Wellington MPhil thesis demo

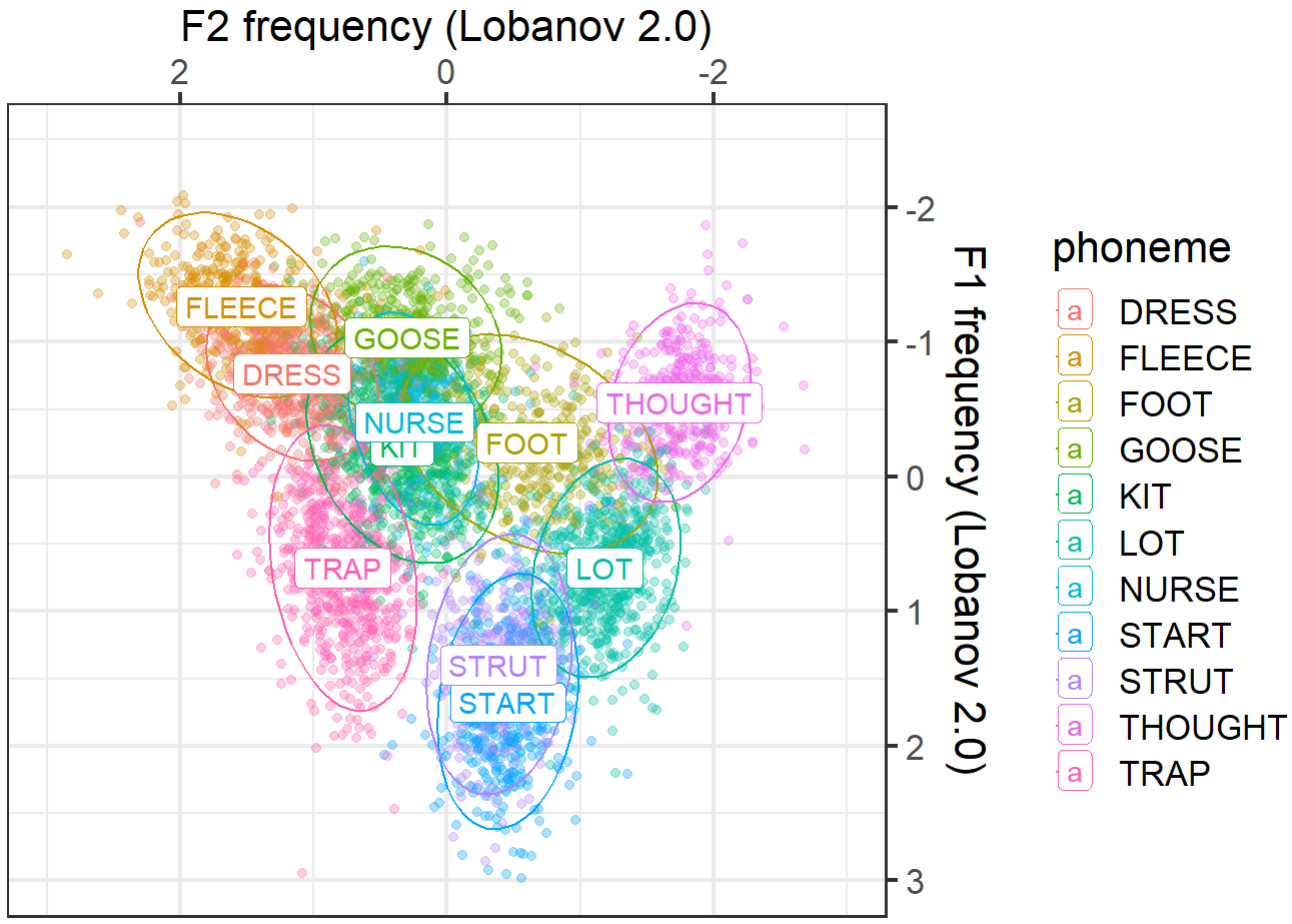
```
library(tidyverse)
library(lme4)
library(gtsummary)
library(mgcv)
library(itsadug)
```

Figure X: Normalised F1 and F2 of all vowel tokens, by phoneme

```
allvowels <- read.csv('allvowels.csv')
allvowels <- filter(allvowels, phoneme != "")

allvowelsmeans <- allvowels %>%
  group_by(phoneme) %>%
  summarise(mean_F1Norm = mean(F1norm),
            mean_F2Norm = mean(F2norm))

allplot <- allvowels %>%
  ggplot(aes(x = F2norm, y = F1norm, colour = phoneme, label = phoneme)) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95) +
  geom_label(data = allvowelsmeans, aes(x = mean_F2Norm, y = mean_F1Norm)) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

allplot
```

```
## Warning: Removed 7 rows containing non-finite outside the scale range
## (`stat_ellipse()`).
```

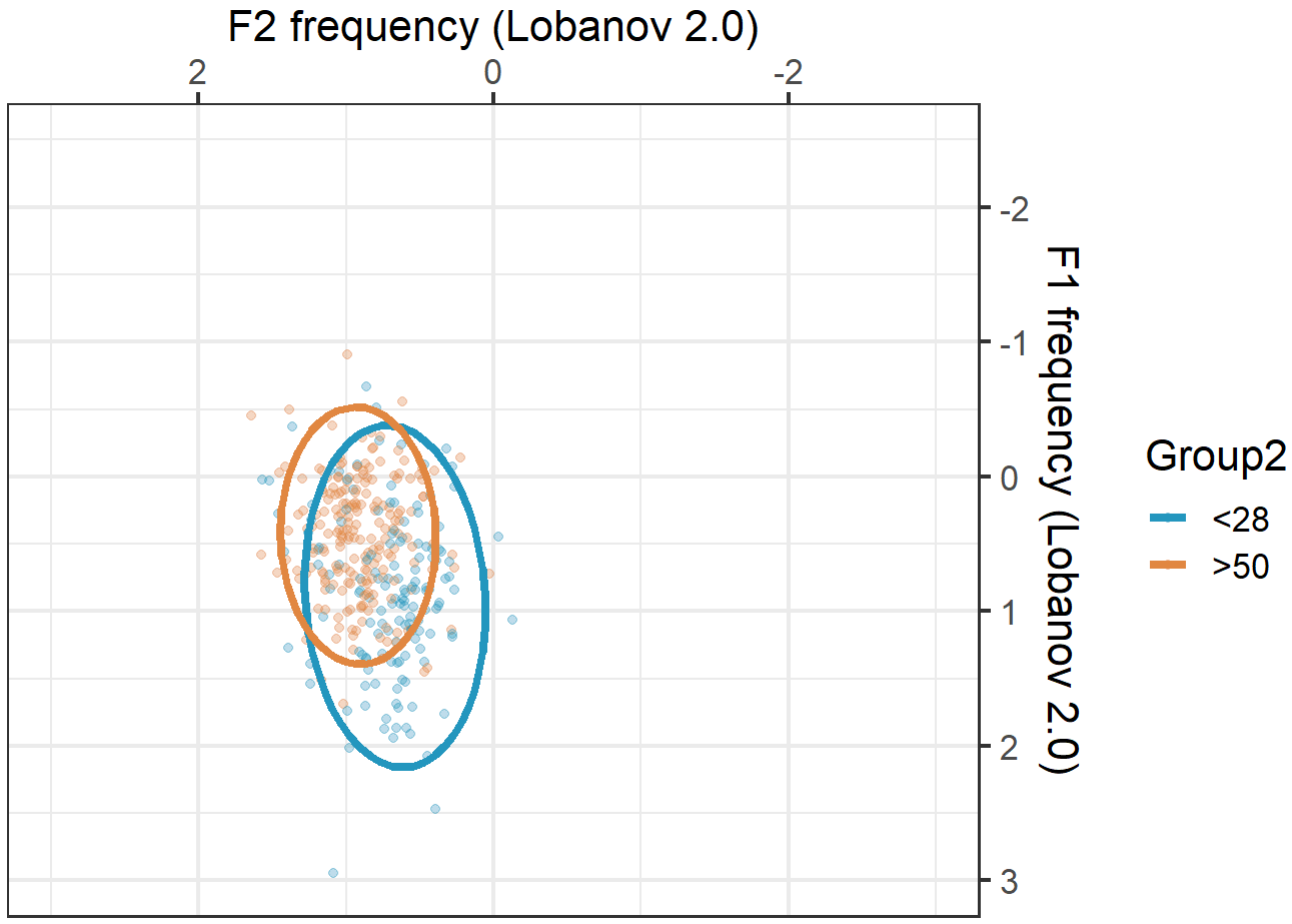
```
## Warning: Removed 7 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



For all of the individual vowel results, the plots generated below were layered over the greyed-out plot of all tokens (at the same scale) to provide visual context.

TRAP results (figure X)

```
TRAPcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                      data = filter(allvowels, phoneme == "TRAP", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

TRAPcompare
```




```
trapF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "TRAP"))

summary(trapF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "TRAP")
##
## REML criterion at convergence: 737.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0742 -0.5357  0.0342  0.5799  4.3999
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.05193  0.2279
##  Residual             0.16768  0.4095
## Number of obs: 617, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  1.175752   0.127578   9.216
## Age         -0.012044   0.002641  -4.560
## GenderM     -0.342688   0.208161  -1.646
## Age:GenderM  0.009671   0.004569   2.117
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.613  0.554
## Age:GenderM  0.522 -0.578 -0.914
```

```
trapF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	-0.01	-0.02, -0.01	<0.001
Gender			0.10
F	—	—	
M	-0.34	-0.75, 0.07	
Age * Gender			0.034
Age * M	0.01	0.00, 0.02	

[1] CI = Confidence Interval

```
trapF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "TR
AP"))
summary(trapF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "TRAP")
##
## REML criterion at convergence: -81.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3752 -0.6486  0.0811  0.6272  3.1940
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01628  0.1276
##  Residual             0.04371  0.2091
## Number of obs: 617, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.510348   0.070358   7.254
## Age          0.005747   0.001458   3.942
## GenderM      0.172226   0.114903   1.499
## Age:GenderM -0.003152   0.002523  -1.249
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.612  0.553
## Age:GenderM  0.522 -0.578 -0.914
```

