baaplots

Massimiliano Canzi

15/03/2018

data <- read\_csv('Madata.csv') %>%  
 filter(F1 < 1500) %>% filter(Vowel != 'u') %>% group\_by(Speaker) %>% mutate(F0\_z = scale(F0)) %>% filter(F0\_z > -2) %>%  
 mutate\_if(is.character, as.factor)

## Parsed with column specification:  
## cols(  
## Speaker = col\_integer(),  
## Gen = col\_character(),  
## Mode = col\_character(),  
## Vowel = col\_character(),  
## Long = col\_character(),  
## Duration = col\_double(),  
## F0 = col\_integer(),  
## F1 = col\_integer(),  
## F2 = col\_integer()  
## )

## Warning in mutate\_impl(.data, dots): Unequal factor levels: coercing to  
## character

## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector

F0\_lmer <- lmer(  
 F0 ~ Vowel \* Mode + (1+Mode|Speaker),  
 data  
)  
summary(F0\_lmer)

## Linear mixed model fit by REML t-tests use Satterthwaite approximations  
## to degrees of freedom [lmerMod]  
## Formula: F0 ~ Vowel \* Mode + (1 + Mode | Speaker)  
## Data: data  
##   
## REML criterion at convergence: 26686.4  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -5.4056 -0.3366 0.0059 0.3730 3.4786   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## Speaker (Intercept) 1934.7 43.98   
## ModeNeutral 511.0 22.60 -0.20   
## ModeSoft 623.5 24.97 -0.20 0.92  
## Residual 335.4 18.31   
## Number of obs: 3079, groups: Speaker, 8  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 192.5104 15.6070 7.1000 12.335 4.72e-06 \*\*\*  
## Vowele 13.7240 1.8691 3040.3000 7.343 2.68e-13 \*\*\*  
## Vowelɛ 3.5156 1.8691 3040.3000 1.881 0.060076 .   
## Voweli 16.7165 1.8766 3040.3000 8.908 < 2e-16 \*\*\*  
## Vowelo 18.0156 1.8691 3040.3000 9.639 < 2e-16 \*\*\*  
## Vowelɔ 17.0104 1.8691 3040.3000 9.101 < 2e-16 \*\*\*  
## ModeNeutral -56.6701 8.2086 7.6000 -6.904 0.000153 \*\*\*  
## ModeSoft -55.1094 9.0572 7.4000 -6.085 0.000394 \*\*\*  
## Vowele:ModeNeutral -6.0001 2.6503 3040.3000 -2.264 0.023647 \*   
## Vowelɛ:ModeNeutral -0.8367 2.6502 3040.3000 -0.316 0.752252   
## Voweli:ModeNeutral -2.5262 2.6685 3040.3000 -0.947 0.343896   
## Vowelo:ModeNeutral -2.8751 2.6485 3040.3000 -1.086 0.277753   
## Vowelɔ:ModeNeutral -7.6842 2.6576 3040.3000 -2.891 0.003862 \*\*   
## Vowele:ModeSoft -12.0580 2.9215 3045.3000 -4.127 3.77e-05 \*\*\*  
## Vowelɛ:ModeSoft -2.1032 2.8674 3043.3000 -0.733 0.463322   
## Voweli:ModeSoft -9.3571 2.9310 3044.9000 -3.193 0.001425 \*\*   
## Vowelo:ModeSoft -12.9348 2.8748 3043.7000 -4.499 7.07e-06 \*\*\*  
## Vowelɔ:ModeSoft -14.0966 2.8663 3042.2000 -4.918 9.21e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Correlation matrix not shown by default, as p = 18 > 12.  
## Use print(x, correlation=TRUE) or  
## vcov(x) if you need it

F1\_lmer <- lmer(  
 F1 ~ Vowel \* Mode + (1+Mode|Speaker),  
 data  
)  
summary(F1\_lmer)

