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

Anonymous

2022-01-02

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](#ref-tidyverse)) and the statistical analysis uses 2-tests (see Cochran [1952](#ref-cochran1952x2)) and configural frequency analysis (see Lienert and Krauth [1975](#ref-lienert1975cfa)) using the *cfa* package (Mair and Funke [2017](#ref-cfa)). Effect sizes are calculated using the *effectsize* package (Ben-Shachar, Lüdecke, and Makowski [2020](#ref-effectsize)).

# 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")

* load packages
* set options

# load packages  
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5 v purrr 0.3.4  
## v tibble 3.1.6 v dplyr 1.0.7  
## v tidyr 1.1.4 v stringr 1.4.0  
## v readr 2.1.1 v forcats 0.5.1

## Warning: Paket 'tibble' wurde unter R Version 4.1.2 erstellt

## Warning: Paket 'readr' wurde unter R Version 4.1.2 erstellt

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(here)

## here() starts at D:/Uni/Projekte/XmuchIrE

library(readxl)  
library(flextable)

## Warning: Paket 'flextable' wurde unter R Version 4.1.2 erstellt

##   
## Attache Paket: 'flextable'

## Das folgende Objekt ist maskiert 'package:purrr':  
##   
## compose

library(quanteda)

## Warning: Paket 'quanteda' wurde unter R Version 4.1.2 erstellt

## Package version: 3.2.0  
## Unicode version: 13.0  
## ICU version: 69.1

## Parallel computing: 8 of 8 threads used.

## See https://quanteda.io for tutorials and examples.

library(tidytext)  
library(cfa)  
library(vcd)

## Lade nötiges Paket: grid

library(effectsize)

## Warning: Paket 'effectsize' wurde unter R Version 4.1.2 erstellt

##   
## Attache Paket: 'effectsize'

## Das folgende Objekt ist maskiert 'package:vcd':  
##   
## oddsratio

# 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  
## kwic  
## 1 lessons . Evi 's first day at home and her first collar ( 8 weeks old ) ./. Cute << much >> ? Evi first seizure . In comparison to what  
## 2 a ring fire with cotton balls and made ' marshmallows on sticks ' cotton balls on pencil tips . Creative << much >> ? ?/? So , we sat fireside &; roasted marshmallows &; all crammed into a play tent to sleep &;  
## 3 the skies will fall if we have a carbon tax . " Skies will fall " ?/? Hype , << much >> ? I do n't think anyone here asserted that . It 's a supid policy that will distort the economy  
## 4 the sleek uniform of the bowl cut which was styled for Lumley by John Frida in 1976 . Twins << much >> ? With the hair cut I became fixed on the clothes  
## 5 powder served alongside it too much lately , I even sneak spoonfuls out of the fridge - pregnancy craving << much >> ? It 's just 1 macadamia nuts shaved on the micro-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 .../... feasible << much >> ? !/! ) ,/, I do n't see any other solution . As someone who has suffered from depression for  
## 7 teach . Can you listen ? Can you hear them ? Airy fairy << much >> ? .../... here 's the biggest joke . The very people who claim they want connection with the  
## 8 Striesand Effect . Do they think anyone would have seen this piddly little video absent their protests ? Ironic , << much >> ? LE 6- A reasonable point of view . But what should we do from a public policy perspective  
## 9 was so lovely to see these girls , hear all about Sharon 's trip to France and Italy ( jealous << much >> ? )/) and talk just a little bit about photography . Must get some practice in before our next catchup  
## 10 Tweeting Without Getting Sued , which I 'll be buying as soon as it comes out next month ( obsessed << much >> ? )/) . If you have any questions , I can try to answer them or at least point  
## 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 ) ./. Cute << much >> ? Evi first seizure"| \_\_truncated\_\_ "a ring fire with cotton balls and made ' marshmallows on sticks ' cotton balls on pencil tips . Creative << muc"| \_\_truncated\_\_ "the skies will fall if we have a carbon tax . \" Skies will fall \" ?/? Hype , << much >> ? I do n't think anyo"| \_\_truncated\_\_ "the sleek uniform of the bowl cut which was styled for Lumley by John Frida in 1976 . Twins << much >> ? With t"| \_\_truncated\_\_ ...  
## $ 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](#ref-stats))

chisq.test(phrase\_x2)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: phrase\_x2  
## X-squared = 2.2635e-30, df = 1, p-value = 1

# effect size  
effectsize::effectsize(chisq.test(phrase\_x2), type = "phi")

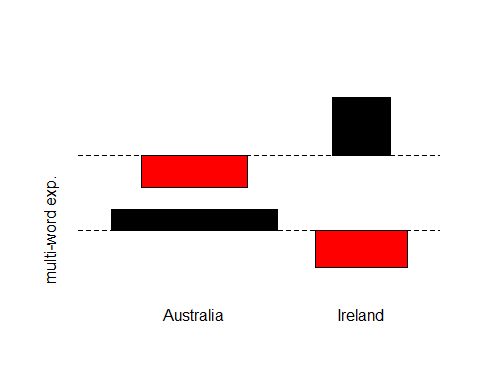
## Phi | 95% CI  
## -------------------  
## 0.02 | [0.00, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

generate visualization

# open window  
png(here::here("images", "multiword\_assoc.png"), width = 500, height = 450)  
# generate plot  
assocplot(phrase\_x2)  
# close window  
dev.off()

## png   
## 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 26 (2.6) 904 (90.4) 70 (7)   
## 2 Australia Xmuch 51 (46.8) 49 (45) 9 (8.3)   
## 3 Ireland control 26 (2.6) 927 (92.7) 47 (4.7)  
## 4 Ireland Xmuch 14 (42.4) 18 (54.5) 1 (3)

Calculate totals

# au  
austot <- 26+904+70+51+49+9  
negau <- round(77/austot\*100, 1)  
nonau <- round(953/austot\*100, 1)  
posau <- round(79/austot\*100, 1)  
# ire  
iretot <- 26+927+47+14+18+1  
negire <- round(40/iretot\*100, 1)  
nonire <- round(945/iretot\*100, 1)  
posire <- round(48/iretot\*100, 1)  
# results  
negau

## [1] 6.9

nonau

## [1] 85.9

posau

## [1] 7.1

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 control 169  
## 2 emotional Xmuch 75  
## 3 nonemotional control 1831  
## 4 nonemotional Xmuch 67

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 169 75  
## nonemotional 1831 67

# perform x2 test  
chisq.test(emo\_x2)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: emo\_x2  
## X-squared = 254.19, df = 1, p-value < 2.2e-16

# effect size  
effectsize::effectsize(chisq.test(emo\_x2), type = "phi")

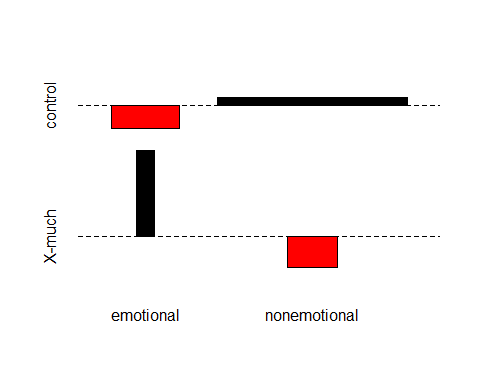
## Phi | 95% CI  
## -------------------  
## 0.35 | [0.31, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

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()

## png   
## 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  
chisq.test(emo\_x22)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: emo\_x22  
## X-squared = 0.58981, df = 1, p-value = 0.4425

# effect size  
effectsize::effectsize(chisq.test(emo\_x22), type = "phi")

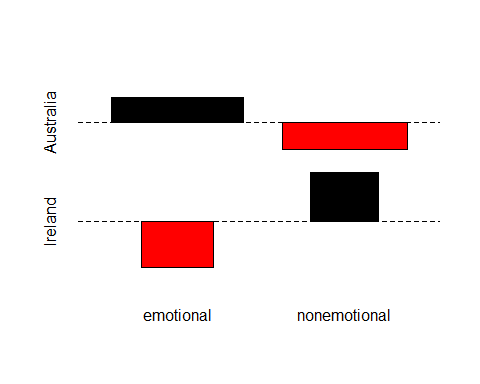
## Phi | 95% CI  
## -------------------  
## 0.08 | [0.00, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

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()

## png   
## 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 control 96  
## 2 emotional Australia Xmuch 60  
## 3 emotional Ireland control 73  
## 4 emotional Ireland Xmuch 15  
## 5 nonemotional Australia control 904  
## 6 nonemotional Australia Xmuch 49  
## 7 nonemotional Ireland control 927  
## 8 nonemotional Ireland Xmuch 18

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 8.374729 0.024196034 318.2393642  
## 2 Ireland nonemotional Xmuch 18 60.680052 0.020506242 30.0195328  
## 3 Ireland emotional control 73 109.870528 0.018143789 12.3730708  
## 4 Ireland emotional Xmuch 15 7.800807 0.003373252 6.6439754  
## 5 Ireland nonemotional control 927 854.648613 0.056201739 6.1250005  
## 6 Australia emotional control 96 117.953935 0.010846559 4.0861314  
## 7 Australia nonemotional Xmuch 49 65.144412 0.007773488 4.0009882  
## 8 Australia nonemotional control 904 917.526924 0.011047138 0.1994248  
## p.chisq sig.chisq z p.z sig.z  
## 1 0.000000e+00 TRUE 17.7011258 0.0000000000 TRUE  
## 2 4.277161e-08 TRUE -5.6234197 0.9999999906 TRUE  
## 3 4.355706e-04 TRUE -3.6603505 0.9998740648 TRUE  
## 4 9.949168e-03 FALSE 2.4029511 0.0081316805 FALSE  
## 5 1.332833e-02 FALSE 3.1703159 0.0007613665 TRUE  
## 6 4.323650e-02 FALSE -2.1268447 0.9832835074 FALSE  
## 7 4.547359e-02 FALSE -2.0942882 0.9818828428 FALSE  
## 8 6.551855e-01 FALSE -0.6124747 0.7298881262 FALSE  
##   
##   
## Summary statistics:  
##   
## Total Chi squared = 381.6875   
## Total degrees of freedom = 4   
## p = 0   
## Sum of counts = 2142   
##   
## Levels:  
##   
## Variety Emotionality Data   
## 2 2 2

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 control 96  
## 2 emotional Australia X-much 60  
## 3 emotional Ireland control 73  
## 4 emotional Ireland X-much 15  
## 5 nonemotional Australia control 904  
## 6 nonemotional Australia X-much 49

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

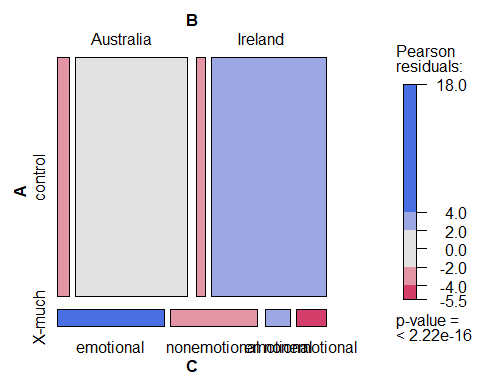
## , , emotional  
##   
## Australia Ireland  
## control 96 73  
## X-much 60 15  
##   
## , , nonemotional  
##   
## Australia Ireland  
## control 904 927  
## X-much 49 18

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()

## png   
## 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 control 52  
## 2 negative Xmuch 65  
## 3 neutral control 1831  
## 4 neutral Xmuch 67  
## 5 positive control 117  
## 6 positive Xmuch 10

Perform x2-test

pol\_x2 <- pol\_tb1 %>%  
 tidyr::spread(Data, Frequency) %>%  
 as.matrix()  
# inspect  
pol\_x2

## Polarity control Xmuch  
## [1,] "negative" " 52" "65"   
## [2,] "neutral" "1831" "67"   
## [3,] "positive" " 117" "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 52 65  
## neutral 1831 67  
## positive 117 10

# perform x2 test  
chisq.test(pol\_x2)

##   
## Pearson's Chi-squared test  
##   
## data: pol\_x2  
## X-squared = 482.24, df = 2, p-value < 2.2e-16

# effect size  
effectsize::effectsize(chisq.test(pol\_x2), type = "phi")

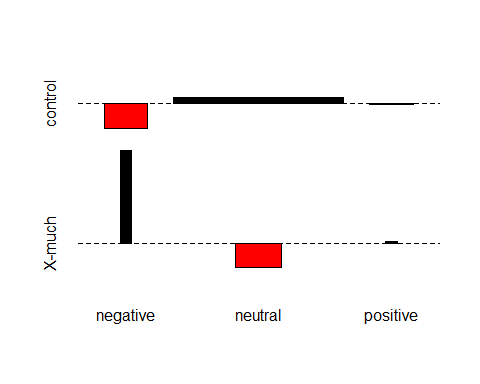
## Phi | 95% CI  
## -------------------  
## 0.47 | [0.44, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