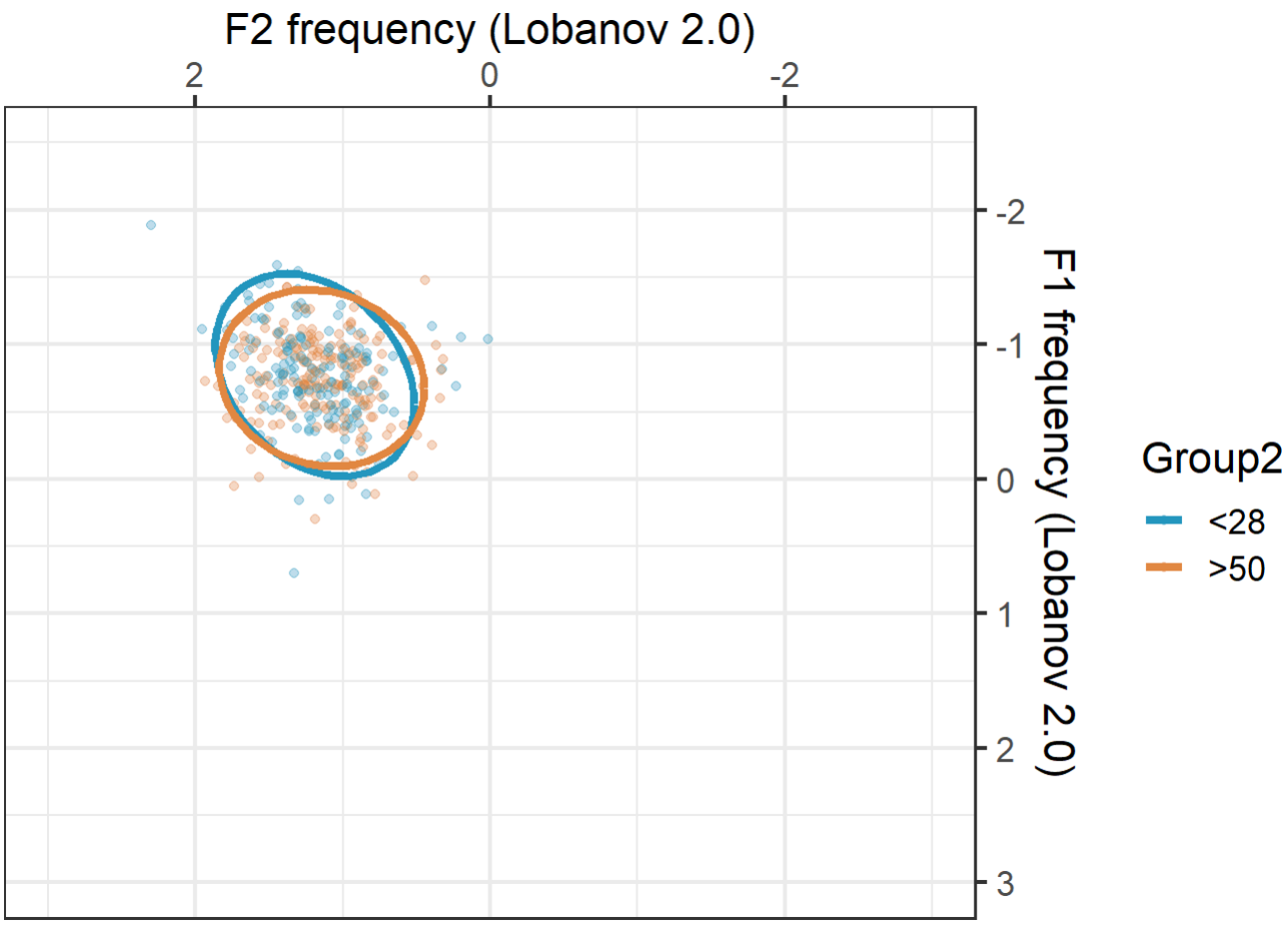
```
trapF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.01	0.00, 0.01	<0.001
Gender			0.13
F	—	—	
M	0.17	-0.05, 0.40	
Age * Gender			0.2
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

DRESS results (figure X)

```
DRESScompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                       data = filter(allvowels, phoneme == "DRESS", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

DRESScompare
```




```
DRESSF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "D
RESS"))
summary(DRESSF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "DRESS")
##
## REML criterion at convergence: 286.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3172 -0.6838 -0.0459  0.6393  4.3632
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01837  0.1355
##  Residual             0.08078  0.2842
## Number of obs: 632, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.838547   0.078564 -10.673
## Age          0.001355   0.001627   0.833
## GenderM     -0.102791   0.129101  -0.796
## Age:GenderM  0.003645   0.002871   1.269
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.609  0.550
## Age:GenderM  0.512 -0.567 -0.915
```

```
DRESSF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.4
Gender			0.4
F	—	—	
M	-0.10	-0.36, 0.15	
Age * Gender			0.2
Age * M	0.00	0.00, 0.01	

[1] CI = Confidence Interval

```
DRESSF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "D
RESS"))
summary(DRESSF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "DRESS")
##
## REML criterion at convergence: 309
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6743 -0.7220  0.0489  0.6312  4.1361
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01066  0.1033
##  Residual             0.08562  0.2926
## Number of obs: 632, groups:  ID, 36
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  1.2042170  0.0651954  18.471
## Age         -0.0011235  0.0013464  -0.834
## GenderM      0.0233838  0.1072865   0.218
## Age:GenderM -0.0004334  0.0023984  -0.181
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.608  0.550
## Age:GenderM  0.508 -0.561 -0.916
```

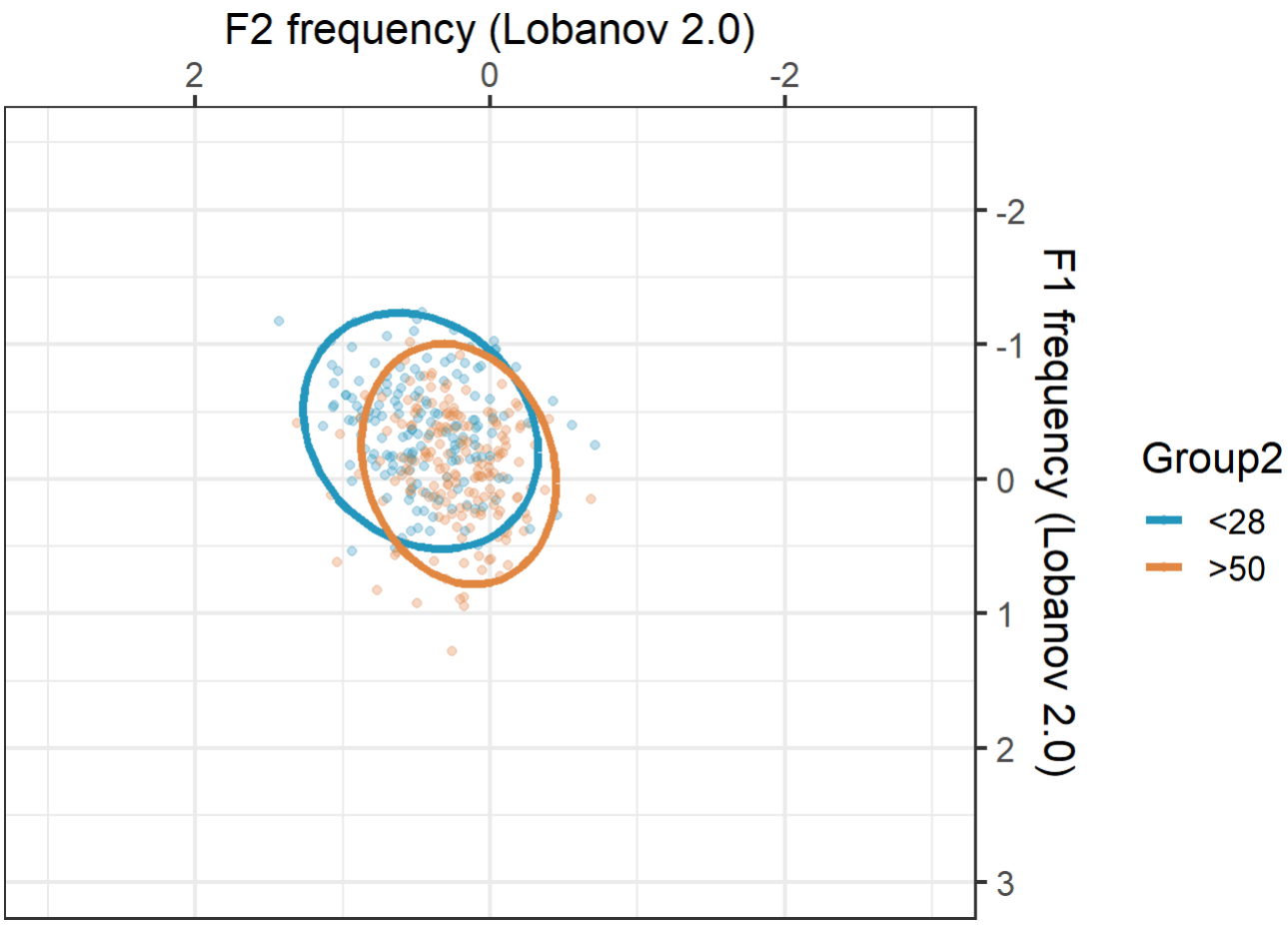
```
DRESSF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.4
Gender			0.8
F	—	—	
M	0.02	-0.19, 0.23	
Age * Gender			0.9
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

KIT results (figure X)

```
KITcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                     data = filter(allvowels, phoneme == "KIT", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

KITcompare
```




```
KITF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "KI
T"))
summary(KITF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "KIT")
##
## REML criterion at convergence: 592.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1476 -0.6964  0.0057  0.6326  3.3618
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.03133  0.1770
##  Residual             0.13039  0.3611
## Number of obs: 636, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.527745   0.101335  -5.208
## Age          0.006883   0.002099   3.279
## GenderM      0.282708   0.165931   1.704
## Age:GenderM -0.007312   0.003660  -1.998
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.611  0.552
## Age:GenderM  0.518 -0.574 -0.914
```

```
KITF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.01	0.00, 0.01	0.001
Gender			0.088
F	—	—	
M	0.28	-0.04, 0.61	
Age * Gender			0.046
Age * M	-0.01	-0.01, 0.00	

