much in IrE and AusE - Part 5: Statistical Analysis

Anonymous

2024-04-15

This document focuses on the use of the X-much structure in Irish and Australian English.

This part of the analysis performs the data analysis. The data is processed using *tidyverse* package(s) (Wickham et al. 2019) and the statistical analysis uses χ^2 -tests (see Cochran 1952) and configural frequency analysis (see Lienert and Krauth 1975) using the *cfa* package (Mair and Funke 2017). Effect sizes are calculated using the *effectsize* package (Ben-Shachar, Lüdtke, and Makowski 2020).

Session preparation

- install packages

```
install.packages("tidyverse")
install.packages("here")
install.packages("readxl")
install.packages("flextable")
install.packages("quanteda")
install.packages("tidytext")
install.packages("cfa")
install.packages("report")
install.packages("effectsize")
install.packages("epitools")
```

- load packages
- set options

```
# load packages
library(tidyverse)
library(here)
library(readxl)
library(flextable)
library(quanteda)
library(tidytext)
library(cfa)
library(vcd)
library(effectsize)
library(epitools)
# setting options
options(stringsAsFactors = F)
```

Load data

```
xmuch <- base::readRDS(file = here::here("data", "xmuch.rda"))
# inspect
head(xmuch, 10)
```

```
##           docname Variety
## 1 wlp_au_b01_##3282941.1 Australia
## 2 wlp_au_b01_##3287006.1 Australia
## 3 wlp_au_b01_##3289741.1 Australia
## 4 wlp_au_b01_##3291338.1 Australia
## 5 wlp_au_b01_##3293441.1 Australia
## 6 wlp_au_b01_##3297127.1 Australia
## 7 wlp_au_b01_##3304019.1 Australia
## 8 wlp_au_b01_##3304123.1 Australia
## 9 wlp_au_b01_##3304829.1 Australia
## 10 wlp_au_b01_##3306913.1 Australia
##
## 1                                     lessons . Evi 's
## 2      a ring fire with cotton balls and made ' marshmallows on sticks ' cotton balls on pencil ti
## 3                                     the skies will fall if we have a carbon tax . " Skies will fall '
## 4                                     the sleek uniform of the bowl cut
## 5      powder served alongside it too much lately , I even sneak spoonfuls out of the fridge - pregn
plane , 1 cup of nutritional yeast and 1 tsp
## 6      the world a round of psychotherapy to open their eyes to the truth of the world ( um ..
## 7                                     teach . Can you listen ? Can you hear th
## 8 Striesand Effect . Do they think anyone would have seen this piddly little video absent their pro
## 9      was so lovely to see these girls , hear all about Sharon 's trip to France and Italy ( je
## 10      Tweeting Without Getting Sued , which I 'll be buying as soon as it co
##      id          X      POS Status Emotionality Data Polarity
## 1  1          cute Adjective   word    emotional Xmuch positive
## 2  5      creative Adjective   word    emotional Xmuch positive
## 3  7          hype      Noun    word nonemotional Xmuch  neutral
## 4  8      twins      Noun    word nonemotional Xmuch  neutral
## 5 11 pregnancy craving      Noun phrase nonemotional Xmuch  neutral
## 6 17      feasible Adjective   word nonemotional Xmuch  neutral
## 7 26      airy fairy      Noun phrase nonemotional Xmuch  neutral
## 8 27      ironic Adjective   word nonemotional Xmuch  neutral
## 9 29      jealous Adjective   word    emotional Xmuch negative
## 10 32      obsessed Adjective   word    emotional Xmuch negative
```

Inspect structure of the data

```
str(xmuch)
```

```
## 'data.frame':   2142 obs. of  10 variables:
## $ docname      : chr  "wlp_au_b01_##3282941.1" "wlp_au_b01_##3287006.1" "wlp_au_b01_##3289741.1" "wlp_au_b01_##3291338.1" ...
## $ Variety      : chr  "Australia" "Australia" "Australia" "Australia" ...
## $ kwic         : chr  "lessons . Evi 's first day at home and her first collar ( 8 weeks old ) ./." "the sleek uniform of the bowl cut" ...
## $ id           : int  1 5 7 8 11 17 26 27 29 32 ...
## $ X            : chr  "cute" "creative" "hype" "twins" ...
```

```
## $ POS      : chr "Adjective" "Adjective" "Noun" "Noun" ...
## $ Status   : chr "word" "word" "word" "word" ...
## $ Emotionality: chr "emotional" "emotional" "nonemotional" "nonemotional" ...
## $ Data     : chr "Xmuch" "Xmuch" "Xmuch" "Xmuch" ...
## $ Polarity  : chr "positive" "positive" "neutral" "neutral" ...
```

corrections

```
xmuch <- xmuch %>%
  dplyr::mutate(POS = ifelse(X == "prfail", "Noun", POS),
               POS = ifelse(X == "ad hominem", "Noun", POS),
               POS = ifelse(X == "phone hacking", "Verb", POS))
# inspect
str(xmuch)
```

```
## 'data.frame': 2142 obs. of 10 variables:
## $ docname : chr "wlp_au_b01_##3282941.1" "wlp_au_b01_##3287006.1" "wlp_au_b01_##3289741.1" "wlp_au_b01_##3289741.1" ...
## $ Variety : chr "Australia" "Australia" "Australia" "Australia" ...
## $ kwic : chr "lessons . Evi 's first day at home and her first collar ( 8 weeks old ) ./." "lessons . Evi 's first day at home and her first collar ( 8 weeks old ) ./." ...
## $ id : int 1 5 7 8 11 17 26 27 29 32 ...
## $ X : chr "cute" "creative" "hype" "twins" ...
## $ POS : chr "Adjective" "Adjective" "Noun" "Noun" ...
## $ Status : chr "word" "word" "word" "word" ...
## $ Emotionality: chr "emotional" "emotional" "nonemotional" "nonemotional" ...
## $ Data : chr "Xmuch" "Xmuch" "Xmuch" "Xmuch" ...
## $ Polarity : chr "positive" "positive" "neutral" "neutral" ...
```

Statistical Analysis

Single vs Multiword

RQ: do ire and aus differ regarding the status of x (single word vs multiword)?

```
phrasetb <- xmuch %>%
  dplyr::filter(Data == "Xmuch") %>%
  dplyr::select(Variety, Status) %>%
  dplyr::group_by(Variety, Status) %>%
  dplyr::summarise(Frequency = n()) %>%
  dplyr::group_by(Variety) %>%
  dplyr::mutate(Total = sum(Frequency)) %>%
  dplyr::ungroup() %>%
  dplyr::rowwise() %>%
  dplyr::mutate(Percent = round(Frequency/Total*100, 2))
```

```
## `summarise()` has grouped output by 'Variety'. You can override using the
## `.groups` argument.
```

```
# inspect
phrasetb
```

```
## # A tibble: 4 x 5
## # Rowwise:
##   Variety   Status Frequency Total Percent
##   <chr>     <chr>      <int> <int>   <dbl>
## 1 Australia phrase        31   109    28.4
## 2 Australia word         78   109    71.6
## 3 Ireland  phrase        10    33    30.3
## 4 Ireland  word         23    33    69.7
```

```
phrase_x2 <- phrasetb %>%
  dplyr::select(-Total, -Percent) %>%
  tidyr::spread(Status, Frequency) %>%
  as.matrix()
# add rownames
rownames(phrase_x2) <- phrase_x2[, 1]
phrase_x2 <- phrase_x2[, 2:3]
# convert to numeric
phrase_x2 <- t(apply(phrase_x2, 1, function(x){
  x <- as.numeric(x) })))
# add column names
# colnames(phrase_x2) <- names(table(phrasetb$element))
colnames(phrase_x2) <- c("single-word", "multi-word exp.")
# inspect
phrase_x2
```

```
##           single-word multi-word exp.
## Australia           31           78
## Ireland             10           23
```

- perform X2-test (R Core Team 2021)

```
fisher.test(phrase_x2)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: phrase_x2
## p-value = 0.8295
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.3661005 2.4125311
## sample estimates:
## odds ratio
##  0.9146895
```

```
# effect size
effectsize::effectsize(fisher.test(phrase_x2), type = "oddsratio")
```

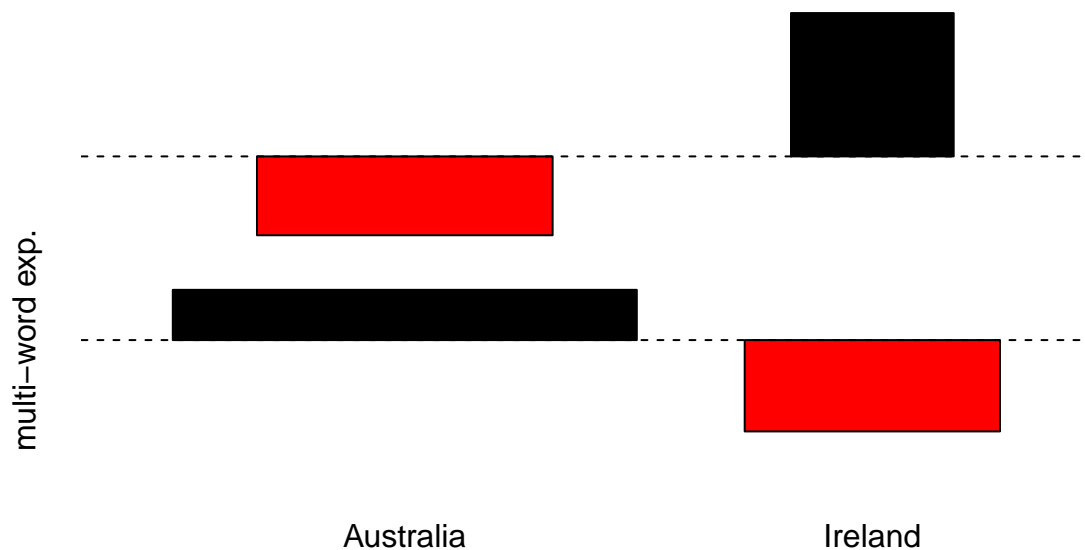
```
## Odds ratio |          95% CI
## -----|-----
## 0.91      | [0.37, 2.41]
```

generate visualization

```
# open window
png(here::here("images", "multiword_assoc.png"), width = 500, height = 450)
# generate plot
assocplot(phrase_x2)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(phrase_x2)
```



Emotionality

tabulation

```
emo_tab <- xmuch %>%
  dplyr::group_by(Variety, Data, Polarity) %>%
  dplyr::summarise(Freq = n()) %>%
  dplyr::group_by(Variety, Data) %>%
  dplyr::mutate(Total = sum(Freq)) %>%
  dplyr::rowwise() %>%
```

```
dplyr::mutate(Percent = round(Freq/Total*100, 1),
              Frequency = paste0(Freq, " (", Percent, ")")) %>%
dplyr::ungroup() %>%
dplyr::select(-Freq, -Total, -Percent) %>%
tidyr::spread(Polarity, Frequency)
```

`summarise()` has grouped output by 'Variety', 'Data'. You can override using
the `.groups` argument.

```
# inspect
emo_tab
```

```
## # A tibble: 4 x 5
##   Variety Data    negative neutral    positive
##   <chr>    <chr>    <chr>    <chr>    <chr>
## 1 Australia control 23 (2.3) 902 (90.2) 75 (7.5)
## 2 Australia Xmuch  51 (46.8) 49 (45)    9 (8.3)
## 3 Ireland  control 25 (2.5) 919 (91.9) 56 (5.6)
## 4 Ireland  Xmuch  14 (42.4) 18 (54.5)  1 (3)
```

Calculate totals

```
# au
austot <- 23+902+75+51+49+9
negau <- round(74/austot*100, 1)
nonau <- round(951/austot*100, 1)
posau <- round(84/austot*100, 1)
# ire
iretot <- 25+919+56+14+18+1
negire <- round(40/iretot*100, 1)
nonire <- round(945/iretot*100, 1)
posire <- round(48/iretot*100, 1)
# results
negau
```

```
## [1] 6.7
```

```
nonau
```

```
## [1] 85.8
```

```
posau
```

```
## [1] 7.6
```

```
negire
```

```
## [1] 3.9
```

```
nonire
```

```
## [1] 91.5
```

```
posire
```

```
## [1] 4.6
```

```
xmdtb <- xmuch %>%  
  dplyr::select(Variety, X, POS, Emotionality, Polarity, Status, Data) %>%  
  dplyr::rename(word = X)  
# inspect  
head(xmdtb)
```

```
##      Variety      word      POS Emotionality Polarity Status  Data  
## 1 Australia      cute Adjective      emotional positive word Xmuch  
## 2 Australia    creative Adjective      emotional positive word Xmuch  
## 3 Australia      hype      Noun nonemotional neutral word Xmuch  
## 4 Australia     twins      Noun nonemotional neutral word Xmuch  
## 5 Australia pregnancy craving      Noun nonemotional neutral phrase Xmuch  
## 6 Australia     feasible Adjective nonemotional neutral word Xmuch
```

- tabulation

```
tb1 <-base::readRDS(file = here::here("tables", "tb1.rda")) %>%  
  dplyr::mutate(Variety = stringr::str_replace_all(text, ".*_([a-z]{2,2})_.*", "\\1"),  
               Variety = ifelse(Variety == "au", "Australia", "Ireland")) %>%  
  dplyr::group_by(Variety) %>%  
  dplyr::summarise(words = sum(words),  
                  texts = n())  
# inspect  
head(tb1)
```

```
## # A tibble: 2 x 3  
##   Variety      words  texts  
##   <chr>      <int> <int>  
## 1 Australia 152077983 129382  
## 2 Ireland  103259885 102426
```

```
xmdtb %>%  
  dplyr::group_by(Variety, Status) %>%  
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Variety'. You can override using the  
## `.groups` argument.
```

```
## # A tibble: 4 x 3  
## # Groups:   Variety [2]  
##   Variety      Status Frequency  
##   <chr>      <chr>      <int>  
## 1 Australia phrase         31  
## 2 Australia word        1078  
## 3 Ireland  phrase         10  
## 4 Ireland  word        1023
```

X2 (wo variety)

RQ: Does the emotionality of words in the test data (in the x-much construction) differ from the emotionality in the control data?

```
emo_tb1 <- xmdtb %>%  
  dplyr::group_by(Emotionality, Data) %>%  
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Emotionality'. You can override using the  
## `.groups` argument.
```

```
# inspect  
head(emo_tb1)
```

```
## # A tibble: 4 x 3  
## # Groups:   Emotionality [2]  
##   Emotionality Data      Frequency  
##   <chr>          <chr>         <int>  
## 1 emotional     Xmuch             75  
## 2 emotional     control          179  
## 3 nonemotional Xmuch             67  
## 4 nonemotional control        1821
```

Perform x2-test

```
emo_x2 <- emo_tb1 %>%  
  tidyr::spread(Data, Frequency) %>%  
  as.matrix()  
rn<- emo_x2[,1]  
emo_x2 = emo_x2[, 2:3]  
emo_x2 <- apply(emo_x2, 2, as.numeric)  
rownames(emo_x2) <- rn  
emo_x2
```

```
##           control Xmuch  
## emotional      179    75  
## nonemotional   1821    67
```

```
# perform x2 test  
fisher.test(emo_x2)
```

```
##  
## Fisher's Exact Test for Count Data  
##  
## data:  emo_x2  
## p-value < 2.2e-16  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
##  0.06006725 0.12855624  
## sample estimates:  
## odds ratio  
##  0.0880061
```



```
# effect size
effectsize::effectsize(fisher.test(emo_x2), type = "oddsratio")
```

```
## Odds ratio |      95% CI
## -----
## 0.09       | [0.06, 0.13]
```

