Retrospective Longitudinal Analysis of Spectral Features Reveals Divergent Vocal Development Patterns for Treble and Non-Treble Singers

Reuben Walker

(2025-07-11)

Contents


```
suppressPackageStartupMessages({
  suppressWarnings({
    library(ggplot2)
    library(dplyr)
    library(ggpubr)
```

```r
    library(sjPlot)
    library(lmerTest)
    library(ggeffects)
    library(ggResidpanel)
    library(performance)
    library(emmeans)
    library(knitr)
    library(patchwork)
    library(viridis)
  })
})
```

Research Questions:

Do classical singers develop similarly over the course of their conservatory training or does their spectral development exhibit differences based on a grouping in female/countertenor or male voice types?

Does breathiness decrease among female students with vocal training?

Statistical Analysis

In each of the nine cases incorporating our three dependent variables and three sung tasks, we performed a linear mixed model with the predictor variables years of study and Voice_Group (male or female/countertenor), with specific emphasis on the interaction between the two predictors. A control variable delta SPL (dSPL) was measured from a reference sample. To account for inter-subject variation in baseline and development, random slopes and intercepts were specified for the individual subjects. We iteratively simplified the models by removing non-significant terms beginning with the interaction term. A linear mixed model was then performed on the female subset for each sung task to investigate increases in CPPs. In order to adjust for multiple comparisons, we adjusted our significance level using the Benjamini-Hochberg correction. We calculated estimated marginal means from the resulting models to examine differences between Voice_Group at different intervals of training.

Sustained high phonation

Load and Prepare Data

```r
#Set path
#setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
#These are the high sustained samples
klang <- read.csv("Klang2_JASA_Final.csv", fileEncoding = "UTF-8")
klang <- klang[c("yearDiff", "H1H2LTAS", "alphaRatio", "CPPs", "Jahr", "Stimmfach", "Voice.Type", "gesch
names(klang)[names(klang) == "geschlecht"] <- "Gender"
names(klang)[names(klang) == "yearDiff"] <- "Years"
names(klang)[names(klang) == "alter"] <- "Age"
klang$Voice_Group[klang$Stimmfach == "Sop/Mezzo/Alt"] <- "Treble"
klang$Voice_Group[klang$Stimmfach == "Ten/Bar/Bass"] <- "Non-Treble"
klang$Gender[klang$Gender == "männl."] <- "Male"
klang$Gender[klang$Gender == "weibl."] <- "Female"
```

Filtering Data

```r
# Step 1: Count total unique male and female students
total_counts <- klang %>%
  group_by(Gender) %>%
  summarise(Total_Students = n_distinct(id))

# Step 2: Filter to only include students with Years <= 4
klang_filtered <- klang %>%
  filter(Years <= 4)

# Count unique male and female students after filtering
filtered_counts <- klang_filtered %>%
  group_by(Gender) %>%
  summarise(Filtered_Students = n_distinct(id))

# Step 3: Remove students with only one recording
id_counts <- klang_filtered %>%
  count(id)  # Count occurrences of each student (id)

valid_ids <- id_counts %>%
  filter(n > 1) %>%
  pull(id)  # Get list of students who appear more than once

klang_final <- klang_filtered %>%
  filter(id %in% valid_ids)  # Keep only students with multiple recordings

# Count unique male and female students after final filtering
final_counts <- klang_final %>%
  group_by(Gender) %>%
  summarise(Final_Students = n_distinct(id))

# Print results
print(total_counts)      # Total students by gender
```

```
## # A tibble: 2 x 2
##   Gender Total_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               49
```

```r
print(filtered_counts)   # After Years <= 4 restriction
```

```
## # A tibble: 2 x 2
##   Gender Filtered_Students
##   <chr>              <int>
## 1 Female                68
## 2 Male                  49
```

```r
print(final_counts)      # After removing single-recording students
```

```
## # A tibble: 2 x 2
##   Gender Final_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               49
```

```r
# Update the klang dataframe to keep only the final filtered version
klang <- klang_final
```

```r
# Split the data into Male/Female and remove Gender column
klang_f <- subset(klang, Gender %in% c("Female"))
# remove Gender
klang_f <- subset(klang_f, select = c("Years", "H1H2LTAS", "alphaRatio", "CPPs", "Jahr", "Gender", "Voi

klang_m <- subset(klang, Gender %in% c("Male"))
klang_m <- subset(klang_m, select = c("Years", "H1H2LTAS", "alphaRatio", "CPPs", "Jahr", "Gender", "Voi
```