[1] CI = Confidence Interval

```
KITF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "KI
T"))
summary(KITF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "KIT")
##
## REML criterion at convergence: 335.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8950 -0.6969 -0.0105  0.6503  3.2041
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01606  0.1267
##  Residual             0.08773  0.2962
## Number of obs: 636, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.610277   0.074892   8.149
## Age         -0.006204   0.001549  -4.004
## GenderM     -0.307988   0.122548  -2.513
## Age:GenderM  0.007452   0.002704   2.756
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.611  0.553
## Age:GenderM  0.518 -0.573 -0.915
```

```
KITF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	-0.01	-0.01, 0.00	<0.001
Gender			0.012
F	—	—	
M	-0.31	-0.55, -0.07	
Age * Gender			0.006
Age * M	0.01	0.00, 0.01	

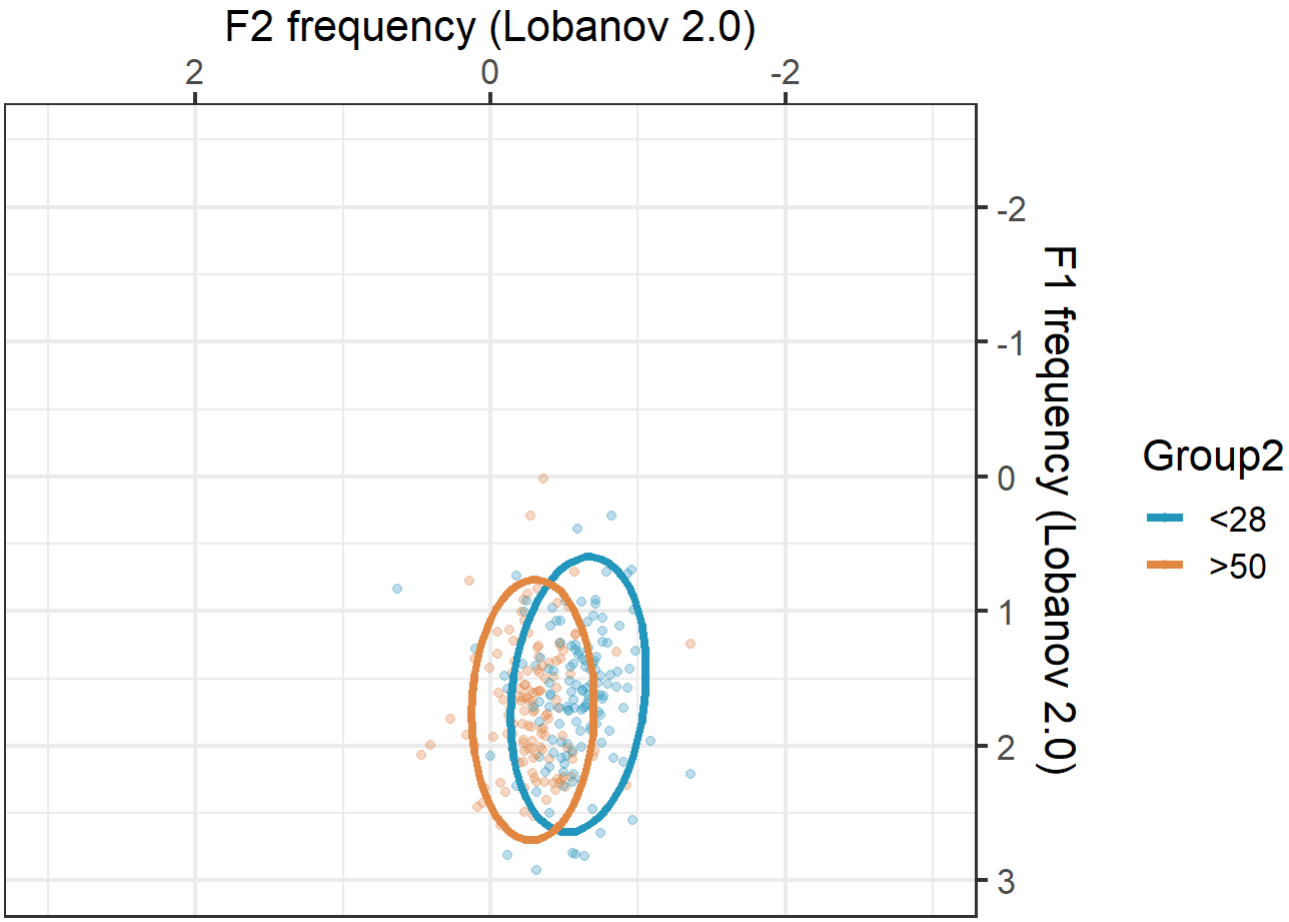
[1] CI = Confidence Interval

START results (figure X)

```
allvowels$Group2[allvowels$Group2 == '25-50'] <- '28-50'

STARTcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                       data = filter(allvowels, phoneme == "START", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

STARTcompare
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_ellipse()`).
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_point()`).
```




```
STARTF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "S
TART"))
summary(STARTF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "START")
##
## REML criterion at convergence: 681.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1046 -0.5295 -0.0204  0.5146  4.7726
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.03189  0.1786
##  Residual             0.20298  0.4505
## Number of obs: 496, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  1.573217   0.110482  14.240
## Age          0.003477   0.002326   1.495
## GenderM      0.141758   0.185755   0.763
## Age:GenderM -0.005936   0.004193  -1.416
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.903
## GenderM     -0.595  0.537
## Age:GenderM  0.501 -0.555 -0.917
```

```
STARTF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.13
Gender			0.4
F	—	—	
M	0.14	-0.22, 0.51	
Age * Gender			0.2
Age * M	-0.01	-0.01, 0.00	

[1] CI = Confidence Interval

```
STARTF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "S
TART"))
summary(STARTF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "START")
##
## REML criterion at convergence: -115
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.2065 -0.4920  0.0143  0.4912  5.0415
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01295  0.1138
##  Residual             0.03879  0.1970
## Number of obs: 496, groups:  ID, 36
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) -0.7552594  0.0640783 -11.787
## Age          0.0066329  0.0013435   4.937
## GenderM      0.1248113  0.1067022   1.170
## Age:GenderM -0.0006346  0.0023857  -0.266
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.903
## GenderM     -0.601  0.542
## Age:GenderM  0.509 -0.563 -0.915
```

```
STARTF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.01	0.00, 0.01	<0.001
Gender			0.2
F	—	—	
M	0.12	-0.08, 0.33	
Age * Gender			0.8
Age * M	0.00	-0.01, 0.00	

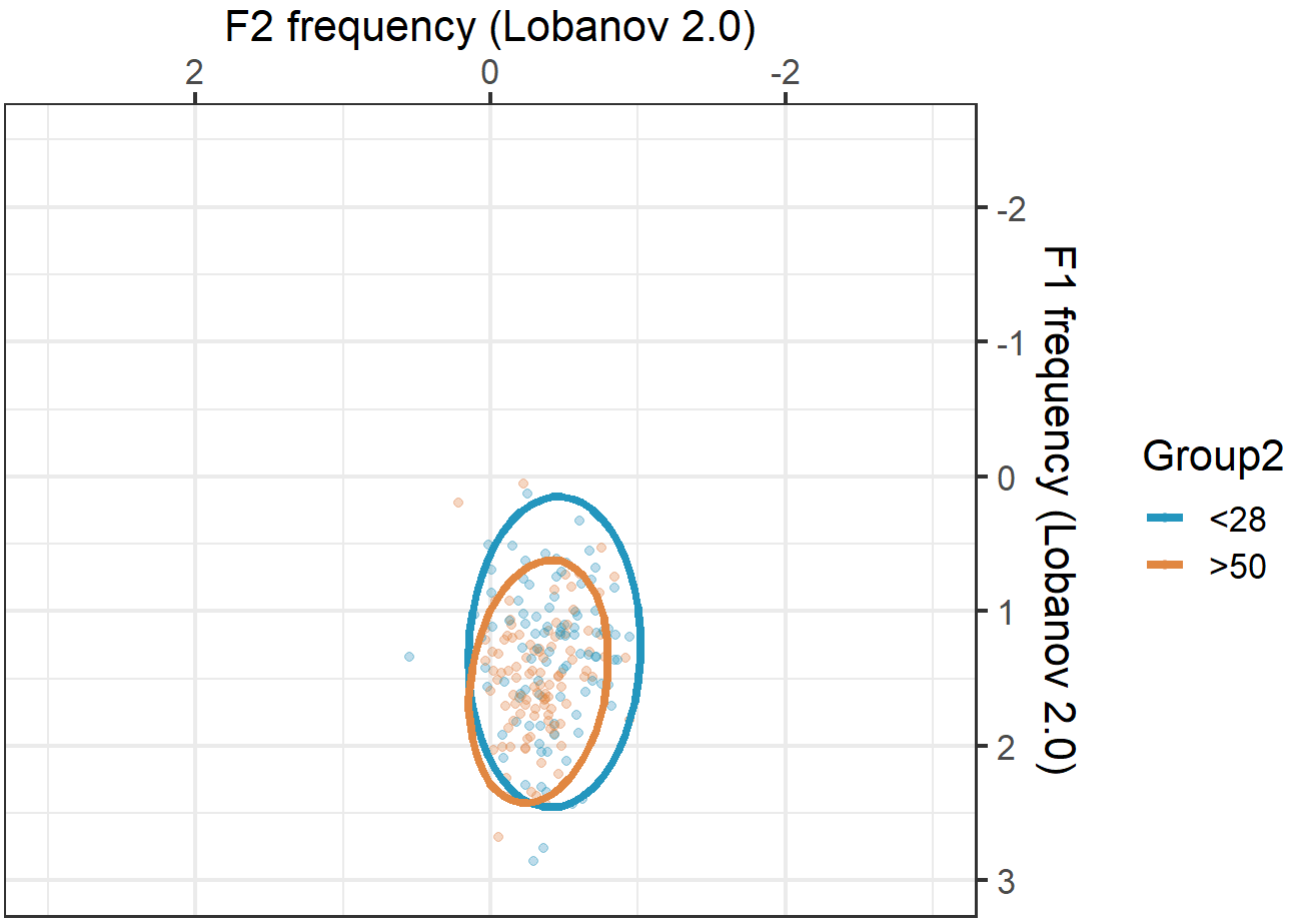
[1] CI = Confidence Interval

STRUT results (figure X)

```
STRUTcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                       data = filter(allvowels, phoneme == "STRUT", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

STRUTcompare
```

```
## Warning: Removed 2 rows containing non-finite outside the scale range
## (`stat_ellipse()`).
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_point()`).
```




```
STRUTF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "S
TRUT"))
summary(STRUTF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "STRUT")
##
## REML criterion at convergence: 501.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9709 -0.5787 -0.0649  0.5382  3.6110
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.003245 0.05696
##  Residual             0.228360 0.47787
## Number of obs: 346, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  1.325862   0.083386  15.900
## Age          0.001915   0.001730   1.107
## GenderM     -0.146199   0.136053  -1.075
## Age:GenderM  0.003436   0.003025   1.136
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.613  0.555
## Age:GenderM  0.518 -0.572 -0.909
```

```
STRUTF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.3
Gender			0.3
F	—	—	
M	-0.15	-0.41, 0.12	
Age * Gender			0.3
Age * M	0.00	0.00, 0.01	