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()

## png   
## 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  
chisq.test(pol\_x22)

## Warning in chisq.test(pol\_x22): Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: pol\_x22  
## X-squared = 1.5819, df = 2, p-value = 0.4534

# effect size  
effectsize::effectsize(chisq.test(pol\_x22), type = "phi")

## Warning in chisq.test(pol\_x22): Chi-squared approximation may be incorrect

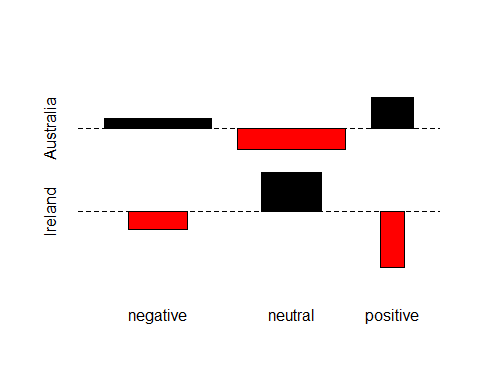
## Phi | 95% CI  
## -------------------  
## 0.11 | [0.00, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

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()

## png   
## 2

assocplot(pol\_x22)



### CFA

pol\_tb1 <- 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\_tb1

## # A tibble: 12 x 4  
## # Groups: Polarity, Variety [6]  
## Polarity Variety Data Frequency  
## <chr> <chr> <chr> <int>  
## 1 negative Australia control 26  
## 2 negative Australia Xmuch 51  
## 3 negative Ireland control 26  
## 4 negative Ireland Xmuch 14  
## 5 neutral Australia control 904  
## 6 neutral Australia Xmuch 49  
## 7 neutral Ireland control 927  
## 8 neutral Ireland Xmuch 18  
## 9 positive Australia control 70  
## 10 positive Australia Xmuch 9  
## 11 positive Ireland control 47  
## 12 positive Ireland Xmuch 1

configs <- pol\_tb1 %>%  
 dplyr::select(Variety, Polarity, Data)  
counts = pol\_tb1$Frequency  
cfa::cfa(configs, counts)

##   
## \*\*\* Analysis of configuration frequencies (CFA) \*\*\*  
##   
## label n expected Q chisq  
## 1 Australia negative Xmuch 51 4.015751 0.021975956 549.7152083  
## 2 Ireland neutral Xmuch 18 60.680052 0.020506242 30.0195328  
## 3 Ireland negative Xmuch 14 3.740551 0.004798037 28.1392468  
## 4 Australia negative control 26 56.559879 0.014653923 16.5118139  
## 5 Ireland negative control 26 52.683819 0.012771556 13.5150829  
## 6 Ireland neutral control 927 854.648613 0.056201739 6.1250005  
## 7 Australia positive Xmuch 9 4.358978 0.002171095 4.9413154  
## 8 Australia neutral Xmuch 49 65.144412 0.007773488 4.0009882  
## 9 Ireland positive Xmuch 1 4.060256 0.001431404 2.3065462  
## 10 Ireland positive control 47 57.186709 0.004886149 1.8145657  
## 11 Australia positive control 70 61.394057 0.004136268 1.2063426  
## 12 Australia neutral control 904 917.526924 0.011047138 0.1994248  
## p.chisq sig.chisq z p.z sig.z  
## 1 0.000000e+00 TRUE 23.4460062 0.000000e+00 TRUE  
## 2 3.029293e-07 TRUE 5.4790084 2.138581e-08 TRUE  
## 3 7.756043e-07 TRUE 5.3046439 5.644658e-08 TRUE  
## 4 2.597199e-04 TRUE 4.0634731 2.417394e-05 TRUE  
## 5 1.162083e-03 TRUE 3.6762866 1.183268e-04 TRUE  
## 6 4.677061e-02 FALSE 2.4748738 6.664163e-03 FALSE  
## 7 8.452925e-02 FALSE 2.2229070 1.311104e-02 FALSE  
## 8 1.352684e-01 FALSE 2.0002470 2.273680e-02 FALSE  
## 9 3.156021e-01 FALSE 1.5187318 6.441501e-02 FALSE  
## 10 4.036194e-01 FALSE 1.3470582 8.898075e-02 FALSE  
## 11 5.470740e-01 FALSE 1.0983363 1.360288e-01 FALSE  
## 12 9.050977e-01 FALSE 0.4465701 3.275927e-01 FALSE  
##   
##   
## Summary statistics:  
##   
## Total Chi squared = 658.4951   
## Total degrees of freedom = 7   
## p = 0   
## Sum of counts = 2142   
##   
## Levels:  
##   
## Variety Polarity Data   
## 2 3 2

### 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 control 26  
## 2 negative Australia Xmuch 51  
## 3 negative Ireland control 26  
## 4 negative Ireland Xmuch 14  
## 5 neutral Australia control 904  
## 6 neutral Australia Xmuch 49  
## 7 neutral Ireland control 927  
## 8 neutral Ireland Xmuch 18  
## 9 positive Australia control 70  
## 10 positive Australia Xmuch 9  
## 11 positive Ireland control 47  
## 12 positive Ireland Xmuch 1

* 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

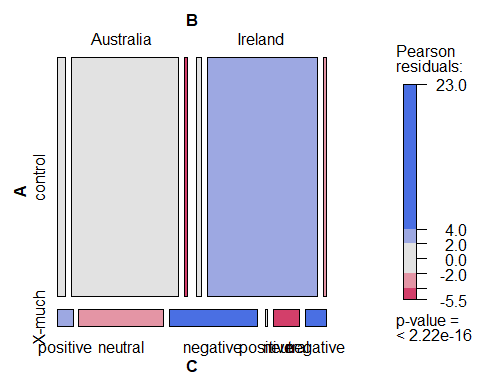
## , , positive  
##   
## Australia Ireland  
## control 70 47  
## X-much 9 1  
##   
## , , neutral  
##   
## Australia Ireland  
## control 904 927  
## X-much 49 18  
##   
## , , negative  
##   
## Australia Ireland  
## control 26 26  
## X-much 51 14

* 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()

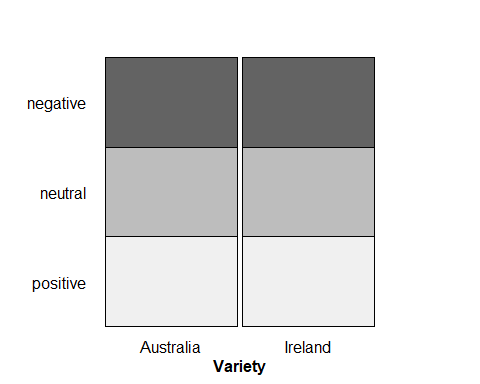
## png   
## 2

# 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 75 (7.5) 322 (32.2) 526 (52.6) 77 (7.7)   
## 2 Australia Xmuch 45 (41.3) 46 (42.2) 2 (1.8) 16 (14.7)  
## 3 Ireland control 79 (7.9) 316 (31.6) 543 (54.3) 62 (6.2)   
## 4 Ireland Xmuch 8 (24.2) 23 (69.7) 1 (3) 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: 8 x 3  
## # Groups: POS [4]  
## POS Data Frequency  
## <chr> <chr> <int>  
## 1 Adjective control 154  
## 2 Adjective Xmuch 53  
## 3 Noun control 638  
## 4 Noun Xmuch 69  
## 5 other control 1069  
## 6 other Xmuch 3  
## 7 Verb control 139  
## 8 Verb Xmuch 17

Perform x2-test

pos\_x2 <- pos\_tb1 %>%  
 tidyr::spread(Data, Frequency) %>%  
 as.matrix()  
# inspect  
pos\_x2

## POS control Xmuch  
## [1,] "Adjective" " 154" "53"   
## [2,] "Noun" " 638" "69"   
## [3,] "other" "1069" " 3"   
## [4,] "Verb" " 139" "17"

rn <- pos\_x2[,1]  
pos\_x2 = pos\_x2[, 2:3]  
pos\_x2 <- apply(pos\_x2, 2, as.numeric)  
rownames(pos\_x2) <- rn  
pos\_x2

## control Xmuch  
## Adjective 154 53  
## Noun 638 69  
## other 1069 3  
## Verb 139 17

# perform x2 test  
chisq.test(pos\_x2)

##   
## Pearson's Chi-squared test  
##   
## data: pos\_x2  
## X-squared = 206.01, df = 3, p-value < 2.2e-16

# effect size  
effectsize::effectsize(chisq.test(pos\_x2), type = "phi")

## Phi | 95% CI  
## -------------------  
## 0.31 | [0.27, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

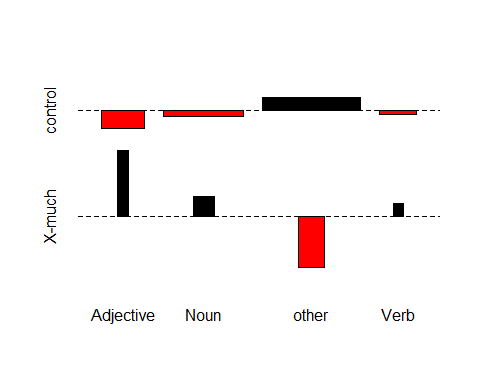
The word classes (parts-of-speech) of words in the x-much structure differ significantly from the word classes in the control data!

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()

## png   
## 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: 8 x 3  
## # Groups: POS [4]  
## POS Variety Frequency  
## <chr> <chr> <int>  
## 1 Adjective Australia 45  
## 2 Adjective Ireland 8  
## 3 Noun Australia 46  
## 4 Noun Ireland 23  
## 5 other Australia 2  
## 6 other Ireland 1  
## 7 Verb Australia 16  
## 8 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" "46" "23"   
## [3,] "other" " 2" " 1"   
## [4,] "Verb" "16" " 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 46 23  
## other 2 1  
## Verb 16 1

# perform x2 test  
chisq.test(pos\_x22)

## Warning in chisq.test(pos\_x22): Chi-squared approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: pos\_x22  
## X-squared = 8.9544, df = 3, p-value = 0.0299

# effect size  
effectsize::effectsize(chisq.test(pos\_x22), type = "phi")

## Warning in chisq.test(pos\_x22): Chi-squared approximation may be incorrect

## Phi | 95% CI  
## -------------------  
## 0.25 | [0.06, 1.00]  
##   
## - One-sided CIs: upper bound fixed at (1).

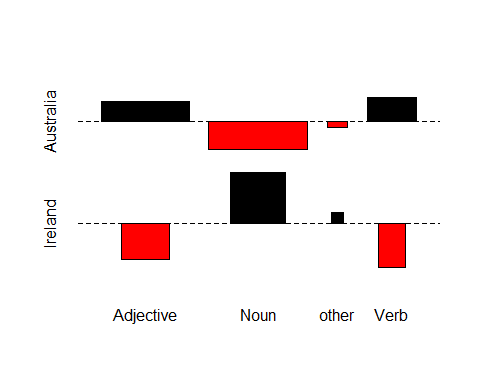
Australia and Ireland differ significantly with respect to the word classes (parts-of-speech) of words in the x-much structure!