H1H2LTAS

```r
# Here we take a look at the correlation matrix for our different metrics.
numeric_columns <- klang_f[, c("Years", "H1H2LTAS", "alphaRatio", "CPPs")]
# Omit na values:
cor(na.omit(numeric_columns))
```

```
##                 Years    H1H2LTAS  alphaRatio       CPPs
## Years      1.00000000  0.04864729  0.02583561  0.1431702
## H1H2LTAS   0.04864729  1.00000000 -0.71761967 -0.1208313
## alphaRatio 0.02583561 -0.71761967  1.00000000  0.2726838
## CPPs       0.14317016 -0.12083135  0.27268375  1.0000000
```

```r
# Repeat for male voices
numeric_columns <- klang_m[, c("Years", "H1H2LTAS", "alphaRatio", "CPPs")]
# Omit na values:
cor(na.omit(numeric_columns))
```

```
##                 Years   H1H2LTAS alphaRatio       CPPs
## Years       1.0000000 -0.1036603  0.1849967  0.1460398
## H1H2LTAS   -0.1036603  1.0000000 -0.2453336 -0.7426606
## alphaRatio  0.1849967 -0.2453336  1.0000000  0.3550278
## CPPs        0.1460398 -0.7426606  0.3550278  1.0000000
```

It makes intuitive sense that H1H2LTAS has a strong negative correlation to alpha ratio in women because H2 is above 1000 Hz. The direction of the Years correlations fits with the hypothesese: H1H2LTAS tends to increase for women, decrease for men Alpha Ratio increases for men, seems to remain constant for women CPPs tends to increase for both men and women

Here we take a look at our linear mixed models:

```r
#Let's make treble the reference group
klang$Voice_Group <- factor(klang$Voice_Group)  # convert to factor
klang$Voice_Group <- relevel(klang$Voice_Group, ref = "Treble")  # now relevel
```

```r
fit0_H <- lmer(H1H2LTAS~Years*Voice_Group + dSPL + (Years | id), data=klang, na.action="na.omit",REML =
anova(fit0_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value    Pr(>F)
## Years                0.60    0.60     1  83.496   0.3768  0.540964
## Voice_Group        461.35  461.35     1  91.190 291.4328 < 2.2e-16 ***
## dSPL                 2.42    2.42     1 231.049   1.5285  0.217589
## Years:Voice_Group   11.47   11.47     1  83.267   7.2439  0.008597 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We remove dSPL from the model as it was non-significant.

```r
fit1_H <- lmer(H1H2LTAS~Years*Voice_Group + (Years | id), data=klang, na.action="na.omit",REML = T)
anova(fit1_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value    Pr(>F)
## Years                0.79    0.79     1  83.211   0.4934  0.484368
## Voice_Group        468.35  468.35     1 100.751 294.2423 < 2.2e-16 ***
## Years:Voice_Group   11.81   11.81     1  83.211   7.4192  0.007862 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As the interaction is significant we leave years

```r
margin1 <- ggpredict(fit1_H, c("Years", "Voice_Group"), ci_level=0.95)
margin1
```

```
## # Predicted values of H1H2LTAS
##
## Voice_Group: Treble
##
## Years | Predicted |      95% CI
## -------------------------------
##     0 |      2.22 | 1.69,  2.76
##     1 |      2.52 | 2.07,  2.97
##     2 |      2.82 | 2.35,  3.28
##     3 |      3.11 | 2.54,  3.69
##     4 |      3.41 | 2.67,  4.15
##
## Voice_Group: Non-Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     -5.39 | -6.08, -4.70
##     1 |     -5.56 | -6.15, -4.97
##     2 |     -5.74 | -6.33, -5.14
##     3 |     -5.91 | -6.61, -5.21
##     4 |     -6.09 | -6.96, -5.21
```

```
##
## Adjusted for:
## * id = 0 (population-level)
```

The directionality of the two groups is clear (female increase, male decrease), though the confidence intervals overlap.

