

# Analysis II

“Aerosol emission is increased in professional singing”

Dirk Mürbe<sup>1</sup>, Martin Kriegel<sup>2</sup>, Julia Lange<sup>2</sup>, Hansjörg Rotheudt<sup>2</sup>, Mario Fleischer<sup>1</sup>

<sup>1</sup>Charité – Universitätsmedizin Berlin, Department of Audiology and Phoniatics, Berlin, Germany,

<sup>2</sup>Technische Universität Berlin, Hermann-Rietschel-Institut, Berlin, Germany

## Purpose

Linear mixed modeling of experiment II  
(sustained vowel /a/ in piano, mezzoforte, forte).

Analysis is inspired by Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499

The advantages are

- 1) Using mixed modeling no (pre-)averaging is necessary
- 2) Condition & Gender (& Interfactorial effects) were modeled as fixed effects
- 3) ID was modeled as random effects (intercept and slope for Gender and Condition)

author: mario.fleischer@charite.de

```
rm(list = ls())
library(ggpubr)

## Loading required package: ggplot2

library(gridExtra)
library(lme4)

## Loading required package: Matrix

## Registered S3 methods overwritten by 'lme4':
##   method                from
##   cooks.distance.influence.merMod car
##   influence.merMod       car
##   dfbeta.influence.merMod car
##   dfbetas.influence.merMod car

library(redres) # https://github.com/goodekat/redres
library(GGally)

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg    ggplot2

library(ggResidpanel)
```

## Loading data

```
# Set working directory according to your file hierarchy.  
orig.data <- read.csv('Data.csv')
```

## log-transform of emission rate and replacing zeros by NaN

```
PM.log <- orig.data$PM  
PM.log[PM.log == 0] <- NA  
orig.data$PM.log = log10(PM.log)  
length(which(is.na(PM.log)==T)) # Number of NaN detected
```

```
## [1] 12
```

## Selection of relevant data

```
exp.data.II <- dplyr::filter(orig.data,  
                             Condition == 'piano'  
                             | Condition == 'mezzoforte'  
                             | Condition == 'forte'  
                             )
```

## Test 0: Deviation from normality and homoscedasticity?

```
II00.model <- lmer(PM~Condition*Gender  
                  +(1+Condition|ID)  
                  +(1+Gender|ID),  
                  exp.data.II,  
                  REML=FALSE,  
                  control = lmerControl(optimizer = 'optimx',  
                                         optCtrl=list(method='nlminb'))  
                  )
```

```
## Loading required namespace: optimx
```

```
## boundary (singular) fit: see ?isSingular
```

```
# (intermediate) results:  
# 1) Deviation from normality and homoscedasticity was observed
```