[1] CI = Confidence Interval

```
STRUTF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "S
TRUT"))

summary(STRUTF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "STRUT")
##
## REML criterion at convergence: 32.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4221 -0.5429 -0.0341  0.5665  3.9377
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01222  0.1106
##  Residual             0.05271  0.2296
## Number of obs: 346, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.524960   0.068107  -7.708
## Age          0.003115   0.001415   2.201
## GenderM      0.146524   0.111914   1.309
## Age:GenderM -0.003571   0.002472  -1.445
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.609  0.550
## Age:GenderM  0.518 -0.573 -0.912
```

```
STRUTF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.028
Gender			0.2
F	—	—	
M	0.15	-0.07, 0.37	
Age * Gender			0.15

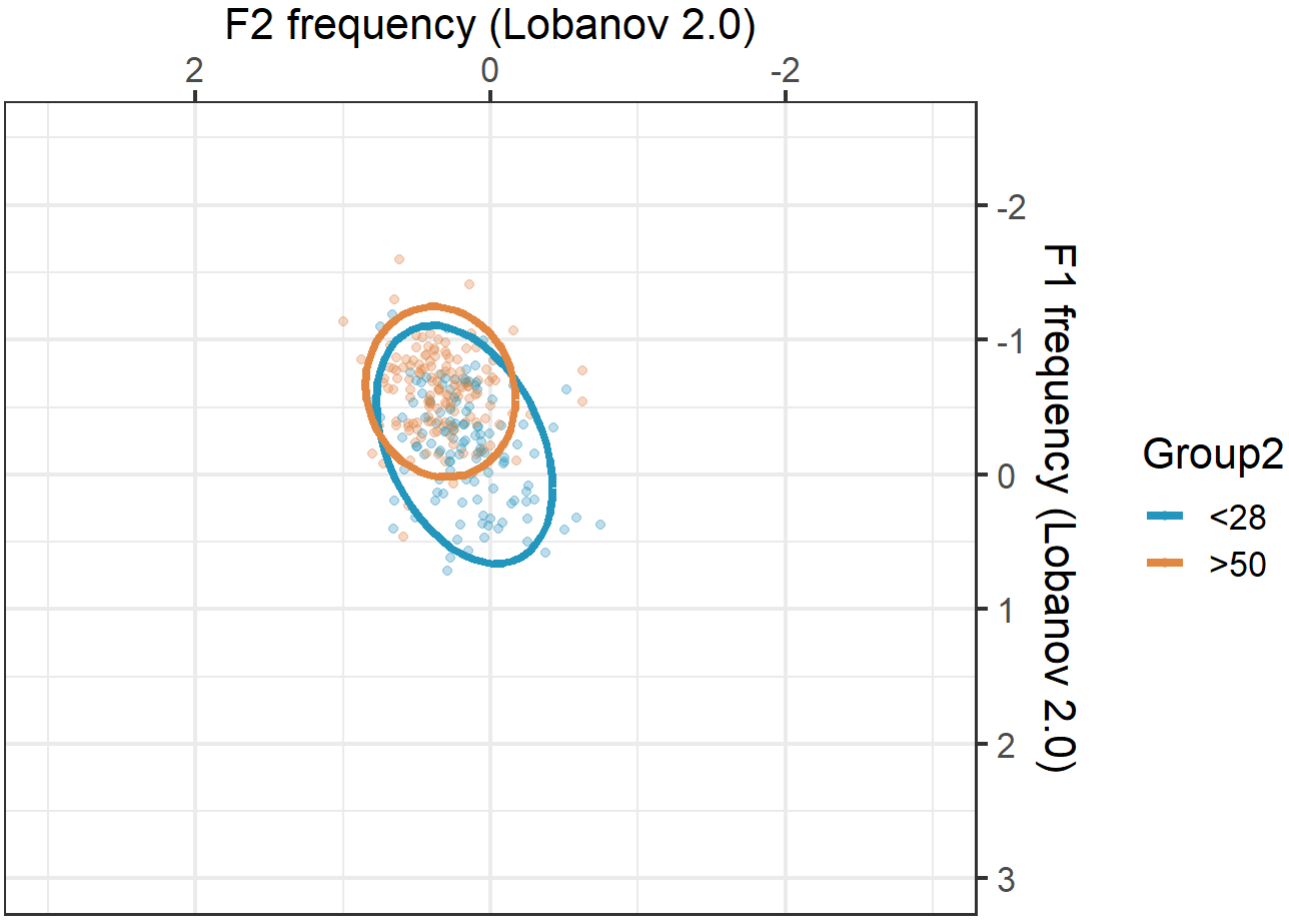
[1] CI = Confidence Interval

Characteristic	Beta	95% CI[1]	p-value
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

NURSE results (figure X)

```
NURSEcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                       data = filter(allvowels, phoneme == "NURSE", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

NURSEcompare
```




```
NURSEF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "N
URSE"))
summary(NURSEF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "NURSE")
##
## REML criterion at convergence: 210.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5668 -0.6366 -0.0788  0.5849  3.1127
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.02945  0.1716
##  Residual             0.07694  0.2774
## Number of obs: 455, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.109631   0.099600   1.101
## Age         -0.011997   0.002049  -5.854
## GenderM     -0.349745   0.160767  -2.175
## Age:GenderM  0.008269   0.003538   2.337
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.907
## GenderM     -0.620  0.562
## Age:GenderM  0.525 -0.579 -0.915
```

```
NURSEF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	-0.01	-0.02, -0.01	<0.001
Gender			0.030
F	—	—	
M	-0.35	-0.66, -0.03	
Age * Gender			0.019
Age * M	0.01	0.00, 0.02	

[1] CI = Confidence Interval

```
NURSEF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "N
URSE"))
summary(NURSEF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "NURSE")
##
## REML criterion at convergence: -111.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0258 -0.5809 -0.0237  0.5977  2.9603
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.02098  0.1449
##  Residual             0.03677  0.1917
## Number of obs: 455, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.123403   0.080977   1.524
## Age          0.003694   0.001670   2.211
## GenderM      0.095914   0.131129   0.731
## Age:GenderM -0.004293   0.002886  -1.488
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.906
## GenderM     -0.618  0.559
## Age:GenderM  0.524 -0.579 -0.915
```

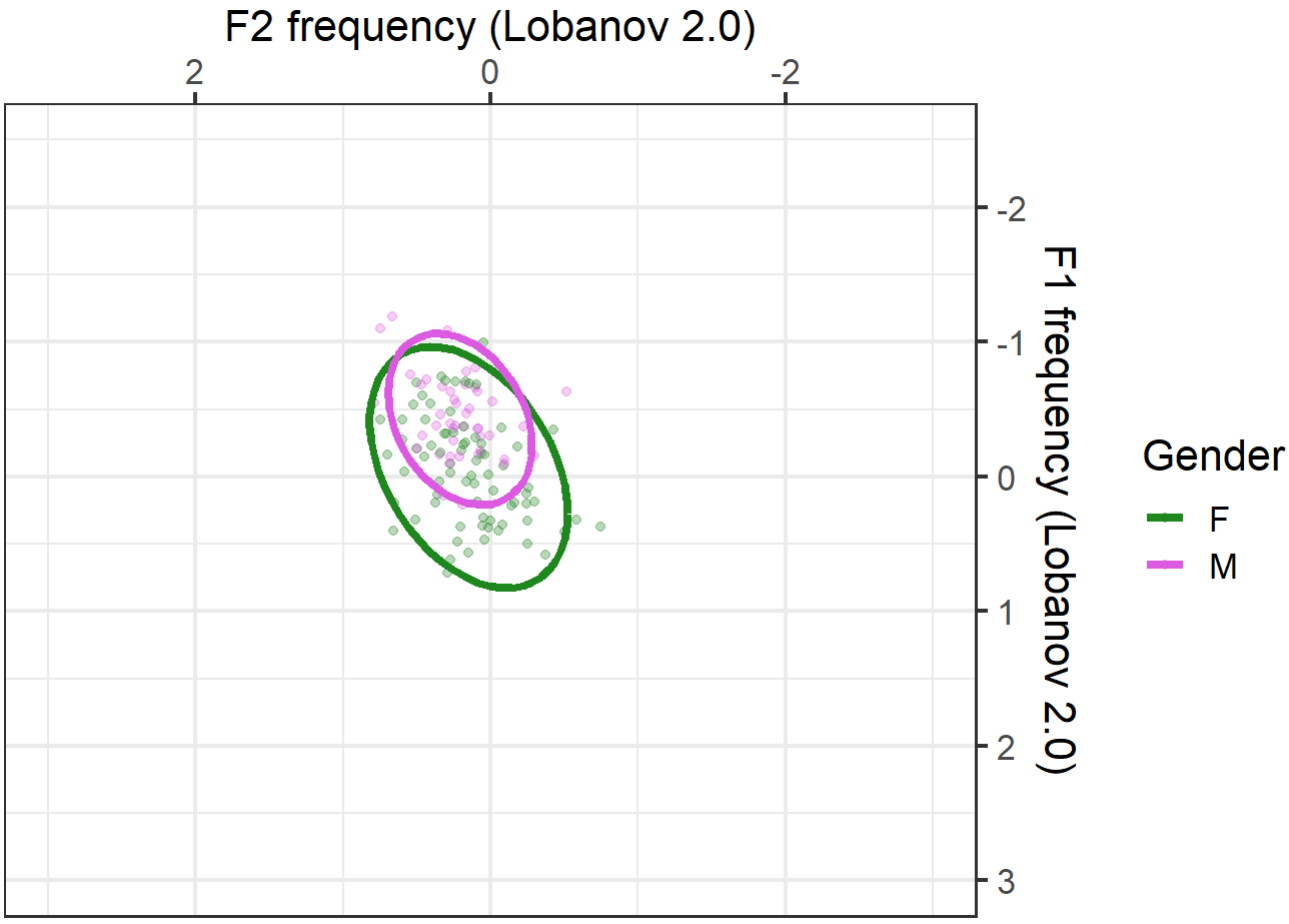
```
NURSEF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.027
Gender			0.5
F	—	—	
M	0.10	-0.16, 0.35	
Age * Gender			0.14
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

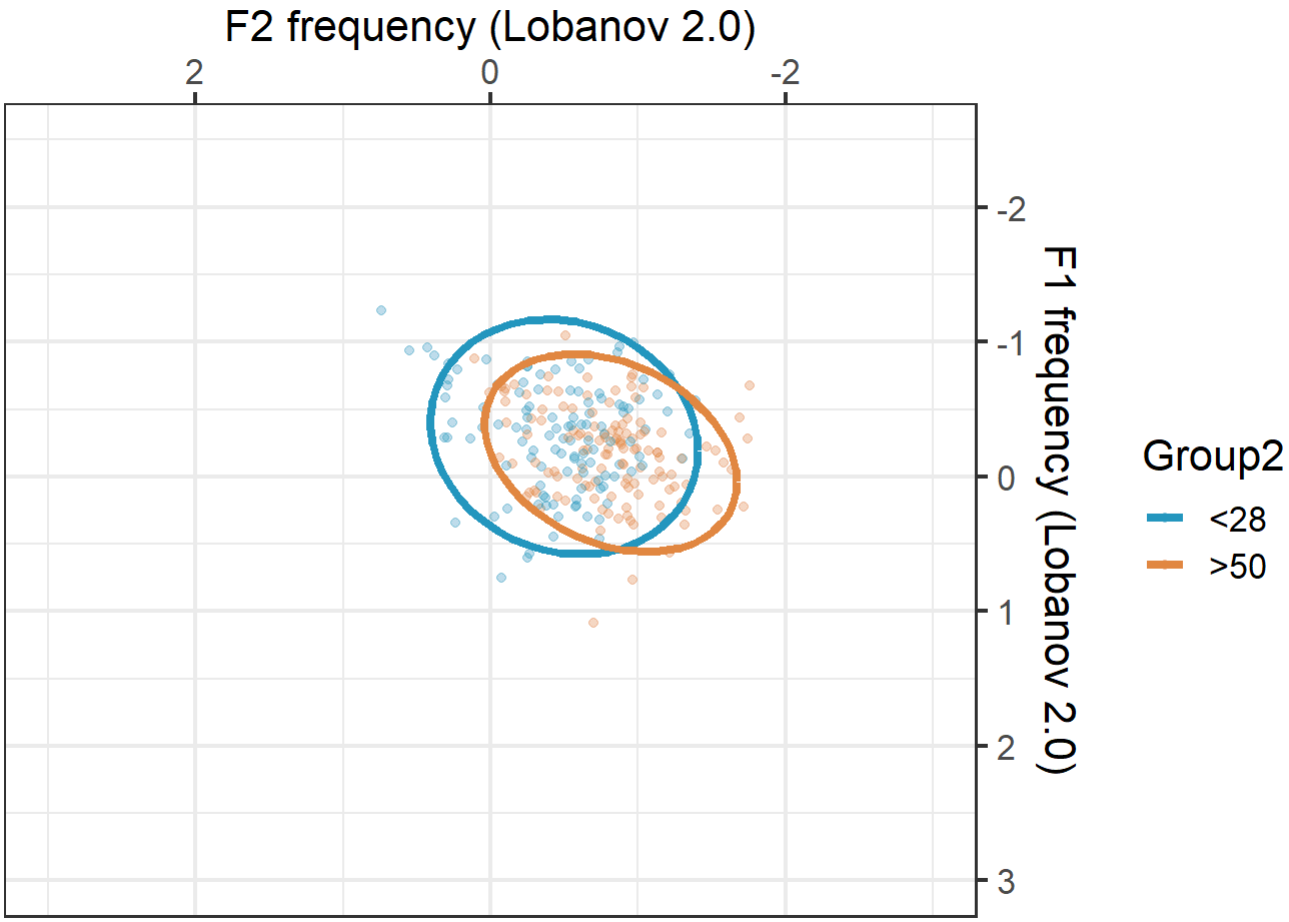
Extra NURSE plot, showing differences between young male and young female speakers:

```
NURSEgender <- ggplot(aes(x = F2norm, y = F1norm, colour = Gender, label = Gender
), data = filter(allvowels, phoneme == "NURSE", Group2 == "<28")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#1e871e", "#dd5ce4")) +
  #geom_label(data = TRAPlabels, aes(x = mean_F2Norm, y = mean_F1Norm)) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

NURSEgender
```



FOOT results (figure X)

```
FOOTcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                      data = filter(allvowels, phoneme == "FOOT", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

FOOTcompare
```




```
FOOTF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "FO
OT"))
summary(FOOTF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "FOOT")
##
## REML criterion at convergence: 376.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6497 -0.6472  0.0097  0.6327  3.2464
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01551  0.1246
##  Residual             0.13049  0.3612
## Number of obs: 401, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.396396   0.084049  -4.716
## Age          0.003486   0.001756   1.985
## GenderM      0.217113   0.141059   1.539
## Age:GenderM -0.004981   0.003062  -1.627
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.904
## GenderM     -0.596  0.539
## Age:GenderM  0.519 -0.573 -0.911
```

```
FOOTF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.047
Gender			0.12
F	—	—	
M	0.22	-0.06, 0.49	
Age * Gender			0.10
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

```
FOOTF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "FO
OT"))
summary(FOOTF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "FOOT")
##
## REML criterion at convergence: 467.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.51294 -0.67740 -0.07775  0.62010  2.79019
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01657  0.1287
##  Residual             0.16526  0.4065
## Number of obs: 401, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.273284   0.090134  -3.032
## Age         -0.007765   0.001885  -4.120
## GenderM     -0.096981   0.151569  -0.640
## Age:GenderM  0.001938   0.003286   0.590
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.595  0.538
## Age:GenderM  0.519 -0.573 -0.911
```

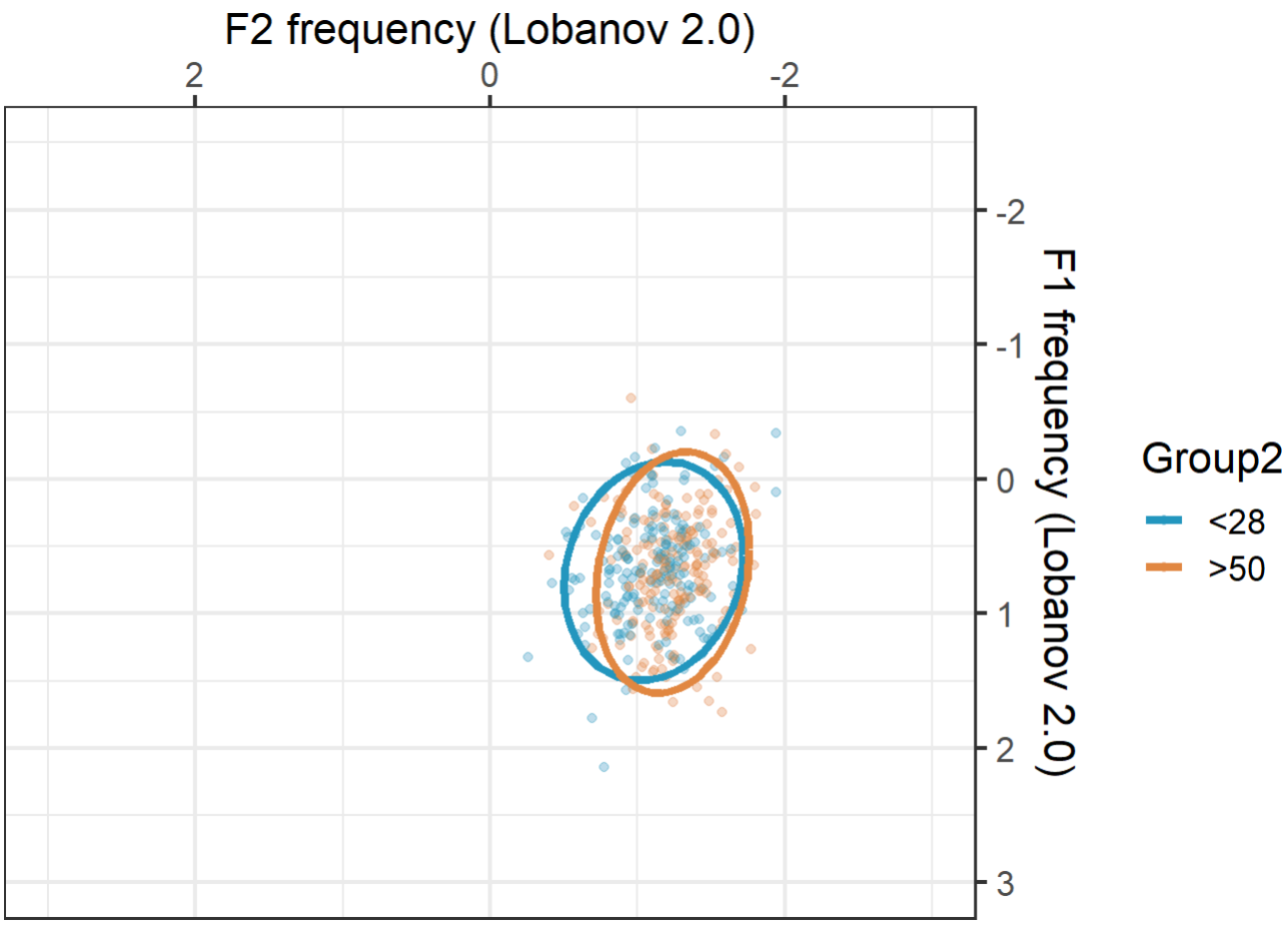
```
FOOTF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	-0.01	-0.01, 0.00	<0.001
Gender			0.5
F	—	—	
M	-0.10	-0.39, 0.20	
Age * Gender			0.6
Age * M	0.00	0.00, 0.01	

[1] CI = Confidence Interval

LOT results (figure X)

```
LOTcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                     data = filter(allvowels, phoneme == "LOT", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

LOTcompare
```




```
LOTF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "LO
T"))
summary(LOTF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "LOT")
##
## REML criterion at convergence: 610.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6601 -0.6253 -0.0276  0.6219  4.0584
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.02024  0.1423
##  Residual             0.13502  0.3674
## Number of obs: 645, groups:  ID, 36
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.573933   0.087125   6.587
## Age         0.001654   0.001832   0.903
## GenderM     0.114854   0.141668   0.811
## Age:GenderM 0.000203   0.003134   0.065
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.615  0.557
## Age:GenderM  0.529 -0.585 -0.915
```

```
LOTF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.01	0.4
Gender			0.4
F	—	—	
M	0.11	-0.16, 0.39	
Age * Gender			>0.9
Age * M	0.00	-0.01, 0.01	

[1] CI = Confidence Interval

```
LOTF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "LO
T"))
summary(LOTF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "LOT")
##
## REML criterion at convergence: 82.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5717 -0.6190 -0.0620  0.5893  4.4740
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01324  0.1151
##  Residual             0.05832  0.2415
## Number of obs: 645, groups:  ID, 36
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) -1.0831507  0.0666890 -16.242
## Age         -0.0021152  0.0013991  -1.512
## GenderM     -0.0007199  0.1087015  -0.007
## Age:GenderM -0.0005951  0.0024018  -0.248
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.614  0.555
## Age:GenderM  0.527 -0.583 -0.915
```