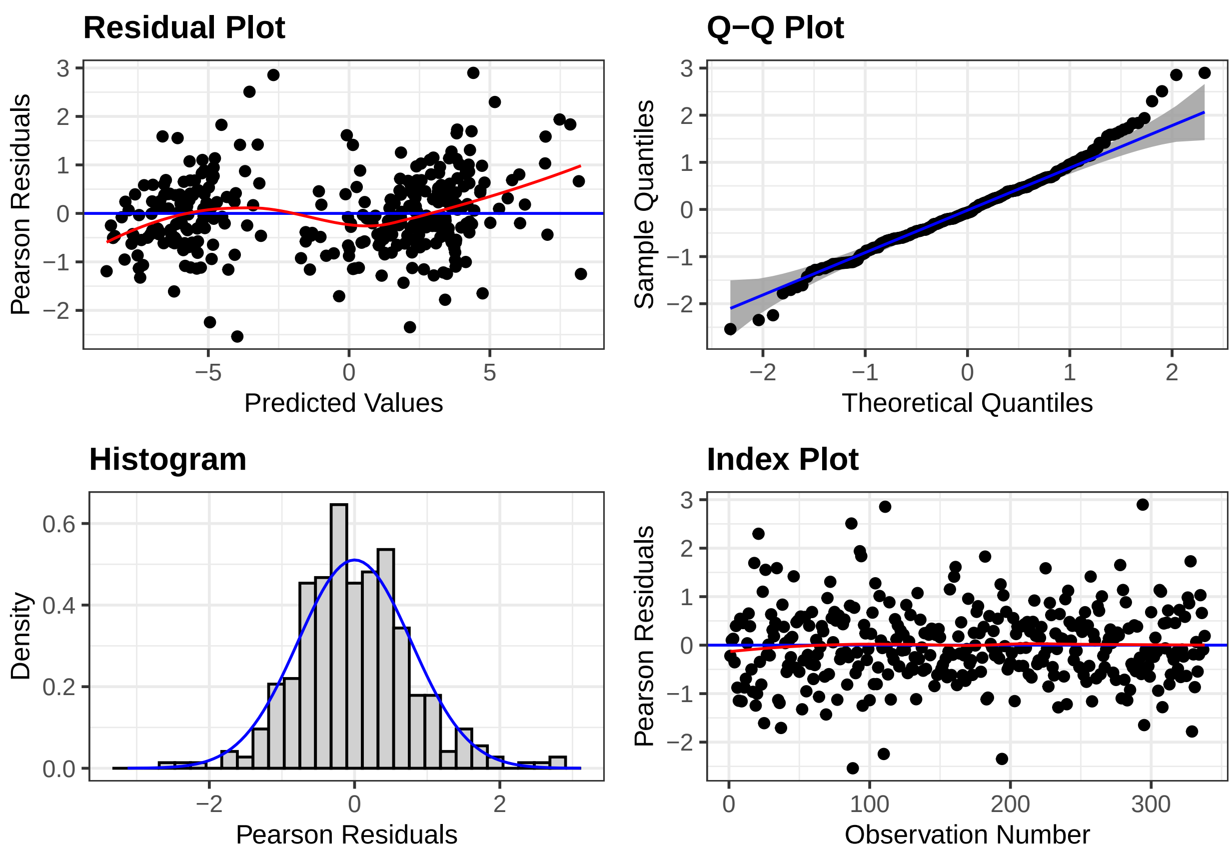
```
emm1_H <- emmeans::emmeans(fit1_H, ~ Years:Voice_Group)
pairs(emm1_H)
```

```
##  contrast                                                          estimate
##  Years1.73769960282078 Treble - (Years1.73769960282078 Non-Treble)     8.43
##     SE  df t.ratio p.value
##  0.374 112  22.513  <.0001
##
## Degrees-of-freedom method: kenward-roger
```

```
p.fit1_H <- ggResidpanel::resid_panel(fit1_H,
              plots = c("resid", "qq", "hist", "index"),
              smoother = TRUE,
              qqbands = TRUE,
              title.opt = TRUE)
```