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()

## png   
## 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: 16 x 4  
## # Groups: POS, Variety [8]  
## POS Variety Data Frequency  
## <chr> <chr> <chr> <int>  
## 1 Adjective Australia control 75  
## 2 Adjective Australia Xmuch 45  
## 3 Adjective Ireland control 79  
## 4 Adjective Ireland Xmuch 8  
## 5 Noun Australia control 322  
## 6 Noun Australia Xmuch 46  
## 7 Noun Ireland control 316  
## 8 Noun Ireland Xmuch 23  
## 9 other Australia control 526  
## 10 other Australia Xmuch 2  
## 11 other Ireland control 543  
## 12 other Ireland Xmuch 1  
## 13 Verb Australia control 77  
## 14 Verb Australia Xmuch 16  
## 15 Verb Ireland control 62  
## 16 Verb Ireland Xmuch 1

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 7.104791 0.0177503837 2.021237e+02  
## 2 Australia other Xmuch 2 36.793893 0.0165275471 3.290261e+01  
## 3 Ireland other Xmuch 1 34.272400 0.0157859109 3.230158e+01  
## 4 Australia Verb Xmuch 16 5.354335 0.0049824194 2.116606e+01  
## 5 Australia Noun Xmuch 46 24.266122 0.0102627996 1.946588e+01  
## 6 Ireland other control 543 482.709860 0.0363348995 7.530198e+00  
## 7 Australia Adjective control 75 100.067478 0.0122763498 6.279547e+00  
## 8 Ireland Verb Xmuch 1 4.987401 0.0018658765 3.187907e+00  
## 9 Ireland Adjective control 79 93.209833 0.0069357190 2.166288e+00  
## 10 Australia Noun control 322 341.776362 0.0109855029 1.144329e+00  
## 11 Ireland Verb control 62 70.245092 0.0039797620 9.677763e-01  
## 12 Ireland Adjective Xmuch 8 6.617898 0.0006472387 2.886423e-01  
## 13 Australia other control 526 518.223847 0.0047889315 1.166842e-01  
## 14 Australia Verb control 77 75.413172 0.0007678498 3.338971e-02  
## 15 Ireland Noun control 316 318.354357 0.0012910167 1.741141e-02  
## 16 Ireland Noun Xmuch 23 22.603159 0.0001872423 6.967278e-03  
## p.chisq sig.chisq z p.z sig.z  
## 1 0.000000e+00 TRUE 14.21702267 0.000000e+00 TRUE  
## 2 3.376675e-07 TRUE 5.73607940 4.844663e-09 TRUE  
## 3 4.520883e-07 TRUE 5.68344772 6.600304e-09 TRUE  
## 4 9.723848e-05 TRUE 4.60065863 2.105786e-06 TRUE  
## 5 2.189855e-04 TRUE 4.41201566 5.120635e-06 TRUE  
## 6 5.678758e-02 FALSE 2.74412071 3.033661e-03 TRUE  
## 7 9.877406e-02 FALSE 2.50590247 6.106965e-03 FALSE  
## 8 3.635511e-01 FALSE 1.78547100 3.709248e-02 FALSE  
## 9 5.386191e-01 FALSE 1.47183155 7.053319e-02 FALSE  
## 10 7.663852e-01 FALSE 1.06973299 1.423698e-01 FALSE  
## 11 8.090485e-01 FALSE 0.98375621 1.626177e-01 FALSE  
## 12 9.621502e-01 FALSE 0.53725444 2.955459e-01 FALSE  
## 13 9.897627e-01 FALSE 0.34159074 3.663295e-01 FALSE  
## 14 9.983935e-01 FALSE 0.18272851 4.275055e-01 FALSE  
## 15 9.993921e-01 FALSE 0.13195229 4.475110e-01 FALSE  
## 16 9.998457e-01 FALSE 0.08347022 4.667388e-01 FALSE  
##   
##   
## Summary statistics:  
##   
## Total Chi squared = 329.699   
## Total degrees of freedom = 10   
## p = 0   
## Sum of counts = 2142   
##   
## Levels:  
##   
## Variety POS Data   
## 2 4 2

### Visualization

* prepare data

pos\_vis <- pos\_tb2 %>%  
 dplyr::ungroup() %>%  
 dplyr::mutate\_if(is.character, factor)  
# inspect  
pos\_vis

## # A tibble: 16 x 4  
## POS Variety Data Frequency  
## <fct> <fct> <fct> <int>  
## 1 Adjective Australia control 75  
## 2 Adjective Australia Xmuch 45  
## 3 Adjective Ireland control 79  
## 4 Adjective Ireland Xmuch 8  
## 5 Noun Australia control 322  
## 6 Noun Australia Xmuch 46  
## 7 Noun Ireland control 316  
## 8 Noun Ireland Xmuch 23  
## 9 other Australia control 526  
## 10 other Australia Xmuch 2  
## 11 other Ireland control 543  
## 12 other Ireland Xmuch 1  
## 13 Verb Australia control 77  
## 14 Verb Australia Xmuch 16  
## 15 Verb Ireland control 62  
## 16 Verb Ireland Xmuch 1

* 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

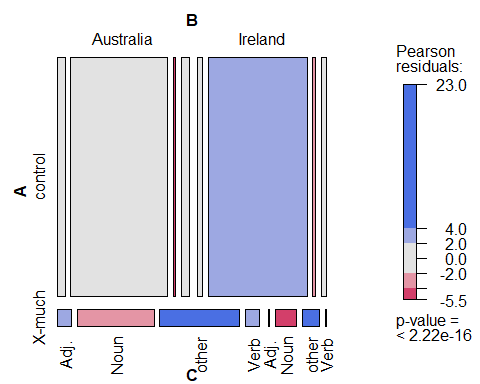
## , , Adj.  
##   
## Australia Ireland  
## control 70 47  
## X-much 9 1  
##   
## , , Noun  
##   
## Australia Ireland  
## control 904 927  
## X-much 49 18  
##   
## , , other  
##   
## Australia Ireland  
## control 26 26  
## X-much 51 14  
##   
## , , Verb  
##   
## Australia Ireland  
## control 70 47  
## X-much 9 1

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()

