

French_Stat_analyses

Arun Prakash Singh

2024-10-24

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

1. Load necessary libraries

```
library(readxl)
library(lme4)
```

```
## Loading required package: Matrix
library(lmerTest)

##
## Attaching package: 'lmerTest'
## The following object is masked from 'package:lme4':
##   lmer
## The following object is masked from 'package:stats':
##   step
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##   filter, lag
## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union
library(crayon)
```

```
##
## Attaching package: 'crayon'
## The following object is masked from 'package:ggplot2':
##   %+%
```

```

library(merTools)

## Loading required package: arm
## Loading required package: MASS
##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##       select
##
## arm (Version 1.14-4, built: 2024-4-1)
## Working directory is C:/Users/arunps/OneDrive/Projects/Scripts/Python/French_IDS_ADS_DayLong_recordin

library/boot)

##
## Attaching package: 'boot'
## The following object is masked from 'package:arm':
##       logit
library/splines)
library/stringr)
library/cowplot)
library/car)

## Loading required package: carData
##
## Attaching package: 'car'
## The following object is masked from 'package:boot':
##       logit
## The following object is masked from 'package:arm':
##       logit
## The following object is masked from 'package:dplyr':
##       recode
library/glmmTMB)

## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.
## TMB was built with Matrix version 1.5.3
## Current Matrix version is 1.5.1
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a
## Warning in checkDepPackageVersion(dep_pkg = "TMB"): Package version inconsistency detected.
## glmmTMB was built with TMB version 1.9.11
## Current TMB version is 1.9.3
## Please re-install glmmTMB from source or restore original 'TMB' package (see '?reinstalling' for more

```

```

library(DHARMA)

## This is DHARMA 0.4.6. For overview type '?DHARMA'. For recent changes, type news(package = 'DHARMA')

#2.Pitch #2.1 ADS

df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_pitch_st.xlsx"))

## New names:
## * `` -> `...` 

output.2.1 <- lmer.full.reduced.null.compare(meanF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowels),
                                               optimizer = "bobyqa",
                                               maxfun = 1000000000,
                                               fullm = TRUE,
                                               redm = TRUE,
                                               redmformula = meanF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowels),
                                               nullm = TRUE,
                                               nulllmformula =meanF0st ~ SES + SEX + (1+AgeInDays|SPK_id) + (1|vowels),
                                               df_ADS)

## 
## Attaching package: 'emmeans'
##
## The following object is masked from 'package:crayon':
## 
##     inverse

## Structure of combined data:
## tibble [1,306 x 7] (S3:tbl_df/tbl/data.frame)
## $ ...1 : num [1:1306] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES    : num [1:1306] 5 5 5 5 5 5 5 5 5 ...
## $ SEX    : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1306] 120 120 120 120 120 120 120 120 120 ...
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...
## $ meanF0st : num [1:1306] 59.9 57.6 53.5 54.3 42.5 ...
## Response variable is: meanF0st

## Structure of combined data after scaling:
## tibble [1,306 x 7] (S3:tbl_df/tbl/data.frame)
## $ ...1 : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 -1.72 ...
## ..- attr(*, "scaled:center")= num 652
## ..- attr(*, "scaled:scale")= num 377
## $ SPK_id : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES    : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX    : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...
## ..- attr(*, "scaled:center")= num 295
## ..- attr(*, "scaled:scale")= num 97.1
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...
## $ meanF0st : num [1:1306] 59.9 57.6 53.5 54.3 42.5 ...

## Warning: `aes_string()` was deprecated in ggplot2 3.0.0.
## i Please use tidy evaluation idioms with `aes()`.

## i See also `vignette("ggplot2-in-packages")` for more information.

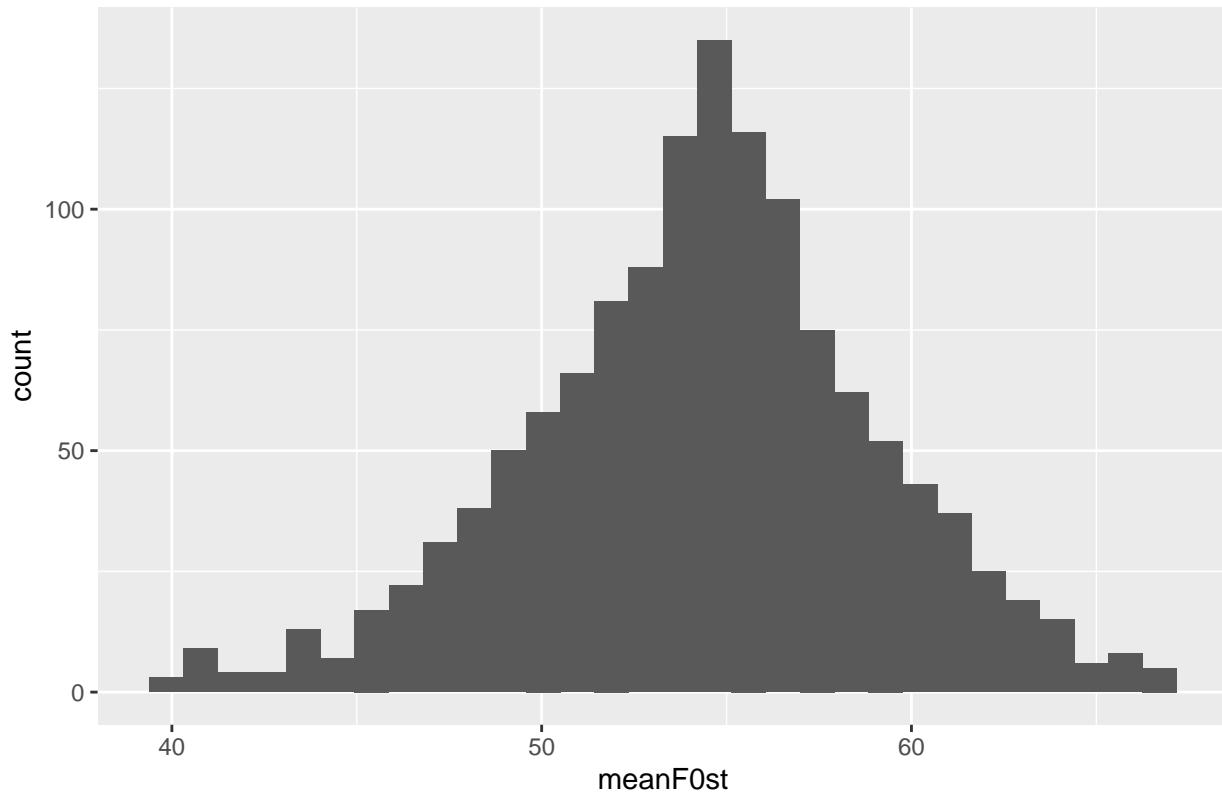
```

```

## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

```

Histogram of Response Variable



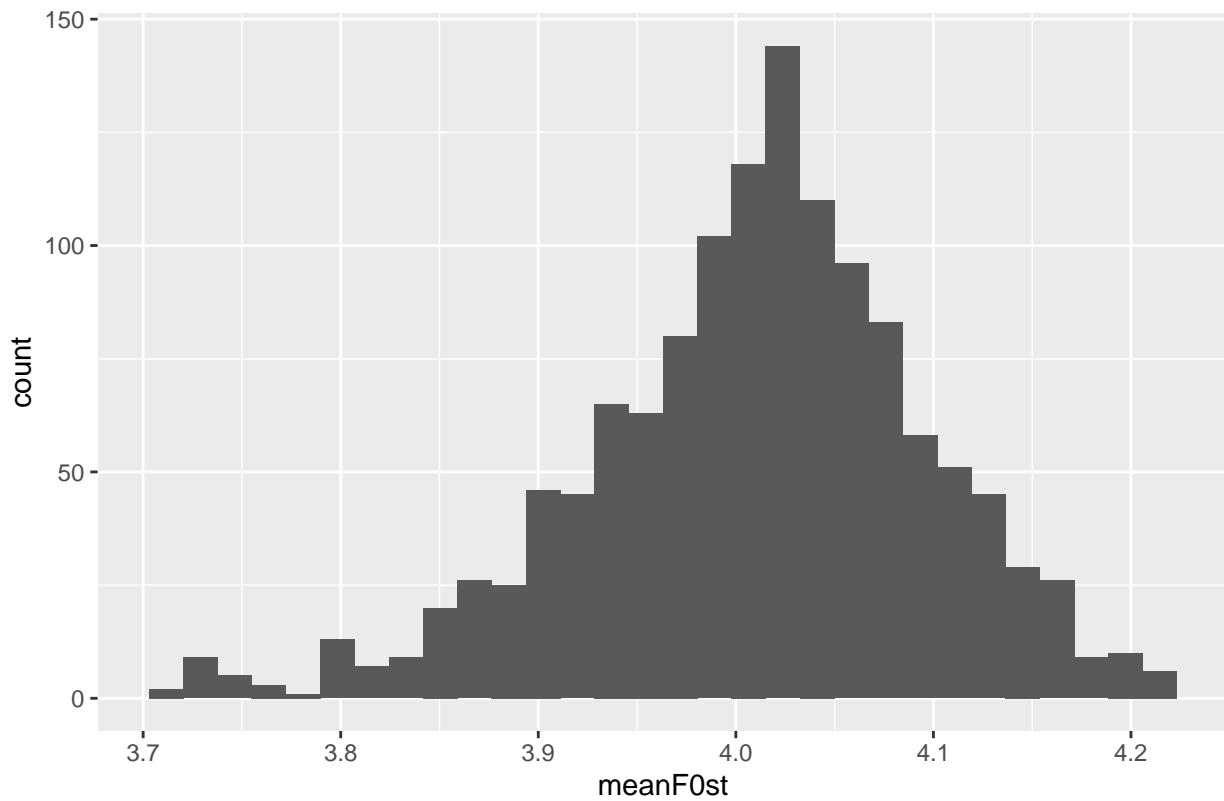
```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 ...
## ..- attr(*, "scaled:center")= num 652
## ..- attr(*, "scaled:scale")= num 377
## $ SPK_id : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...
## ..- attr(*, "scaled:center")= num 295
## ..- attr(*, "scaled:scale")= num 97.1
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...
## $ meanF0st : num [1:1306] 4.11 4.07 4 4.01 3.77 ...
## Fitting Full Model.

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

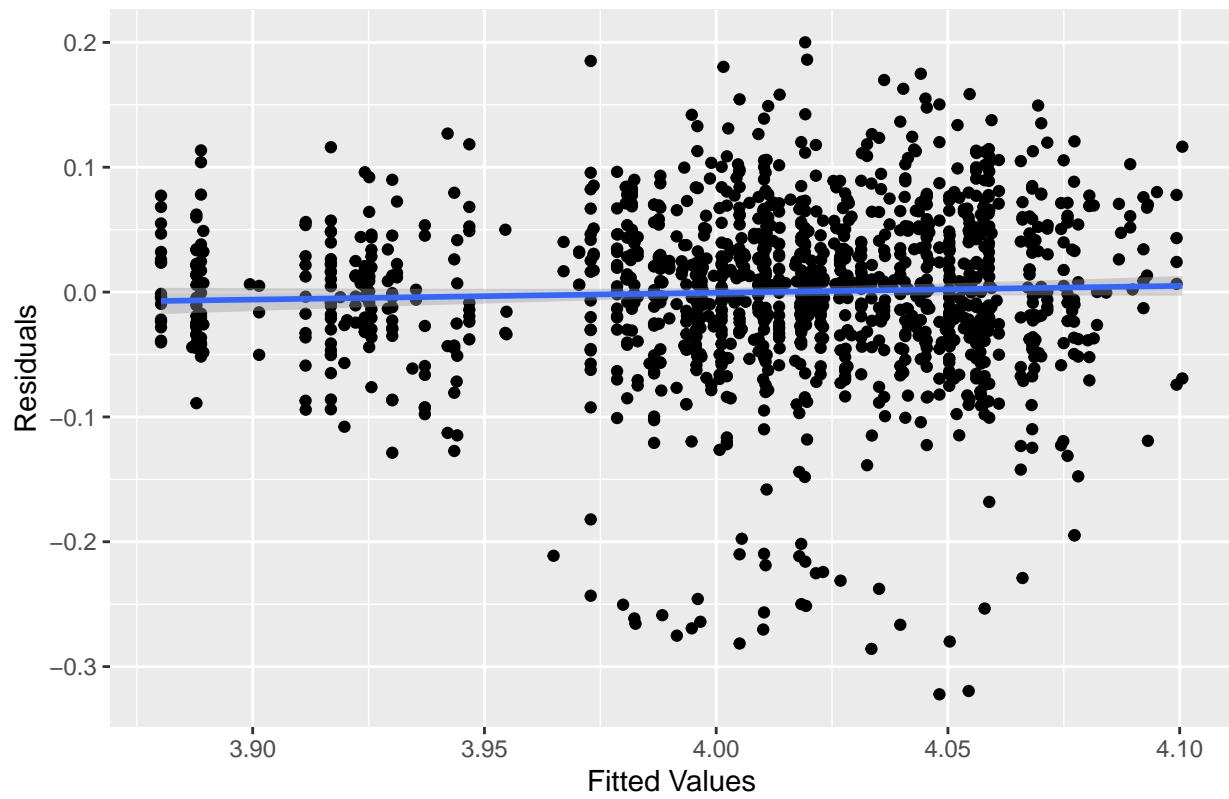
```

Histogram of Log-transformed Response Variable

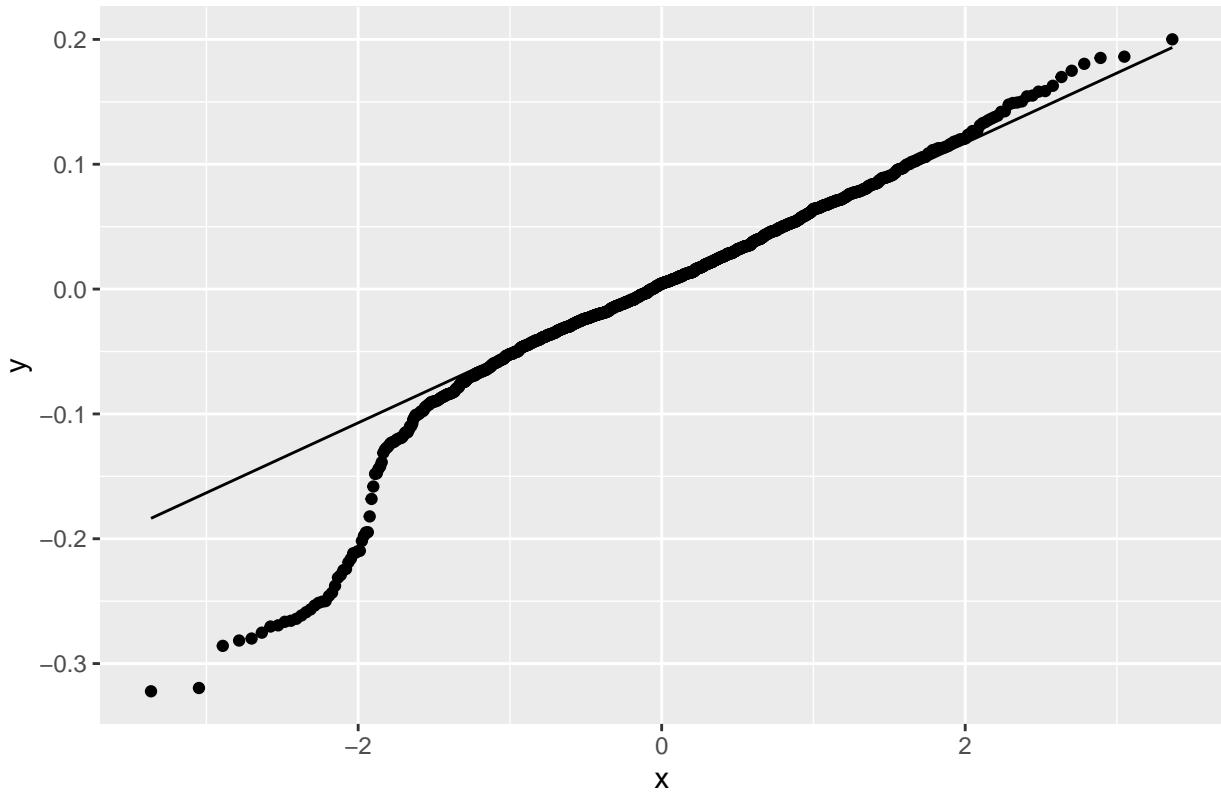


```
## Full model is singular. Consider simplifying the random effects structure.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: -3098.4
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -4.5885 -0.4688  0.0599  0.6087  2.8450 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 2.051e-03 0.0452899
##          AgeInDays   1.055e-03 0.0324821  0.82
## vowels   (Intercept) 2.264e-04 0.0150461
##          AgeInDays   5.676e-07 0.0007534 -1.00
## Residual            4.939e-03 0.0702753
## Number of obs: 1306, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept)  4.018744  0.013528 21.675539 297.075 <2e-16 ***
## SES         0.012752  0.008209 13.768394   1.553    0.143  

```

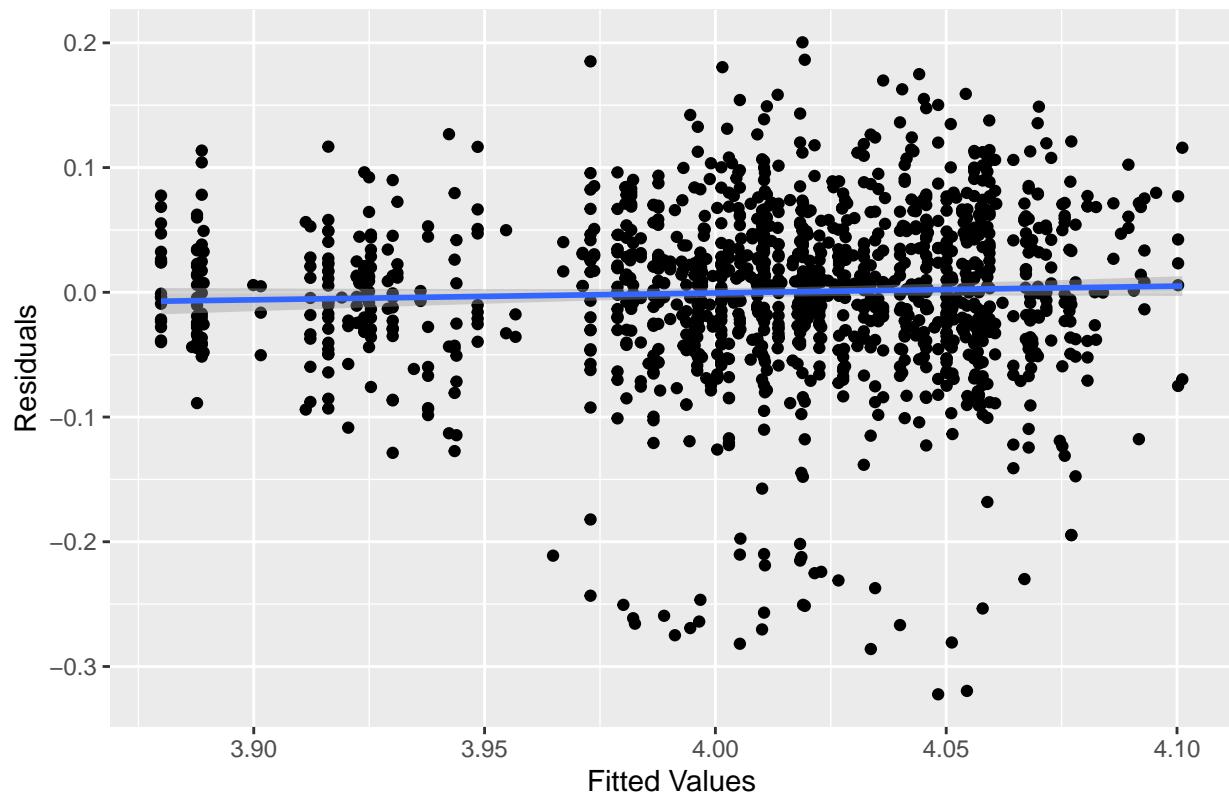
```

## SEXM          -0.009758  0.023517 19.157512 -0.415      0.683
## AgeInDays    -0.003723  0.010091 13.323403 -0.369      0.718
## SEXM:AgeInDays -0.003620  0.017394 12.379678 -0.208      0.839
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM   AgInDy
## SES        -0.126
## SEXM       -0.518  0.302
## AgeInDays   0.633 -0.087 -0.391
## SEXM:AgInDy -0.369  0.018  0.719 -0.577
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = 1567.36 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 3.03358639732413
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model     8 -3118.2 -3076.8 1567.1   -3134.2
## fit_model         12 -3110.7 -3048.6 1567.4   -3134.7 0.5091  4      0.9726
## Fitting Reduced Model.

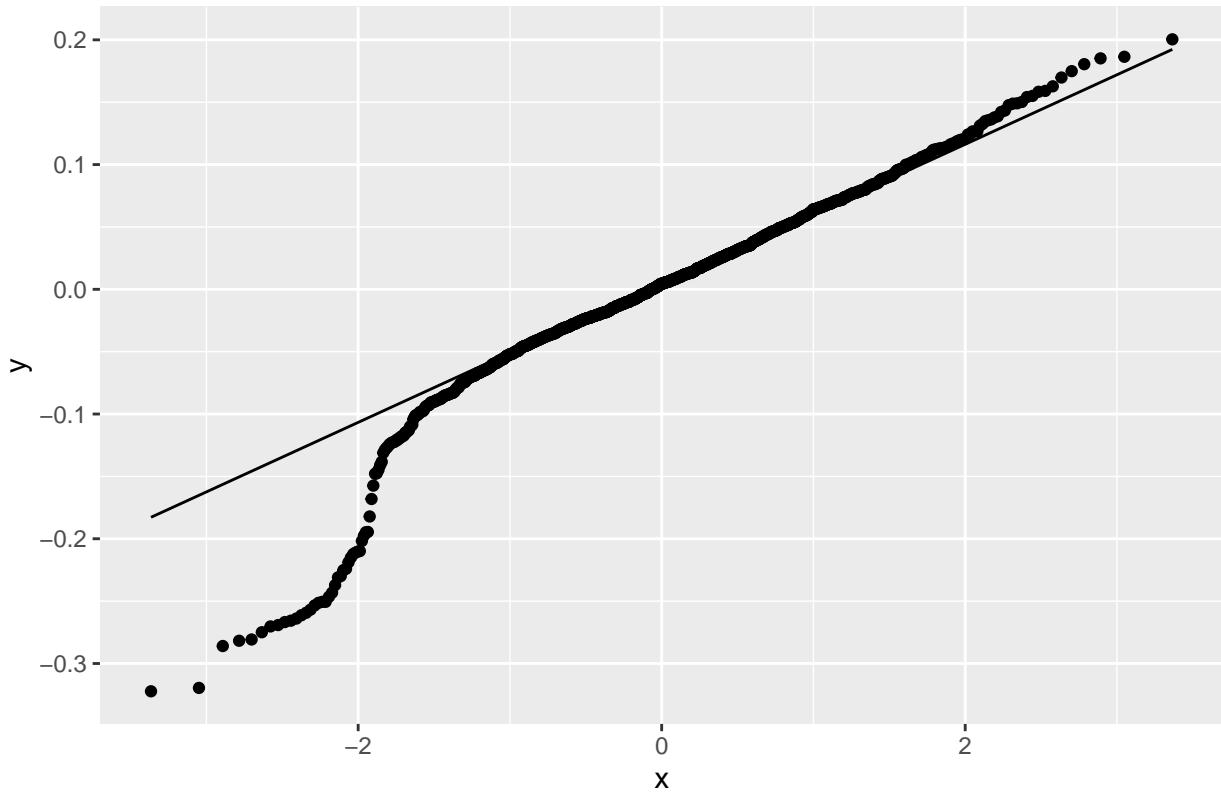
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: -3098.3
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -4.5899 -0.4700  0.0590  0.6088  2.8511 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.0020466 0.04524
##          AgeInDays   0.0010498 0.03240  0.82
## vowels   (Intercept) 0.0002261 0.01504
## Residual           0.0049392 0.07028
## Number of obs: 1306, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  4.018707  0.013514 21.700036 297.375 <2e-16 ***
## SES          0.012760  0.008211 13.791623   1.554    0.143    
## SEXM         -0.009837  0.023495 19.183404  -0.419    0.680    
## 
```

```

## AgeInDays      -0.003379   0.010069 13.350996  -0.336     0.742
## SEXM:AgeInDays -0.003798   0.017362 12.408191  -0.219     0.830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES    SEXM   AgInDy
## SES       -0.126
## SEXM      -0.518  0.303
## AgeInDays  0.643 -0.087 -0.390
## SEXM:AgInDy -0.369  0.018  0.718 -0.577
## Reduced Model: Log-Likelihood = 1567.33 , Degrees of Freedom = 10
## Max VIF value of the Reduced Model: 3.03255702176095
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##           npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model     8 -3118.2 -3076.8 1567.1   -3134.2
## red_fit_model     10 -3114.7 -3062.9 1567.3   -3134.7 0.4526  2     0.7975
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).

pred.data.2.1 <- boot.ci.predict.lmer(
  m = output.2.1$reduced_model,
  optimizer = output.2.1$optimizer,
  maxfun = output.2.1$maxfun,
  data = output.2.1$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.2.1$fit_data$AgeInDays), max(output.2.1$fit_data$AgeInDays),
  reqcol = c("AgeInDays"),
  centercol = c("SES", "SEX"),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## Number of singular fits during bootstrapping: 0 out of 1000 bootstrap iterations.

#2.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_pitch_st.xlsx"))

## New names:
## * ` ` -> `...`1` 

output.2.2 <- lmer.full.reduced.null.compare(meanF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id),
                                               optimizer = "bobyqa",
                                               maxfun = 10000000000,
                                               fullm = TRUE,
                                               redm = TRUE,
                                               redmformula = meanF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id),
                                               nullm = TRUE,
                                               nullmformula = meanF0st ~ SES + SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowel),
                                               df_IDS)

## Structure of combined data:

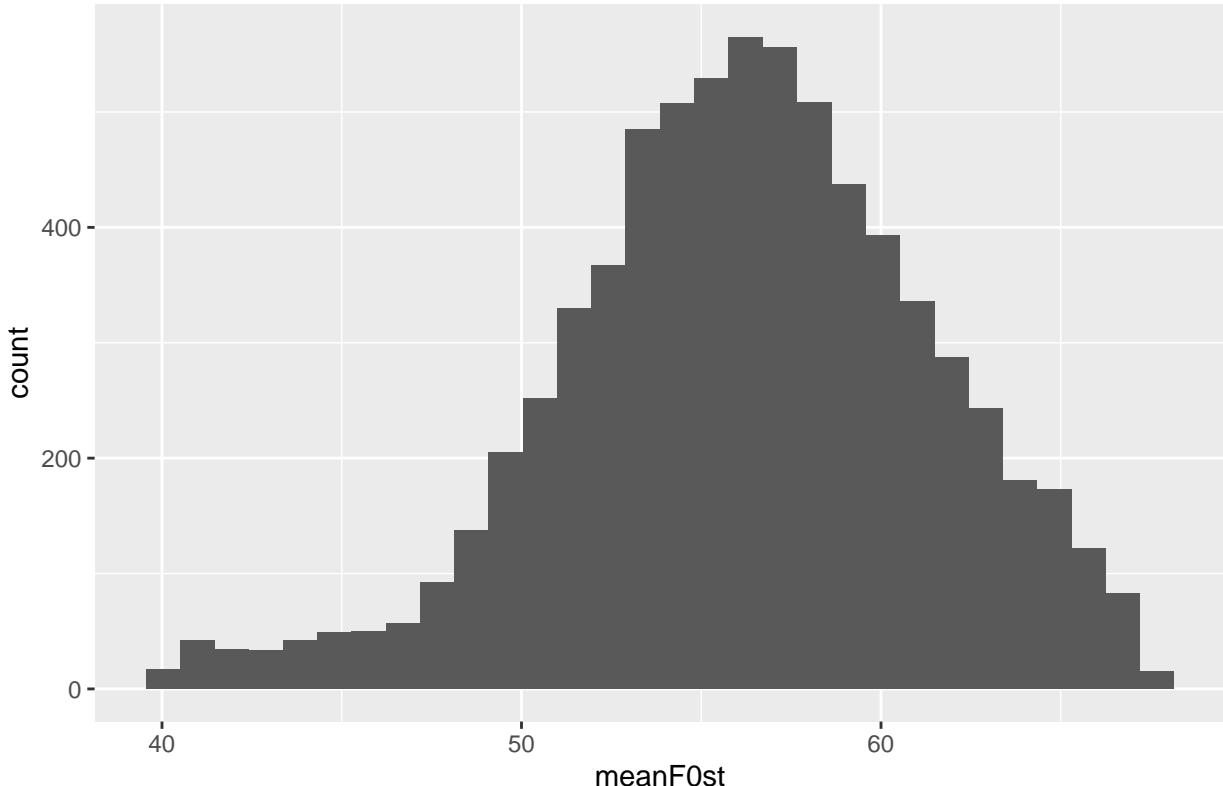
```

```

## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:7127] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES : num [1:7127] 5 5 5 5 5 5 5 5 5 5 ...
## $ SEX : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:7127] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels : chr [1:7127] "a" "eu" "a" "a" ...
## $ meanF0st : num [1:7127] 62.3 65.5 57.4 52.9 58.5 ...
## Response variable is: meanF0st
## Structure of combined data after scaling:
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3563
## ..- attr(*, "scaled:scale")= num 2058
## $ SPK_id : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.38
## $ SEX : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 244
## ..- attr(*, "scaled:scale")= num 100
## $ vowels : chr [1:7127] "a" "eu" "a" "a" ...
## $ meanF0st : num [1:7127] 62.3 65.5 57.4 52.9 58.5 ...

```

Histogram of Response Variable



```

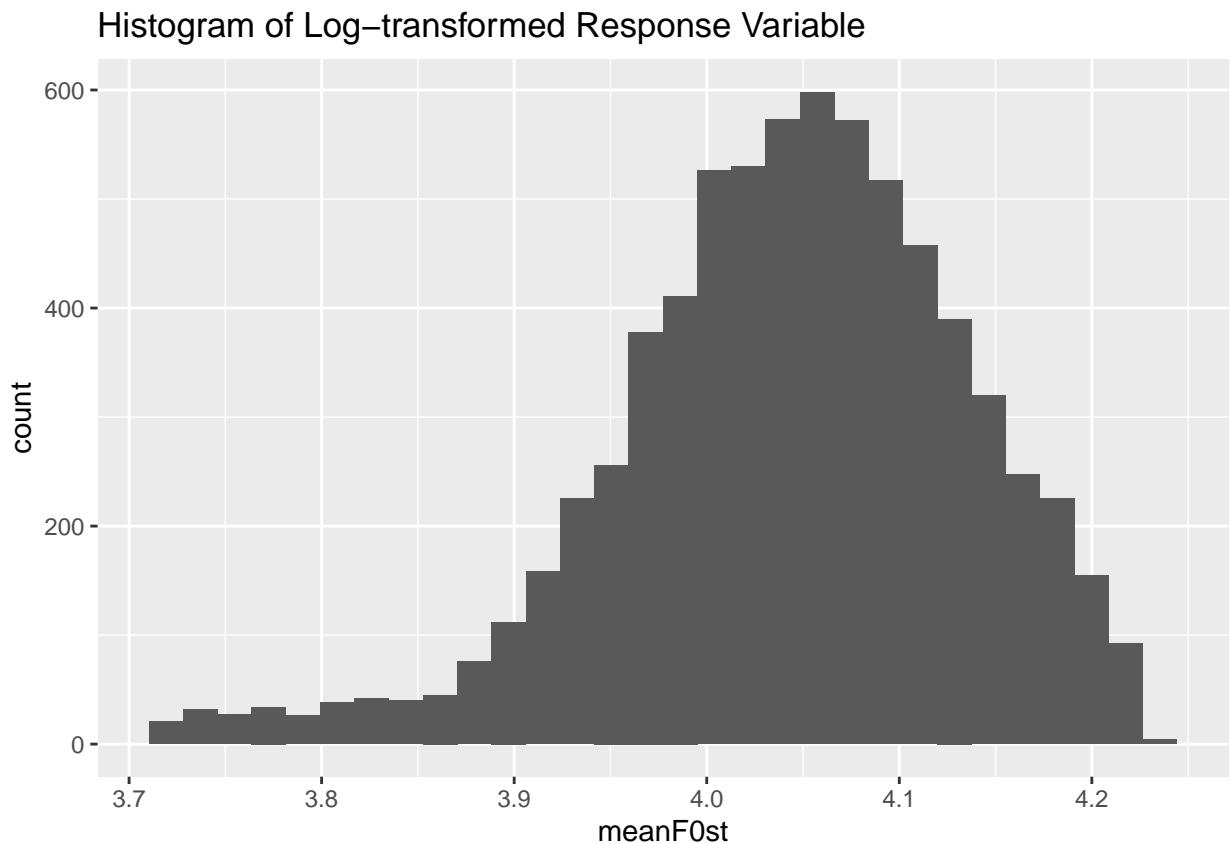
## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:

```

```

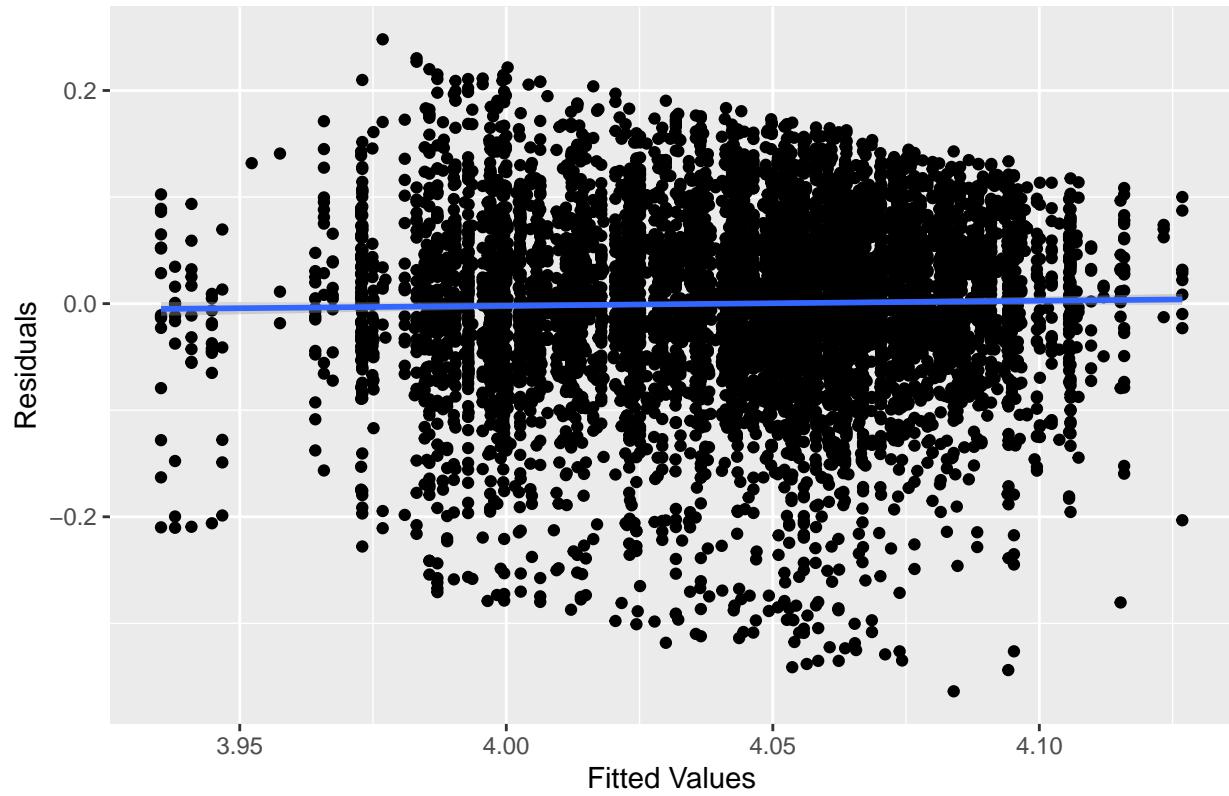
### tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
### $ ...1 : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...
### ..- attr(*, "scaled:center")= num 3563
### ..- attr(*, "scaled:scale")= num 2058
### $ SPK_id : chr [1:7127] "C083" "C083" "C083" "C083" ...
### $ SES     : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 0 ...
### ..- attr(*, "scaled:center")= num 5
### ..- attr(*, "scaled:scale")= num 1.38
### $ SEX     : chr [1:7127] "F" "F" "F" "F" ...
### $ AgeInDays: num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
### ..- attr(*, "scaled:center")= num 244
### ..- attr(*, "scaled:scale")= num 100
### $ vowels   : chr [1:7127] "a" "eu" "a" "a" ...
### $ meanF0st : num [1:7127] 4.15 4.2 4.07 3.99 4.09 ...

```

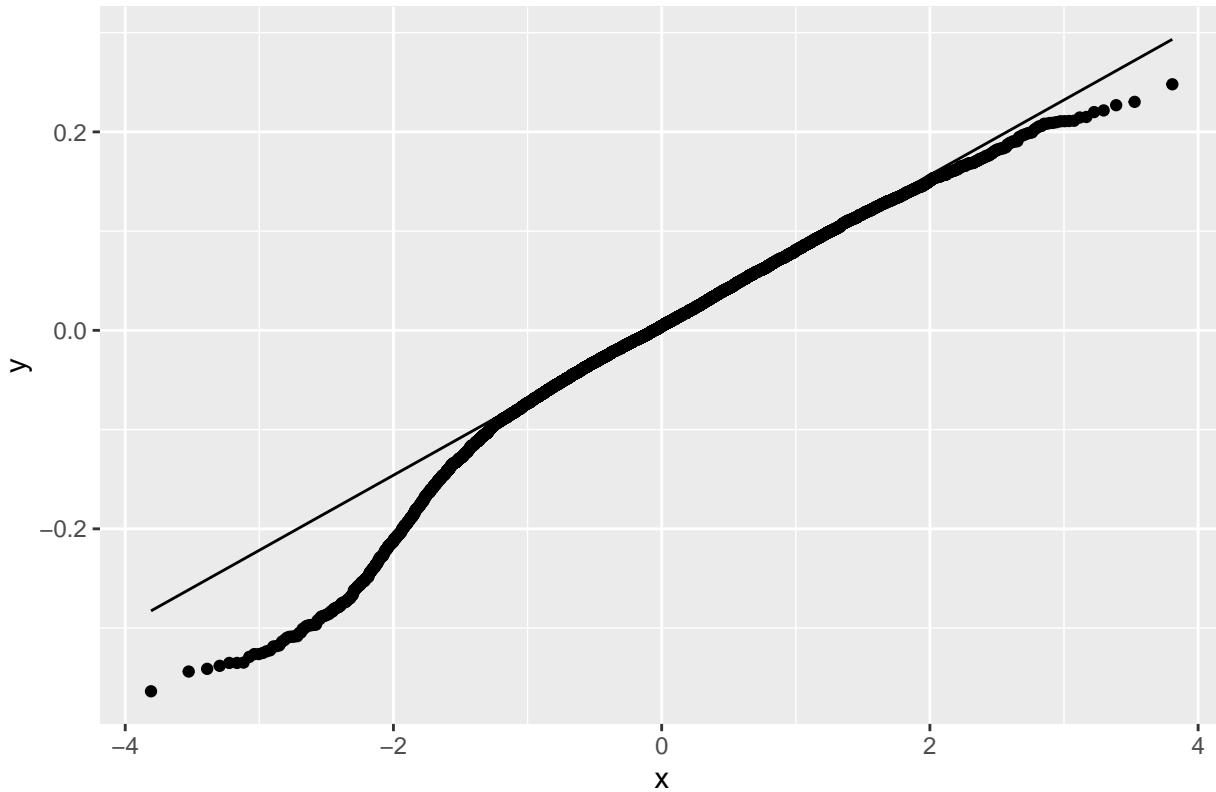


```
## Fitting Full Model.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: -14727.5
##
## Scaled residuals:
##    Min     1Q  Median     3Q    Max 
## -4.2868 -0.5383  0.0532  0.6637  2.9310 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 1.148e-03 0.033881
##          AgeInDays   4.472e-04 0.021147 0.22
## vowels   (Intercept) 1.933e-04 0.013902
##          AgeInDays   6.823e-06 0.002612 0.42
## Residual            7.201e-03 0.084856
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept)  4.048065  0.009886 30.649767 409.464 <2e-16 ***
## SES         0.011106  0.007830 20.734143   1.419    0.171  

```

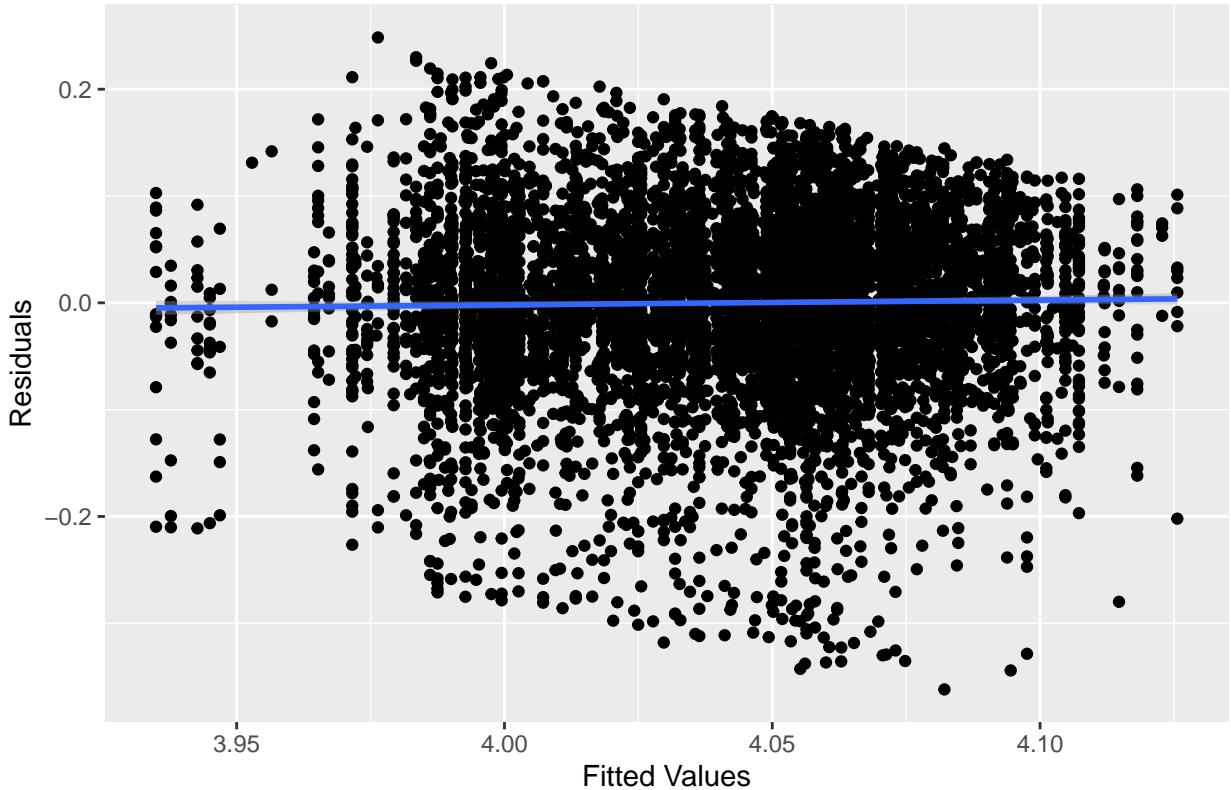
```

## SEXM          -0.010480  0.015693 23.970986 -0.668    0.511
## AgeInDays     0.002270  0.005990 19.557671  0.379    0.709
## SEXM:AgeInDays -0.002748  0.010199 19.367559 -0.269    0.790
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES      SEXM     AgInDy
## SES        -0.259
## SEXM       -0.545  0.381
## AgeInDays   0.170  0.026 -0.083
## SEXM:AgInDy -0.082 -0.018  0.193 -0.572
## Full Model: Log-Likelihood = 7383.14 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.54982556963116
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model 10 -14746 -14677 7383.1    -14766
## fit_model       12 -14742 -14660 7383.1    -14766 0.1693 2      0.9188
## Fitting Reduced Model.

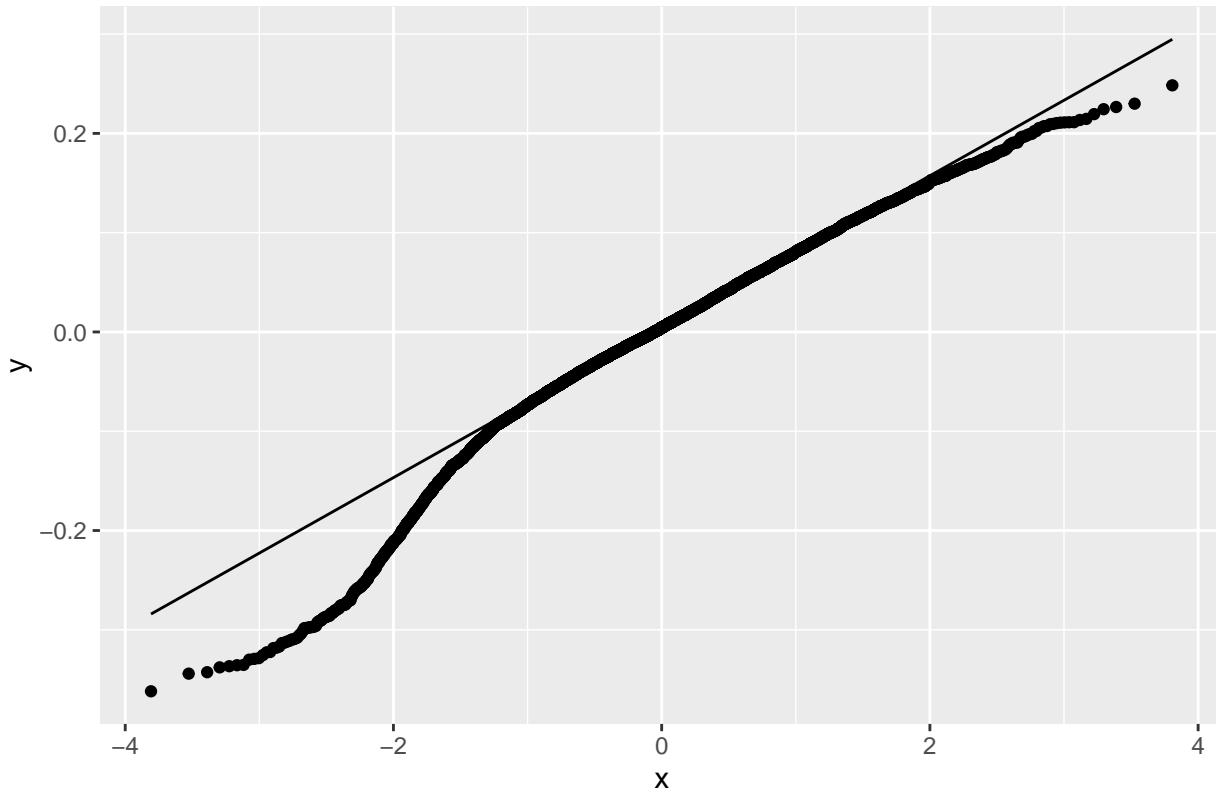
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: -14725.5
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -4.2641 -0.5420  0.0527  0.6658  2.9351
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.0011517 0.03394
##          AgeInDays   0.0004472 0.02115  0.22
## vowels   (Intercept) 0.0001910 0.01382
## Residual           0.0072061 0.08489
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  4.048047  0.009886 30.621486 409.483 <2e-16 ***
## SES          0.011129  0.007847 20.756645   1.418   0.171
## SEXM         -0.010437  0.015717 23.958720  -0.664   0.513

```



```

## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 50 out of 1000 bootstrap iterations.

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.2.1$fit_data$AgeInDays, "scaled:scale") + attr(output.2.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.2.2$fit_data$AgeInDays, "scaled:scale") + attr(output.2.2$fit_data$AgeInDays, "scaled")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

Pitch_semitones <- continuous_fit_ci_plot(
  plot.data = list(output.2.1$fit_data, output.2.2$fit_data),
  coefs = list(pred.data.2.1,pred.data.2.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "meanF0st",
  x.labs = "Age in Days",
  y.labs = "Pitch(semitones)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = TRUE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  legend.labels = c("ADS", "IDS")
)

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.

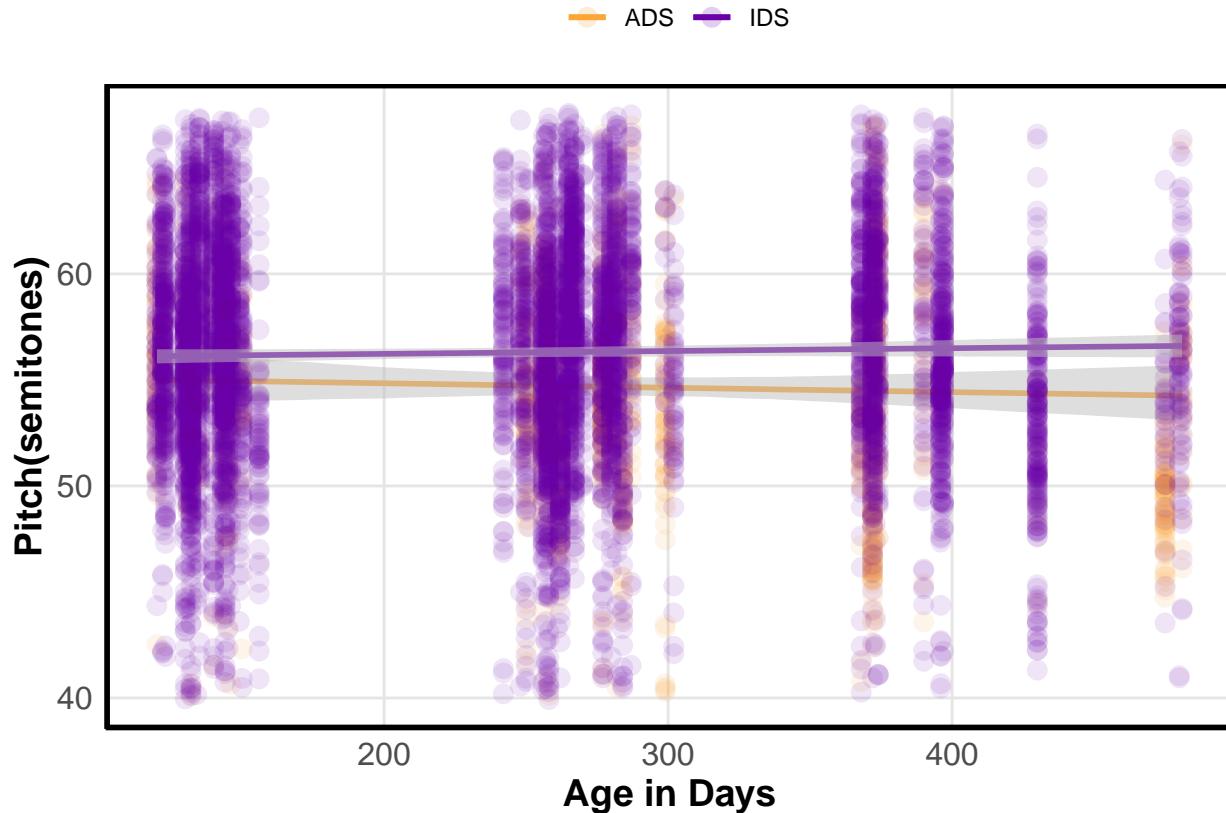
```

```

## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

print(Pitch_semitones)

```



```

#3.Range #3.1 ADS
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_range_st.xlsx"))

## New names:
## * `` -> `...` 

output.3.1 <- lmer.full.reduced.null.compare(RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+Age
optimizer = "bobyqa",
maxfun = 1000000000,
fullm = TRUE,
redm = TRUE,
redmformula = RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id) +
nullm = TRUE,
nullmformula = RangeF0st ~ SES + SEX + (1|SPK_id) + (1|vowels),
df_ADS)

## Structure of combined data:
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:1306] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:1306] 5 5 5 5 5 5 5 5 5 5 ...
## $ SEX       : chr [1:1306] "F" "F" "F" "F" ...

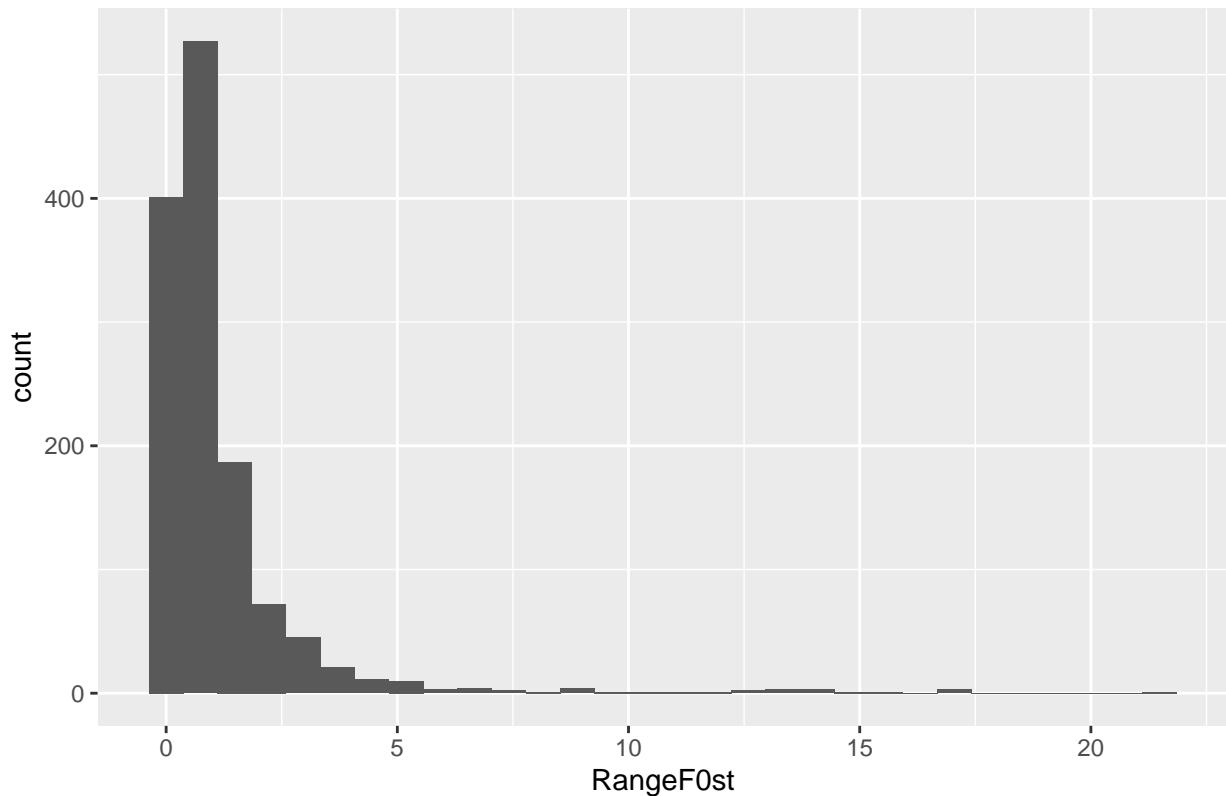
```

```

## $ AgeInDays: num [1:1306] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...
## $ RangeF0st: num [1:1306] 0.511 0.114 0.787 0.872 0.697 ...
## Response variable is: RangeF0st
## Structure of combined data after scaling:
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 -1.72 ...
## ..- attr(*, "scaled:center")= num 652
## ..- attr(*, "scaled:scale")= num 377
## $ SPK_id : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...
## ..- attr(*, "scaled:center")= num 295
## ..- attr(*, "scaled:scale")= num 97.1
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...
## $ RangeF0st: num [1:1306] 0.511 0.114 0.787 0.872 0.697 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 -1.72 ...
## ..- attr(*, "scaled:center")= num 652
## ..- attr(*, "scaled:scale")= num 377