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

```
p.fit1_H
```




```r
summary(fit1_H)
```

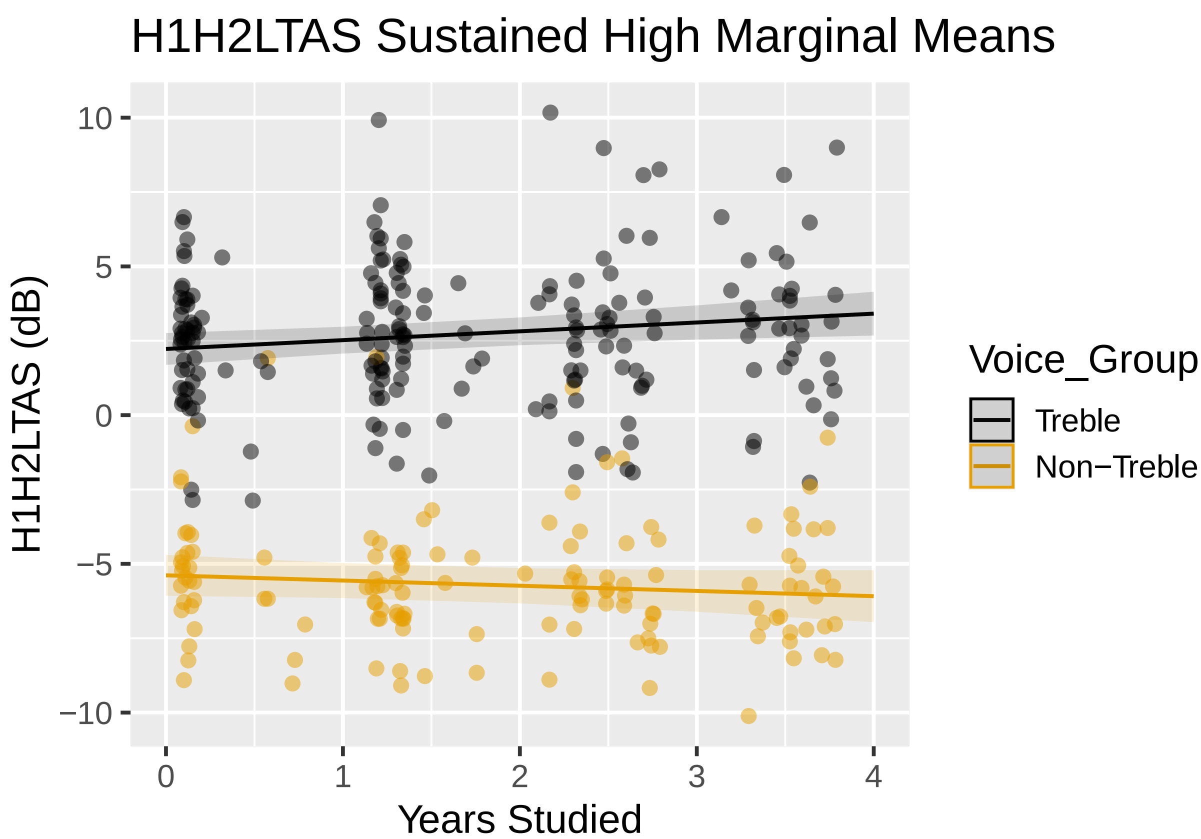
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: H1H2LTAS ~ Years * Voice_Group + (Years | id)
##    Data: klang
##
## REML criterion at convergence: 1381.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.53889 -0.48969 -0.06774  0.45594  2.89801
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 3.1747   1.7818
##           Years       0.2688   0.5185   -0.27
##  Residual             1.5917   1.2616
## Number of obs: 338, groups:  id, 117
##
## Fixed effects:
##                             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)                   2.2241     0.2719 102.2620   8.179 8.22e-13 ***
## Years                         0.2968     0.1127  98.2008   2.632  0.00985 **
## Voice_GroupNon-Treble        -7.6096     0.4436 100.7506 -17.153  < 2e-16 ***
```