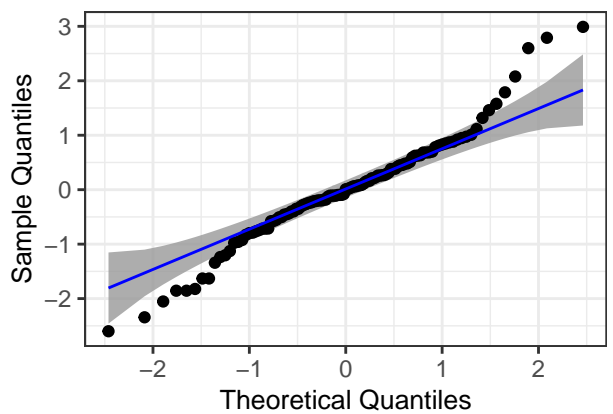
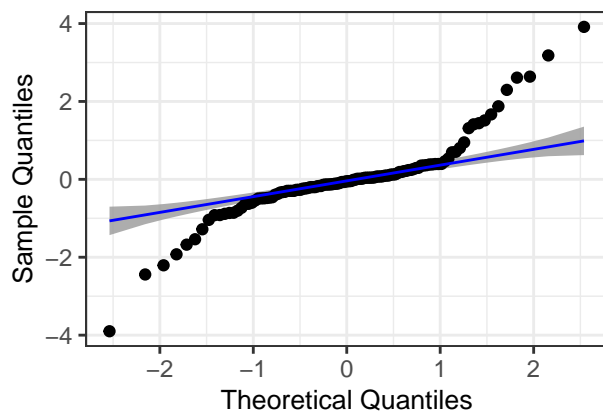
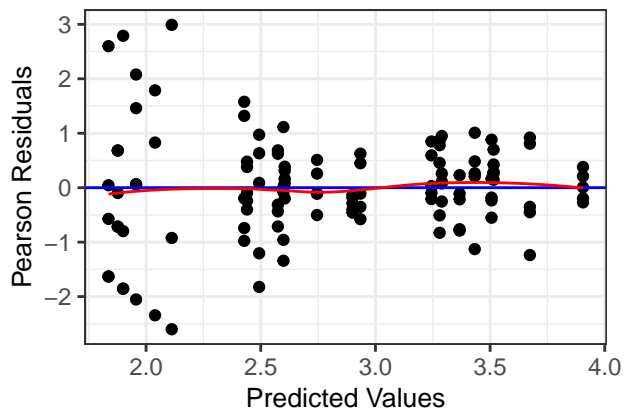
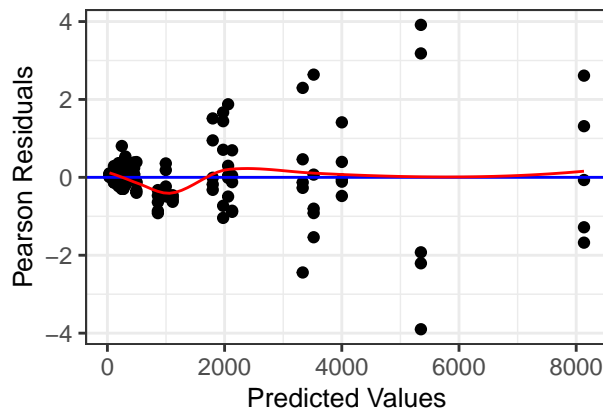
```
II0.model <- lmer(PM.log~Condition*Gender  
                 +(1+Condition|ID)  
                 +(1+Gender|ID),  
                 exp.data.II,  
                 REML=FALSE,  
                 control = lmerControl(optimizer = 'optimx',  
                                         optCtrl=list(method='nlminb'))  
                 )
```

```
## boundary (singular) fit: see ?isSingular
```

## Comparing model with and without log-transform of $P_M$

```
resid_compare(models = list(II00.model,
                            II0.model
                          ),
              plots = c("resid", "qq"),
              smoother = TRUE,
              qqbands = TRUE,
              title.opt = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

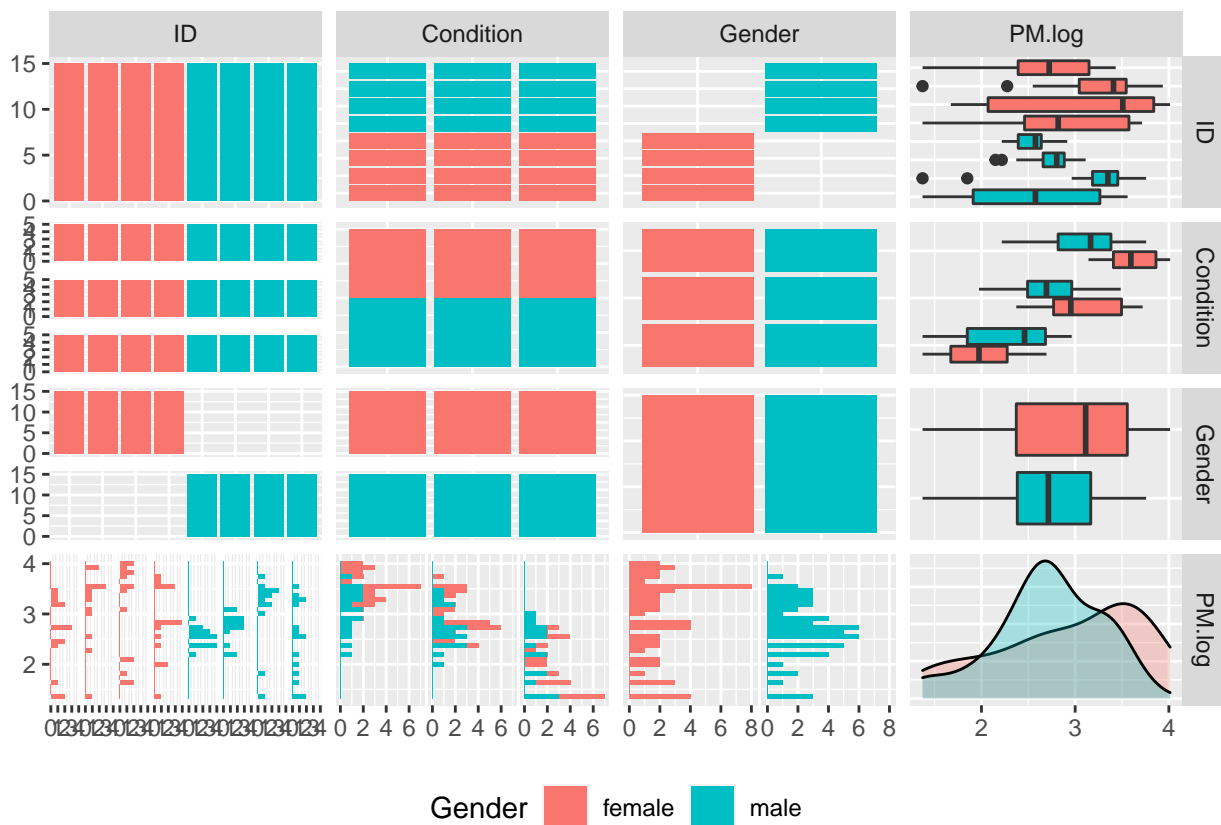


## Pair plot of the data