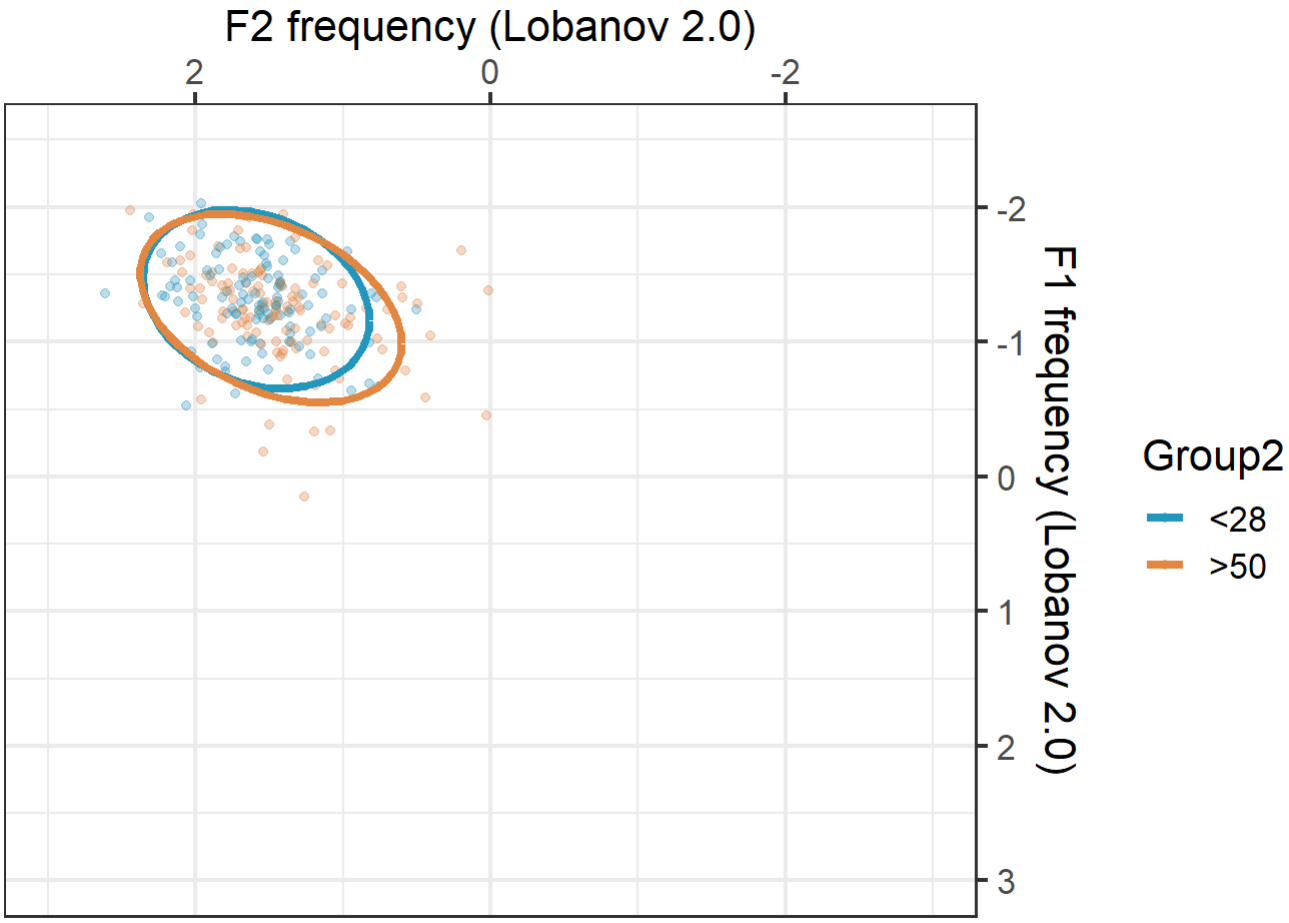
```
LOTF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.13
Gender			>0.9
F	—	—	
M	0.00	-0.21, 0.21	
Age * Gender			0.8
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

FLEECE results (figure X)

```
FLEECEcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                        data = filter(allvowels, phoneme == "FLEECE", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

FLEECEcompare
```




```
FLEECEF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme ==
"FLEECE"))
summary(FLEECEF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "FLEECE")
##
## REML criterion at convergence: 256.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8186 -0.5859 -0.0700  0.5462  4.3585
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01321  0.1150
##  Residual             0.09648  0.3106
## Number of obs: 395, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -1.309309   0.074434 -17.590
## Age          0.001709   0.001538   1.111
## GenderM     -0.007155   0.128180  -0.056
## Age:GenderM -0.001714   0.002850  -0.601
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.900
## GenderM     -0.581  0.523
## Age:GenderM  0.486 -0.540 -0.915
```

```
FLEECEF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.3
Gender			>0.9
F	—	—	
M	-0.01	-0.26, 0.24	
Age * Gender			0.5
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

```
FLEECEF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme ==
"FLEECE"))
summary(FLEECEF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "FLEECE")
##
## REML criterion at convergence: 348.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9323 -0.5238  0.0523  0.6321  3.2867
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.02592  0.1610
##  Residual             0.11933  0.3454
## Number of obs: 395, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  1.709184   0.096682  17.678
## Age         -0.004057   0.002003  -2.026
## GenderM     -0.128560   0.164639  -0.781
## Age:GenderM  0.003960   0.003650   1.085
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.900
## GenderM     -0.587  0.529
## Age:GenderM  0.494 -0.549 -0.915
```

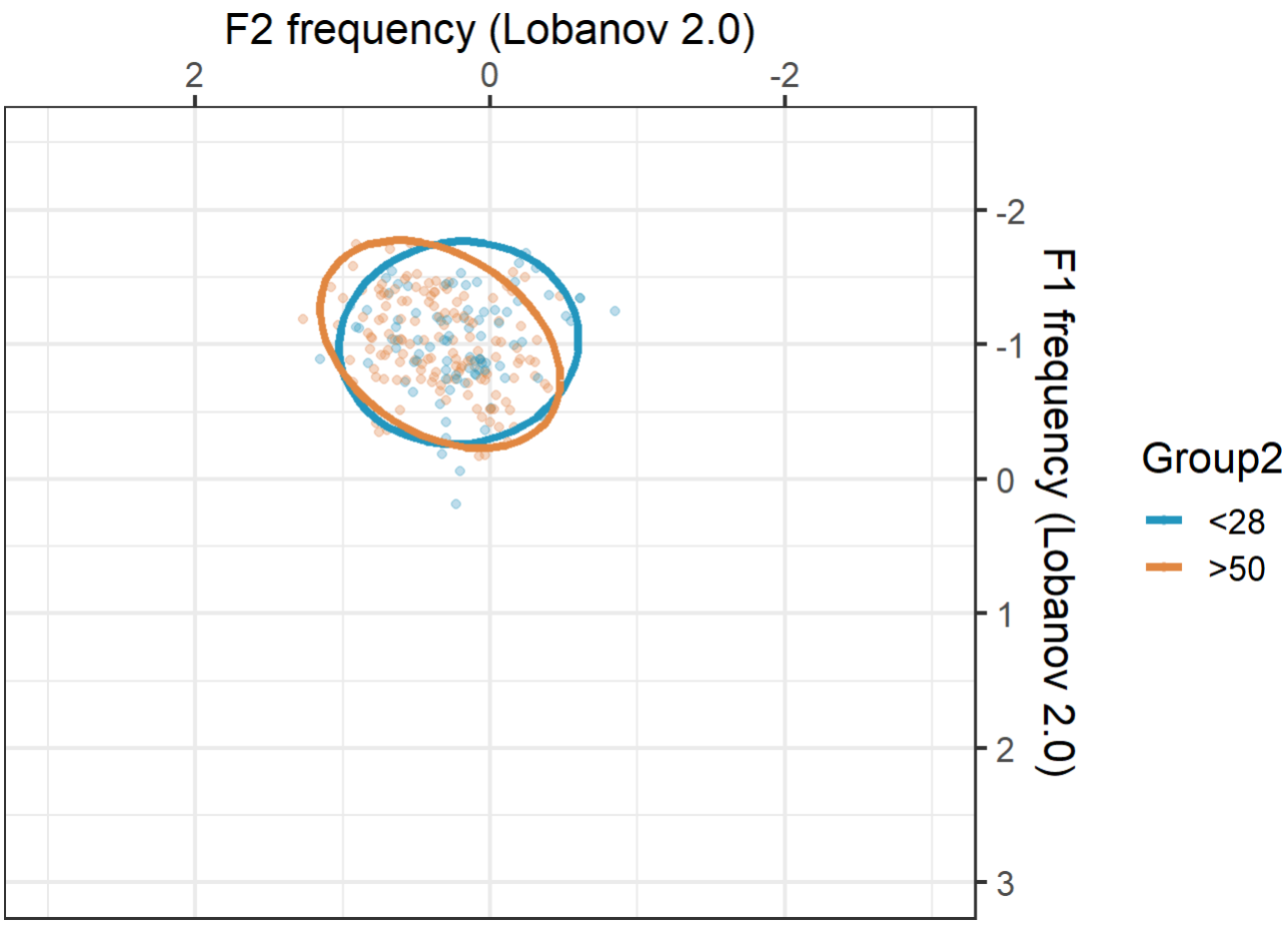
```
FLEECEF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	-0.01, 0.00	0.043
Gender			0.4
F	—	—	
M	-0.13	-0.45, 0.19	
Age * Gender			0.3
Age * M	0.00	0.00, 0.01	

[1] CI = Confidence Interval

GOOSE results (figure X)

```
GOOSEcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                       data = filter(allvowels, phoneme == "GOOSE", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

GOOSEcompare
```




```
GOOSEF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "G
OOSE"))
summary(GOOSEF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "GOOSE")
##
## REML criterion at convergence: 266.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5509 -0.6743 -0.0134  0.5648  3.7368
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.009287 0.09637
##  Residual             0.095270 0.30866
## Number of obs: 435, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -1.097326   0.070167 -15.639
## Age          0.002082   0.001441   1.445
## GenderM      0.064282   0.116377   0.552
## Age:GenderM -0.002775   0.002551  -1.088
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.909
## GenderM     -0.603  0.548
## Age:GenderM  0.513 -0.565 -0.921
```

```
GOOSEF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.15
Gender			0.6
F	—	—	
M	0.06	-0.16, 0.29	
Age * Gender			0.3
Age * M	0.00	-0.01, 0.00	

[1] CI = Confidence Interval

```
GOOSEF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme == "G
OOSE"))
summary(GOOSEF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "GOOSE")
##
## REML criterion at convergence: 306.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.9846 -0.5910 -0.0005  0.5947  3.2974
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01963  0.1401
##  Residual             0.10133  0.3183
## Number of obs: 435, groups:  ID, 36
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  0.2514427  0.0885981   2.838
## Age          0.0010551  0.0018259   0.578
## GenderM     -0.0012307  0.1463249  -0.008
## Age:GenderM -0.0009026  0.0032135  -0.281
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.907
## GenderM     -0.605  0.549
## Age:GenderM  0.515 -0.568 -0.919
```