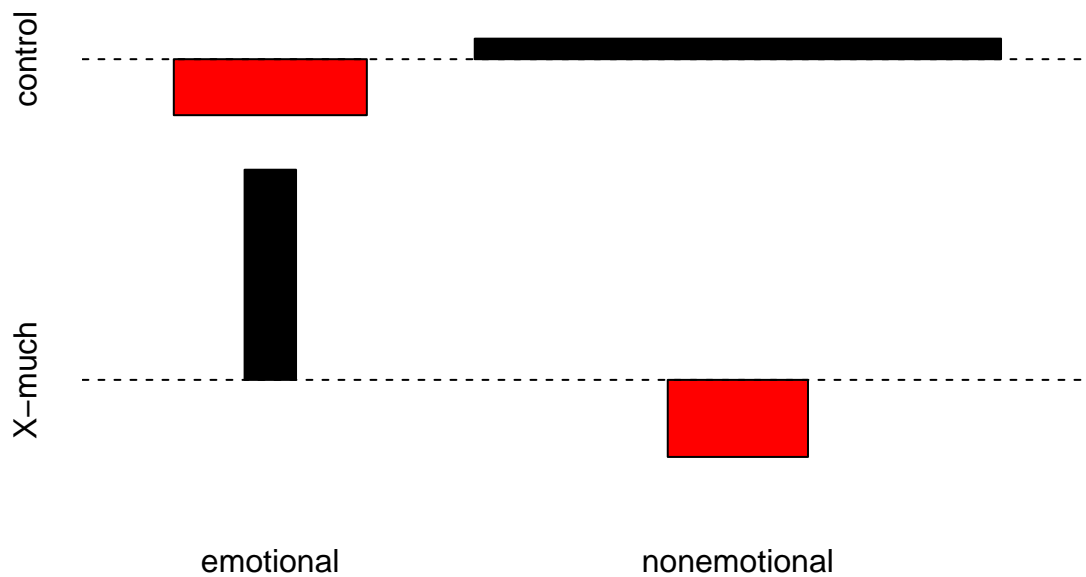
The words in the x-much structure are significantly more emotional compared to the words in the control data.

```
# adapt column names
colnames(emo_x2) <- c("control", "X-much")

# open window
png(here::here("images", "emo_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(emo_x2)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(emo_x2)
```



X2 (vareity)

RQ: Does the emotionality of words in the X-much structure in Australia differ from the emotionality of words in the X-much structure in Ireland?

```
emo_tb2 <- xmuch %>%
  dplyr::filter(Data == "Xmuch") %>%
  dplyr::group_by(Emotionality, Variety) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Emotionality'. You can override using the
## `.groups` argument.
```

```
# inspect
head(emo_tb2)
```

```
## # A tibble: 4 x 3
## # Groups:   Emotionality [2]
##   Emotionality Variety   Frequency
##   <chr>          <chr>         <int>
## 1 emotional      Australia         60
## 2 emotional      Ireland          15
## 3 nonemotional   Australia         49
## 4 nonemotional   Ireland           18
```

Perform x2-test

```
emo_x22 <- emo_tb2 %>%
  tidyr::spread(Variety, Frequency) %>%
  as.matrix()
rn<- emo_x22[,1]
emo_x22 = emo_x22[, 2:3]
emo_x22 <- apply(emo_x22, 2, as.numeric)
rownames(emo_x22) <- rn
emo_x22
```

```
##           Australia Ireland
## emotional          60      15
## nonemotional       49      18
```

```
# perform x2 test
fisher.test(emo_x22)
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  emo_x22
## p-value = 0.4264
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.6250054 3.4794174
## sample estimates:
## odds ratio
##  1.465359
```

```
# effect size
effectsize::effectsize(fisher.test(emo_x22), type = "oddsratio")
```

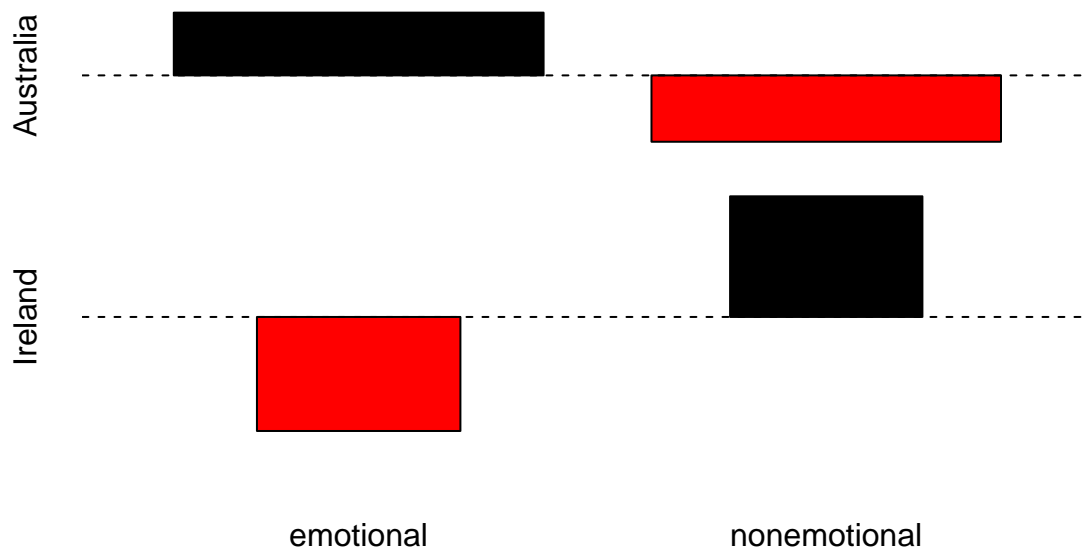
```
## Odds ratio |      95% CI
## -----
## 1.47      | [0.63, 3.48]
```

Australia and Ireland differ significantly with respect to the emotionality of words in the x-much structure.

```
# open window
png(here::here("images", "emo_var_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(emo_x22)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(emo_x22)
```



CFA

```
emo_tb3 <- xmdtb %>%
  dplyr::group_by(Emotionality, Variety, Data) %>%
  dplyr::summarise(Frequency = n())
```

`summarise()` has grouped output by 'Emotionality', 'Variety'. You can override
using the `.groups` argument.

```
# inspect
emo_tb3
```

```
## # A tibble: 8 x 4
## # Groups:   Emotionality, Variety [4]
##   Emotionality Variety   Data   Frequency
##   <chr>          <chr>   <chr>     <int>
## 1 emotional      Australia Xmuch        60
## 2 emotional      Australia control     98
## 3 emotional      Ireland   Xmuch        15
## 4 emotional      Ireland   control     81
## 5 nonemotional   Australia Xmuch        49
## 6 nonemotional   Australia control    902
## 7 nonemotional   Ireland   Xmuch        18
## 8 nonemotional   Ireland   control    919
```

```
configs <- emo_tb3 %>%
  dplyr::select(Variety, Emotionality, Data)
counts = emo_tb3$Frequency
cfa::cfa(configs, counts)
```

```
##
## *** Analysis of configuration frequencies (CFA) ***
##
##               label    n  expected      Q      chisq
## 1   Australia emotional Xmuch  60   8.717956 0.024039036 301.658786
## 2   Ireland nonemotional Xmuch  18  60.360346 0.020349510  29.728109
## 3   Ireland emotional control  81 114.373418 0.016459351   9.738146
## 4   Ireland emotional Xmuch  15   8.120513 0.003223934   5.828123
## 5   Ireland nonemotional control 919 850.145723 0.053298796   5.576587
## 6   Australia emotional control  98 122.788113 0.012276133   5.004153
## 7   Australia nonemotional Xmuch  49  64.801185 0.007606968   3.852977
## 8   Australia nonemotional control 902 912.692746 0.008698188   0.125272
##
##      p.chisq sig.chisq      z      p.z sig.z
## 1 0.000000e+00      TRUE 17.2340931 0.00000000 TRUE
## 2 4.970880e-08      TRUE -5.5961168 0.99999999 TRUE
## 3 1.804828e-03      TRUE -3.2554589 0.99943395 TRUE
## 4 1.577194e-02     FALSE  2.2429461 0.01245014 FALSE
## 5 1.820216e-02     FALSE  3.0187141 0.00126925 TRUE
## 6 2.528657e-02     FALSE -2.3504826 0.99062546 FALSE
## 7 4.965779e-02     FALSE -2.0563565 0.98012592 FALSE
## 8 7.233855e-01     FALSE -0.4890507 0.68759709 FALSE
##
##
```

```
## Summary statistics:
##
## Total Chi squared      = 361.5122
## Total degrees of freedom = 4
## p                      = 0
## Sum of counts          = 2142
##
## Levels:
##
##      Variety Emotionality      Data
##      2          2          2
```

```
# save
cfaemo <- cfa::cfa(configs, counts)
cfaemo <- cfaemo$table %>%
  as.data.frame() %>%
  dplyr::filter(sig.z == TRUE)
write.table(cfaemo, here::here("tables", "cfaemo.txt"), sep = "\t", row.names = F)
```

There are significantly more negative adjectives in the x-much construction in the Australian data after corrections compared to any other configuration.

Visualization

- prepare data

```
emo_vis <- emo_tb3 %>%
  dplyr::ungroup() %>%
  dplyr::mutate_if(is.character, factor) %>%
  dplyr::mutate(Data = ifelse(Data == "Xmuch", "X-much", "control"))
# inspect
head(emo_vis)
```

```
## # A tibble: 6 x 4
##   Emotionality Variety   Data   Frequency
##   <fct>         <fct>   <chr>     <int>
## 1 emotional    Australia X-much      60
## 2 emotional    Australia control    98
## 3 emotional    Ireland   X-much     15
## 4 emotional    Ireland   control    81
## 5 nonemotional Australia X-much     49
## 6 nonemotional Australia control   902
```

```
emo1 <- emo_vis %>%
  dplyr::filter(Emotionality == "emotional") %>%
  dplyr::pull()
emo2 <- emo_vis %>%
  dplyr::filter(Emotionality == "nonemotional") %>%
  dplyr::pull()
# add dimnames
column.names <- c("Australia", "Ireland")
row.names <- c("control", "X-much")
```

```

matrix.names <- c("emotional", "nonemotional")
# generate matrix
emo_mx <- array(c(emo1, emo2), dim = c(2, 2, 2),
               dimnames = list(row.names,
                               column.names,
                               matrix.names))

# inspect
emo_mx

```

```

## , , emotional
##
##      Australia Ireland
## control      60      15
## X-much       98      81
##
## , , nonemotional
##
##      Australia Ireland
## control      49      18
## X-much      902     919

```

Generate mosaic plot

```

# open connection
png(here::here("images", "emo_mosaic.png"), width = 750, height = 300)
# generate plot
mosaic(emo_mx,
       axis.cex = 15,
       shade = TRUE,
       direction = c("h", "v", "v"),
       just_labels = c("center", "center", "center", "center"))
# close window
dev.off()

```

```

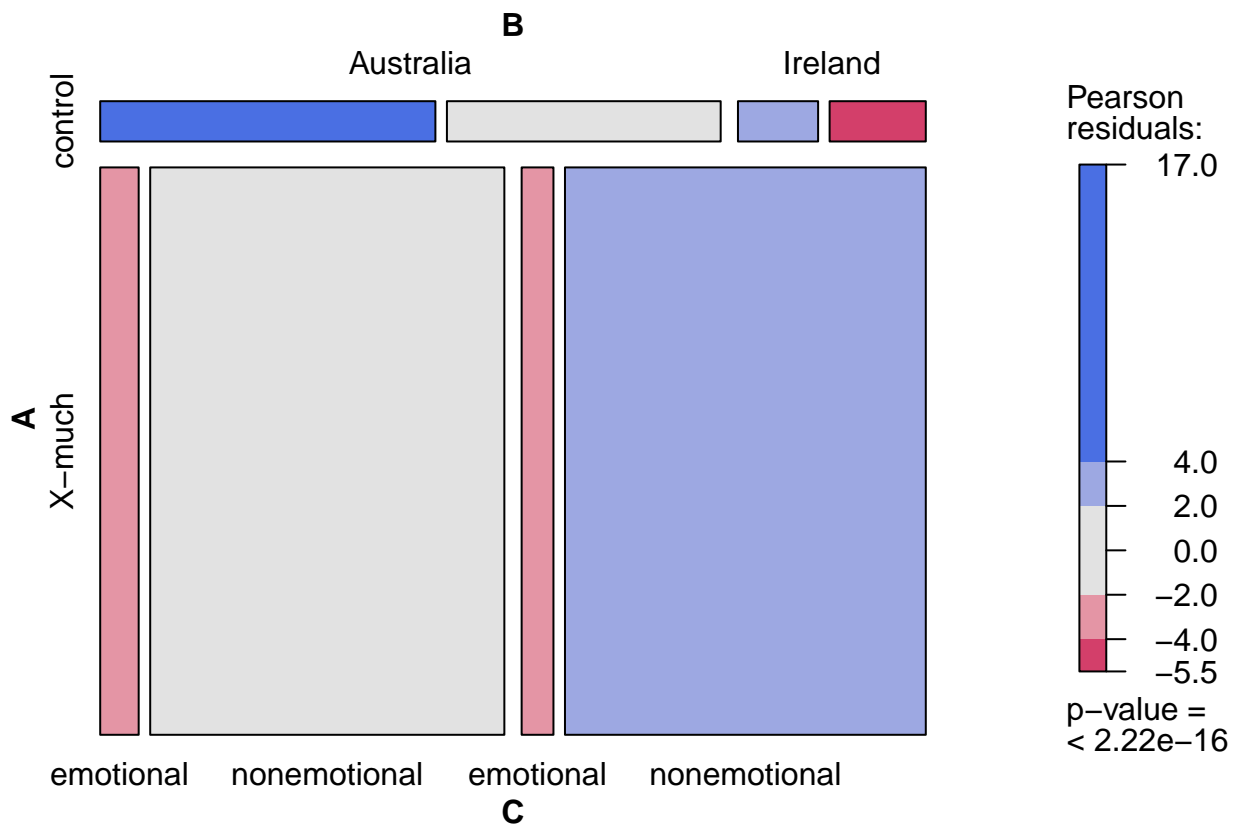
## pdf
## 2

```

```

# show plot
mosaic(emo_mx,
       shade = TRUE,
       direction = c("h", "v", "v"),
       just_labels = c("center", "center", "center", "center"))

```



Polarity

X2 (wo variety)

RQ: Does the polarity of words in the x-much structure differ from the polarity in the control data?

Prepare data

```
pol_tb1 <- xmuch %>%
  dplyr::group_by(Polarity, Data) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Polarity'. You can override using the
## `.groups` argument.
```

```
# inspect
pol_tb1
```

```
## # A tibble: 6 x 3
## # Groups:   Polarity [3]
##   Polarity Data    Frequency
##   <chr>    <chr>         <int>
## 1 negative Xmuch          65
## 2 negative control       48
```

```
## 3 neutral Xmuch 67
## 4 neutral control 1821
## 5 positive Xmuch 10
## 6 positive control 131
```

Perform x2-test

```
pol_x2 <- pol_tb1 %>%
  tidyr::spread(Data, Frequency) %>%
  as.matrix()
# inspect
pol_x2
```

```
##      Polarity control Xmuch
## [1,] "negative" " 48" "65"
## [2,] "neutral" "1821" "67"
## [3,] "positive" " 131" "10"
```

```
rn <- pol_x2[,1]
pol_x2 = pol_x2[, 2:3]
pol_x2 <- apply(pol_x2, 2, as.numeric)
rownames(pol_x2) <- rn
pol_x2
```

```
##      control Xmuch
## negative    48    65
## neutral    1821    67
## positive    131    10
```

```
# perform x2 test
fisher.test(pol_x2)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: pol_x2
## p-value < 2.2e-16
## alternative hypothesis: two.sided
```

```
# effect size
effectsize::effectsize(fisher.test(pol_x2))
```

```
## Cramer's V (adj.) |      95% CI
## -----
## 0.48              | [0.44, 0.53]
```

The polarity of words in the X-much structure is significantly different from the polarity of words in the control data.


```

# adapt column names
colnames(pol_x2) <- c("control", "X-much")

# open window
png(here::here("images", "pol_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(pol_x2)
# close window
dev.off()

```

```

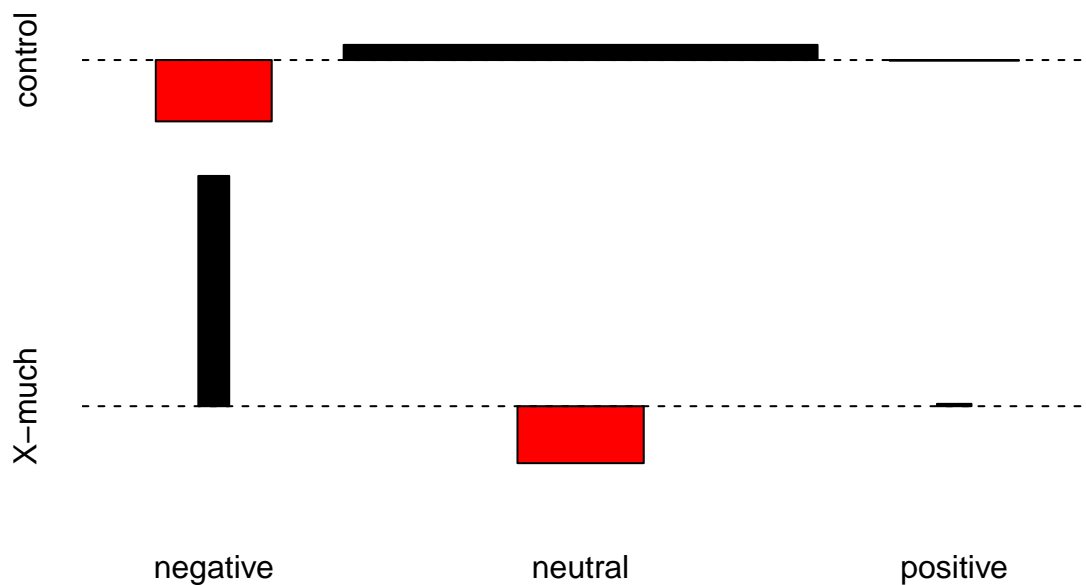
## pdf
## 2

```

```

assocplot(pol_x2)

```



X2 (wo variety)

RQ: Does the polarity of words in the x-much structure in Australia differ from the polarity of words in the x-much structure in Ireland?