## png   
## 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 0.4647031 3.055969e-03 91.908365104 1  
## 2 Xmuch paranoid 4 0.2655446 1.746105e-03 52.519065774 1  
## 3 Xmuch agenda 4 0.2655446 1.746105e-03 52.519065774 1  
## 4 Xmuch threatened 3 0.1991585 1.309538e-03 39.389299330 1  
## 5 Xmuch other 104 58.5525947 2.184502e-02 35.275407739 1  
## 6 Xmuch hypocrite 2 0.1327723 8.729984e-04 26.259532887 1  
## 7 Xmuch hipster 2 0.1327723 8.729984e-04 26.259532887 1  
## 8 Xmuch generalise 2 0.1327723 8.729984e-04 26.259532887 1  
## 9 Xmuch excuses 2 0.1327723 8.729984e-04 26.259532887 1  
## 10 Xmuch excited 2 0.1327723 8.729984e-04 26.259532887 1  
## 11 Xmuch entitled 2 0.1327723 8.729984e-04 26.259532887 1  
## 12 Xmuch creepy 2 0.1327723 8.729984e-04 26.259532887 1  
## 13 Xmuch coincidence 2 0.1327723 8.729984e-04 26.259532887 1  
## 14 Xmuch bitter 2 0.1327723 8.729984e-04 26.259532887 1  
## 15 Xmuch biased 2 0.1327723 8.729984e-04 26.259532887 1  
## 16 control other 778 823.4474053 3.454625e-02 2.508316424 1  
## 17 control the 123 114.8345021 4.034007e-03 0.580621283 1  
## 18 control and 53 49.4815334 1.683865e-03 0.250186406 1  
## 19 control a 53 49.4815334 1.683865e-03 0.250186406 1  
## 20 control to 51 47.6143058 1.618876e-03 0.240745410 1  
## 21 control in 51 47.6143058 1.618876e-03 0.240745410 1  
## 22 control of 41 38.2781674 1.295665e-03 0.193540428 1  
## 23 control is 33 30.8092567 1.039158e-03 0.155776442 1  
## 24 control i 23 21.4731183 7.210684e-04 0.108571459 1  
## 25 control for 23 21.4731183 7.210684e-04 0.108571459 1  
## 26 control was 21 19.6058906 6.577868e-04 0.099130463 1  
## 27 control on 17 15.8714353 5.315574e-04 0.080248470 1  
## 28 control are 17 15.8714353 5.315574e-04 0.080248470 1  
## 29 control you 16 14.9378214 5.000694e-04 0.075527972 1  
## 30 control that 15 14.0042076 4.686091e-04 0.070807473 1  
## 31 control as 15 14.0042076 4.686091e-04 0.070807473 1  
## 32 control not 14 13.0705937 4.371764e-04 0.066086975 1  
## 33 control it 14 13.0705937 4.371764e-04 0.066086975 1  
## 34 control with 13 12.1369799 4.057714e-04 0.061366477 1  
## 35 control from 12 11.2033661 3.743938e-04 0.056645979 1  
## 36 control he 11 10.2697522 3.430438e-04 0.051925481 1  
## 37 control be 11 10.2697522 3.430438e-04 0.051925481 1  
## 38 control this 10 9.3361384 3.117213e-04 0.047204982 1  
## 39 control they 10 9.3361384 3.117213e-04 0.047204982 1  
## 40 control there 10 9.3361384 3.117213e-04 0.047204982 1  
## 41 control at 9 8.4025245 2.804262e-04 0.042484484 1  
## 42 control an 9 8.4025245 2.804262e-04 0.042484484 1  
## 43 control which 8 7.4689107 2.491586e-04 0.037763986 1  
## 44 control so 8 7.4689107 2.491586e-04 0.037763986 1  
## 45 control or 8 7.4689107 2.491586e-04 0.037763986 1  
## 46 control no 8 7.4689107 2.491586e-04 0.037763986 1  
## 47 control have 8 7.4689107 2.491586e-04 0.037763986 1  
## 48 control do 8 7.4689107 2.491586e-04 0.037763986 1  
## 49 control would 7 6.5352969 2.179183e-04 0.033043488 1  
## 50 control up 7 6.5352969 2.179183e-04 0.033043488 1  
## 51 control their 7 6.5352969 2.179183e-04 0.033043488 1  
## 52 control some 7 6.5352969 2.179183e-04 0.033043488 1  
## 53 control out 7 6.5352969 2.179183e-04 0.033043488 1  
## 54 control my 7 6.5352969 2.179183e-04 0.033043488 1  
## 55 control if 7 6.5352969 2.179183e-04 0.033043488 1  
## 56 control documentary 7 6.5352969 2.179183e-04 0.033043488 1  
## 57 control can 7 6.5352969 2.179183e-04 0.033043488 1  
## 58 control but 7 6.5352969 2.179183e-04 0.033043488 1  
## 59 control about 7 6.5352969 2.179183e-04 0.033043488 1  
## 60 control who 6 5.6016830 1.867054e-04 0.028322989 1  
## 61 control what 6 5.6016830 1.867054e-04 0.028322989 1  
## 62 control time 6 5.6016830 1.867054e-04 0.028322989 1  
## 63 control may 6 5.6016830 1.867054e-04 0.028322989 1  
## 64 control like 6 5.6016830 1.867054e-04 0.028322989 1  
## 65 control has 6 5.6016830 1.867054e-04 0.028322989 1  
## 66 control by 6 5.6016830 1.867054e-04 0.028322989 1  
## 67 control your 5 4.6680692 1.555198e-04 0.023602491 1  
## 68 control we 5 4.6680692 1.555198e-04 0.023602491 1  
## 69 control way 5 4.6680692 1.555198e-04 0.023602491 1  
## 70 control still 5 4.6680692 1.555198e-04 0.023602491 1  
## 71 control our 5 4.6680692 1.555198e-04 0.023602491 1  
## 72 control much 5 4.6680692 1.555198e-04 0.023602491 1  
## 73 control his 5 4.6680692 1.555198e-04 0.023602491 1  
## 74 control get 5 4.6680692 1.555198e-04 0.023602491 1  
## 75 control any 5 4.6680692 1.555198e-04 0.023602491 1  
## 76 control want 4 3.7344554 1.243614e-04 0.018881993 1  
## 77 control us 4 3.7344554 1.243614e-04 0.018881993 1  
## 78 control think 4 3.7344554 1.243614e-04 0.018881993 1  
## 79 control new 4 3.7344554 1.243614e-04 0.018881993 1  
## 80 control had 4 3.7344554 1.243614e-04 0.018881993 1  
## 81 control community 4 3.7344554 1.243614e-04 0.018881993 1  
## 82 control being 4 3.7344554 1.243614e-04 0.018881993 1  
## 83 control before 4 3.7344554 1.243614e-04 0.018881993 1  
## 84 control years 3 2.8008415 9.323030e-05 0.014161495 1  
## 85 control within 3 2.8008415 9.323030e-05 0.014161495 1  
## 86 control will 3 2.8008415 9.323030e-05 0.014161495 1  
## 87 control were 3 2.8008415 9.323030e-05 0.014161495 1  
## 88 control very 3 2.8008415 9.323030e-05 0.014161495 1  
## 89 control two 3 2.8008415 9.323030e-05 0.014161495 1  
## 90 control though 3 2.8008415 9.323030e-05 0.014161495 1  
## 91 control these 3 2.8008415 9.323030e-05 0.014161495 1  
## 92 control state 3 2.8008415 9.323030e-05 0.014161495 1  
## 93 control services 3 2.8008415 9.323030e-05 0.014161495 1  
## 94 control said 3 2.8008415 9.323030e-05 0.014161495 1  
## 95 control rt 3 2.8008415 9.323030e-05 0.014161495 1  
## 96 control people 3 2.8008415 9.323030e-05 0.014161495 1  
## 97 control part 3 2.8008415 9.323030e-05 0.014161495 1  
## 98 control one 3 2.8008415 9.323030e-05 0.014161495 1  
## 99 control more 3 2.8008415 9.323030e-05 0.014161495 1  
## 100 control months 3 2.8008415 9.323030e-05 0.014161495 1  
## 101 control lead 3 2.8008415 9.323030e-05 0.014161495 1  
## 102 control just 3 2.8008415 9.323030e-05 0.014161495 1  
## 103 control into 3 2.8008415 9.323030e-05 0.014161495 1  
## 104 control how 3 2.8008415 9.323030e-05 0.014161495 1  
## 105 control hours 3 2.8008415 9.323030e-05 0.014161495 1  
## 106 control government 3 2.8008415 9.323030e-05 0.014161495 1  
## 107 control god 3 2.8008415 9.323030e-05 0.014161495 1  
## 108 control found 3 2.8008415 9.323030e-05 0.014161495 1  
## 109 control early 3 2.8008415 9.323030e-05 0.014161495 1  
## 110 control different 3 2.8008415 9.323030e-05 0.014161495 1  
## 111 control day 3 2.8008415 9.323030e-05 0.014161495 1  
## 112 control business 3 2.8008415 9.323030e-05 0.014161495 1  
## 113 control bored 3 2.8008415 9.323030e-05 0.014161495 1  
## 114 control also 3 2.8008415 9.323030e-05 0.014161495 1  
## 115 control all 3 2.8008415 9.323030e-05 0.014161495 1  
## 116 control after 3 2.8008415 9.323030e-05 0.014161495 1  
## 117 control 2011 3 2.8008415 9.323030e-05 0.014161495 1  
## 118 control young 2 1.8672277 6.212638e-05 0.009440996 1  
## 119 control world 2 1.8672277 6.212638e-05 0.009440996 1  
## 120 control work 2 1.8672277 6.212638e-05 0.009440996 1  
## 121 control while 2 1.8672277 6.212638e-05 0.009440996 1  
## 122 control where 2 1.8672277 6.212638e-05 0.009440996 1  
## 123 control well 2 1.8672277 6.212638e-05 0.009440996 1  
## 124 control ways 2 1.8672277 6.212638e-05 0.009440996 1  
## 125 control water 2 1.8672277 6.212638e-05 0.009440996 1  
## 126 control vulnerable 2 1.8672277 6.212638e-05 0.009440996 1  
## 127 control unless 2 1.8672277 6.212638e-05 0.009440996 1  
## 128 control unique 2 1.8672277 6.212638e-05 0.009440996 1  
## 129 control types 2 1.8672277 6.212638e-05 0.009440996 1  
## 130 control turn 2 1.8672277 6.212638e-05 0.009440996 1  
## 131 control town 2 1.8672277 6.212638e-05 0.009440996 1  
## 132 control total 2 1.8672277 6.212638e-05 0.009440996 1  
## 133 control too 2 1.8672277 6.212638e-05 0.009440996 1  
## 134 control those 2 1.8672277 6.212638e-05 0.009440996 1  
## 135 control thanks 2 1.8672277 6.212638e-05 0.009440996 1  
## 136 control system 2 1.8672277 6.212638e-05 0.009440996 1  
## 137 control such 2 1.8672277 6.212638e-05 0.009440996 1  
## 138 control stay 2 1.8672277 6.212638e-05 0.009440996 1  
## 139 control school 2 1.8672277 6.212638e-05 0.009440996 1  
## 140 control return 2 1.8672277 6.212638e-05 0.009440996 1  
## 141 control really 2 1.8672277 6.212638e-05 0.009440996 1  
## 142 control race 2 1.8672277 6.212638e-05 0.009440996 1  
## 143 control put 2 1.8672277 6.212638e-05 0.009440996 1  
## 144 control position 2 1.8672277 6.212638e-05 0.009440996 1  
## 145 control politics 2 1.8672277 6.212638e-05 0.009440996 1  
## 146 control played 2 1.8672277 6.212638e-05 0.009440996 1  
## 147 control paragraph 2 1.8672277 6.212638e-05 0.009440996 1  
## 148 control over 2 1.8672277 6.212638e-05 0.009440996 1  
## 149 control outing 2 1.8672277 6.212638e-05 0.009440996 1  
## 150 control number 2 1.8672277 6.212638e-05 0.009440996 1  
## 151 control music 2 1.8672277 6.212638e-05 0.009440996 1  
## 152 control most 2 1.8672277 6.212638e-05 0.009440996 1  
## 153 control model 2 1.8672277 6.212638e-05 0.009440996 1  
## 154 control minutes 2 1.8672277 6.212638e-05 0.009440996 1  
## 155 control minimum 2 1.8672277 6.212638e-05 0.009440996 1  
## 156 control memorable 2 1.8672277 6.212638e-05 0.009440996 1  
## 157 control me 2 1.8672277 6.212638e-05 0.009440996 1  
## 158 control many 2 1.8672277 6.212638e-05 0.009440996 1  
## 159 control making 2 1.8672277 6.212638e-05 0.009440996 1  
## 160 control make 2 1.8672277 6.212638e-05 0.009440996 1  
## 161 control made 2 1.8672277 6.212638e-05 0.009440996 1  
## 162 control location 2 1.8672277 6.212638e-05 0.009440996 1  
## 163 control little 2 1.8672277 6.212638e-05 0.009440996 1  
## 164 control life 2 1.8672277 6.212638e-05 0.009440996 1  
## 165 control known 2 1.8672277 6.212638e-05 0.009440996 1  
## 166 control know 2 1.8672277 6.212638e-05 0.009440996 1  
## 167 control james 2 1.8672277 6.212638e-05 0.009440996 1  
## 168 control ireland 2 1.8672277 6.212638e-05 0.009440996 1  
## 169 control include 2 1.8672277 6.212638e-05 0.009440996 1  
## 170 control impressive 2 1.8672277 6.212638e-05 0.009440996 1  
## 171 control important 2 1.8672277 6.212638e-05 0.009440996 1  
## 172 control http 2 1.8672277 6.212638e-05 0.009440996 1  
## 173 control home 2 1.8672277 6.212638e-05 0.009440996 1  
## 174 control her 2 1.8672277 6.212638e-05 0.009440996 1  
## 175 control hair 2 1.8672277 6.212638e-05 0.009440996 1  
## 176 control good 2 1.8672277 6.212638e-05 0.009440996 1  
## 177 control go 2 1.8672277 6.212638e-05 0.009440996 1  
## 178 control further 2 1.8672277 6.212638e-05 0.009440996 1  
## 179 control four 2 1.8672277 6.212638e-05 0.009440996 1  
## 180 control following 2 1.8672277 6.212638e-05 0.009440996 1  
## 181 control first 2 1.8672277 6.212638e-05 0.009440996 1  
## 182 control find 2 1.8672277 6.212638e-05 0.009440996 1  
## 183 control fact 2 1.8672277 6.212638e-05 0.009440996 1  
## 184 control every 2 1.8672277 6.212638e-05 0.009440996 1  
## 185 control else 2 1.8672277 6.212638e-05 0.009440996 1  
## 186 control ear 2 1.8672277 6.212638e-05 0.009440996 1  
## 187 control down 2 1.8672277 6.212638e-05 0.009440996 1  
## 188 control director 2 1.8672277 6.212638e-05 0.009440996 