```

```

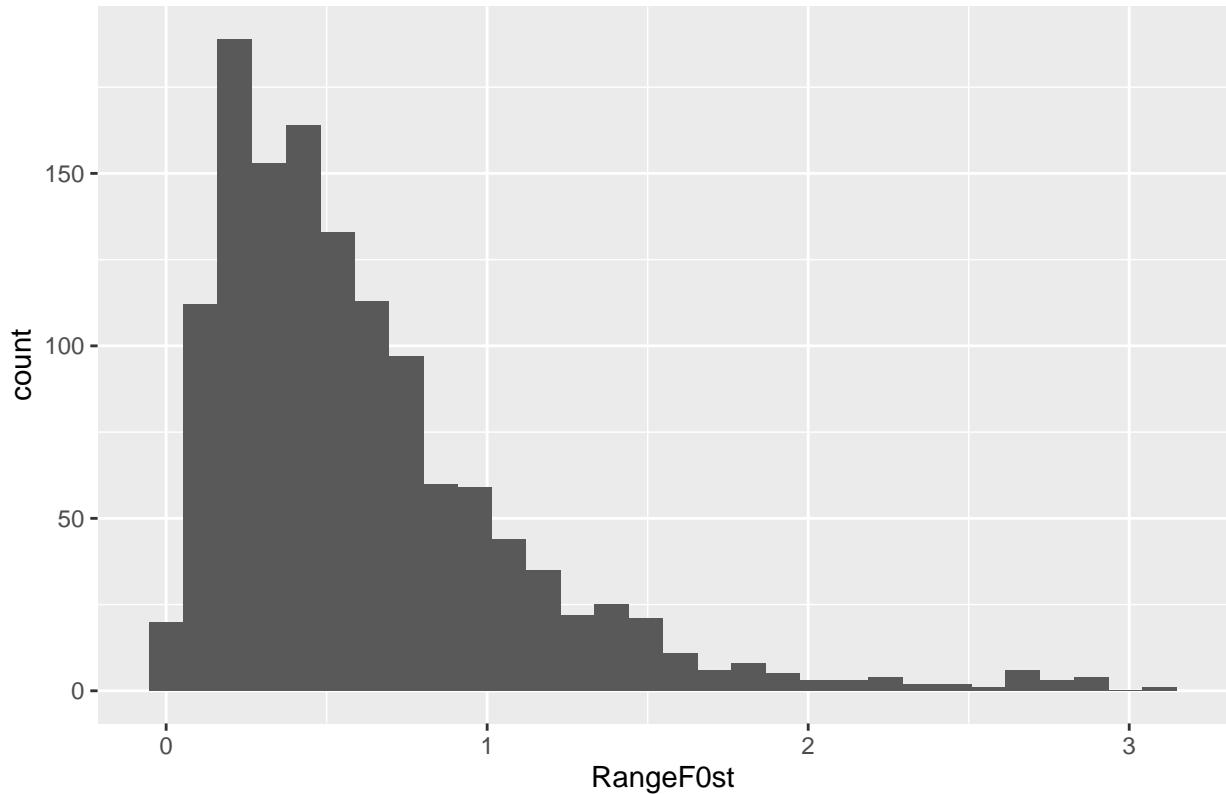
## $ SPK_id    : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX       : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...
## ..- attr(*, "scaled:center")= num 295
## ..- attr(*, "scaled:scale")= num 97.1
## $ vowels   : chr [1:1306] "o" "a" "i" "a" ...
## $ RangeF0st: num [1:1306] 0.413 0.108 0.58 0.627 0.529 ...

## Fitting Full Model.

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

```

Histogram of Log-transformed Response Variable

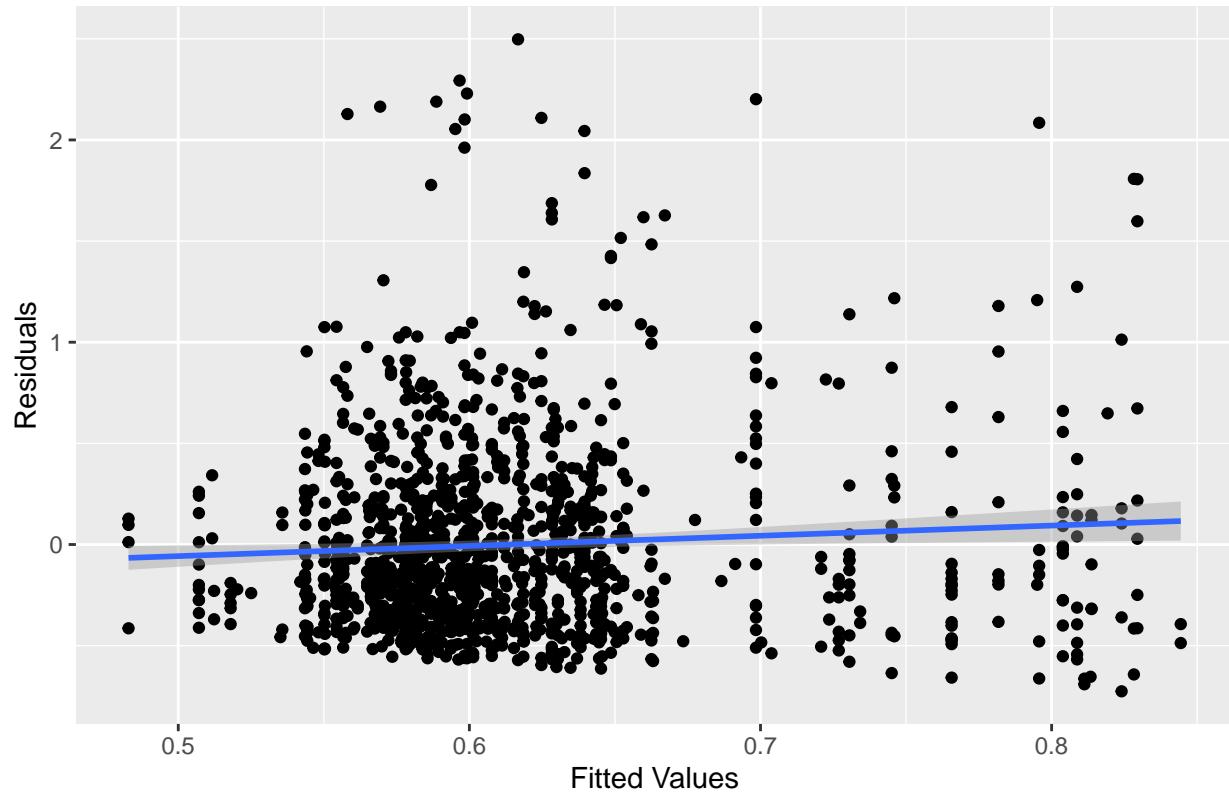


```

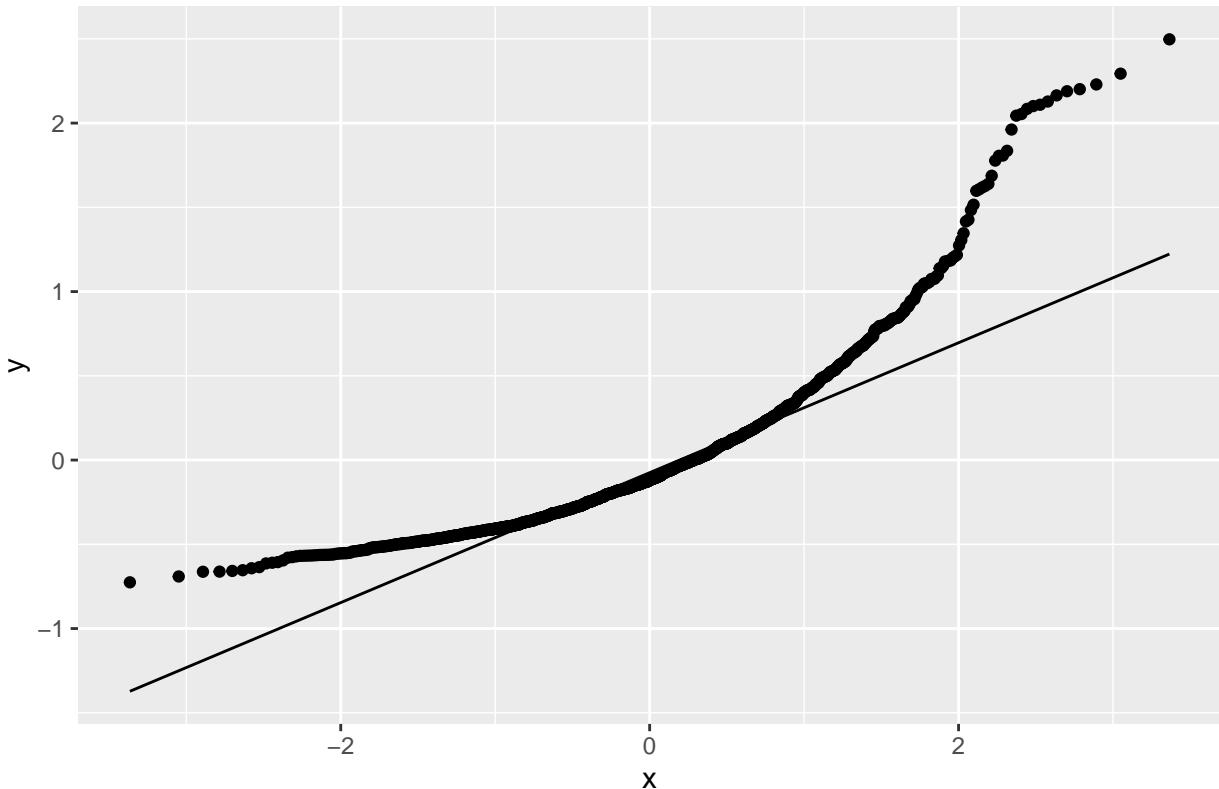
## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



Q-Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1815.1
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.5424 -0.6949 -0.2523  0.3840  5.2242 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 3.724e-03 0.061026 
##          AgeInDays   9.521e-04 0.030856 -1.00
## vowels   (Intercept) 7.902e-04 0.028110 
##          AgeInDays   4.983e-05 0.007059 -1.00
## Residual            2.275e-01 0.476954 
## Number of obs: 1306, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  0.617453  0.026804 15.940828 23.036 1.16e-13 ***
## SES         -0.003358  0.019089 18.962244 -0.176    0.862  

```

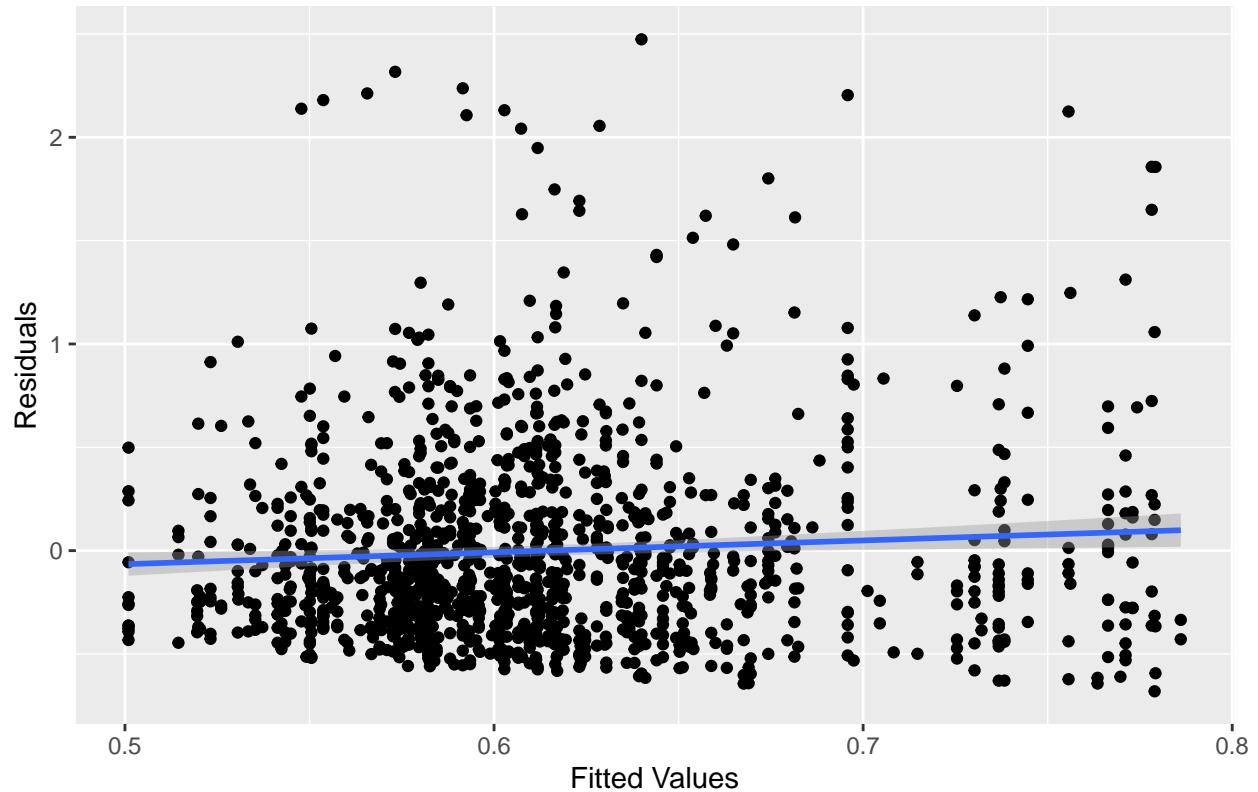
```

## SEXM          -0.032984  0.044759 17.548543 -0.737    0.471
## AgeInDays    -0.023562  0.024215 27.463068 -0.973    0.339
## SEXM:AgeInDays 0.022132  0.033110 24.713233  0.668    0.510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM   AgInDy
## SES        -0.255
## SEXM       -0.563  0.404
## AgeInDays  -0.187 -0.196  0.031
## SEXM:AgInDy 0.114  0.098 -0.295 -0.711
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -893.38 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 2.36279862992802
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model  6 1805.0 1836.0 -896.48    1793.0
## fit_model      12 1810.8 1872.9 -893.38    1786.8 6.1954  6     0.4017
## Fitting Reduced Model.

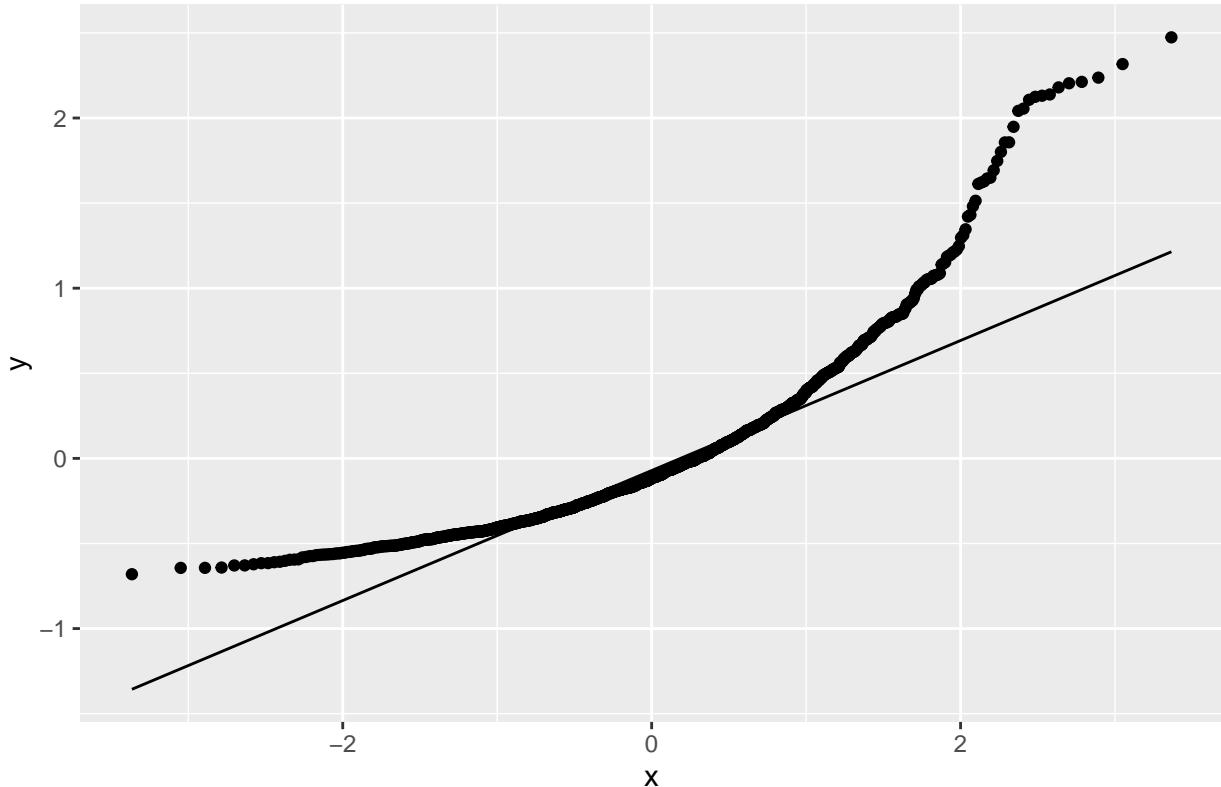
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1817.8
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.4523 -0.6913 -0.2472  0.3915  5.1631 
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.0044492 0.06670
## vowels   (Intercept) 0.0008225 0.02868
## Residual           0.2281111 0.47761
## Number of obs: 1306, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept) 6.187e-01 2.800e-02 1.686e+01 22.094 6.94e-14 ***
## SES         4.183e-04 2.154e-02 2.003e+01  0.019   0.985    
## SEXM        -3.366e-02 4.798e-02 1.678e+01 -0.701   0.493    
## AgeInDays   -3.529e-02 2.096e-02 4.204e+02 -1.684   0.093 .  
## 
```

```

## SEXM:AgeInDays  3.908e-02  3.023e-02  3.337e+02    1.293    0.197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.238
## SEXM     -0.557  0.413
## AgeInDays  0.138 -0.170 -0.131
## SEXM:AgInDy -0.089  0.082  0.062 -0.687
## Reduced Model: Log-Likelihood = -894.93 , Degrees of Freedom = 8
## Max VIF value of the Reduced Model: 1.96914020088207
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##          npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model  6 1805.0 1836.0 -896.48   1793.0
## red_fit_model   8 1805.8 1847.2 -894.93   1789.8 3.1114  2      0.211
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_range_st.xlsx"))

## New names:
## * `` -> `...` 

# Calculate the quartiles
Q1 <- quantile(df_ADS$RangeF0st, 0.25)
Q3 <- quantile(df_ADS$RangeF0st, 0.75)
# Calculate the interquartile range (IQR)
IQR <- Q3 - Q1
# Define the lower and upper bounds for non-outliers
lower <- Q1 - 1.5 * IQR
upper <- Q3 + 1.5 * IQR

# Remove outliers and create a new dataframe
df_no_outliers_ADS <- df_ADS[df_ADS$RangeF0st >= lower & df_ADS$RangeF0st <= upper, ]

output.3.1_no_outlier <- lmer.full.reduced.null.compare(RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX +
  optimizer = "bobyqa",
  maxfun = 1000000000,
  fullm = TRUE,
  redm = TRUE,
  redmformula = RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id) +
  nullm = TRUE,
  nullmformula = RangeF0st ~ SES + SEX + (1+AgeInDays|SPK_id) + (1|SPK_id) + (1|v
  df_no_outliers_ADS)

## Structure of combined data:
## tibble [1,187 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:1187] 0 1 2 3 4 6 7 8 9 10 ...
## $ SPK_id    : chr [1:1187] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:1187] 5 5 5 5 5 5 5 5 5 5 ...

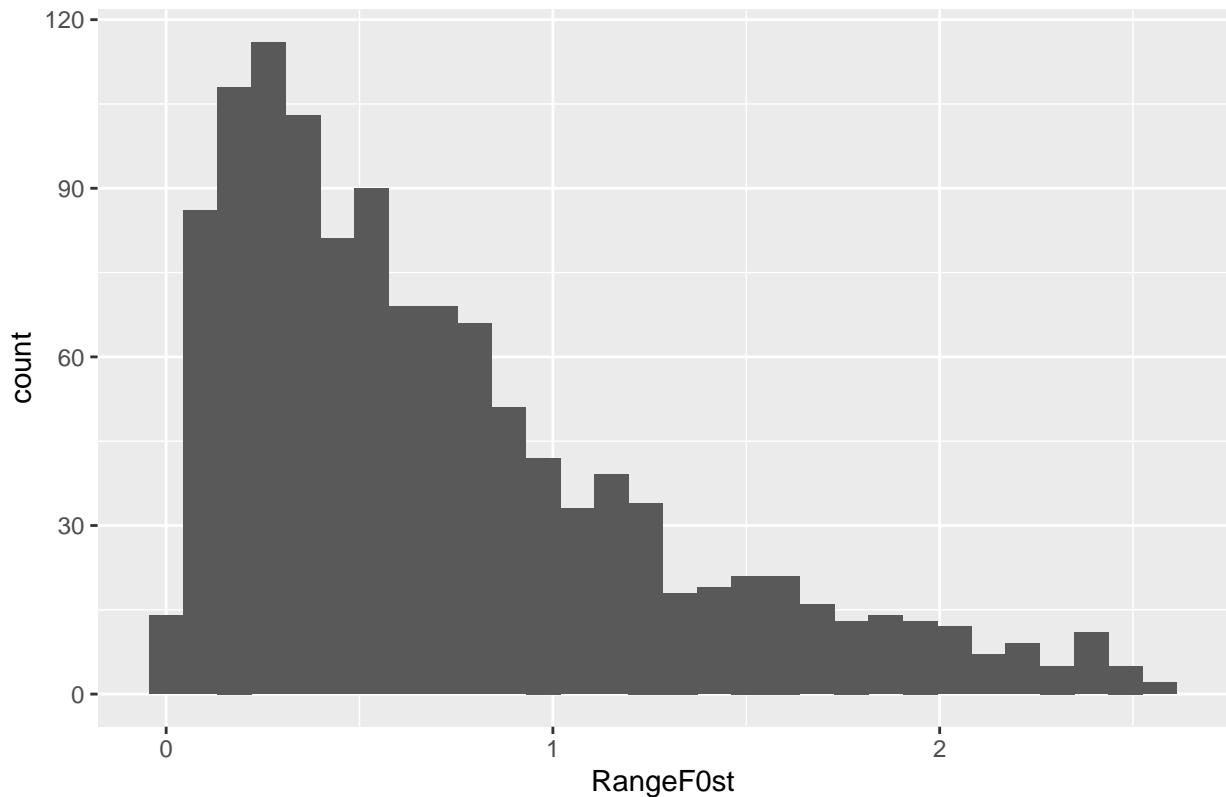
```

```

## $ SEX      : chr [1:1187] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1187] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels   : chr [1:1187] "o" "a" "i" "a" ...
## $ RangeF0st: num [1:1187] 0.511 0.114 0.787 0.872 0.697 ...
## Response variable is: RangeF0st
## Structure of combined data after scaling:
## tibble [1,187 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:1187, 1] -1.76 -1.76 -1.75 -1.75 -1.75 ...
## ..- attr(*, "scaled:center")= num 660
## ..- attr(*, "scaled:scale")= num 375
## $ SPK_id   : chr [1:1187] "C083" "C083" "C083" "C083" ...
## $ SES      : num [1:1187, 1] -0.101 -0.101 -0.101 -0.101 -0.101 ...
## ..- attr(*, "scaled:center")= num 5.13
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX      : chr [1:1187] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:1187, 1] -1.82 -1.82 -1.82 -1.82 -1.82 ...
## ..- attr(*, "scaled:center")= num 296
## ..- attr(*, "scaled:scale")= num 96.7
## $ vowels   : chr [1:1187] "o" "a" "i" "a" ...
## $ RangeF0st: num [1:1187] 0.511 0.114 0.787 0.872 0.697 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [1,187 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:1187, 1] -1.76 -1.76 -1.75 -1.75 -1.75 ...
## ..- attr(*, "scaled:center")= num 660

```

```

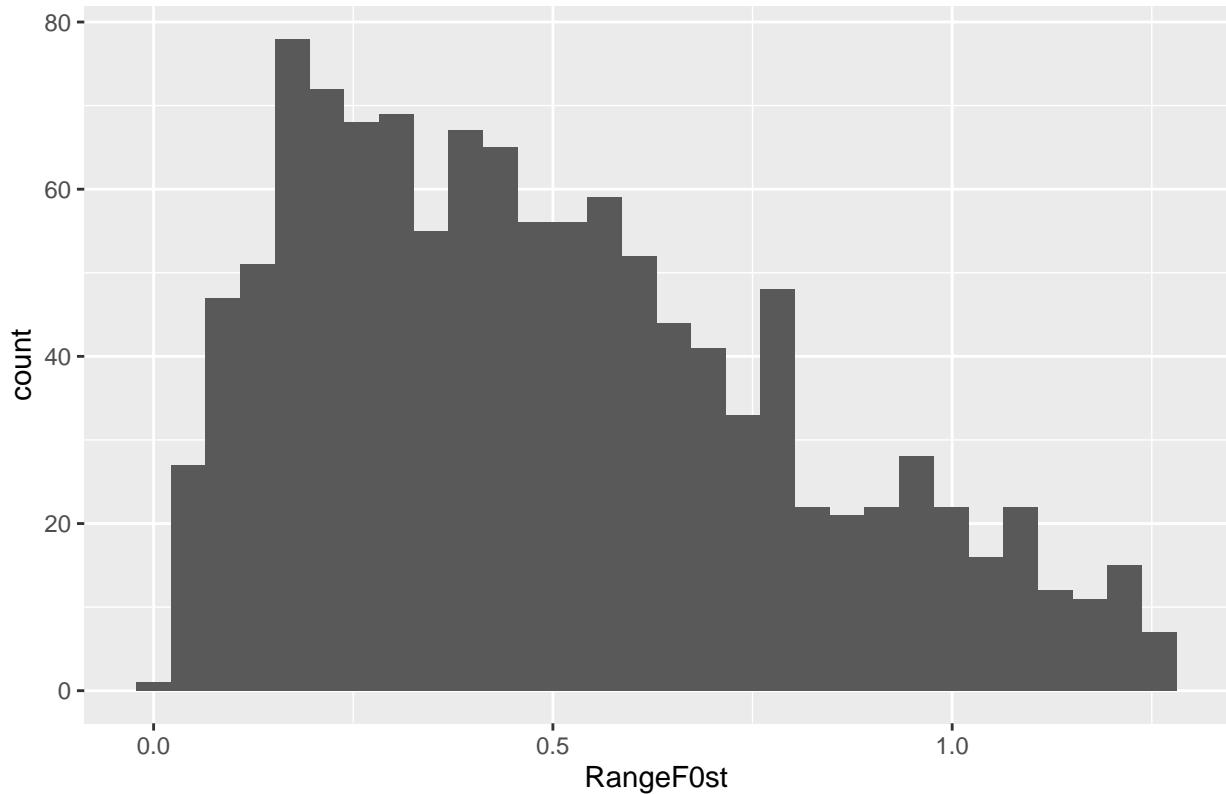
##  ..- attr(*, "scaled:center")= num 5.13
##  ..- attr(*, "scaled:scale")= num 1.3
##  $ SEX      : chr [1:1187] "F" "F" "F" "F" ...
##  $ AgeInDays: num [1:1187, 1] -1.82 -1.82 -1.82 -1.82 -1.82 ...
##  ..- attr(*, "scaled:center")= num 296
##  ..- attr(*, "scaled:scale")= num 96.7
##  $ vowels   : chr [1:1187] "o" "a" "i" "a" ...
##  $ RangeF0st: num [1:1187] 0.413 0.108 0.58 0.627 0.529 ...

## Fitting Full Model.

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

```

Histogram of Log-transformed Response Variable

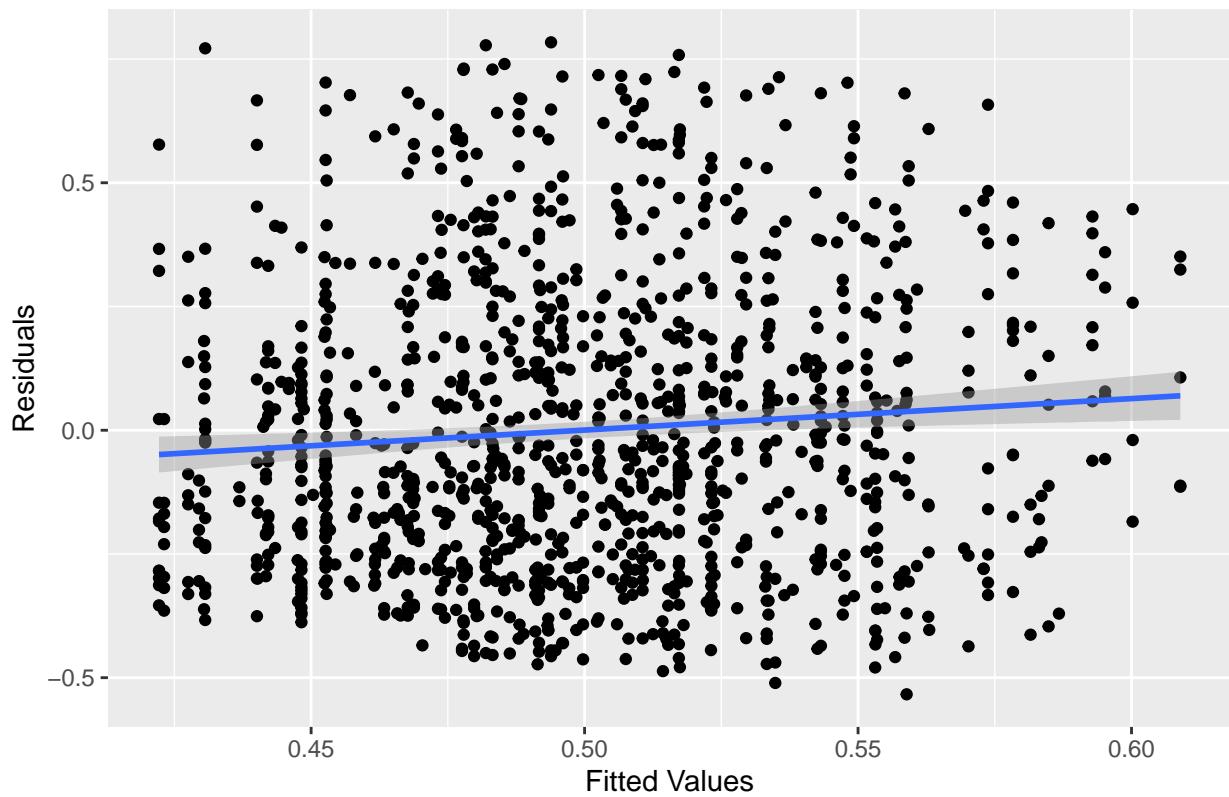


```

## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 526.1
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -1.8278 -0.8152 -0.1726  0.6342  2.6533
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 1.947e-03 0.044126
##          AgeInDays   1.541e-05 0.003926 1.00
## vowels   (Intercept) 1.228e-03 0.035049
##          AgeInDays   7.939e-06 0.002818 1.00
## Residual            8.740e-02 0.295633
## Number of obs: 1187, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 0.505723  0.021166 15.425276 23.893 1.32e-13 ***
## SES         0.003868  0.014169 21.069990   0.273    0.788

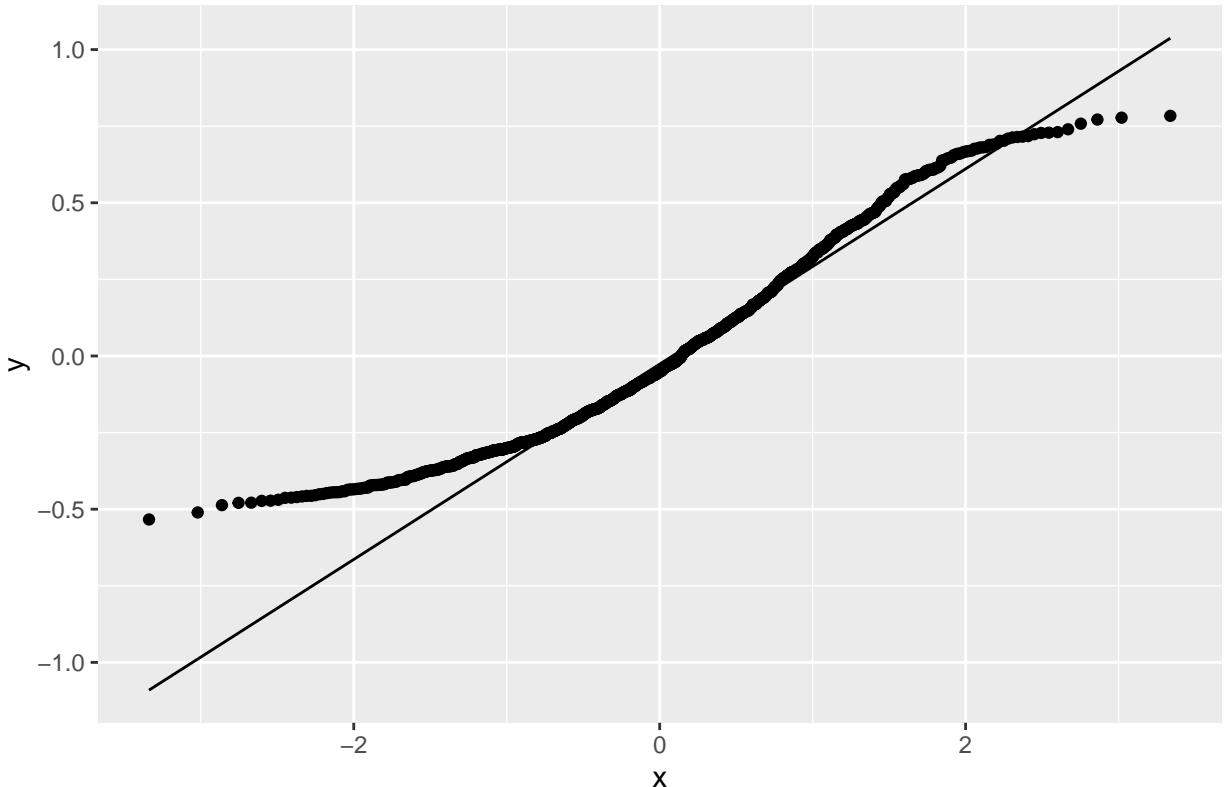
```

```

##  SEXM          -0.003356   0.031560  16.580699  -0.106   0.917
##  AgeInDays     -0.011812   0.013704 340.734062  -0.862   0.389
##  SEXM:AgeInDays 0.016621   0.019876  81.866245   0.836   0.405
##  ---
##  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES      SEXM    AgInDy
##  SES      -0.201
##  SEXM     -0.481   0.408
##  AgeInDays  0.215  -0.167 -0.167
##  SEXM:AgInDy -0.113   0.084  0.153 -0.681
##  optimizer (bobyqa) convergence code: 0 (OK)
##  boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -247.18 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.95580740223561
## Fitting Null Model.
##
## boundary (singular) fit: see help('isSingular')

```

Q–Q Plot of Residuals for Full Model



```

## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)

```

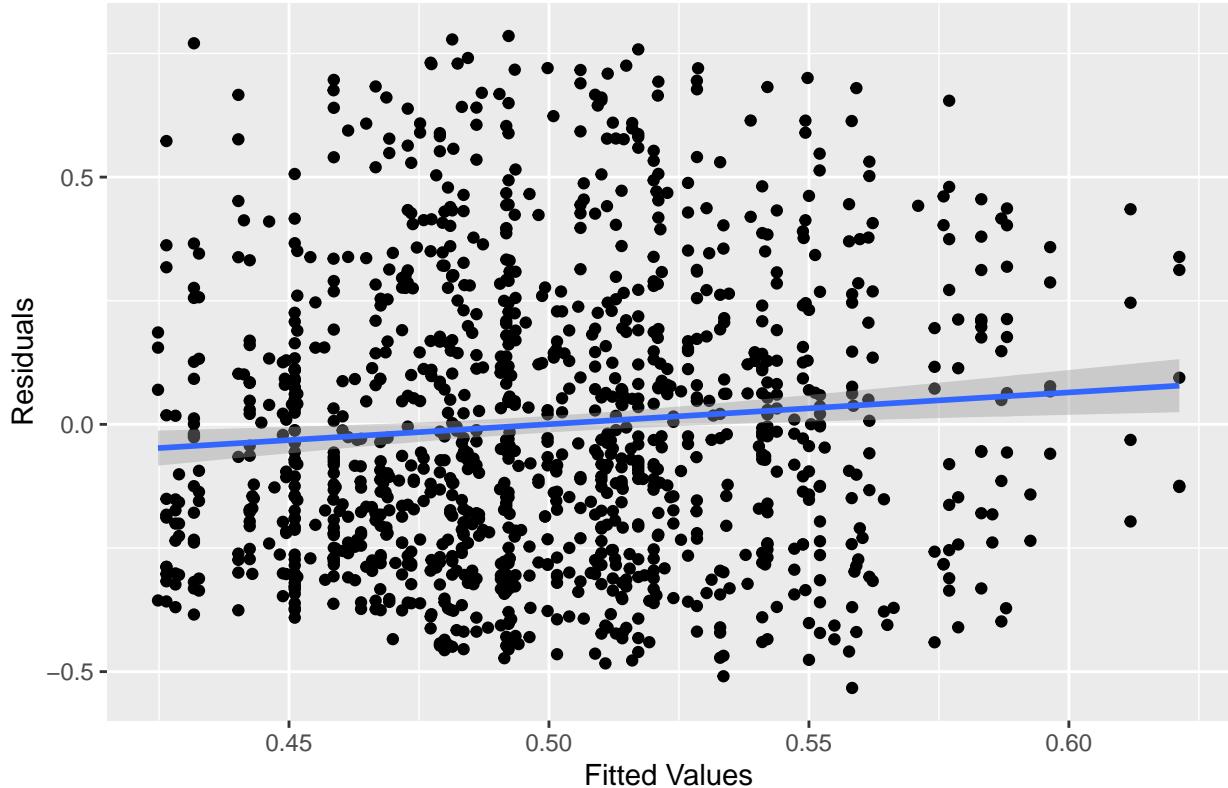
```

## null_fit_model      9 513.31 559.03 -247.66    495.31
## fit_model          12 518.35 579.30 -247.18    494.35 0.9618  3     0.8105
## Fitting Reduced Model.

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

```

Residuals vs Fitted Values for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 526.3
##
## Scaled residuals:
##    Min     1Q   Median     3Q    Max
## -1.8254 -0.8160 -0.1822  0.6387  2.6572
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.001909 0.04370
## vowels   (Intercept) 0.001248 0.03532
## Residual            0.087410 0.29565
## Number of obs: 1187, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:

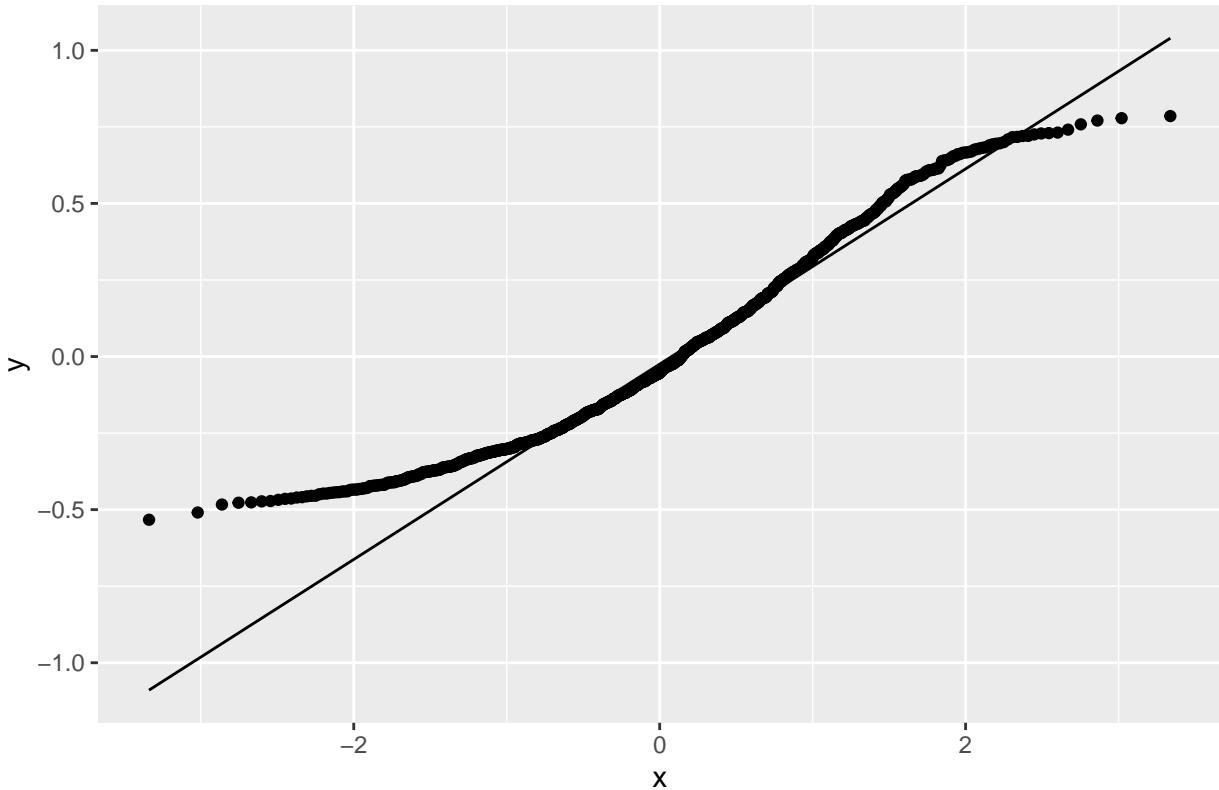
```

```

##                               Estimate Std. Error      df t value Pr(>|t|) 
## (Intercept)            0.505199  0.021138  15.524455 23.900 1.15e-13 ***
## SES                  0.002876  0.014033  20.939452  0.205   0.840
## SEXM                 -0.003992  0.031214  17.499827 -0.128   0.900
## AgeInDays             -0.010969  0.013739 410.358681 -0.798   0.425
## SEXM:AgeInDays       0.014298  0.019579 344.770658  0.730   0.466
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Correlation of Fixed Effects:
##          (Intr) SES    SEXM   AgInDy
## SES     -0.206
## SEXM    -0.481  0.412
## AgeInDays 0.116 -0.169 -0.132
## SEXM:AgInDy -0.076  0.086  0.070 -0.696
## Reduced Model: Log-Likelihood = -247.24 , Degrees of Freedom = 8
## Max VIF value of the Reduced Model: 2.00640790262258
## Fitting Null Model.
## boundary (singular) fit: see help('isSingular')

```

Q–Q Plot of Residuals for Reduced Model



```

## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## red_fit_model: redmformula
## null_fit_model: nullmformula
##           npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)

```



```

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

## singular fit
## Number of singular fits during bootstrapping: 68 out of 1000 bootstrap iterations.

pred.data.3.1_no_outlier <- boot.ci.predict.lmer(
  m = output.3.1_no_outlier$reduced_model,
  optimizer = output.3.1_no_outlier$optimizer,
  maxfun = output.3.1_no_outlier$maxfun,
  data = output.3.1_no_outlier$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.3.1_no_outlier$fit_data$AgeInDays), max(output.3.1_no_outlier$fit_data$AgeInDays)),
  reqcol = c("AgeInDays"),
  centercol = c("SES", 'SEX'),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 5 out of 1000 bootstrap iterations.

#3.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_range_st.xlsx"))

## New names:
## * ` ` -> `...`1` 

output.3.2 <- lmer.full.reduced.null.compare(RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays:SEX|SPK_id),
  optimizer = "bobyqa",
  maxfun = 1000000000,
  fullm = TRUE,
  redm = TRUE,
  redmformula = RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays:SEX|SPK_id),
  nullm = TRUE,
  nullmformula = RangeF0st ~ SES + SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowel),
  df_IDS)

## Structure of combined data:
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:7127] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:7127] "C083" "C083" "C083" "C083" ...

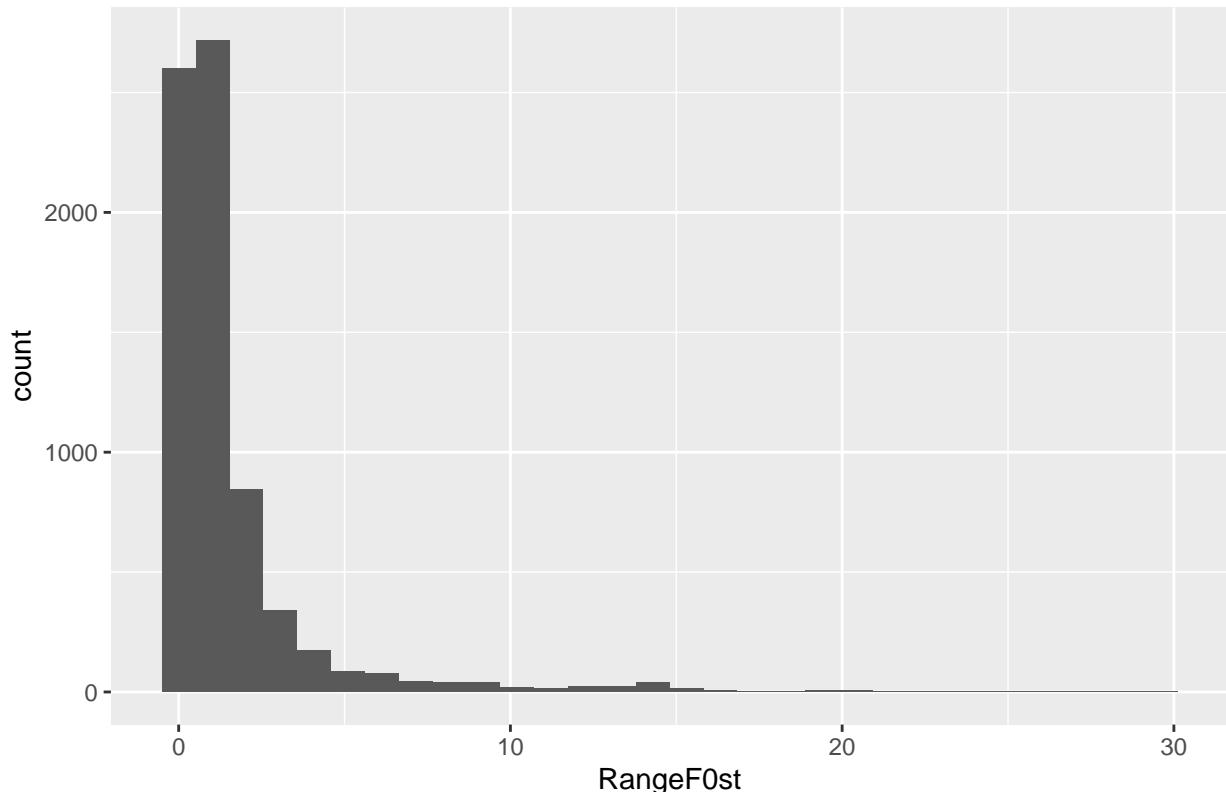
```

```

## $ SES      : num [1:7127] 5 5 5 5 5 5 5 5 5 ...
## $ SEX      : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:7127] 120 120 120 120 120 120 120 120 120 ...
## $ vowels   : chr [1:7127] "a" "eu" "a" "a" ...
## $ RangeF0st: num [1:7127] 3.881 1.649 0.824 0.301 0.178 ...
## Response variable is: RangeF0st
## Structure of combined data after scaling:
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3563
## ..- attr(*, "scaled:scale")= num 2058
## $ SPK_id   : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES      : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.38
## $ SEX      : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 244
## ..- attr(*, "scaled:scale")= num 100
## $ vowels   : chr [1:7127] "a" "eu" "a" "a" ...
## $ RangeF0st: num [1:7127] 3.881 1.649 0.824 0.301 0.178 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...