Prepare data

```

pol_tb2 <- xmuch %>%
  dplyr::filter(Data == "Xmuch") %>%

```

```
dplyr::group_by(Polarity, Variety) %>%
dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Polarity'. You can override using the
## `.groups` argument.
```

```
# inspect
pol_tb2
```

```
## # A tibble: 6 x 3
## # Groups:   Polarity [3]
##   Polarity Variety   Frequency
##   <chr>    <chr>         <int>
## 1 negative Australia     51
## 2 negative Ireland      14
## 3 neutral  Australia     49
## 4 neutral  Ireland      18
## 5 positive Australia      9
## 6 positive Ireland       1
```

Perform x2-test

```
pol_x22 <- pol_tb2 %>%
  tidyr::spread(Variety, Frequency) %>%
  as.matrix()
# inspect
pol_x22
```

```
##      Polarity  Australia Ireland
## [1,] "negative" "51"      "14"
## [2,] "neutral"  "49"      "18"
## [3,] "positive" " 9"      " 1"
```

```
rn <- pol_x22[,1]
pol_x22 = pol_x22[, 2:3]
pol_x22 <- apply(pol_x22, 2, as.numeric)
rownames(pol_x22) <- rn
pol_x22
```

```
##      Australia Ireland
## negative      51      14
## neutral       49      18
## positive       9       1
```

```
# perform x2 test
fisher.test(pol_x22)
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  pol_x22
## p-value = 0.5498
## alternative hypothesis: two.sided
```

```
# effect size
effectsize::effectsize(fisher.test(pol_x22))
```

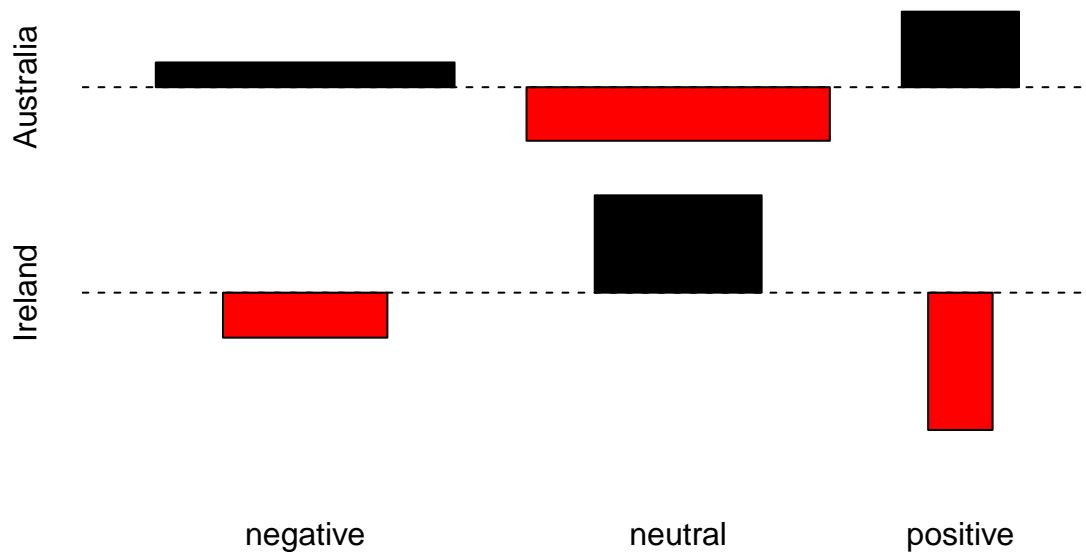
```
## Cramer's V (adj.) |          95% CI
## -----
## 0.00              | [0.00, 0.22]
```

Australia and Ireland do **not** differ significantly with respect to the polarity of the words in the X-much structure.

```
# open window
png(here::here("images", "pol_var_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(pol_x22)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(pol_x22)
```



CFA

```
pol_tbt1 <- xmuch %>%
  dplyr::group_by(Polarity, Variety, Data) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'Polarity', 'Variety'. You can override
## using the `.groups` argument.
```

```
# inspect
pol_tbt1
```

```
## # A tibble: 12 x 4
## # Groups:   Polarity, Variety [6]
##   Polarity Variety Data Frequency
##   <chr>    <chr>   <chr>      <int>
## 1 negative Australia Xmuch         51
## 2 negative Australia control        23
## 3 negative Ireland Xmuch         14
## 4 negative Ireland control        25
## 5 neutral Australia Xmuch         49
## 6 neutral Australia control       902
## 7 neutral Ireland Xmuch         18
## 8 neutral Ireland control       919
## 9 positive Australia Xmuch          9
## 10 positive Australia control        75
## 11 positive Ireland Xmuch          1
## 12 positive Ireland control        56
```

```
configs <- pol_tbt1 %>%
  dplyr::select(Variety, Polarity, Data)
counts = pol_tbt1$Frequency
cfa::cfa(configs, counts)
```

```
##
## *** Analysis of configuration frequencies (CFA) ***
##
##           label    n  expected      Q      chisq
## 1  Australia negative Xmuch  51   3.878461 0.022038756 572.5053321
## 2    Ireland negative Xmuch  14   3.612669 0.004857554 29.8661858
## 3    Ireland neutral Xmuch  18  60.360346 0.020349510 29.7281088
## 4 Australia negative control  23  54.626208 0.015151195 18.3102042
## 5    Ireland negative control  25  50.882662 0.012377432 13.1658247
## 6    Ireland neutral control 919 850.145723 0.053298796  5.5765869
## 7  Australia neutral Xmuch  49  64.801185 0.007606968  3.8529766
## 8  Australia positive Xmuch   9   4.839495 0.001946744  3.5767779
## 9    Ireland positive Xmuch   1   4.507844 0.001641102  2.7296792
## 10 Ireland positive control  56  63.490756 0.003603908  0.8837731
## 11 Australia positive control  75  68.161905 0.003297314  0.6860069
## 12 Australia neutral control 902 912.692746 0.008698188  0.1252720
##           p.chisq sig.chisq      z      p.z sig.z
```

```
## 1 0.000000e+00      TRUE 23.9270837 0.000000e+00      TRUE
## 2 3.270696e-07      TRUE  5.4649964 2.314581e-08      TRUE
## 3 3.504477e-07      TRUE  5.4523489 2.485440e-08      TRUE
## 4 1.056792e-04      TRUE  4.2790424 9.384954e-06      TRUE
## 5 1.383813e-03      TRUE  3.6284742 1.425506e-04      TRUE
## 6 6.152612e-02     FALSE  2.3614798 9.101082e-03     FALSE
## 7 1.456588e-01     FALSE  1.9629000 2.482890e-02     FALSE
## 8 1.672294e-01     FALSE  1.8912371 2.929635e-02     FALSE
## 9 2.554216e-01     FALSE  1.6521741 4.924953e-02     FALSE
## 10 6.428226e-01     FALSE  0.9400921 1.735852e-01     FALSE
## 11 7.096358e-01     FALSE  0.8282554 2.037630e-01     FALSE
## 12 9.392853e-01     FALSE  0.3539378 3.616927e-01     FALSE
```

```
##
```

```
##
```

```
## Summary statistics:
```

```
##
```

```
## Total Chi squared      = 681.0067
```

```
## Total degrees of freedom = 7
```

```
## p                      = 0
```

```
## Sum of counts          = 2142
```

```
##
```

```
## Levels:
```

```
##
```

```
## Variety Polarity      Data
```

```
##          2          3          2
```

```
# save
```

```
cfapol <- cfa::cfa(configs, counts)
```

```
cfapol <- cfapol$table %>%
```

```
  as.data.frame() %>%
```

```
  dplyr::filter(sig.z == TRUE)
```

```
write.table(cfapol, here::here("tables", "cfapol.txt"), sep = "\t", row.names = F)
```

Visualization

- prepare data

```
pol_vis <- pol_tb1 %>%
```

```
  dplyr::ungroup() %>%
```

```
  dplyr::mutate_if(is.character, factor)
```

```
# inspect
```

```
pol_vis
```

```
## # A tibble: 12 x 4
##   Polarity Variety      Data      Frequency
##   <fct>      <fct>      <fct>          <int>
## 1 negative Australia Xmuch           51
## 2 negative Australia control        23
## 3 negative Ireland  Xmuch           14
## 4 negative Ireland  control        25
## 5 neutral  Australia Xmuch           49
## 6 neutral  Australia control        902
```

```
## 7 neutral Ireland Xmuch 18
## 8 neutral Ireland control 919
## 9 positive Australia Xmuch 9
## 10 positive Australia control 75
## 11 positive Ireland Xmuch 1
## 12 positive Ireland control 56
```

- convert to matrix

```
pol1 <- pol_vis %>%
  dplyr::filter(Polarity == "positive") %>%
  dplyr::pull(Frequency)
pol2 <- pol_vis %>%
  dplyr::filter(Polarity == "neutral") %>%
  dplyr::pull(Frequency)
pol3 <- pol_vis %>%
  dplyr::filter(Polarity == "negative") %>%
  dplyr::pull(Frequency)
# add dimnames
column.names <- c("Australia", "Ireland")
row.names <- c("control", "X-much")
matrix.names <- c("positive", "neutral", "negative")
# generate matrix
pol_mx <- array(c(pol1, pol2, pol3), dim = c(2, 2, 3),
               dimnames = list(row.names,
                               column.names,
                               matrix.names))

# inspect
pol_mx
```

```
## , , positive
##
##      Australia Ireland
## control      9      1
## X-much      75     56
##
## , , neutral
##
##      Australia Ireland
## control     49     18
## X-much     902    919
##
## , , negative
##
##      Australia Ireland
## control     51     14
## X-much     23     25
```

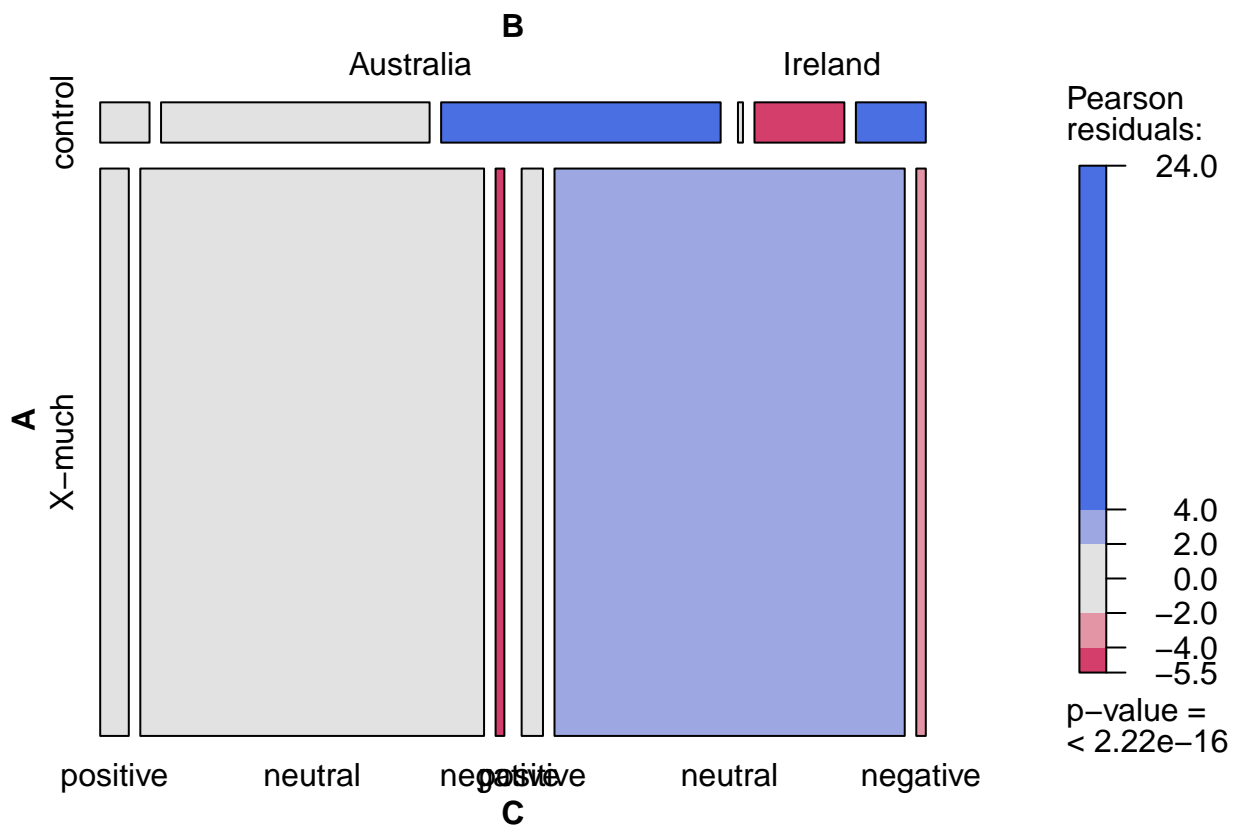
- mosaic plot

```
# open window
png(here::here("images", "pol_mosaic.png"), width = 750, height = 300)
```

```
# generate plot
mosaic(pol_mx,
  shade = TRUE,
  direction = c("h", "v", "v"),
  just_labels = c("center", "center", "center", "center"))
# close window
dev.off()
```

```
## pdf
## 2
```

```
# show plot
mosaic(pol_mx,
  shade = TRUE,
  direction = c("h", "v", "v"),
  just_labels = c("center", "center", "center", "center"))
```



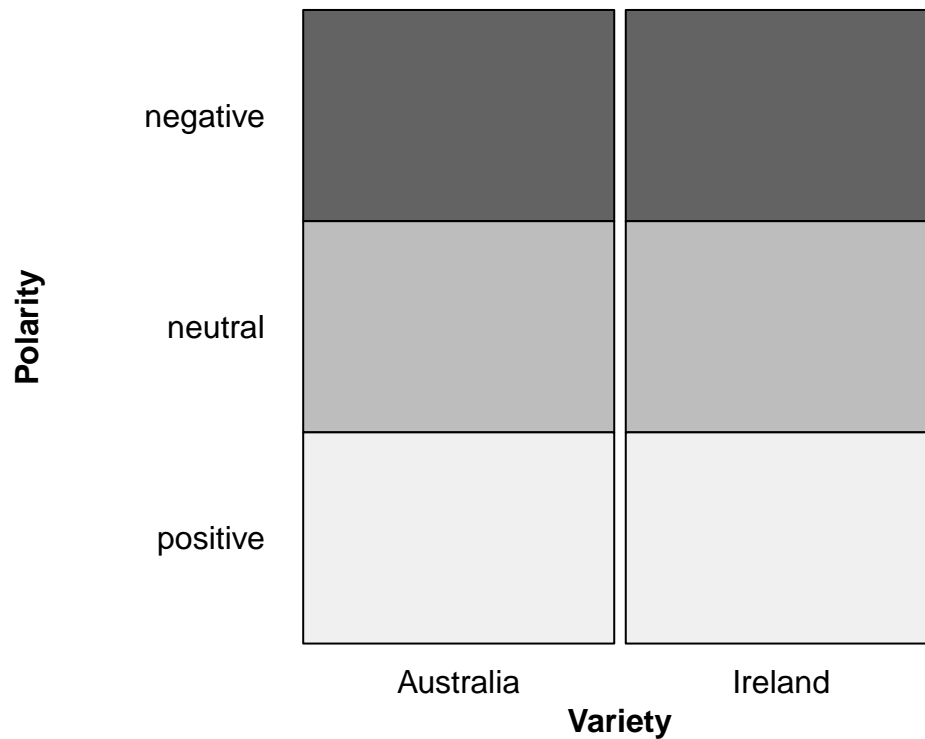
- alternative mosaic plot

```
# define cols
mycols <- rev(RColorBrewer::brewer.pal(3, "Greys"))
vcd::mosaic(Polarity ~ Variety, data = pol_vis,
  direction = c("v", "h"),
  rot_labels = c(0, 0, 0, 0),
```

```

highlighting_fill = mycols,
labeling_args = list(tl_labels = c(F, T),
                      set_varnames = c(Polarity = "Polarity\n\n\n\n\n\n\n"),
                      offset_varnames = c(0, 0, 0, 1),
                      just_labels = c("center", "center", "center", "right")))

```



```

# save plot
#ggsave(here::here("images", "pol_var.png"), units = "cm", width = 8, height = 6)

```

POS

tabulation

```

pos_tab <- xmuch %>%
  dplyr::group_by(Variety, Data, POS) %>%
  dplyr::summarise(Freq= n()) %>%
  dplyr::group_by(Variety, Data) %>%
  dplyr::mutate(Total = sum(Freq)) %>%
  dplyr::rowwise() %>%
  dplyr::mutate(Percent = round(Freq/Total*100, 1),
                Frequency = paste0(Freq, " (", Percent, "%)") %>%
  dplyr::ungroup() %>%
  dplyr::select(-Freq, -Total, -Percent) %>%
  tidyr::spread(POS, Frequency)

```



```
## `summarise()` has grouped output by 'Variety', 'Data'. You can override using
## the `.groups` argument.
```

```
# inspect
pos_tab
```

```
## # A tibble: 4 x 6
##   Variety Data   Adjective Noun      other      Verb
##   <chr>   <chr>   <chr>   <chr>   <chr>   <chr>
## 1 Australia control 79 (7.9) 313 (31.3) 537 (53.7) 71 (7.1)
## 2 Australia Xmuch 45 (41.3) 47 (43.1) <NA>      17 (15.6)
## 3 Ireland control 82 (8.2) 327 (32.7) 517 (51.7) 74 (7.4)
## 4 Ireland Xmuch 8 (24.2) 24 (72.7) <NA>      1 (3)
```

RQ: Do the word classes (parts-of-speech) of words in the x-much structure differ from the word classes in the control data?