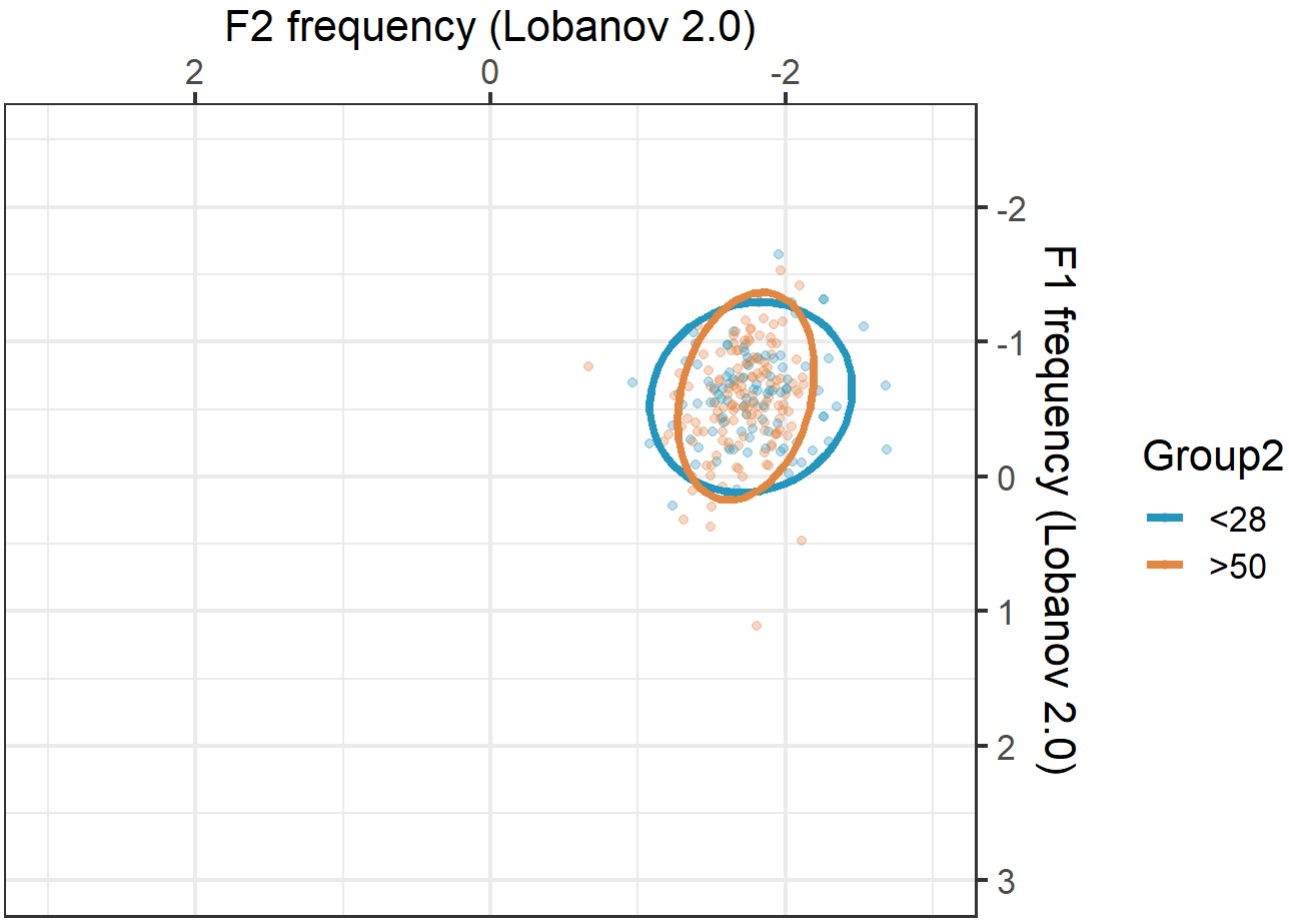
```
GOOSEF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.6
Gender			>0.9
F	—	—	
M	0.00	-0.29, 0.29	
Age * Gender			0.8
Age * M	0.00	-0.01, 0.01	

[1] CI = Confidence Interval

THOUGHT results (figure X)

```
THOUGHTcompare <- ggplot(aes(x = F2norm, y = F1norm, colour = Group2),
                         data = filter(allvowels, phoneme == "THOUGHT", Group2 != "28-50")) +
  geom_point(alpha=0.3) +
  stat_ellipse(level =  0.95, linewidth = 1.5) +
  scale_colour_manual(values=c("#2596be", "#e28743")) +
  scale_x_reverse(name = "F2 frequency (Lobanov 2.0)", position = "top", limit=c(3, -3)) +
  scale_y_reverse(name = "F1 frequency (Lobanov 2.0)", position = "right", limit=c(3, -2.5))
+
  theme_bw(base_size=16)

THOUGHTcompare
```




```
THOUGHTF1_Age_Gender <- lmer(F1norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme ==
"THOUGHT"))
summary(THOUGHTF1_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F1norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "THOUGHT")
##
## REML criterion at convergence: 335.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7538 -0.5884 -0.0326  0.6662  4.5248
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01391  0.1180
##  Residual             0.11649  0.3413
## Number of obs: 406, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.596411   0.082109  -7.264
## Age          0.001216   0.001667   0.729
## GenderM      0.023238   0.135814   0.171
## Age:GenderM -0.000440   0.002941  -0.150
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.908
## GenderM     -0.605  0.549
## Age:GenderM  0.515 -0.567 -0.917
```

```
THOUGHTF1_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.5
Gender			0.9
F	—	—	
M	0.02	-0.24, 0.29	
Age * Gender			0.9
Age * M	0.00	-0.01, 0.01	

[1] CI = Confidence Interval

```
THOUGHTF2_Age_Gender <- lmer(F2norm~Age*Gender + (1|ID), data = filter(allvowels, phoneme ==
"THOUGHT"))
summary(THOUGHTF2_Age_Gender)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: F2norm ~ Age * Gender + (1 | ID)
##    Data: filter(allvowels, phoneme == "THOUGHT")
##
## REML criterion at convergence: 16.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7521 -0.6585 -0.0005  0.5528  4.8904
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 0.01897  0.1377
##  Residual             0.04947  0.2224
## Number of obs: 406, groups:  ID, 36
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -1.791730   0.079194 -22.625
## Age          0.001143   0.001630   0.702
## GenderM     -0.036066   0.130698  -0.276
## Age:GenderM  0.001445   0.002851   0.507
##
## Correlation of Fixed Effects:
##             (Intr) Age    GendrM
## Age         -0.905
## GenderM     -0.606  0.548
## Age:GenderM  0.517 -0.572 -0.915
```

```
THOUGHTF2_Age_Gender %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Age	0.00	0.00, 0.00	0.5
Gender			0.8
F	—	—	
M	-0.04	-0.29, 0.22	
Age * Gender			0.6
Age * M	0.00	0.00, 0.01	

[1] CI = Confidence Interval

Figure 3.14: Predicted vowel movement by year of birth. Code is adapted from this tutorial by Joshua Wilson Black : https://joshua.wilsonblack.nz/posts/visualising-vowel-space-gamms/ (https://joshua.wilsonblack.nz/posts/visualising-vowel-space-gamms/)

```
vowelsbyage <- read.csv("allvowels_timenorm.csv")

vowelsbyage <- vowelsbyage %>%
  mutate(
    phoneme <- as.factor(phoneme))

vowelsbyage <- vowelsbyage %>%
  mutate(
    Gender <- as.factor(Gender))


vowelsbyage <- vowelsbyage %>%
  pivot_longer(
    cols = F1norm:F2norm,
    names_to = "Formant",
    values_to = "Frequency"
  )

vowelsbyage <- vowelsbyage %>%
  group_by(phoneme, Formant) %>%
  nest()

vowelsbyage <- vowelsbyage %>%
  mutate(
    model = map(
      data,
      ~ bam(
        Frequency ~ Gender +
          s(Age, k=3),
        data = .x,
        method = 'fREML',
        discrete = TRUE,
        nthreads = 2
      )
    )
  )


to_predict <- list(
  "Age" = seq(19, 84, 1)) # complete age range in 1yr increments

vowelsbyage <- vowelsbyage %>%
  mutate(
    prediction = map(
      model,
      ~ get_predictions(model = .x, cond = to_predict, print.summary = FALSE)
    )
  )

predictions <- vowelsbyage %>%
  select(
    phoneme, Formant, prediction
  ) %>%
  unnest(prediction)
```

```
predictions <- predictions %>%
  select(-CI) %>%
  pivot_wider(
    names_from = Formant,
    values_from = fit
  )