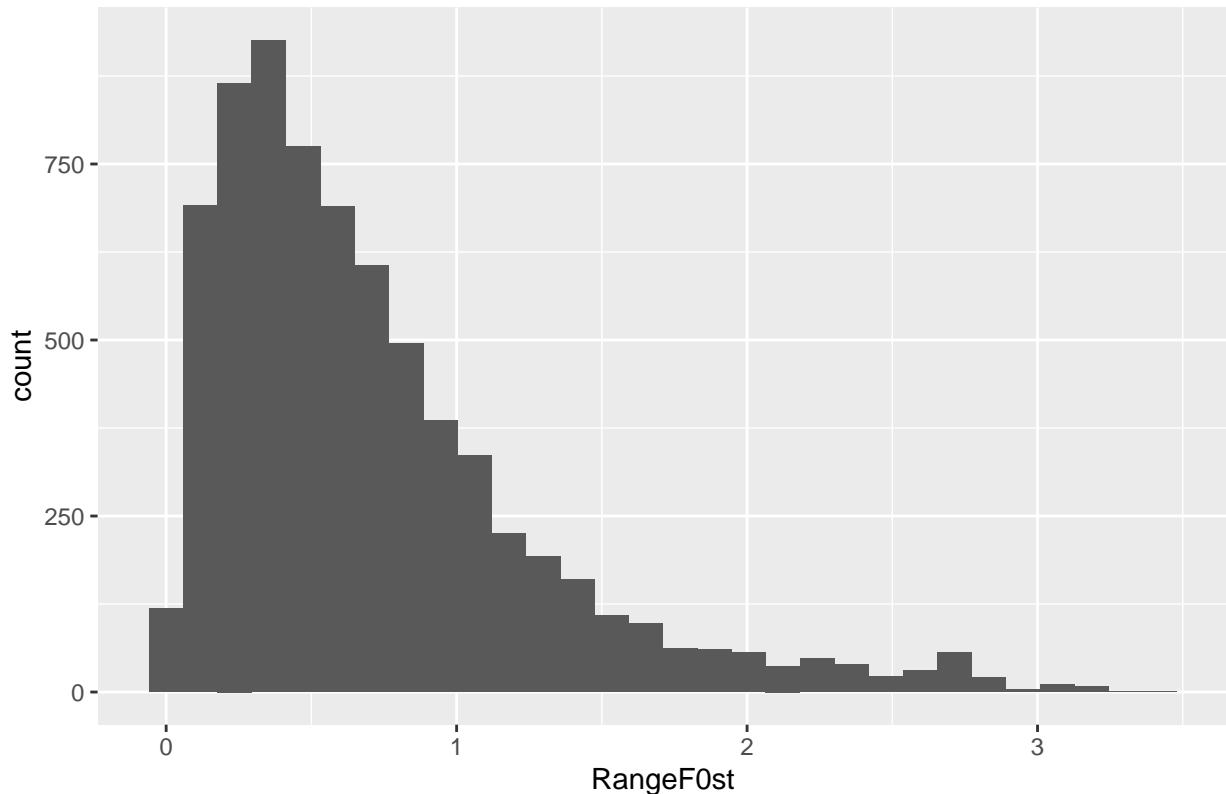
```

```

## ..- attr(*, "scaled:center")= num 3563
## ..- attr(*, "scaled:scale")= num 2058
## $ SPK_id    : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.38
## $ SEX       : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 244
## ..- attr(*, "scaled:scale")= num 100
## $ vowels   : chr [1:7127] "a" "eu" "a" "a" ...
## $ RangeF0st: num [1:7127] 1.585 0.974 0.601 0.263 0.163 ...

```

Histogram of Log-transformed Response Variable

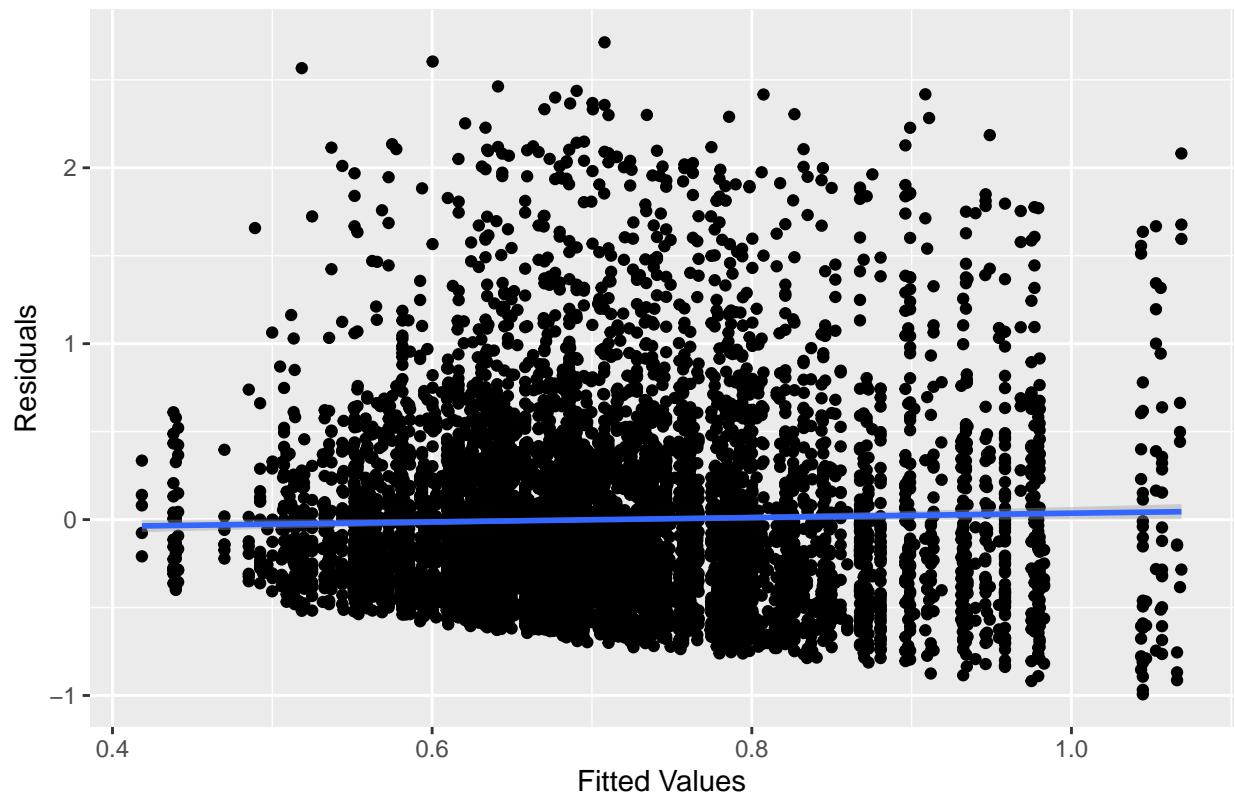


```

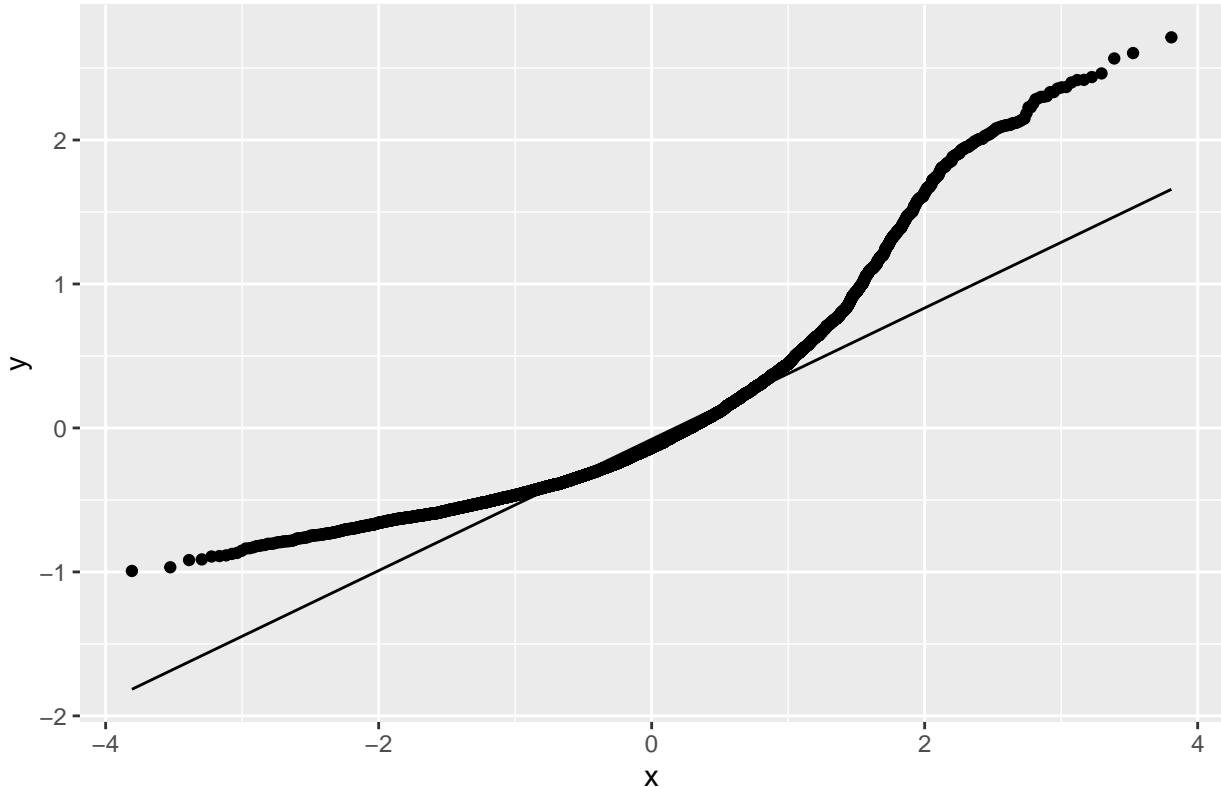
## Fitting Full Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



Q-Q Plot of Residuals for Full Model



```

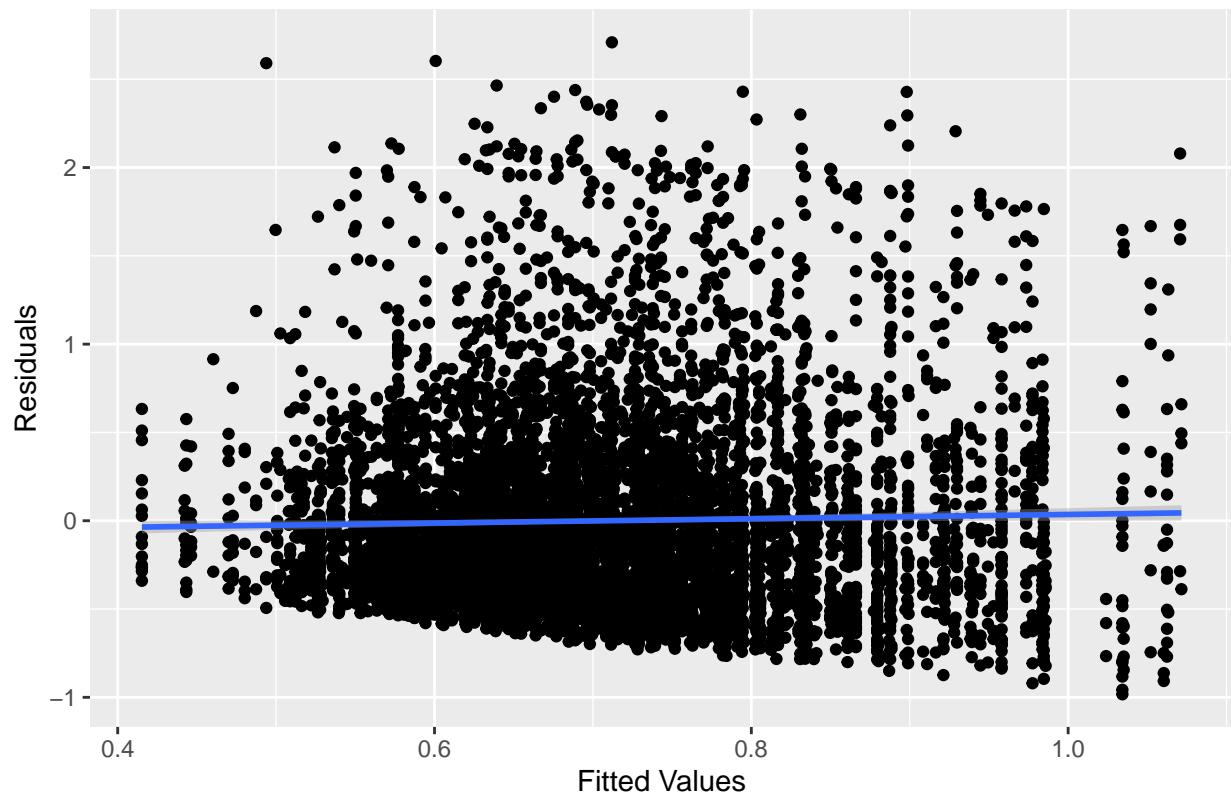
## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 11820.9
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.8101 -0.7019 -0.2480  0.4155  4.9329 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.0158559 0.12592
##          AgeInDays   0.0018765 0.04332 -0.80
## vowels   (Intercept) 0.0065396 0.08087
##          AgeInDays   0.0002553 0.01598 -0.79
## Residual            0.3022076 0.54973
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept)  0.74398   0.04234 24.09095 17.571 3.05e-15 ***
## SES         -0.04728   0.02558 14.33268 -1.849   0.0852 .  
## 
```

```

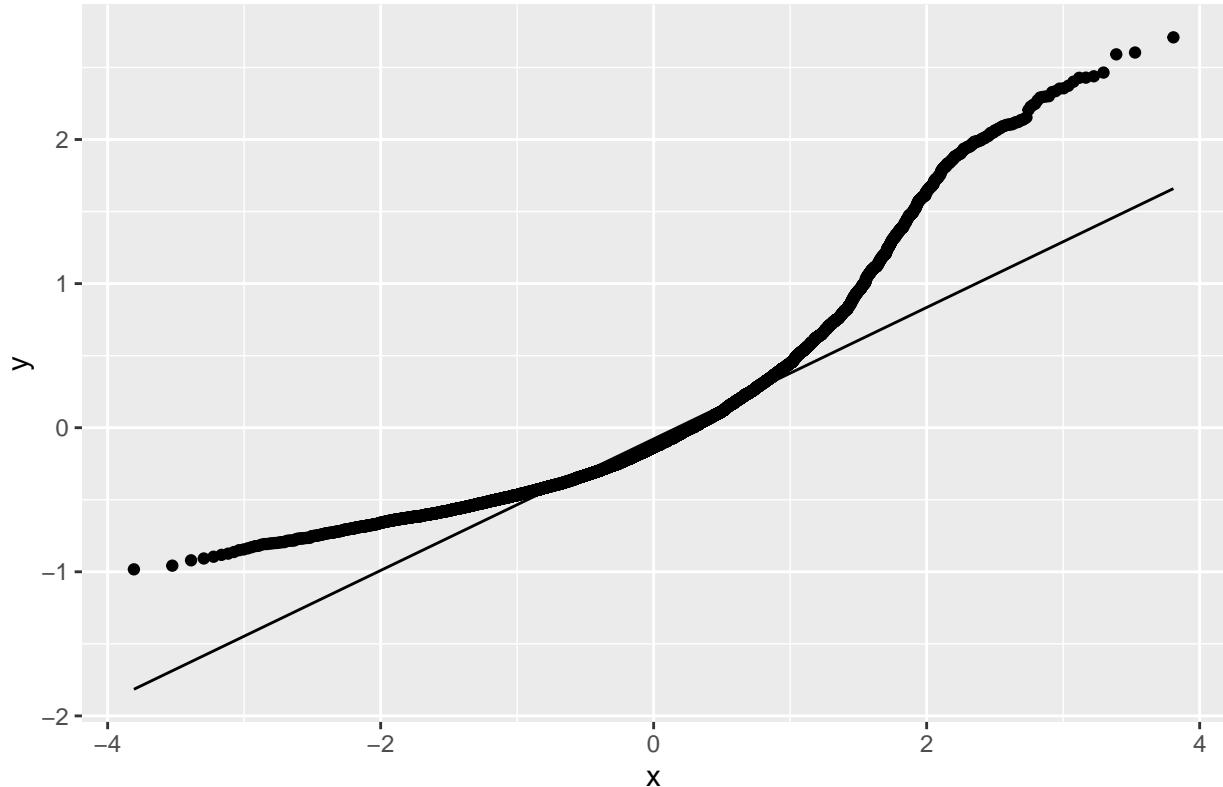
## SEXM          -0.05889   0.05733 21.79540  -1.027   0.3156
## AgeInDays    -0.04573   0.01568  9.32488  -2.916   0.0165 *
## SEXM:AgeInDays 0.01390   0.02602 10.01873   0.534   0.6049
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM    AgInDy
## SES        -0.171
## SEXM       -0.463  0.314
## AgeInDays  -0.574 -0.014  0.293
## SEXM:AgInDy 0.233  0.051 -0.494 -0.522
## Full Model: Log-Likelihood = -5896.67 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.78626297321747
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model 10 11822 11890 -5900.9      11802
## fit_model      12 11817 11900 -5896.7      11793 8.4057  2   0.01495 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 11823.1
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.7934 -0.7012 -0.2446  0.4179  4.9253 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.015969 0.12637  
##          AgeInDays   0.001924 0.04386  -0.82 
## vowels   (Intercept) 0.006401 0.08001  
## Residual            0.302388 0.54990  
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept)  0.74419   0.04221 24.25814 17.629 2.45e-15 ***
## SES         -0.04762   0.02533 14.14254 -1.880  0.0809 .  
## SEXM        -0.05831   0.05741 21.76184 -1.016  0.3210  
## 
```

```

## AgeInDays      -0.04731   0.01467  8.06751  -3.226   0.0120 *
## SEXM:AgeInDays  0.01360   0.02612  9.88208   0.521   0.6140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES    SEXM   AgInDy
## SES       -0.169
## SEXM      -0.464  0.309
## AgeInDays -0.444 -0.016  0.324
## SEXM:AgInDy  0.241  0.052 -0.510 -0.562
## Reduced Model: Log-Likelihood = -5897.79 , Degrees of Freedom = 10
## Max VIF value of the Reduced Model: 1.85460696024849
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## red_fit_model: redmformula
## null_fit_model: nullmformula
##           npar  AIC  BIC logLik deviance Chisq Df Pr(>Chisq)
## red_fit_model   10 11816 11884 -5897.8     11796
## null_fit_model   10 11822 11890 -5900.9     11802     0 0
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_range_st.xlsx"))

## New names:
## * `` -> `...`1` 

# Calculate the quartiles
Q1 <- quantile(df_IDS$RangeF0st, 0.25)
Q3 <- quantile(df_IDS$RangeF0st, 0.75)
# Calculate the interquartile range (IQR)
IQR <- Q3 - Q1
# Define the lower and upper bounds for non-outliers
lower <- Q1 - 1.5 * IQR
upper <- Q3 + 1.5 * IQR

# Remove outliers and create a new dataframe
df_no_outliers_IDS <- df_IDS[df_IDS$RangeF0st >= lower & df_IDS$RangeF0st <= upper, ]

output.3.2_no_outlier <- lmer.full.reduced.null.compare(RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX
                                                       optimizer = "bobyqa",
                                                       maxfun = 1000000000,
                                                       fullm = TRUE,
                                                       redm = TRUE,
                                                       redmformula = RangeF0st ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id) +
                                                       nullm = TRUE,
                                                       nullmformula = RangeF0st ~ SES + SEX + (1|SPK_id) + (1|vowels),
                                                       df_no_outliers_IDS)

## Structure of combined data:
## tibble [6,456 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:6456] 1 2 3 4 5 6 7 9 10 11 ...
## $ SPK_id    : chr [1:6456] "C083" "C083" "C083" "C083" ...

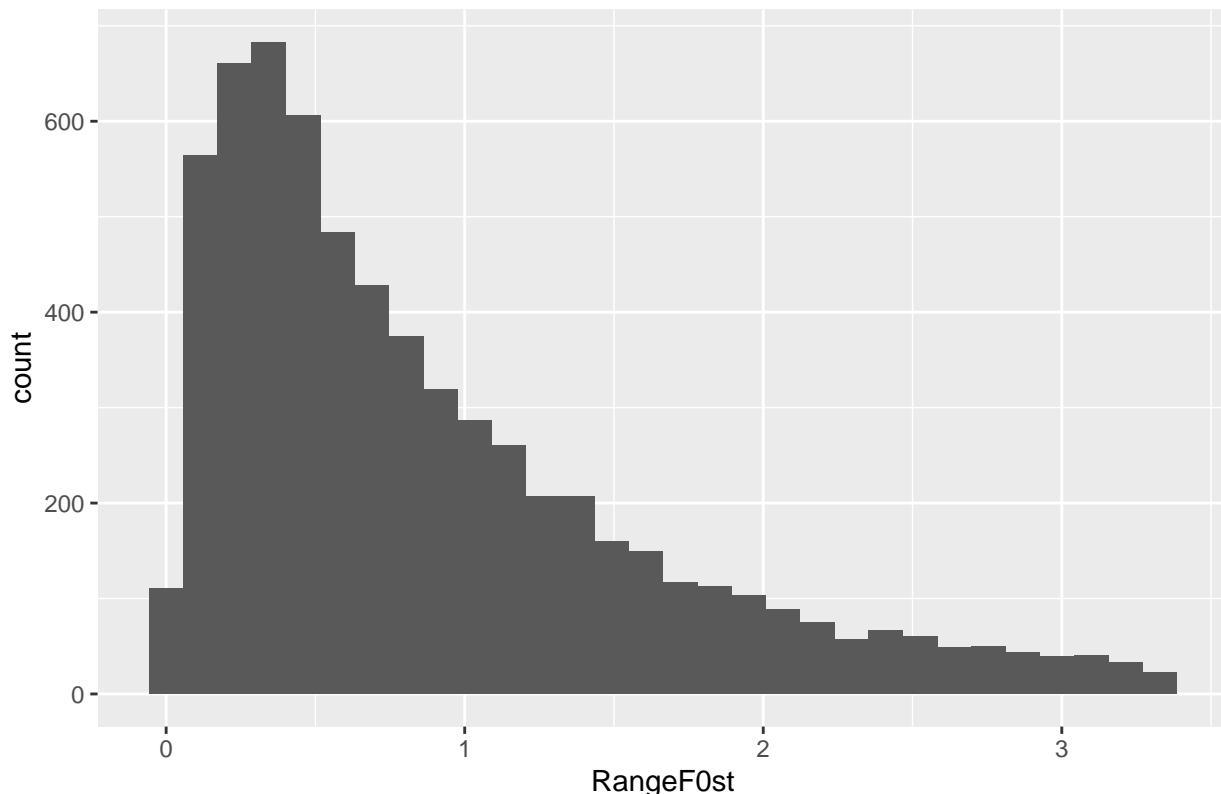
```

```

## $ SES      : num [1:6456] 5 5 5 5 5 5 5 5 5 ...
## $ SEX      : chr [1:6456] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:6456] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels   : chr [1:6456] "eu" "a" "a" "e" ...
## $ RangeF0st: num [1:6456] 1.649 0.824 0.301 0.178 1.275 ...
## Response variable is: RangeF0st
## Structure of combined data after scaling:
## tibble [6,456 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:6456, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3567
## ..- attr(*, "scaled:scale")= num 2060
## $ SPK_id   : chr [1:6456] "C083" "C083" "C083" "C083" ...
## $ SES      : num [1:6456, 1] -0.00458 -0.00458 -0.00458 -0.00458 -0.00458 ...
## ..- attr(*, "scaled:center")= num 5.01
## ..- attr(*, "scaled:scale")= num 1.39
## $ SEX      : chr [1:6456] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:6456, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 245
## ..- attr(*, "scaled:scale")= num 101
## $ vowels   : chr [1:6456] "eu" "a" "a" "e" ...
## $ RangeF0st: num [1:6456] 1.649 0.824 0.301 0.178 1.275 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [6,456 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1     : num [1:6456, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...

```

```

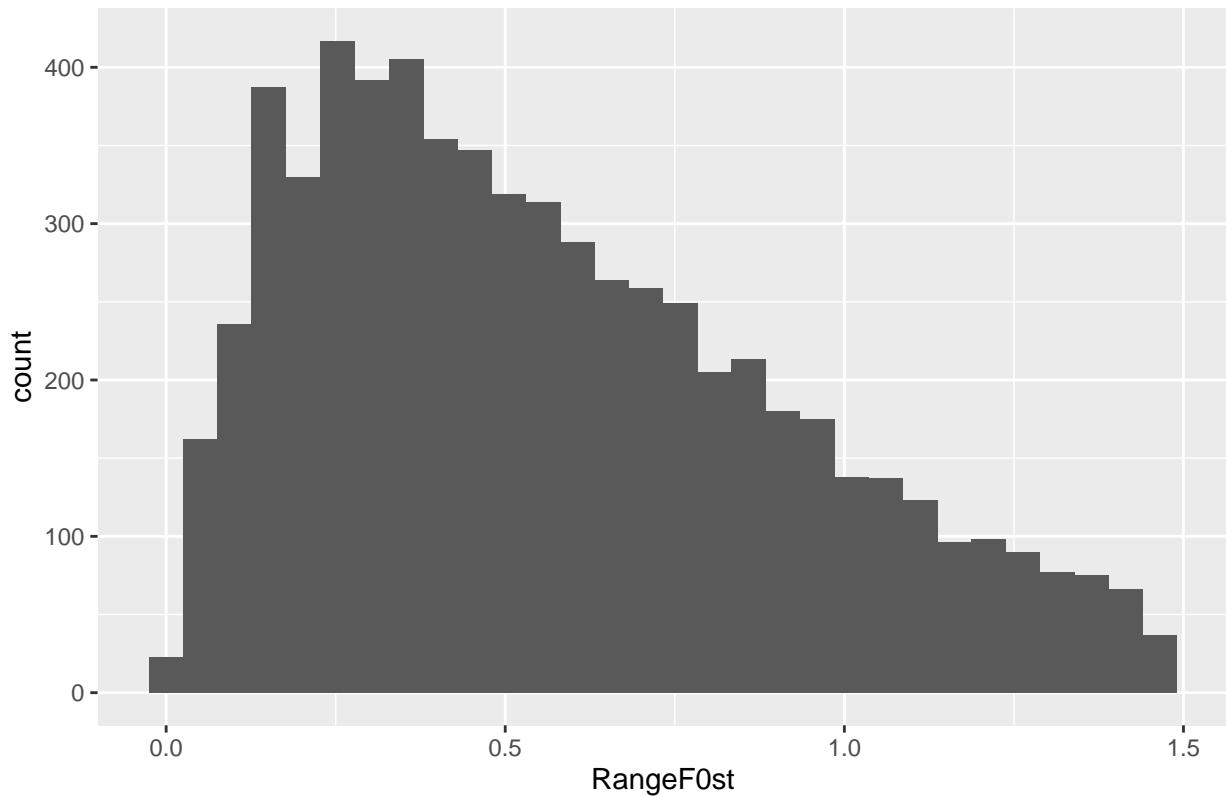
## ..- attr(*, "scaled:center")= num 3567
## ..- attr(*, "scaled:scale")= num 2060
## $ SPK_id    : chr [1:6456] "C083" "C083" "C083" "C083" ...
## $ SES        : num [1:6456, 1] -0.00458 -0.00458 -0.00458 -0.00458 -0.00458 ...
## ..- attr(*, "scaled:center")= num 5.01
## ..- attr(*, "scaled:scale")= num 1.39
## $ SEX        : chr [1:6456] "F" "F" "F" "F" ...
## $ AgeInDays: num [1:6456, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 245
## ..- attr(*, "scaled:scale")= num 101
## $ vowels    : chr [1:6456] "eu" "a" "a" "e" ...
## $ RangeF0st: num [1:6456] 0.974 0.601 0.263 0.163 0.822 ...

## Fitting Full Model.

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

```

Histogram of Log-transformed Response Variable

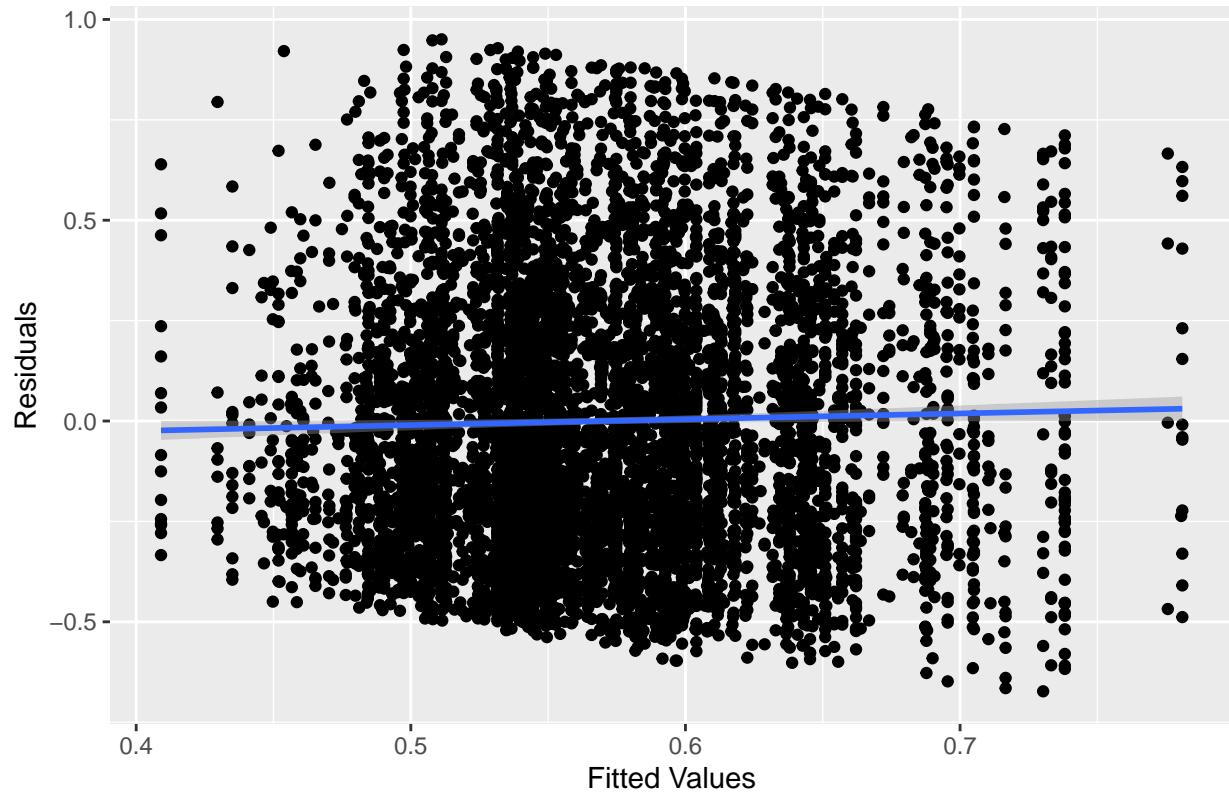


```

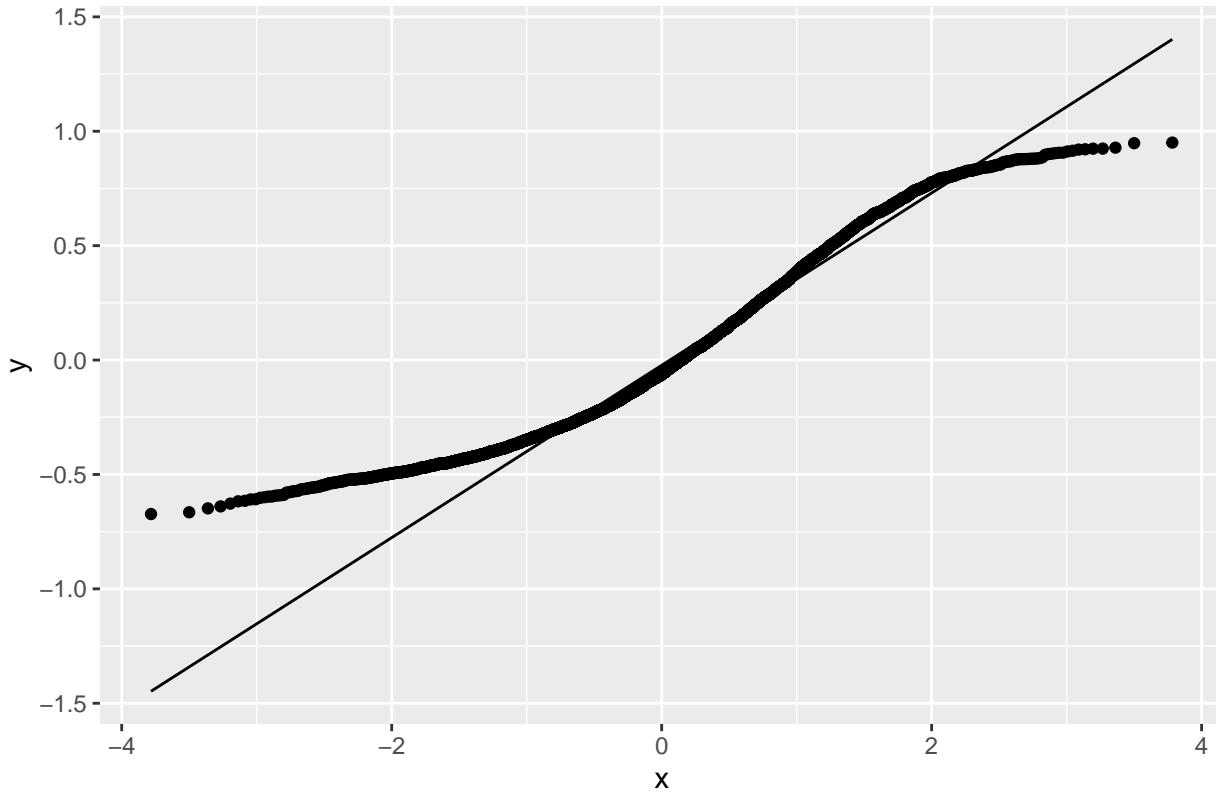
## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 4653.4
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.9705 -0.8057 -0.1812  0.6700  2.7671 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 4.221e-03 0.064966 
##          AgeInDays   5.294e-05 0.007276 -0.23 
## vowels   (Intercept) 1.486e-03 0.038550 
##          AgeInDays   4.687e-06 0.002165 -0.52 
## Residual           1.184e-01 0.344165 
## Number of obs: 6456, groups: SPK_id, 27; vowels, 9 
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  0.586363  0.021826 24.378824 26.865 < 2e-16 ***
## SES         -0.014276  0.016039 20.071964 -0.890  0.38396  

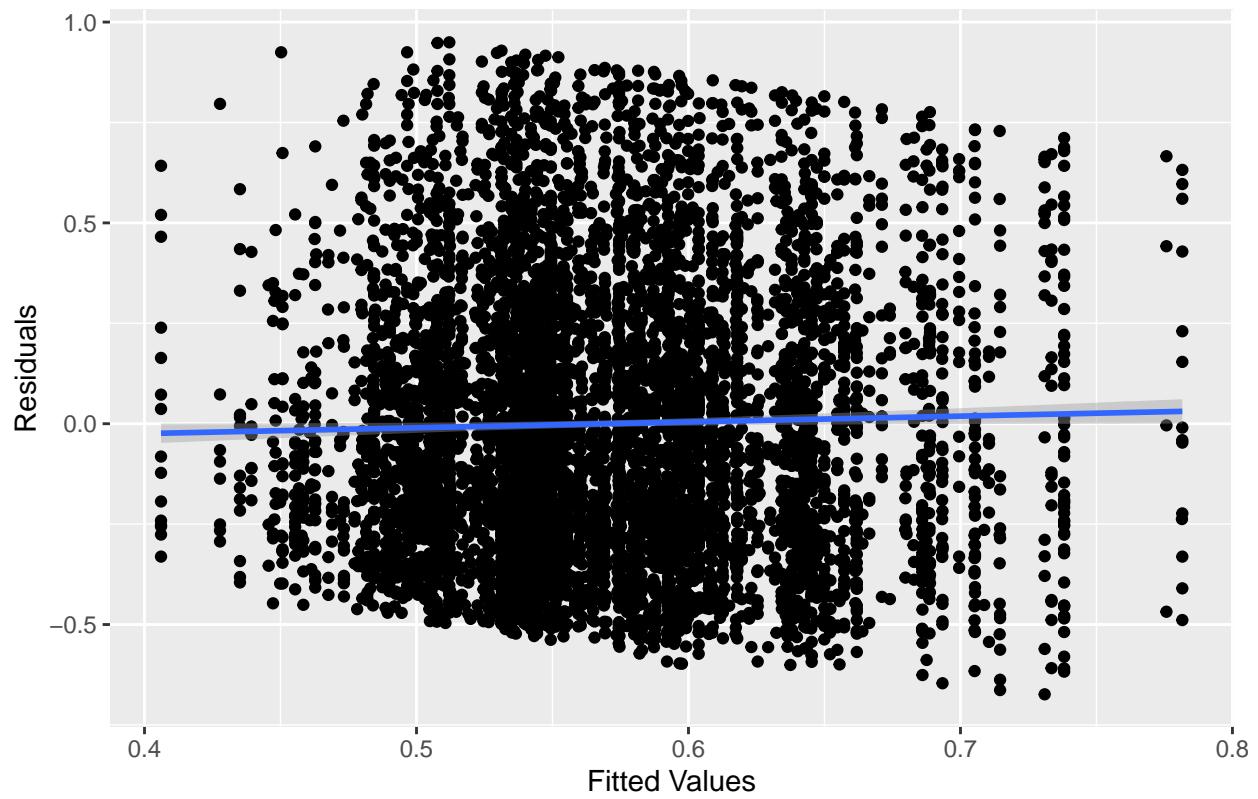
```

```

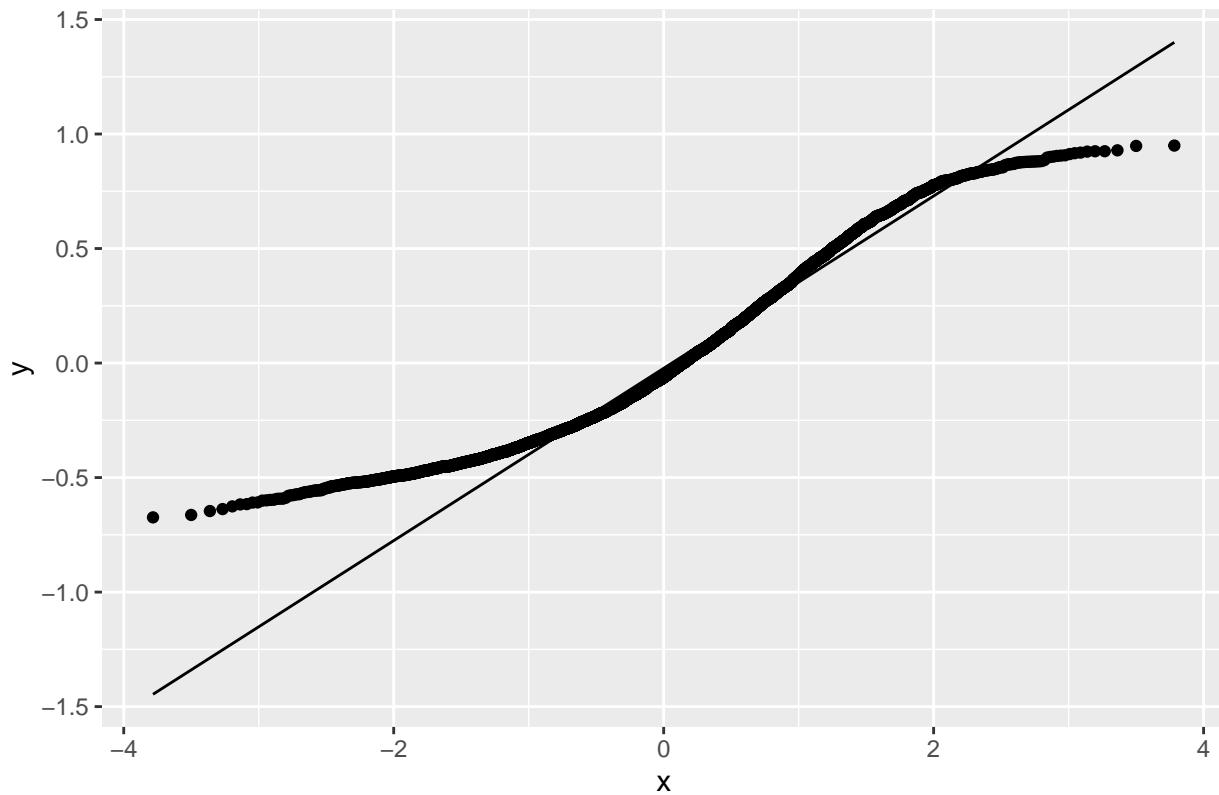
## SEXM          -0.042063  0.031259 22.293051 -1.346  0.19195
## AgeInDays    -0.023197  0.006244  6.221648 -3.715  0.00928 **
## SEXM:AgeInDays 0.017796  0.011568 12.710861  1.538  0.14846
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM   AgInDy
## SES        -0.229
## SEXM       -0.496  0.387
## AgeInDays  -0.106  0.041  0.058
## SEXM:AgInDy 0.039 -0.033 -0.045 -0.530
## Full Model: Log-Likelihood = -2309.77 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.3898767051865
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model    6 4648.1 4688.8 -2318.1    4636.1
## fit_model        12 4643.5 4724.8 -2309.8    4619.5 16.611  6   0.01082 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 4653.6
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.9712 -0.8033 -0.1792  0.6710  2.7607 
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.004173 0.06460
## vowels   (Intercept) 0.001487 0.03856
## Residual            0.118491 0.34423
## Number of obs: 6456, groups:  SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept) 5.863e-01 2.175e-02 2.476e+01 26.958 < 2e-16 ***
## SES         -1.305e-02 1.591e-02 2.107e+01 -0.820  0.4214    
## SEXM        -4.201e-02 3.108e-02 2.298e+01 -1.352  0.1896    
## AgeInDays   -2.324e-02 5.752e-03 5.018e+03 -4.040 5.42e-05 ***

```

```

## SEXM:AgeInDays  1.921e-02  1.087e-02  3.914e+03    1.767   0.0773 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.230
## SEXM     -0.495  0.388
## AgeInDays -0.022  0.047  0.027
## SEXM:AgInDy  0.016 -0.046  0.006 -0.529
## Reduced Model: Log-Likelihood = -2309.79 , Degrees of Freedom = 8
## Max VIF value of the Reduced Model: 1.39358191156262
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model    6 4648.1 4688.8 -2318.1    4636.1
## red_fit_model     8 4635.6 4689.8 -2309.8    4619.6 16.565  2  0.0002529 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).

pred.data.3.2_no_outlier <- boot.ci.predict.lmer(
  m = output.3.2_no_outlier$reduced_model,
  optimizer = output.3.2_no_outlier$optimizer,
  maxfun = output.3.2_no_outlier$maxfun,
  data = output.3.2_no_outlier$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.3.2_no_outlier$fit_data$AgeInDays), max(output.3.2_no_outlier$fit_data$AgeInDays)),
  reqcol = c("AgeInDays"),
  centercol = c("SES", "SEX"),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## Number of singular fits during bootstrapping: 0 out of 1000 bootstrap iterations.

pred.data.3.2 <- boot.ci.predict.lmer(
  m = output.3.2$full_model,
  optimizer = output.3.2$optimizer,
  maxfun = output.3.2$maxfun,
  data = output.3.2$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.3.2$fit_data$AgeInDays), max(output.3.2$fit_data$AgeInDays)),
  reqcol = c("AgeInDays"),
  centercol = c("SES", "SEX"),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

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

```



```

## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 193 out of 1000 bootstrap iterations.
inv_x_transform_fun_1 <- function(x) {
  x * attr(output.3.1$fit_data$AgeInDays, "scaled:scale") + attr(output.3.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.3.2$fit_data$AgeInDays, "scaled:scale") + attr(output.3.2$fit_data$AgeInDays, "scaled")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

pitch_range_semitones <- continuous_fit_ci_plot(
  plot.data = list(output.3.1$fit_data, output.3.2$fit_data),
  coefs = list(pred.data.3.1,pred.data.3.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "RangeFOst",
  x.labs = "Age in Days",
  y.labs = "Pitch Range(semitones)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),

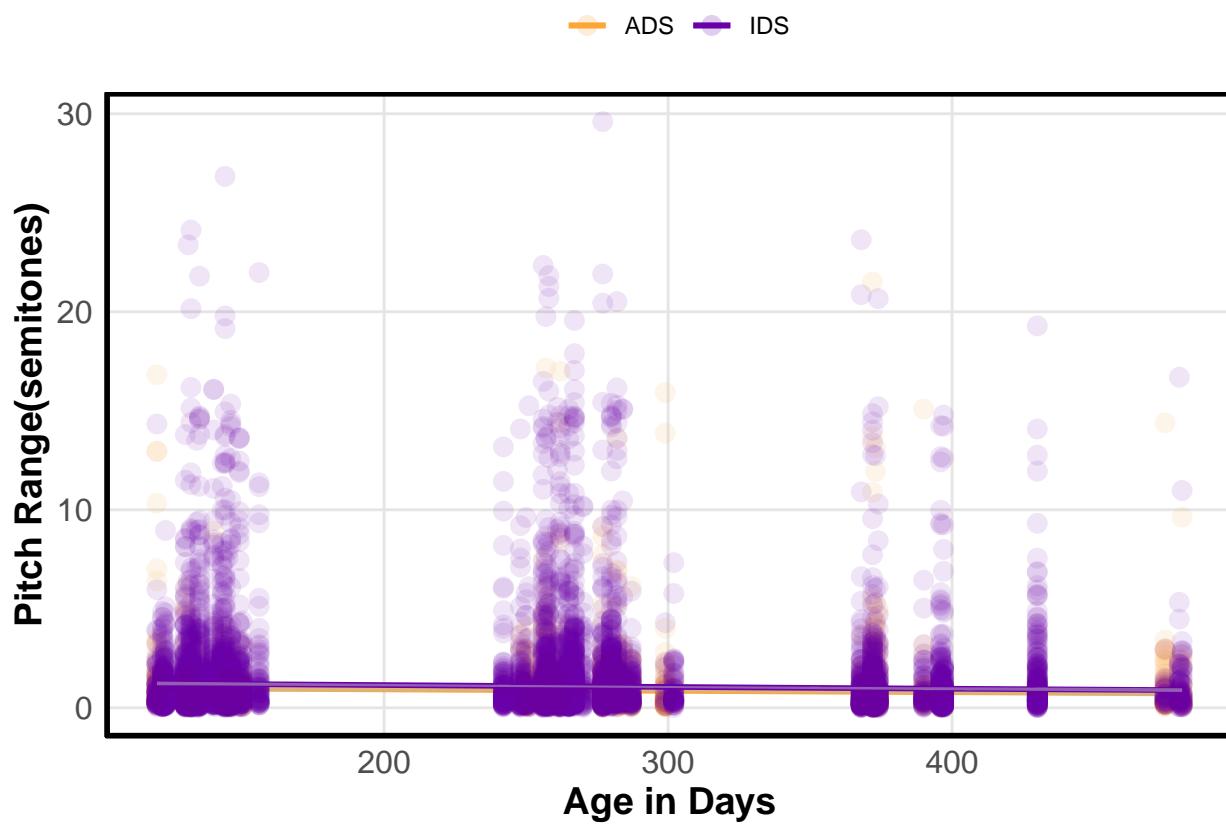
```

```

#y.lim = c(-1, ),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
x.ax.transformation = TRUE,
inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
y.ax.transformation = TRUE,
inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
#y_grid_log = TRUE,
#x.breaks = seq(100, 400, by = 100),
y.breaks = seq(0, 20, by = 10),
legend.labels = c("ADS", "IDS")
)

print(pitch_range_semitones)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.3.1$fit_data$AgeInDays, "scaled:scale") + attr(output.3.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.3.2$fit_data$AgeInDays, "scaled:scale") + attr(output.3.2$fit_data$AgeInDays, "scaled")
}
inv_y_transform_fun_1 <- function(x) { exp(x)-1 + (x = ifelse(x <= 0, 0.1, x)) }
inv_y_transform_fun_2 <- function(x) { exp(x)-1 + (x = ifelse(x <= 0, 0.1, x)) }

pitch_range_semitones_log_grid <- continuous_fit_ci_plot(
  plot.data = list(output.3.1$fit_data, output.3.2$fit_data),
  coefs = list(pred.data.3.1,pred.data.3.2),
  plot.data.col.x = "AgeInDays",

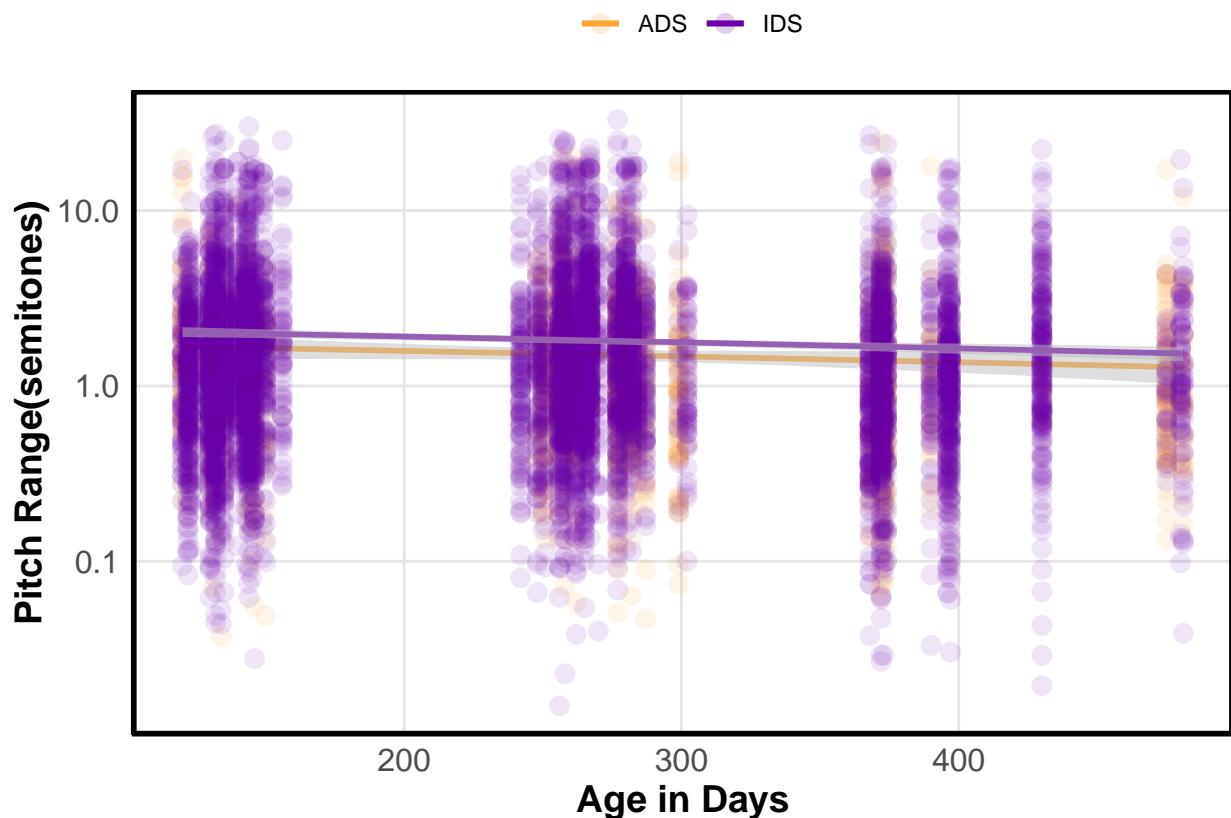
```

```

plot.data.col.y = "RangeF0st",
x.labs = "Age in Days",
y.labs = "Pitch Range(semitones)",
#x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
#y.lim = c(-1, ),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
x.ax.transformation = TRUE,
inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
y.ax.transformation = TRUE,
inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
y_grid_log = TRUE,
#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(0, 20, by = 10),
legend.labels = c("ADS", "IDS")
)

print(pitch_range_semitones_log_grid)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.3.1_no_outlier$fit_data$AgeInDays, "scaled:scale") + attr(output.3.1_no_outlier$fit_d
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.3.2_no_outlier$fit_data$AgeInDays, "scaled:scale") + attr(output.3.2_no_outlier$fit_d
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

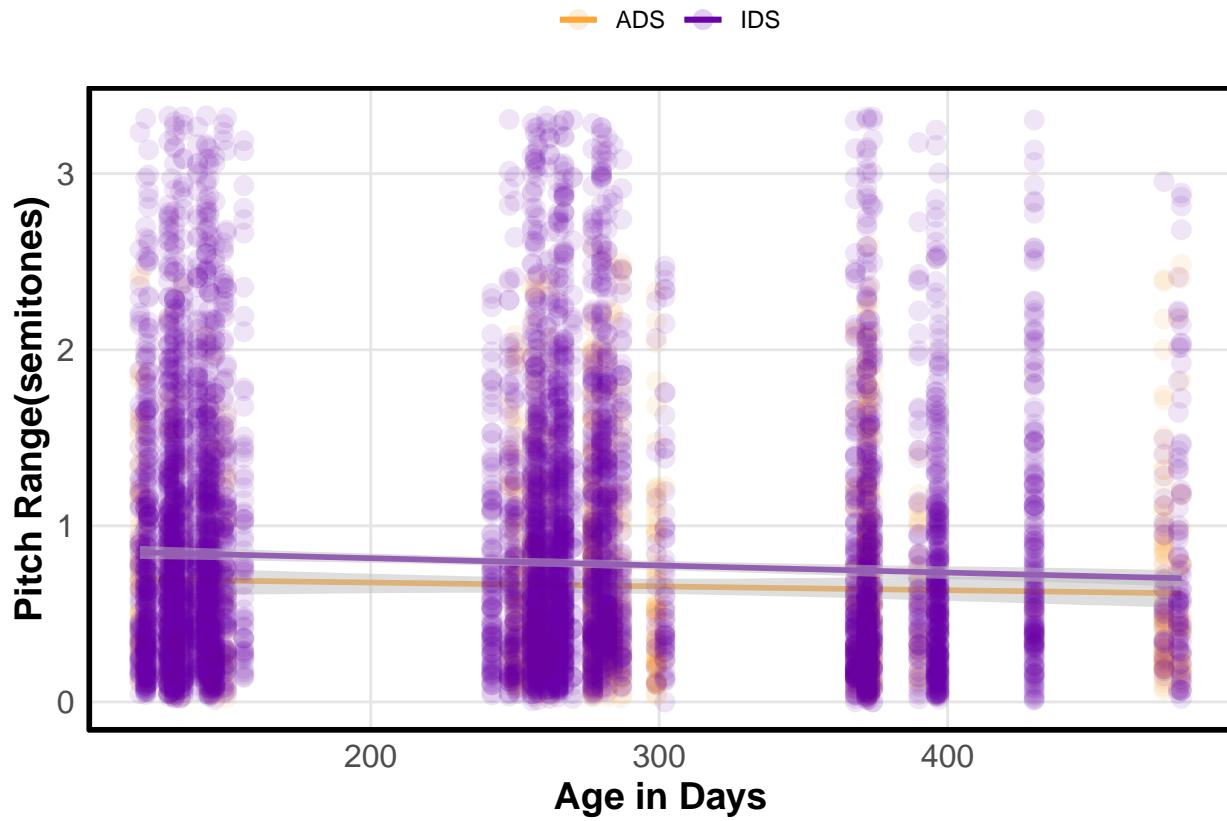
```

```

pitch_range_semitones_no_outlier <- continuous_fit_ci_plot(
  plot.data = list(output.3.1_no_outlier$fit_data, output.3.2_no_outlier$fit_data),
  coefs = list(pred.data.3.1_no_outlier, pred.data.3.2_no_outlier),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "RangeF0st",
  x.labs = "Age in Days",
  y.labs = "Pitch Range(semitones)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = TRUE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  legend.labels = c("ADS", "IDS")
)

print(pitch_range_semitones_no_outlier)

```



```

#4.Duration #4.1 ADS
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_dur_ms.xlsx"))

## New names:
## * `` -> `...`1`
```

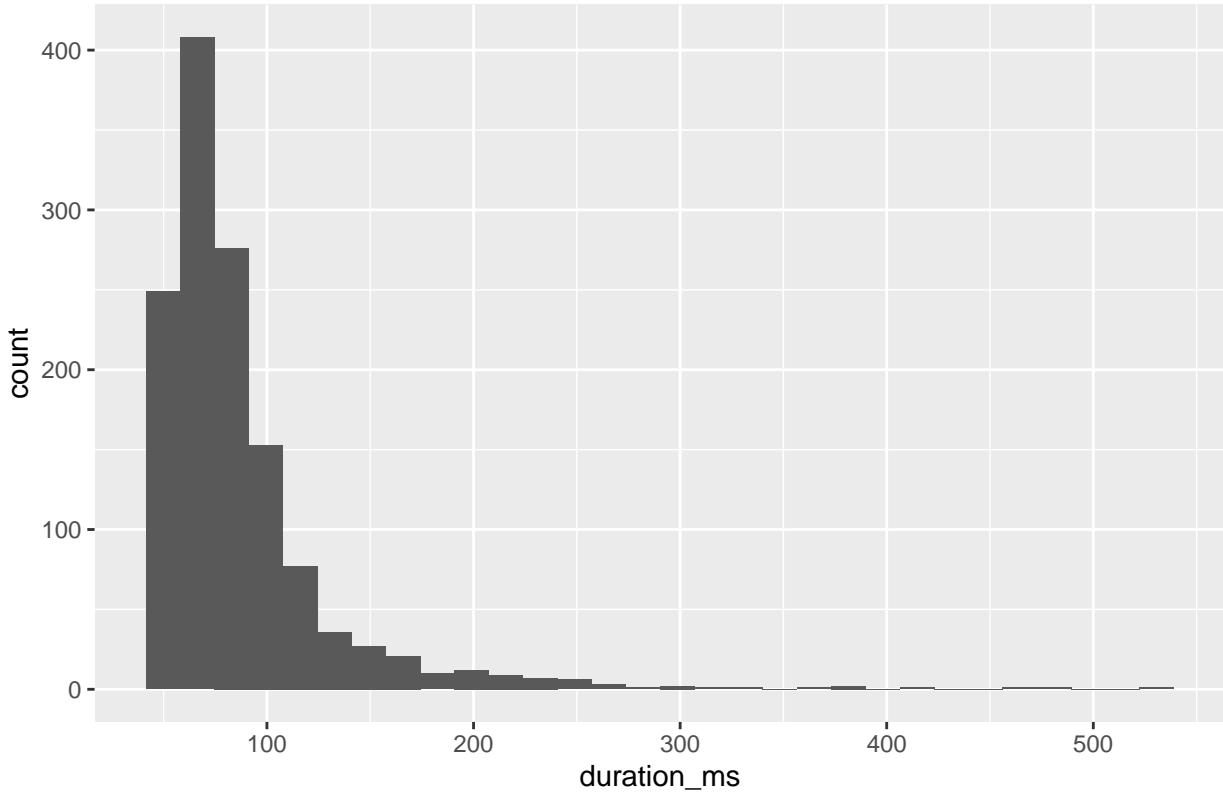
```

output.4.1 <- lmer.full.reduced.null.compare(duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|AgeInDays:SES:SPK_id),
                                             optimizer = "bobyqa",
                                             maxfun = 1000000000,
                                             fullm = TRUE,
                                             redm = TRUE,
                                             redmformula = duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                             nullm = TRUE,
                                             nullmformula = duration_ms ~ SES + SEX + (1|SPK_id) + (1|vowels),
                                             df_ADS)

## Structure of combined data:
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:1306] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES        : num [1:1306] 5 5 5 5 5 5 5 5 5 5 ...
## $ SEX        : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays  : num [1:1306] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels     : chr [1:1306] "o" "a" "i" "a" ...
## $ duration_ms: num [1:1306] 66.4 82.2 106.3 86.4 84.6 ...
## Response variable is: duration_ms
## Structure of combined data after scaling:
## tibble [1,306 x 7] (S3:tbl_df/tbl/data.frame)
## $ ...1      : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 -1.72 ...
## $ ..- attr(*, "scaled:center")= num 652
## $ ..- attr(*, "scaled:scale")= num 377
## $ SPK_id    : chr [1:1306] "C083" "C083" "C083" "C083" ...
## $ SES        : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...
## $ ..- attr(*, "scaled:center")= num 5.14
## $ ..- attr(*, "scaled:scale")= num 1.3
## $ SEX        : chr [1:1306] "F" "F" "F" "F" ...
## $ AgeInDays  : num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...
## $ ..- attr(*, "scaled:center")= num 295
## $ ..- attr(*, "scaled:scale")= num 97.1
## $ vowels     : chr [1:1306] "o" "a" "i" "a" ...
## $ duration_ms: num [1:1306] 66.4 82.2 106.3 86.4 84.6 ...

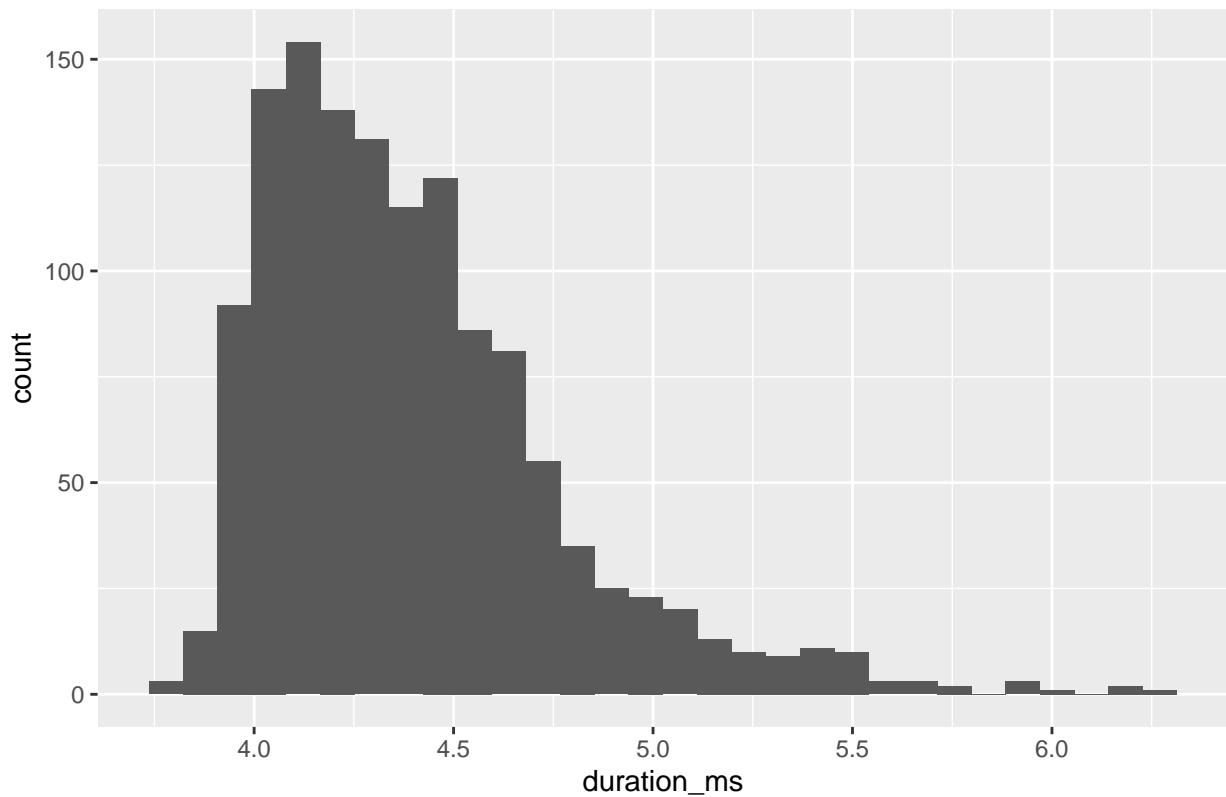
```