X2 (wo variety)

Prepare data

```
pos_tb1 <- xmuch %>%
  dplyr::group_by(POS, Data) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'POS'. You can override using the `.groups`
## argument.
```

```
# inspect
pos_tb1
```

```
## # A tibble: 7 x 3
## # Groups:   POS [4]
##   POS      Data Frequency
##   <chr>   <chr>   <int>
## 1 Adjective Xmuch      53
## 2 Adjective control   161
## 3 Noun      Xmuch      71
## 4 Noun      control   640
## 5 Verb      Xmuch      18
## 6 Verb      control   145
## 7 other     control  1054
```

Perform x2-test

```
pos_x2 <- pos_tb1 %>%
  tidyr::spread(Data, Frequency) %>%
  as.matrix()
# inspect
pos_x2
```

```
##      POS      control Xmuch
## [1,] "Adjective" " 161" "53"
## [2,] "Noun"      " 640" "71"
## [3,] "other"     "1054" NA
## [4,] "Verb"      " 145" "18"
```

```
rn <- pos_x2[,1]
pos_x2 = pos_x2[, 2:3]
pos_x2 <- apply(pos_x2, 2, as.numeric)
rownames(pos_x2) <- rn
pos_x2[3, 2] <- 0
pos_x2
```

```
##      control Xmuch
## Adjective    161    53
## Noun         640    71
## other       1054     0
## Verb        145    18
```

```
# perform x2 test
fisher.test(pos_x2)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: pos_x2
## p-value < 2.2e-16
## alternative hypothesis: two.sided
```

```
# effect size
effectsize::effectsize(fisher.test(pos_x2))
```

```
## Cramer's V (adj.) |      95% CI
## -----
## 0.31              | [0.26, 0.35]
```

The word classes (parts-of-speech) of words in the x-much structure differ significantly from the word classes in the control data!

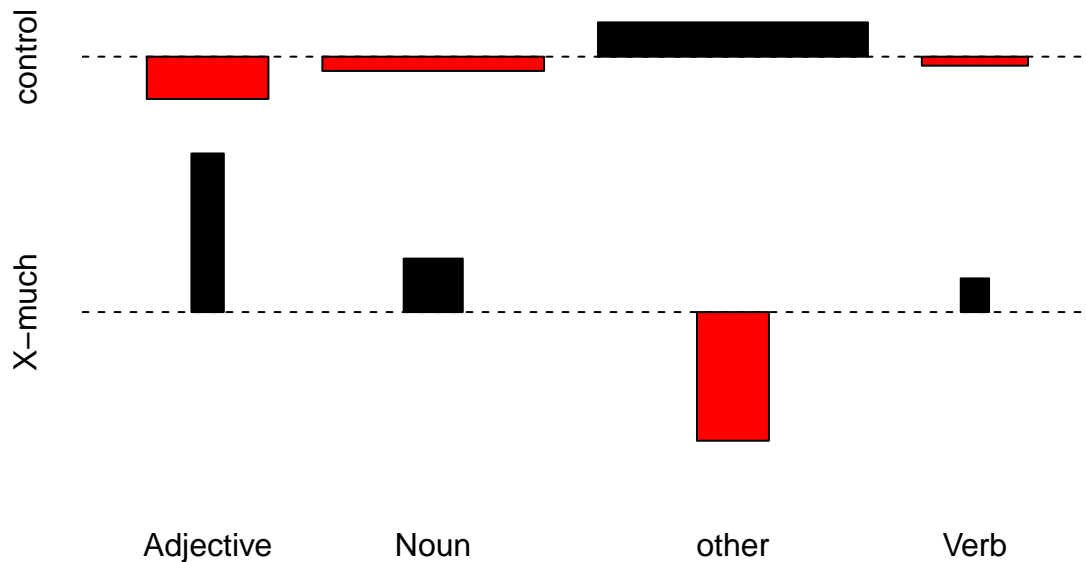
generate association plot

```
# adapt column names
colnames(pos_x2) <- c("control", "X-much")

# open window
png(here::here("images", "pos_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(pos_x2)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(pos_x2)
```



X2 (variety)

RQ: Do Australia and Ireland differ with respect to the word classes (parts-of-speech) of words in the x-much structure?

Prepare data

```
pos_tb2 <- xmuch %>%  
  dplyr::filter(Data == "Xmuch") %>%  
  dplyr::group_by(POS, Variety) %>%  
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'POS'. You can override using the `.groups`  
## argument.
```

```
# inspect  
pos_tb2
```

```
## # A tibble: 6 x 3  
## # Groups:   POS [3]  
##   POS      Variety  Frequency
```

```
##    <chr>      <chr>      <int>
## 1 Adjective Australia    45
## 2 Adjective Ireland      8
## 3 Noun        Australia   47
## 4 Noun        Ireland    24
## 5 Verb        Australia   17
## 6 Verb        Ireland     1
```

Perform x2-test

```
pos_x22 <- pos_tb2 %>%
  tidyr::spread(Variety, Frequency) %>%
  as.matrix()
# inspect
pos_x22
```

```
##      POS      Australia Ireland
## [1,] "Adjective" "45"        " 8"
## [2,] "Noun"      "47"        "24"
## [3,] "Verb"      "17"        " 1"
```

```
rn <- pos_x22[,1]
pos_x22 = pos_x22[, 2:3]
pos_x22 <- apply(pos_x22, 2, as.numeric)
rownames(pos_x22) <- rn
pos_x22
```

```
##      Australia Ireland
## Adjective      45      8
## Noun           47     24
## Verb          17      1
```

```
# perform x2 test
fisher.test(pos_x22)
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  pos_x22
## p-value = 0.008735
## alternative hypothesis: two.sided
```

```
# effect size
effectsize::effectsize(fisher.test(pos_x22))
```

```
## Cramer's V (adj.) |      95% CI
## -----
## 0.23              | [0.00, 0.40]
```

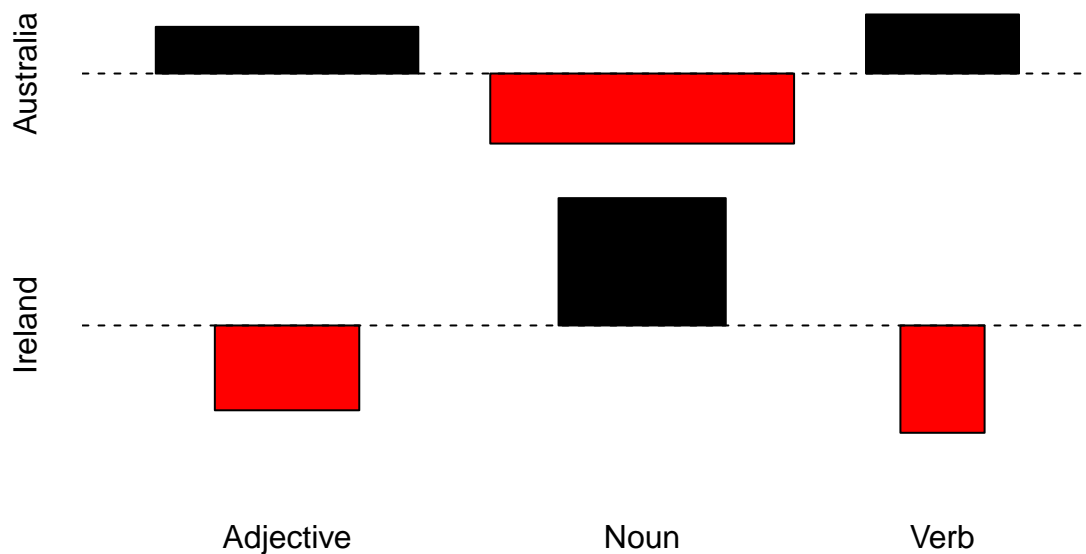
Australia and Ireland differ significantly with respect to the word classes (parts-of-speech) of words in the x-much structure!

generate association plot

```
# open window
png(here::here("images", "pos_var_assoc.png"), width = 500, height = 350)
# generate plot
assocplot(pos_x22)
# close window
dev.off()
```

```
## pdf
## 2
```

```
assocplot(pos_x22)
```



CFA

process data

```
pos_tb2 <- xmuch %>%
  dplyr::group_by(POS, Variety, Data) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'POS', 'Variety'. You can override using
## the `.groups` argument.
```

```
# inspect
pos_tb2
```

```
## # A tibble: 14 x 4
## # Groups:   POS, Variety [8]
##   POS      Variety Data      Frequency
##   <chr>    <chr>   <chr>         <int>
## 1 Adjective Australia Xmuch          45
## 2 Adjective Australia control       79
## 3 Adjective Ireland  Xmuch           8
## 4 Adjective Ireland  control       82
## 5 Noun      Australia Xmuch          47
## 6 Noun      Australia control      313
## 7 Noun      Ireland  Xmuch          24
## 8 Noun      Ireland  control      327
## 9 Verb      Australia Xmuch          17
## 10 Verb     Australia control       71
## 11 Verb     Ireland  Xmuch           1
## 12 Verb     Ireland  control       74
## 13 other    Australia control      537
## 14 other    Ireland  control      517
```

perform CFA

```
configs <- pos_tb2 %>%
  dplyr::select(Variety, POS, Data)
counts = pos_tb2$Frequency
cfa::cfa(configs, counts)
```

```
##
## *** Analysis of configuration frequencies (CFA) ***
##
##           label    n  expected      Q      chisq
## 1  Australia Adjective Xmuch  45  7.345050 0.0176398300 1.930409e+02
## 2      Australia Verb Xmuch  17  5.594594 0.0053385964 2.325161e+01
## 3      Australia Noun Xmuch  47 24.403412 0.0106708652 2.092354e+01
## 4 Australia Adjective control  79 103.451402 0.0119945153 5.779246e+00
## 5      Ireland other control 517 474.604657 0.0254260896 3.787079e+00
## 6      Ireland Verb Xmuch   1  5.211195 0.0019708055 3.403090e+00
## 7      Australia Noun control 313 343.710033 0.0170773533 2.743900e+00
## 8      Ireland Adjective control  82 96.361856 0.0070207218 2.140504e+00
## 9      Australia other control 537 509.522327 0.0168318826 1.481824e+00
## 10      Australia Verb control  71 78.797096 0.0037791223 7.715349e-01
## 11      Ireland Adjective Xmuch   8  6.841692 0.0005424929 1.961032e-01
## 12      Ireland Noun control 327 320.155513 0.0037568996 1.463258e-01
## 13      Ireland Noun Xmuch  24 22.731041 0.0005987718 7.083951e-02
## 14      Ireland Verb control  74 73.397115 0.0002914456 4.952109e-03
##
##      p.chisq sig.chisq      z      p.z sig.z
## 1  0.000000e+00      TRUE 13.89391747 0.000000e+00      TRUE
## 2  3.578855e-05      TRUE  4.82199188 7.106587e-07      TRUE
## 3  1.091962e-04      TRUE  4.57422561 2.389923e-06      TRUE
## 4  1.228585e-01     FALSE  2.40400619 8.108249e-03     FALSE
```

```
## 5 2.853927e-01 FALSE 1.94604178 2.582486e-02 FALSE
## 6 3.335503e-01 FALSE 1.84474652 3.253721e-02 FALSE
## 7 4.328183e-01 FALSE 1.65647230 4.881310e-02 FALSE
## 8 5.437621e-01 FALSE 1.46304606 7.172739e-02 FALSE
## 9 6.864716e-01 FALSE 1.21730200 1.117447e-01 FALSE
## 10 8.562627e-01 FALSE 0.87837059 1.898713e-01 FALSE
## 11 9.782159e-01 FALSE 0.44283545 3.289424e-01 FALSE
## 12 9.857500e-01 FALSE 0.38252551 3.510358e-01 FALSE
## 13 9.950907e-01 FALSE 0.26615693 3.950592e-01 FALSE
## 14 9.999075e-01 FALSE 0.07037122 4.719491e-01 FALSE
##
##
## Summary statistics:
##
## Total Chi squared      = 257.7415
## Total degrees of freedom = 10
## p                     = 0
## Sum of counts          = 2142
##
## Levels:
##
## Variety      POS      Data
##      2        4        2
```

```
# save
cfapos <- cfa::cfa(configs, counts)
cfapos <- cfapos$table %>%
  as.data.frame() %>%
  dplyr::filter(sig.z == TRUE)
write.table(cfapos, here::here("tables", "cfapos.txt"), sep = "\t", row.names = F)
```

Visualization

- prepare data

```
pos_vis <- pos_tb2 %>%
  dplyr::ungroup() %>%
  dplyr::mutate_if(is.character, factor)
# inspect
pos_vis
```

```
## # A tibble: 14 x 4
##   POS      Variety Data      Frequency
##   <fct>    <fct>    <fct>         <int>
## 1 Adjective Australia Xmuch          45
## 2 Adjective Australia control        79
## 3 Adjective Ireland  Xmuch           8
## 4 Adjective Ireland  control         82
## 5 Noun      Australia Xmuch          47
## 6 Noun      Australia control        313
## 7 Noun      Ireland  Xmuch          24
## 8 Noun      Ireland  control        327
```

| | | | |
|-------------|-----------|---------|-----|
| ## 9 Verb | Australia | Xmuch | 17 |
| ## 10 Verb | Australia | control | 71 |
| ## 11 Verb | Ireland | Xmuch | 1 |
| ## 12 Verb | Ireland | control | 74 |
| ## 13 other | Australia | control | 537 |
| ## 14 other | Ireland | control | 517 |

- prepare data

```
pos1 <- pos_vis %>%
  dplyr::filter(POS == "Adjective") %>%
  dplyr::pull(Frequency)
pos2 <- pos_vis %>%
  dplyr::filter(POS == "Noun") %>%
  dplyr::pull(Frequency)
pos3 <- pos_vis %>%
  dplyr::filter(POS == "other") %>%
  dplyr::pull(Frequency)
pos4 <- pos_vis %>%
  dplyr::filter(POS == "Verb") %>%
  dplyr::pull(Frequency)
# add dimnames
column.names <- c("Australia", "Ireland")
row.names <- c("control", "X-much")
matrix.names <- c("Adj.", "Noun", "other", "Verb")
# generate matrix
pos_mx <- array(c(pol1, pol2, pol3), dim = c(2, 2, 4),
               dimnames = list(row.names,
                               column.names,
                               matrix.names))

# inspect
pos_mx
```

```
## , , Adj.
##
##      Australia Ireland
## control      9      1
## X-much     75     56
##
## , , Noun
##
##      Australia Ireland
## control     49     18
## X-much    902    919
##
## , , other
##
##      Australia Ireland
## control     51     14
## X-much     23     25
##
## , , Verb
##
```