## Linear mixed model fit by REML t-tests use Satterthwaite approximations  
## to degrees of freedom [lmerMod]  
## Formula: F1 ~ Vowel \* Mode + (1 + Mode | Speaker)  
## Data: data  
##   
## REML criterion at convergence: 36092.6  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -6.4261 -0.5720 -0.0300 0.5208 7.9318   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## Speaker (Intercept) 7352 85.74   
## ModeNeutral 1878 43.33 -0.77   
## ModeSoft 2018 44.92 -0.67 0.82  
## Residual 7335 85.64   
## Number of obs: 3079, groups: Speaker, 8  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 952.188 30.938 7.500 30.777 3.69e-09 \*\*\*  
## Vowele -439.401 8.741 3040.000 -50.269 < 2e-16 \*\*\*  
## Vowelɛ -179.870 8.741 3040.000 -20.578 < 2e-16 \*\*\*  
## Voweli -588.110 8.776 3040.000 -67.011 < 2e-16 \*\*\*  
## Vowelo -313.167 8.741 3040.000 -35.827 < 2e-16 \*\*\*  
## Vowelɔ -146.380 8.741 3040.000 -16.746 < 2e-16 \*\*\*  
## ModeNeutral -134.418 17.650 11.100 -7.616 9.92e-06 \*\*\*  
## ModeSoft -195.653 18.461 10.800 -10.598 4.84e-07 \*\*\*  
## Vowele:ModeNeutral 81.251 12.394 3040.000 6.556 6.49e-11 \*\*\*  
## Vowelɛ:ModeNeutral 21.375 12.394 3040.000 1.725 0.08470 .   
## Voweli:ModeNeutral 120.436 12.480 3040.200 9.651 < 2e-16 \*\*\*  
## Vowelo:ModeNeutral 29.128 12.386 3040.000 2.352 0.01875 \*   
## Vowelɔ:ModeNeutral 13.334 12.429 3040.100 1.073 0.28342   
## Vowele:ModeSoft 108.906 13.649 3047.300 7.979 2.22e-15 \*\*\*  
## Vowelɛ:ModeSoft 58.889 13.403 3047.100 4.394 1.15e-05 \*\*\*  
## Voweli:ModeSoft 161.336 13.695 3047.900 11.781 < 2e-16 \*\*\*  
## Vowelo:ModeSoft -2.955 13.436 3047.500 -0.220 0.82597   
## Vowelɔ:ModeSoft -47.322 13.401 3045.200 -3.531 0.00042 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Correlation matrix not shown by default, as p = 18 > 12.  
## Use print(x, correlation=TRUE) or  
## vcov(x) if you need it

F2\_lmer <- lmer(  
 F2 ~ Vowel \* Mode + (1+Mode|Speaker),  
 data  
)  
summary(F2\_lmer)

## Linear mixed model fit by REML t-tests use Satterthwaite approximations  
## to degrees of freedom [lmerMod]  
## Formula: F2 ~ Vowel \* Mode + (1 + Mode | Speaker)  
## Data: data  
##   
## REML criterion at convergence: 41141.5  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -7.7498 -0.4535 -0.0406 0.3950 8.6977   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## Speaker (Intercept) 9877 99.38   
## ModeNeutral 4707 68.61 -0.82   
## ModeSoft 5235 72.35 -0.27 0.52  
## Residual 48210 219.57   
## Number of obs: 3028, groups: Speaker, 8  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 1409.005 38.546 9.400 36.554 1.72e-11 \*\*\*  
## Vowele 740.240 22.410 2996.500 33.032 < 2e-16 \*\*\*  
## Vowelɛ 465.687 22.410 2996.500 20.781 < 2e-16 \*\*\*  
## Voweli 897.494 23.360 3001.400 38.421 < 2e-16 \*\*\*  
## Vowelo -257.115 22.410 2996.500 -11.473 < 2e-16 \*\*\*  
## Vowelɔ -265.000 22.410 2996.500 -11.825 < 2e-16 \*\*\*  
## ModeNeutral -100.608 33.065 18.300 -3.043 0.0069 \*\*   
## ModeSoft -75.034 35.149 18.400 -2.135 0.0465 \*   
## Vowele:ModeNeutral 44.785 31.775 2996.500 1.409 0.1588   
## Vowelɛ:ModeNeutral -26.661 31.775 2996.500 -0.839 0.4015   
## Voweli:ModeNeutral -5.623 32.789 3000.100 -0.171 0.8639   
## Vowelo:ModeNeutral -37.771 31.754 2996.500 -1.189 0.2343   
## Vowelɔ:ModeNeutral 7.449 31.863 2996.700 0.234 0.8152   
## Vowele:ModeSoft 29.722 35.170 3001.300 0.845 0.3981   
## Vowelɛ:ModeSoft -5.473 34.356 3002.300 -0.159 0.8734   
## Voweli:ModeSoft -63.339 36.554 3004.500 -1.733 0.0832 .   
## Vowelo:ModeSoft -197.061 34.441 3002.500 -5.722 1.16e-08 \*\*\*  
## Vowelɔ:ModeSoft -68.641 34.352 3000.900 -1.998 0.0458 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## Correlation matrix not shown by default, as p = 18 > 12.  
## Use print(x, correlation=TRUE) or  
## vcov(x) if you need it

durData <- read\_csv('Madata.csv') %>% group\_by(Speaker) %>%  
 mutate\_if(is.character, as.factor)