Histogram of Response Variable



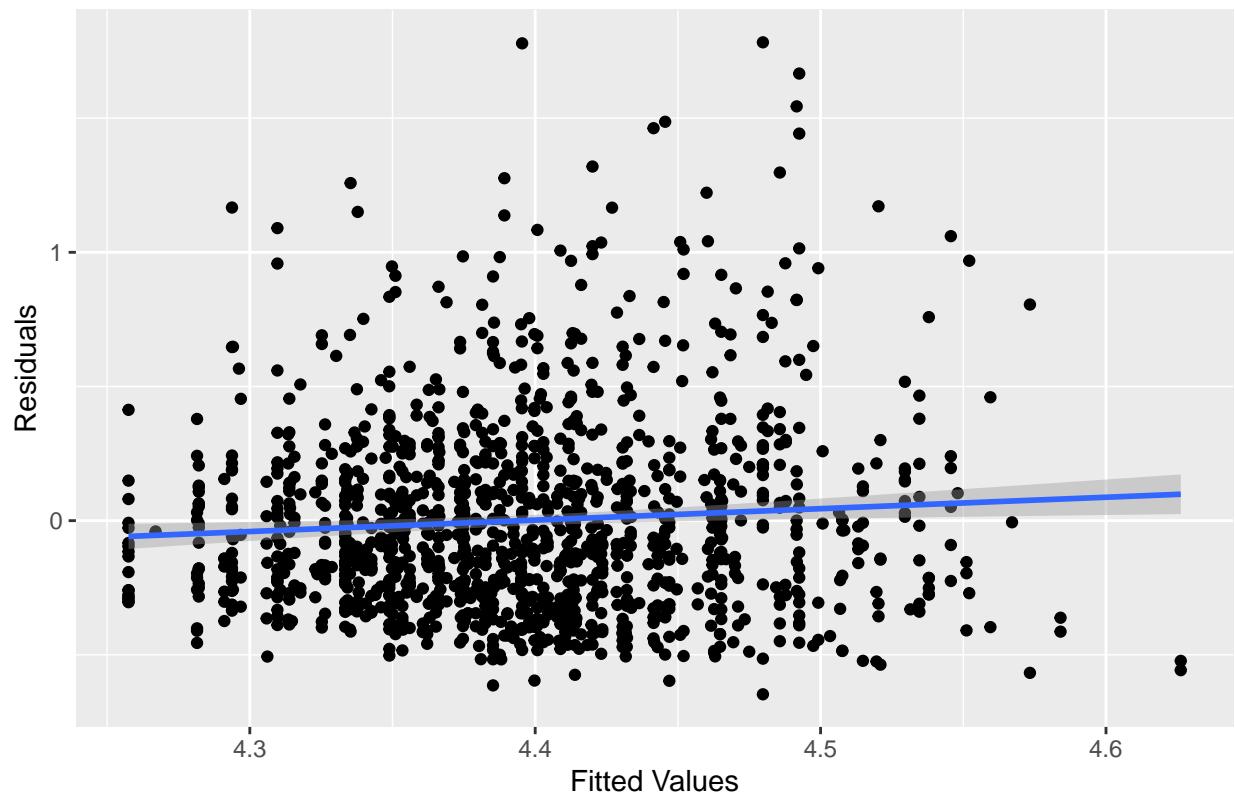
```
## Response variable is not normally distributed. Log transformation applied.  
## Structure of combined data after transformation:  
## tibble [1,306 x 7] (S3: tbl_df/tbl/data.frame)  
## $ ...1 : num [1:1306, 1] -1.73 -1.73 -1.72 -1.72 -1.72 ...  
## ..- attr(*, "scaled:center")= num 652  
## ..- attr(*, "scaled:scale")= num 377  
## $ SPK_id : chr [1:1306] "C083" "C083" "C083" "C083" ...  
## $ SES : num [1:1306, 1] -0.106 -0.106 -0.106 -0.106 -0.106 ...  
## ..- attr(*, "scaled:center")= num 5.14  
## ..- attr(*, "scaled:scale")= num 1.3  
## $ SEX : chr [1:1306] "F" "F" "F" "F" ...  
## $ AgeInDays : num [1:1306, 1] -1.8 -1.8 -1.8 -1.8 -1.8 ...  
## ..- attr(*, "scaled:center")= num 295  
## ..- attr(*, "scaled:scale")= num 97.1  
## $ vowels : chr [1:1306] "o" "a" "i" "a" ...  
## $ duration_ms: num [1:1306] 4.21 4.42 4.68 4.47 4.45 ...  
## Fitting Full Model.  
## boundary (singular) fit: see help('isSingular')
```

Histogram of Log-transformed Response Variable

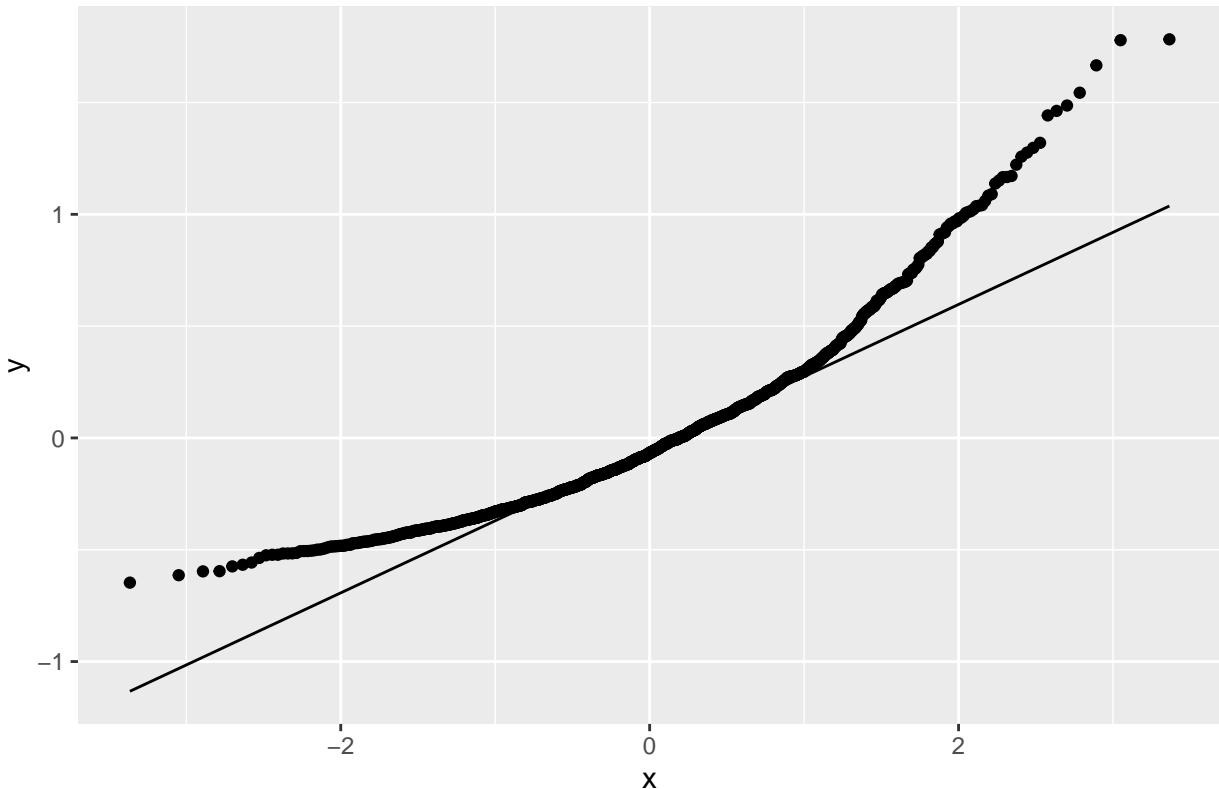


```
## Full model is singular. Consider simplifying the random effects structure.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1119.1
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.7943 -0.7156 -0.1914  0.4608  4.8878 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.0036514 0.06043
##          AgeInDays   0.0001921 0.01386 -0.52
## vowels   (Intercept) 0.0011087 0.03330
##          AgeInDays   0.0006129 0.02476 -0.47
## Residual            0.1322722 0.36369
## Number of obs: 1306, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept)  4.41103   0.02489 17.33898 177.226 <2e-16 ***
## SES         -0.02131   0.01790 18.73211  -1.191    0.249  
## 
```

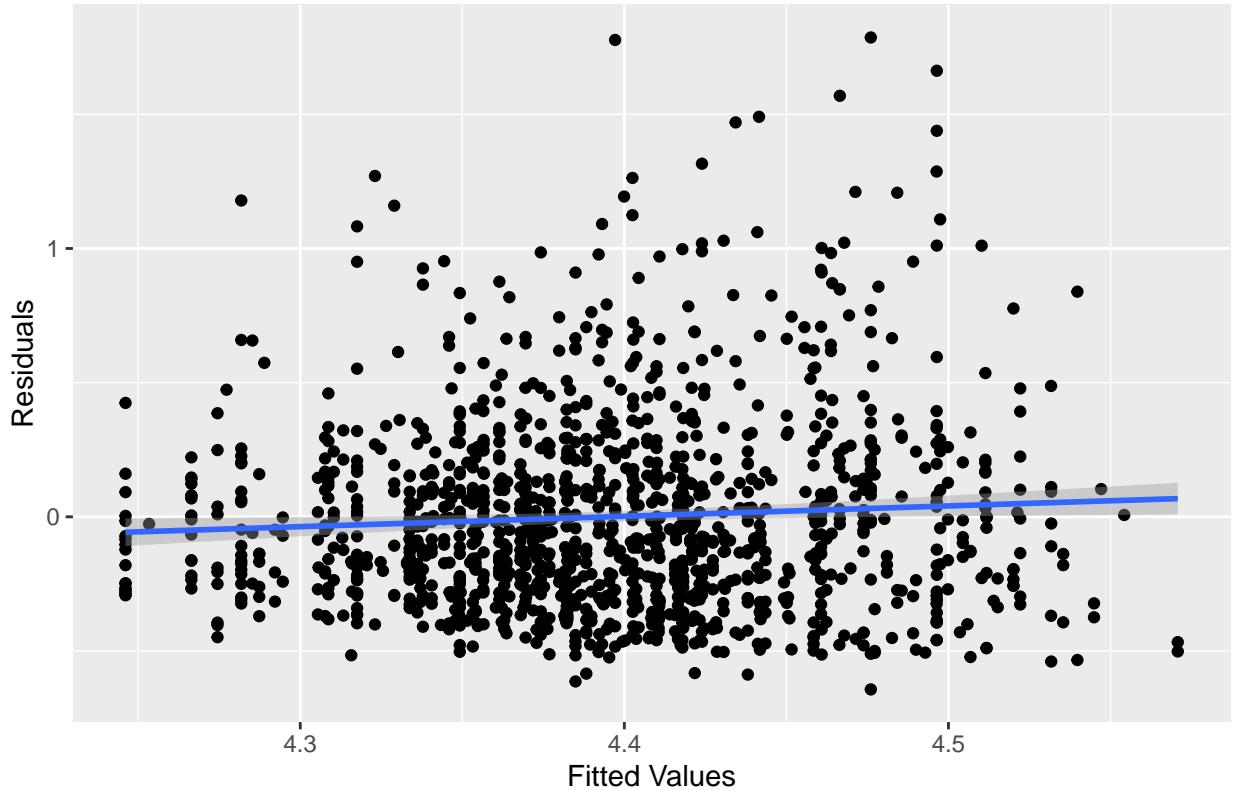
```

##  SEXM      -0.03924   0.04027 18.05808  -0.974    0.343
##  AgeInDays -0.02407   0.01992  6.19520  -1.209    0.271
##  SEXM:AgeInDays -0.01583   0.02479  4.98258  -0.639    0.551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.226
## SEXM     -0.525  0.412
## AgeInDays -0.083 -0.138 -0.056
## SEXM:AgInDy -0.014  0.076 -0.045 -0.616
## Full Model: Log-Likelihood = -544.81 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.75267660356173
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model   6 1110.6 1141.6 -549.28    1098.6
## fit_model        12 1113.6 1175.7 -544.81    1089.6 8.939  6       0.177
## Fitting Reduced Model.

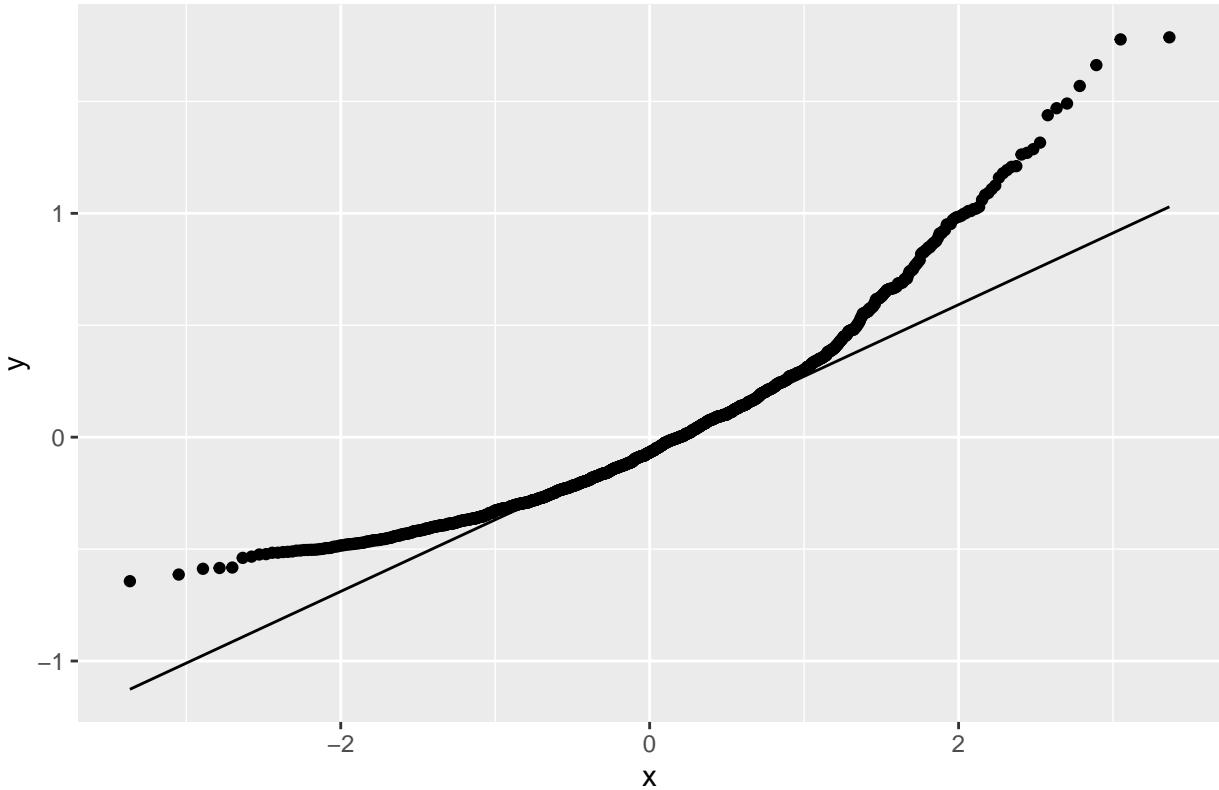
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1120.4
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.7816 -0.7196 -0.1889  0.4591  4.8836 
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.003704 0.06086
## vowels   (Intercept) 0.001003 0.03167
## Residual            0.132879 0.36453
## Number of obs: 1306, groups:  SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  4.41200  0.02470 17.30070 178.659 <2e-16 ***
## SES          -0.02297  0.01808 20.48770 -1.270   0.2182    
## SEXM         -0.04123  0.04058 18.05439 -1.016   0.3230    
## AgeInDays    -0.02725  0.01628 504.57079 -1.674   0.0948 .  

```

```

## SEXM:AgeInDays -0.01204    0.02354 414.75926 -0.511    0.6094
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.223
## SEXM     -0.529  0.411
## AgeInDays  0.122 -0.153 -0.122
## SEXM:AgInDy -0.079  0.073  0.069 -0.686
## Reduced Model: Log-Likelihood = -545.3 , Degrees of Freedom = 8
## Max VIF value of the Reduced Model: 1.94669004716573
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model   6 1110.6 1141.6 -549.28    1098.6
## red_fit_model    8 1106.6 1148.0 -545.30    1090.6 7.9681  2    0.01861 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_dur_ms.xlsx"))

## New names:
## * `` -> `...`1` 

# Calculate the quartiles
Q1 <- quantile(df_ADS$duration_ms, 0.25)
Q3 <- quantile(df_ADS$duration_ms, 0.75)
# Calculate the interquartile range (IQR)
IQR <- Q3 - Q1
# Define the lower and upper bounds for non-outliers
lower <- Q1 - 1.5 * IQR
upper <- Q3 + 1.5 * IQR

# Remove outliers and create a new dataframe
df_no_outliers_ADS <- df_ADS[df_ADS$duration_ms >= lower & df_ADS$duration_ms <= upper, ]

output.4.1_no_outliers <- lmer.full.reduced.null.compare(duration_ms ~ SES + SEX + AgeInDays + AgeInDay,
                                                          optimizer = "bobyqa",
                                                          maxfun = 10000000000,
                                                          fullm = TRUE,
                                                          redm = TRUE,
                                                          redmformula = duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                                          nullm = TRUE,
                                                          nullmformula = duration_ms ~ SES + SEX + (1|SPK_id) + (1|vowels),
                                                          df_no_outliers_ADS)

## Structure of combined data:
## tibble [1,209 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:1209] 0 1 2 3 4 6 7 8 9 10 ...

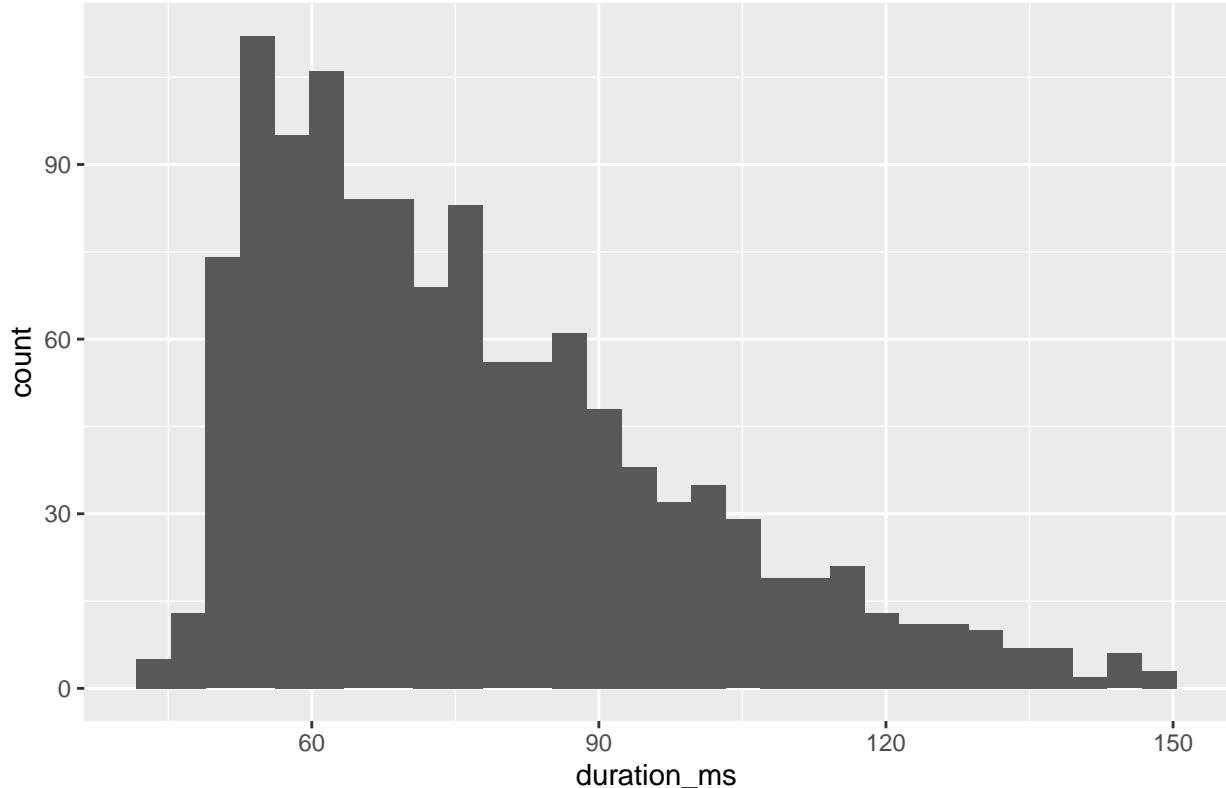
```

```

## $ SPK_id      : chr [1:1209] "C083" "C083" "C083" "C083" ...
## $ SES         : num [1:1209] 5 5 5 5 5 5 5 5 5 ...
## $ SEX         : chr [1:1209] "F" "F" "F" "F" ...
## $ AgeInDays   : num [1:1209] 120 120 120 120 120 120 120 120 120 ...
## $ vowels      : chr [1:1209] "o" "a" "i" "a" ...
## $ duration_ms: num [1:1209] 66.4 82.2 106.3 86.4 84.6 ...
## Response variable is: duration_ms
## Structure of combined data after scaling:
## tibble [1,209 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1        : num [1:1209, 1] -1.75 -1.75 -1.75 -1.74 -1.74 ...
## ..- attr(*, "scaled:center")= num 660
## ..- attr(*, "scaled:scale")= num 377
## $ SPK_id      : chr [1:1209] "C083" "C083" "C083" ...
## $ SES         : num [1:1209, 1] -0.108 -0.108 -0.108 -0.108 -0.108 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.3
## $ SEX         : chr [1:1209] "F" "F" "F" "F" ...
## $ AgeInDays   : num [1:1209, 1] -1.82 -1.82 -1.82 -1.82 -1.82 ...
## ..- attr(*, "scaled:center")= num 296
## ..- attr(*, "scaled:scale")= num 97
## $ vowels      : chr [1:1209] "o" "a" "i" "a" ...
## $ duration_ms: num [1:1209] 66.4 82.2 106.3 86.4 84.6 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [1,209 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1        : num [1:1209, 1] -1.75 -1.75 -1.75 -1.74 -1.74 ...

```

```

##  ..- attr(*, "scaled:center")= num 660
##  ..- attr(*, "scaled:scale")= num 377
## $ SPK_id      : chr [1:1209] "C083" "C083" "C083" "C083" ...
## $ SES         : num [1:1209, 1] -0.108 -0.108 -0.108 -0.108 -0.108 ...
##  ..- attr(*, "scaled:center")= num 5.14
##  ..- attr(*, "scaled:scale")= num 1.3
## $ SEX         : chr [1:1209] "F" "F" "F" "F" ...
## $ AgeInDays   : num [1:1209, 1] -1.82 -1.82 -1.82 -1.82 -1.82 ...
##  ..- attr(*, "scaled:center")= num 296
##  ..- attr(*, "scaled:scale")= num 97
## $ vowels       : chr [1:1209] "o" "a" "i" "a" ...
## $ duration_ms: num [1:1209] 4.21 4.42 4.68 4.47 4.45 ...

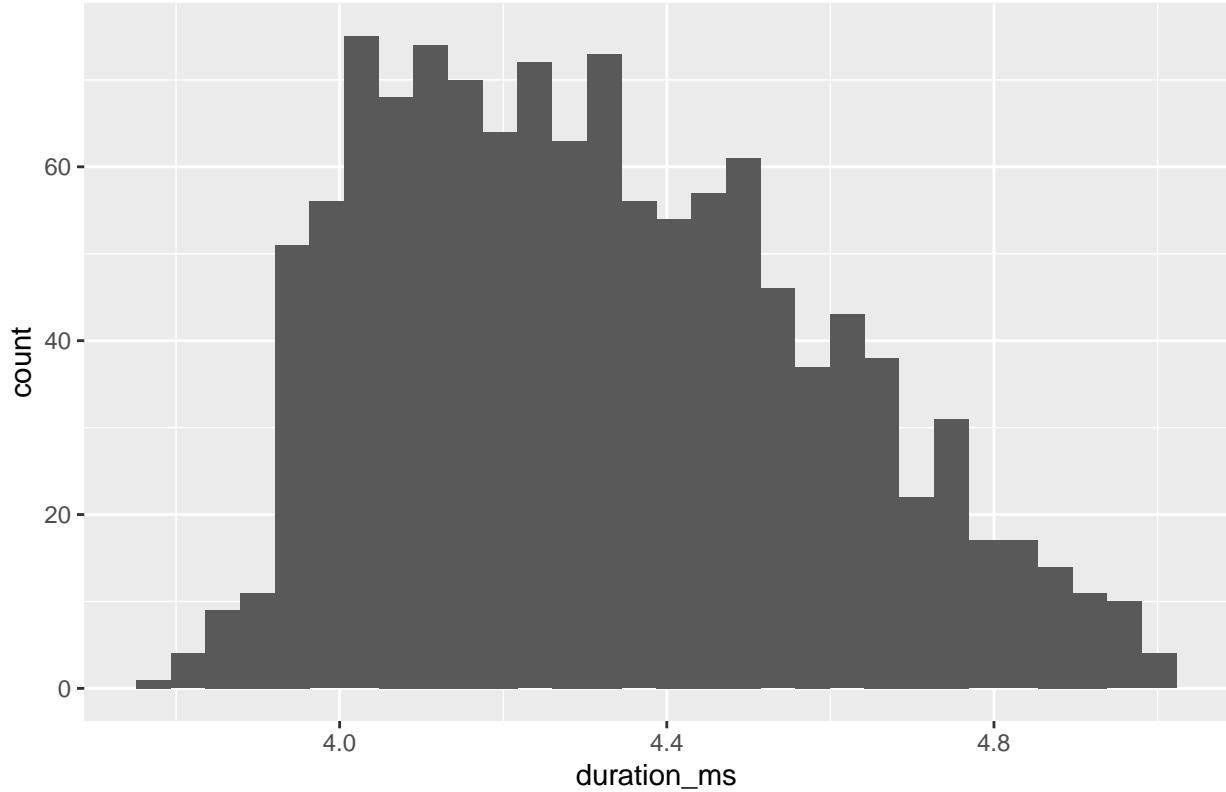
## Fitting Full Model.

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

## Warning: Model failed to converge with 1 negative eigenvalue: -2.5e-02

```

Histogram of Log-transformed Response Variable

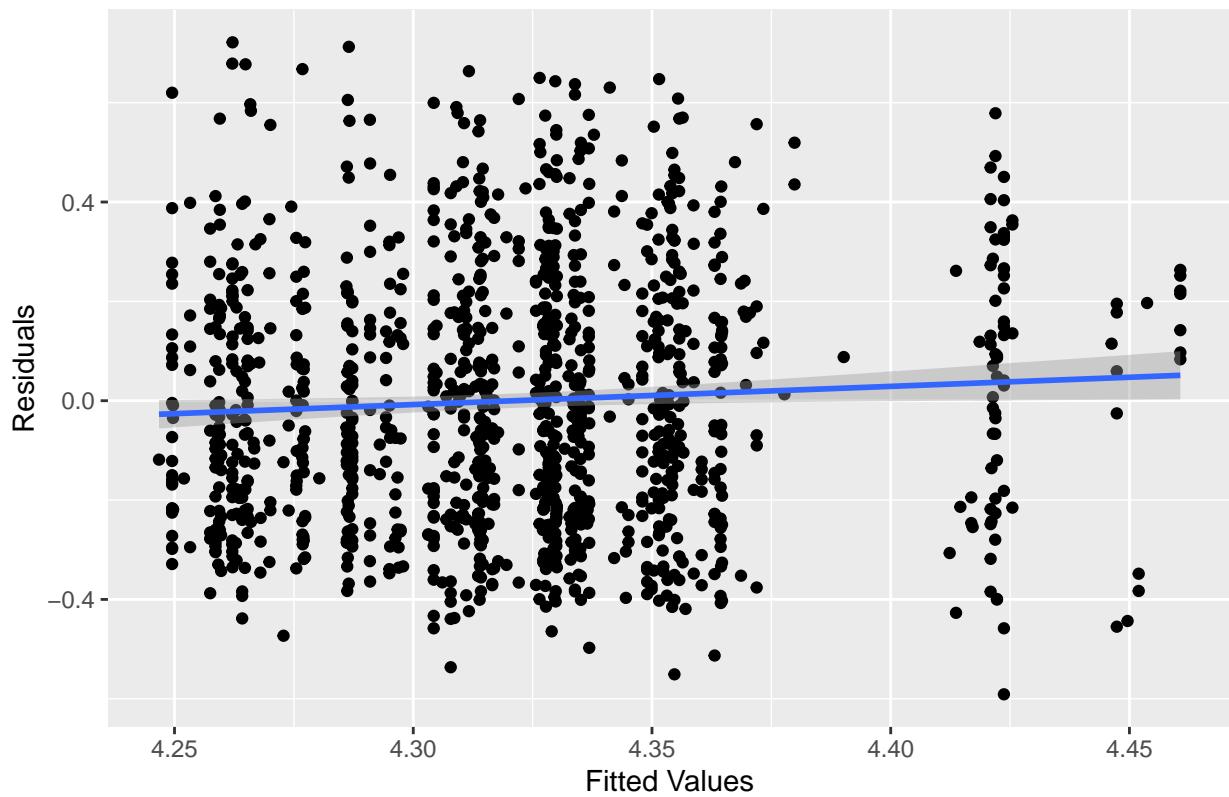


```

## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 212
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -2.07034 -0.82029 -0.09888  0.69846  2.71176 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.000e+00 0.000000 
##          AgeInDays  2.202e-03 0.046923  NaN
## vowels   (Intercept) 3.237e-06 0.001799 
##          AgeInDays  1.294e-04 0.011376  1.00
## Residual            6.707e-02 0.258987 
## Number of obs: 1209, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept)  4.32576   0.01029 238.73818 420.335 <2e-16 ***
## SES         -0.01650   0.01101 13.16666 -1.498    0.158  
## 
```

```

## SEXM          -0.02741   0.01906 164.77605 -1.438    0.152
## AgeInDays     0.00391   0.01936 10.10234  0.202    0.844
## SEXM:AgeInDays -0.04252  0.03007  6.11629 -1.414    0.206
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES      SEXM     AgInDy
## SES        -0.083
## SEXM       -0.564  0.364
## AgeInDays   0.093 -0.221 -0.113
## SEXM:AgInDy -0.050  0.126  0.125 -0.605
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -84.34 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 2.24285369630436
## Fitting Null Model.

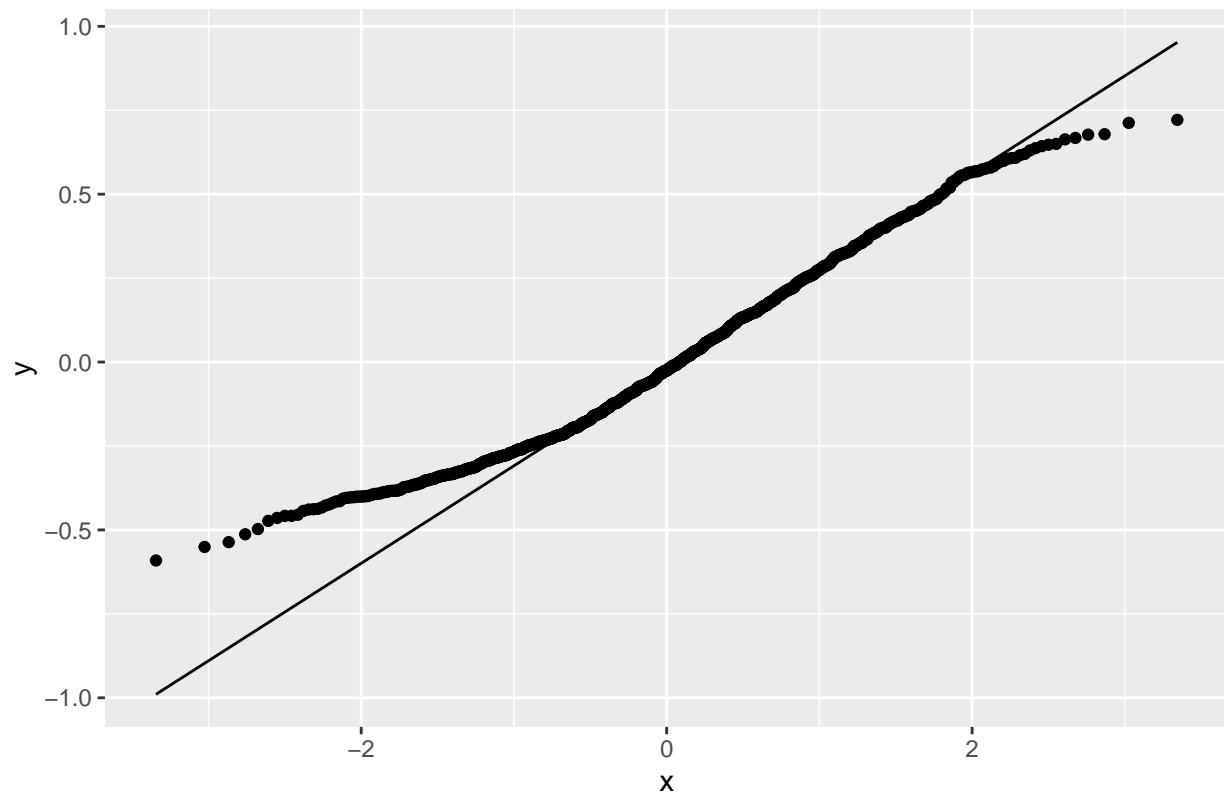
## boundary (singular) fit: see help('isSingular')

## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar     AIC     BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model   6 190.04 220.62 -89.018    178.04
## fit_model        12 192.68 253.85 -84.342    168.68 9.353  6     0.1547
## Fitting Reduced Model.

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

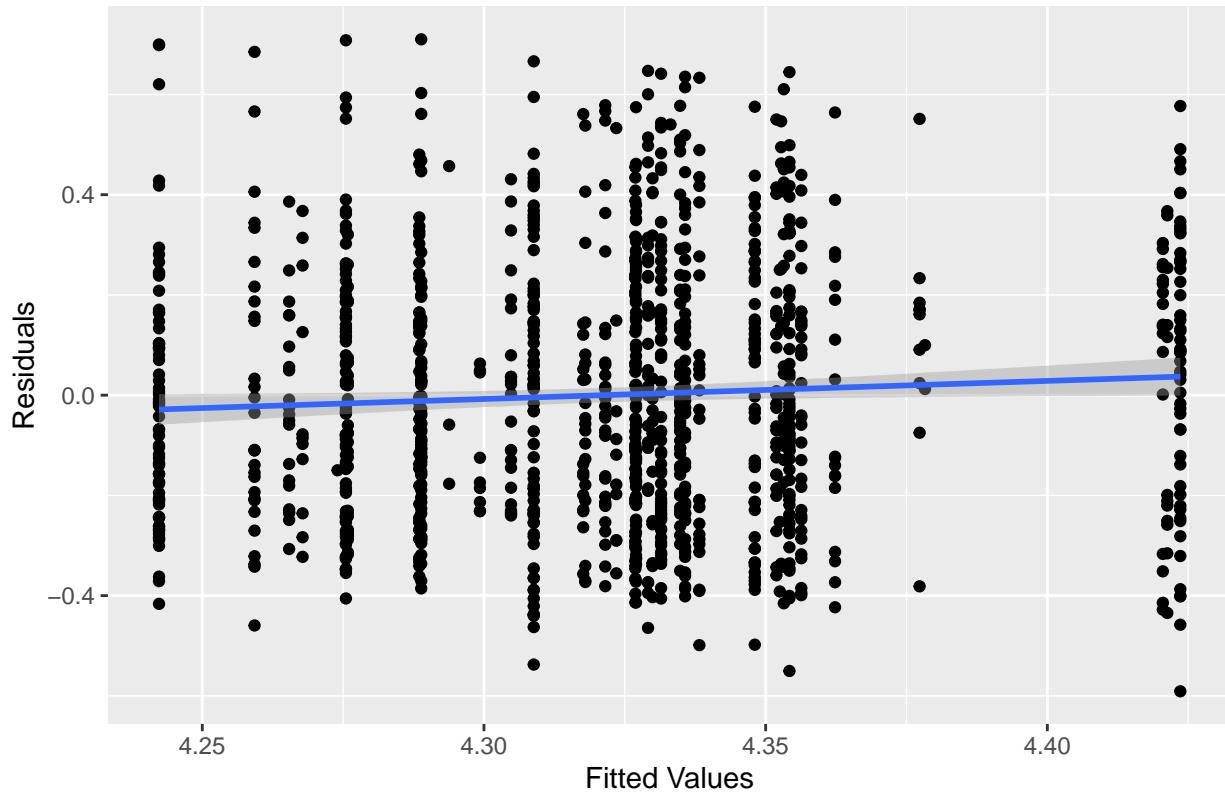
```

Q–Q Plot of Residuals for Full Model



```
## Reduced model is singular. Consider simplifying the random effects structure.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 203.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.32611 -0.82088 -0.09901  0.67456  2.76862
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.002294 0.0479
## vowels   (Intercept) 0.000000 0.0000
## Residual            0.066508 0.2579
## Number of obs: 1209, groups: SPK_id, 24; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 4.324281  0.016640 19.677584 259.880 <2e-16 ***
## SES         -0.016851  0.013831 20.446754 -1.218  0.2370
## SEXM        -0.012789  0.031046 18.380139 -0.412  0.6852
## AgeInDays    0.002971  0.012068 535.326397  0.246  0.8056

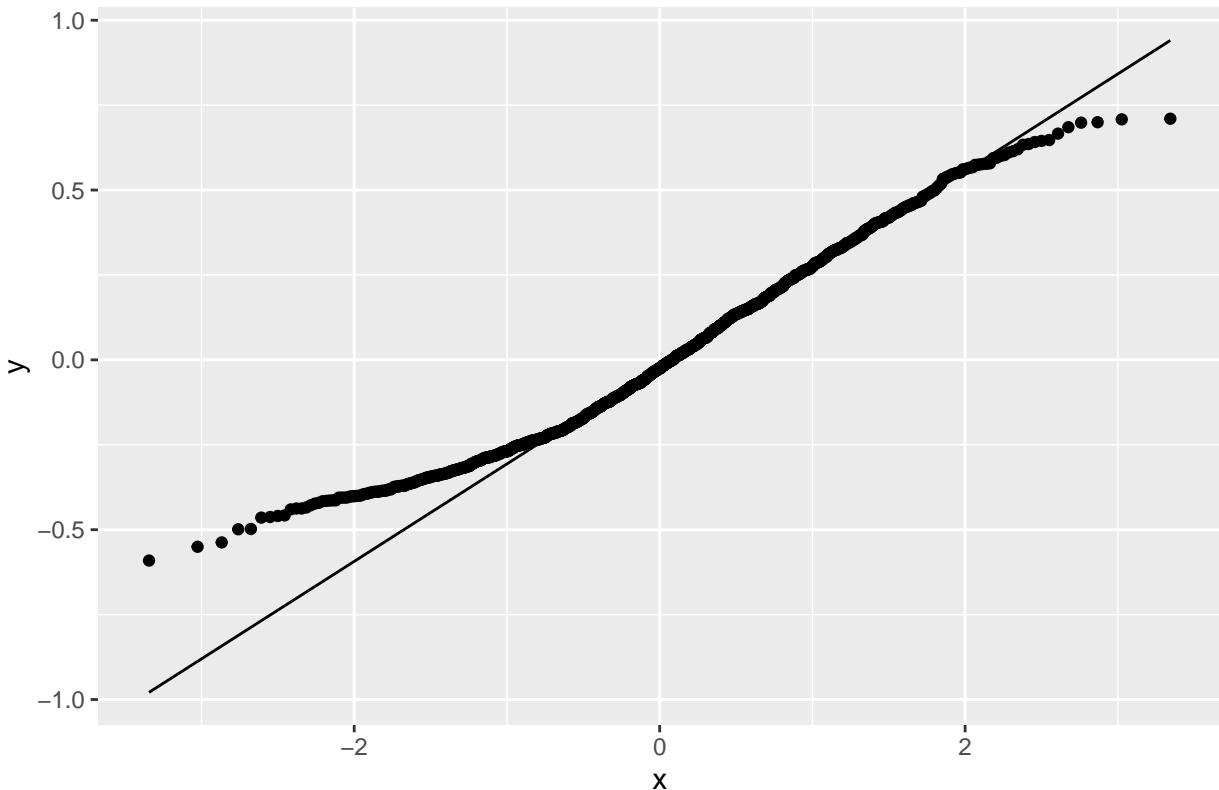
```

```

## SEXM:AgeInDays -0.037044  0.017384 437.372586 -2.131  0.0337 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.253
## SEXM     -0.605  0.409
## AgeInDays  0.145 -0.143 -0.117
## SEXM:AgInDy -0.093  0.068  0.074 -0.690
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Reduced Model: Log-Likelihood = -85.26 , Degrees of Freedom = 8
## Max VIF value of the Reduced Model: 1.95760570779353
## Fitting Null Model.
##
## boundary (singular) fit: see help('isSingular')

```

Q-Q Plot of Residuals for Reduced Model



```

## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##           npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model    6 190.04 220.62 -89.018    178.04
## red_fit_model     8 186.53 227.31 -85.264    170.53 7.5092  2   0.02341 *

```



```

## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 321 out of 1000 bootstrap iterations.

#4.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_dur_ms.xlsx"))

## New names:
## * ` ` -> `...1`

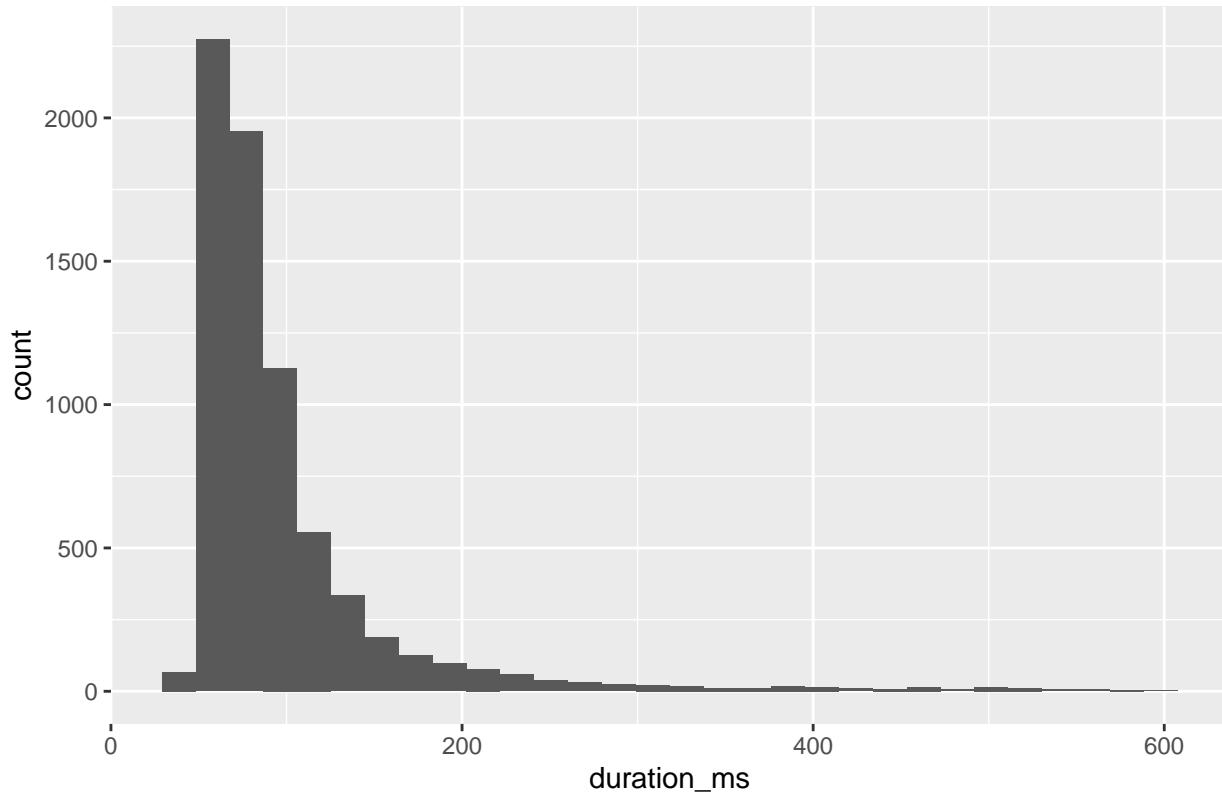
output.4.2 <- lmer.full.reduced.null.compare(duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowel|SPK_id),
                                              optimizer = "bobyqa",
                                              maxfun = 10000000000,
                                              fullm = TRUE,
                                              redm = TRUE,
                                              redmformula = duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowel|SPK_id),
                                              nullm = TRUE,
                                              nullmformula =duration_ms ~ SES + SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowel|SPK_id),
                                              df_IDS)

## Structure of combined data:
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:7127] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:7127] 5 5 5 5 5 5 5 5 5 5 ...
## $ SEX       : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays : num [1:7127] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels    : chr [1:7127] "a" "eu" "a" "a" ...
## $ duration_ms: num [1:7127] 118.9 169.1 86.9 67 91.1 ...
## Response variable is: duration_ms

## Structure of combined data after scaling:
## tibble [7,127 x 7] (S3:tbl_df/tbl/data.frame)
## $ ...1      : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3563
## ..- attr(*, "scaled:scale")= num 2058
## $ SPK_id    : chr [1:7127] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.38
## $ SEX       : chr [1:7127] "F" "F" "F" "F" ...
## $ AgeInDays : num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 244
## ..- attr(*, "scaled:scale")= num 100
## $ vowels    : chr [1:7127] "a" "eu" "a" "a" ...
## $ duration_ms: num [1:7127] 118.9 169.1 86.9 67 91.1 ...

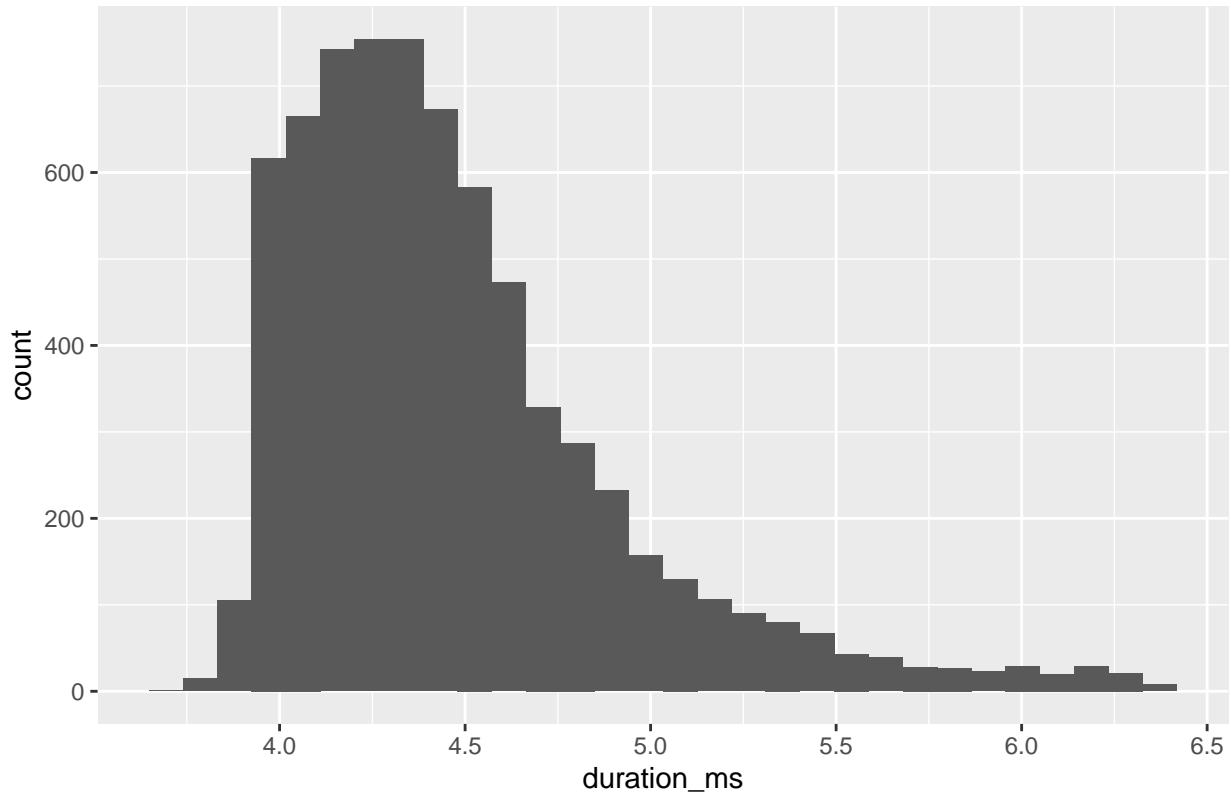
```

Histogram of Response Variable



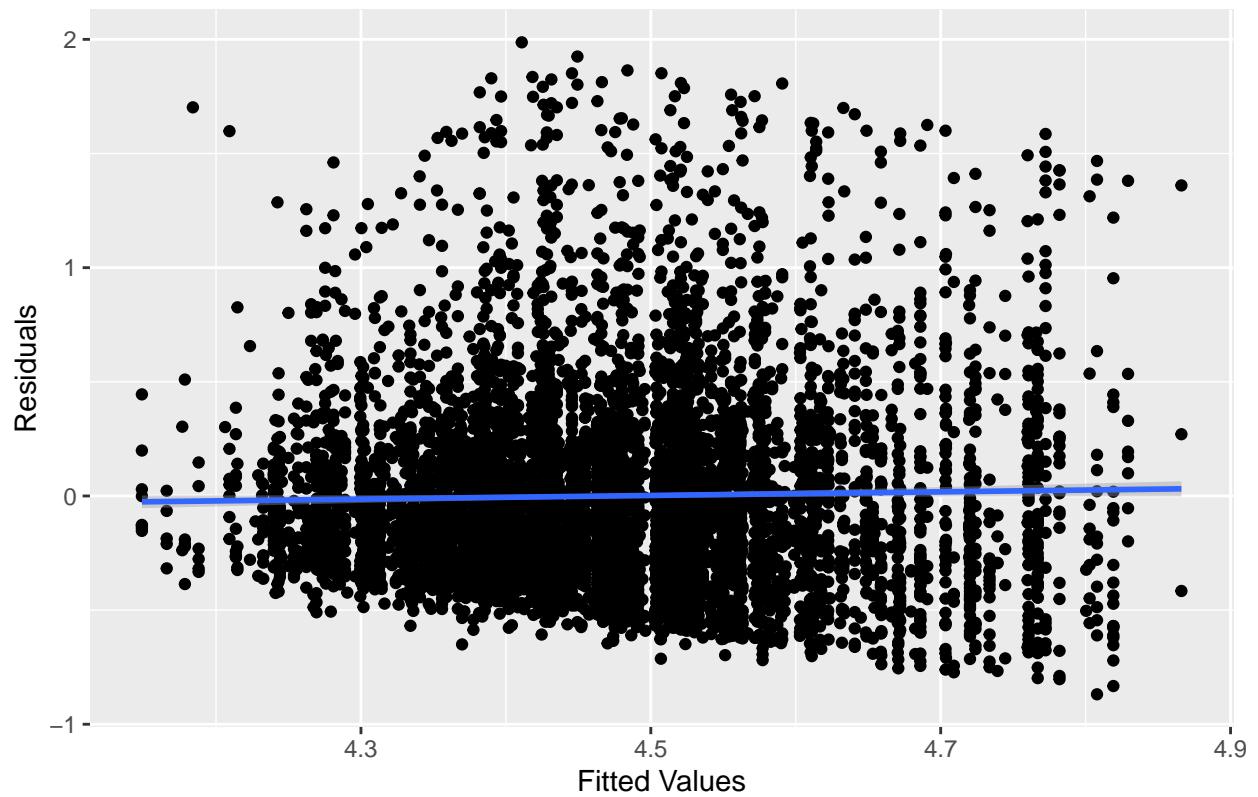
```
## Response variable is not normally distributed. Log transformation applied.  
## Structure of combined data after transformation:  
## tibble [7,127 x 7] (S3: tbl_df/tbl/data.frame)  
## $ ...1 : num [1:7127, 1] -1.73 -1.73 -1.73 -1.73 -1.73 ...  
## ..- attr(*, "scaled:center")= num 3563  
## ..- attr(*, "scaled:scale")= num 2058  
## $ SPK_id : chr [1:7127] "C083" "C083" "C083" "C083" ...  
## $ SES : num [1:7127, 1] 0 0 0 0 0 0 0 0 0 0 ...  
## ..- attr(*, "scaled:center")= num 5  
## ..- attr(*, "scaled:scale")= num 1.38  
## $ SEX : chr [1:7127] "F" "F" "F" "F" ...  
## $ AgeInDays : num [1:7127, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...  
## ..- attr(*, "scaled:center")= num 244  
## ..- attr(*, "scaled:scale")= num 100  
## $ vowels : chr [1:7127] "a" "eu" "a" "a" ...  
## $ duration_ms: num [1:7127] 4.79 5.14 4.48 4.22 4.52 ...
```