1  
## 189 control didn 2 1.8672277 6.212638e-05 0.009440996 1  
## 190 control did 2 1.8672277 6.212638e-05 0.009440996 1  
## 191 control current 2 1.8672277 6.212638e-05 0.009440996 1  
## 192 control country 2 1.8672277 6.212638e-05 0.009440996 1  
## 193 control cork 2 1.8672277 6.212638e-05 0.009440996 1  
## 194 control computer 2 1.8672277 6.212638e-05 0.009440996 1  
## 195 control common 2 1.8672277 6.212638e-05 0.009440996 1  
## 196 control comment 2 1.8672277 6.212638e-05 0.009440996 1  
## 197 control come 2 1.8672277 6.212638e-05 0.009440996 1  
## 198 control church 2 1.8672277 6.212638e-05 0.009440996 1  
## 199 control brand 2 1.8672277 6.212638e-05 0.009440996 1  
## 200 control bill 2 1.8672277 6.212638e-05 0.009440996 1  
## 201 control begin 2 1.8672277 6.212638e-05 0.009440996 1  
## 202 control been 2 1.8672277 6.212638e-05 0.009440996 1  
## 203 control because 2 1.8672277 6.212638e-05 0.009440996 1  
## 204 control beautiful 2 1.8672277 6.212638e-05 0.009440996 1  
## 205 control bear 2 1.8672277 6.212638e-05 0.009440996 1  
## 206 control bank 2 1.8672277 6.212638e-05 0.009440996 1  
## 207 control ball 2 1.8672277 6.212638e-05 0.009440996 1  
## 208 control back 2 1.8672277 6.212638e-05 0.009440996 1  
## 209 control australia 2 1.8672277 6.212638e-05 0.009440996 1  
## 210 control ask 2 1.8672277 6.212638e-05 0.009440996 1  
## 211 control another 2 1.8672277 6.212638e-05 0.009440996 1  
## 212 control aid 2 1.8672277 6.212638e-05 0.009440996 1  
## 213 control adam 2 1.8672277 6.212638e-05 0.009440996 1  
## 214 control act 2 1.8672277 6.212638e-05 0.009440996 1  
## 215 control academic 2 1.8672277 6.212638e-05 0.009440996 1  
## 216 control 4 2 1.8672277 6.212638e-05 0.009440996 1  
## sig.chisq z p.z sig.z  
## 1 FALSE 9.58688506 0.000000e+00 TRUE  
## 2 FALSE 7.24700392 2.130518e-13 TRUE  
## 3 FALSE 7.24700392 2.130518e-13 TRUE  
## 4 FALSE 6.27608949 1.735972e-10 TRUE  
## 5 FALSE 5.93931038 1.431118e-09 TRUE  
## 6 FALSE 5.12440561 1.492389e-07 TRUE  
## 7 FALSE 5.12440561 1.492389e-07 TRUE  
## 8 FALSE 5.12440561 1.492389e-07 TRUE  
## 9 FALSE 5.12440561 1.492389e-07 TRUE  
## 10 FALSE 5.12440561 1.492389e-07 TRUE  
## 11 FALSE 5.12440561 1.492389e-07 TRUE  
## 12 FALSE 5.12440561 1.492389e-07 TRUE  
## 13 FALSE 5.12440561 1.492389e-07 TRUE  
## 14 FALSE 5.12440561 1.492389e-07 TRUE  
## 15 FALSE 5.12440561 1.492389e-07 TRUE  
## 16 FALSE 1.58376653 5.662343e-02 FALSE  
## 17 FALSE 0.76198509 2.230345e-01 FALSE  
## 18 FALSE 0.50018637 3.084719e-01 FALSE  
## 19 FALSE 0.50018637 3.084719e-01 FALSE  
## 20 FALSE 0.49065814 3.118341e-01 FALSE  
## 21 FALSE 0.49065814 3.118341e-01 FALSE  
## 22 FALSE 0.43993230 3.299931e-01 FALSE  
## 23 FALSE 0.39468524 3.465376e-01 FALSE  
## 24 FALSE 0.32950184 3.708882e-01 FALSE  
## 25 FALSE 0.32950184 3.708882e-01 FALSE  
## 26 FALSE 0.31484991 3.764378e-01 FALSE  
## 27 FALSE 0.28328161 3.884805e-01 FALSE  
## 28 FALSE 0.28328161 3.884805e-01 FALSE  
## 29 FALSE 0.27482353 3.917259e-01 FALSE  
## 30 FALSE 0.26609674 3.950824e-01 FALSE  
## 31 FALSE 0.26609674 3.950824e-01 FALSE  
## 32 FALSE 0.25707387 3.985609e-01 FALSE  
## 33 FALSE 0.25707387 3.985609e-01 FALSE  
## 34 FALSE 0.24772258 4.021745e-01 FALSE  
## 35 FALSE 0.23800416 4.059389e-01 FALSE  
## 36 FALSE 0.22787163 4.098730e-01 FALSE  
## 37 FALSE 0.22787163 4.098730e-01 FALSE  
## 38 FALSE 0.21726708 4.140001e-01 FALSE  
## 39 FALSE 0.21726708 4.140001e-01 FALSE  
## 40 FALSE 0.21726708 4.140001e-01 FALSE  
## 41 FALSE 0.20611765 4.183495e-01 FALSE  
## 42 FALSE 0.20611765 4.183495e-01 FALSE  
## 43 FALSE 0.19432958 4.229589e-01 FALSE  
## 44 FALSE 0.19432958 4.229589e-01 FALSE  
## 45 FALSE 0.19432958 4.229589e-01 FALSE  
## 46 FALSE 0.19432958 4.229589e-01 FALSE  
## 47 FALSE 0.19432958 4.229589e-01 FALSE  
## 48 FALSE 0.19432958 4.229589e-01 FALSE  
## 49 FALSE 0.18177868 4.278782e-01 FALSE  
## 50 FALSE 0.18177868 4.278782e-01 FALSE  
## 51 FALSE 0.18177868 4.278782e-01 FALSE  
## 52 FALSE 0.18177868 4.278782e-01 FALSE  
## 53 FALSE 0.18177868 4.278782e-01 FALSE  
## 54 FALSE 0.18177868 4.278782e-01 FALSE  
## 55 FALSE 0.18177868 4.278782e-01 FALSE  
## 56 FALSE 0.18177868 4.278782e-01 FALSE  
## 57 FALSE 0.18177868 4.278782e-01 FALSE  
## 58 FALSE 0.18177868 4.278782e-01 FALSE  
## 59 FALSE 0.18177868 4.278782e-01 FALSE  
## 60 FALSE 0.16829435 4.331759e-01 FALSE  
## 61 FALSE 0.16829435 4.331759e-01 FALSE  
## 62 FALSE 0.16829435 4.331759e-01 FALSE  
## 63 FALSE 0.16829435 4.331759e-01 FALSE  
## 64 FALSE 0.16829435 4.331759e-01 FALSE  
## 65 FALSE 0.16829435 4.331759e-01 FALSE  
## 66 FALSE 0.16829435 4.331759e-01 FALSE  
## 67 FALSE 0.15363102 4.389503e-01 FALSE  
## 68 FALSE 0.15363102 4.389503e-01 FALSE  
## 69 FALSE 0.15363102 4.389503e-01 FALSE  
## 70 FALSE 0.15363102 4.389503e-01 FALSE  
## 71 FALSE 0.15363102 4.389503e-01 FALSE  
## 72 FALSE 0.15363102 4.389503e-01 FALSE  
## 73 FALSE 0.15363102 4.389503e-01 FALSE  
## 74 FALSE 0.15363102 4.389503e-01 FALSE  
## 75 FALSE 0.15363102 4.389503e-01 FALSE  
## 76 FALSE 0.13741176 4.453527e-01 FALSE  
## 77 FALSE 0.13741176 4.453527e-01 FALSE  
## 78 FALSE 0.13741176 4.453527e-01 FALSE  
## 79 FALSE 0.13741176 4.453527e-01 FALSE  
## 80 FALSE 0.13741176 4.453527e-01 FALSE  
## 81 FALSE 0.13741176 4.453527e-01 FALSE  
## 82 FALSE 0.13741176 4.453527e-01 FALSE  
## 83 FALSE 0.13741176 4.453527e-01 FALSE  
## 84 FALSE 0.11900208 4.526369e-01 FALSE  
## 85 FALSE 0.11900208 4.526369e-01 FALSE  
## 86 FALSE 0.11900208 4.526369e-01 FALSE  
## 87 FALSE 0.11900208 4.526369e-01 FALSE  
## 88 FALSE 0.11900208 4.526369e-01 FALSE  
## 89 FALSE 0.11900208 4.526369e-01 FALSE  
## 90 FALSE 0.11900208 4.526369e-01 FALSE  
## 91 FALSE 0.11900208 4.526369e-01 FALSE  
## 92 FALSE 0.11900208 4.526369e-01 FALSE  
## 93 FALSE 0.11900208 4.526369e-01 FALSE  
## 94 FALSE 0.11900208 4.526369e-01 FALSE  
## 95 FALSE 0.11900208 4.526369e-01 FALSE  
## 96 FALSE 0.11900208 4.526369e-01 FALSE  
## 97 FALSE 0.11900208 4.526369e-01 FALSE  
## 98 FALSE 0.11900208 4.526369e-01 FALSE  
## 99 FALSE 0.11900208 4.526369e-01 FALSE  
## 100 FALSE 0.11900208 4.526369e-01 FALSE  
## 101 FALSE 0.11900208 4.526369e-01 FALSE  
## 102 FALSE 0.11900208 4.526369e-01 FALSE  
## 103 FALSE 0.11900208 4.526369e-01 FALSE  
## 104 FALSE 0.11900208 4.526369e-01 FALSE  
## 105 FALSE 0.11900208 4.526369e-01 FALSE  
## 106 FALSE 0.11900208 4.526369e-01 FALSE  
## 107 FALSE 0.11900208 4.526369e-01 FALSE  
## 108 FALSE 0.11900208 4.526369e-01 FALSE  
## 109 FALSE 0.11900208 4.526369e-01 FALSE  
## 110 FALSE 0.11900208 4.526369e-01 FALSE  
## 111 FALSE 0.11900208 4.526369e-01 FALSE  
## 112 FALSE 0.11900208 4.526369e-01 FALSE  
## 113 FALSE 0.11900208 4.526369e-01 FALSE  
## 114 FALSE 0.11900208 4.526369e-01 FALSE  
## 115 FALSE 0.11900208 4.526369e-01 FALSE  
## 116 FALSE 0.11900208 4.526369e-01 FALSE  
## 117 FALSE 0.11900208 4.526369e-01 FALSE  
## 118 FALSE 0.09716479 4.612978e-01 FALSE  
## 119 FALSE 0.09716479 4.612978e-01 FALSE  
## 120 FALSE 0.09716479 4.612978e-01 FALSE  
## 121 FALSE 0.09716479 4.612978e-01 FALSE  
## 122 FALSE 0.09716479 4.612978e-01 FALSE  
## 123 FALSE 0.09716479 4.612978e-01 FALSE  
## 124 FALSE 0.09716479 4.612978e-01 FALSE  
## 125 FALSE 0.09716479 4.612978e-01 FALSE  
## 126 FALSE 0.09716479 4.612978e-01 FALSE  
## 127 FALSE 0.09716479 4.612978e-01 FALSE  
## 128 FALSE 0.09716479 4.612978e-01 FALSE  
## 129 FALSE 0.09716479 4.612978e-01 FALSE  
## 130 FALSE 0.09716479 4.612978e-01 FALSE  
## 131 FALSE 0.09716479 4.612978e-01 FALSE  
## 132 FALSE 0.09716479 4.612978e-01 FALSE  
## 133 FALSE 0.09716479 4.612978e-01 FALSE  
## 134 FALSE 0.09716479 4.612978e-01 FALSE  
## 135 FALSE 0.09716479 4.612978e-01 FALSE  
## 136 FALSE 0.09716479 4.612978e-01 FALSE  
## 137 FALSE 0.09716479 4.612978e-01 FALSE  
## 138 FALSE 0.09716479 4.612978e-01 FALSE  
## 139 FALSE 0.09716479 4.612978e-01 FALSE  
## 140 FALSE 0.09716479 4.612978e-01 FALSE  
## 141 FALSE 0.09716479 4.612978e-01 FALSE  
## 142 FALSE 0.09716479 4.612978e-01 FALSE  
## 143 FALSE 0.09716479 4.612978e-01 FALSE  
## 144 FALSE 0.09716479 4.612978e-01 FALSE  
## 145 FALSE 0.09716479 4.612978e-01 FALSE  
## 146 FALSE 0.09716479 4.612978e-01 FALSE  
## 147 FALSE 0.09716479 4.612978e-01 FALSE  
## 148 FALSE 0.09716479 4.612978e-01 FALSE  
## 149 FALSE 0.09716479 4.612978e-01 FALSE  
## 150 FALSE 0.09716479 4.612978e-01 FALSE  
## 151 FALSE 0.09716479 4.612978e-01 FALSE  
## 152 FALSE 0.09716479 4.612978e-01 FALSE  
## 153 FALSE 0.09716479 4.612978e-01 FALSE  
## 154 FALSE 0.09716479 4.612978e-01 FALSE  
## 155 FALSE 0.09716479 4.612978e-01 FALSE  
## 156 FALSE 0.09716479 4.612978e-01 FALSE  
## 157 FALSE 0.09716479 4.612978e-01 FALSE  
## 158 FALSE 0.09716479 4.612978e-01 FALSE  
## 159 FALSE 0.09716479 4.612978e-01 FALSE  
## 160 FALSE 0.09716479 4.612978e-01 FALSE  
## 161 FALSE 0.09716479 4.612978e-01 FALSE  
## 162 FALSE 0.09716479 4.612978e-01 FALSE  
## 163 FALSE 0.09716479 4.612978e-01 FALSE  
## 164 FALSE 0.09716479 4.612978e-01 FALSE  
## 165 FALSE 0.09716479 4.612978e-01 FALSE  
## 166 FALSE 0.09716479 4.612978e-01 FALSE  
## 167 FALSE 0.09716479 4.612978e-01 FALSE  
## 168 FALSE 0.09716479 4.612978e-01 FALSE  
## 169 FALSE 0.09716479 4.612978e-01 FALSE  
## 170 FALSE 0.09716479 4.612978e-01 FALSE  
## 171 FALSE 0.09716479 4.612978e-01 FALSE  
## 172 FALSE 0.09716479 4.612978e-01 FALSE  
## 173 FALSE 0.09716479 4.612978e-01 FALSE  
## 174 FALSE 0.09716479 4.612978e-01 FALSE  
## 175 FALSE 0.09716479 4.612978e-01 FALSE  
## 176 FALSE 0.09716479 4.612978e-01 FALSE  
## 177 FALSE 0.09716479 4.612978e-01 FALSE  
## 178 FALSE 0.09716479 4.612978e-01 FALSE  
## 179 FALSE 0.09716479 4.612978e-01 FALSE  
## 180 FALSE 0.09716479 4.612978e-01 FALSE  
## 181 FALSE 0.09716479 4.612978e-01 FALSE  
## 182 FALSE 0.09716479 4.612978e-01 FALSE  
## 183 FALSE 0.09716479 4.612978e-01 FALSE  
## 184 FALSE 0.09716479 4.612978e-01 FALSE  
## 185 FALSE 0.09716479 4.612978e-01 FALSE  
## 186 FALSE 0.09716479 4.612978e-01 FALSE  
## 187 FALSE 0.09716479 4.612978e-01 FALSE  
## 188 FALSE 0.09716479 4.612978e-01 FALSE  
## 189 FALSE 0.09716479 4.612978e-01 FALSE  
## 190 FALSE 0.09716479 4.612978e-01 FALSE  
## 191 FALSE 0.09716479 4.612978e-01 FALSE  
## 192 FALSE 0.09716479 4.612978e-01 FALSE  
## 193 FALSE 0.09716479 4.612978e-01 FALSE  
## 194 FALSE 0.09716479 4.612978e-01 FALSE  
## 195 FALSE 0.09716479 4.612978e-01 FALSE  
## 196 FALSE 0.09716479 4.612978e-01 FALSE  
## 197 FALSE 0.09716479 4.612978e-01 FALSE  
## 198 FALSE 0.09716479 4.612978e-01 FALSE  
## 199 FALSE 0.09716479 4.612978e-01 FALSE  
## 200 FALSE 0.09716479 4.612978e-01 FALSE  
## 201 FALSE 0.09716479 4.612978e-01 FALSE  
## 202 FALSE 0.09716479 4.612978e-01 FALSE  
## 203 FALSE 0.09716479 4.612978e-01 FALSE  
## 204 FALSE 0.09716479 4.612978e-01 FALSE  
## 205 FALSE 0.09716479 4.612978e-01 FALSE  
## 206 FALSE 0.09716479 4.612978e-01 FALSE  
## 207 FALSE 0.09716479 4.612978e-01 FALSE  
## 208 FALSE 0.09716479 4.612978e-01 FALSE  
## 209 FALSE 0.09716479 4.612978e-01 FALSE  
## 210 FALSE 0.09716479 4.612978e-01 FALSE  
## 211 FALSE 0.09716479 4.612978e-01 FALSE  
## 212 FALSE 0.09716479 4.612978e-01 FALSE  
## 213 FALSE 0.09716479 4.612978e-01 FALSE  
## 214 FALSE 0.09716479 4.612978e-01 FALSE  
## 215 FALSE 0.09716479 4.612978e-01 FALSE  
## 216 FALSE 0.09716479 4.612978e-01 FALSE  
##   
##   
## Summary statistics:  
##   
## Total Chi squared = 542.4691   
## Total degrees of freedom = 214   
## p = 0   
## Sum of counts = 2139   
##   
## Levels:  
##   
## Data X   
## 2 215