```
exp.data.II %>% ggpairs(.,
                      mapping = ggplot2::aes(colour=Gender),
                      legend = 1,
                      column = c('ID', 'Condition', 'Gender', 'PM.log'),
                      lower = list(continuous = wrap("smooth", alpha = 0.3, size=0.1)),
                      diag = list(discrete="barDiag",
                                  continuous = wrap("densityDiag", alpha=0.3))
                      ) + theme(legend.position = "bottom")
```

```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
## Warning: Removed 5 rows containing non-finite values (stat_boxplot).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 5 rows containing non-finite values (stat_bin).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 5 rows containing non-finite values (stat_bin).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 5 rows containing non-finite values (stat_bin).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 5 rows containing non-finite values (stat_bin).
## Warning: Removed 5 rows containing non-finite values (stat_bin).
## Warning: Removed 5 rows containing non-finite values (stat_density).
```



**Test 1: Is the interaction between Condition and Gender significant?**

```
II1.model <- II0.model

II1.null <- lmer(PM.log~Condition+Gender
                +(1+Condition|ID)
                +(1+Gender|ID),
                exp.data.II,
```

```

      REML=FALSE,
      control = lmerControl(optimizer = 'optimx',
                            optCtrl=list(method='nlminb'))
    )

## boundary (singular) fit: see ?isSingular
anova(II1.null, II1.model)

## Data: exp.data.II
## Models:
## II1.null: PM.log ~ Condition + Gender + (1 + Condition | ID) + (1 + Gender |
## II1.null:      ID)
## II1.model: PM.log ~ Condition * Gender + (1 + Condition | ID) + (1 + Gender |
## II1.model:      ID)
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## II1.null   14 102.53 140.96 -37.266   74.531
## II1.model   16 102.64 146.55 -35.317   70.635 3.8961 2     0.1426

# (intermediate) results:
# 1) interaction between Condition and Gender IS NOT significant.
# 2) Condition:Gender CAN be omitted!
#####

```

## Test 2: Is Condition significant?

```

II2.model <- II1.null

II2.null <- lmer(PM.log~Gender
                +(1+Condition|ID)
                +(1+Gender|ID),
                exp.data.II,
                REML=FALSE,
                control = lmerControl(optimizer = 'optimx',
                                      optCtrl=list(method='nlminb'))
    )

## boundary (singular) fit: see ?isSingular
anova(II2.null, II2.model)

## Data: exp.data.II
## Models:
## II2.null: PM.log ~ Gender + (1 + Condition | ID) + (1 + Gender | ID)
## II2.model: PM.log ~ Condition + Gender + (1 + Condition | ID) + (1 + Gender |
## II2.model:      ID)
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## II2.null   12 111.00 143.94 -43.500   87.001
## II2.model   14 102.53 140.96 -37.266   74.531 12.47 2     0.00196 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(II2.model)