Histogram of Log-transformed Response Variable

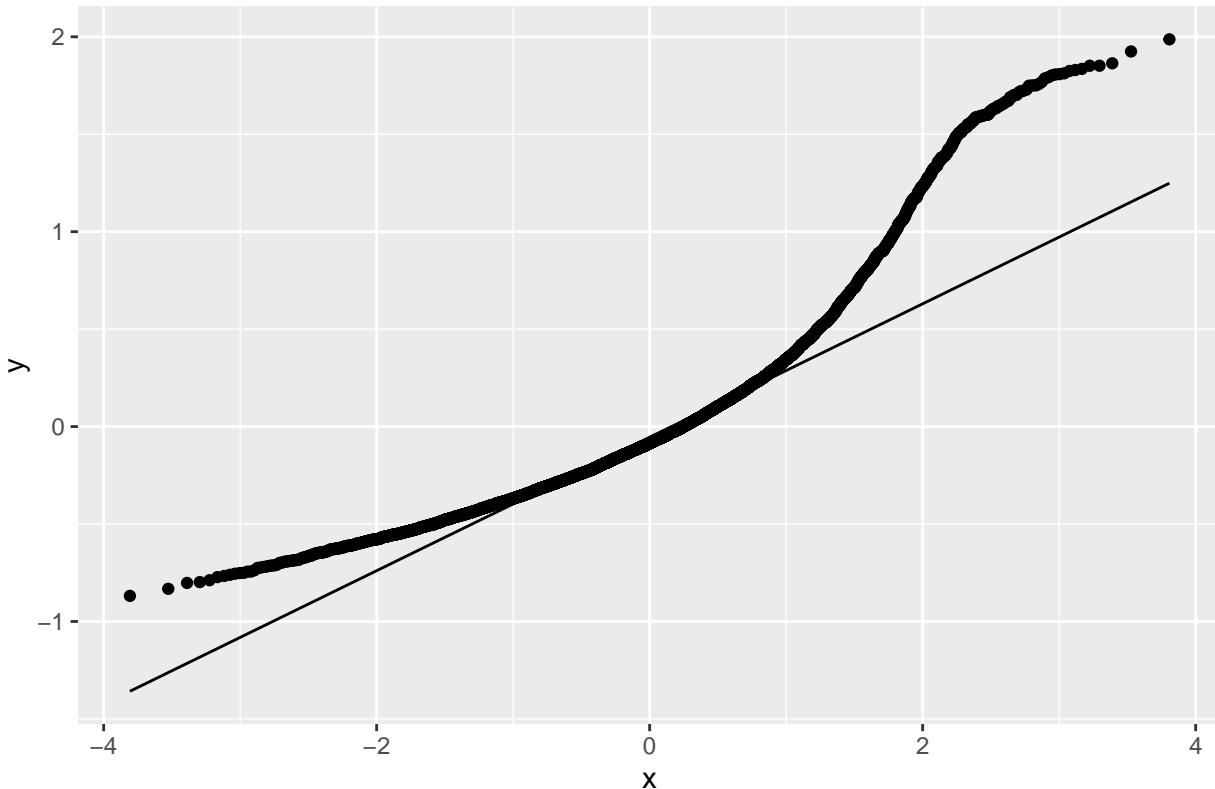


```
## Fitting Full Model.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



Q-Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 8206.2
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -2.0477 -0.6710 -0.2016  0.4157  4.6659 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.016562 0.12869
##          AgeInDays   0.002256 0.04750 -0.24
## vowels   (Intercept) 0.006311 0.07944
##          AgeInDays   0.000148 0.01216 -0.65
## Residual            0.181142 0.42561
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept)  4.47761   0.04247 26.81675 105.427 < 2e-16 ***
## SES         -0.01920   0.03017 23.04816  -0.637  0.53066

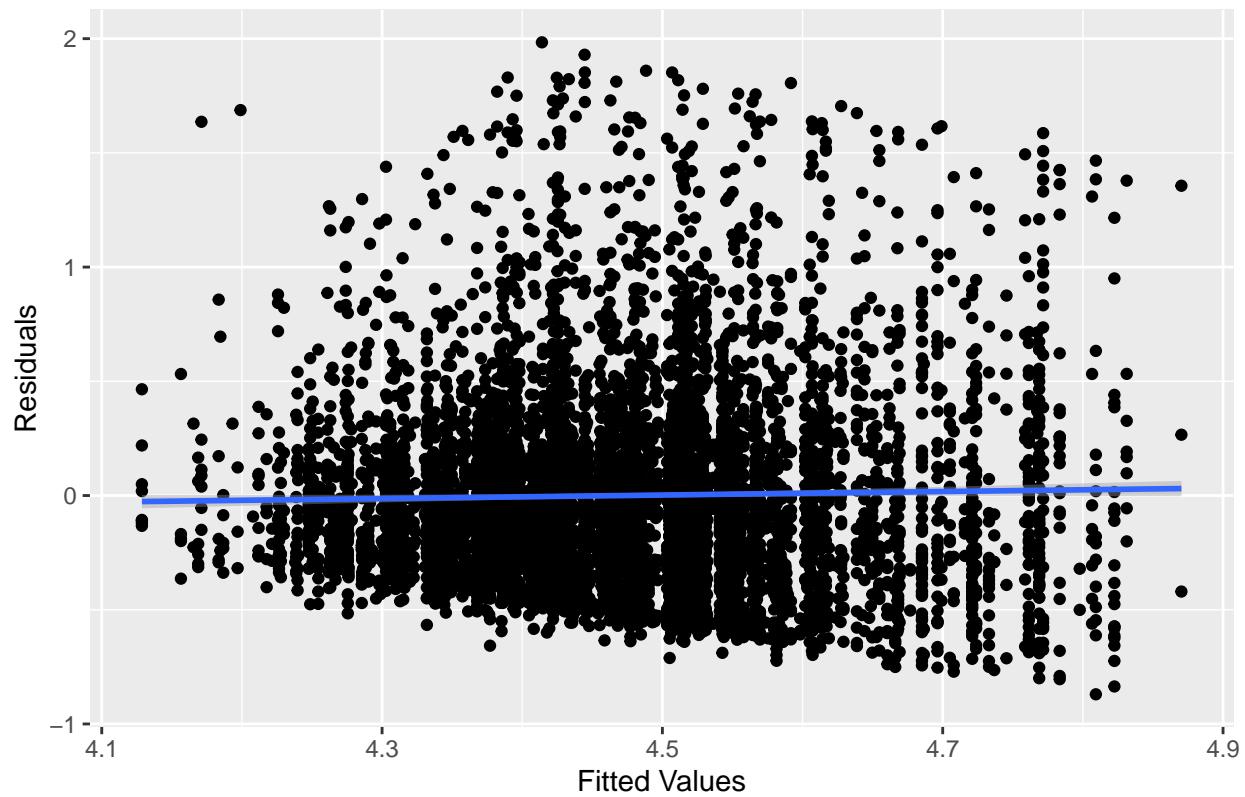
```

```

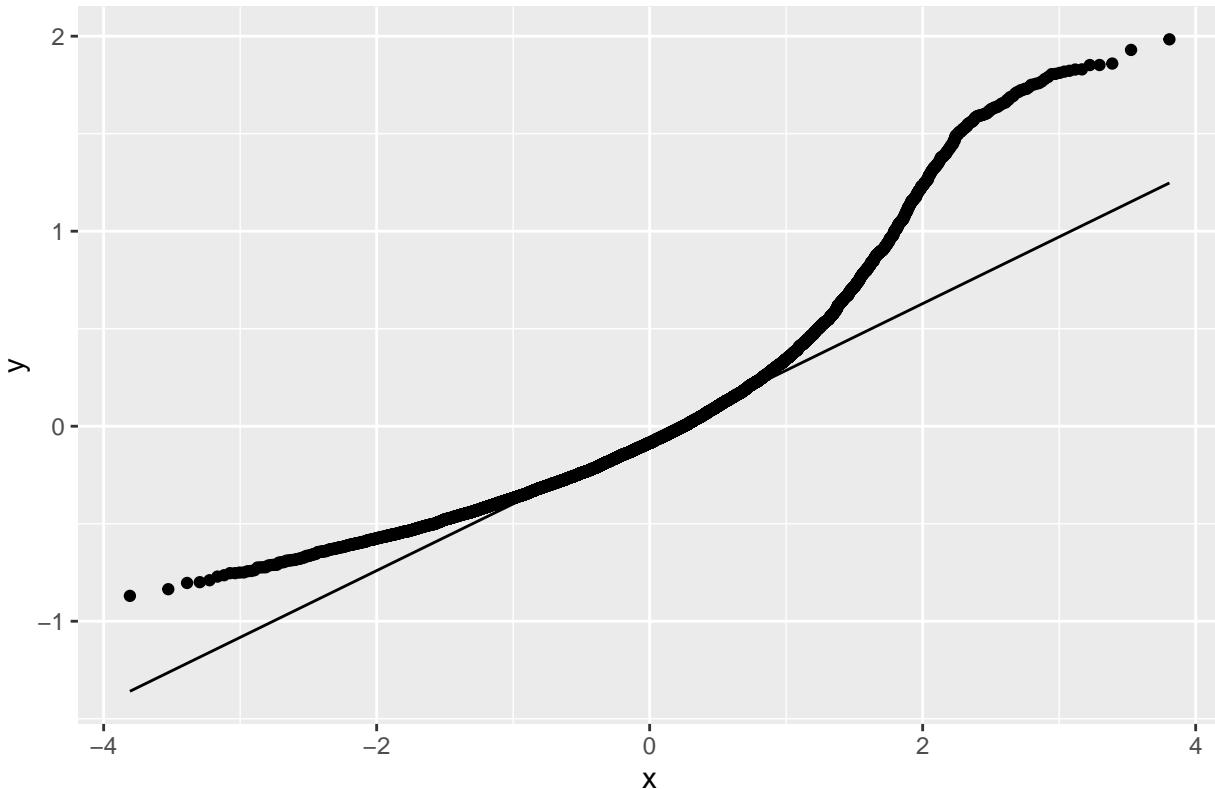
## SEXM          -0.02539   0.05926 23.92894 -0.428  0.67212
## AgeInDays    -0.05060   0.01596 13.41745 -3.170  0.00713 **
## SEXM:AgeInDays 0.00211   0.02677 13.62783  0.079  0.93832
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM    AgInDy
## SES        -0.218
## SEXM       -0.480  0.374
## AgeInDays  -0.239  0.006  0.098
## SEXM:AgInDy 0.076  0.012 -0.136 -0.550
## Full Model: Log-Likelihood = -4089.86 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.45653891343266
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar      AIC      BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model 10 8207.3 8276.0 -4093.7    8187.3
## fit_model      12 8203.7 8286.2 -4089.9    8179.7 7.5749  2  0.02265 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 8208.4
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -2.0502 -0.6723 -0.2012  0.4129  4.6574 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.016574 0.12874
##          AgeInDays   0.002252 0.04746 -0.24
## vowels   (Intercept) 0.006307 0.07942
## Residual           0.181257 0.42574
## Number of obs: 7127, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  4.477964  0.042473 26.833023 105.432 < 2e-16 ***
## SES         -0.019105  0.030169 23.054269  -0.633  0.53280  
## SEXM        -0.025029  0.059279 23.943072  -0.422  0.67663  
## 
```

```

## AgeInDays      -0.053822   0.015328 12.048209  -3.511  0.00427 ***
## SEXM:AgeInDays  0.002142   0.026751 13.594014    0.080  0.93733
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES       -0.218
## SEXM      -0.480  0.374
## AgeInDays  -0.141  0.006  0.102
## SEXM:AgInDy  0.077  0.012 -0.137 -0.573
## Reduced Model: Log-Likelihood = -4090.92 , Degrees of Freedom = 10
## Max VIF value of the Reduced Model: 1.50887691192787
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## red_fit_model: redmformula
## null_fit_model: nullmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## red_fit_model   10 8201.8 8270.6 -4090.9    8181.8
## null_fit_model   10 8207.3 8276.0 -4093.7    8187.3    0 0
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_dur_ms.xlsx"))

## New names:
## * `` -> `...`1` 

# Calculate the quartiles
Q1 <- quantile(df_IDS$duration_ms, 0.25)
Q3 <- quantile(df_IDS$duration_ms, 0.75)
# Calculate the interquartile range (IQR)
IQR <- Q3 - Q1
# Define the lower and upper bounds for non-outliers
lower <- Q1 - 1.5 * IQR
upper <- Q3 + 1.5 * IQR

# Remove outliers and create a new dataframe
df_no_outliers_IDS <- df_IDS[df_IDS$duration_ms >= lower & df_IDS$duration_ms <= upper, ]

output.4.2_no_outliers <- lmer.full.reduced.null.compare(duration_ms ~ SES + SEX + AgeInDays + AgeInDay
  optimizer = "bobyqa",
  maxfun = 1000000000,
  fullm = TRUE,
  redm = TRUE,
  redmformula = duration_ms ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id),
  nullm = TRUE,
  nullmformula =duration_ms ~ SES + SEX + (1+AgeInDays|SPK_id) + (1|vowels),
  df_no_outliers_IDS)

## Structure of combined data:
## tibble [6,509 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:6509] 0 2 3 4 5 6 7 9 10 11 ...
## $ SPK_id    : chr [1:6509] "C083" "C083" "C083" "C083" ...

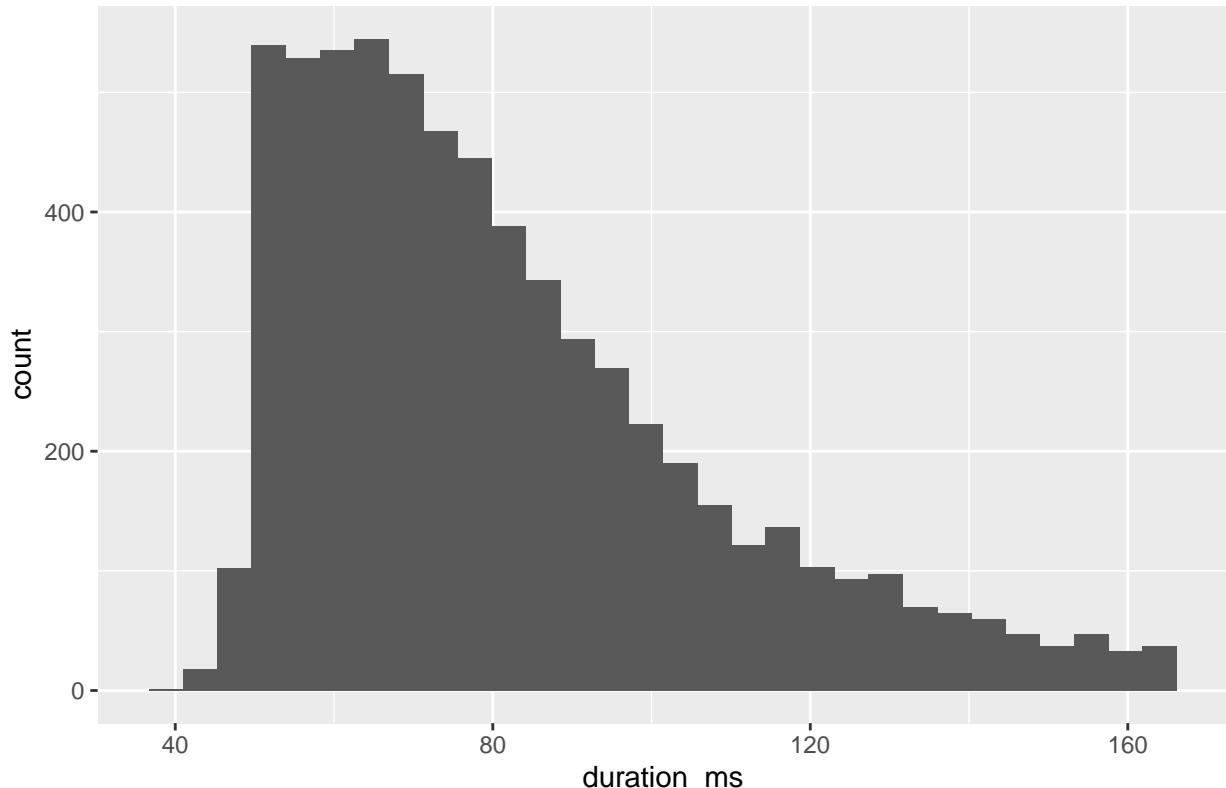
```

```

## $ SES      : num [1:6509] 5 5 5 5 5 5 5 5 5 ...
## $ SEX      : chr [1:6509] "F" "F" "F" "F" ...
## $ AgeInDays : num [1:6509] 120 120 120 120 120 120 120 120 120 120 ...
## $ vowels   : chr [1:6509] "a" "a" "a" "e" ...
## $ duration_ms: num [1:6509] 118.9 86.9 67 91.1 83 ...
## Response variable is: duration_ms
## Structure of combined data after scaling:
## tibble [6,509 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:6509, 1] -1.74 -1.74 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3576
## ..- attr(*, "scaled:scale")= num 2060
## $ SPK_id    : chr [1:6509] "C083" "C083" "C083" "C083" ...
## $ SES       : num [1:6509, 1] -0.000554 -0.000554 -0.000554 -0.000554 -0.000554 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.39
## $ SEX       : chr [1:6509] "F" "F" "F" "F" ...
## $ AgeInDays : num [1:6509, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
## ..- attr(*, "scaled:center")= num 246
## ..- attr(*, "scaled:scale")= num 101
## $ vowels   : chr [1:6509] "a" "a" "a" "e" ...
## $ duration_ms: num [1:6509] 118.9 86.9 67 91.1 83 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [6,509 x 7] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:6509, 1] -1.74 -1.74 -1.73 -1.73 -1.73 ...
## ..- attr(*, "scaled:center")= num 3576

```

```

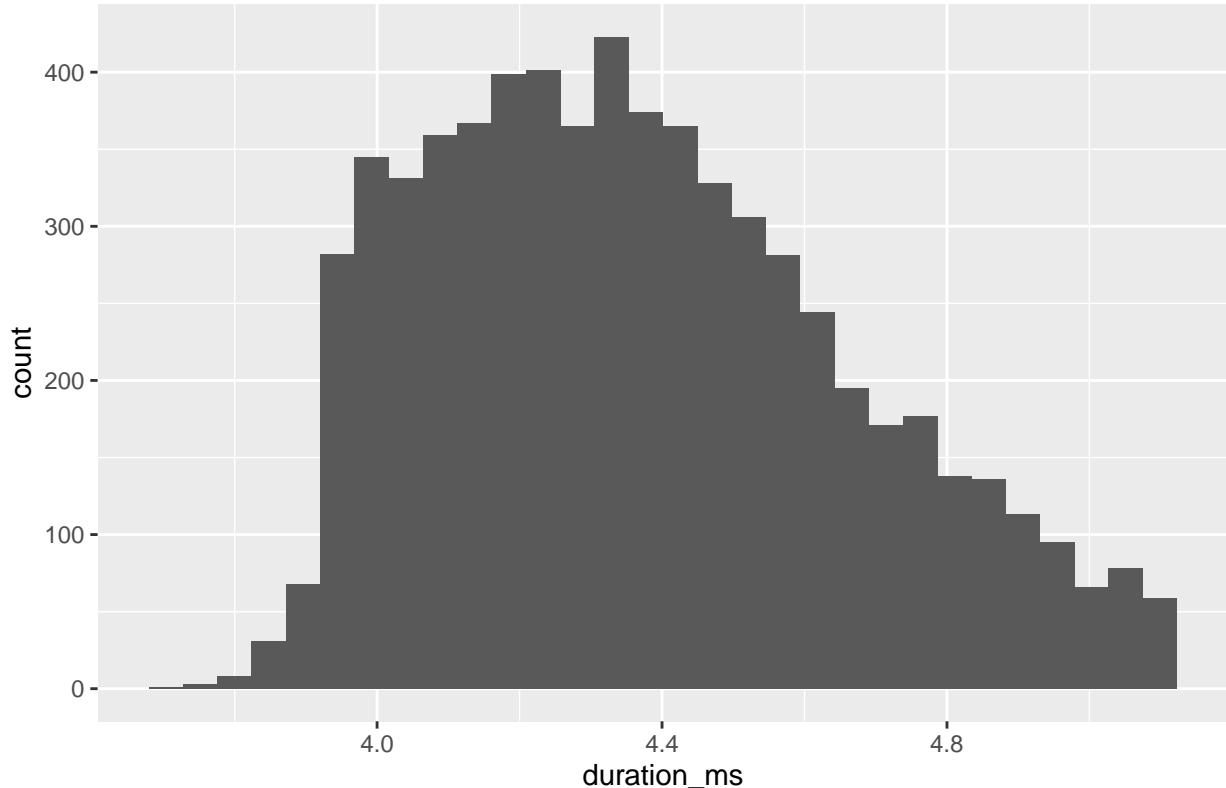
##  ..- attr(*, "scaled:center")= num 5
##  ..- attr(*, "scaled:scale")= num 1.39
##  $ SEX      : chr [1:6509] "F" "F" "F" "F" ...
##  $ AgeInDays : num [1:6509, 1] -1.24 -1.24 -1.24 -1.24 -1.24 ...
##  ..- attr(*, "scaled:center")= num 246
##  ..- attr(*, "scaled:scale")= num 101
##  $ vowels   : chr [1:6509] "a" "a" "a" "e" ...
##  $ duration_ms: num [1:6509] 4.79 4.48 4.22 4.52 4.43 ...

## Fitting Full Model.

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

```

Histogram of Log-transformed Response Variable

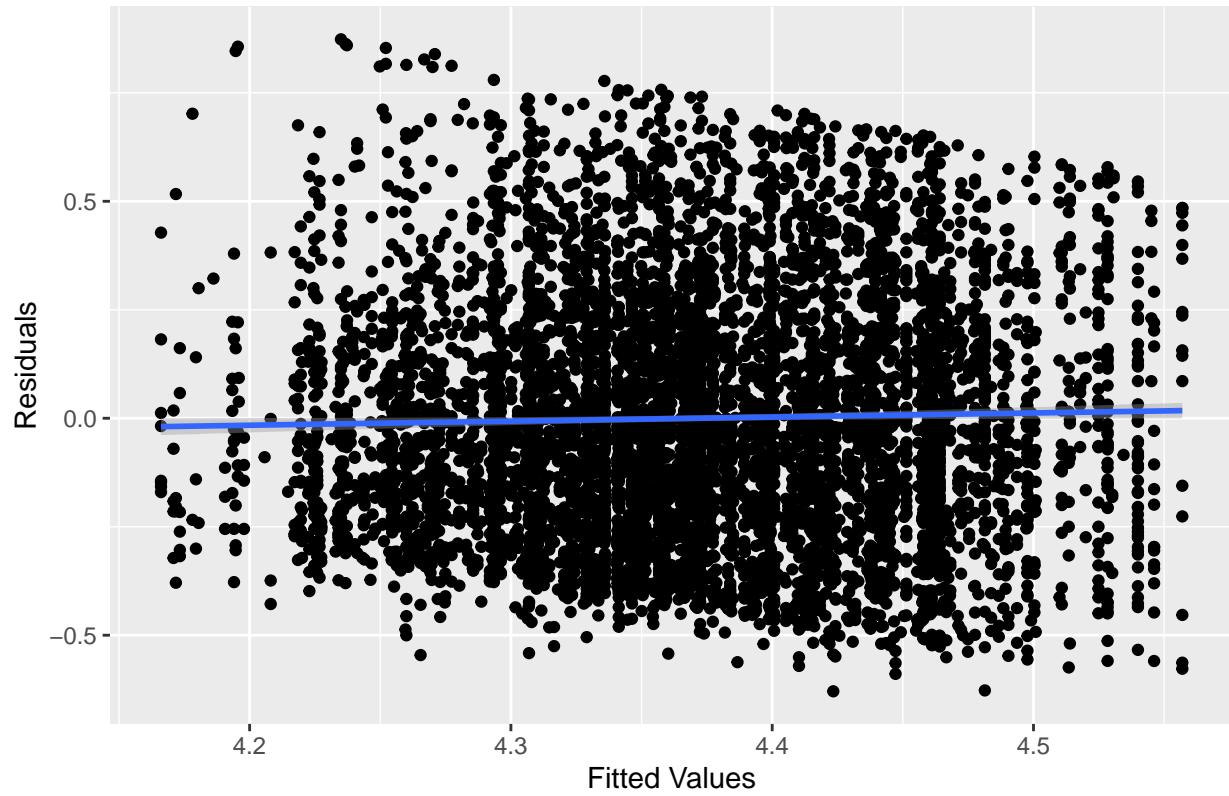


```

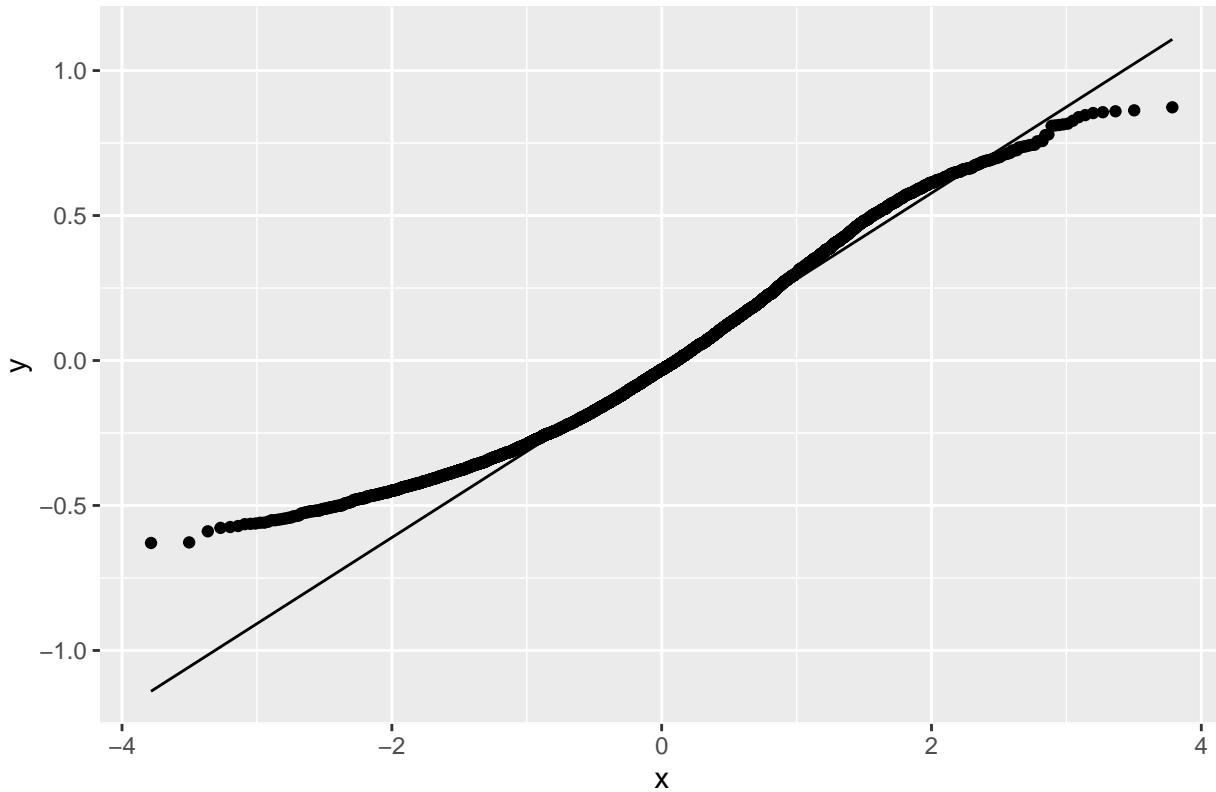
## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 2100.8
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -2.2467 -0.7721 -0.1142  0.6534  3.1090 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 5.670e-03 0.075297 
##          AgeInDays   1.583e-03 0.039786 -0.03
## vowels   (Intercept) 2.365e-03 0.048631 
##          AgeInDays   3.098e-05 0.005566 -1.00
## Residual            7.895e-02 0.280979 
## Number of obs: 6509, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept)  4.374192  0.025712 25.359757 170.125 <2e-16 ***
## SES         -0.010914  0.018230 20.268731  -0.599  0.5560

```

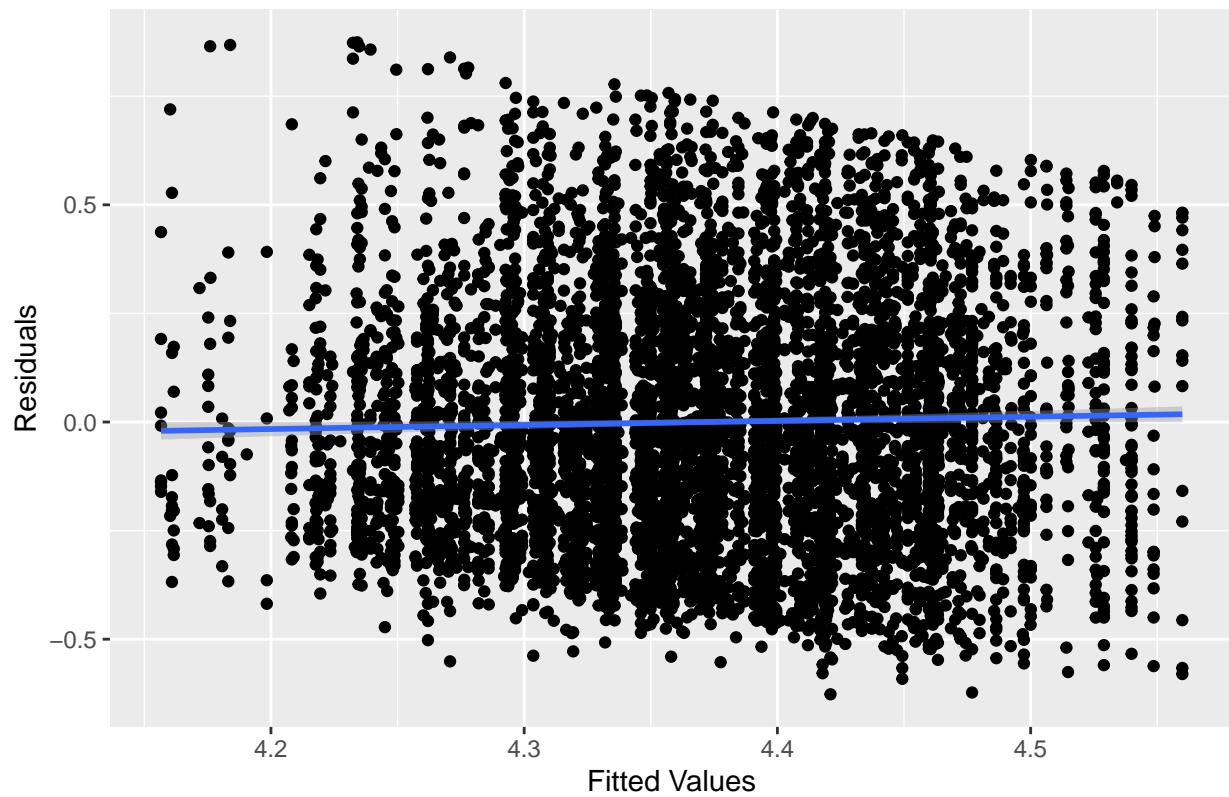
```

## SEXM          -0.031198  0.035746 22.971442 -0.873   0.3918
## AgeInDays    -0.030721  0.012387 10.963557 -2.480   0.0306 *
## SEXM:AgeInDays 0.003265  0.021253 11.336630  0.154   0.8806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM   AgInDy
## SES        -0.230
## SEXM       -0.479  0.391
## AgeInDays  -0.122  0.025  0.024
## SEXM:AgInDy 0.014 -0.011  0.006 -0.569
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -1035.18 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.47533199130487
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model  8 2093.6 2147.9 -1038.8    2077.6
## fit_model      12 2094.4 2175.7 -1035.2    2070.4 7.257  4     0.1229
## Fitting Reduced Model.

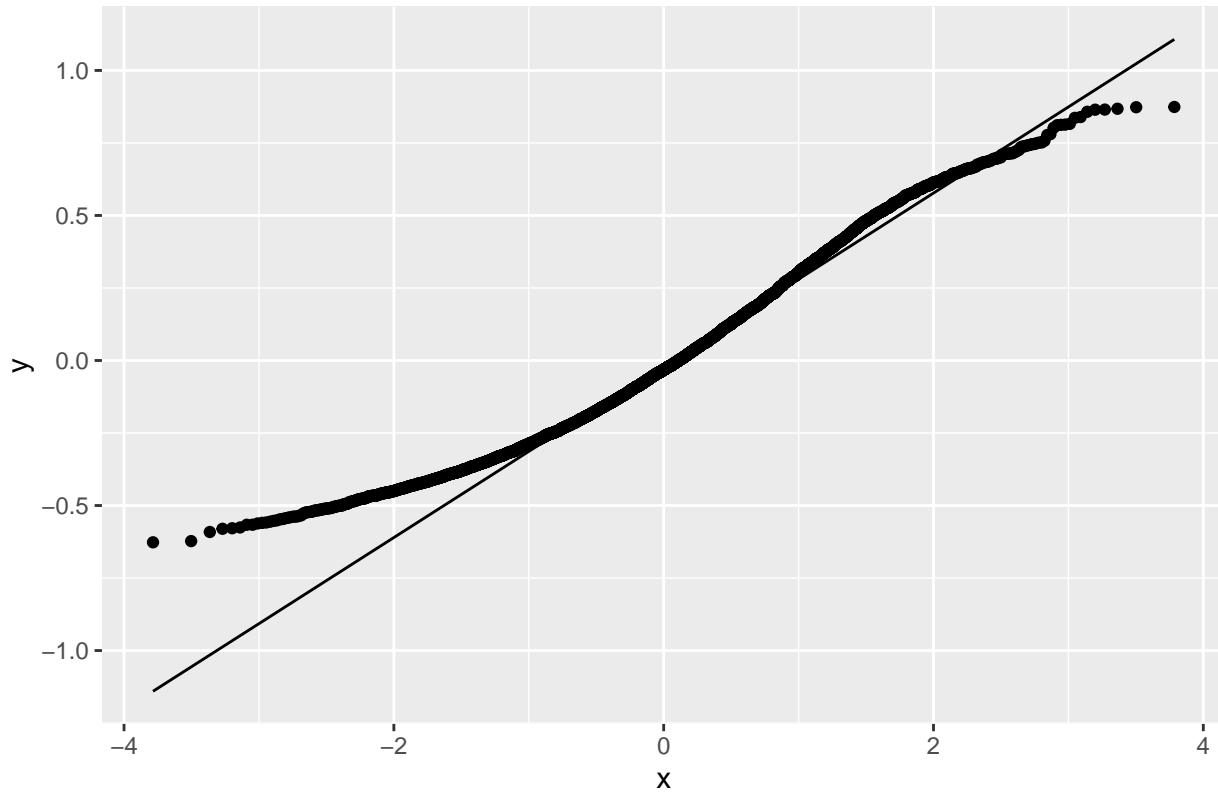
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 2102.4
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -2.2373 -0.7717 -0.1151  0.6541  3.1118 
## 
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.005663 0.07525 
##          AgeInDays   0.001572 0.03964 -0.03 
## vowels   (Intercept) 0.002400 0.04899 
## Residual            0.078970 0.28102 
## Number of obs: 6509, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept)  4.374449  0.025778 25.267359 169.697 <2e-16 ***
## SES         -0.010858  0.018218 20.261707  -0.596  0.5578    
## SEXM        -0.030995  0.035724 22.983347  -0.868  0.3946    
## 
```



```

## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 83 out of 1000 bootstrap iterations.

pred.data.4.2_no_outliers <- boot.ci.predict.lmer(
  m = output.4.2_no_outliers$reduced_model,
  optimizer = output.4.2_no_outliers$optimizer,
  maxfun = output.4.2_no_outliers$maxfun,
  data = output.4.2$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.4.2_no_outliers$fit_data$AgeInDays), max(output.4.2_no_outliers$fit_data$AgeInDays)),
  reqcol = c("AgeInDays"),
  centercol = c("SES", "SEX"),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## Number of singular fits during bootstrapping: 0 out of 1000 bootstrap iterations.

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.4.1$fit_data$AgeInDays, "scaled:scale") + attr(output.4.1$fit_data$AgeInDays, "scaled:center")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.4.2$fit_data$AgeInDays, "scaled:scale") + attr(output.4.2$fit_data$AgeInDays, "scaled:center")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

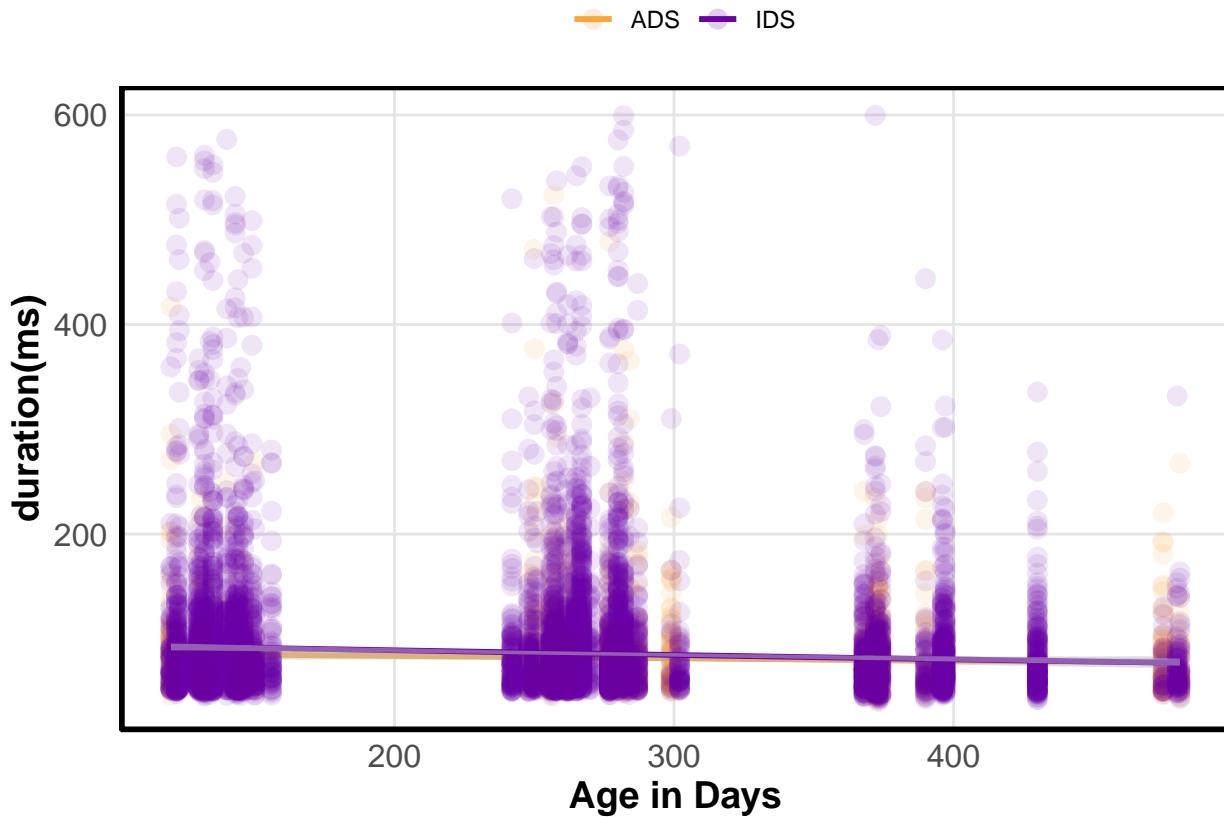
duration_ms <- continuous_fit_ci_plot(
  plot.data = list(output.4.1$fit_data, output.4.2$fit_data),
  coefs = list(pred.data.4.1,pred.data.4.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "duration_ms",
  x.labs = "Age in Days",
  y.labs = "duration(ms)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = TRUE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),

```

```

#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(40, 60, by = 10),
legend.labels = c("ADS", "IDS")
)
print(duration_ms)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.4.1$fit_data$AgeInDays, "scaled:scale") + attr(output.4.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.4.2$fit_data$AgeInDays, "scaled:scale") + attr(output.4.2$fit_data$AgeInDays, "scaled")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

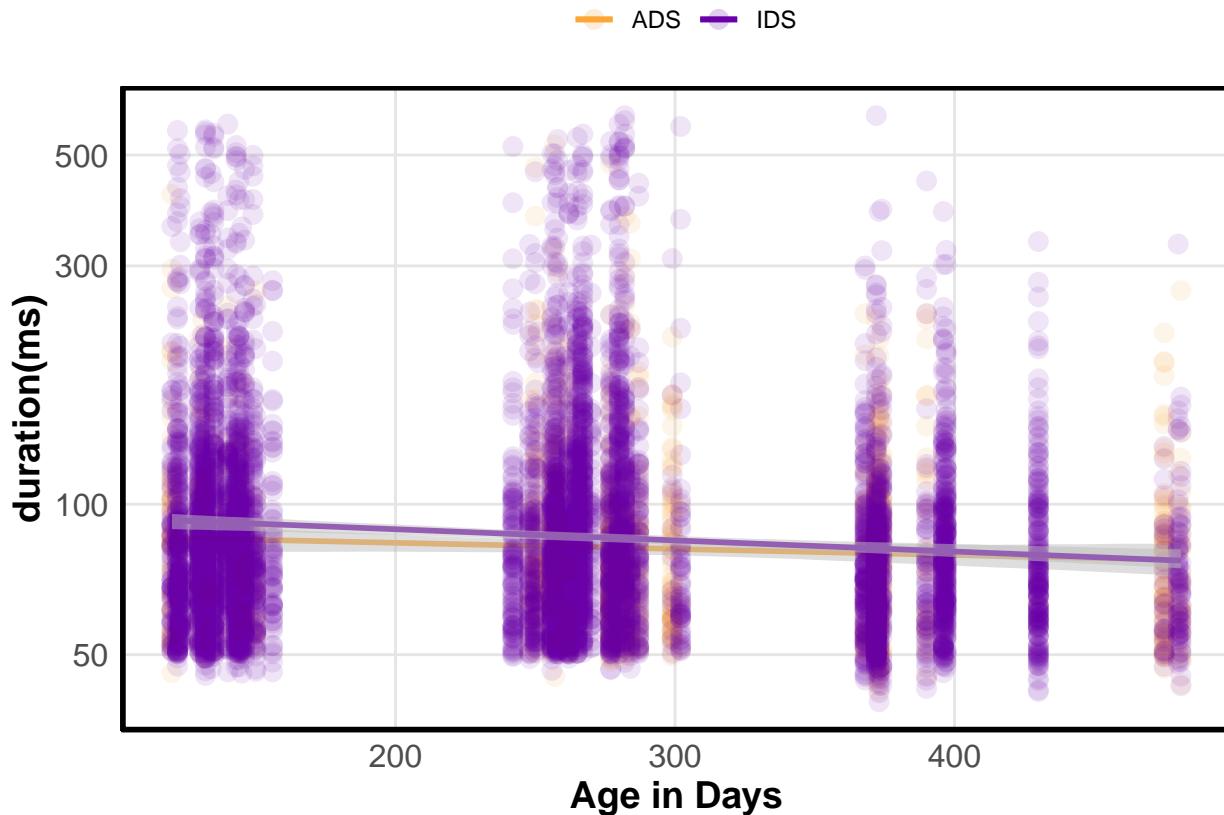
duration_ms_log_grid <- continuous_fit_ci_plot(
  plot.data = list(output.4.1$fit_data, output.4.2$fit_data),
  coefs = list(pred.data.4.1,pred.data.4.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "duration_ms",
  x.labs = "Age in Days",
  y.labs = "duration(ms)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),

```

```

y.ax.transformation = TRUE,
inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
y_grid_log = TRUE,
#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(40, 60, by = 10),
legend.labels = c("ADS", "IDS")
)
print(duration_ms_log_grid)

```



#5. Vowel space Area #5.1 Corner vowels #5.1.1 ADS

```
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_area_cor.xlsx"))
```

```

## New names:
## * ` ` -> `...` 

output.5.1.1 <- lmer.full.reduced.null.compare(ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX +
                                               optimizer = "bobyqa",
                                               maxfun = 10000000000,
                                               fullm = TRUE,
                                               redm = TRUE,
                                               redmformula = ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                               nullm = TRUE,
                                               nullmformula =ConvexHullArea ~ SES + SEX + (1|SPK_id),
                                               df_ADS)

## Structure of combined data:
## tibble [20 x 6] (S3: tbl_df/tbl/data.frame)

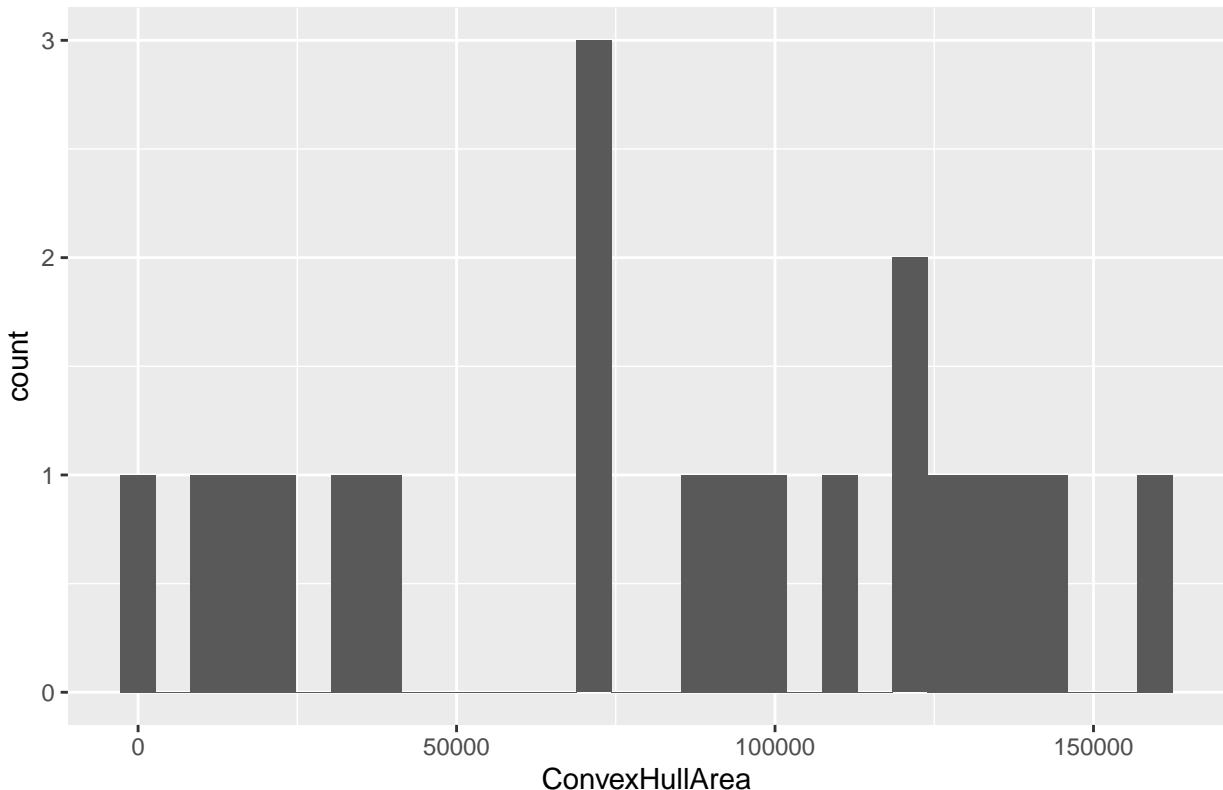
```

```

## $ ...1 : num [1:20] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:20] "C023" "C028" "C031" "C044" ...
## $ AgeInDays : num [1:20] 251 287 373 250 390 131 481 277 258 256 ...
## $ SES : num [1:20] 5 5 8 5 5 4 3 5 2 5 ...
## $ SEX : chr [1:20] "M" "F" "F" "F" ...
## $ ConvexHullArea: num [1:20] 101217 123873 112454 38742 34293 ...
## Response variable is: ConvexHullArea
## Structure of combined data after scaling:
## tibble [20 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:20, 1] -1.61 -1.44 -1.27 -1.1 -0.93 ...
## ..- attr(*, "scaled:center")= num 9.5
## ..- attr(*, "scaled:scale")= num 5.92
## $ SPK_id : chr [1:20] "C023" "C028" "C031" "C044" ...
## $ AgeInDays : num [1:20, 1] -0.575 -0.197 0.708 -0.586 0.887 ...
## ..- attr(*, "scaled:center")= num 306
## ..- attr(*, "scaled:scale")= num 95.1
## $ SES : num [1:20, 1] -0.0674 -0.0674 1.9552 -0.0674 -0.0674 ...
## ..- attr(*, "scaled:center")= num 5.1
## ..- attr(*, "scaled:scale")= num 1.48
## $ SEX : chr [1:20] "M" "F" "F" "F" ...
## $ ConvexHullArea: num [1:20] 101217 123873 112454 38742 34293 ...

```

Histogram of Response Variable

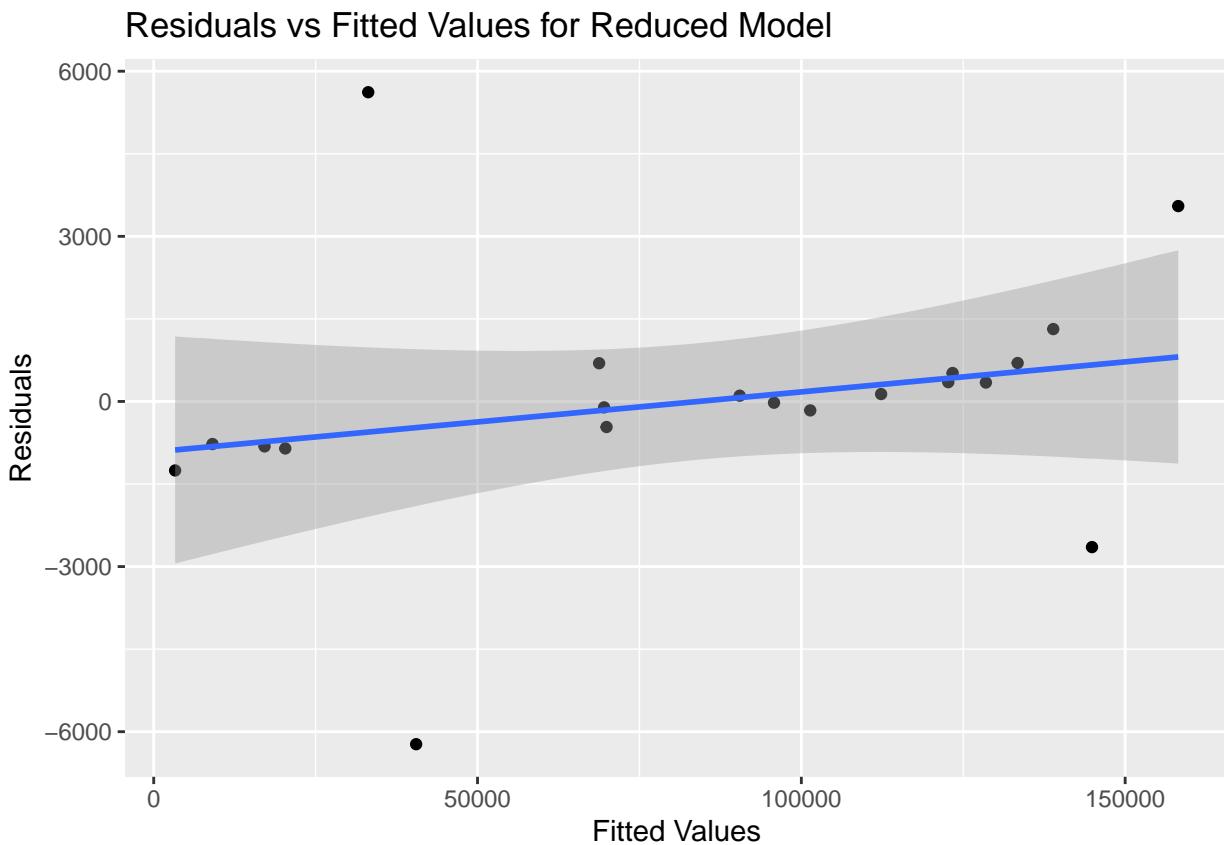


```

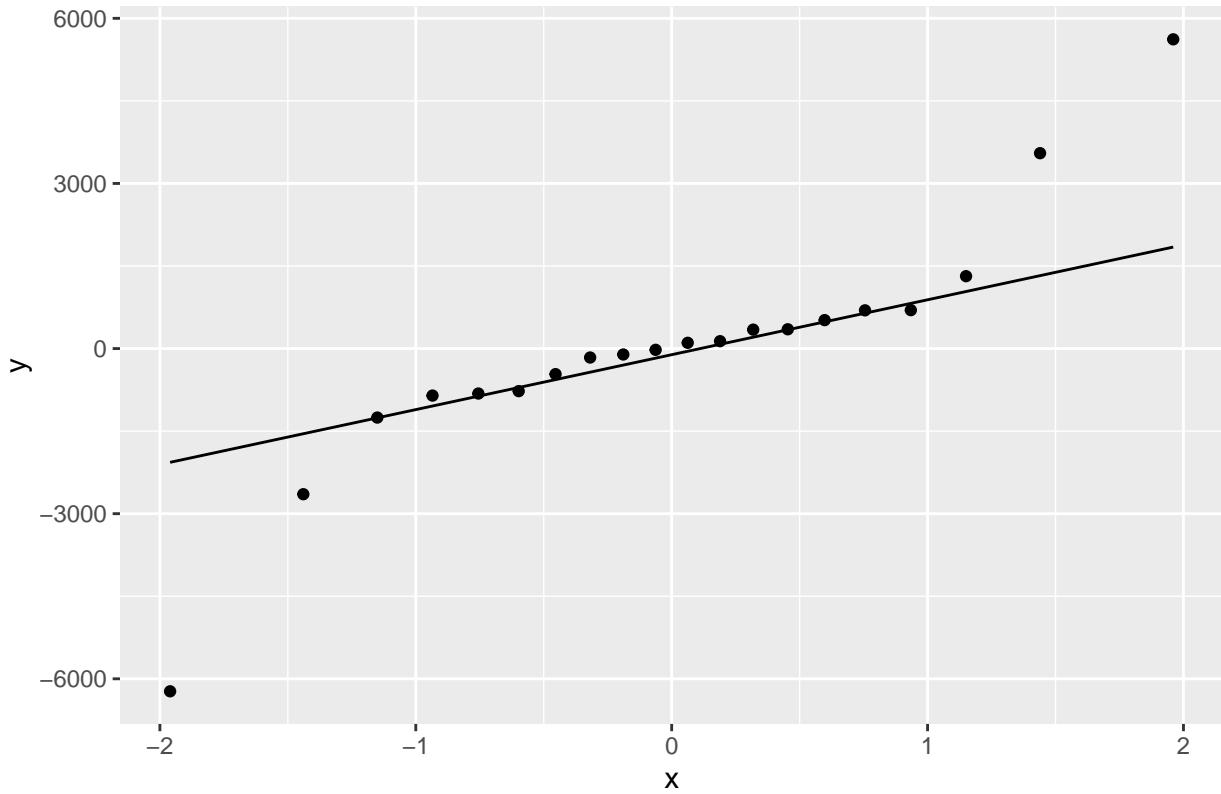
## Response variable is normally distributed; no transformation applied.
## Fitting Full Model.
## Error in fitting the full model: number of observations (=20) <= number of random effects (=34) for ...
## Full model failed. Consider using a different optimizer or increasing maxfun.
## Fitting Null Model.

```

```
## Fitting Reduced Model.  
## `geom_smooth()` using formula = 'y ~ x'
```



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 377.6
##
## Scaled residuals:
##      Min       1Q     Median       3Q      Max
## -0.70890 -0.18621  0.01232  0.14434  0.54813
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 2.829e+09 53193
## Residual           9.234e+07  9609
## Number of obs: 20, groups: SPK_id, 17
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)
## (Intercept) 84327.725 17064.380 13.890 4.942 0.000222 ***
## SES          7115.913 13913.834 14.005 0.511 0.617020
## SEXM         13460.915 30727.000 13.926 0.438 0.668046
## AgeInDays    5180.906  4409.580   1.008 1.175 0.447780
## SEXM:AgeInDays -32385.177 7249.268   1.030 -4.467 0.134708

```

```

## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.297
## SEXM     -0.640  0.449
## AgeInDays  0.068 -0.076 -0.059
## SEXM:AgInDy -0.044  0.056 -0.004 -0.609
## Reduced Model: Log-Likelihood = -238.88 , Degrees of Freedom = 7
## Max VIF value of the Reduced Model: 1.58640044807658
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## red_fit_model: redmformula
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model 5 495.33 500.31 -242.66    485.33
## red_fit_model  7 491.76 498.73 -238.88    477.76 7.5665  2    0.02275 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).

pred.data.5.1.1 <- NULL
for (i in seq(1, 1000, by = 100)) {
  batch_result <- boot.ci.predict.lmer(
    m = output.5.1.1$reduced_model,
    optimizer = output.5.1.1$optimizer,
    maxfun = output.5.1.1$maxfun,
    data = output.5.1.1$fit_data,
    pred.data = expand.grid(AgeInDays = seq(min(output.5.1.1$fit_data$AgeInDays), max(output.5.1.1$fit_data$AgeInDays)),
    reqcol = c("AgeInDays"),
    centercol = c("SES", "SEX"),
    nboots = min(100, 1000 - (i - 1)),
    link = "identity",
    keep.boots = FALSE
  )
  pred.data.5.1.1 <- rbind(pred.data.5.1.1, batch_result)
}

## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.