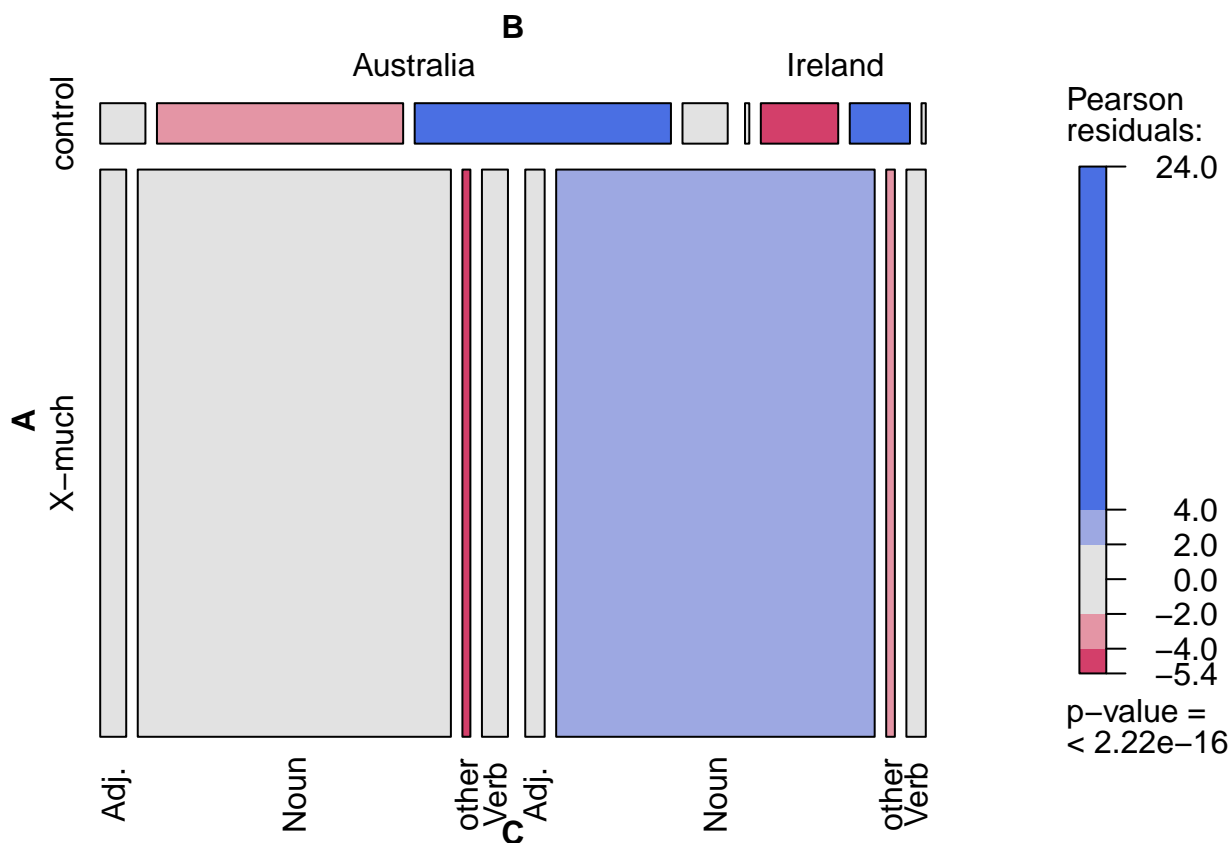
```
##           Australia Ireland
## control           9         1
## X-much           75        56
```

Generate mosaic plot

```
# open window
png(here::here("images", "pos_mosaic.png"), width = 750, height = 300)
# generate plot
mosaic(pos_mx,
       shade = TRUE,
       direction = c("h", "v", "v"),
       just_labels = c("center", "center", "center", "center"),
       rot_labels=c(0,0,90,90),
       offset_labels = c(0, 0, 0.5,0))
# close window
dev.off()
```

```
## pdf
## 2
```

```
# show plot
mosaic(pos_mx,
       shade = TRUE,
       direction = c("h", "v", "v"),
       just_labels = c("center", "center", "center", "center"),
       rot_labels=c(0,0,90,90),
       offset_labels = c(0, 0, 0.5,0))
```



Words

RQ: Are there Words that are significantly over-represented in the X-much structure?

Constructionalization more advanced in AusE

CFA (wo vareity)

RQ: Regardless of variety, are there words that are significantly attracted by the X-much structure?

```
wordcat_tb <- xmuch %>%
  dplyr::group_by(Data, X) %>%
  dplyr::mutate(Frequency = n()) %>%
  dplyr::mutate(X = ifelse(Frequency < 2, "other", X)) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(X, Data) %>%
  dplyr::summarise(Frequency = n())
```

`summarise()` has grouped output by 'X'. You can override using the `.groups`
argument.

```
configs <- wordcat_tb %>%
  dplyr::select(Data, X)
counts = wordcat_tb$Frequency
cfa::cfa(configs, counts)
```

```

##
## *** Analysis of configuration frequencies (CFA) ***
##
##          label      n      expected      Q      chisq p.chisq
## 1      Xmuch jealous   7      0.4649205 3.057297e-03 91.859286684      1
## 2      Xmuch paranoid   4      0.2656688 1.746864e-03 52.491020962      1
## 3      Xmuch agenda     4      0.2656688 1.746864e-03 52.491020962      1
## 4      Xmuch threatened  3      0.1992516 1.310107e-03 39.368265722      1
## 5      Xmuch other 104 60.3732460 2.099836e-02 31.525448569      1
## 6      Xmuch hypocrite   2      0.1328344 8.733777e-04 26.245510481      1
## 7      Xmuch hipster     2      0.1328344 8.733777e-04 26.245510481      1
## 8      Xmuch generalise   2      0.1328344 8.733777e-04 26.245510481      1
## 9      Xmuch excuses     2      0.1328344 8.733777e-04 26.245510481      1
## 10     Xmuch excited     2      0.1328344 8.733777e-04 26.245510481      1
## 11     Xmuch entitled    2      0.1328344 8.733777e-04 26.245510481      1
## 12     Xmuch creepy      2      0.1328344 8.733777e-04 26.245510481      1
## 13     Xmuch coincidence  2      0.1328344 8.733777e-04 26.245510481      1
## 14     Xmuch bitter      2      0.1328344 8.733777e-04 26.245510481      1
## 15     Xmuch biased      2      0.1328344 8.733777e-04 26.245510481      1
## 16     control other 805 848.6267540 3.383563e-02 2.242792433      1
## 17     control the 137 127.9008419 4.526721e-03 0.647334894      1
## 18     control of 63 58.8157156 2.012464e-03 0.297679550      1
## 19     control and 56 52.2806361 1.783252e-03 0.264604044      1
## 20     control to 46 42.9448082 1.458287e-03 0.217353322      1
## 21     control in 39 36.4097287 1.232529e-03 0.184277817      1
## 22     control a 36 33.6089804 1.136205e-03 0.170102600      1
## 23     control that 22 20.5388213 6.900616e-04 0.103951589      1
## 24     control is 22 20.5388213 6.900616e-04 0.103951589      1
## 25     control this 19 17.7380730 5.951751e-04 0.089776372      1
## 26     control it 17 15.8709074 5.320565e-04 0.080326228      1
## 27     control you 16 14.9373246 5.005389e-04 0.075601156      1
## 28     control on 16 14.9373246 5.005389e-04 0.075601156      1
## 29     control i 16 14.9373246 5.005389e-04 0.075601156      1
## 30     control for 15 14.0037418 4.690489e-04 0.070876083      1
## 31     control was 14 13.0701590 4.375867e-04 0.066151011      1
## 32     control have 13 12.1365762 4.061520e-04 0.061425939      1
## 33     control with 12 11.2029935 3.747450e-04 0.056700867      1
## 34     control so 11 10.2694107 3.433655e-04 0.051975794      1
## 35     control from 10 9.3358279 3.120136e-04 0.047250722      1
## 36     control be 10 9.3358279 3.120136e-04 0.047250722      1
## 37     control as 10 9.3358279 3.120136e-04 0.047250722      1
## 38     control they 9 8.4022451 2.806891e-04 0.042525650      1
## 39     control which 8 7.4686623 2.493921e-04 0.037800578      1
## 40     control what 8 7.4686623 2.493921e-04 0.037800578      1
## 41     control or 8 7.4686623 2.493921e-04 0.037800578      1
## 42     control one 8 7.4686623 2.493921e-04 0.037800578      1
## 43     control had 8 7.4686623 2.493921e-04 0.037800578      1
## 44     control by 8 7.4686623 2.493921e-04 0.037800578      1
## 45     control are 8 7.4686623 2.493921e-04 0.037800578      1
## 46     control would 7 6.5350795 2.181225e-04 0.033075506      1
## 47     control will 7 6.5350795 2.181225e-04 0.033075506      1
## 48     control no 7 6.5350795 2.181225e-04 0.033075506      1
## 49     control can 7 6.5350795 2.181225e-04 0.033075506      1
## 50     control at 7 6.5350795 2.181225e-04 0.033075506      1

```

| | | | | | | |
|--------|---------------------|---|-----------|--------------|-------------|---|
| ## 51 | control about | 7 | 6.5350795 | 2.181225e-04 | 0.033075506 | 1 |
| ## 52 | control we | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 53 | control there | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 54 | control their | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 55 | control people | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 56 | control out | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 57 | control not | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 58 | control more | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 59 | control his | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 60 | control but | 6 | 5.6014967 | 1.868803e-04 | 0.028350433 | 1 |
| ## 61 | control then | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 62 | control radio | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 63 | control make | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 64 | control into | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 65 | control he | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 66 | control go | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 67 | control could | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 68 | control because | 5 | 4.6679139 | 1.556654e-04 | 0.023625361 | 1 |
| ## 69 | control your | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 70 | control world | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 71 | control who | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 72 | control time | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 73 | control see | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 74 | control probably | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 75 | control part | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 76 | control my | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 77 | control me | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 78 | control many | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 79 | control life | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 80 | control http | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 81 | control has | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 82 | control free | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 83 | control down | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 84 | control do | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 85 | control company | 4 | 3.7343312 | 1.244779e-04 | 0.018900289 | 1 |
| ## 86 | control years | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 87 | control year | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 88 | control work | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 89 | control while | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 90 | control war | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 91 | control want | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 92 | control use | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 93 | control too | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 94 | control these | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 95 | control them | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 96 | control than | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 97 | control she | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 98 | control our | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 99 | control opportunity | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 100 | control number | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 101 | control new | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 102 | control most | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 103 | control like | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 104 | control last | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |

| | | | | | | |
|--------|----------------------|---|-----------|--------------|-------------|---|
| ## 105 | control instead | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 106 | control home | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 107 | control history | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 108 | control government | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 109 | control get | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 110 | control first | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 111 | control even | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 112 | control dublin | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 113 | control documentary | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 114 | control between | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 115 | control before | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 116 | control bank | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 117 | control an | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 118 | control also | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 119 | control 2012 | 3 | 2.8007484 | 9.331758e-05 | 0.014175217 | 1 |
| ## 120 | control working | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 121 | control winter | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 122 | control were | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 123 | control went | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 124 | control up | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 125 | control today | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 126 | control three | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 127 | control though | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 128 | control think | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 129 | control technical | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 130 | control take | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 131 | control support | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 132 | control such | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 133 | control subject | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 134 | control story | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 135 | control still | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 136 | control song | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 137 | control someone | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 138 | control some | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 139 | control social | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 140 | control should | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 141 | control seen | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 142 | control sector | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 143 | control seat | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 144 | control science | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 145 | control school | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 146 | control same | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 147 | control right | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 148 | control registration | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 149 | control refused | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 150 | control really | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 151 | control put | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 152 | control professional | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 153 | control president | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 154 | control potential | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 155 | control position | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 156 | control policy | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 157 | control police | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 158 | control past | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |

| | | | | | | |
|--------|---------------------|---|-----------|--------------|-------------|---|
| ## 159 | control parents | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 160 | control own | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 161 | control order | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 162 | control only | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 163 | control now | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 164 | control nobody | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 165 | control nice | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 166 | control next | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 167 | control much | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 168 | control minutes | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 169 | control management | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 170 | control man | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 171 | control made | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 172 | control long | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 173 | control likely | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 174 | control law | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 175 | control justice | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 176 | control just | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 177 | control its | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 178 | control island | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 179 | control irish | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 180 | control ireland | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 181 | control information | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 182 | control industrial | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 183 | control immediately | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 184 | control if | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 185 | control however | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 186 | control how | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 187 | control grateful | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 188 | control getting | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 189 | control form | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 190 | control following | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 191 | control fix | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 192 | control few | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 193 | control february | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 194 | control fast | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 195 | control exercise | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 196 | control enough | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 197 | control end | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 198 | control education | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 199 | control drug | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 200 | control done | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 201 | control disclaimer | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 202 | control despite | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 203 | control decided | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 204 | control dead | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 205 | control curious | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 206 | control contribute | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 207 | control completely | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 208 | control communion | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 209 | control clear | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 210 | control children | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 211 | control change | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 212 | control case | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |

| | | | | | | |
|--------|--------------------|------------|--------------|--------------|-------------|---|
| ## 213 | control care | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 214 | control bus | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 215 | control building | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 216 | control better | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 217 | control best | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 218 | control being | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 219 | control been | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 220 | control back | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 221 | control baby | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 222 | control australian | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 223 | control amount | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 224 | control all | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 225 | control agreement | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 226 | control against | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 227 | control again | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 228 | control add | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## 229 | control action | 2 | 1.8671656 | 6.218453e-05 | 0.009450144 | 1 |
| ## | sig.chisq | z | p.z | sig.z | | |
| ## 1 | FALSE | 9.58432505 | 0.000000e+00 | TRUE | | |
| ## 2 | FALSE | 7.24506873 | 2.161604e-13 | TRUE | | |
| ## 3 | FALSE | 7.24506873 | 2.161604e-13 | TRUE | | |
| ## 4 | FALSE | 6.27441358 | 1.754773e-10 | TRUE | | |
| ## 5 | FALSE | 5.61475276 | 9.842175e-09 | TRUE | | |
| ## 6 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 7 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 8 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 9 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 10 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 11 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 12 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 13 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 14 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 15 | FALSE | 5.12303723 | 1.503265e-07 | TRUE | | |
| ## 16 | FALSE | 1.49759555 | 6.711918e-02 | FALSE | | |
| ## 17 | FALSE | 0.80457125 | 2.105336e-01 | FALSE | | |
| ## 18 | FALSE | 0.54560017 | 2.926704e-01 | FALSE | | |
| ## 19 | FALSE | 0.51439678 | 3.034873e-01 | FALSE | | |
| ## 20 | FALSE | 0.46621167 | 3.205320e-01 | FALSE | | |
| ## 21 | FALSE | 0.42927592 | 3.338612e-01 | FALSE | | |
| ## 22 | FALSE | 0.41243496 | 3.400103e-01 | FALSE | | |
| ## 23 | FALSE | 0.32241524 | 3.735691e-01 | FALSE | | |
| ## 24 | FALSE | 0.32241524 | 3.735691e-01 | FALSE | | |
| ## 25 | FALSE | 0.29962706 | 3.822308e-01 | FALSE | | |
| ## 26 | FALSE | 0.28341882 | 3.884279e-01 | FALSE | | |
| ## 27 | FALSE | 0.27495664 | 3.916748e-01 | FALSE | | |
| ## 28 | FALSE | 0.27495664 | 3.916748e-01 | FALSE | | |
| ## 29 | FALSE | 0.27495664 | 3.916748e-01 | FALSE | | |
| ## 30 | FALSE | 0.26622562 | 3.950327e-01 | FALSE | | |
| ## 31 | FALSE | 0.25719839 | 3.985128e-01 | FALSE | | |
| ## 32 | FALSE | 0.24784257 | 4.021281e-01 | FALSE | | |
| ## 33 | FALSE | 0.23811944 | 4.058942e-01 | FALSE | | |
| ## 34 | FALSE | 0.22798200 | 4.098301e-01 | FALSE | | |
| ## 35 | FALSE | 0.21737231 | 4.139591e-01 | FALSE | | |
| ## 36 | FALSE | 0.21737231 | 4.139591e-01 | FALSE | | |

```

## 37      FALSE 0.21737231 4.139591e-01 FALSE
## 38      FALSE 0.20621748 4.183105e-01 FALSE
## 39      FALSE 0.19442371 4.229221e-01 FALSE
## 40      FALSE 0.19442371 4.229221e-01 FALSE
## 41      FALSE 0.19442371 4.229221e-01 FALSE
## 42      FALSE 0.19442371 4.229221e-01 FALSE
## 43      FALSE 0.19442371 4.229221e-01 FALSE
## 44      FALSE 0.19442371 4.229221e-01 FALSE
## 45      FALSE 0.19442371 4.229221e-01 FALSE
## 46      FALSE 0.18186672 4.278437e-01 FALSE
## 47      FALSE 0.18186672 4.278437e-01 FALSE
## 48      FALSE 0.18186672 4.278437e-01 FALSE
## 49      FALSE 0.18186672 4.278437e-01 FALSE
## 50      FALSE 0.18186672 4.278437e-01 FALSE
## 51      FALSE 0.18186672 4.278437e-01 FALSE
## 52      FALSE 0.16837587 4.331438e-01 FALSE
## 53      FALSE 0.16837587 4.331438e-01 FALSE
## 54      FALSE 0.16837587 4.331438e-01 FALSE
## 55      FALSE 0.16837587 4.331438e-01 FALSE
## 56      FALSE 0.16837587 4.331438e-01 FALSE
## 57      FALSE 0.16837587 4.331438e-01 FALSE
## 58      FALSE 0.16837587 4.331438e-01 FALSE
## 59      FALSE 0.16837587 4.331438e-01 FALSE
## 60      FALSE 0.16837587 4.331438e-01 FALSE
## 61      FALSE 0.15370544 4.389210e-01 FALSE
## 62      FALSE 0.15370544 4.389210e-01 FALSE
## 63      FALSE 0.15370544 4.389210e-01 FALSE
## 64      FALSE 0.15370544 4.389210e-01 FALSE
## 65      FALSE 0.15370544 4.389210e-01 FALSE
## 66      FALSE 0.15370544 4.389210e-01 FALSE
## 67      FALSE 0.15370544 4.389210e-01 FALSE
## 68      FALSE 0.15370544 4.389210e-01 FALSE
## 69      FALSE 0.13747832 4.453264e-01 FALSE
## 70      FALSE 0.13747832 4.453264e-01 FALSE
## 71      FALSE 0.13747832 4.453264e-01 FALSE
## 72      FALSE 0.13747832 4.453264e-01 FALSE
## 73      FALSE 0.13747832 4.453264e-01 FALSE
## 74      FALSE 0.13747832 4.453264e-01 FALSE
## 75      FALSE 0.13747832 4.453264e-01 FALSE
## 76      FALSE 0.13747832 4.453264e-01 FALSE
## 77      FALSE 0.13747832 4.453264e-01 FALSE
## 78      FALSE 0.13747832 4.453264e-01 FALSE
## 79      FALSE 0.13747832 4.453264e-01 FALSE
## 80      FALSE 0.13747832 4.453264e-01 FALSE
## 81      FALSE 0.13747832 4.453264e-01 FALSE
## 82      FALSE 0.13747832 4.453264e-01 FALSE
## 83      FALSE 0.13747832 4.453264e-01 FALSE
## 84      FALSE 0.13747832 4.453264e-01 FALSE
## 85      FALSE 0.13747832 4.453264e-01 FALSE
## 86      FALSE 0.11905972 4.526140e-01 FALSE
## 87      FALSE 0.11905972 4.526140e-01 FALSE
## 88      FALSE 0.11905972 4.526140e-01 FALSE
## 89      FALSE 0.11905972 4.526140e-01 FALSE
## 90      FALSE 0.11905972 4.526140e-01 FALSE