## Linear mixed model fit by maximum likelihood ['lmerMod']

```

```

## Formula: PM.log ~ Condition + Gender + (1 + Condition | ID) + (1 + Gender |
## ID)
## Data: exp.data.II
## Control: lmerControl(optimizer = "optimx", optCtrl = list(method = "nlminb"))
##
## AIC BIC logLik deviance df.resid
## 102.5 141.0 -37.3 74.5 101
##
## Scaled residuals:
## Min 1Q Median 3Q Max
## -2.44826 -0.44266 -0.02624 0.48208 3.13991
##
## Random effects:
## Groups Name Variance Std.Dev. Corr
## ID (Intercept) 1.410e-01 3.756e-01
## Conditionmezzoforte 6.785e-02 2.605e-01 -0.39
## Conditionpiano 4.155e-01 6.446e-01 -0.97 0.61
## ID.1 (Intercept) 7.848e-12 2.801e-06
## Gendermale 8.520e-12 2.919e-06 -1.00
## Residual 8.107e-02 2.847e-01
## Number of obs: 115, groups: ID, 8
##
## Fixed effects:
## Estimate Std. Error t value
## (Intercept) 3.38416 0.14737 22.964
## Conditionmezzoforte -0.45994 0.11196 -4.108
## Conditionpiano -1.25514 0.23734 -5.288
## Gendermale -0.07071 0.09078 -0.779
##
## Correlation of Fixed Effects:
## (Intr) Cndtnm Cndtnp
## Cndtnmzzfirt -0.408
## Conditionpn -0.892 0.559
## Gendermale -0.308 0.000 -0.004
## convergence code: 0
## boundary (singular) fit: see ?isSingular

```

```

# (intermediate) results:
# 1) Condition IS significant. Condition affected log10(PM) ( $\chi^2(2)=12.47, p=0.00196$ ),
# lowering it by a factor of -0.45994 P/s+-0.11196 (standard errors)
# from forte to mezzoforte and by a factor of -1.25514+-0.23734 from forte to
# piano
#####

```

### Test 3: Is Gender significant?

```

II3.model <- II1.null

II3.null <- lmer(PM.log~Condition
                +(1+Condition|ID)
                +(1+Gender|ID),
                exp.data.II,

```

```

REML=FALSE,
control = lmerControl(optimizer = 'optimx',
                      optCtrl=list(method='nlminb'))
)

```

```
## boundary (singular) fit: see ?isSingular
```

```
anova(II3.null, II3.model)
```

```

## Data: exp.data.II
## Models:
## II3.null: PM.log ~ Condition + (1 + Condition | ID) + (1 + Gender | ID)
## II3.model: PM.log ~ Condition + Gender + (1 + Condition | ID) + (1 + Gender |
## II3.model: ID)
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## II3.null   13 100.90 136.58 -37.448   74.896
## II3.model  14 102.53 140.96 -37.266   74.531 0.3644  1    0.546

```

```

# (intermediate) results:
# 1) Gender IS NOT significant. Gender affected log10(PM) ( $X_i \sim 2(1)=0.3644, p=0.546$ )
#####

```

## By-subject analysis

```
coef(II3.model)
```

```

## $ID
## (Intercept) Conditionmezzoforte Conditionpiano Gendermale
## S1 3.314580 -0.7749093 -1.42340719 -0.07071285
## S2 3.925158 -0.3739972 -1.59322863 -0.07071285
## S3 4.412974 -0.4655094 -2.01584325 -0.07071285
## S4 3.615122 -0.6011372 -1.52341821 -0.07071285
## S5 1.942733 -0.1532245 0.01893116 -0.07071285
## S6 2.635843 -0.1935906 -0.51906772 -0.07071285
## S7 3.831866 -0.2933289 -1.46839368 -0.07071285
## S8 3.395002 -0.8238402 -1.51666077 -0.07071285
##
## attr(,"class")
## [1] "coef.mer"

```

```

# 1) all subjects
# showed an decrease of emitted Particles from forte to mezzoforte
# 2) all subjects (except for S5 - see Fig. 3 in manuscript)
# show an decrease of emitted Particles from forte to piano
# 3) all subjects show, that females emitted more particles than males (but
# this finding was NOT significant)

```