#5.1.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_area_cor.xlsx"))

```

```

## New names:
## * `` -> `...` 

output.5.1.2 <- lmer.full.reduced.null.compare(ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX +
                                               optimizer = "bobyqa",
                                               maxfun = 1000000000,
                                               fullm = TRUE,
                                               redm = TRUE,
                                               redmformula = ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                               nullm = TRUE,
                                               nullmformula =ConvexHullArea ~ SES + SEX + (1+AgeInDays|SPK_id),
                                               df_IDS)

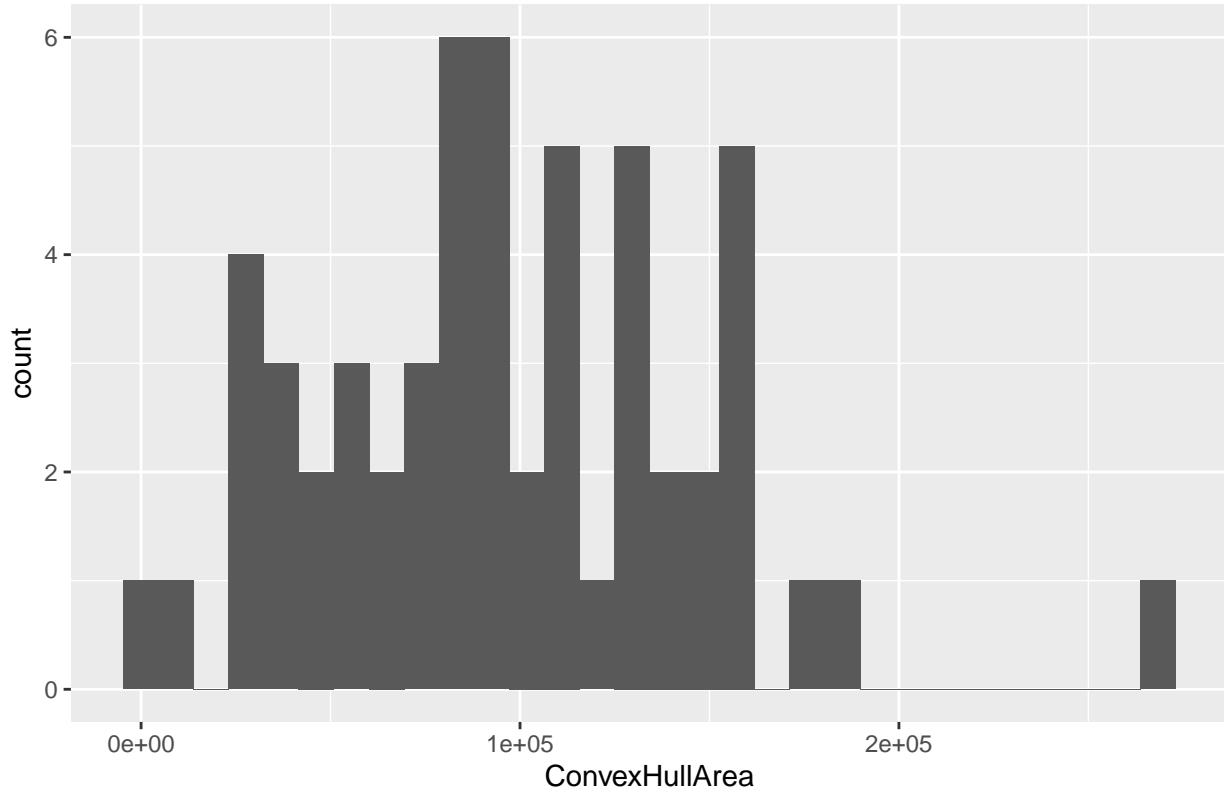
## Structure of combined data:
## tibble [56 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:56] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id     : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays  : num [1:56] 280 267 135 251 140 287 135 373 143 282 ...
## $ SES        : num [1:56] 5 5 5 5 5 8 8 5 5 ...
## $ SEX        : chr [1:56] "M" "F" "M" "M" ...
## $ ConvexHullArea: num [1:56] 93430 113635 90749 123468 135382 ...
## Response variable is: ConvexHullArea
## Structure of combined data after scaling:
## tibble [56 x 6] (S3:tbl_df/tbl/data.frame)
## $ ...1      : num [1:56, 1] -1.69 -1.62 -1.56 -1.5 -1.44 ...
## ..- attr(*, "scaled:center")= num 27.5
## ..- attr(*, "scaled:scale")= num 16.3
## $ SPK_id     : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays  : num [1:56, 1] 0.2314 0.1093 -1.1303 -0.0409 -1.0834 ...
## ..- attr(*, "scaled:center")= num 255
## ..- attr(*, "scaled:scale")= num 106
## $ SES        : num [1:56, 1] -0.0814 -0.0814 -0.0814 -0.0814 -0.0814 ...
## ..- attr(*, "scaled:center")= num 5.11
## ..- attr(*, "scaled:scale")= num 1.32
## $ SEX        : chr [1:56] "M" "F" "M" "M" ...
## $ ConvexHullArea: num [1:56] 93430 113635 90749 123468 135382 ...

## Response variable is normally distributed; no transformation applied.
## Fitting Full Model.

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

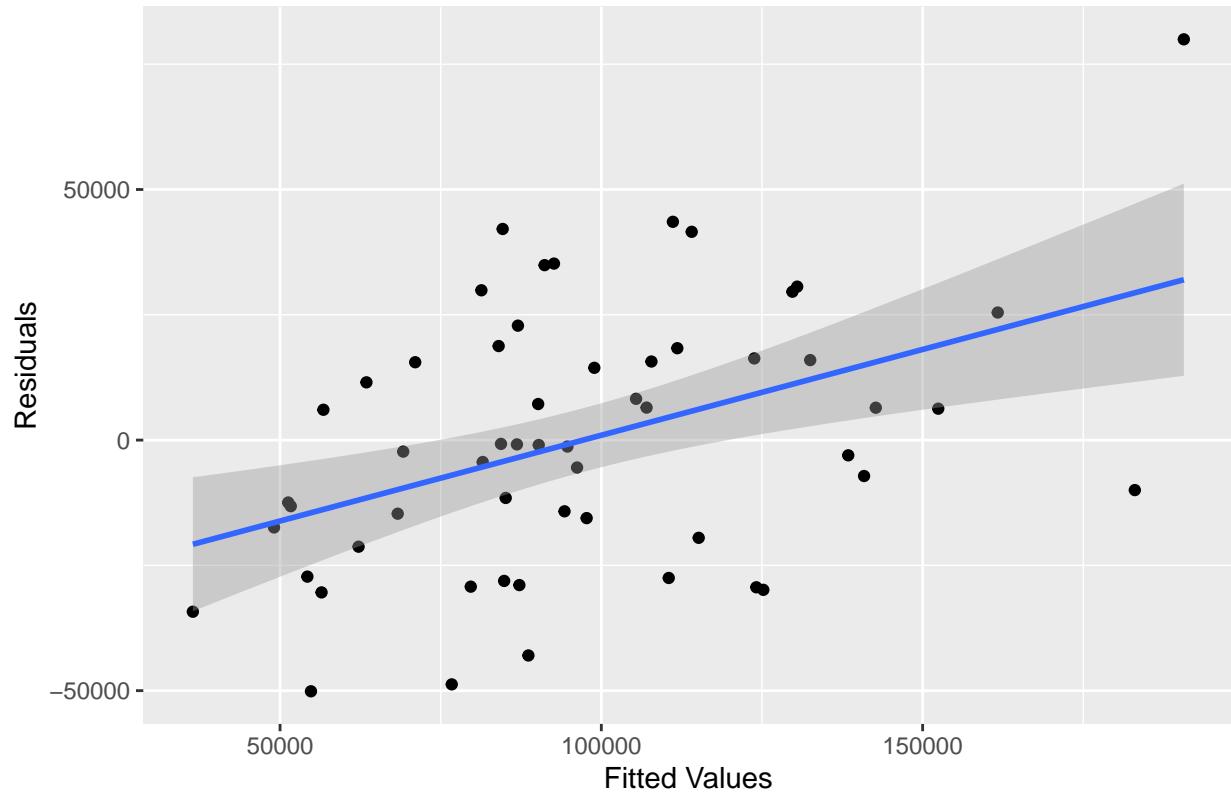
```

Histogram of Response Variable



```
## Full model is singular. Consider simplifying the random effects structure.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1257.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.49876 -0.52266 -0.03429  0.47340  2.33716
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 1.787e+09 42269
##          AgeInDays    1.203e+07 3469     -1.00
## Residual            1.108e+09 33288
## Number of obs: 56, groups: SPK_id, 27
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 93506.84  11734.96  22.33  7.968 5.67e-08 ***
## SES         -1007.68  10550.23  22.12 -0.096   0.925
## SEXM        197.92   21483.75  23.05  0.009   0.993
## AgeInDays   -3202.57   6101.96  29.56 -0.525   0.604

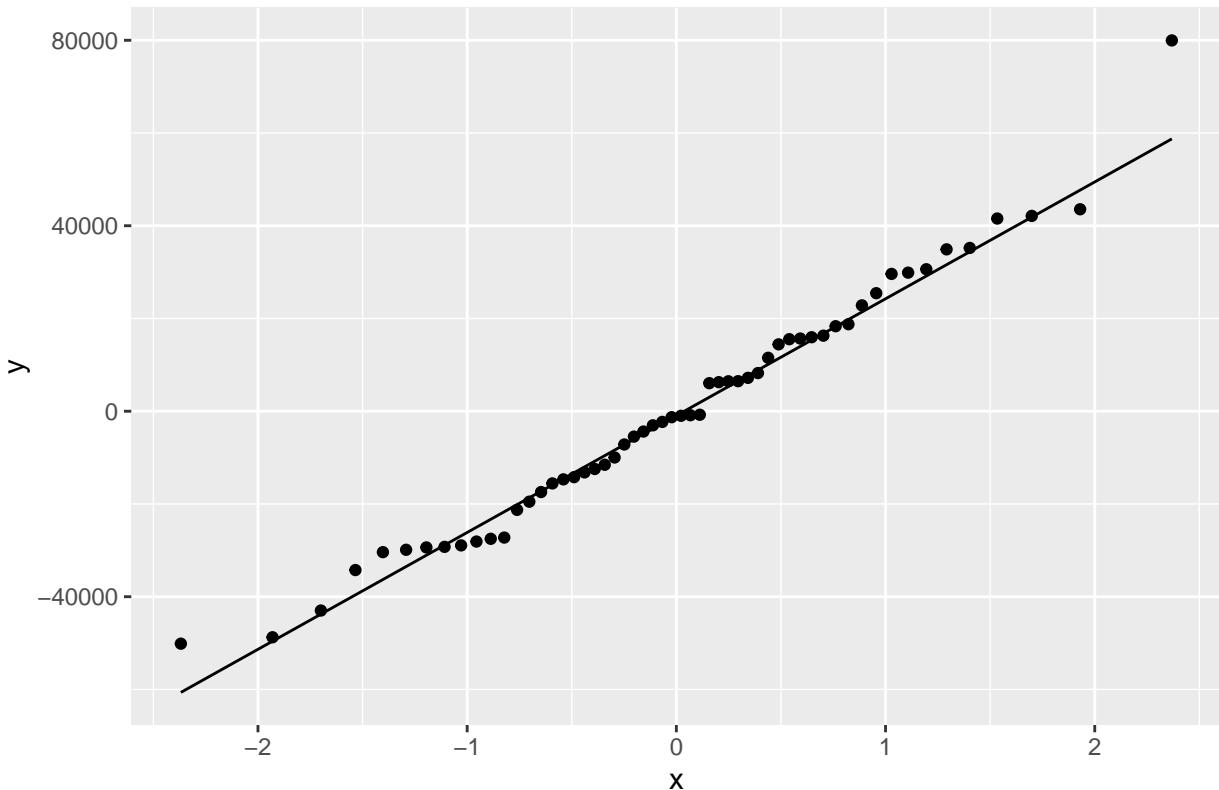
```

```

## SEXM:AgeInDays 15062.39    10096.53     30.49    1.492     0.146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES      SEXM   AgInDy
## SES       -0.208
## SEXM      -0.600   0.372
## AgeInDays -0.126   0.027   0.076
## SEXM:AgInDy  0.076  -0.014  -0.102  -0.604
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -679.37 , Degrees of Freedom = 9
## Max VIF value of the Full Model: 1.58583321165876
## Fitting Null Model.
##
## boundary (singular) fit: see help('isSingular')

```

Q–Q Plot of Residuals for Full Model



```

## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model    7 1375.0 1389.2 -680.52    1361.0
## fit_model         9 1376.7 1395.0 -679.37    1358.7 2.3156  2     0.3142

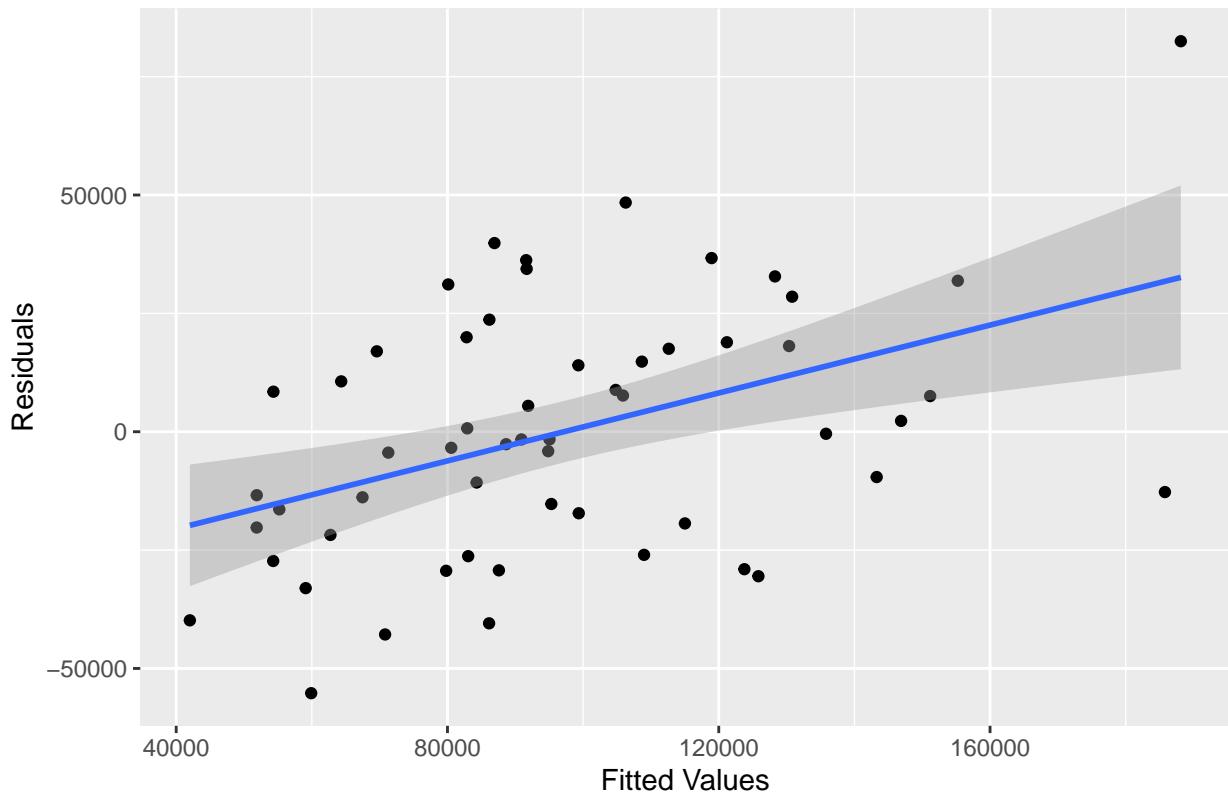
```

```

## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 1258
##
## Scaled residuals:
##      Min     1Q   Median     3Q    Max 
## -1.62044 -0.57101 -0.04773  0.51202  2.37384 
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 1.734e+09 41635 
## Residual           1.141e+09 33786 
## Number of obs: 56, groups:  SPK_id, 27
##
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept) 92790.97  11662.53  22.30  7.956 5.86e-08 ***
## SES         -19.99   10430.53  21.37 -0.002   0.998  

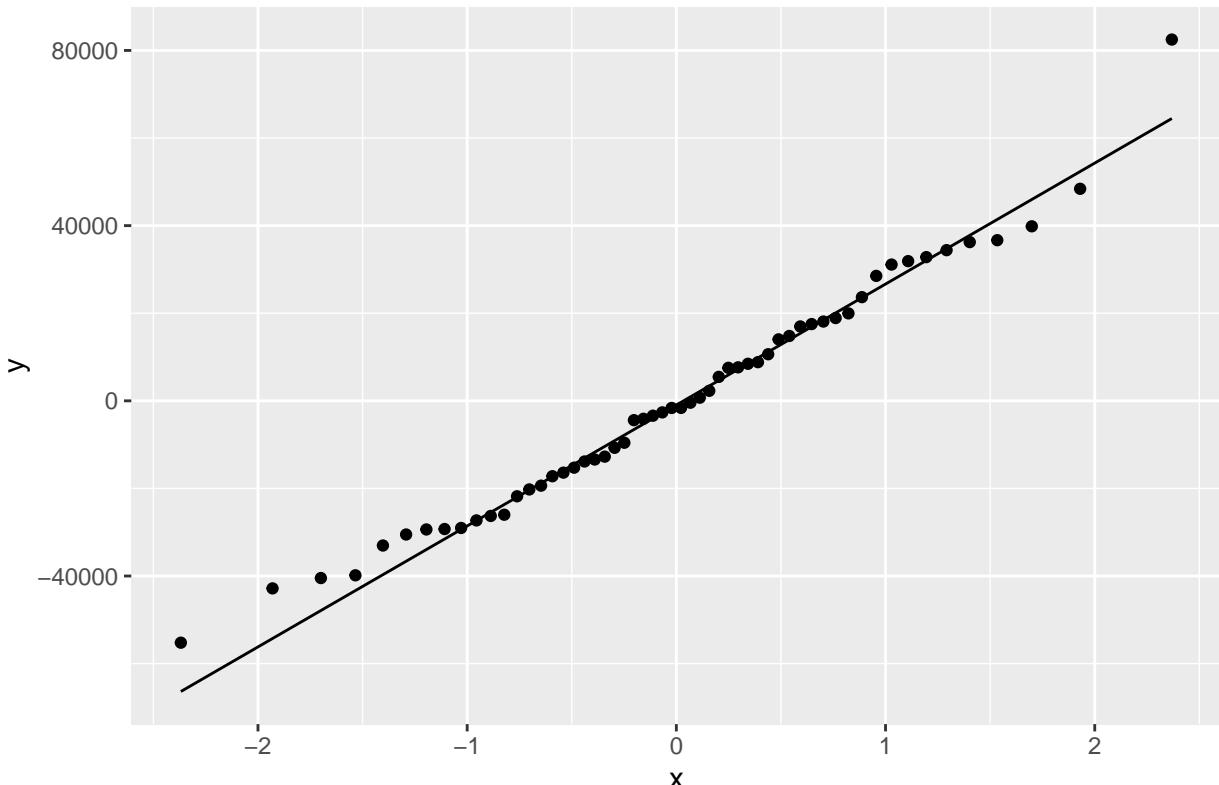
```

```

##  SEXM          1515.99   21371.65    22.96   0.071   0.944
##  AgeInDays     -3812.33    6163.65    28.90  -0.619   0.541
##  SEXM:AgeInDays 16447.92   10249.07    29.82   1.605   0.119
##  ---
##  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES      SEXM    AgInDy
##  SES      -0.213
##  SEXM     -0.601  0.374
##  AgeInDays -0.002  0.036  0.010
##  SEXM:AgInDy  0.003 -0.032  0.020 -0.602
## Reduced Model: Log-Likelihood = -679.49 , Degrees of Freedom = 7
## Max VIF value of the Reduced Model: 1.57160210187435
## Fitting Null Model.
## boundary (singular) fit: see help('isSingular')

```

Q–Q Plot of Residuals for Reduced Model



```

## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## red_fit_model: redmformula
## null_fit_model: nullmformula
##           npar  AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## red_fit_model    7 1373 1387.2 -679.49      1359
## null_fit_model   7 1375 1389.2 -680.52      1361      0 0
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).

```

```

pred.data.5.1.2 <- boot.ci.predict.lmer(
  m = output.5.1.2$reduced_model,
  optimizer = output.5.1.2$optimizer,
  maxfun = output.5.1.2$maxfun,
  data = output.5.1.2$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.5.1.2$fit_data$AgeInDays), max(output.5.1.2$fit_data$AgeInDays),
  reqcol = c("AgeInDays"),
  centercol = c("SES", "SEX"),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## Number of singular fits during bootstrapping: 0 out of 1000 bootstrap iterations.

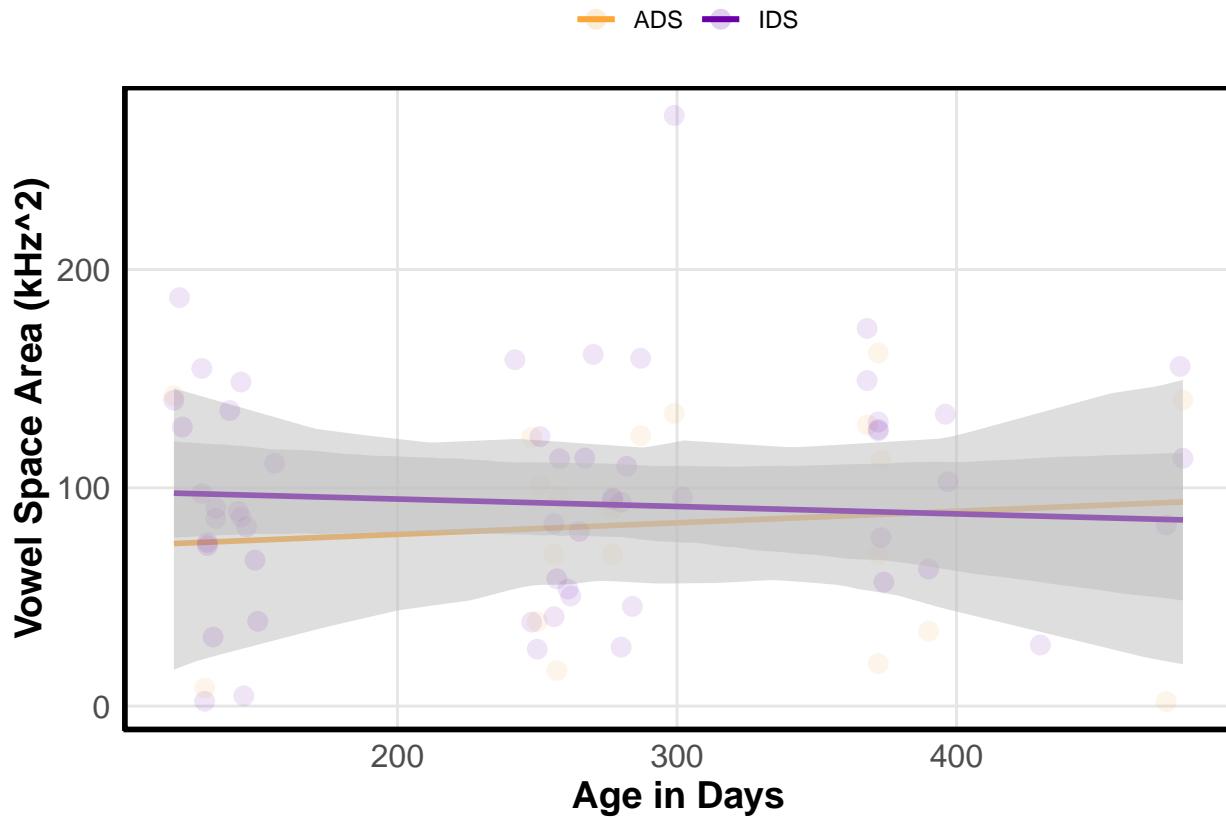
inv_x_transform_fun_1 <- function(x) {
  x * attr(output.5.1.1$fit_data$AgeInDays, "scaled:scale") + attr(output.5.1.1$fit_data$AgeInDays, "sc
}

inv_x_transform_fun_2 <- function(x) {
  x * attr(output.5.1.2$fit_data$AgeInDays, "scaled:scale") + attr(output.5.1.2$fit_data$AgeInDays, "sc
}

inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

Area_cor <- continuous_fit_ci_plot(
  plot.data = list(output.5.1.1$fit_data, output.5.1.2$fit_data),
  coefs = list(pred.data.5.1.1, pred.data.5.1.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "ConvexHullArea",
  x.labs = "Age in Days",
  y.labs = "Vowel Space Area (kHz^2)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  div = 1000,
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = FALSE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  #colors = c("#6A00A8"),
  #alpha_level = 0.3,
  legend.labels = c("ADS", "IDS")
)
print(Area_cor)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.5.1.1$fit_data$AgeInDays, "scaled:scale") + attr(output.5.1.1$fit_data$AgeInDays, "sc")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.5.1.2$fit_data$AgeInDays, "scaled:scale") + attr(output.5.1.2$fit_data$AgeInDays, "sc")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

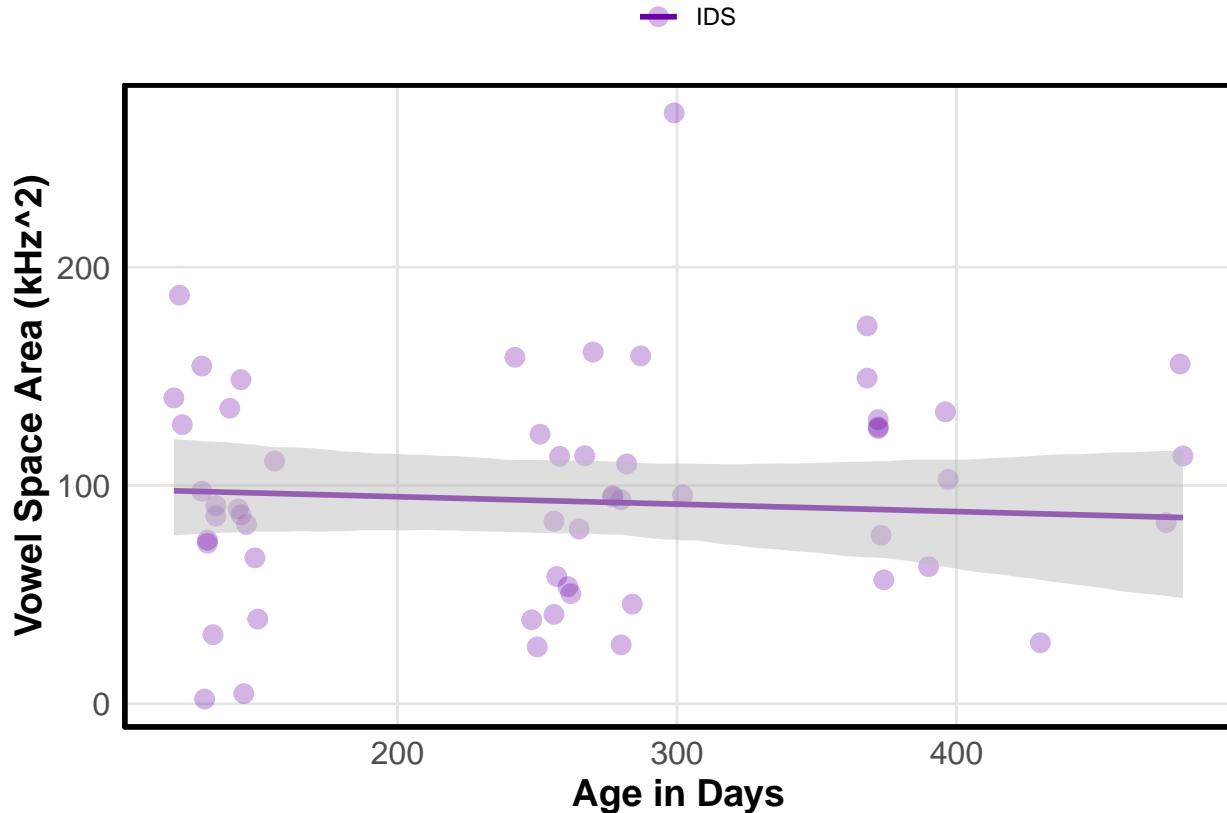
Area_cor_IDS <- continuous_fit_ci_plot(
  plot.data = list(output.5.1.2$fit_data),
  coefs = list(pred.data.5.1.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "ConvexHullArea",
  x.labs = "Age in Days",
  y.labs = "Vowel Space Area (kHz^2)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  div = 1000,
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_2),
  y.ax.transformation = FALSE,
  inv_y_transform_fun = list(inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  colors = c("#6A00A8"),
  )

```

```

alpha_level = 0.3,
legend.labels = c("IDS")
)
print(Area_cor_IDS)

```



#5.2 Border vowels #5.2.1 ADS

```
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_area_full.xlsx"))
```

```
## New names:
```

```
## * ` ` -> `...`
```

```
output.5.2.1 <- lmer.full.reduced.null.compare(ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX +
  optimizer = "bobyqa",
  maxfun = 10000000000,
  fullm = TRUE,
  redm = TRUE,
  redmformula = ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
  nullm = TRUE,
  nullmformula =ConvexHullArea ~ SES + SEX + (1|SPK_id),
  df_ADS)
```

```
## Structure of combined data:
```

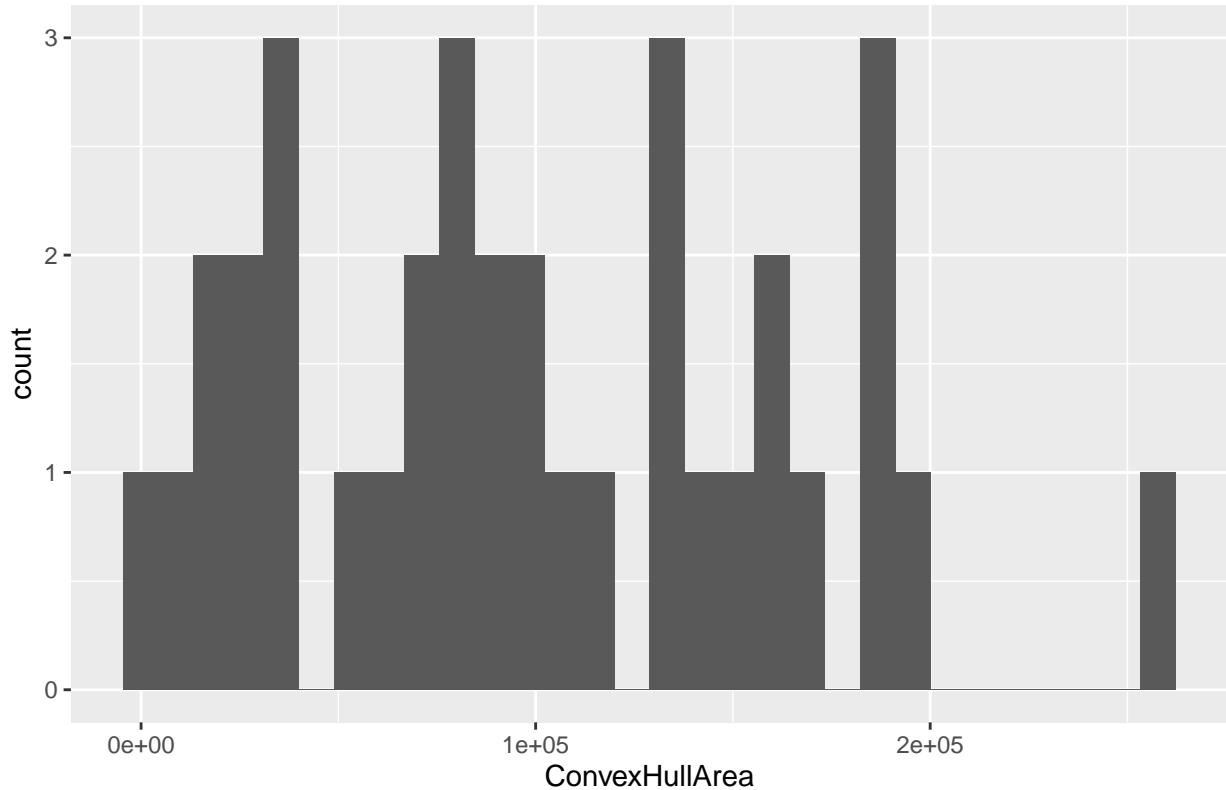
```
## tibble [35 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:35] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:35] "C023" "C028" "C028" "C031" ...
## $ AgeInDays : num [1:35] 251 140 287 373 143 282 374 250 390 131 ...
## $ SES : num [1:35] 5 5 5 8 5 5 5 5 5 4 ...
```

```

## $ SEX : chr [1:35] "M" "F" "F" "F" ...
## $ ConvexHullArea: num [1:35] 101217 74559 182709 158591 38700 ...
## Response variable is: ConvexHullArea
## Structure of combined data after scaling:
## tibble [35 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:35, 1] -1.66 -1.56 -1.46 -1.37 -1.27 ...
## ..- attr(*, "scaled:center")= num 17
## ..- attr(*, "scaled:scale")= num 10.2
## $ SPK_id : chr [1:35] "C023" "C028" "C028" "C031" ...
## $ AgeInDays : num [1:35, 1] -0.219 -1.278 0.124 0.945 -1.25 ...
## ..- attr(*, "scaled:center")= num 274
## ..- attr(*, "scaled:scale")= num 105
## $ SES : num [1:35, 1] 0 0 0 2.19 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.37
## $ SEX : chr [1:35] "M" "F" "F" "F" ...
## $ ConvexHullArea: num [1:35] 101217 74559 182709 158591 38700 ...

```

Histogram of Response Variable

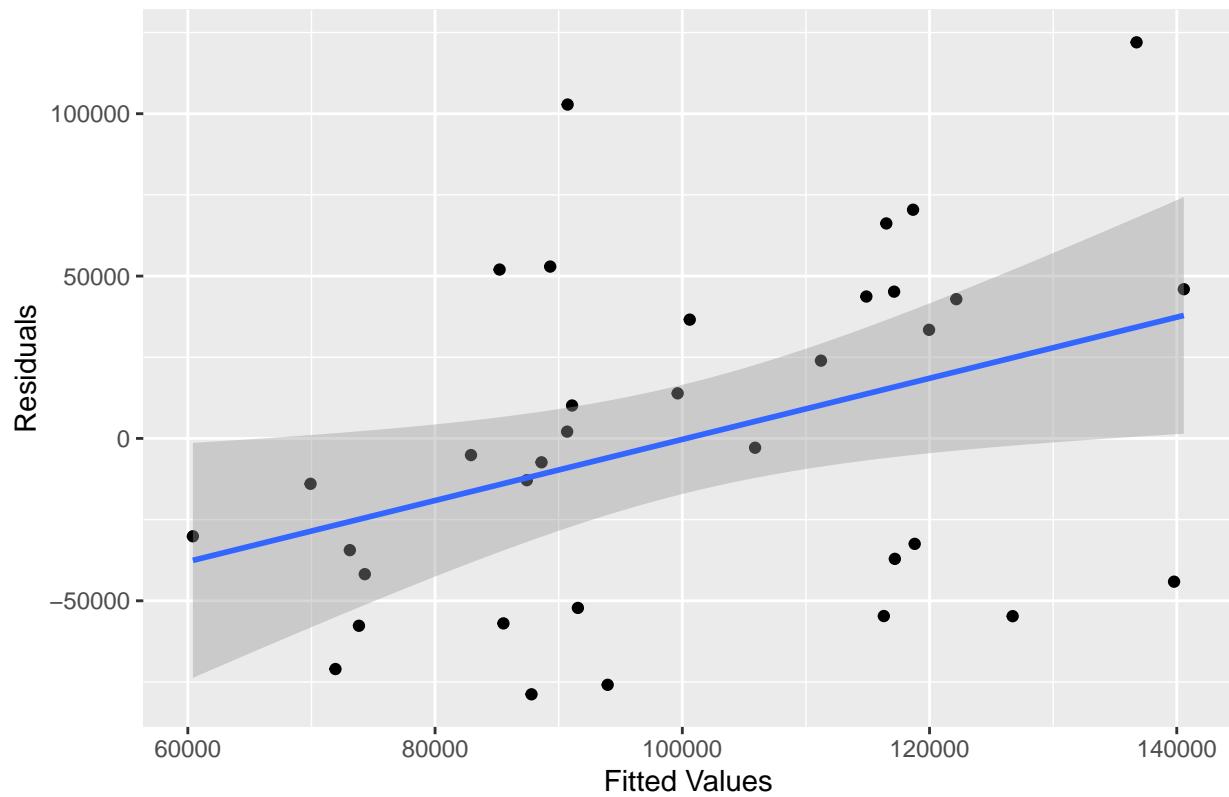


```

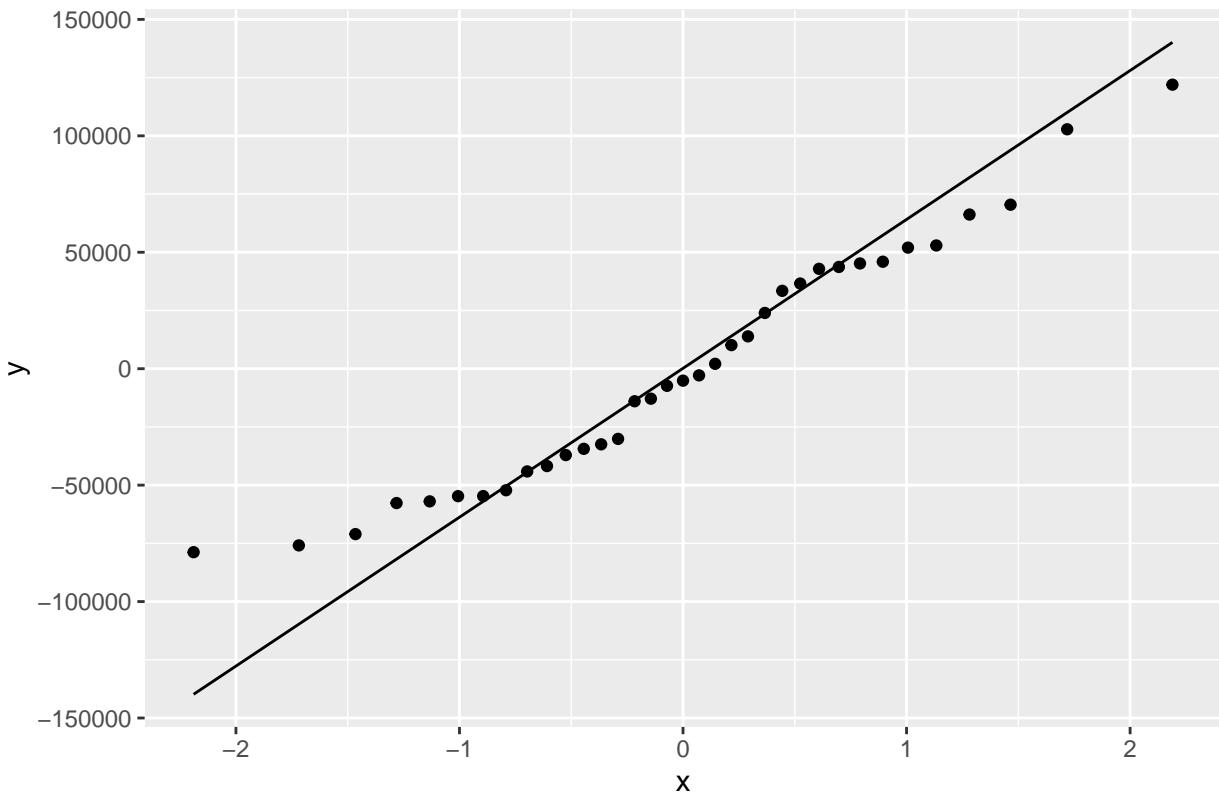
## Response variable is normally distributed; no transformation applied.
## Fitting Full Model.
## Error in fitting the full model: number of observations (=35) <= number of random effects (=44) for ...
## Full model failed. Consider using a different optimizer or increasing maxfun.
## Fitting Null Model.
## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q–Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 763.8
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.28031 -0.68654 -0.04103  0.67726  1.93723
## 
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 9.981e+08 31592
## Residual            3.266e+09 57152
## Number of obs: 35, groups:  SPK_id, 22
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept) 104149.94  15154.55  13.74  6.873 8.44e-06 ***
## SES          -7139.50   13448.38  16.59 -0.531   0.603    
## SEXM         -14082.25   28660.45  15.86 -0.491   0.630    
## AgeInDays    20206.53   13838.39  24.26  1.460   0.157    
## SEXM:AgeInDays -16486.36  21734.02  24.06 -0.759   0.455  
## 
```



```

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

## singular fit
## Number of singular fits during bootstrapping: 8 out of 1000 bootstrap iterations.

#5.2.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_area_full.xlsx"))

## New names:
## * `` -> `...` 

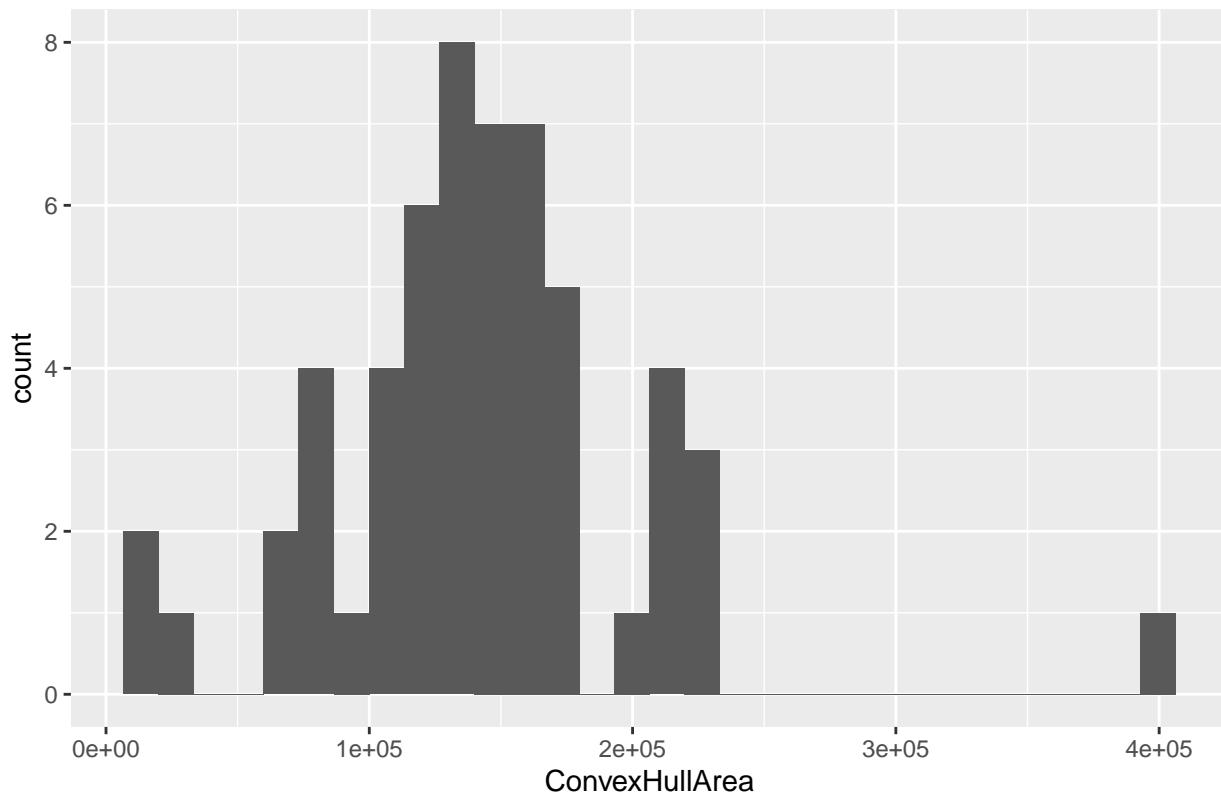
output.5.2.2 <- lmer.full.reduced.null.compare(ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX +
                                                 optimizer = "bobyqa",
                                                 maxfun = 1000000000,
                                                 fullm = TRUE,
                                                 redm = TRUE,
                                                 redmformula = ConvexHullArea ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                                 nullm = TRUE,
                                                 nullmformula =ConvexHullArea ~ SES + SEX + (1+AgeInDays|SPK_id),
                                                 df_IDS)

## Structure of combined data:
## tibble [56 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:56] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays : num [1:56] 280 267 135 251 140 287 135 373 143 282 ...
## $ SES : num [1:56] 5 5 5 5 5 8 8 5 5 ...
## $ SEX : chr [1:56] "M" "F" "M" "M" ...
## $ ConvexHullArea: num [1:56] 127667 167220 123643 147945 156801 ...

## Response variable is: ConvexHullArea
## Structure of combined data after scaling:
## tibble [56 x 6] (S3:tbl_df/tbl/data.frame)
## $ ...1 : num [1:56, 1] -1.69 -1.62 -1.56 -1.5 -1.44 ...
## ..- attr(*, "scaled:center")= num 27.5
## ..- attr(*, "scaled:scale")= num 16.3
## $ SPK_id : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays : num [1:56, 1] 0.2314 0.1093 -1.1303 -0.0409 -1.0834 ...
## ..- attr(*, "scaled:center")= num 255
## ..- attr(*, "scaled:scale")= num 106
## $ SES : num [1:56, 1] -0.0814 -0.0814 -0.0814 -0.0814 -0.0814 ...
## ..- attr(*, "scaled:center")= num 5.11
## ..- attr(*, "scaled:scale")= num 1.32
## $ SEX : chr [1:56] "M" "F" "M" "M" ...
## $ ConvexHullArea: num [1:56] 127667 167220 123643 147945 156801 ...

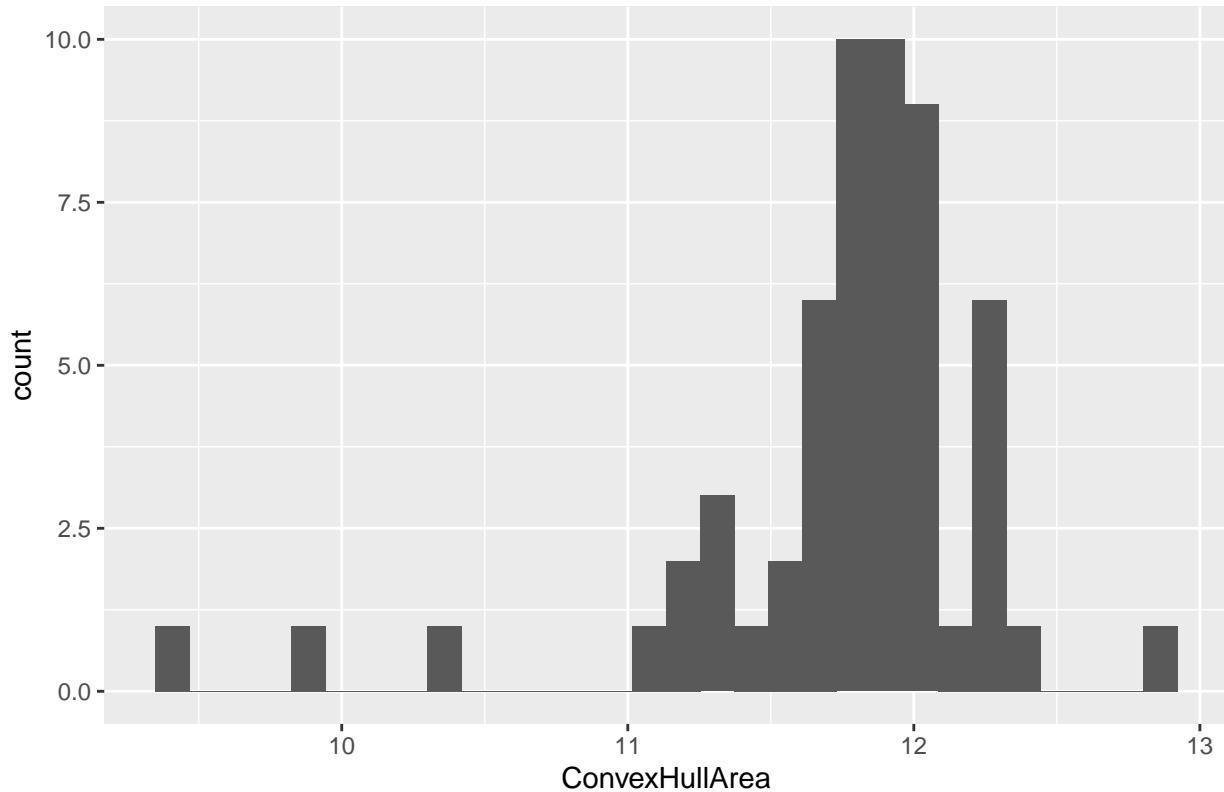
```

Histogram of Response Variable



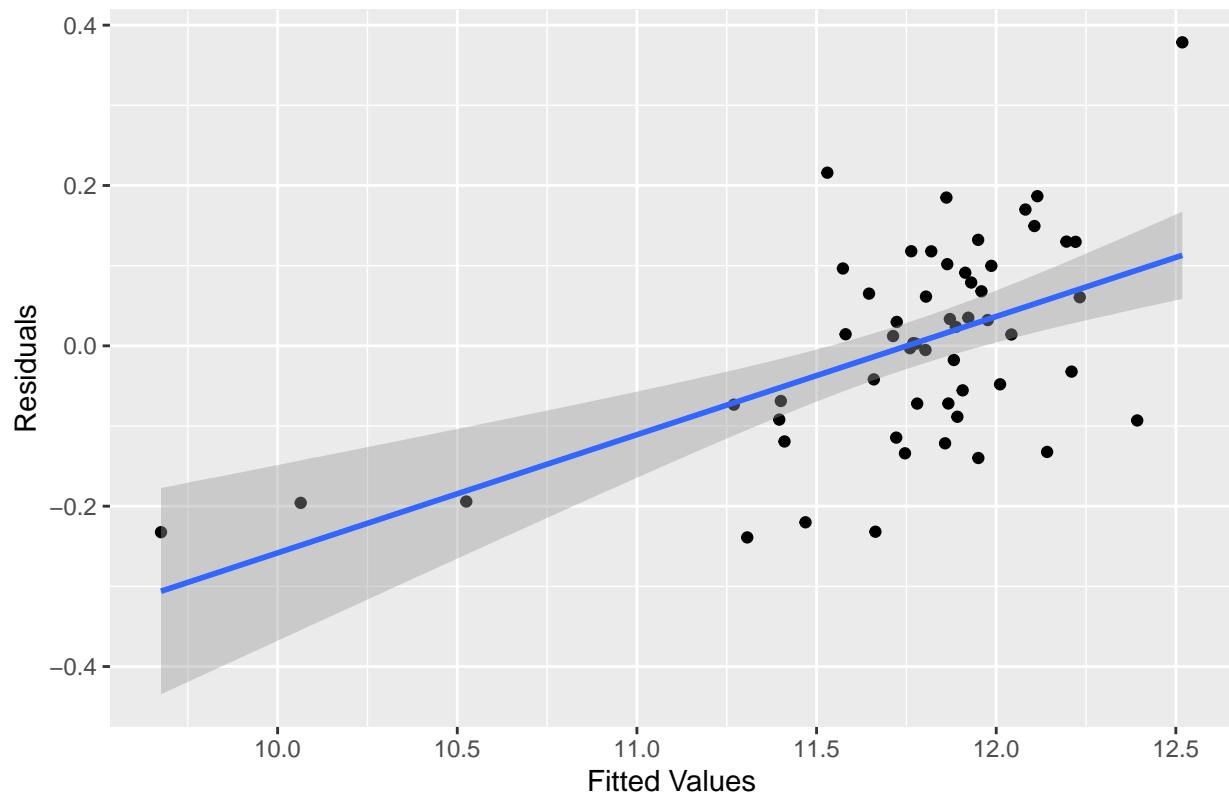
```
## Response variable is not normally distributed. Log transformation applied.  
## Structure of combined data after transformation:  
## tibble [56 x 6] (S3: tbl_df/tbl/data.frame)  
## $ ...1 : num [1:56, 1] -1.69 -1.62 -1.56 -1.5 -1.44 ...  
## ..- attr(*, "scaled:center")= num 27.5  
## ..- attr(*, "scaled:scale")= num 16.3  
## $ SPK_id : chr [1:56] "C010" "C012" "C023" "C023" ...  
## $ AgeInDays : num [1:56, 1] 0.2314 0.1093 -1.1303 -0.0409 -1.0834 ...  
## ..- attr(*, "scaled:center")= num 255  
## ..- attr(*, "scaled:scale")= num 106  
## $ SES : num [1:56, 1] -0.0814 -0.0814 -0.0814 -0.0814 -0.0814 ...  
## ..- attr(*, "scaled:center")= num 5.11  
## ..- attr(*, "scaled:scale")= num 1.32  
## $ SEX : chr [1:56] "M" "F" "M" "M" ...  
## $ ConvexHullArea: num [1:56] 11.8 12 11.7 11.9 12 ...
```

Histogram of Log-transformed Response Variable

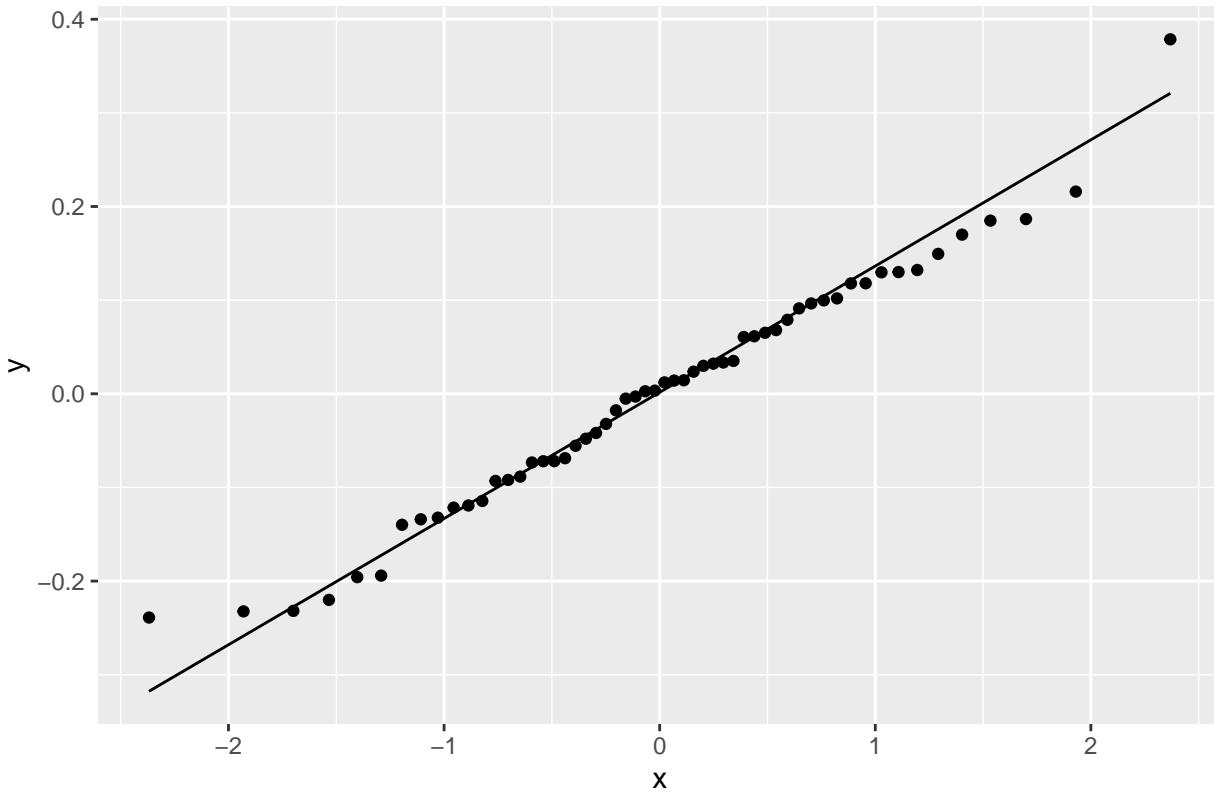


```
## Fitting Full Model.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



Q–Q Plot of Residuals for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 89.7
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -1.0479 -0.3767  0.0314  0.3998  1.5771 
##
## Random effects:
##   Groups   Name        Variance Std.Dev. Corr
##   SPK_id (Intercept) 0.18512  0.4303  
##          AgeInDays   0.13091  0.3618  -0.36 
##   Residual           0.04849  0.2202  
## Number of obs: 56, groups: SPK_id, 27
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept) 11.76053  0.11337 19.21946 103.735 <2e-16 ***
## SES          0.02007  0.10011 20.62677   0.200   0.843    
## SEXM         -0.04138  0.20841 20.44306  -0.199   0.845    
## AgeInDays   -0.03936  0.10627 17.86691  -0.370   0.715    
## 
```

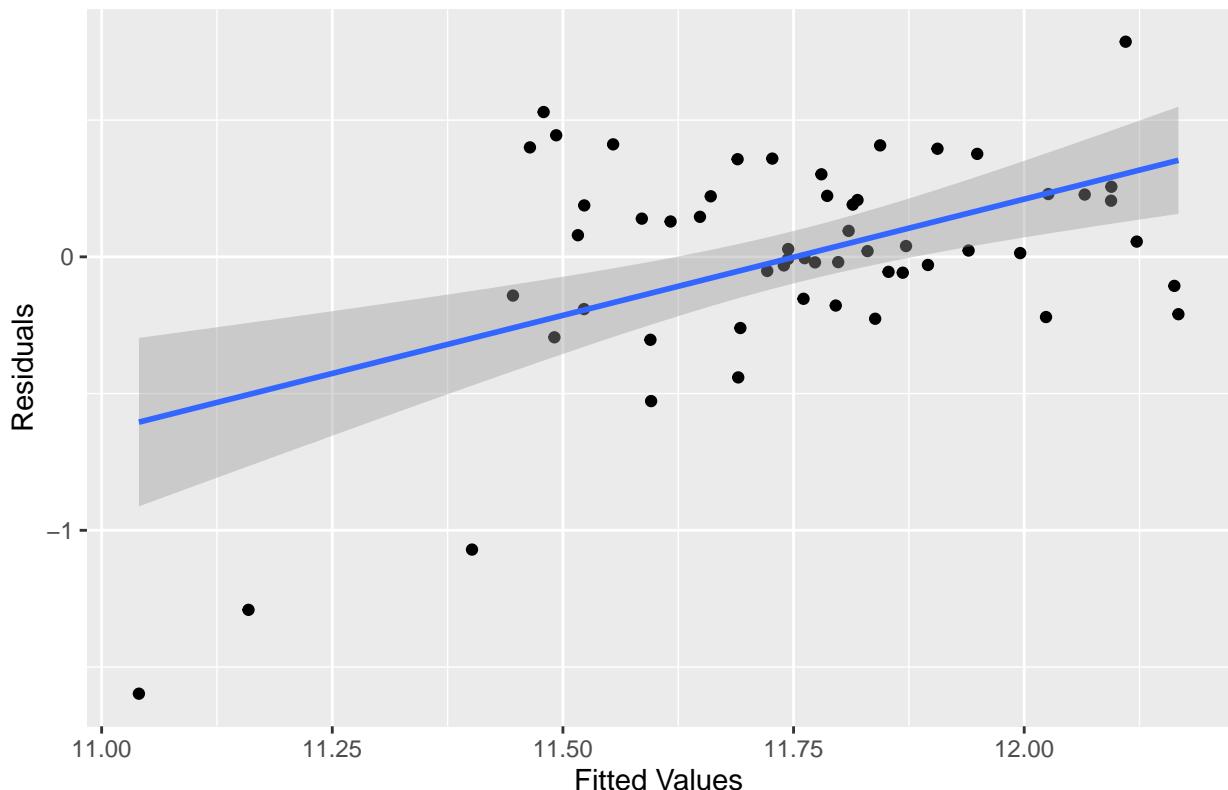
```

## SEXM:AgeInDays  0.26687    0.18194 18.34782   1.467    0.159
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES      -0.216
## SEXM     -0.598  0.370
## AgeInDays -0.265  0.000  0.144
## SEXM:AgInDy  0.155  0.001 -0.222 -0.584
## Full Model: Log-Likelihood = -38.71 , Degrees of Freedom = 9
## Max VIF value of the Full Model: 1.57087063350108
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##          npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model    7 93.922 108.10 -39.961    79.922
## fit_model         9 95.414 113.64 -38.707    77.414 2.508  2     0.2854
## Fitting Reduced Model.