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 0.24027382 3.159105e-03 1.901743e+02  
## 2 Xmuch Australia threatened 3 0.10297450 1.353815e-03 8.150326e+01  
## 3 Xmuch Australia paranoid 3 0.10297450 1.353815e-03 8.150326e+01  
## 4 Xmuch Australia other 86 34.29050703 2.455680e-02 7.797702e+01  
## 5 Xmuch Ireland excuses 2 0.06406062 9.046717e-04 5.850492e+01  
## 6 Xmuch Ireland coincidence 2 0.06406062 9.046717e-04 5.850492e+01  
## 7 Xmuch Australia hypocrite 2 0.06864966 9.025291e-04 5.433550e+01  
## 8 Xmuch Australia generalise 2 0.06864966 9.025291e-04 5.433550e+01  
## 9 Xmuch Australia entitled 2 0.06864966 9.025291e-04 5.433550e+01  
## 10 Xmuch Australia creepy 2 0.06864966 9.025291e-04 5.433550e+01  
## 11 Xmuch Ireland agenda 2 0.12812123 8.747621e-04 2.734855e+01  
## 12 Xmuch Australia agenda 2 0.13729933 8.704767e-04 2.527073e+01  
## 13 control Ireland no 7 3.15475980 1.799494e-03 4.686846e+00  
## 14 control Ireland documentary 7 3.15475980 1.799494e-03 4.686846e+00  
## 15 control Ireland the 71 55.43363656 7.467435e-03 4.371203e+00  
## 16 control Ireland on 13 7.66155952 2.503562e-03 3.719732e+00  
## 17 control Australia like 6 2.89778933 1.451597e-03 3.321053e+00  
## 18 control Australia if 6 2.89778933 1.451597e-03 3.321053e+00  
## 19 control Ireland of 26 18.47787885 3.545625e-03 3.062165e+00  
## 20 control Ireland a 32 23.88603852 3.834369e-03 2.756270e+00  
## 21 control Ireland get 4 1.80271989 1.027632e-03 2.678197e+00  
## 22 control Australia he 9 5.31261377 1.727366e-03 2.559346e+00  
## 23 control Australia think 4 1.93185955 9.672940e-04 2.214035e+00  
## 24 control Australia community 4 1.93185955 9.672940e-04 2.214035e+00  
## 25 control Australia on 4 8.21040309 1.975056e-03 2.159150e+00  
## 26 control Australia they 8 4.82964888 1.484824e-03 2.081130e+00  
## 27 control Ireland rt 3 1.35203992 7.705616e-04 2.008648e+00  
## 28 control Ireland part 3 1.35203992 7.705616e-04 2.008648e+00  
## 29 control Ireland new 3 1.35203992 7.705616e-04 2.008648e+00  
## 30 control Ireland bored 3 1.35203992 7.705616e-04 2.008648e+00  
## 31 control Ireland before 3 1.35203992 7.705616e-04 2.008648e+00  
## 32 control Ireland 2011 3 1.35203992 7.705616e-04 2.008648e+00  
## 33 control Ireland he 2 4.95747969 1.385209e-03 1.764341e+00  
## 34 control Australia very 3 1.44889466 7.253066e-04 1.660526e+00  
## 35 control Australia us 3 1.44889466 7.253066e-04 1.660526e+00  
## 36 control Australia services 3 1.44889466 7.253066e-04 1.660526e+00  
## 37 control Australia said 3 1.44889466 7.253066e-04 1.660526e+00  
## 38 control Australia lead 3 1.44889466 7.253066e-04 1.660526e+00  
## 39 control Australia had 3 1.44889466 7.253066e-04 1.660526e+00  
## 40 control Australia government 3 1.44889466 7.253066e-04 1.660526e+00  
## 41 control Australia being 3 1.44889466 7.253066e-04 1.660526e+00  
## 42 control Australia and 32 25.59713905 3.028212e-03 1.601610e+00  
## 43 control Ireland have 6 3.60543978 1.120842e-03 1.590352e+00  
## 44 control Australia other 456 482.48192288 1.597685e-02 1.453510e+00  
## 45 control Ireland they 2 4.50679972 1.173874e-03 1.394347e+00  
## 46 control Ireland years 2 0.90135994 5.135995e-04 1.339099e+00  
## 47 control Ireland world 2 0.90135994 5.135995e-04 1.339099e+00  
## 48 control Ireland work 2 0.90135994 5.135995e-04 1.339099e+00  
## 49 control Ireland within 2 0.90135994 5.135995e-04 1.339099e+00  
## 50 control Ireland well 2 0.90135994 5.135995e-04 1.339099e+00  
## 51 control Ireland vulnerable 2 0.90135994 5.135995e-04 1.339099e+00  
## 52 control Ireland two 2 0.90135994 5.135995e-04 1.339099e+00  
## 53 control Ireland town 2 0.90135994 5.135995e-04 1.339099e+00  
## 54 control Ireland total 2 0.90135994 5.135995e-04 1.339099e+00  
## 55 control Ireland too 2 0.90135994 5.135995e-04 1.339099e+00  
## 56 control Ireland though 2 0.90135994 5.135995e-04 1.339099e+00  
## 57 control Ireland such 2 0.90135994 5.135995e-04 1.339099e+00  
## 58 control Ireland stay 2 0.90135994 5.135995e-04 1.339099e+00  
## 59 control Ireland state 2 0.90135994 5.135995e-04 1.339099e+00  
## 60 control Ireland put 2 0.90135994 5.135995e-04 1.339099e+00  
## 61 control Ireland over 2 0.90135994 5.135995e-04 1.339099e+00  
## 62 control Ireland one 2 0.90135994 5.135995e-04 1.339099e+00  
## 63 control Ireland number 2 0.90135994 5.135995e-04 1.339099e+00  
## 64 control Ireland more 2 0.90135994 5.135995e-04 1.339099e+00  
## 65 control Ireland life 2 0.90135994 5.135995e-04 1.339099e+00  
## 66 control Ireland just 2 0.90135994 5.135995e-04 1.339099e+00  
## 67 control Ireland ireland 2 0.90135994 5.135995e-04 1.339099e+00  
## 68 control Ireland http 2 0.90135994 5.135995e-04 1.339099e+00  
## 69 control Ireland home 2 0.90135994 5.135995e-04 1.339099e+00  
## 70 control Ireland her 2 0.90135994 5.135995e-04 1.339099e+00  
## 71 control Ireland god 2 0.90135994 5.135995e-04 1.339099e+00  
## 72 control Ireland four 2 0.90135994 5.135995e-04 1.339099e+00  
## 73 control Ireland find 2 0.90135994 5.135995e-04 1.339099e+00  
## 74 control Ireland fact 2 0.90135994 5.135995e-04 1.339099e+00  
## 75 control Ireland early 2 0.90135994 5.135995e-04 1.339099e+00  
## 76 control Ireland ear 2 0.90135994 5.135995e-04 1.339099e+00  
## 77 control Ireland different 2 0.90135994 5.135995e-04 1.339099e+00  
## 78 control Ireland didn 2 0.90135994 5.135995e-04 1.339099e+00  
## 79 control Ireland did 2 0.90135994 5.135995e-04 1.339099e+00  
## 80 control Ireland day 2 0.90135994 5.135995e-04 1.339099e+00  
## 81 control Ireland cork 2 0.90135994 5.135995e-04 1.339099e+00  
## 82 control Ireland computer 2 0.90135994 5.135995e-04 1.339099e+00  
## 83 control Ireland brand 2 0.90135994 5.135995e-04 1.339099e+00  
## 84 control Ireland bill 2 0.90135994 5.135995e-04 1.339099e+00  
## 85 control Ireland begin 2 0.90135994 5.135995e-04 1.339099e+00  
## 86 control Ireland been 2 0.90135994 5.135995e-04 1.339099e+00  
## 87 control Ireland bear 2 0.90135994 5.135995e-04 1.339099e+00  
## 88 control Ireland another 2 0.90135994 5.135995e-04 1.339099e+00  
## 89 control Ireland all 2 0.90135994 5.135995e-04 1.339099e+00  
## 90 control Ireland aid 2 0.90135994 5.135995e-04 1.339099e+00  
## 91 control Ireland after 2 0.90135994 5.135995e-04 1.339099e+00  
## 92 control Ireland was 13 9.46427941 1.659545e-03 1.320895e+00  
## 93 control Ireland for 14 10.36563936 1.706566e-03 1.274266e+00  
## 94 control Australia of 15 19.80156040 2.264675e-03 1.164301e+00  
## 95 control Australia will 2 0.96592978 4.834286e-04 1.107018e+00  
## 96 control Australia were 2 0.96592978 4.834286e-04 1.107018e+00  
## 97 control Australia ways 2 0.96592978 4.834286e-04 1.107018e+00  
## 98 control Australia unless 2 0.96592978 4.834286e-04 1.107018e+00  
## 99 control Australia types 2 0.96592978 4.834286e-04 1.107018e+00  
## 100 control Australia turn 2 0.96592978 4.834286e-04 1.107018e+00  
## 101 control Australia these 2 0.96592978 4.834286e-04 1.107018e+00  
## 102 control Australia thanks 2 0.96592978 4.834286e-04 1.107018e+00  
## 103 control Australia system 2 0.96592978 4.834286e-04 1.107018e+00  
## 104 control Australia school 2 0.96592978 4.834286e-04 1.107018e+00  
## 105 control Australia played 2 0.96592978 4.834286e-04 1.107018e+00  
## 106 control Australia people 2 0.96592978 4.834286e-04 1.107018e+00  
## 107 control Australia paragraph 2 0.96592978 4.834286e-04 1.107018e+00  
## 108 control Australia most 2 0.96592978 4.834286e-04 1.107018e+00  
## 109 control Australia months 2 0.96592978 4.834286e-04 1.107018e+00  
## 110 control Australia making 2 0.96592978 4.834286e-04 1.107018e+00  
## 111 control Australia made 2 0.96592978 4.834286e-04 1.107018e+00  
## 112 control Australia location 2 0.96592978 4.834286e-04 1.107018e+00  
## 113 control Australia known 2 0.96592978 4.834286e-04 1.107018e+00  
## 114 control Australia know 2 0.96592978 4.834286e-04 1.107018e+00  
## 115 control Australia into 2 0.96592978 4.834286e-04 1.107018e+00  
## 116 control Australia impressive 2 0.96592978 4.834286e-04 1.107018e+00  
## 117 control Australia important 2 0.96592978 4.834286e-04 1.107018e+00  
## 118 control Australia how 2 0.96592978 4.834286e-04 1.107018e+00  
## 119 control Australia hours 2 0.96592978 4.834286e-04 1.107018e+00  
## 120 control Australia hair 2 0.96592978 4.834286e-04 1.107018e+00  
## 121 control Australia good 2 0.96592978 4.834286e-04 1.107018e+00  
## 122 control Australia further 2 0.96592978 4.834286e-04 1.107018e+00  
## 123 control Australia found 2 0.96592978 4.834286e-04 1.107018e+00  
## 124 control Australia every 2 0.96592978 4.834286e-04 1.107018e+00  
## 125 control Australia else 2 0.96592978 4.834286e-04 1.107018e+00  
## 126 control Australia down 2 0.96592978 4.834286e-04 1.107018e+00  
## 127 control Australia church 2 0.96592978 4.834286e-04 1.107018e+00  
## 128 control Australia business 2 0.96592978 4.834286e-04 1.107018e+00  
## 129 control Australia because 2 0.96592978 4.834286e-04 1.107018e+00  
## 130 control Australia beautiful 2 0.96592978 4.834286e-04 1.107018e+00  
## 131 control Australia ball 2 0.96592978 4.834286e-04 1.107018e+00  
## 132 control Australia back 2 0.96592978 4.834286e-04 1.107018e+00  
## 133 control Australia australia 2 0.96592978 4.834286e-04 1.107018e+00  
## 134 control Australia also 2 0.96592978 4.834286e-04 1.107018e+00  
## 135 control Australia act 2 0.96592978 4.834286e-04 1.107018e+00  
## 136 control Ireland in 28 22.98467857 2.369053e-03 1.094357e+00  
## 137 control Ireland can 5 3.15475980 8.635348e-04 1.079293e+00  
## 138 control Australia that 10 7.24447332 1.292003e-03 1.048099e+00  
## 139 control Ireland at 6 4.05611975 9.100802e-04 9.315974e-01  
## 140 control Australia the 52 59.40468119 3.558924e-03 9.229795e-01  
## 141 control Ireland other 430 450.22929208 1.197162e-02 9.089241e-01  
## 142 control Australia have 2 3.86371910 8.724720e-04 8.989910e-01  
## 143 control Australia a 21 25.59713905 2.174202e-03 8.256269e-01  
## 144 Xmuch Ireland other 27 31.99827802 2.371098e-03 7.807540e-01  
## 145 control Australia their 5 3.38075421 7.578542e-04 7.755538e-01  
## 146 control Australia some 5 3.38075421 7.578542e-04 7.755538e-01  
## 147 control Australia out 5 3.38075421 7.578542e-04 7.755538e-01  
## 148 control Australia about 5 3.38075421 7.578542e-04 7.755538e-01  
## 149 control Ireland are 10 7.66155952 1.096655e-03 7.137325e-01  
## 150 control Australia you 10 7.72743820 1.065793e-03 6.683375e-01  
## 151 control Ireland is 18 14.87243908 1.471705e-03 6.577023e-01  
## 152 control Ireland who 4 2.70407983 6.063363e-04 6.210649e-01  
## 153 control Australia can 2 3.38075421 6.462332e-04 5.639222e-01  
## 154 control Ireland or 5 3.60543978 6.527634e-04 5.394067e-01  
## 155 control Ireland there 6 4.50679972 6.992297e-04 4.947296e-01  
## 156 control Ireland from 7 5.40815966 7.457352e-04 4.685431e-01  
## 157 control Australia to 28 24.63120928 1.592531e-03 4.607468e-01  
## 158 control Ireland that 5 6.76019958 8.251297e-04 4.583153e-01  
## 159 control Australia was 8 10.14226264 1.005824e-03 4.524917e-01  
## 160 control Ireland their 2 3.15475980 5.404040e-04 4.226852e-01  
## 161 control Ireland some 2 3.15475980 5.404040e-04 4.226852e-01  
## 162 control Ireland out 2 3.15475980 5.404040e-04 4.226852e-01  
## 163 control Ireland about 2 3.15475980 5.404040e-04 4.226852e-01  
## 164 control Australia by 4 2.89778933 5.157501e-04 4.192397e-01  
## 165 control Australia at 3 4.34668399 6.305724e-04 4.172279e-01  
## 166 control Australia for 9 11.10819242 9.902769e-04 4.001079e-01  
## 167 control Ireland and 21 23.88603852 1.363839e-03 3.487066e-01  
## 168 control Australia which 5 3.86371910 5.319328e-04 3.341688e-01  
## 169 control Australia do 5 3.86371910 5.319328e-04 3.341688e-01  
## 170 control Australia i 13 11.10819242 8.886349e-04 3.221889e-01  
## 171 control Australia who 2 2.89778933 4.200966e-04 2.781519e-01  
## 172 control Ireland we 3 2.25339986 3.492463e-04 2.473648e-01  
## 173 control Ireland would 4 3.15475980 3.955552e-04 2.264613e-01  
## 174 control Ireland up 4 3.15475980 3.955552e-04 2.264613e-01  
## 175 control Ireland my 4 3.15475980 3.955552e-04 2.264613e-01  
## 176 control Ireland but 4 3.15475980 3.955552e-04 2.264613e-01  
## 177 control Ireland with 7 5.85883964 5.347164e-04 2.222705e-01  
## 178 control Ireland you 6 7.21087955 5.677446e-04 2.033357e-01  
## 179 control Australia or 3 3.86371910 4.043371e-04 1.930810e-01  
## 180 control Ireland by 2 2.70407983 3.294255e-04 1.833261e-01  
## 181 control Australia are 7 8.21040309 5.677873e-04 1.784414e-01  
## 182 control Australia there 4 4.82964888 3.885633e-04 1.425191e-01  
## 183 control Australia your 3 2.41482444 2.737554e-04 1.418035e-01  
## 184 control Australia way 3 2.41482444 2.737554e-04 1.418035e-01  
## 185 control Australia still 3 2.41482444 2.737554e-04 1.418035e-01  
## 186 control Australia our 3 2.41482444 2.737554e-04 1.418035e-01  
## 187 control Australia much 3 2.41482444 2.737554e-04 1.418035e-01  
## 188 control Australia his 3 2.41482444 2.737554e-04 1.418035e-01  
## 189 control Australia any 3 2.41482444 2.737554e-04 1.418035e-01  
## 190 control Australia from 5 5.79557865 3.727753e-04 1.092118e-01  
## 191 control Australia in 23 24.63120928 7.711229e-04 1.080273e-01  
## 192 control Ireland which 3 3.60543978 2.833932e-04 1.016679e-01  
## 193 control Ireland do 3 3.60543978 2.833932e-04 1.016679e-01  
## 194 control Australia an 5 4.34668399 3.059092e-04 9.819481e-02  
## 195 control Australia be 6 5.31261377 3.220079e-04 8.893924e-02  
## 196 control Australia as 8 7.24447332 3.542491e-04 7.879394e-02  
## 197 control Ireland not 7 6.30951961 3.236085e-04 7.556252e-02  
## 198 control Ireland it 7 6.30951961 3.236085e-04 7.556252e-02  
## 199 control Australia we 2 2.41482444 1.940622e-04 7.125956e-02  
## 200 control Australia is 15 15.93784130 4.415319e-04 5.518604e-02  
## 201 control Ireland this 5 4.50679972 2.309538e-04 5.397323e-02  
## 202 control Ireland so 4 3.60543978 1.846851e-04 4.317858e-02  
## 203 control Australia would 3 3.38075421 1.782041e-04 4.288208e-02  
## 204 control Australia up 3 3.38075421 1.782041e-04 4.288208e-02  
## 205 control Australia my 3 3.38075421 1.782041e-04 4.288208e-02  
## 206 control Australia but 3 3.38075421 1.782041e-04 4.288208e-02  
## 207 control Ireland what 3 2.70407983 1.384554e-04 3.238394e-02  
## 208 control Ireland time 3 2.70407983 1.384554e-04 3.238394e-02  
## 209 control Ireland may 3 2.70407983 1.384554e-04 3.238394e-02  
## 210 control Ireland has 3 2.70407983 1.384554e-04 3.238394e-02  
## 211 control Ireland your 2 2.25339986 1.185360e-04 2.849538e-02  
## 212 control Ireland way 2 2.25339986 1.185360e-04 2.849538e-02  