## Parsed with column specification:  
## cols(  
## Speaker = col\_integer(),  
## Gen = col\_character(),  
## Mode = col\_character(),  
## Vowel = col\_character(),  
## Long = col\_character(),  
## Duration = col\_double(),  
## F0 = col\_integer(),  
## F1 = col\_integer(),  
## F2 = col\_integer()  
## )

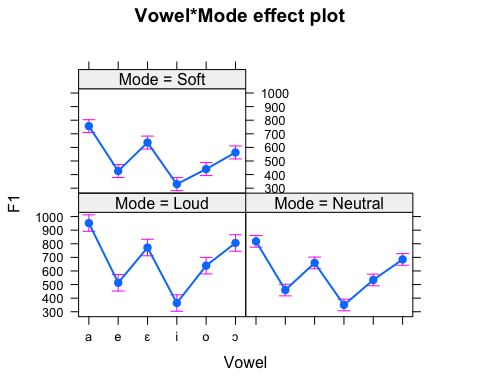
## Warning in mutate\_impl(.data, dots): Unequal factor levels: coercing to  
## character

## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector  
  
## Warning in mutate\_impl(.data, dots): binding character and factor vector,  
## coercing into character vector

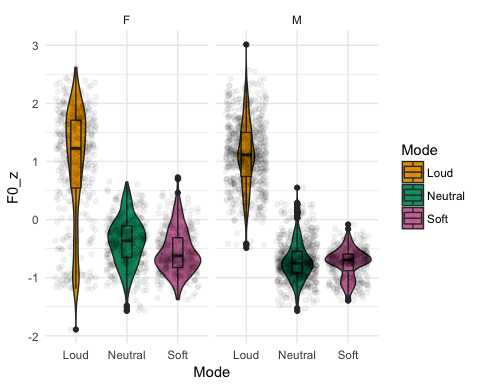
Dur\_lmer <- lmer(  
 Duration ~ Long \* Mode + (1+Mode|Speaker),  
 durData  
)  
summary(Dur\_lmer)

## Linear mixed model fit by REML t-tests use Satterthwaite approximations  
## to degrees of freedom [lmerMod]  
## Formula: Duration ~ Long \* Mode + (1 + Mode | Speaker)  
## Data: durData  
##   
## REML criterion at convergence: -16482.6  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -4.1920 -0.6253 -0.0781 0.5815 4.6230   
##   
## Random effects:  
## Groups Name Variance Std.Dev. Corr   
## Speaker (Intercept) 3.079e-04 0.017548   
## ModeNeutral 7.236e-05 0.008506 -0.11   
## ModeSoft 3.077e-04 0.017542 -0.72 0.20  
## Residual 9.501e-04 0.030824   
## Number of obs: 4032, groups: Speaker, 8  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)  
## (Intercept) 1.744e-01 6.317e-03 7.000e+00 27.609 1.29e-08  
## LongShort -3.052e-02 1.682e-03 4.005e+03 -18.150 < 2e-16  
## ModeNeutral -3.309e-02 3.446e-03 9.000e+00 -9.605 4.92e-06  
## ModeSoft -3.554e-02 6.426e-03 8.000e+00 -5.530 0.000691  
## LongShort:ModeNeutral 5.595e-03 2.378e-03 4.005e+03 2.353 0.018680  
## LongShort:ModeSoft 2.292e-03 2.378e-03 4.005e+03 0.964 0.335277  
##   
## (Intercept) \*\*\*  
## LongShort \*\*\*  
## ModeNeutral \*\*\*  
## ModeSoft \*\*\*  
## LongShort:ModeNeutral \*   
## LongShort:ModeSoft   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) LngShr MdNtrl ModSft LnS:MN  
## LongShort -0.133   
## ModeNeutral -0.156 0.244   
## ModeSoft -0.718 0.131 0.230   
## LngShrt:MdN 0.094 -0.707 -0.345 -0.093   
## LngShrt:MdS 0.094 -0.707 -0.173 -0.185 0.500