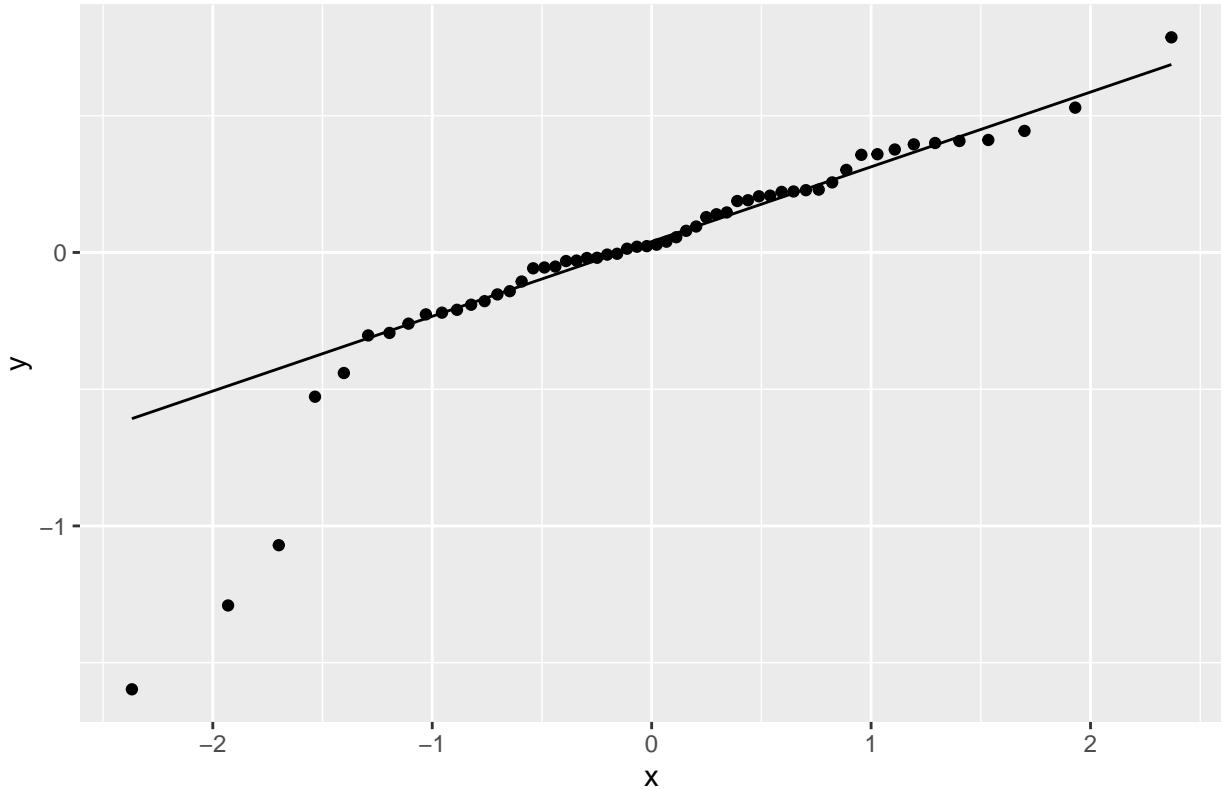
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 101.2
##
## Scaled residuals:
##      Min     1Q Median     3Q    Max 
## -3.3001 -0.3245  0.0408  0.4554  1.5790 
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.1188   0.3447 
## Residual           0.2172   0.4660 
## Number of obs: 56, groups: SPK_id, 27
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)    
## (Intercept) 11.731566  0.115893 17.490984 101.228 <2e-16 ***
## SES          0.056732  0.102938 16.709272   0.551   0.589    
## SEXM         -0.009958  0.213727 18.629970  -0.047   0.963    
## AgeInDays    -0.029123  0.083222 27.464467  -0.350   0.729    
## SEXM:AgeInDays 0.236465  0.137495 29.227092   1.720   0.096 .  

```



```

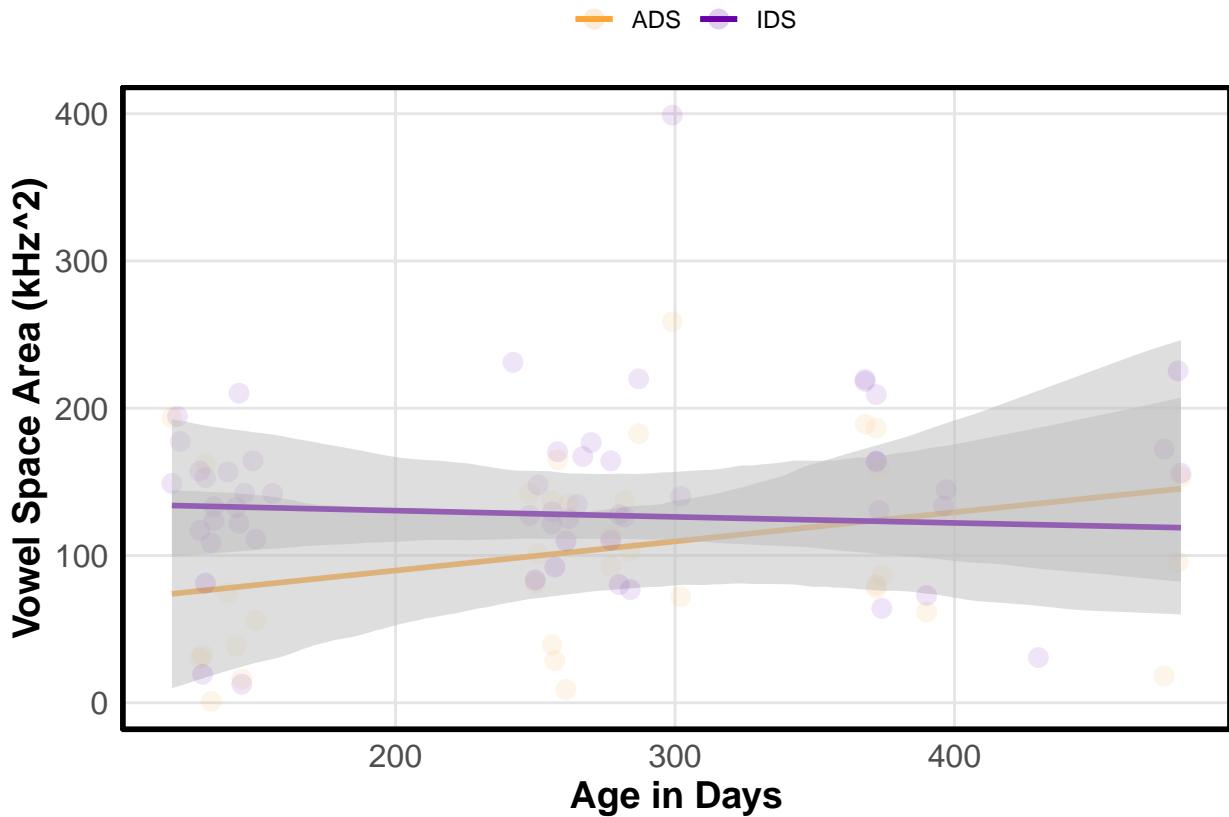
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 14 out of 1000 bootstrap iterations.

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.5.2.1$fit_data$AgeInDays, "scaled:scale") + attr(output.5.2.1$fit_data$AgeInDays, "sc
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.5.2.2$fit_data$AgeInDays, "scaled:scale") + attr(output.5.2.2$fit_data$AgeInDays, "sc
}
inv_y_transform_fun_1 <- function(x) { x }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

Area_full <- continuous_fit_ci_plot(
  plot.data = list(output.5.2.1$fit_data, output.5.2.2$fit_data),
  coefs = list(pred.data.5.2.1,pred.data.5.2.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "ConvexHullArea",
  x.labs = "Age in Days",
  y.labs = "Vowel Space Area (kHz^2)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  div = 1000,
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = TRUE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  #colors = c("#6A00A8"),
  #alpha_level = 0.3,
  #legend.labels = c("IDS")
  legend.labels = c( "ADS","IDS")
)

```

```
print(Area_full)
```



#6. Variability #6.1 ADS

```
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_var.xlsx"))
```

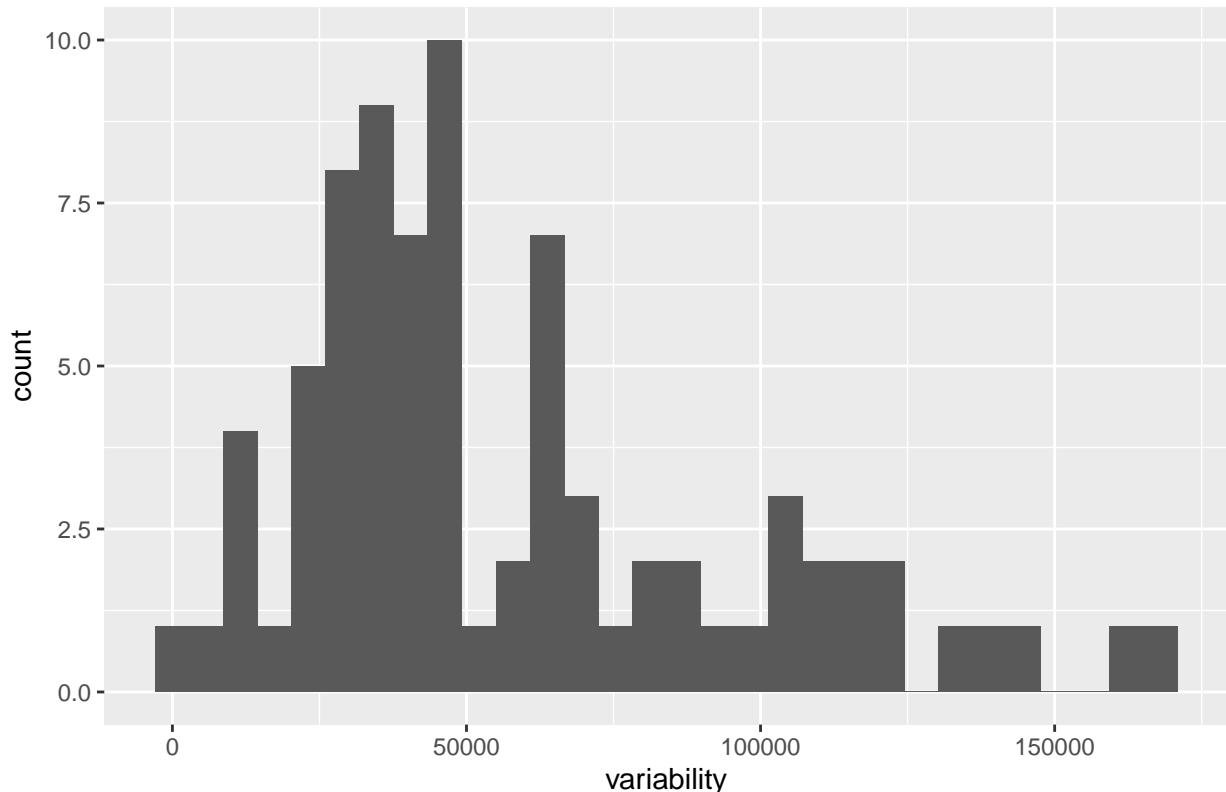
```
## New names:  
## * `` -> `...`  
output.6.1 <- lmer.full.reduced.null.compare(variability ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|AgeInDays:SPK_id),  
                                              optimizer = "bobyqa",  
                                              maxfun = 1000000000,  
                                              fullm = TRUE,  
                                              redm = TRUE,  
                                              redmformula = variability ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),  
                                              nullm = TRUE,  
                                              nullmformula = variability ~ SES + SEX + (1|SPK_id) + (1|vowels),  
                                              df_ADS)  
  
## Structure of combined data:  
## tibble [80 x 8] (S3: tbl_df/tbl/data.frame)  
## $ ...1 : num [1:80] 0 1 2 3 4 5 6 7 8 9 ...  
## $ SPK_id : chr [1:80] "C028" "C028" "C028" "C031" ...  
## $ AgeInDays : num [1:80] 287 287 287 373 373 373 373 282 282 282 ...  
## $ SES : num [1:80] 5 5 5 8 8 8 8 5 5 5 ...  
## $ SEX : chr [1:80] "F" "F" "F" "F" ...  
## $ vowels : chr [1:80] "a" "ai" "e" "a" ...
```

```

## $ no_Samples : num [1:80] 6 3 6 14 7 10 3 4 7 3 ...
## $ variability: num [1:80] 49214 91935 28978 64856 89508 ...
## Response variable is: variability
## Structure of combined data after scaling:
## tibble [80 x 8] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:80, 1] -1.7 -1.66 -1.61 -1.57 -1.53 ...
## ..- attr(*, "scaled:center")= num 39.5
## ..- attr(*, "scaled:scale")= num 23.2
## $ SPK_id    : chr [1:80] "C028" "C028" "C028" "C031" ...
## $ AgeInDays : num [1:80, 1] -0.13 -0.13 -0.13 0.757 0.757 ...
## ..- attr(*, "scaled:center")= num 300
## ..- attr(*, "scaled:scale")= num 97.1
## $ SES       : num [1:80, 1] -0.119 -0.119 -0.119 2.477 2.477 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.16
## $ SEX       : chr [1:80] "F" "F" "F" "F" ...
## $ vowels    : chr [1:80] "a" "ai" "e" "a" ...
## $ no_Samples: num [1:80, 1] 0.187 -0.906 0.187 3.099 0.551 ...
## ..- attr(*, "scaled:center")= num 5.49
## ..- attr(*, "scaled:scale")= num 2.75
## $ variability: num [1:80] 49214 91935 28978 64856 89508 ...

```

Histogram of Response Variable



```

## Response variable is not normally distributed. Log transformation applied.
## Structure of combined data after transformation:
## tibble [80 x 8] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:80, 1] -1.7 -1.66 -1.61 -1.57 -1.53 ...

```

```

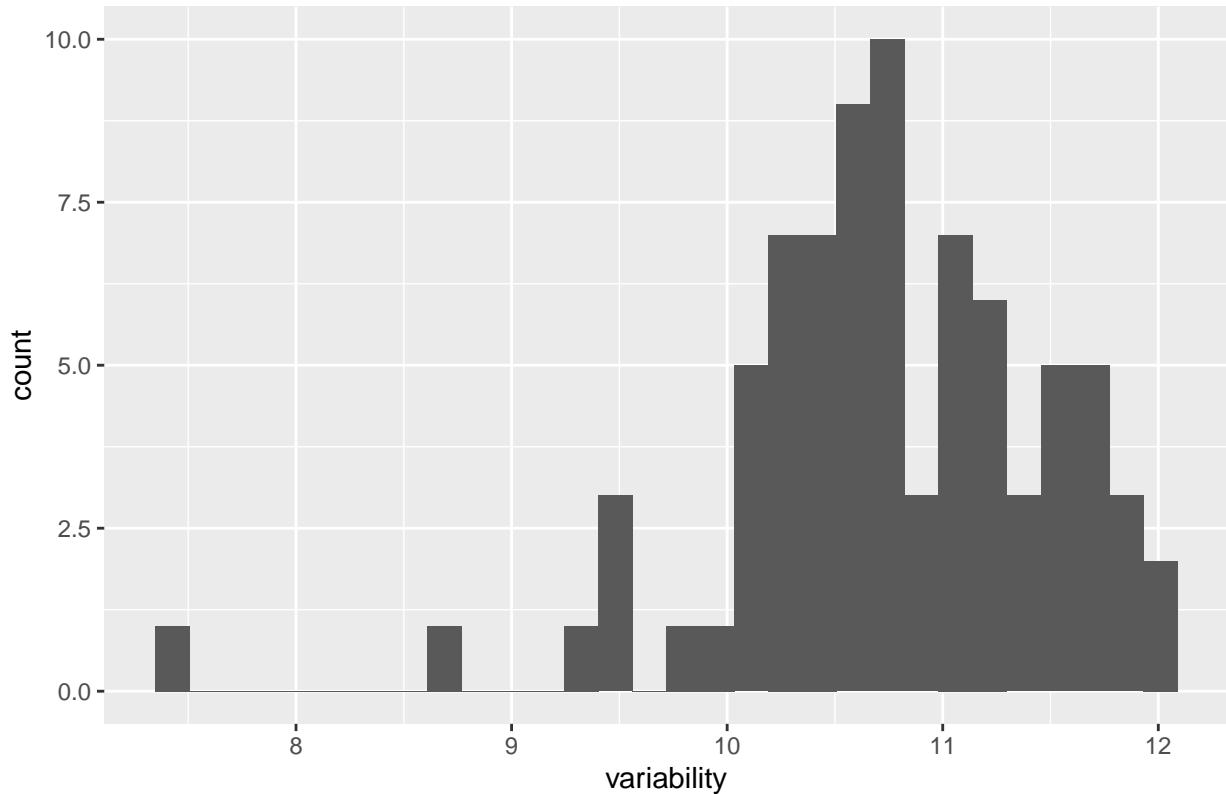
## ..- attr(*, "scaled:center")= num 39.5
## ..- attr(*, "scaled:scale")= num 23.2
## $ SPK_id      : chr [1:80] "C028" "C028" "C028" "C031" ...
## $ AgeInDays   : num [1:80, 1] -0.13 -0.13 -0.13 0.757 0.757 ...
## ..- attr(*, "scaled:center")= num 300
## ..- attr(*, "scaled:scale")= num 97.1
## $ SES         : num [1:80, 1] -0.119 -0.119 -0.119 2.477 2.477 ...
## ..- attr(*, "scaled:center")= num 5.14
## ..- attr(*, "scaled:scale")= num 1.16
## $ SEX         : chr [1:80] "F" "F" "F" "F" ...
## $ vowels      : chr [1:80] "a" "ai" "e" "a" ...
## $ no_Samples  : num [1:80, 1] 0.187 -0.906 0.187 3.099 0.551 ...
## ..- attr(*, "scaled:center")= num 5.49
## ..- attr(*, "scaled:scale")= num 2.75
## $ variability: num [1:80] 10.8 11.4 10.3 11.1 11.4 ...

## Fitting Full Model.

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

```

Histogram of Log-transformed Response Variable

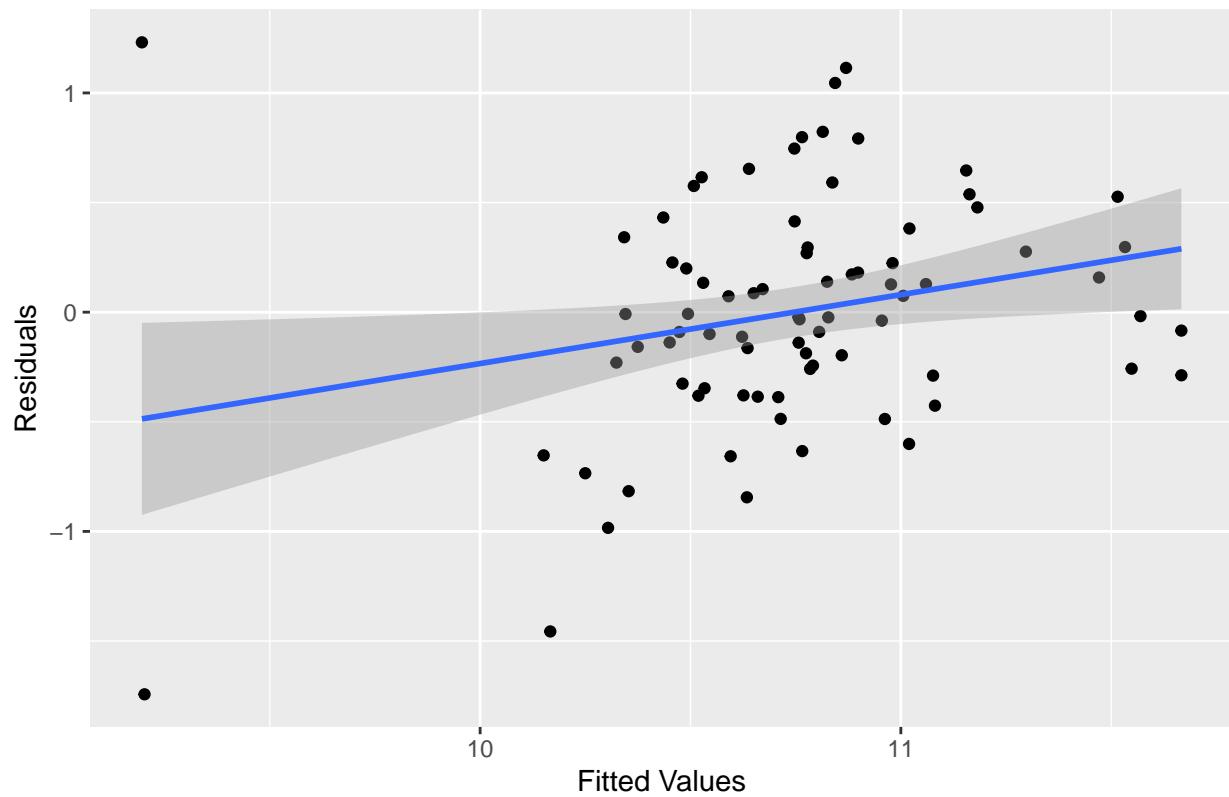


```

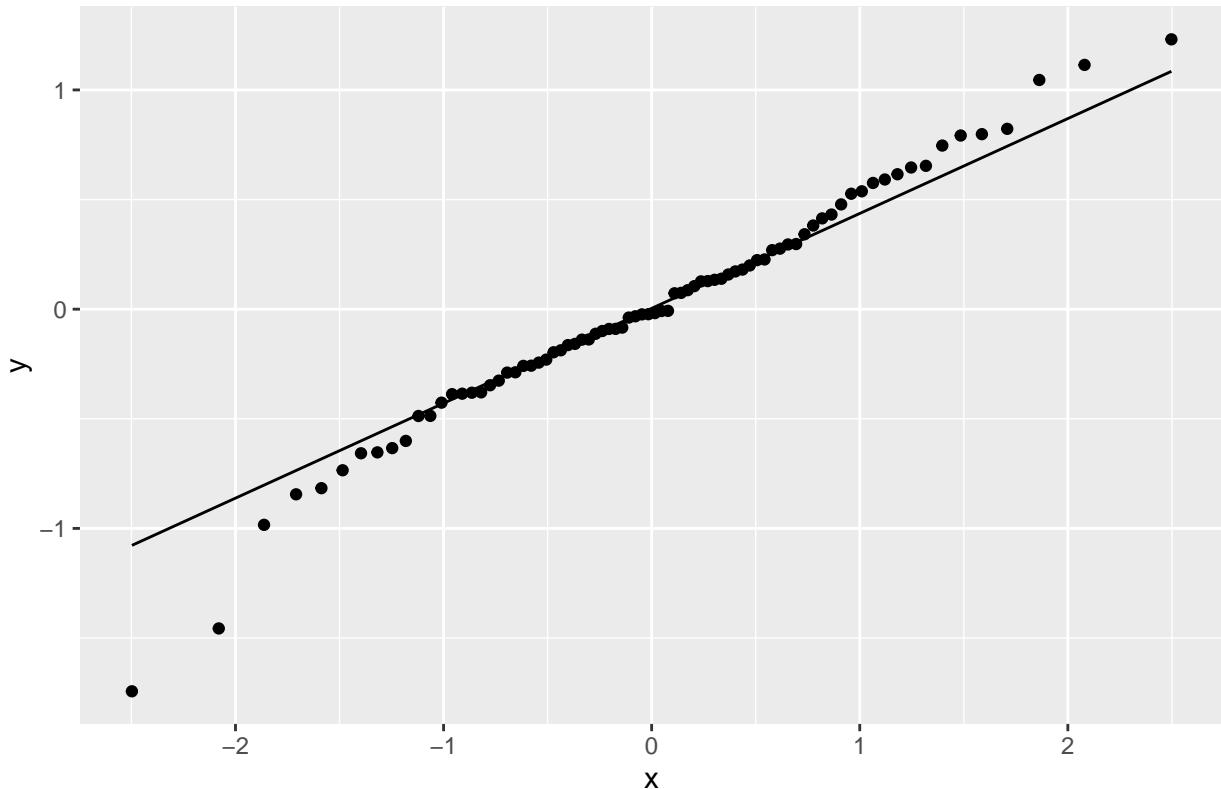
## Full model is singular. Consider simplifying the random effects structure.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Full Model



Q-Q Plot of Residuals for Full Model



```

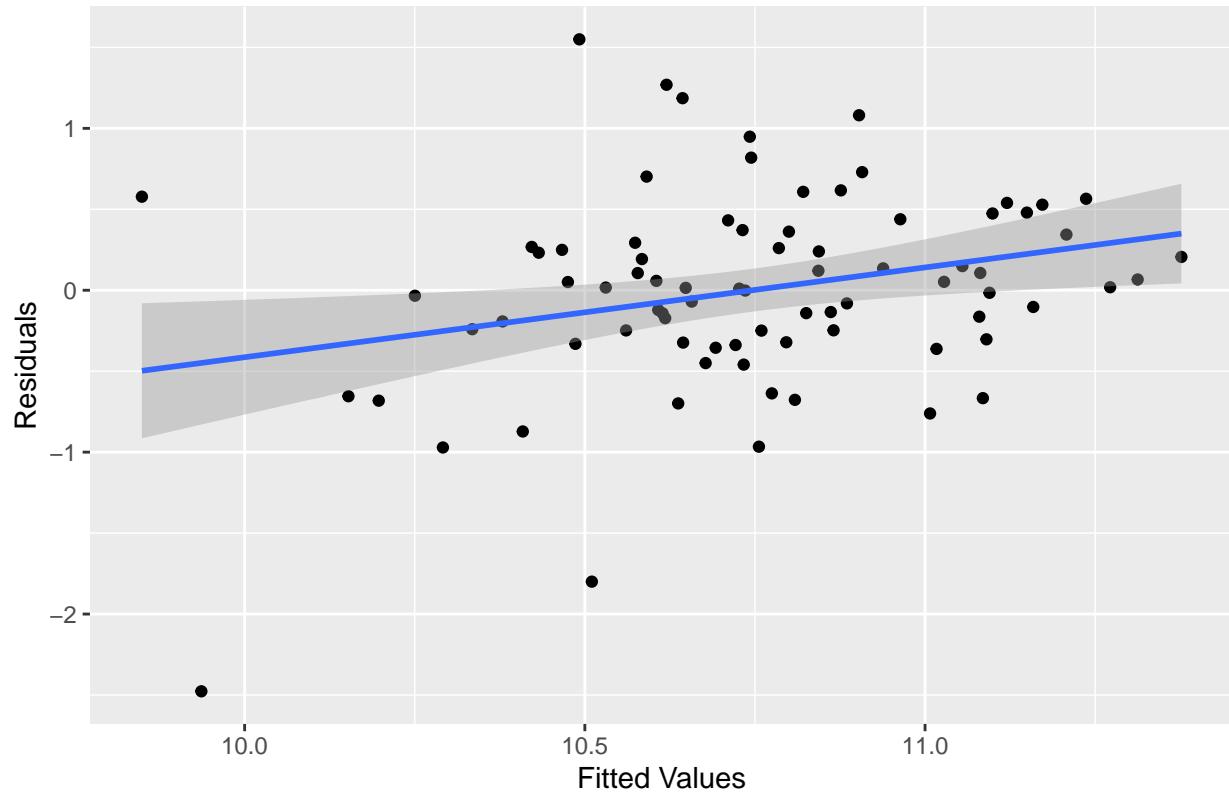
## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 171.6
##
## Scaled residuals:
##      Min     1Q   Median     3Q    Max 
## -2.94963 -0.50462 -0.02355  0.45492  2.23744
## 
## Random effects:
##   Groups   Name        Variance Std.Dev. Corr
##   SPK_id   (Intercept) 0.04914  0.2217    
##          AgeInDays   0.18179  0.4264   -0.76  
##   vowels   (Intercept) 0.04843  0.2201    
##          AgeInDays   0.01532  0.1238   1.00  
##   Residual           0.32923  0.5738    
## 
## Number of obs: 80, groups:  SPK_id, 18; vowels, 8
## 
## Fixed effects:
##             Estimate Std. Error    df t value Pr(>|t|)    
## (Intercept) 10.63310  0.14181 10.19442 74.982 2.57e-15 ***
## SES         0.06235  0.09440 10.35050  0.660    0.523  
## 
```

```

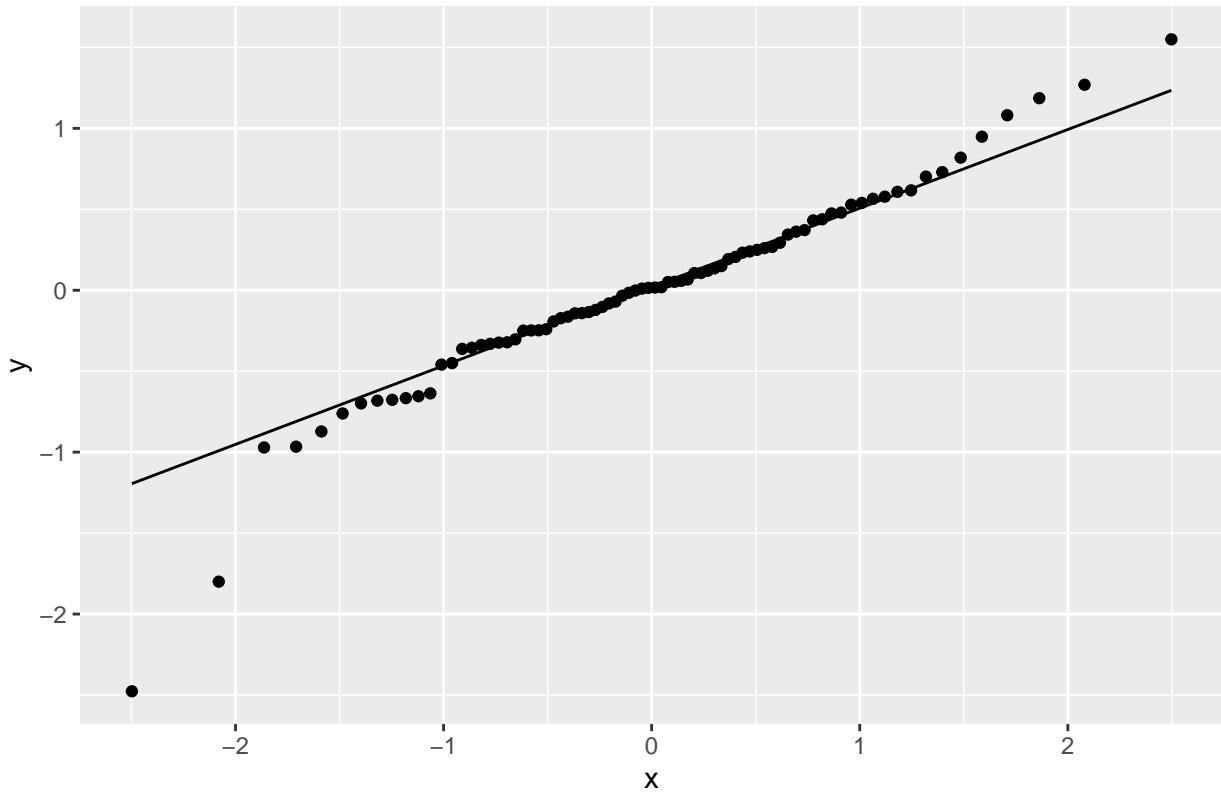
## SEXM          -0.02818   0.21355 10.77911 -0.132    0.897
## AgeInDays    -0.08869   0.20583 12.44555 -0.431    0.674
## SEXM:AgeInDays 0.34151   0.30502  8.10520  1.120    0.295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES     SEXM   AgInDy
## SES        -0.124
## SEXM       -0.438  0.316
## AgeInDays  -0.025 -0.350  0.033
## SEXM:AgInDy 0.108  0.277 -0.226 -0.652
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -80.86 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 2.12297050607084
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##           npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
## null_fit_model  6 186.74 201.03 -87.372   174.74
## fit_model      12 185.71 214.30 -80.856   161.71 13.031  6   0.04254 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Fitting Reduced Model.
## `geom_smooth()` using formula = 'y ~ x'

```

Residuals vs Fitted Values for Reduced Model



Q-Q Plot of Residuals for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 182
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -3.6194 -0.4399  0.0075  0.5201  2.2150
##
## Random effects:
## Groups   Name        Variance Std.Dev.
## SPK_id   (Intercept) 0.11862  0.3444
## vowels   (Intercept) 0.03728  0.1931
## Residual            0.42852  0.6546
## Number of obs: 80, groups: SPK_id, 18; vowels, 8
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 10.70595  0.16521 10.58177 64.802 4.22e-15 ***
## SES          0.02215  0.11553 15.01752  0.192   0.851
## SEXM         -0.19063  0.26807 10.60225 -0.711   0.492
## AgeInDays    -0.09802  0.12476 71.04282 -0.786   0.435

```



```
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## boundary (singular) fit: see help('isSingular')
## singular fit
## Number of singular fits during bootstrapping: 78 out of 1000 bootstrap iterations.
```

```

#6.2 IDS

df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_var.xlsx"))

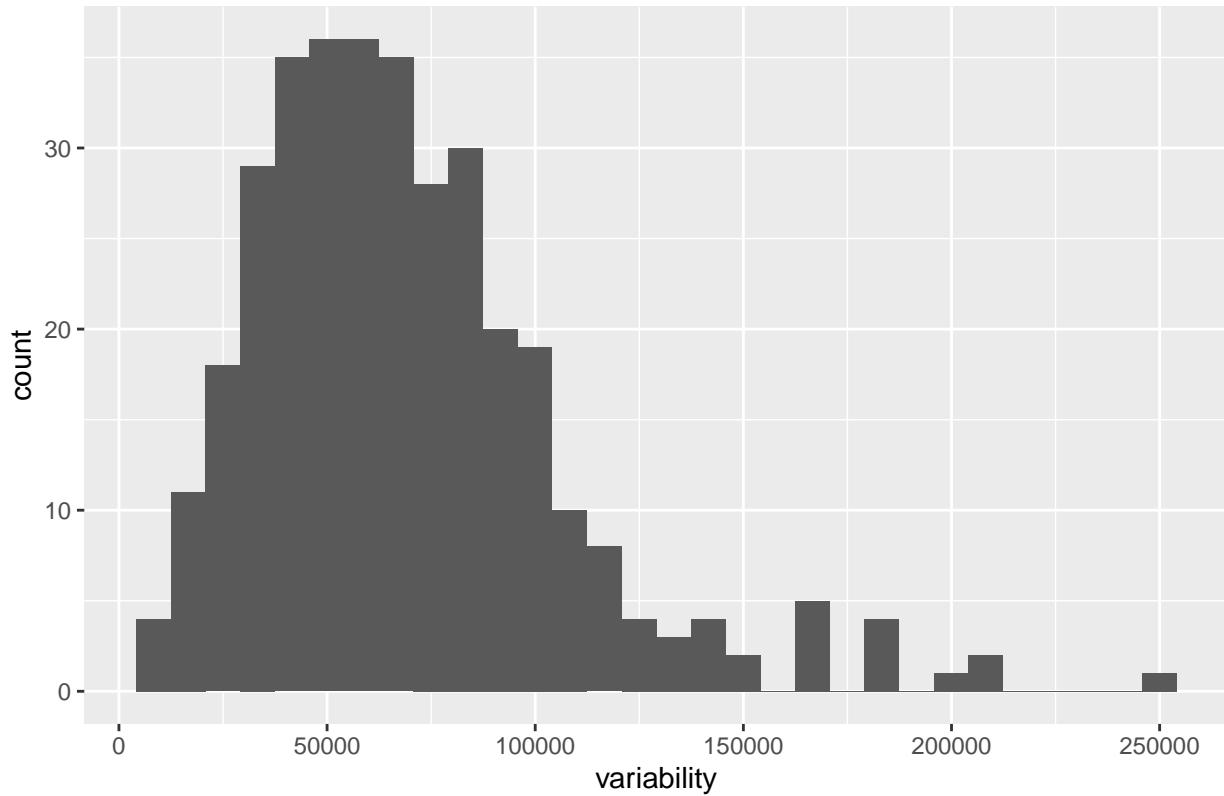
## New names:
## * `` -> `...` 

output.6.2 <- lmer.full.reduced.null.compare(variability ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowels),
                                               optimizer = "bobyqa",
                                               maxfun = 1000000000,
                                               fullm = TRUE,
                                               redm = TRUE,
                                               redmformula = variability ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowels),
                                               nullm = TRUE,
                                               nullmformula = variability ~ SES + SEX + (1+AgeInDays|SPK_id) + (1+AgeInDays|vowels),
                                               df_IDS)

## Structure of combined data:
## tibble [345 x 8] (S3:tbl_df/tbl/data.frame)
## $ ...1      : num [1:345] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:345] "C010" "C010" "C010" "C010" ...
## $ AgeInDays : num [1:345] 280 280 280 280 280 280 280 267 267 267 ...
## $ SES       : num [1:345] 5 5 5 5 5 5 5 5 5 5 ...
## $ SEX       : chr [1:345] "M" "M" "M" "M" ...
## $ vowels    : chr [1:345] "a" "ai" "e" "i" ...
## $ no_Samples: num [1:345] 31 15 11 5 3 11 5 27 16 25 ...
## $ variability: num [1:345] 64479 54889 31907 50607 51055 ...
## Response variable is: variability
## Structure of combined data after scaling:
## tibble [345 x 8] (S3:tbl_df/tbl/data.frame)
## $ ...1      : num [1:345, 1] -1.72 -1.71 -1.7 -1.69 -1.68 ...
## $ ..- attr(*, "scaled:center")= num 172
## $ ..- attr(*, "scaled:scale")= num 99.7
## $ SPK_id    : chr [1:345] "C010" "C010" "C010" "C010" ...
## $ AgeInDays : num [1:345, 1] 0.302 0.302 0.302 0.302 0.302 ...
## $ ..- attr(*, "scaled:center")= num 249
## $ ..- attr(*, "scaled:scale")= num 102
## $ SES       : num [1:345, 1] -0.0893 -0.0893 -0.0893 -0.0893 -0.0893 ...
## $ ..- attr(*, "scaled:center")= num 5.12
## $ ..- attr(*, "scaled:scale")= num 1.36
## $ SEX       : chr [1:345] "M" "M" "M" "M" ...
## $ vowels    : chr [1:345] "a" "ai" "e" "i" ...
## $ no_Samples: num [1:345, 1] 2.771 0.706 0.19 -0.584 -0.842 ...
## $ ..- attr(*, "scaled:center")= num 9.53
## $ ..- attr(*, "scaled:scale")= num 7.75
## $ variability: num [1:345] 64479 54889 31907 50607 51055 ...

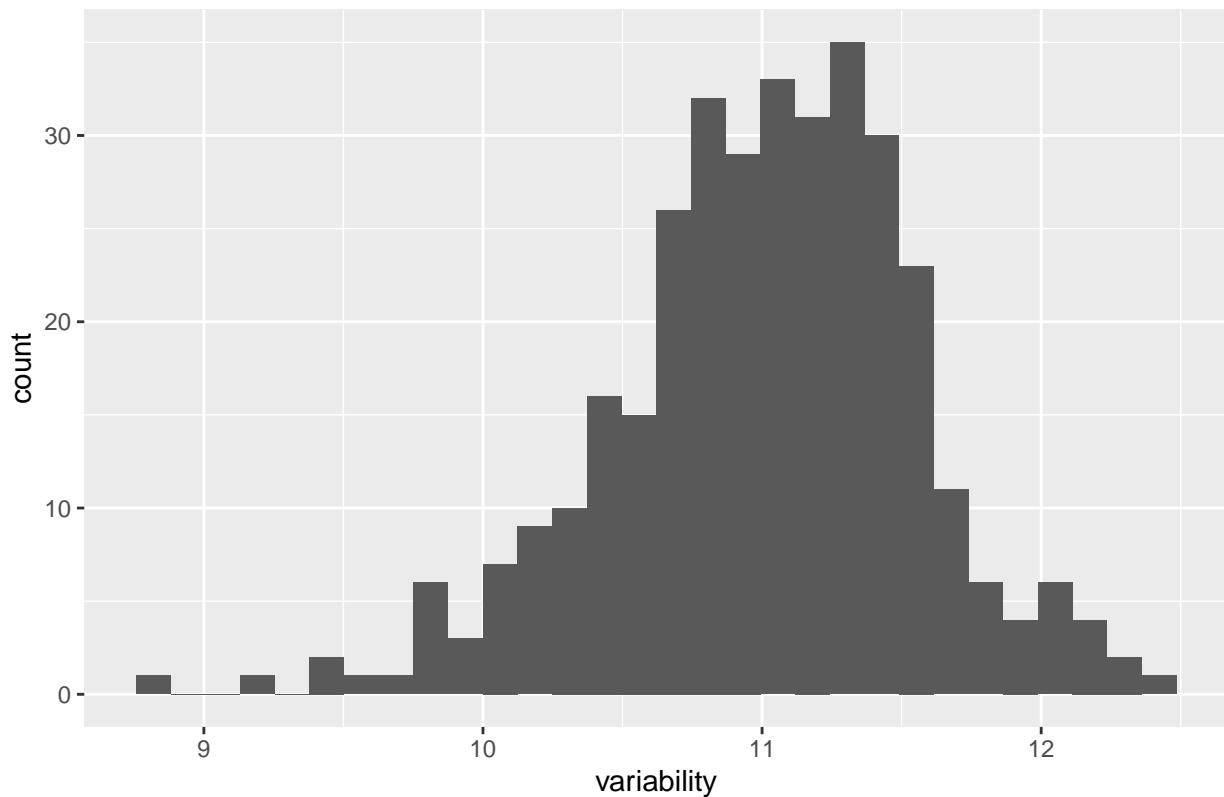
```

Histogram of Response Variable



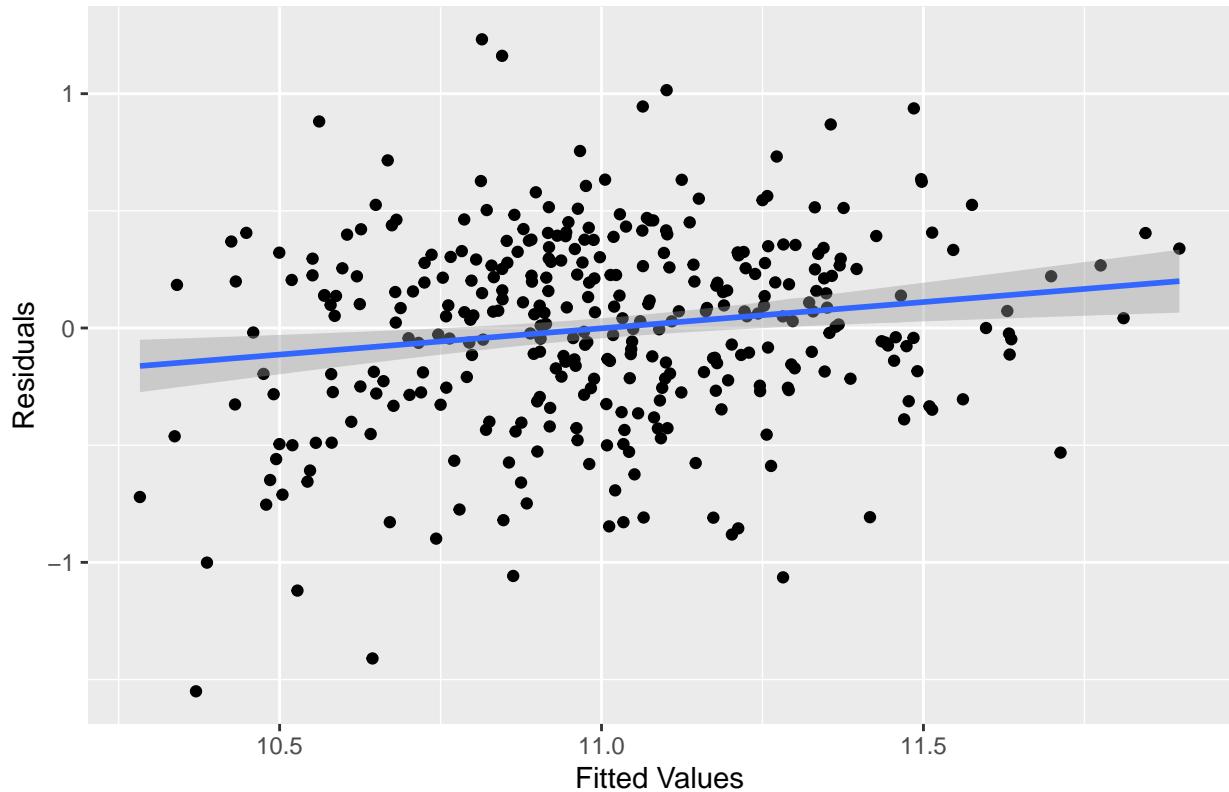
```
## Response variable is not normally distributed. Log transformation applied.  
## Structure of combined data after transformation:  
## tibble [345 x 8] (S3: tbl_df/tbl/data.frame)  
## $ ...1 : num [1:345, 1] -1.72 -1.71 -1.7 -1.69 -1.68 ...  
## ..- attr(*, "scaled:center")= num 172  
## ..- attr(*, "scaled:scale")= num 99.7  
## $ SPK_id : chr [1:345] "C010" "C010" "C010" "C010" ...  
## $ AgeInDays : num [1:345, 1] 0.302 0.302 0.302 0.302 0.302 ...  
## ..- attr(*, "scaled:center")= num 249  
## ..- attr(*, "scaled:scale")= num 102  
## $ SES : num [1:345, 1] -0.0893 -0.0893 -0.0893 -0.0893 -0.0893 ...  
## ..- attr(*, "scaled:center")= num 5.12  
## ..- attr(*, "scaled:scale")= num 1.36  
## $ SEX : chr [1:345] "M" "M" "M" "M" ...  
## $ vowels : chr [1:345] "a" "ai" "e" "i" ...  
## $ no_Samples : num [1:345, 1] 2.771 0.706 0.19 -0.584 -0.842 ...  
## ..- attr(*, "scaled:center")= num 9.53  
## ..- attr(*, "scaled:scale")= num 7.75  
## $ variability: num [1:345] 11.1 10.9 10.4 10.8 10.8 ...  
## Fitting Full Model.  
## boundary (singular) fit: see help('isSingular')  
## boundary (singular) fit: see help('isSingular')
```

Histogram of Log-transformed Response Variable



```
## Full model is singular. Consider simplifying the random effects structure.  
## `geom_smooth()` using formula = 'y ~ x'
```

Residuals vs Fitted Values for Full Model



```

## Summary of the Full Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: fullmformula
##   Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 491.9
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max 
## -3.5410 -0.5776  0.1210  0.6482  2.8552 
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.0448441 0.21176
##          AgeInDays   0.0304913 0.17462 -0.12
## vowels   (Intercept) 0.0414932 0.20370
##          AgeInDays   0.0003634 0.01906 -1.00
## Residual           0.1882704 0.43390
## Number of obs: 345, groups: SPK_id, 27; vowels, 9
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)    
## (Intercept) 10.96081   0.09286 17.42340 118.030 <2e-16 ***
## SES         -0.04022   0.05742 20.68256  -0.700   0.4915  

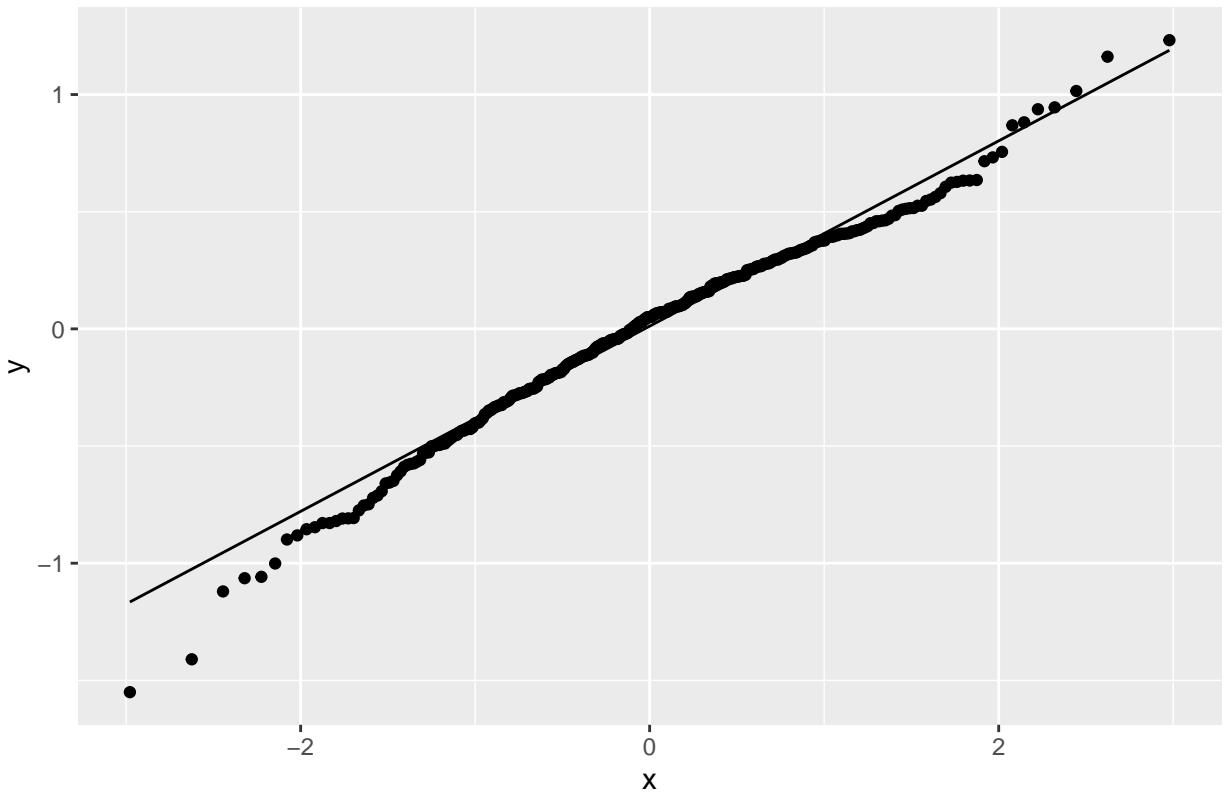
```

```

## SEXM          -0.14156   0.11770 24.87008  -1.203   0.2404
## AgeInDays     0.02194   0.05798 19.76119   0.378   0.7092
## SEXM:AgeInDays -0.18067  0.10099 22.52137  -1.789   0.0871 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) SES    SEXM  AgInDy
## SES        -0.171
## SEXM       -0.406  0.414
## AgeInDays  -0.158  0.050  0.074
## SEXM:AgInDy  0.045 -0.043 -0.045 -0.567
## optimizer (bobyqa) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
##
## Full Model: Log-Likelihood = -237.52 , Degrees of Freedom = 12
## Max VIF value of the Full Model: 1.47634076766356
## Fitting Null Model.
##
## boundary (singular) fit: see help('isSingular')

```

Q–Q Plot of Residuals for Full Model



```

## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: nullmformula
## fit_model: fullmformula
##             npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)

```

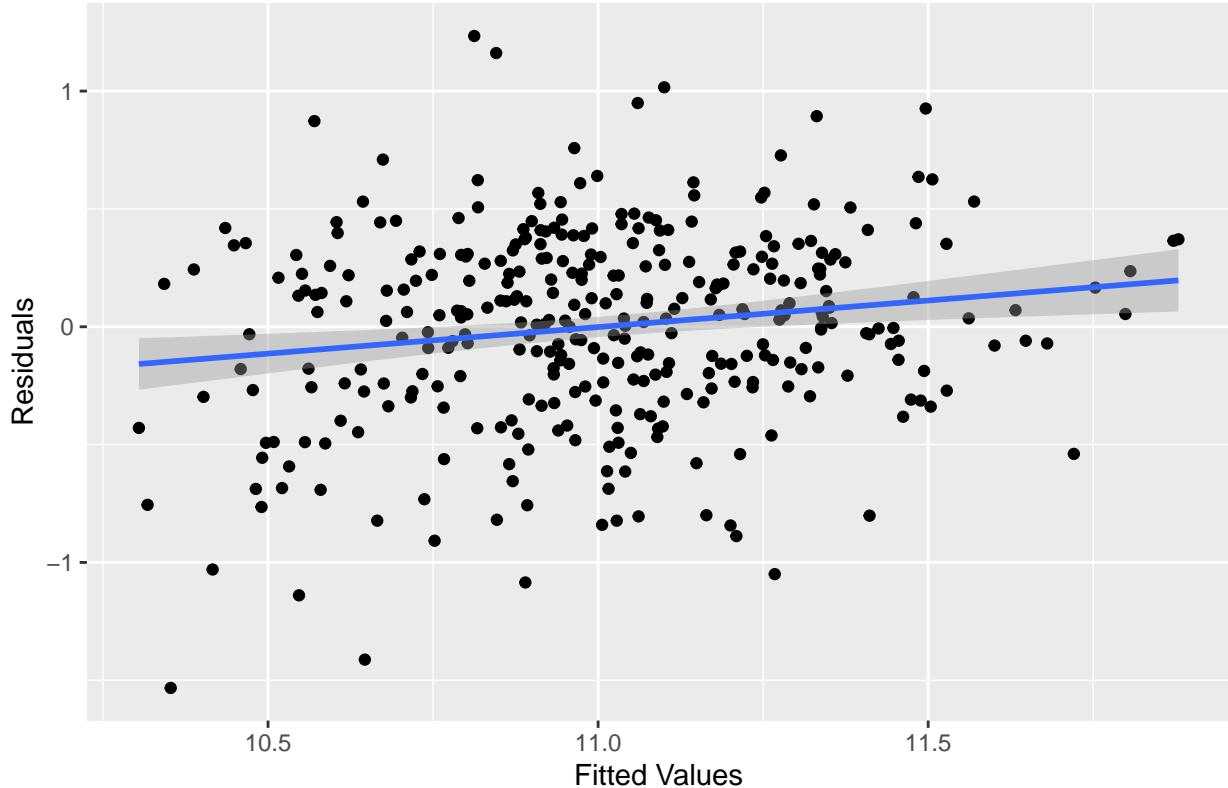
```

## null_fit_model    10 498.92 537.35 -239.46    478.92
## fit_model        12 499.04 545.16 -237.52    475.04 3.8765  2      0.144
## Fitting Reduced Model.