predictions %>%
  ggplot(
    aes(
      x = F2norm,
      y = F1norm,
      colour = phoneme
    )
  ) +
  geom_point() +
  #geom_path(arrow=arrow(length= unit(2, "mm")), linewidth = 1) +
  # the arrows resulted in the wrong direction being implied (towards increased age)
  # so were added manually later in the correct direction
  scale_x_reverse(limit=c(2, -2)) +
  scale_y_reverse(limit=c(2, -2)) +
  theme_bw()
```

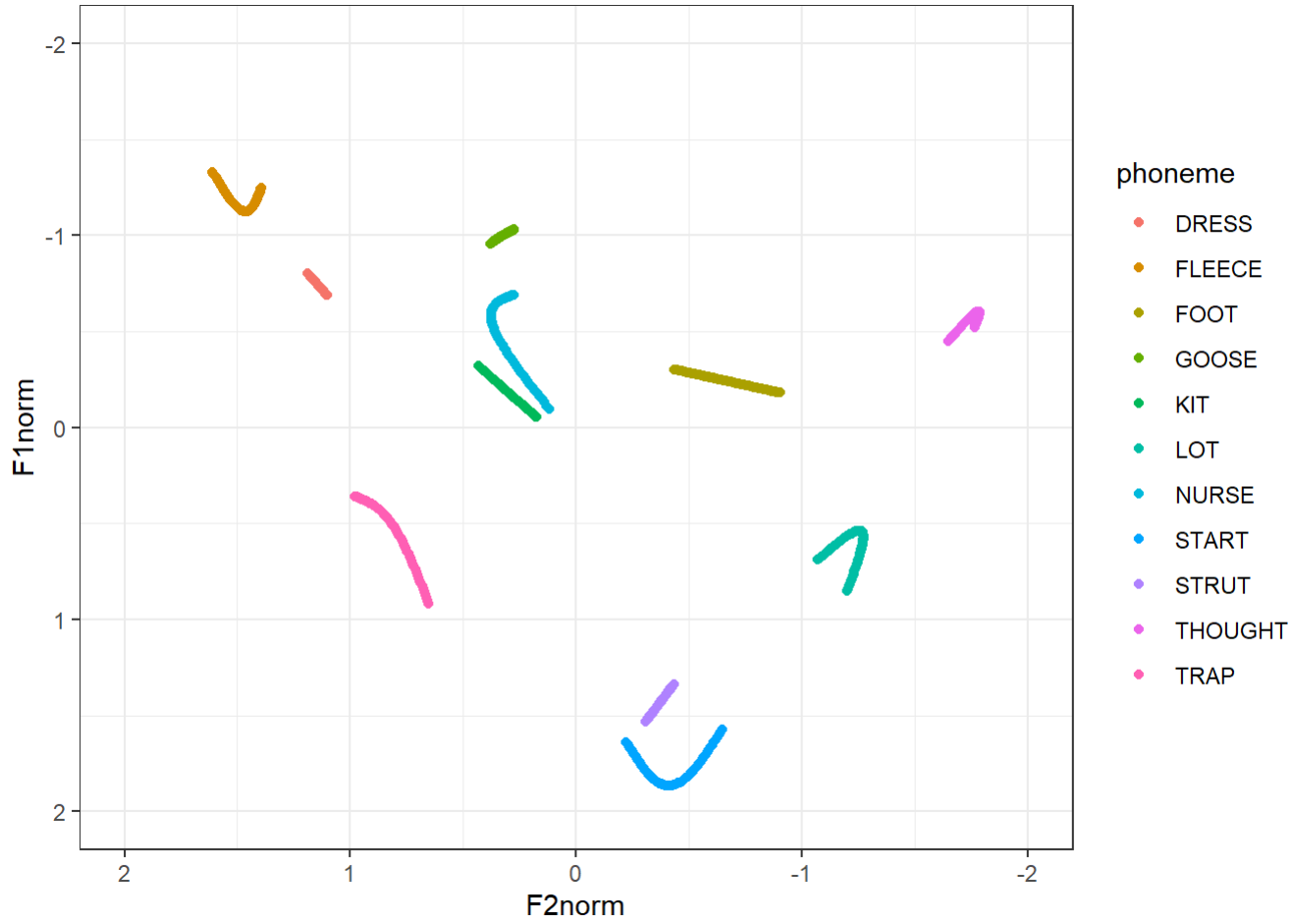
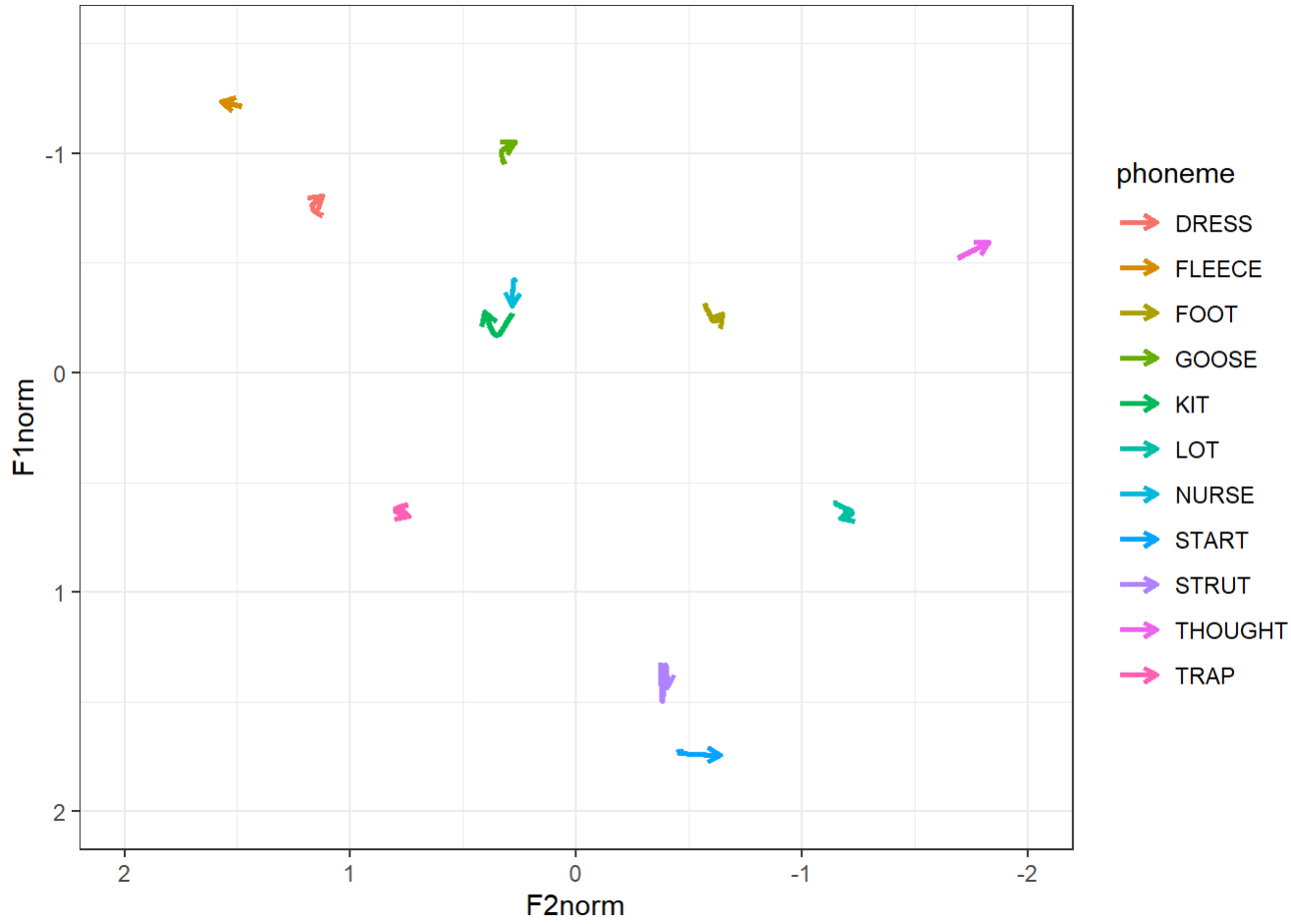


Figure X: predicted movements of monophthongs over the course of the interview

```
vowelsbytime <- read.csv("allvowels_timenorm.csv")

vowelsbytime <- vowelsbytime %>%
  mutate(
    phoneme <- as.factor(phoneme))

vowelsbytime <- vowelsbytime %>%
  mutate(
    Gender <- as.factor(Gender))


vowelsbytime <- vowelsbytime %>%
  pivot_longer(
    cols = F1norm:F2norm,
    names_to = "Formant",
    values_to = "Frequency"
  )

vowelsbytime <- vowelsbytime %>%
  group_by(phoneme, Formant) %>%
  nest()

vowelsbytime <- vowelsbytime %>%
  mutate(
    model = map(
      data,
      ~ bam(Frequency ~ Gender +
          s(Time_norm, k=5),
        data = .x,
        method = 'fREML',
        discrete = TRUE,
        nthreads = 2
      )
    )
  )



to_predict <- list(
  "Time_norm" = seq(0.0, 1.0, 0.0001))

vowelsbytime <- vowelsbytime %>%
  mutate(
    prediction = map(
      model,
      ~ get_predictions(model = .x, cond = to_predict, print.summary = FALSE)
    )
  )

predictions <- vowelsbytime %>%
  select(
    phoneme, Formant, prediction
  ) %>%
  unnest(prediction)
```

```
predictions <- predictions %>%
  select(
    -CI
  ) %>%
  pivot_wider( # Pivot
    names_from = Formant,
    values_from = fit
  )

predictions %>%
  ggplot(
    aes(
      x = F2norm,
      y = F1norm,
      colour = phoneme
    )
  ) +
  geom_path(arrow=arrow(length= unit(2, "mm")), linewidth = 1) +
  scale_x_reverse(limit=c(2, -2)) +
  scale_y_reverse(limit=c(2, -1.5)) +
  theme_bw()
```



Stats for change in normalised F1 and F2 over the course of the interviews:

```
forcheck <- read.csv("allvowels.csv")
library(lme4)
library(gtsummary)

dressF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"DRESS"))

dressF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.08	-0.16, 0.01	0.089

[1] CI = Confidence Interval

```
dressF2accom <- lmer(F2norm~Time_norm + (1|ID), data = filter(forcheck, phoneme == "DRESS"))

dressF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.01	-0.08, 0.10	0.9

[1] CI = Confidence Interval

TRAP:

```
trapF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"TRAP"))

trapF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.06	-0.08, 0.21	0.4

[1] CI = Confidence Interval

```
trapF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"TRAP"))

trapF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.04	-0.03, 0.11	0.3

[1] CI = Confidence Interval

KIT:

```
kitF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "K
IT"))

kitF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.07	-0.03, 0.17	0.2

[1] CI = Confidence Interval

```
kitF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "K
IT"))

kitF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.08	0.00, 0.17	0.056

[1] CI = Confidence Interval

START:

```
startF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"START"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
startF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.04	-0.10, 0.17	0.6

Characteristic	Beta	95% CI[1]	p-value

[1] CI = Confidence Interval

```
startF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"START"))

startF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.08	-0.14, -0.02	0.010

[1] CI = Confidence Interval

STRUT:

```
strutF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"STRUT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
strutF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.07	-0.25, 0.10	0.4

[1] CI = Confidence Interval

```
strutF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"STRUT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
strutF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.03	-0.12, 0.06	0.5

[1] CI = Confidence Interval

NURSE:

```
nurseF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "NURSE"))

nurseF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.02	-0.08, 0.12	0.7

[1] CI = Confidence Interval

```
nurseF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "NURSE"))

nurseF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.04	-0.03, 0.11	0.2

[1] CI = Confidence Interval

FOOT:

```
footF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "FOOT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
footF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.06	-0.08, 0.20	0.4

[1] CI = Confidence Interval

```
footF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "FOOT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
footF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.08	-0.23, 0.08	0.3

[1] CI = Confidence Interval

LOT:

```
lotF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "L
OT"))

lotF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.07	-0.04, 0.18	0.2

[1] CI = Confidence Interval

```
lotF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme == "L
OT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
lotF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.09	-0.16, -0.01	0.020

[1] CI = Confidence Interval

FLEECE:

```
fleeceF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"FLEECE"))

fleeceF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.04	-0.16, 0.08	0.6

[1] CI = Confidence Interval

```
fleeceF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"FLEECE"))

fleeceF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	0.12	-0.02, 0.25	0.090

[1] CI = Confidence Interval

GOOSE:

```
gooseF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"GOOSE"))

gooseF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.10	-0.21, 0.01	0.077

[1] CI = Confidence Interval

```
gooseF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme ==
"GOOSE"))

gooseF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.06	-0.18, 0.05	0.3

[1] CI = Confidence Interval

THOUGHT:

```
thoughtF1accom <- lmer(F1norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme =
= "THOUGHT"))
```

```
## boundary (singular) fit: see help('isSingular')
```

```
thoughtF1accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.03	-0.15, 0.10	0.7

[1] CI = Confidence Interval

```
thoughtF2accom <- lmer(F2norm~Time_norm + (1|ID) + (1|Age), data = filter(forcheck, phoneme =
= "THOUGHT"))

thoughtF2accom %>%
  tbl_regression() %>%
  add_global_p()
```

Characteristic	Beta	95% CI[1]	p-value
Time_norm	-0.09	-0.17, -0.01	0.037

[1] CI = Confidence Interval