```
## Years:Voice_GroupNon-Treble  -0.4719     0.1732  83.2113  -2.724  0.00786 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Years  V_GN-T
## Years       -0.558
## Vc_GrpNn-Tr -0.613  0.342
## Yrs:Vc_GN-T  0.363 -0.651 -0.552
```

```
fig_h1 <- plot_model(fit1_H, type = "pred", terms = c("Years", "Voice_Group"), show.data= TRUE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
        y = "H1H2LTAS (dB)",
        title = "H1H2LTAS Sustained High Marginal Means")
```

```
## Data points may overlap. Use the 'jitter' argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```
fig_h1
```



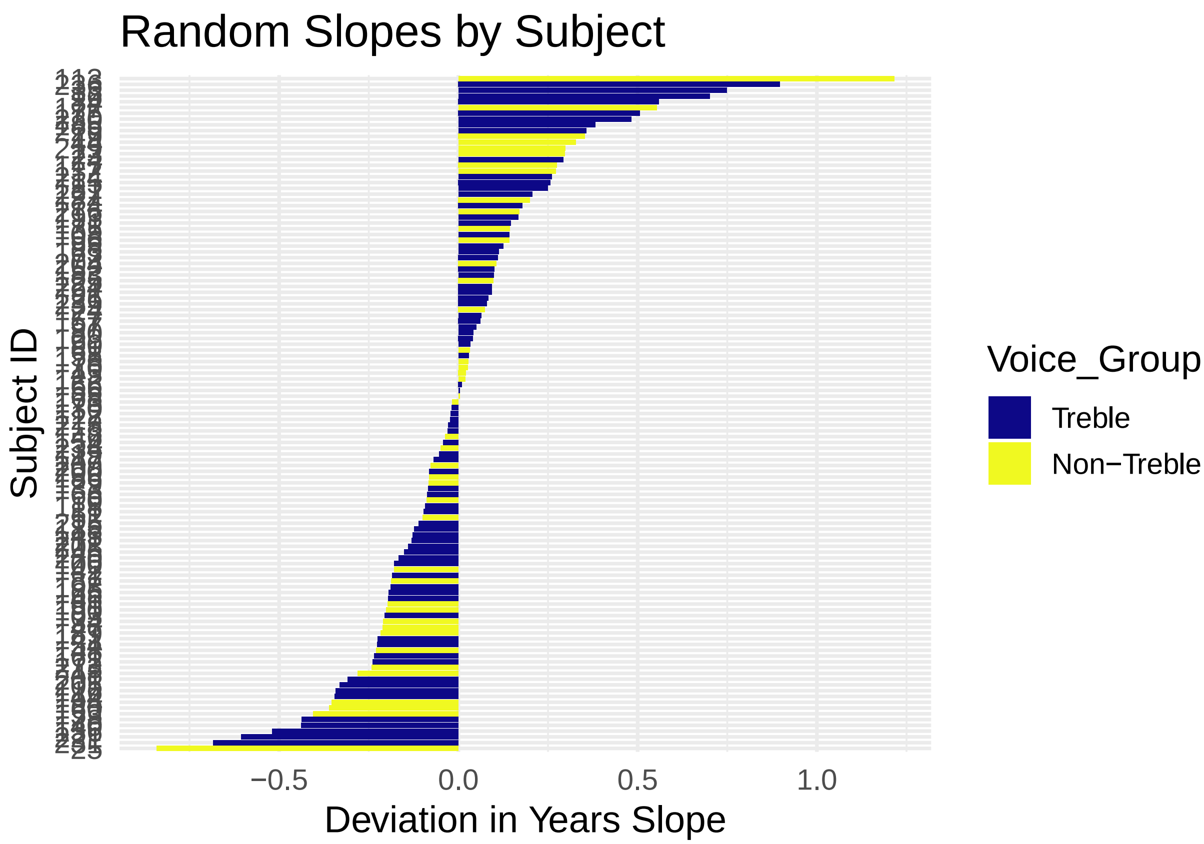

```
ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)
```

```r
performance::icc(fit1_H)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.688
##   Unadjusted ICC: 0.158
```