## 213 control Ireland still 2 2.25339986 1.185360e-04 2.849538e-02  
## 214 control Ireland our 2 2.25339986 1.185360e-04 2.849538e-02  
## 215 control Ireland much 2 2.25339986 1.185360e-04 2.849538e-02  
## 216 control Ireland his 2 2.25339986 1.185360e-04 2.849538e-02  
## 217 control Ireland any 2 2.25339986 1.185360e-04 2.849538e-02  
## 218 control Ireland want 2 1.80271989 9.226469e-05 2.158929e-02  
## 219 control Ireland i 10 10.36563936 1.716911e-04 1.289763e-02  
## 220 control Australia with 6 6.27854354 1.305435e-04 1.235740e-02  
## 221 control Ireland as 7 6.76019958 1.124114e-04 8.506293e-03  
## 222 control Australia not 7 6.76150843 1.117979e-04 8.412062e-03  
## 223 control Australia it 7 6.76150843 1.117979e-04 8.412062e-03  
## 224 control Australia this 5 4.82964888 7.978339e-05 6.008616e-03  
## 225 control Australia so 4 3.86371910 6.379785e-05 4.806893e-03  
## 226 control Australia what 3 2.89778933 4.782676e-05 3.605170e-03  
## 227 control Australia time 3 2.89778933 4.782676e-05 3.605170e-03  
## 228 control Australia may 3 2.89778933 4.782676e-05 3.605170e-03  
## 229 control Australia has 3 2.89778933 4.782676e-05 3.605170e-03  
## 230 control Australia want 2 1.93185955 3.187010e-05 2.403446e-03  
## 231 control Ireland an 4 4.05611975 2.627398e-05 7.764628e-04  
## 232 control Ireland be 5 4.95747969 1.991544e-05 3.646967e-04  
## 233 control Ireland to 23 22.98467857 7.237277e-06 1.021315e-05  
## p.chisq sig.chisq z p.z sig.z  
## 1 0.1761959 FALSE 13.790368603 0.000000e+00 TRUE  
## 2 1.0000000 FALSE 9.027915427 0.000000e+00 TRUE  
## 3 1.0000000 FALSE 9.027915427 0.000000e+00 TRUE  
## 4 1.0000000 FALSE 8.830459967 0.000000e+00 TRUE  
## 5 1.0000000 FALSE 7.648850939 1.010303e-14 TRUE  
## 6 1.0000000 FALSE 7.648850939 1.010303e-14 TRUE  
## 7 1.0000000 FALSE 7.371262079 8.448797e-14 TRUE  
## 8 1.0000000 FALSE 7.371262079 8.448797e-14 TRUE  
## 9 1.0000000 FALSE 7.371262079 8.448797e-14 TRUE  
## 10 1.0000000 FALSE 7.371262079 8.448797e-14 TRUE  
## 11 1.0000000 FALSE 5.229584235 8.494584e-08 TRUE  
## 12 1.0000000 FALSE 5.026999782 2.491065e-07 TRUE  
## 13 1.0000000 FALSE 2.164912381 1.519720e-02 FALSE  
## 14 1.0000000 FALSE 2.164912381 1.519720e-02 FALSE  
## 15 1.0000000 FALSE 2.090742179 1.827559e-02 FALSE  
## 16 1.0000000 FALSE 1.928660627 2.688650e-02 FALSE  
## 17 1.0000000 FALSE 1.822375565 3.419901e-02 FALSE  
## 18 1.0000000 FALSE 1.822375565 3.419901e-02 FALSE  
## 19 1.0000000 FALSE 1.749904169 4.006743e-02 FALSE  
## 20 1.0000000 FALSE 1.660201789 4.843693e-02 FALSE  
## 21 1.0000000 FALSE 1.636519934 5.086541e-02 FALSE  
## 22 1.0000000 FALSE 1.599795639 5.482196e-02 FALSE  
## 23 1.0000000 FALSE 1.487963418 6.838027e-02 FALSE  
## 24 1.0000000 FALSE 1.487963418 6.838027e-02 FALSE  
## 25 1.0000000 FALSE 1.469404739 7.086152e-02 FALSE  
## 26 1.0000000 FALSE 1.442612147 7.456488e-02 FALSE  
## 27 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 28 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 29 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 30 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 31 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 32 1.0000000 FALSE 1.417267837 7.820232e-02 FALSE  
## 33 1.0000000 FALSE 1.328285095 9.204197e-02 FALSE  
## 34 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 35 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 36 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 37 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 38 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 39 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 40 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 41 1.0000000 FALSE 1.288614120 9.876614e-02 FALSE  
## 42 1.0000000 FALSE 1.265547227 1.028376e-01 FALSE  
## 43 1.0000000 FALSE 1.261091624 1.036379e-01 FALSE  
## 44 1.0000000 FALSE 1.205615973 1.139828e-01 FALSE  
## 45 1.0000000 FALSE 1.180824916 1.188361e-01 FALSE  
## 46 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 47 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 48 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 49 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 50 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 51 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 52 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 53 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 54 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 55 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 56 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 57 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 58 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 59 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 60 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 61 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 62 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 63 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 64 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 65 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 66 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 67 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 68 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 69 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 70 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 71 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 72 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 73 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 74 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 75 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 76 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 77 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 78 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 79 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 80 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 81 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 82 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 83 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 84 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 85 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 86 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 87 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 88 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 89 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 90 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 91 1.0000000 FALSE 1.157194343 1.235965e-01 FALSE  
## 92 1.0000000 FALSE 1.149301995 1.252157e-01 FALSE  
## 93 1.0000000 FALSE 1.128833719 1.294840e-01 FALSE  
## 94 1.0000000 FALSE 1.079027936 1.402876e-01 FALSE  
## 95 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 96 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 97 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 98 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 99 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 100 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 101 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 102 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 103 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 104 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 105 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 106 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 107 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 108 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 109 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 110 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 111 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 112 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 113 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 114 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 115 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 116 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 117 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 118 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 119 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 120 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 121 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 122 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 123 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 124 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 125 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 126 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 127 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 128 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 129 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 130 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 131 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 132 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 133 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 134 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 135 1.0000000 FALSE 1.052149023 1.463656e-01 FALSE  
## 136 1.0000000 FALSE 1.046115301 1.477539e-01 FALSE  
## 137 1.0000000 FALSE 1.038890457 1.494278e-01 FALSE  
## 138 1.0000000 FALSE 1.023767203 1.529726e-01 FALSE  
## 139 1.0000000 FALSE 0.965192911 1.672241e-01 FALSE  
## 140 1.0000000 FALSE 0.960718227 1.683469e-01 FALSE  
## 141 1.0000000 FALSE 0.953375114 1.702000e-01 FALSE  
## 142 1.0000000 FALSE 0.948151385 1.715262e-01 FALSE  
## 143 1.0000000 FALSE 0.908640157 1.817701e-01 FALSE  
## 144 1.0000000 FALSE 0.883602844 1.884553e-01 FALSE  
## 145 1.0000000 FALSE 0.880655347 1.892522e-01 FALSE  
## 146 1.0000000 FALSE 0.880655347 1.892522e-01 FALSE  
## 147 1.0000000 FALSE 0.880655347 1.892522e-01 FALSE  
## 148 1.0000000 FALSE 0.880655347 1.892522e-01 FALSE  
## 149 1.0000000 FALSE 0.844826892 1.991037e-01 FALSE  
## 150 1.0000000 FALSE 0.817519143 2.068159e-01 FALSE  
## 151 1.0000000 FALSE 0.810988468 2.086861e-01 FALSE  
## 152 1.0000000 FALSE 0.788076712 2.153259e-01 FALSE  
## 153 1.0000000 FALSE 0.750947504 2.263421e-01 FALSE  
## 154 1.0000000 FALSE 0.734443093 2.313394e-01 FALSE  
## 155 1.0000000 FALSE 0.703370149 2.409126e-01 FALSE  
## 156 1.0000000 FALSE 0.684502049 2.468291e-01 FALSE  
## 157 1.0000000 FALSE 0.678783326 2.486376e-01 FALSE  
## 158 1.0000000 FALSE 0.676989845 2.492062e-01 FALSE  
## 159 1.0000000 FALSE 0.672675001 2.505770e-01 FALSE  
## 160 1.0000000 FALSE 0.650142428 2.578001e-01 FALSE  
## 161 1.0000000 FALSE 0.650142428 2.578001e-01 FALSE  
## 162 1.0000000 FALSE 0.650142428 2.578001e-01 FALSE  
## 163 1.0000000 FALSE 0.650142428 2.578001e-01 FALSE  
## 164 1.0000000 FALSE 0.647487231 2.586583e-01 FALSE  
## 165 1.0000000 FALSE 0.645931796 2.591618e-01 FALSE  
## 166 1.0000000 FALSE 0.632540810 2.635168e-01 FALSE  
## 167 1.0000000 FALSE 0.590513810 2.774231e-01 FALSE  
## 168 1.0000000 FALSE 0.578073330 2.816073e-01 FALSE  
## 169 1.0000000 FALSE 0.578073330 2.816073e-01 FALSE  
## 170 1.0000000 FALSE 0.567616831 2.851476e-01 FALSE  
## 171 1.0000000 FALSE 0.527401103 2.989575e-01 FALSE  
## 172 1.0000000 FALSE 0.497357802 3.094684e-01 FALSE  
## 173 1.0000000 FALSE 0.475879496 3.170801e-01 FALSE  
## 174 1.0000000 FALSE 0.475879496 3.170801e-01 FALSE  
## 175 1.0000000 FALSE 0.475879496 3.170801e-01 FALSE  
## 176 1.0000000 FALSE 0.475879496 3.170801e-01 FALSE  
## 177 1.0000000 FALSE 0.471455681 3.186577e-01 FALSE  
## 178 1.0000000 FALSE 0.450927607 3.260209e-01 FALSE  
## 179 1.0000000 FALSE 0.439409814 3.301823e-01 FALSE  
## 180 1.0000000 FALSE 0.428165973 3.342651e-01 FALSE  
## 181 1.0000000 FALSE 0.422423222 3.363581e-01 FALSE  
## 182 1.0000000 FALSE 0.377517033 3.528947e-01 FALSE  
## 183 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 184 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 185 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 186 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 187 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 188 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 189 1.0000000 FALSE 0.376567989 3.532473e-01 FALSE  
## 190 1.0000000 FALSE 0.330472039 3.705217e-01 FALSE  
## 191 1.0000000 FALSE 0.328675109 3.712006e-01 FALSE  
## 192 1.0000000 FALSE 0.318853969 3.749186e-01 FALSE  
## 193 1.0000000 FALSE 0.318853969 3.749186e-01 FALSE  
## 194 1.0000000 FALSE 0.313360511 3.770034e-01 FALSE  
## 195 1.0000000 FALSE 0.298226828 3.827650e-01 FALSE  
## 196 1.0000000 FALSE 0.280702576 3.894693e-01 FALSE  
## 197 1.0000000 FALSE 0.274886368 3.917018e-01 FALSE  
## 198 1.0000000 FALSE 0.274886368 3.917018e-01 FALSE  
## 199 1.0000000 FALSE 0.266944854 3.947558e-01 FALSE  
## 200 1.0000000 FALSE 0.234917085 4.071365e-01 FALSE  
## 201 1.0000000 FALSE 0.232321383 4.081442e-01 FALSE  
## 202 1.0000000 FALSE 0.207794562 4.176947e-01 FALSE  
## 203 1.0000000 FALSE 0.207079887 4.179737e-01 FALSE  
## 204 1.0000000 FALSE 0.207079887 4.179737e-01 FALSE  
## 205 1.0000000 FALSE 0.207079887 4.179737e-01 FALSE  
## 206 1.0000000 FALSE 0.207079887 4.179737e-01 FALSE  
## 207 1.0000000 FALSE 0.179955370 4.285938e-01 FALSE  
## 208 1.0000000 FALSE 0.179955370 4.285938e-01 FALSE  
## 209 1.0000000 FALSE 0.179955370 4.285938e-01 FALSE  
## 210 1.0000000 FALSE 0.179955370 4.285938e-01 FALSE  
## 211 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 212 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 213 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 214 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 215 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 216 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 217 1.0000000 FALSE 0.168805751 4.329747e-01 FALSE  
## 218 1.0000000 FALSE 0.146932944 4.415925e-01 FALSE  
## 219 1.0000000 FALSE 0.113567715 4.547902e-01 FALSE  
## 220 1.0000000 FALSE 0.111163865 4.557432e-01 FALSE  
## 221 1.0000000 FALSE 0.092229569 4.632578e-01 FALSE  
## 222 1.0000000 FALSE 0.091717295 4.634613e-01 FALSE  
## 223 1.0000000 FALSE 0.091717295 4.634613e-01 FALSE  
## 224 1.0000000 FALSE 0.077515262 4.691068e-01 FALSE  
## 225 1.0000000 FALSE 0.069331758 4.723628e-01 FALSE  
## 226 1.0000000 FALSE 0.060043064 4.760607e-01 FALSE  
## 227 1.0000000 FALSE 0.060043064 4.760607e-01 FALSE  
## 228 1.0000000 FALSE 0.060043064 4.760607e-01 FALSE  
## 229 1.0000000 FALSE 0.060043064 4.760607e-01 FALSE  
## 230 1.0000000 FALSE 0.049024956 4.804497e-01 FALSE  
## 231 1.0000000 FALSE 0.027865082 4.888849e-01 FALSE  
## 232 1.0000000 FALSE 0.019097034 4.923818e-01 FALSE  
## 233 1.0000000 FALSE 0.003195803 4.987251e-01 FALSE  
##   
##   
## Summary statistics:  
##   
## Total Chi squared = 1033.014   
## Total degrees of freedom = 520   
## p = 0   
## Sum of counts = 2140   
##   
## Levels:  
##   
## Data Variety X   
## 2 2 174