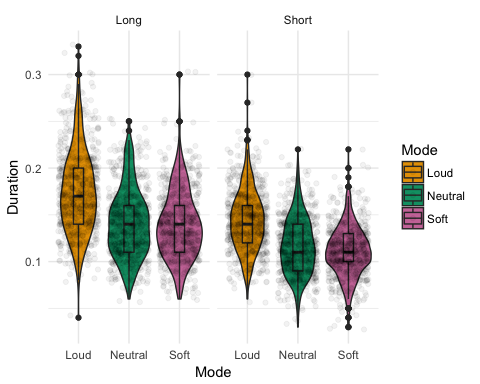
plot(allEffects(F1\_lmer))



F0\_plot <- ggplot(data, aes(Mode, F0\_z, fill = Mode)) +  
 geom\_violin() +  
 geom\_boxplot(width=0.2) +  
 geom\_jitter(alpha=0.05) +  
 facet\_grid(~Gen)  
  
print(F0\_plot + scale\_fill\_manual(values = cbPalette))



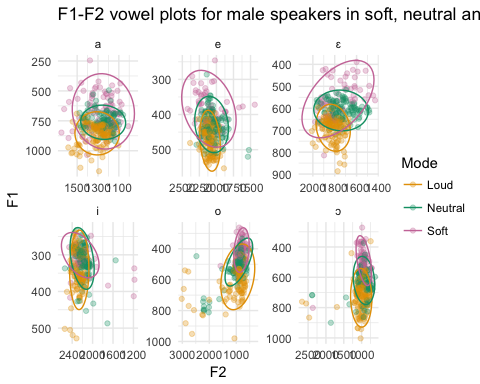
Dur\_plot <- ggplot(durData, aes(Mode, Duration, fill = Mode)) +  
 geom\_violin() +  
 geom\_boxplot(width=0.2) +  
 geom\_jitter(alpha=0.05) +  
 facet\_grid(~Long)  
  
print(Dur\_plot + scale\_fill\_manual(values = cbPalette))



FM\_plot <- filter(data, Gen == "M") %>%  
ggplot(aes(F2, F1, colour = Mode)) +  
 geom\_point(alpha = 0.3) +  
 stat\_ellipse() +  
 scale\_x\_reverse() +  
 scale\_y\_reverse() +   
 facet\_wrap(~Vowel, scales = 'free') +  
 labs(title = "F1-F2 vowel plots for male speakers in soft, neutral and loud speech modes")  
  
print(FM\_plot + scale\_colour\_manual(values = cbPalette))

## Warning: Removed 13 rows containing non-finite values (stat\_ellipse).

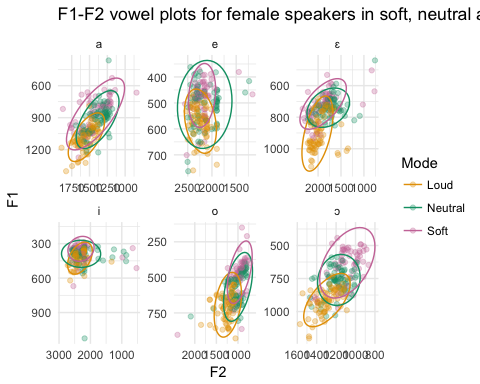
## Warning: Removed 13 rows containing missing values (geom\_point).



FF\_plot <- filter(data, Gen == "F") %>%  
ggplot(aes(F2, F1, colour = Mode)) +  
 geom\_point(alpha = 0.3) +  
 stat\_ellipse() +  
 scale\_x\_reverse() +  
 scale\_y\_reverse() +   
 facet\_wrap(~Vowel, scales = 'free') +  
 labs(title = "F1-F2 vowel plots for female speakers in soft, neutral and loud speech modes")  
  
print(FF\_plot + scale\_colour\_manual(values = cbPalette))

## Warning: Removed 38 rows containing non-finite values (stat\_ellipse).

## Warning: Removed 38 rows containing missing values (geom\_point).



Dur\_hist <- read\_csv('Madata.csv') %>% ggplot(aes(Vowel, Duration, fill = Vowel)) +  
 geom\_histogram(stat = 'identity') +  
 facet\_grid(~Mode)

## Parsed with column specification:  
## cols(  
## Speaker = col\_integer(),  
## Gen = col\_character(),  
## Mode = col\_character(),  
## Vowel = col\_character(),  
## Long = col\_character(),  
## Duration = col\_double(),  
## F0 = col\_integer(),  
## F1 = col\_integer(),  
## F2 = col\_integer()  
## )

## Warning: Ignoring unknown parameters: binwidth, bins, pad

print(Dur\_hist + scale\_fill\_manual(values = cbbPalette))