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

```

Residuals vs Fitted Values for Reduced Model



```

## Summary of the Reduced Model:
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: redmformula
## Data: combined_data
## Control: lmerControl(optimizer = optimizer, optCtrl = list(maxfun = maxfun))
##
## REML criterion at convergence: 492.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.4969 -0.5602  0.0953  0.6472  2.8563
##
## Random effects:
## Groups   Name        Variance Std.Dev. Corr
## SPK_id   (Intercept) 0.04456  0.2111
##          AgeInDays   0.03076  0.1754  -0.12
## vowels   (Intercept) 0.04204  0.2050
## Residual            0.18857  0.4342
## Number of obs: 345, groups:  SPK_id, 27; vowels, 9
##

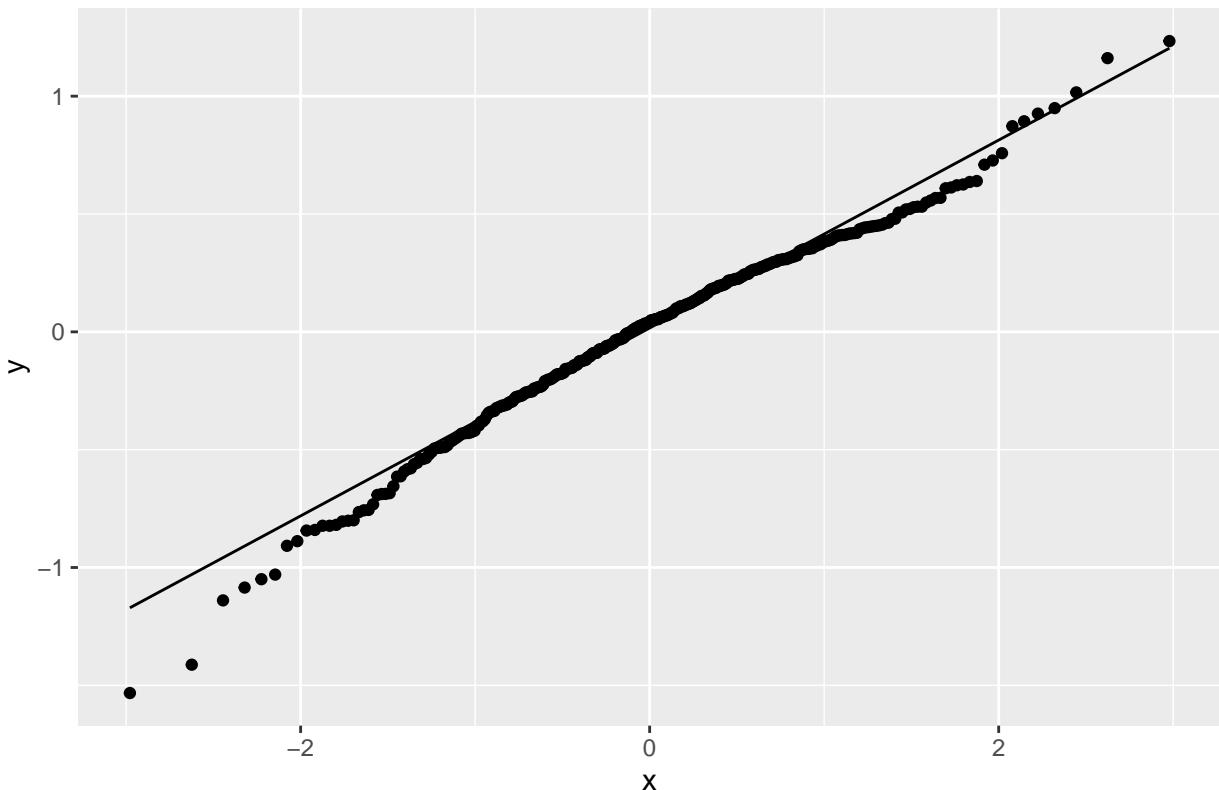
```

```

## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 10.96049   0.09312 17.26193 117.698 <2e-16 ***
## SES         -0.03967   0.05731 20.65243  -0.692  0.4965
## SEXM        -0.14221   0.11754 24.90314  -1.210  0.2377
## AgeInDays    0.01760   0.05781 19.47623   0.304  0.7640
## SEXM:AgeInDays -0.18319   0.10128 22.58276  -1.809  0.0838 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SES     SEXM   AgInDy
## SES       -0.171
## SEXM      -0.404  0.415
## AgeInDays -0.074  0.050  0.073
## SEXM:AgInDy  0.044 -0.043 -0.042 -0.571
## Reduced Model: Log-Likelihood = -237.78 , Degrees of Freedom = 10
## Max VIF value of the Reduced Model: 1.48601134384429
## Fitting Null Model.
## boundary (singular) fit: see help('isSingular')

```

Q-Q Plot of Residuals for Reduced Model



```

## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## red_fit_model: redmformula
## null_fit_model: nullmformula

```

```

##          npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)
## red_fit_model    10 495.56 534.00 -237.78    475.56
## null_fit_model   10 498.92 537.35 -239.46    478.92     0  0
## Returning the fitted models, data, optimizer settings, and emmeans (if computed).

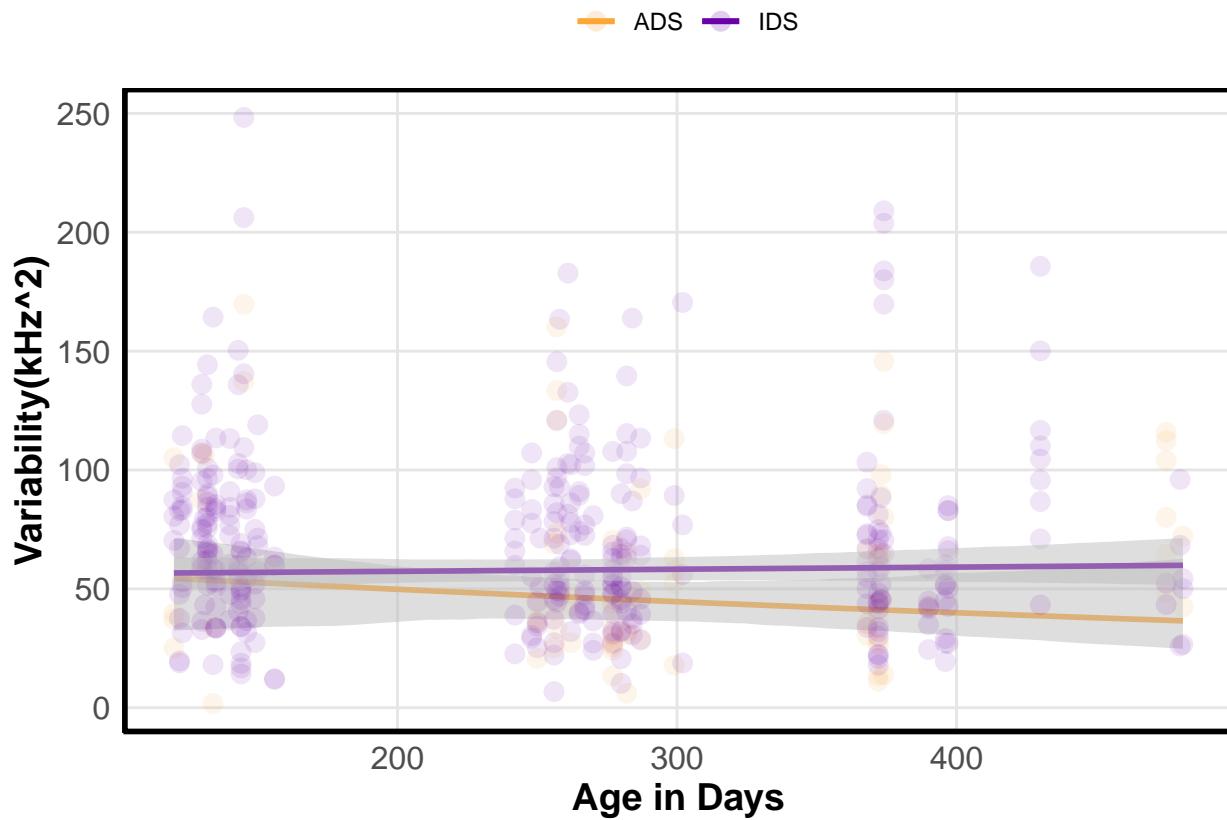
pred.data.6.2 <- boot.ci.predict.lmer(
  m = output.6.2$reduced_model,
  optimizer = output.6.2$optimizer,
  maxfun = output.6.2$maxfun,
  data = output.6.2$fit_data,
  pred.data = expand.grid(AgeInDays = seq(min(output.6.2$fit_data$AgeInDays), max(output.6.2$fit_data$AgeInDays)),
  reqcol = c("AgeInDays"),
  centercol = c("SES", 'SEX'),
  nboots = 1000,
  link = "identity",
  keep.boots = FALSE
)

## Number of singular fits during bootstrapping: 0 out of 1000 bootstrap iterations.

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.6.1$fit_data$AgeInDays, "scaled:scale") + attr(output.6.1$fit_data$AgeInDays, "scaled:center")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.6.2$fit_data$AgeInDays, "scaled:scale") + attr(output.6.2$fit_data$AgeInDays, "scaled:center")
}
inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

var_vow <- continuous_fit_ci_plot(
  plot.data = list(output.6.1$fit_data, output.6.2$fit_data),
  coefs = list(pred.data.6.1,pred.data.6.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "variability",
  x.labs = "Age in Days",
  y.labs = "Variability(kHz^2)",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  div = 1000,
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
  y.ax.transformation = TRUE,
  inv_y_transform_fun = list(inv_y_transform_fun_1, inv_y_transform_fun_2),
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  #colors = c("#6A00A8"),
  #alpha_level = 0.3,
  legend.labels = c( "ADS","IDS")
)
print(var_vow)

```



#7. Distinctiveness #7.1 ADS

```
df_ADS <- read_excel(file.path(getwd(), "acoustic_measures", "ADS_dist.xlsx"))

## New names:
## * ` ` -> `...1`

output.7.1 <- glmmTB.full.reduced.null.compare(Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                                 family = beta_family(link = "logit"),
                                                 zi_formula = ~1,
                                                 fullm = FALSE,
                                                 redm = TRUE,
                                                 redmformula = Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
                                                 nullm = TRUE,
                                                 nullmformula = Vowel_Dist ~ SES + SEX + (1|SPK_id),
                                                 df_ADS)

## 
## Attaching package: 'performance'
## 
## The following object is masked from 'package:arm':
## 
##     display

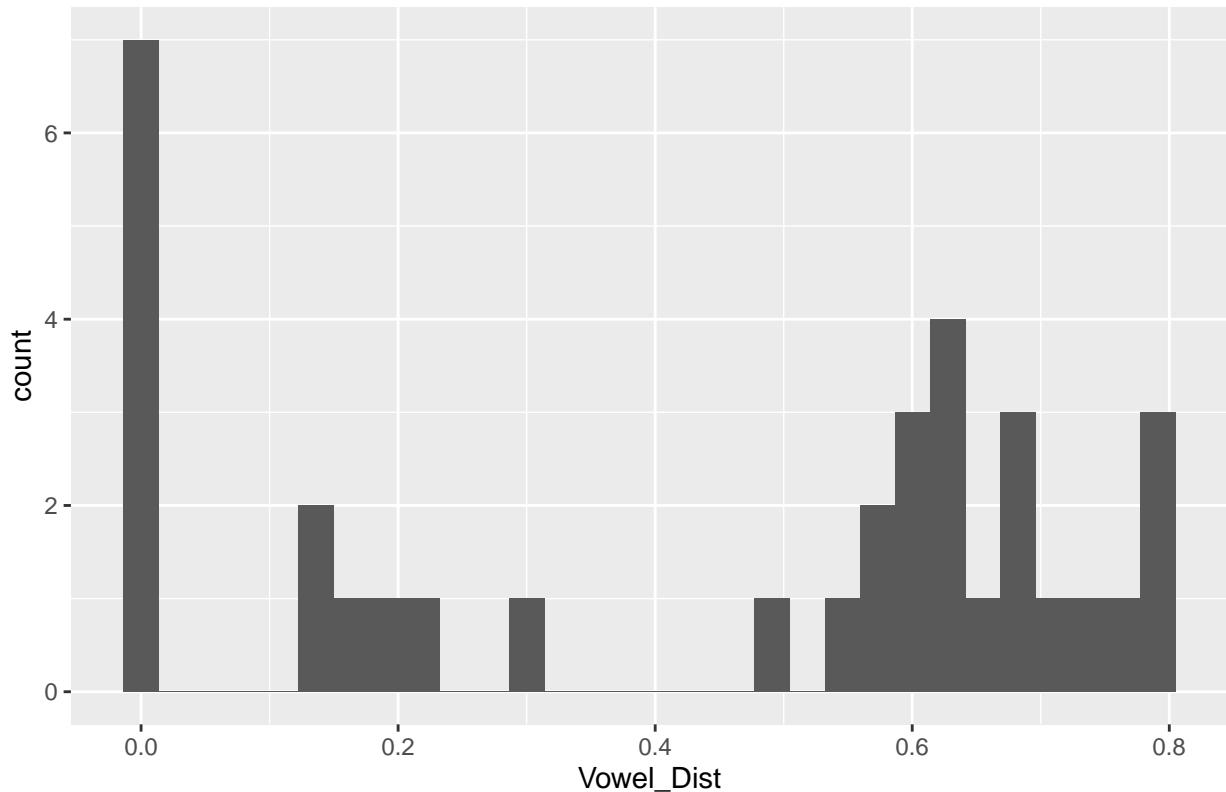
## Structure of combined data:
## tibble [34 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:34] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id    : chr [1:34] "C023" "C028" "C028" "C031" ...
```

```

## $ AgeInDays : num [1:34] 251 140 287 373 143 282 374 250 390 131 ...
## $ SES       : num [1:34] 5 5 5 8 5 5 5 5 5 4 ...
## $ SEX       : chr [1:34] "M" "F" "F" "F" ...
## $ Vowel_Dist: chr [1:34] "0.7914" "0.0000" "0.6346" "0.2877" ...
## [1] "Vowel_Dist"
## Structure of combined data after scaling:
## tibble [34 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1      : num [1:34, 1] -1.66 -1.56 -1.46 -1.36 -1.26 ...
## ..- attr(*, "scaled:center")= num 16.5
## ..- attr(*, "scaled:scale")= num 9.96
## $ SPK_id    : chr [1:34] "C023" "C028" "C028" "C031" ...
## $ AgeInDays : num [1:34, 1] -0.121 -1.21 0.232 1.076 -1.181 ...
## ..- attr(*, "scaled:center")= num 263
## ..- attr(*, "scaled:scale")= num 102
## $ SES       : num [1:34, 1] 0 0 0 2.15 0 ...
## ..- attr(*, "scaled:center")= num 5
## ..- attr(*, "scaled:scale")= num 1.39
## $ SEX       : chr [1:34] "M" "F" "F" "F" ...
## $ Vowel_Dist: num [1:34] 0.791 0 0.635 0.288 0 ...

```

Histogram of Response Variable



```

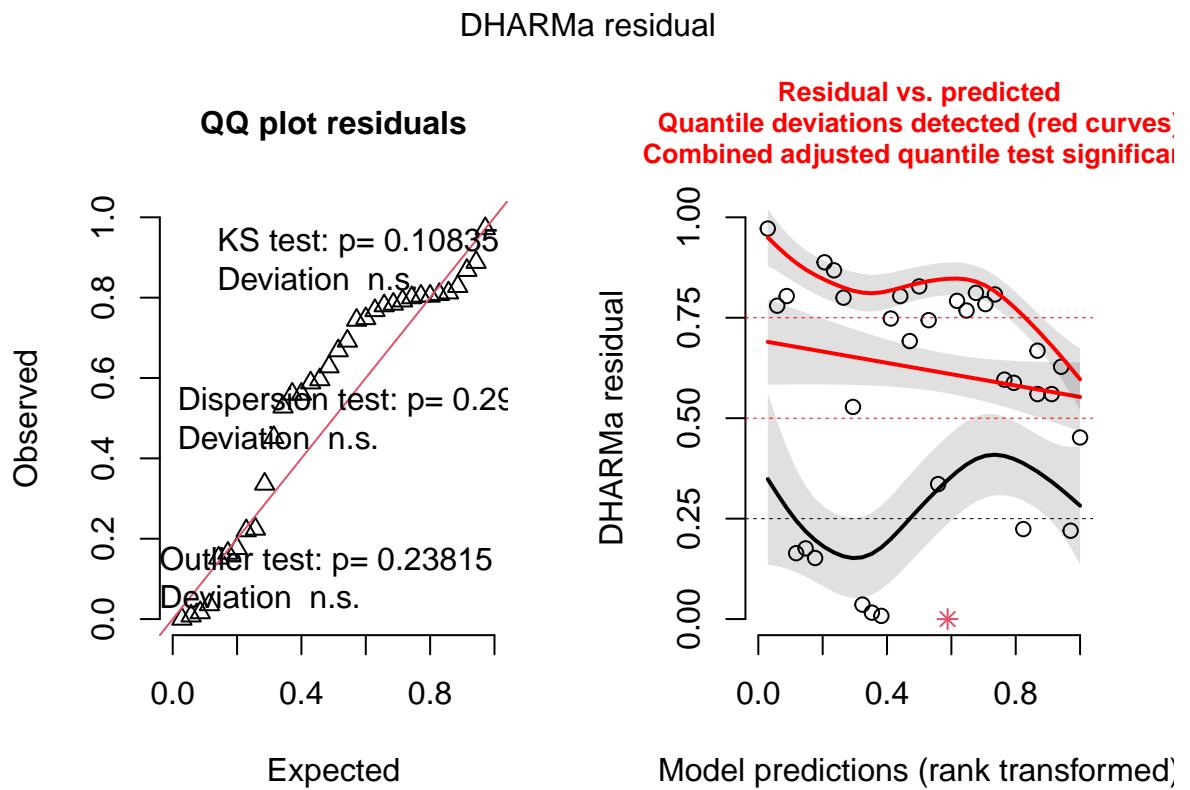
## Fitting Reduced Model.
## Summary of the reduced model:
## Family: beta  ( logit )
## Formula:
## Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1 | SPK_id)
## Zero inflation:           ~1
## Data: combined_data

```

```

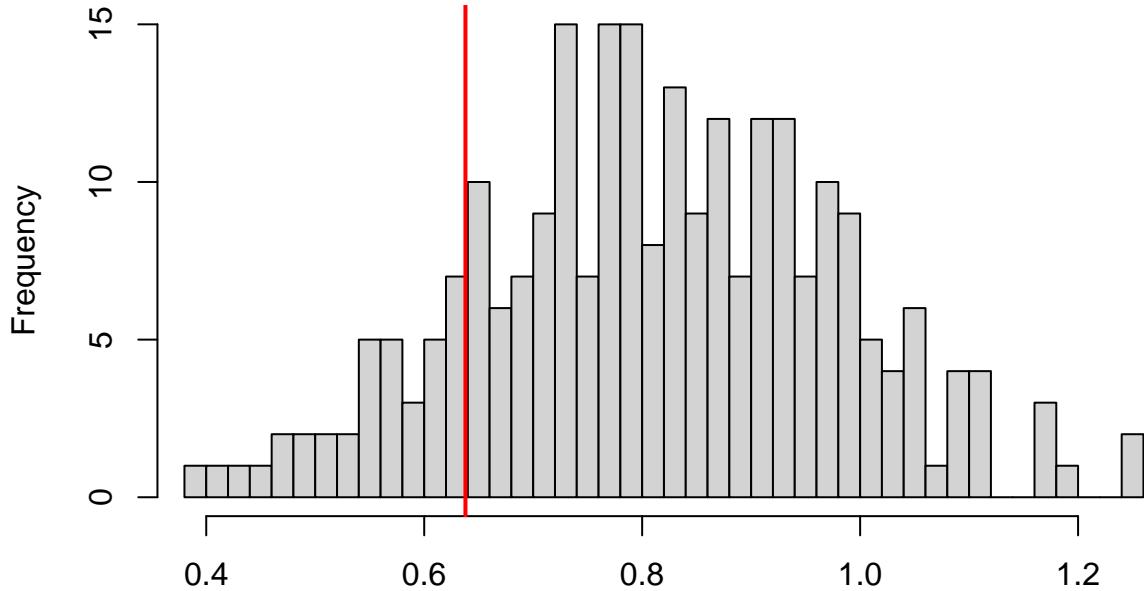
##          AIC      BIC logLik deviance df.resid
##     -83.6    -71.4    49.8    -99.6      26
##
## Random effects:
## 
## Conditional model:
##   Groups Name        Variance Std.Dev.
##   SPK_id (Intercept) 0.6488   0.8055
##   Number of obs: 34, groups: SPK_id, 22
##
## Dispersion parameter for beta family (): 1.76
## 
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.92033  0.35799 -2.571  0.0101 *
## SES          0.07304  0.30780  0.237  0.8124
## SEXM         0.43059  0.67760  0.635  0.5251
## AgeInDays    1.08379  0.34721  3.121  0.0018 **
## SEXM:AgeInDays -0.93345  0.51780 -1.803  0.0714 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -22.87    15871.68 -0.001   0.999
## 
## Simulate residuals for the fitted reduced glmmTMB model.

```



```
## Check for overdispersion with DHARMA test for reduced model.
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated



Simulated values, red line = fitted model. p-value (two.sided) = 0.296

```

## DHARMA dispersion for reduced model test results:
## Dispersion: 0.7838414
## p-value: 0.296
## Alternative hypothesis: two.sided
## VIF for the reduced model:
## # Check for Multicollinearity
##
## * conditional component:
##
## Low Correlation
##
##           Term      VIF      VIF 95% CI Increased SE Tolerance Tolerance 95% CI
##           SES 1.37 [1.12, 2.18]      1.17      0.73 [0.46, 0.90]
##           SEX 1.31 [1.08, 2.12]      1.14      0.76 [0.47, 0.92]
##           AgeInDays 2.20 [1.59, 3.41]      1.48      0.46 [0.29, 0.63]
##           SEX:AgeInDays 2.14 [1.56, 3.32]      1.46      0.47 [0.30, 0.64]
## Reduced Model: Log-Likelihood = 49.81 , Degrees of Freedom = 8
## Fitting Null Model.
## Comparing Reduced and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: Vowel_Dist ~ SES + SEX + (1 | SPK_id), zi=~1, disp=~1
## red_fit_model: Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1 | SPK_id), zi=~1, disp=~1
##               Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## null_fit_model 6 -77.499 -68.340 44.749 -89.499
## red_fit_model  8 -83.626 -71.416 49.813 -99.626 10.128      2 0.006321 **

```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
pred.data.7.1 <- NULL
for (i in seq(1, 1000, by = 100)) {
  batch_result <- boot.ci.predict.glmmTMB(
    m = output.7.1$reduced_model,
    data = output.7.1$fit_data,
    pred.data = expand.grid(AgeInDays = seq(min(output.7.1$fit_data$AgeInDays), max(output.7.1$fit_data$AgeInDays),
      reqcol = c("AgeInDays"),
      centercol = c("SES", "SEX"),
      nboots = min(100, 1000 - (i - 1)),
      keep.boots = FALSE
    )
    pred.data.7.1 <- rbind(pred.data.7.1, batch_result)
  }
}

## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 100 bootstrap iterations.

#7.2 IDS
df_IDS <- read_excel(file.path(getwd(), "acoustic_measures", "IDS_dist.xlsx"))

## New names:
## * ` ` -> `...`1` 

output.7.2 <- glmmTB.full.reduced.null.compare(Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1+SPK_id|AgeInDays),
  family = beta_family(link = "logit"),
  zi_formula = ~1,
  fullm = TRUE,
  redm = FALSE,
  redmformula = Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1|SPK_id),
  nullm = TRUE,
  nullmformula = Vowel_Dist ~ SES + SEX + (1+AgeInDays|SPK_id),
  df_IDS)

## Structure of combined data:
## tibble [56 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:56] 0 1 2 3 4 5 6 7 8 9 ...
## $ SPK_id : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays : num [1:56] 280 267 135 251 140 287 135 373 143 282 ...
## $ SES : num [1:56] 5 5 5 5 5 5 8 8 5 5 ...
## $ SEX : chr [1:56] "M" "F" "M" "M" ...
## $ Vowel_Dist: chr [1:56] "0.5685" "0.5816" "0.6137" "0.6127" ...
## [1] "Vowel_Dist"
## Structure of combined data after scaling:
## tibble [56 x 6] (S3: tbl_df/tbl/data.frame)
## $ ...1 : num [1:56, 1] -1.69 -1.62 -1.56 -1.5 -1.44 ...

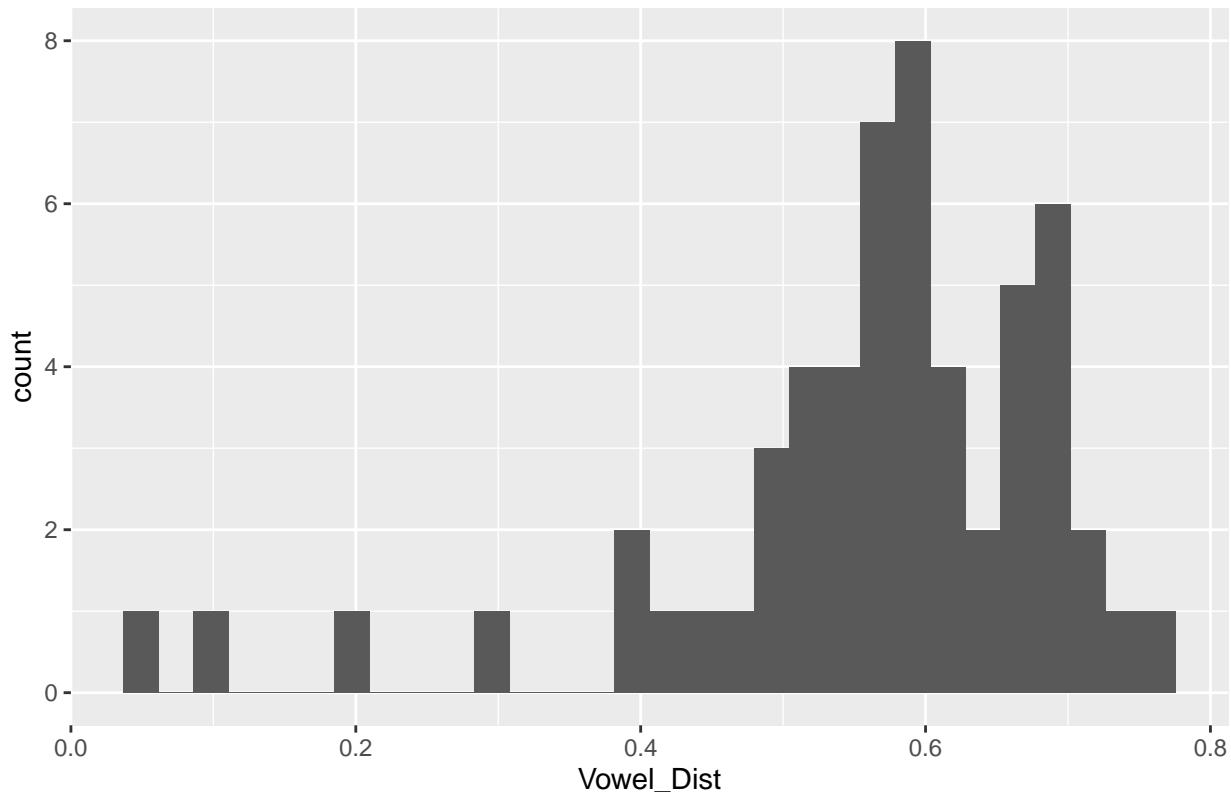
```

```

## ..- attr(*, "scaled:center")= num 27.5
## ..- attr(*, "scaled:scale")= num 16.3
## $ SPK_id : chr [1:56] "C010" "C012" "C023" "C023" ...
## $ AgeInDays : num [1:56, 1] 0.2314 0.1093 -1.1303 -0.0409 -1.0834 ...
## ..- attr(*, "scaled:center")= num 255
## ..- attr(*, "scaled:scale")= num 106
## $ SES      : num [1:56, 1] -0.0814 -0.0814 -0.0814 -0.0814 -0.0814 ...
## ..- attr(*, "scaled:center")= num 5.11
## ..- attr(*, "scaled:scale")= num 1.32
## $ SEX      : chr [1:56] "M" "F" "M" "M" ...
## $ Vowel_Dist: num [1:56] 0.568 0.582 0.614 0.613 0.583 ...

```

Histogram of Response Variable



```

## Fitting Full Model.
## Summary of the full model:
## Family: beta  ( logit )
## Formula:
## Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1 + AgeInDays |
##           SPK_id)
## Zero inflation:          ~1
## Data: combined_data
##
##      AIC      BIC   logLik deviance df.resid
##     -52.5    -32.3     36.3    -72.5      46
## 
## Random effects:
## 
## Conditional model:

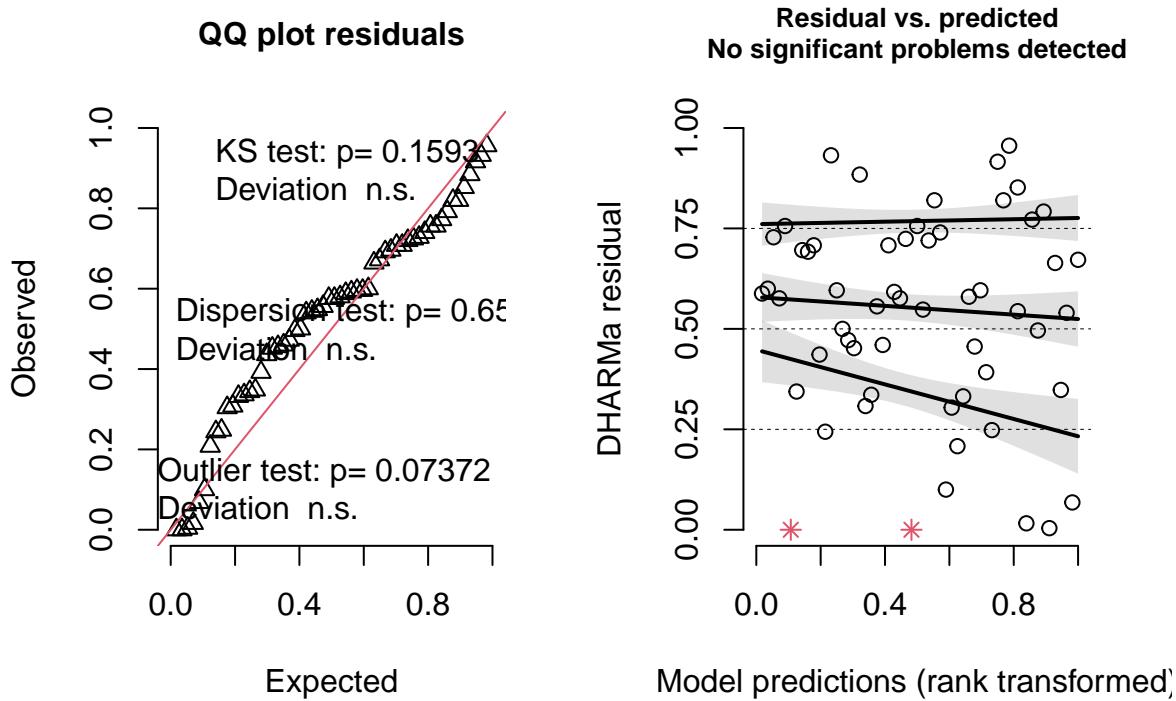
```

```

## Groups Name      Variance Std.Dev. Corr
## SPK_id (Intercept) 0.1883   0.4339
##           AgeInDays  0.1638   0.4048  -0.46
## Number of obs: 56, groups: SPK_id, 27
##
## Dispersion parameter for beta family (): 59.2
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 0.27106  0.11835  2.290  0.022 *
## SES        -0.01312  0.10266 -0.128  0.898
## SEXM       -0.07049  0.21747 -0.324  0.746
## AgeInDays   0.02747  0.12321  0.223  0.824
## SEXM:AgeInDays 0.21240  0.20776  1.022  0.307
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -23.3     15344.0 -0.002    0.999
## Simulate residuals for the fitted full glmmTMB model.

```

DHARMA residual

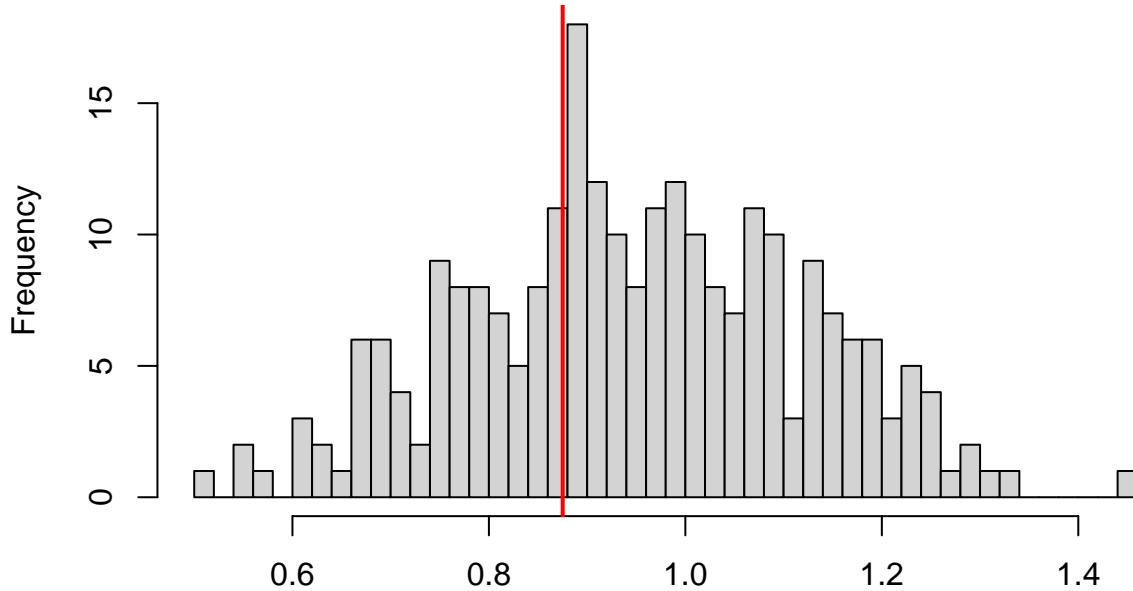


```

## Check for overdispersion with DHARMA test for full model.

```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated



Simulated values, red line = fitted model. p-value (two.sided) = 0.656

```

## DHARMA dispersion for full model test results:
## Dispersion: 0.9229295
## p-value: 0.656
## Alternative hypothesis: two.sided
## VIF for the full model:
## # Check for Multicollinearity
##
## * conditional component:
##
## Low Correlation
##
##           Term   VIF   VIF 95% CI Increased SE Tolerance Tolerance 95% CI
##           SES 1.18 [1.03, 1.94]      1.09      0.85 [0.51, 0.97]
##           SEX 1.28 [1.08, 1.95]      1.13      0.78 [0.51, 0.92]
##           AgeInDays 1.55 [1.24, 2.26]     1.24      0.65 [0.44, 0.81]
##           SEX:AgeInDays 1.64 [1.30, 2.40]     1.28      0.61 [0.42, 0.77]
## Full Model: Log-Likelihood = 36.27 , Degrees of Freedom = 10
## Fitting Null Model.
## Comparing Full and Null Models using ANOVA (Likelihood Ratio Test):
## Data: combined_data
## Models:
## null_fit_model: Vowel_Dist ~ SES + SEX + (1 + AgeInDays | SPK_id), zi=~1, disp=~1
## fit_model: Vowel_Dist ~ SES + SEX + AgeInDays + AgeInDays:SEX + (1 + AgeInDays | , zi=~1, disp=~1
## fit_model:      SPK_id), zi=~1, disp=~1
##          Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## null_fit_model 8 -54.517 -38.314 35.258 -70.517

```

```

## fit_model      10 -52.542 -32.289 36.271  -72.542 2.0255      2      0.3632
pred.data.7.2 <- NULL
for (i in seq(1, 1000, by = 50)) {
  batch_result <- boot.ci.predict.glmmTMB(
    m = output.7.2$full_model,
    data = output.7.2$fit_data,
    pred.data = expand.grid(AgeInDays = seq(min(output.7.2$fit_data$AgeInDays), max(output.7.2$fit_data$AgeInDays),
    reqcol = c("AgeInDays"),
    centercol = c("SES", 'SEX'),
    nboots = min(50, 1000 - (i - 1)),
    keep.boots = FALSE
  )
  pred.data.7.2 <- rbind(pred.data.7.2, batch_result)
}

## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.
## Number of singular fits during bootstrapping: 0 out of 50 bootstrap iterations.

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.7.1$fit_data$AgeInDays, "scaled:scale") + attr(output.7.1$fit_data$AgeInDays, "scaled:center")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.7.2$fit_data$AgeInDays, "scaled:scale") + attr(output.7.2$fit_data$AgeInDays, "scaled:center")
}
#inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
#inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

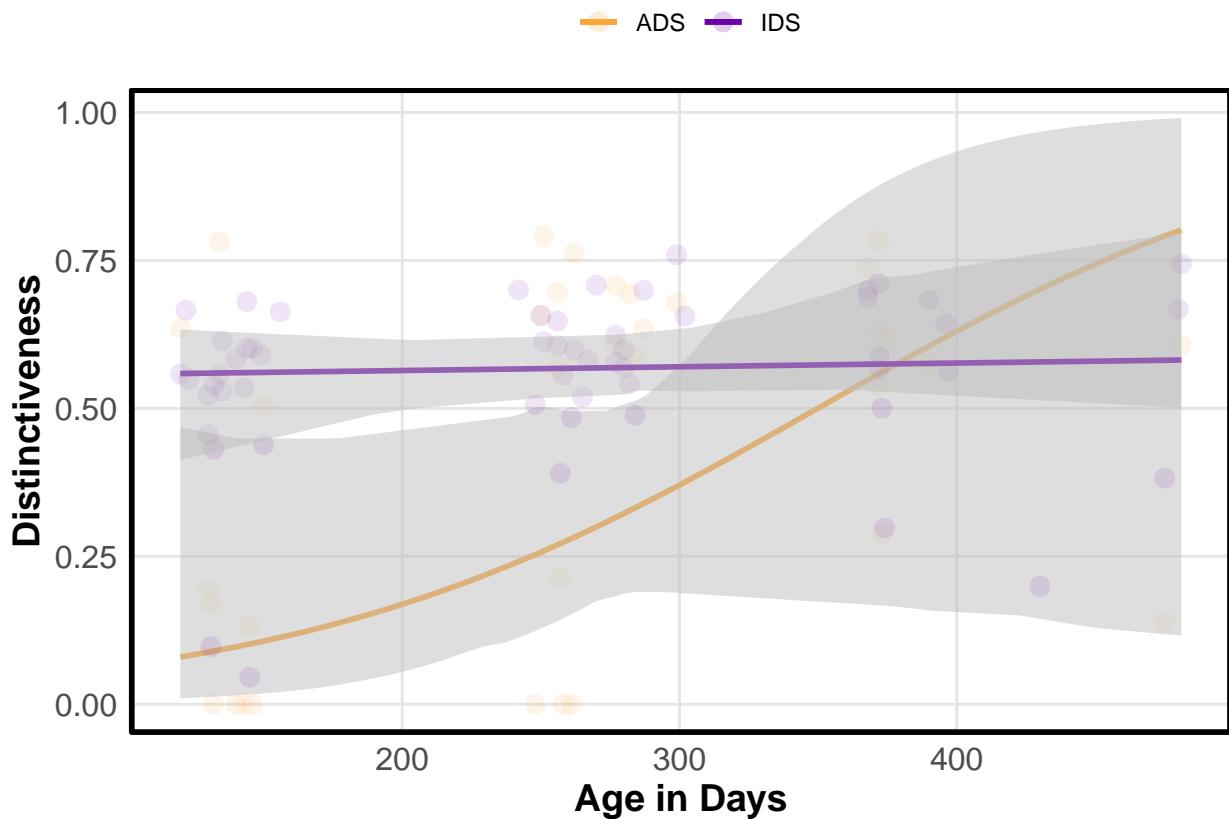
distinctiveness <- continuous_fit_ci_plot(
  plot.data = list(output.7.1$fit_data, output.7.2$fit_data),
  coefs = list(pred.data.7.1, pred.data.7.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "Vowel_Dist",
  x.labs = "Age in Days",
  y.labs = "Distinctiveness",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanFOst), max(output.2.1$fit_data$meanFOst)),

```

```

x.ax.transformation = TRUE,
inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
y.ax.transformation = FALSE,
inv_y_transform_fun = NULL,
#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(40, 60, by = 10),
#colors = c("#6A00A8"),
#alpha_level = 0.3,
legend.labels = c("ADS", "IDS")
)
print(distinctiveness)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.7.1$fit_data$AgeInDays, "scaled:scale") + attr(output.7.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.7.2$fit_data$AgeInDays, "scaled:scale") + attr(output.7.2$fit_data$AgeInDays, "scaled")
}
#inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
#inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

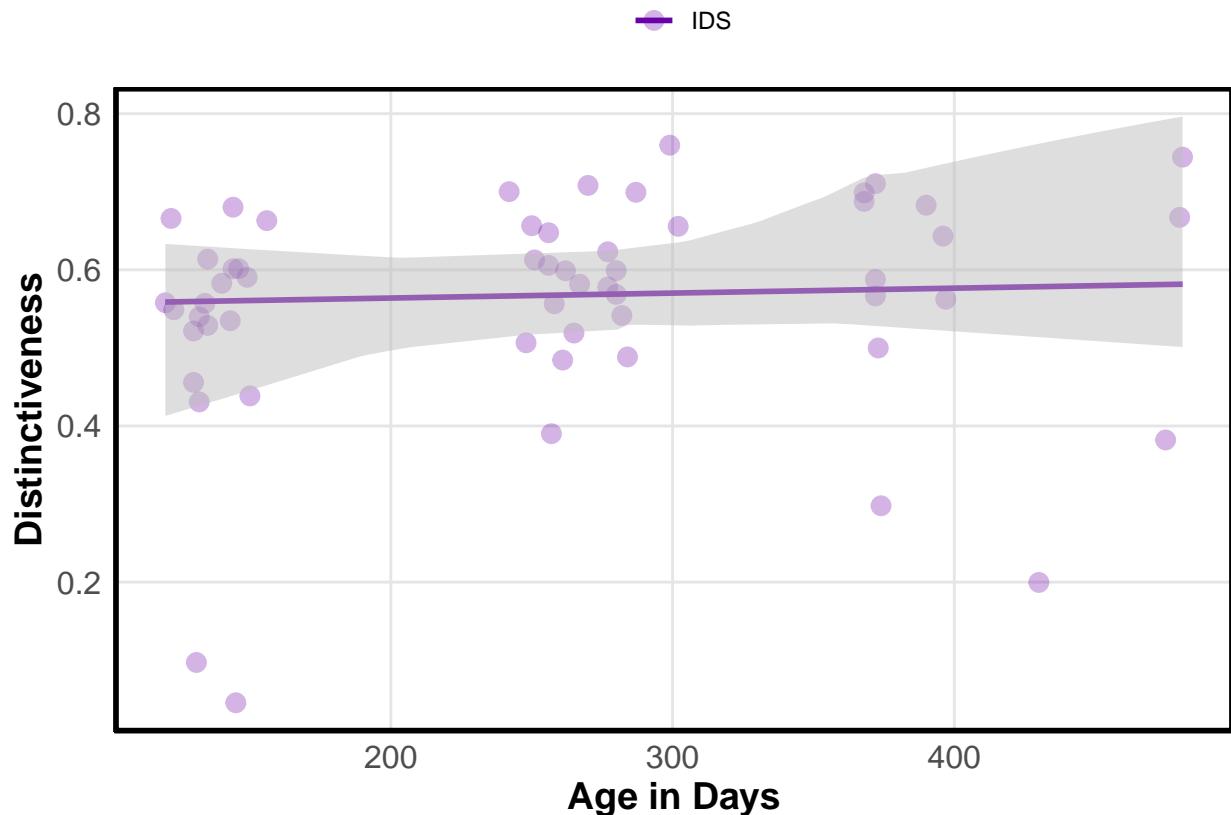
distinctiveness_IDS <- continuous_fit_ci_plot(
  plot.data = list(output.7.2$fit_data),
  coefs = list(pred.data.7.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "Vowel_Dist",
  x_label = "Age in Days",
  y_label = "Distinctiveness"
)

```

```

x.labs = "Age in Days",
y.labs = "Distinctiveness",
#x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
#y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
x.ax.transformation = TRUE,
inv_x_transform_fun = list(inv_x_transform_fun_2),
y.ax.transformation = FALSE,
inv_y_transform_fun = NULL,
#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(40, 60, by = 10),
colors = c("#6A00A8"),
alpha_level = 0.3,
legend.labels = c("IDS")
)
print(distinctiveness_IDS)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.7.1$fit_data$AgeInDays, "scaled:scale") + attr(output.7.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.7.2$fit_data$AgeInDays, "scaled:scale") + attr(output.7.2$fit_data$AgeInDays, "scaled")
}
#inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
#inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

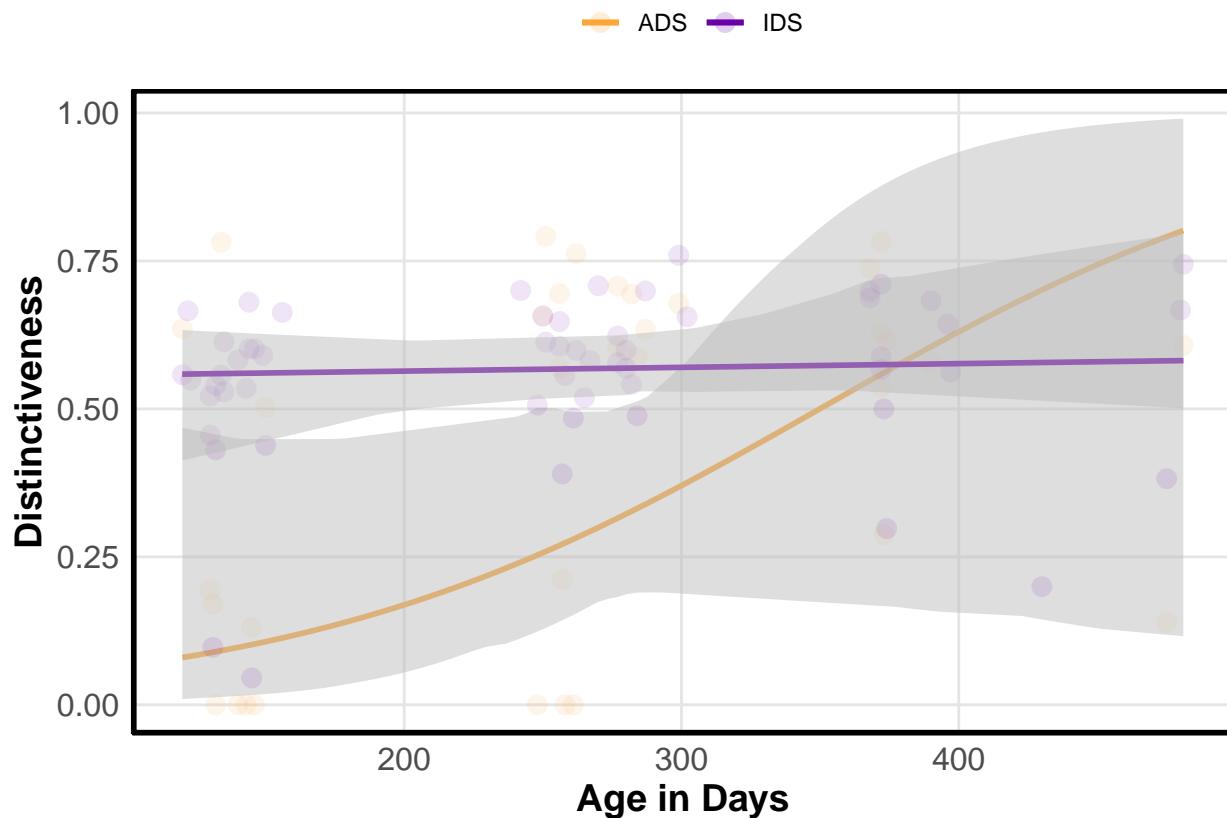
distinctiveness <- continuous_fit_ci_plot(

```

```

plot.data = list(output.7.1$fit_data, output.7.2$fit_data),
coefs = list(pred.data.7.1, pred.data.7.2),
plot.data.col.x = "AgeInDays",
plot.data.col.y = "Vowel_Dist",
x.labs = "Age in Days",
y.labs = "Distinctiveness",
#x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
#y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
x.ax.transformation = TRUE,
inv_x_transform_fun = list(inv_x_transform_fun_1, inv_x_transform_fun_2),
y.ax.transformation = FALSE,
inv_y_transform_fun = NULL,
#x.breaks = seq(100, 400, by = 100),
#y.breaks = seq(40, 60, by = 10),
#colors = c("#6A00A8"),
#alpha_level = 0.3,
legend.labels = c("ADS", "IDS")
)
print(distinctiveness)

```



```

inv_x_transform_fun_1 <- function(x) {
  x * attr(output.7.1$fit_data$AgeInDays, "scaled:scale") + attr(output.7.1$fit_data$AgeInDays, "scaled")
}
inv_x_transform_fun_2 <- function(x) {
  x * attr(output.7.2$fit_data$AgeInDays, "scaled:scale") + attr(output.7.2$fit_data$AgeInDays, "scaled")
}

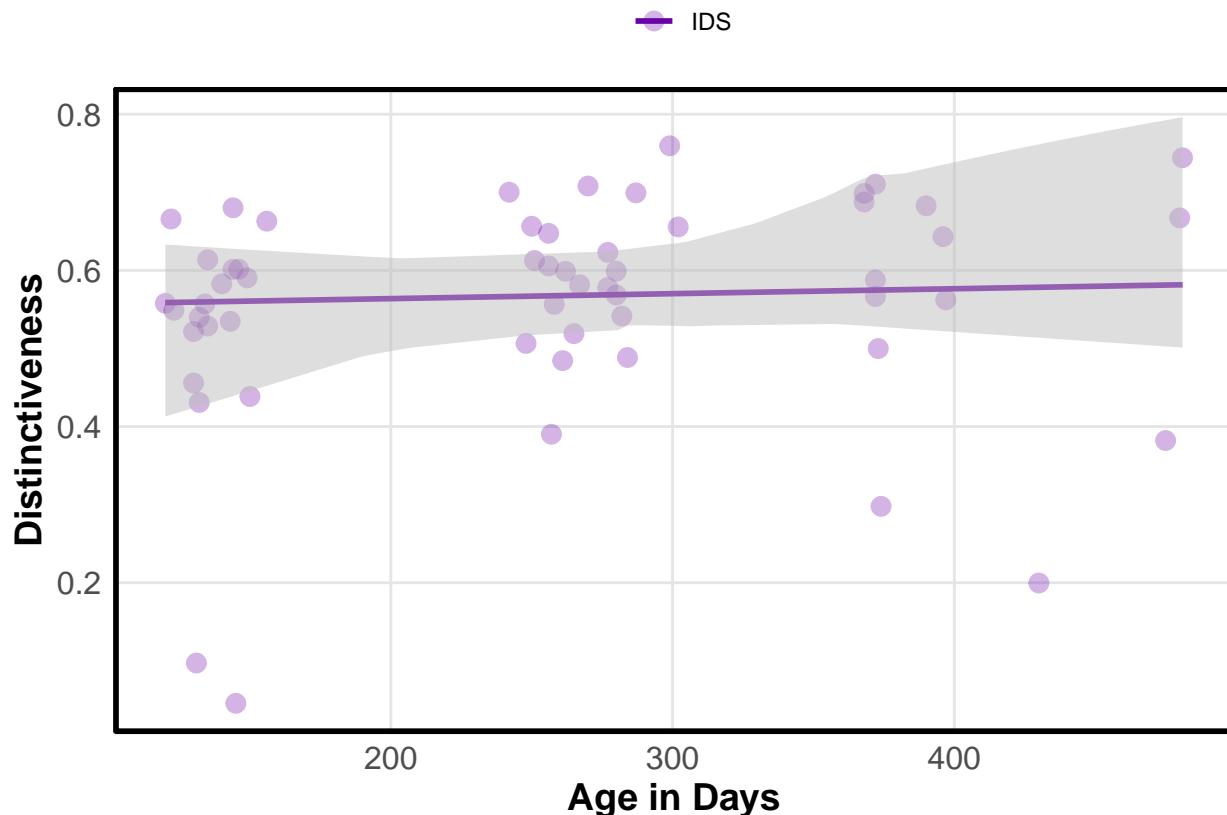
```

```

#inv_y_transform_fun_1 <- function(x) { exp(x) - 1 }
#inv_y_transform_fun_2 <- function(x) { exp(x) - 1 }

distinctiveness_IDS <- continuous_fit_ci_plot(
  plot.data = list(output.7.2$fit_data),
  coefs = list(pred.data.7.2),
  plot.data.col.x = "AgeInDays",
  plot.data.col.y = "Vowel_Dist",
  x.labs = "Age in Days",
  y.labs = "Distinctiveness",
  #x.lim = c(120,500),#c(min(pred.data.2.1$AgeInDays), max(pred.data.2.1$AgeInDays)),
  #y.lim = c(40,100),#c(min(output.2.1$fit_data$meanF0st), max(output.2.1$fit_data$meanF0st)),
  x.ax.transformation = TRUE,
  inv_x_transform_fun = list(inv_x_transform_fun_2),
  y.ax.transformation = FALSE,
  inv_y_transform_fun = NULL,
  #x.breaks = seq(100, 400, by = 100),
  #y.breaks = seq(40, 60, by = 10),
  colors = c("#6A00A8"),
  alpha_level = 0.3,
  legend.labels = c("IDS")
)
print(distinctiveness_IDS)

```



duration_ms_log_grid

```

ggsave(file.path(getwd(), "StatPlots", "pitch.pdf"), plot = Pitch_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_outlier.pdf"), plot = pitch_range_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_outlier_log_grid.pdf"), plot = pitch_range_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_no_outlier.pdf"), plot = pitch_range_semitones_no_outlier, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "duration.pdf"), plot = duration_ms, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "duration_log_grid.pdf"), plot = duration_ms_log_grid, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Area_corner_vowels.pdf"), plot = Area_cor, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Are_full.pdf"), plot = Area_full, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Variability.pdf"), plot = var_vow, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Distinctiveness.pdf"), plot = distinctiveness, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch.jpg"), plot = Pitch_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_outlier.jpg"), plot = pitch_range_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_outlier_log_grid.jpg"), plot = pitch_range_semitones, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "pitch_range_no_outlier.jpg"), plot = pitch_range_semitones_no_outlier, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "duration.jpg"), plot = duration_ms, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "duration_log_grid.jpg"), plot = duration_ms_log_grid, height = 5)

```

```
ggsave(file.path(getwd(), "StatPlots", "Area_corner_vowels.jpg"), plot = Area_cor, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Are_full.jpg"), plot = Area_full, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Variability.jpg"), plot = var_vow, height = 5)

## Saving 6.5 x 5 in image
ggsave(file.path(getwd(), "StatPlots", "Distinctiveness.jpg"), plot = distinctiveness, height = 5)

## Saving 6.5 x 5 in image
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.