```


[illegible]

[illegible]

```

## 199      FALSE 0.09721185 4.612791e-01 FALSE
## 200      FALSE 0.09721185 4.612791e-01 FALSE
## 201      FALSE 0.09721185 4.612791e-01 FALSE
## 202      FALSE 0.09721185 4.612791e-01 FALSE
## 203      FALSE 0.09721185 4.612791e-01 FALSE
## 204      FALSE 0.09721185 4.612791e-01 FALSE
## 205      FALSE 0.09721185 4.612791e-01 FALSE
## 206      FALSE 0.09721185 4.612791e-01 FALSE
## 207      FALSE 0.09721185 4.612791e-01 FALSE
## 208      FALSE 0.09721185 4.612791e-01 FALSE
## 209      FALSE 0.09721185 4.612791e-01 FALSE
## 210      FALSE 0.09721185 4.612791e-01 FALSE
## 211      FALSE 0.09721185 4.612791e-01 FALSE
## 212      FALSE 0.09721185 4.612791e-01 FALSE
## 213      FALSE 0.09721185 4.612791e-01 FALSE
## 214      FALSE 0.09721185 4.612791e-01 FALSE
## 215      FALSE 0.09721185 4.612791e-01 FALSE
## 216      FALSE 0.09721185 4.612791e-01 FALSE
## 217      FALSE 0.09721185 4.612791e-01 FALSE
## 218      FALSE 0.09721185 4.612791e-01 FALSE
## 219      FALSE 0.09721185 4.612791e-01 FALSE
## 220      FALSE 0.09721185 4.612791e-01 FALSE
## 221      FALSE 0.09721185 4.612791e-01 FALSE
## 222      FALSE 0.09721185 4.612791e-01 FALSE
## 223      FALSE 0.09721185 4.612791e-01 FALSE
## 224      FALSE 0.09721185 4.612791e-01 FALSE
## 225      FALSE 0.09721185 4.612791e-01 FALSE
## 226      FALSE 0.09721185 4.612791e-01 FALSE
## 227      FALSE 0.09721185 4.612791e-01 FALSE
## 228      FALSE 0.09721185 4.612791e-01 FALSE
## 229      FALSE 0.09721185 4.612791e-01 FALSE
##
##
## Summary statistics:
##
## Total Chi squared      = 538.0605
## Total degrees of freedom = 227
## p                      = 0
## Sum of counts          = 2138
##
## Levels:
##
## Data      X
##      2    228

```

```

r7 <- cfa::cfa(configs, counts)
# save to disc
sig_words <- r7$table %>%
  as.data.frame() %>%
  dplyr::mutate(expected = round(expected, 1),
                Q = round(Q, 3),
                chisq = round(chisq, 3),
                z = round(z, 3),
                p.z = round(p.z, 3),

```

```

        label = stringr::str_remove_all(label, "Xmuch ") %>%
dplyr::select(-p.chisq, -sig.chisq)
# save to disc
write.table(sig_words,
            here::here("tables", "xmuch_words.txt"),
            sep = "\t",
            row.names = F)
# extract words
words <- r7$table %>%
  as.data.frame() %>%
  dplyr::mutate(Type = ifelse(n > expected, "Type", "Antitype")) %>%
  dplyr::filter(sig.z == T,
                Type == "Type") %>%
  dplyr::pull(label) %>%
  stringr::str_remove_all("Xmuch ")
words

```

```

## [1] "jealous"      "paranoid"      "agenda"        "threatened"    "other"
## [6] "hypocrite"    "hipster"       "generalise"    "excuses"       "excited"
## [11] "entitled"     "creepy"        "coincidence"   "bitter"        "biased"

```

Words that are significantly over-represented in the X-much structure: jealous, paranoid, agenda, threatened, other, hypocrite, hipster, generalise, excuses, excited, entitled, creepy, coincidence, bitter, biased

X2 (w variety)

RQ: Do the words that are significantly attracted by the x-much construction differ across varieties?

```

wordcat_tb <- xmuch %>%
  dplyr::filter(Data == "Xmuch") %>%
  dplyr::group_by(Variety, X) %>%
  dplyr::mutate(Frequency = n()) %>%
  dplyr::mutate(X = ifelse(Frequency < 2, "other", X)) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(X, Variety) %>%
  dplyr::summarise(Frequency = n())

```

```

## `summarise()` has grouped output by 'X'. You can override using the `.groups`
## argument.

```

```

configs <- wordcat_tb %>%
  dplyr::select(Variety, X)
counts = wordcat_tb$Frequency
cfa::cfa(configs, counts)

```

```

##
## *** Analysis of configuration frequencies (CFA) ***
##
##           label  n  expected      Q      chisq  p.chisq
## 1      Ireland excuses  2  0.4647887 0.010846850 5.070849338 0.8863952
## 2      Ireland coincidence  2  0.4647887 0.010846850 5.070849338 0.8863952

```

```
## 3      Ireland agenda 2 0.9295775 0.007587859 1.232607768 0.9995549
## 4      Australia jealous 7 5.3732394 0.011906603 0.492505492 0.9999939
## 5      Australia agenda 2 3.0704225 0.007704785 0.373174829 0.9999984
## 6      Australia threatened 3 2.3028169 0.004990674 0.211073782 0.9999999
## 7      Australia paranoid 3 2.3028169 0.004990674 0.211073782 0.9999999
## 8      Australia hypocrite 2 1.5352113 0.003308934 0.140715855 1.0000000
## 9      Australia generalise 2 1.5352113 0.003308934 0.140715855 1.0000000
## 10     Australia entitled 2 1.5352113 0.003308934 0.140715855 1.0000000
## 11     Australia creepy 2 1.5352113 0.003308934 0.140715855 1.0000000
## 12      Ireland other 27 26.2605634 0.006388804 0.020820822 1.0000000
## 13     Australia other 86 86.7394366 0.008524803 0.006303552 1.0000000
##      sig.chisq      z      p.z sig.z
## 1      FALSE 2.25185464 0.01216573 FALSE
## 2      FALSE 2.25185464 0.01216573 FALSE
## 3      FALSE 1.11022870 0.13345024 FALSE
## 4      FALSE 0.70178736 0.24140589 FALSE
## 5      FALSE 0.61088037 0.27063939 FALSE
## 6      FALSE 0.45942767 0.32296354 FALSE
## 7      FALSE 0.45942767 0.32296354 FALSE
## 8      FALSE 0.37512112 0.35378520 FALSE
## 9      FALSE 0.37512112 0.35378520 FALSE
## 10     FALSE 0.37512112 0.35378520 FALSE
## 11     FALSE 0.37512112 0.35378520 FALSE
## 12     FALSE 0.14429422 0.44263407 FALSE
## 13     FALSE 0.07939491 0.46835926 FALSE
##
##
## Summary statistics:
##
## Total Chi squared      = 13.25212
## Total degrees of freedom = 10
## p                      = 0.209915
## Sum of counts          = 142
##
## Levels:
##
## Variety      X
##      2      11
```

```
r8 <- cfa::cfa(configs, counts)
words <- r8$table %>%
  as.data.frame() %>%
  dplyr::filter(sig.z == T) %>%
  dplyr::pull(label) %>%
  stringr::str_remove_all("Xmuch ")
words
```

```
## character(0)
```

The analysis does not confirm variety specific attraction of types to the X-much structure.

CFA (w variety)

RQ: Do the words that are significantly attracted by the x-much construction differ across varieties when we consider the control data?

```
wordcat_tb <- xmuch %>%
  dplyr::group_by(Data, Variety, X) %>%
  dplyr::mutate(Frequency = n()) %>%
  dplyr::mutate(X = ifelse(Frequency < 2, "other", X)) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(X, Variety, Data) %>%
  dplyr::summarise(Frequency = n())
```

```
## `summarise()` has grouped output by 'X', 'Variety'. You can override using the
## `.groups` argument.
```

```
configs <- wordcat_tb %>%
  dplyr::select(Data, Variety, X)
counts = wordcat_tb$Frequency
cfa::cfa(configs, counts)
```

```
##
## *** Analysis of configuration frequencies (CFA) ***
##
##              label    n   expected      Q      chisq
## 1      Xmuch Australia jealous    7  0.24072356 3.161851e-03 1.897937e+02
## 2      Xmuch Australia threatened    3  0.10316724 1.354992e-03 8.134016e+01
## 3      Xmuch Australia paranoid    3  0.10316724 1.354992e-03 8.134016e+01
## 4      Xmuch Australia other   86 37.17459600 2.324106e-02 6.412767e+01
## 5      Xmuch Ireland excuses    2  0.06405626 9.055200e-04 5.850916e+01
## 6      Xmuch Ireland coincidence    2  0.06405626 9.055200e-04 5.850916e+01
## 7      Xmuch Australia hypocrite    2  0.06877816 9.033134e-04 5.422677e+01
## 8      Xmuch Australia generalise    2  0.06877816 9.033134e-04 5.422677e+01
## 9      Xmuch Australia entitled    2  0.06877816 9.033134e-04 5.422677e+01
## 10     Xmuch Australia creepy    2  0.06877816 9.033134e-04 5.422677e+01
## 11     Xmuch Ireland agenda    2  0.12811253 8.755845e-04 2.735066e+01
## 12     Xmuch Australia agenda    2  0.13755632 8.711710e-04 2.521655e+01
## 13     control Ireland radio    5  2.25099124 1.287140e-03 3.357209e+00
## 14     control Ireland people    5  2.25099124 1.287140e-03 3.357209e+00
## 15     control Ireland on   12  7.20317197 2.251190e-03 3.194365e+00
## 16     control Australia we    5  2.41692270 1.209542e-03 2.760654e+00
## 17     control Australia could    5  2.41692270 1.209542e-03 2.760654e+00
## 18     control Australia that   16 10.63445987 2.522152e-03 2.707145e+00
## 19     control Australia other 485 522.53868744 2.323713e-02 2.696744e+00
## 20     control Ireland life    4  1.80079299 1.029495e-03 2.685768e+00
## 21     control Ireland http    4  1.80079299 1.029495e-03 2.685768e+00
## 22     control Australia you   12  7.73415264 2.002495e-03 2.352870e+00
## 23     control Ireland this   13  8.55376671 2.087976e-03 2.311144e+00
## 24     control Australia world    4  1.93353816 9.674146e-04 2.208524e+00
## 25     control Australia as    8  4.83384540 1.484251e-03 2.073822e+00
## 26     control Ireland your    3  1.35059474 7.719588e-04 2.014326e+00
## 27     control Ireland than    3  1.35059474 7.719588e-04 2.014326e+00
## 28     control Ireland most    3  1.35059474 7.719588e-04 2.014326e+00
```

| | | | | | |
|-------|------------------------------|----|-------------|--------------|--------------|
| ## 29 | control Ireland many | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 30 | control Ireland last | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 31 | control Ireland instead | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 32 | control Ireland first | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 33 | control Ireland dublin | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 34 | control Ireland down | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 35 | control Ireland documentary | 3 | 1.35059474 | 7.719588e-04 | 2.014326e+00 |
| ## 36 | control Australia on | 4 | 7.73415264 | 1.752905e-03 | 1.802899e+00 |
| ## 37 | Xmuch Ireland other | 27 | 34.62241055 | 3.623891e-03 | 1.678137e+00 |
| ## 38 | control Australia part | 3 | 1.45015362 | 7.253968e-04 | 1.656393e+00 |
| ## 39 | control Australia our | 3 | 1.45015362 | 7.253968e-04 | 1.656393e+00 |
| ## 40 | control Australia me | 3 | 1.45015362 | 7.253968e-04 | 1.656393e+00 |
| ## 41 | control Australia free | 3 | 1.45015362 | 7.253968e-04 | 1.656393e+00 |
| ## 42 | control Australia also | 3 | 1.45015362 | 7.253968e-04 | 1.656393e+00 |
| ## 43 | control Ireland one | 6 | 3.60158598 | 1.123696e-03 | 1.597182e+00 |
| ## 44 | control Ireland of | 35 | 28.36248962 | 3.146280e-03 | 1.553338e+00 |
| ## 45 | control Ireland that | 6 | 9.90436145 | 1.834674e-03 | 1.539124e+00 |
| ## 46 | control Ireland you | 4 | 7.20317197 | 1.503274e-03 | 1.424416e+00 |
| ## 47 | control Ireland a | 21 | 16.20713693 | 2.258874e-03 | 1.417372e+00 |
| ## 48 | control Ireland as | 2 | 4.50198248 | 1.172714e-03 | 1.390480e+00 |
| ## 49 | control Ireland years | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 50 | control Ireland year | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 51 | control Ireland work | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 52 | control Ireland while | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 53 | control Ireland were | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 54 | control Ireland use | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 55 | control Ireland too | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 56 | control Ireland song | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 57 | control Ireland should | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 58 | control Ireland she | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 59 | control Ireland sector | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 60 | control Ireland registration | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 61 | control Ireland refused | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 62 | control Ireland professional | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 63 | control Ireland past | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 64 | control Ireland number | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 65 | control Ireland nice | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 66 | control Ireland new | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 67 | control Ireland management | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 68 | control Ireland like | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 69 | control Ireland its | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 70 | control Ireland irish | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 71 | control Ireland ireland | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 72 | control Ireland how | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 73 | control Ireland history | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 74 | control Ireland government | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 75 | control Ireland form | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 76 | control Ireland following | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 77 | control Ireland despite | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 78 | control Ireland dead | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 79 | control Ireland curious | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 80 | control Ireland change | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 81 | control Ireland case | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 82 | control Ireland building | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |

| | | | | | |
|--------|-------------------------------|----|-------------|--------------|--------------|
| ## 83 | control Ireland an | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 84 | control Ireland 2012 | 2 | 0.90039650 | 5.145308e-04 | 1.342884e+00 |
| ## 85 | control Australia by | 6 | 3.86707632 | 9.994334e-04 | 1.176435e+00 |
| ## 86 | control Australia have | 9 | 6.28399902 | 1.274091e-03 | 1.173880e+00 |
| ## 87 | control Australia winter | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 88 | control Australia went | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 89 | control Australia war | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 90 | control Australia want | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 91 | control Australia think | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 92 | control Australia these | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 93 | control Australia them | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 94 | control Australia support | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 95 | control Australia story | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 96 | control Australia police | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 97 | control Australia opportunity | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 98 | control Australia law | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 99 | control Australia information | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 100 | control Australia however | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 101 | control Australia home | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 102 | control Australia get | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 103 | control Australia fix | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 104 | control Australia february | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 105 | control Australia exercise | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 106 | control Australia even | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 107 | control Australia enough | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 108 | control Australia end | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 109 | control Australia education | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 110 | control Australia contribute | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 111 | control Australia completely | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 112 | control Australia children | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 113 | control Australia between | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 114 | control Australia being | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 115 | control Australia before | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 116 | control Australia been | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 117 | control Australia bank | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 118 | control Australia australian | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 119 | control Australia all | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 120 | control Australia against | 2 | 0.96676908 | 4.834885e-04 | 1.104262e+00 |
| ## 121 | control Australia this | 6 | 9.18430625 | 1.495811e-03 | 1.104036e+00 |
| ## 122 | control Ireland about | 5 | 3.15138774 | 8.659219e-04 | 1.084401e+00 |
| ## 123 | control Australia one | 2 | 3.86707632 | 8.748641e-04 | 9.014495e-01 |
| ## 124 | control Australia and | 32 | 27.06953422 | 2.335684e-03 | 8.980388e-01 |
| ## 125 | control Ireland to | 25 | 20.70911940 | 2.026590e-03 | 8.890603e-01 |
| ## 126 | control Australia would | 5 | 3.38369178 | 7.571891e-04 | 7.720716e-01 |
| ## 127 | control Australia was | 9 | 6.76738356 | 1.047571e-03 | 7.365588e-01 |
| ## 128 | control Ireland it | 10 | 7.65337021 | 1.101525e-03 | 7.195093e-01 |
| ## 129 | control Ireland by | 2 | 3.60158598 | 7.503688e-04 | 7.122078e-01 |
| ## 130 | control Australia they | 6 | 4.35046086 | 7.731069e-04 | 6.254462e-01 |
| ## 131 | control Ireland out | 4 | 2.70118949 | 6.082570e-04 | 6.245059e-01 |
| ## 132 | control Ireland but | 4 | 2.70118949 | 6.082570e-04 | 6.245059e-01 |
| ## 133 | control Ireland have | 4 | 5.85257722 | 8.688786e-04 | 5.864156e-01 |
| ## 134 | control Australia about | 2 | 3.38369178 | 6.482157e-04 | 5.658325e-01 |
| ## 135 | control Ireland or | 5 | 3.60158598 | 6.551795e-04 | 5.429724e-01 |
| ## 136 | control Ireland are | 5 | 3.60158598 | 6.551795e-04 | 5.429724e-01 |

| | | | | | |
|--------|---------------------------|----|-------------|--------------|--------------|
| ## 137 | control Australia the | 72 | 66.22368194 | 2.788099e-03 | 5.038356e-01 |
| ## 138 | control Ireland with | 7 | 5.40237898 | 7.491432e-04 | 4.724572e-01 |
| ## 139 | control Ireland would | 2 | 3.15138774 | 5.393299e-04 | 4.206698e-01 |
| ## 140 | control Australia there | 4 | 2.90030724 | 5.150545e-04 | 4.169642e-01 |
| ## 141 | control Australia more | 4 | 2.90030724 | 5.150545e-04 | 4.169642e-01 |
| ## 142 | control Australia which | 5 | 3.86707632 | 5.308590e-04 | 3.319086e-01 |
| ## 143 | control Australia what | 5 | 3.86707632 | 5.308590e-04 | 3.319086e-01 |
| ## 144 | control Australia a | 15 | 17.40184343 | 1.132625e-03 | 3.315081e-01 |
| ## 145 | control Australia be | 6 | 4.83384540 | 5.466778e-04 | 2.813322e-01 |
| ## 146 | control Australia out | 2 | 2.90030724 | 4.216699e-04 | 2.794715e-01 |
| ## 147 | control Australia but | 2 | 2.90030724 | 4.216699e-04 | 2.794715e-01 |
| ## 148 | control Ireland they | 3 | 4.05178423 | 4.928818e-04 | 2.730279e-01 |
| ## 149 | control Ireland was | 5 | 6.30277547 | 6.111447e-04 | 2.692820e-01 |
| ## 150 | control Ireland he | 3 | 2.25099124 | 3.507007e-04 | 2.492298e-01 |
| ## 151 | control Ireland go | 3 | 2.25099124 | 3.507007e-04 | 2.492298e-01 |
| ## 152 | control Ireland because | 3 | 2.25099124 | 3.507007e-04 | 2.492298e-01 |
| ## 153 | control Australia in | 21 | 18.85199705 | 1.013616e-03 | 2.447442e-01 |
| ## 154 | control Ireland for | 8 | 6.75297372 | 5.851158e-04 | 2.302800e-01 |
| ## 155 | control Ireland no | 4 | 3.15138774 | 3.975047e-04 | 2.285161e-01 |
| ## 156 | control Ireland can | 4 | 3.15138774 | 3.975047e-04 | 2.285161e-01 |
| ## 157 | control Australia i | 9 | 7.73415264 | 5.942204e-04 | 2.071810e-01 |
| ## 158 | control Australia of | 28 | 30.45322600 | 1.164020e-03 | 1.976250e-01 |
| ## 159 | control Australia or | 3 | 3.86707632 | 4.062897e-04 | 1.944160e-01 |
| ## 160 | control Australia are | 3 | 3.86707632 | 4.062897e-04 | 1.944160e-01 |
| ## 161 | control Ireland there | 2 | 2.70118949 | 3.283800e-04 | 1.820186e-01 |
| ## 162 | control Ireland more | 2 | 2.70118949 | 3.283800e-04 | 1.820186e-01 |
| ## 163 | control Australia it | 7 | 8.21753718 | 5.716721e-04 | 1.803943e-01 |
| ## 164 | control Ireland the | 65 | 61.67715997 | 1.600348e-03 | 1.790171e-01 |
| ## 165 | control Australia then | 3 | 2.41692270 | 2.730296e-04 | 1.406661e-01 |
| ## 166 | control Australia make | 3 | 2.41692270 | 2.730296e-04 | 1.406661e-01 |
| ## 167 | control Australia into | 3 | 2.41692270 | 2.730296e-04 | 1.406661e-01 |
| ## 168 | control Ireland is | 11 | 9.90436145 | 5.148446e-04 | 1.212015e-01 |
| ## 169 | control Australia will | 4 | 3.38369178 | 2.887208e-04 | 1.122549e-01 |
| ## 170 | control Australia at | 4 | 3.38369178 | 2.887208e-04 | 1.122549e-01 |
| ## 171 | control Australia with | 5 | 5.80061448 | 3.754876e-04 | 1.105027e-01 |
| ## 172 | control Ireland which | 3 | 3.60158598 | 2.818527e-04 | 1.004851e-01 |
| ## 173 | control Ireland what | 3 | 3.60158598 | 2.818527e-04 | 1.004851e-01 |
| ## 174 | control Australia so | 6 | 5.31722994 | 3.201461e-04 | 8.767252e-02 |
| ## 175 | control Australia he | 2 | 2.41692270 | 1.952266e-04 | 7.191977e-02 |
| ## 176 | control Australia go | 2 | 2.41692270 | 1.952266e-04 | 7.191977e-02 |
| ## 177 | control Australia because | 2 | 2.41692270 | 1.952266e-04 | 7.191977e-02 |
| ## 178 | control Australia to | 21 | 22.23568883 | 5.840390e-04 | 6.867010e-02 |
| ## 179 | control Ireland and | 24 | 25.21110188 | 5.732243e-04 | 5.817944e-02 |
| ## 180 | control Ireland be | 4 | 4.50198248 | 2.352861e-04 | 5.597232e-02 |
| ## 181 | control Ireland from | 5 | 4.50198248 | 2.334277e-04 | 5.509161e-02 |
| ## 182 | control Ireland had | 4 | 3.60158598 | 1.866634e-04 | 4.407329e-02 |
| ## 183 | control Australia no | 3 | 3.38369178 | 1.797474e-04 | 4.350851e-02 |
| ## 184 | control Australia can | 3 | 3.38369178 | 1.797474e-04 | 4.350851e-02 |
| ## 185 | control Ireland their | 3 | 2.70118949 | 1.399385e-04 | 3.305496e-02 |
| ## 186 | control Ireland not | 3 | 2.70118949 | 1.399385e-04 | 3.305496e-02 |
| ## 187 | control Ireland his | 3 | 2.70118949 | 1.399385e-04 | 3.305496e-02 |
| ## 188 | control Ireland then | 2 | 2.25099124 | 1.175191e-04 | 2.798616e-02 |
| ## 189 | control Ireland make | 2 | 2.25099124 | 1.175191e-04 | 2.798616e-02 |
| ## 190 | control Ireland into | 2 | 2.25099124 | 1.175191e-04 | 2.798616e-02 |

| | | | | | |
|--------|----------------------------|-----------|--------------|--------------|--------------|
| ## 191 | control Ireland other | 483 | 486.66430601 | 2.218995e-03 | 2.759014e-02 |
| ## 192 | control Ireland who | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 193 | control Ireland time | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 194 | control Ireland see | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 195 | control Ireland probably | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 196 | control Ireland my | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 197 | control Ireland has | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 198 | control Ireland do | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 199 | control Ireland company | 2 | 1.80079299 | 9.325301e-05 | 2.203664e-02 |
| ## 200 | control Australia is | 11 | 10.63445987 | 1.718276e-04 | 1.256477e-02 |
| ## 201 | control Ireland in | 18 | 17.55773167 | 2.085736e-04 | 1.114046e-02 |
| ## 202 | control Australia for | 7 | 7.25076810 | 1.176901e-04 | 8.672824e-03 |
| ## 203 | control Ireland will | 3 | 3.15138774 | 7.091263e-05 | 7.272430e-03 |
| ## 204 | control Ireland at | 3 | 3.15138774 | 7.091263e-05 | 7.272430e-03 |
| ## 205 | control Ireland i | 7 | 7.20317197 | 9.535023e-05 | 5.730649e-03 |
| ## 206 | control Australia from | 5 | 4.83384540 | 7.789107e-05 | 5.711261e-03 |
| ## 207 | control Australia had | 4 | 3.86707632 | 6.228463e-05 | 4.569009e-03 |
| ## 208 | control Australia their | 3 | 2.90030724 | 4.669232e-05 | 3.426757e-03 |
| ## 209 | control Australia not | 3 | 2.90030724 | 4.669232e-05 | 3.426757e-03 |
| ## 210 | control Australia his | 3 | 2.90030724 | 4.669232e-05 | 3.426757e-03 |
| ## 211 | control Australia who | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 212 | control Australia time | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 213 | control Australia see | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 214 | control Australia probably | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 215 | control Australia my | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 216 | control Australia has | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 217 | control Australia do | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 218 | control Australia company | 2 | 1.93353816 | 3.111413e-05 | 2.284504e-03 |
| ## 219 | control Ireland so | 5 | 4.95218073 | 2.241828e-05 | 4.617527e-04 |
| ## | p.chisq | sig.chisq | z | p.z | sig.z |
| ## 1 | 0.04291081 | FALSE | 13.77656376 | 0.000000e+00 | TRUE |
| ## 2 | 0.99999993 | FALSE | 9.01887804 | 0.000000e+00 | TRUE |
| ## 3 | 0.99999993 | FALSE | 9.01887804 | 0.000000e+00 | TRUE |
| ## 4 | 1.00000000 | FALSE | 8.00797515 | 5.551115e-16 | TRUE |
| ## 5 | 1.00000000 | FALSE | 7.64912802 | 1.010303e-14 | TRUE |
| ## 6 | 1.00000000 | FALSE | 7.64912802 | 1.010303e-14 | TRUE |
| ## 7 | 1.00000000 | FALSE | 7.36388308 | 8.937295e-14 | TRUE |
| ## 8 | 1.00000000 | FALSE | 7.36388308 | 8.937295e-14 | TRUE |
| ## 9 | 1.00000000 | FALSE | 7.36388308 | 8.937295e-14 | TRUE |
| ## 10 | 1.00000000 | FALSE | 7.36388308 | 8.937295e-14 | TRUE |
| ## 11 | 1.00000000 | FALSE | 5.22978624 | 8.485307e-08 | TRUE |
| ## 12 | 1.00000000 | FALSE | 5.02160873 | 2.562024e-07 | TRUE |
| ## 13 | 1.00000000 | FALSE | 1.83226895 | 3.345568e-02 | FALSE |
| ## 14 | 1.00000000 | FALSE | 1.83226895 | 3.345568e-02 | FALSE |
| ## 15 | 1.00000000 | FALSE | 1.78727861 | 3.694623e-02 | FALSE |
| ## 16 | 1.00000000 | FALSE | 1.66152172 | 4.830436e-02 | FALSE |
| ## 17 | 1.00000000 | FALSE | 1.66152172 | 4.830436e-02 | FALSE |
| ## 18 | 1.00000000 | FALSE | 1.64534028 | 4.994983e-02 | FALSE |
| ## 19 | 1.00000000 | FALSE | 1.64217660 | 5.027671e-02 | FALSE |
| ## 20 | 1.00000000 | FALSE | 1.63883117 | 5.062421e-02 | FALSE |
| ## 21 | 1.00000000 | FALSE | 1.63883117 | 5.062421e-02 | FALSE |
| ## 22 | 1.00000000 | FALSE | 1.53390669 | 6.252630e-02 | FALSE |
| ## 23 | 1.00000000 | FALSE | 1.52024487 | 6.422472e-02 | FALSE |
| ## 24 | 1.00000000 | FALSE | 1.48611020 | 6.862499e-02 | FALSE |

| | | | | | |
|-------|------------|-------|------------|--------------|-------|
| ## 25 | 1.00000000 | FALSE | 1.44007708 | 7.492280e-02 | FALSE |
| ## 26 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 27 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 28 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 29 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 30 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 31 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 32 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 33 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 34 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 35 | 1.00000000 | FALSE | 1.41926942 | 7.791024e-02 | FALSE |
| ## 36 | 1.00000000 | FALSE | 1.34272074 | 8.968121e-02 | FALSE |
| ## 37 | 1.00000000 | FALSE | 1.29542921 | 9.758610e-02 | FALSE |
| ## 38 | 1.00000000 | FALSE | 1.28700919 | 9.904554e-02 | FALSE |
| ## 39 | 1.00000000 | FALSE | 1.28700919 | 9.904554e-02 | FALSE |
| ## 40 | 1.00000000 | FALSE | 1.28700919 | 9.904554e-02 | FALSE |
| ## 41 | 1.00000000 | FALSE | 1.28700919 | 9.904554e-02 | FALSE |
| ## 42 | 1.00000000 | FALSE | 1.28700919 | 9.904554e-02 | FALSE |
| ## 43 | 1.00000000 | FALSE | 1.26379682 | 1.031515e-01 | FALSE |
| ## 44 | 1.00000000 | FALSE | 1.24632997 | 1.063216e-01 | FALSE |
| ## 45 | 1.00000000 | FALSE | 1.24061428 | 1.073741e-01 | FALSE |
| ## 46 | 1.00000000 | FALSE | 1.19348884 | 1.163390e-01 | FALSE |
| ## 47 | 1.00000000 | FALSE | 1.19053419 | 1.169183e-01 | FALSE |
| ## 48 | 1.00000000 | FALSE | 1.17918613 | 1.191620e-01 | FALSE |
| ## 49 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 50 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 51 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 52 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 53 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 54 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 55 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 56 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 57 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 58 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 59 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 60 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 61 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 62 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 63 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 64 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 65 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 66 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 67 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 68 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 69 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 70 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 71 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 72 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 73 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 74 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 75 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 76 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 77 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 78 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |