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 4.1.1 (2021-08-10)  
## Platform: x86\_64-w64-mingw32/x64 (64-bit)  
## Running under: Windows 10 x64 (build 19043)  
##   
## Matrix products: default  
##   
## locale:  
## [1] LC\_COLLATE=German\_Germany.1252 LC\_CTYPE=German\_Germany.1252   
## [3] LC\_MONETARY=German\_Germany.1252 LC\_NUMERIC=C   
## [5] LC\_TIME=German\_Germany.1252   
##   
## attached base packages:  
## [1] grid stats graphics grDevices utils datasets methods   
## [8] base   
##   
## other attached packages:  
## [1] effectsize\_0.5 vcd\_1.4-9 cfa\_0.10-0 tidytext\_0.3.2   
## [5] quanteda\_3.2.0 flextable\_0.6.10 readxl\_1.3.1 here\_1.0.1   
## [9] forcats\_0.5.1 stringr\_1.4.0 dplyr\_1.0.7 purrr\_0.3.4   
## [13] readr\_2.1.1 tidyr\_1.1.4 tibble\_3.1.6 ggplot2\_3.3.5   
## [17] tidyverse\_1.3.1   
##   
## loaded via a namespace (and not attached):  
## [1] fs\_1.5.2 lubridate\_1.8.0 RColorBrewer\_1.1-2 insight\_0.14.5   
## [5] httr\_1.4.2 rprojroot\_2.0.2 SnowballC\_0.7.0 tools\_4.1.1   
## [9] backports\_1.4.1 utf8\_1.2.2 R6\_2.5.1 DBI\_1.1.2   
## [13] colorspace\_2.0-2 withr\_2.4.3 tidyselect\_1.1.1 emmeans\_1.7.1-1   
## [17] compiler\_4.1.1 performance\_0.8.0 cli\_3.1.0 rvest\_1.0.2   
## [21] xml2\_1.3.3 sandwich\_3.0-1 officer\_0.4.1 bayestestR\_0.11.5   
## [25] scales\_1.1.1 mvtnorm\_1.1-3 lmtest\_0.9-39 systemfonts\_1.0.3   
## [29] digest\_0.6.28 rmarkdown\_2.11 base64enc\_0.1-3 pkgconfig\_2.0.3   
## [33] htmltools\_0.5.2 highr\_0.9 dbplyr\_2.1.1 fastmap\_1.1.0   
## [37] rlang\_0.4.12 rstudioapi\_0.13 generics\_0.1.1 zoo\_1.8-9   
## [41] jsonlite\_1.7.2 zip\_2.2.0 tokenizers\_0.2.1 magrittr\_2.0.1   
## [45] parameters\_0.15.0 Matrix\_1.4-0 Rcpp\_1.0.7 munsell\_0.5.0   
## [49] fansi\_0.5.0 gdtools\_0.2.3 lifecycle\_1.0.1 multcomp\_1.4-17   
## [53] stringi\_1.7.6 yaml\_2.2.1 MASS\_7.3-54 crayon\_1.4.2   
## [57] lattice\_0.20-45 splines\_4.1.1 haven\_2.4.3 hms\_1.1.1   
## [61] knitr\_1.37 pillar\_1.6.4 uuid\_1.0-3 estimability\_1.3   
## [65] codetools\_0.2-18 stopwords\_2.3 fastmatch\_1.1-3 reprex\_2.0.1.9000   
## [69] glue\_1.6.0 evaluate\_0.14 data.table\_1.14.2 RcppParallel\_5.1.4  
## [73] modelr\_0.1.8 vctrs\_0.3.8 tzdb\_0.2.0 cellranger\_1.1.0   
## [77] gtable\_0.3.0 datawizard\_0.2.1 assertthat\_0.2.1 xfun\_0.29   
## [81] xtable\_1.8-4 broom\_0.7.10 coda\_0.19-4 janeaustenr\_0.1.5   
## [85] survival\_3.2-13 TH.data\_1.1-0 ellipsis\_0.3.2

# References

Ben-Shachar, Mattan S., Daniel Lüdecke, 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 χ2 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 Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.