| | | | | | |
|--------|------------|-------|------------|--------------|-------|
| ## 79 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 80 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 81 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 82 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 83 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 84 | 1.00000000 | FALSE | 1.15882863 | 1.232630e-01 | FALSE |
| ## 85 | 1.00000000 | FALSE | 1.08463582 | 1.390415e-01 | FALSE |
| ## 86 | 1.00000000 | FALSE | 1.08345747 | 1.393027e-01 | FALSE |
| ## 87 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 88 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 89 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 90 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 91 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 92 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 93 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 94 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 95 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 96 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 97 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 98 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 99 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 100 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 101 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 102 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 103 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 104 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 105 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 106 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 107 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 108 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 109 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 110 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 111 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 112 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 113 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 114 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 115 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 116 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 117 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 118 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 119 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 120 | 1.00000000 | FALSE | 1.05083860 | 1.466664e-01 | FALSE |
| ## 121 | 1.00000000 | FALSE | 1.05073126 | 1.466910e-01 | FALSE |
| ## 122 | 1.00000000 | FALSE | 1.04134566 | 1.488576e-01 | FALSE |
| ## 123 | 1.00000000 | FALSE | 0.94944694 | 1.711967e-01 | FALSE |
| ## 124 | 1.00000000 | FALSE | 0.94764911 | 1.716541e-01 | FALSE |
| ## 125 | 1.00000000 | FALSE | 0.94289995 | 1.728660e-01 | FALSE |
| ## 126 | 1.00000000 | FALSE | 0.87867604 | 1.897885e-01 | FALSE |
| ## 127 | 1.00000000 | FALSE | 0.85823006 | 1.953827e-01 | FALSE |
| ## 128 | 1.00000000 | FALSE | 0.84823896 | 1.981525e-01 | FALSE |
| ## 129 | 1.00000000 | FALSE | 0.84392405 | 1.993559e-01 | FALSE |
| ## 130 | 1.00000000 | FALSE | 0.79085159 | 2.145153e-01 | FALSE |
| ## 131 | 1.00000000 | FALSE | 0.79025685 | 2.146889e-01 | FALSE |
| ## 132 | 1.00000000 | FALSE | 0.79025685 | 2.146889e-01 | FALSE |

| | | | | | |
|--------|------------|-------|------------|--------------|-------|
| ## 133 | 1.00000000 | FALSE | 0.76577775 | 2.219043e-01 | FALSE |
| ## 134 | 1.00000000 | FALSE | 0.75221842 | 2.259599e-01 | FALSE |
| ## 135 | 1.00000000 | FALSE | 0.73686660 | 2.306017e-01 | FALSE |
| ## 136 | 1.00000000 | FALSE | 0.73686660 | 2.306017e-01 | FALSE |
| ## 137 | 1.00000000 | FALSE | 0.70981380 | 2.389098e-01 | FALSE |
| ## 138 | 1.00000000 | FALSE | 0.68735523 | 2.459295e-01 | FALSE |
| ## 139 | 1.00000000 | FALSE | 0.64859064 | 2.583015e-01 | FALSE |
| ## 140 | 1.00000000 | FALSE | 0.64572762 | 2.592279e-01 | FALSE |
| ## 141 | 1.00000000 | FALSE | 0.64572762 | 2.592279e-01 | FALSE |
| ## 142 | 1.00000000 | FALSE | 0.57611513 | 2.822687e-01 | FALSE |
| ## 143 | 1.00000000 | FALSE | 0.57611513 | 2.822687e-01 | FALSE |
| ## 144 | 1.00000000 | FALSE | 0.57576739 | 2.823862e-01 | FALSE |
| ## 145 | 1.00000000 | FALSE | 0.53040762 | 2.979147e-01 | FALSE |
| ## 146 | 1.00000000 | FALSE | 0.52865061 | 2.985239e-01 | FALSE |
| ## 147 | 1.00000000 | FALSE | 0.52865061 | 2.985239e-01 | FALSE |
| ## 148 | 1.00000000 | FALSE | 0.52252070 | 3.006539e-01 | FALSE |
| ## 149 | 1.00000000 | FALSE | 0.51892388 | 3.019069e-01 | FALSE |
| ## 150 | 1.00000000 | FALSE | 0.49922922 | 3.088090e-01 | FALSE |
| ## 151 | 1.00000000 | FALSE | 0.49922922 | 3.088090e-01 | FALSE |
| ## 152 | 1.00000000 | FALSE | 0.49922922 | 3.088090e-01 | FALSE |
| ## 153 | 1.00000000 | FALSE | 0.49471627 | 3.104002e-01 | FALSE |
| ## 154 | 1.00000000 | FALSE | 0.47987495 | 3.156582e-01 | FALSE |
| ## 155 | 1.00000000 | FALSE | 0.47803356 | 3.163132e-01 | FALSE |
| ## 156 | 1.00000000 | FALSE | 0.47803356 | 3.163132e-01 | FALSE |
| ## 157 | 1.00000000 | FALSE | 0.45517141 | 3.244930e-01 | FALSE |
| ## 158 | 1.00000000 | FALSE | 0.44455030 | 3.283224e-01 | FALSE |
| ## 159 | 1.00000000 | FALSE | 0.44092625 | 3.296332e-01 | FALSE |
| ## 160 | 1.00000000 | FALSE | 0.44092625 | 3.296332e-01 | FALSE |
| ## 161 | 1.00000000 | FALSE | 0.42663637 | 3.348221e-01 | FALSE |
| ## 162 | 1.00000000 | FALSE | 0.42663637 | 3.348221e-01 | FALSE |
| ## 163 | 1.00000000 | FALSE | 0.42472849 | 3.355173e-01 | FALSE |
| ## 164 | 1.00000000 | FALSE | 0.42310412 | 3.361096e-01 | FALSE |
| ## 165 | 1.00000000 | FALSE | 0.37505482 | 3.538098e-01 | FALSE |
| ## 166 | 1.00000000 | FALSE | 0.37505482 | 3.538098e-01 | FALSE |
| ## 167 | 1.00000000 | FALSE | 0.37505482 | 3.538098e-01 | FALSE |
| ## 168 | 1.00000000 | FALSE | 0.34814011 | 3.638675e-01 | FALSE |
| ## 169 | 1.00000000 | FALSE | 0.33504455 | 3.687957e-01 | FALSE |
| ## 170 | 1.00000000 | FALSE | 0.33504455 | 3.687957e-01 | FALSE |
| ## 171 | 1.00000000 | FALSE | 0.33241946 | 3.697863e-01 | FALSE |
| ## 172 | 1.00000000 | FALSE | 0.31699383 | 3.756241e-01 | FALSE |
| ## 173 | 1.00000000 | FALSE | 0.31699383 | 3.756241e-01 | FALSE |
| ## 174 | 1.00000000 | FALSE | 0.29609546 | 3.835786e-01 | FALSE |
| ## 175 | 1.00000000 | FALSE | 0.26817863 | 3.942809e-01 | FALSE |
| ## 176 | 1.00000000 | FALSE | 0.26817863 | 3.942809e-01 | FALSE |
| ## 177 | 1.00000000 | FALSE | 0.26817863 | 3.942809e-01 | FALSE |
| ## 178 | 1.00000000 | FALSE | 0.26204980 | 3.966415e-01 | FALSE |
| ## 179 | 1.00000000 | FALSE | 0.24120414 | 4.046984e-01 | FALSE |
| ## 180 | 1.00000000 | FALSE | 0.23658470 | 4.064895e-01 | FALSE |
| ## 181 | 1.00000000 | FALSE | 0.23471601 | 4.072146e-01 | FALSE |
| ## 182 | 1.00000000 | FALSE | 0.20993639 | 4.168587e-01 | FALSE |
| ## 183 | 1.00000000 | FALSE | 0.20858693 | 4.173854e-01 | FALSE |
| ## 184 | 1.00000000 | FALSE | 0.20858693 | 4.173854e-01 | FALSE |
| ## 185 | 1.00000000 | FALSE | 0.18181024 | 4.278658e-01 | FALSE |
| ## 186 | 1.00000000 | FALSE | 0.18181024 | 4.278658e-01 | FALSE |

```
## 187 1.00000000    FALSE  0.18181024 4.278658e-01 FALSE
## 188 1.00000000    FALSE  0.16729065 4.335707e-01 FALSE
## 189 1.00000000    FALSE  0.16729065 4.335707e-01 FALSE
## 190 1.00000000    FALSE  0.16729065 4.335707e-01 FALSE
## 191 1.00000000    FALSE  0.16610281 4.340380e-01 FALSE
## 192 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 193 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 194 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 195 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 196 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 197 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 198 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 199 1.00000000    FALSE  0.14844744 4.409948e-01 FALSE
## 200 1.00000000    FALSE  0.11209270 4.553750e-01 FALSE
## 201 1.00000000    FALSE  0.10554840 4.579703e-01 FALSE
## 202 1.00000000    FALSE  0.09312800 4.629009e-01 FALSE
## 203 1.00000000    FALSE  0.08527854 4.660200e-01 FALSE
## 204 1.00000000    FALSE  0.08527854 4.660200e-01 FALSE
## 205 1.00000000    FALSE  0.07570105 4.698285e-01 FALSE
## 206 1.00000000    FALSE  0.07557288 4.698795e-01 FALSE
## 207 1.00000000    FALSE  0.06759444 4.730542e-01 FALSE
## 208 1.00000000    FALSE  0.05853850 4.766598e-01 FALSE
## 209 1.00000000    FALSE  0.05853850 4.766598e-01 FALSE
## 210 1.00000000    FALSE  0.05853850 4.766598e-01 FALSE
## 211 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 212 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 213 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 214 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 215 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 216 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 217 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 218 1.00000000    FALSE  0.04779649 4.809392e-01 FALSE
## 219 1.00000000    FALSE  0.02148843 4.914280e-01 FALSE
```

```
##
```

```
##
```

```
## Summary statistics:
```

```
##
```

```
## Total Chi squared      = 989.2796
```

```
## Total degrees of freedom = 475
```

```
## p                      = 0
```

```
## Sum of counts          = 2138
```

```
##
```

```
## Levels:
```

```
##
```

```
##      Data Variety      X
```

```
##      2      2      159
```

```
# save
```

```
cfawords <- cfa::cfa(configs, counts)
```

```
cfawords <- cfawords$table %>%
```

```
  as.data.frame() %>%
```

```
  dplyr::filter(sig.z == TRUE,
```

```
    str_detect("control", label) == F)
```

```
write.table(cfawords, here::here("tables", "cfawords.txt"), sep = "\t", row.names = F)
```

```

r8 <- cfa::cfa(configs, counts)
# save to disc
config_words <- r8$table %>%
  as.data.frame() %>%
  dplyr::mutate(expected = round(expected, 1),
                Q = round(Q, 3),
                chisq = round(chisq, 3),
                z = round(z, 3),
                p.z = round(p.z, 3),
                label = stringr::str_remove_all(label, "Xmuch ")) %>%
  dplyr::select(-p.chisq, -sig.chisq)
# save to disc
write.table(config_words,
            here::here("tables", "xmuch_wordconfigs.txt"),
            sep = "\t",
            row.names = F)
# save to disc
words <- r8$table %>%
  as.data.frame() %>%
  dplyr::filter(stringr::str_detect(label, "control", negate = TRUE),
                sig.z == T) %>%
  dplyr::pull(label) %>%
  stringr::str_remove_all("Xmuch ")
words

```

```

## [1] "Australia jealous"      "Australia threatened" "Australia paranoid"
## [4] "Australia other"        "Ireland excuses"      "Ireland coincidence"
## [7] "Australia hypocrite"    "Australia generalise" "Australia entitled"
## [10] "Australia creepy"       "Ireland agenda"      "Australia agenda"

```

The words that are attracted to the X-much structure in Ireland are less negative compared to the words that are attracted to the x-much construction in Oz. Compare *excuses*, *coincidence*, or *agenda* (Ireland) to *jealous*, *paranoid*, *hypocrite*, *creepy*, *threatened*, or *entitled* (Australia).

Outro

```
sessionInfo()
```

```

## R version 4.3.2 (2023-10-31 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 11 x64 (build 22621)
##
## Matrix products: default
##
##
## locale:
## [1] LC_COLLATE=English_Australia.utf8  LC_CTYPE=English_Australia.utf8
## [3] LC_MONETARY=English_Australia.utf8 LC_NUMERIC=C
## [5] LC_TIME=English_Australia.utf8
##

```

```

## time zone: Australia/Brisbane
## tzcode source: internal
##
## attached base packages:
## [1] grid      stats      graphics  grDevices utils      datasets  methods
## [8] base
##
## other attached packages:
## [1] epitools_0.5-10.1 effectsize_0.8.6 vcd_1.4-12      cfa_0.10-1
## [5] tidytext_0.4.1    quanteda_3.3.1    flextable_0.9.4 readxl_1.4.3
## [9] here_1.0.1        lubridate_1.9.3   forcats_1.0.0   stringr_1.5.1
## [13] dplyr_1.1.4       purrr_1.0.2       readr_2.1.5     tidyr_1.3.0
## [17] tibble_3.2.1      ggplot2_3.5.0     tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] sandwich_3.1-0      rlang_1.1.3        magrittr_2.0.3
## [4] multcomp_1.4-25     compiler_4.3.2     systemfonts_1.0.5
## [7] vctr_0.6.5          httpcode_0.3.0     pkgconfig_2.0.3
## [10] crayon_1.5.2        fastmap_1.1.1      ellipsis_0.3.2
## [13] utf8_1.2.4          promises_1.2.1     rmarkdown_2.25
## [16] tzdb_0.4.0          ragg_1.2.7         xfun_0.41
## [19] jsonlite_1.8.8      SnowballC_0.7.1    highr_0.10
## [22] later_1.3.2         uuid_1.2-0         stopwords_2.3
## [25] R6_2.5.1            RColorBrewer_1.1-3 stringi_1.8.3
## [28] lmtest_0.9-40       cellranger_1.1.0   estimability_1.4.1
## [31] Rcpp_1.0.12         knitr_1.45         zoo_1.8-12
## [34] parameters_0.21.3   httpuv_1.6.13      Matrix_1.6-5
## [37] splines_4.3.2       timechange_0.3.0   tidyselect_1.2.0
## [40] rstudioapi_0.15.0   yaml_2.3.8         codetools_0.2-19
## [43] curl_5.2.0          lattice_0.21-9     shiny_1.8.0
## [46] withr_2.5.2         bayestestR_0.13.1  askpass_1.2.0
## [49] coda_0.19-4.1       evaluate_0.23      survival_3.5-7
## [52] RcppParallel_5.1.7 zip_2.3.0          xml2_1.3.6
## [55] pillar_1.9.0        janeaustenr_1.0.0  insight_0.19.8
## [58] generics_0.1.3      rprojroot_2.0.4    hms_1.1.3
## [61] munsell_0.5.0       scales_1.3.0       xtable_1.8-4
## [64] glue_1.7.0          gdtools_0.3.5      emmeans_1.9.0
## [67] tools_4.3.2         gfonts_0.2.0       data.table_1.14.10
## [70] tokenizers_0.3.0    mvtnorm_1.2-4      fastmatch_1.1-4
## [73] datawizard_0.9.1    colorspace_2.1-0   cli_3.6.2
## [76] textshaping_0.3.7   officer_0.6.4      fontBitstreamVera_0.1.1
## [79] fansi_1.0.6         gtable_0.3.4       digest_0.6.33
## [82] fontquiver_0.2.1    crul_1.4.0         TH.data_1.1-2
## [85] htmltools_0.5.7     lifecycle_1.0.4    mime_0.12
## [88] fontLiberation_0.1.0 openssl_2.1.1      MASS_7.3-60

```

References

- Ben-Shachar, Mattan S., Daniel Lüdtke, and Dominique Makowski. 2020. “effectsize: Estimation of Effect Size Indices and Standardized Parameters.” *Journal of Open Source Software* 5 (56): 2815. <https://doi.org/10.21105/joss.02815>.
- Cochran, William G. 1952. “The X² Test of Goodness of Fit.” *The Annals of Mathematical Statistics* 23 (3): 315–45.

- Lienert, Gustav A., and Joachim Krauth. 1975. "Configural Frequency Analysis as a Statistical Tool for Defining Types." *Educational and Psychological Measurement* 35 (2): 231–38.
- Mair, Patrick, and Stefan Funke. 2017. *Cfa: Configural Frequency Analysis (CFA)*. <https://CRAN.R-project.org/package=cfa>.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.