```r
#Individual Trajectories

# Clean and prep
klang_clean <- klang[!is.na(klang$H1H2LTAS), ]
klang_clean$fitted <- fitted(fit1_H)

# Ensure factors
klang_clean$id <- as.factor(klang_clean$id)
klang_clean$Voice_Group <- as.factor(klang_clean$Voice_Group)

fig_h1 <- plot_model(fit1_H, type = "pred", terms = c("Years", "Voice_Group"), show.data= FALSE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
        y = "H1H2LTAS (dB)",
        title = "H1H2LTAS Sustained High Individual Slopes")
#fig_h1
#ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)

# Plot all subjects, colored by Voice_Group
fig_slopes1 <- fig_h1 + geom_line(data = klang_clean,
            aes(x = Years, y = fitted, group = id, color = Voice_Group),
            alpha = 0.3, linewidth = 0.7, inherit.aes = FALSE) #+

ggsave("H1H2LTASMarginal_JASA_Individual_Final.pdf", width=9.25, height=5.71)
```

```r
library(lme4)
library(broom.mixed)

# Extract random slopes
re_slopes <- ranef(fit1_H)$id
re_df <- data.frame(id = rownames(re_slopes), slope = re_slopes$Years)

# Join Voice_Group info
re_df <- left_join(re_df, unique(klang_clean[, c("id", "Voice_Group")]), by = "id")

# Plot random slopes by Voice Group
ggplot(re_df, aes(x = reorder(id, slope), y = slope, fill = Voice_Group)) +
  geom_col(show.legend = TRUE) +
  coord_flip() +
  scale_fill_viridis_d(option = "C") +
  labs(title = "Random Slopes by Subject",
       x = "Subject ID", y = "Deviation in Years Slope") +
  theme_minimal(base_size = 14)
```



Alpha-Ratio

```
#Let's make non-treble the reference group
#klang$Voice_Group <- factor(klang$Voice_Group)  # convert to factor
klang$Voice_Group <- relevel(klang$Voice_Group, ref = "Non-Treble")  # now relevel
```

```
#Same process for Alpha Ratio.
fit0_a <- lmer(alphaRatio~Years*Voice_Group +dSPL + (Years | id), data=klang, na.action="na.omit",REML=
anova(fit0_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF  DenDF F value  Pr(>F)
## Years             1.08227 1.08227     1 101.00  2.5553 0.11305
## Voice_Group       0.14056 0.14056     1 105.91  0.3319 0.56578
## dSPL              0.35294 0.35294     1 253.53  0.8333 0.36218
## Years:Voice_Group 2.48834 2.48834     1 100.78  5.8752 0.01714 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Simplify removing the insignificant dSPL term

```
#Same process for Alpha Ratio.
fit1_a <- lmer(alphaRatio~Years*Voice_Group + (Years | id), data=klang, na.action="na.omit",REML=T)
anova(fit1_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
```

```
##                     Sum Sq Mean Sq NumDF  DenDF F value Pr(>F)
## Years             1.11339 1.11339     1 101.21  2.6538 0.1064
## Voice_Group       0.14107 0.14107     1 110.52  0.3362 0.5632
## Years:Voice_Group 2.35587 2.35587     1 101.21  5.6153 0.0197 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Again the interaction term is significant, so we don't reduce further.

```
margin1 <- ggpredict(fit1_a, c("Years", "Voice_Group"), ci_level=0.95)
margin1
```

```
## # Predicted values of alphaRatio
##
## Voice_Group: Non-Treble
##
## Years | Predicted |     95% CI
## ------------------------------
##     0 |      0.50 | 0.15, 0.85
##     1 |      0.68 | 0.38, 0.98
##     2 |      0.86 | 0.56, 1.16
##     3 |      1.04 | 0.69, 1.39
##     4 |      1.22 | 0.78, 1.66
##
## Voice_Group: Treble
##
## Years | Predicted |      95% CI
## -------------------------------
##     0 |      0.37 |  0.09, 0.64
##     1 |      0.34 |  0.11, 0.56
##     2 |      0.30 |  0.07, 0.54
##     3 |      0.27 | -0.02, 0.56
##     4 |      0.24 | -0.14, 0.61
##
## Adjusted for:
## * id = 0 (population-level)
```

For alpha ratio, the male means in year four have increased beyond the confidence interval at the beginning of studies. The female means have not changed.

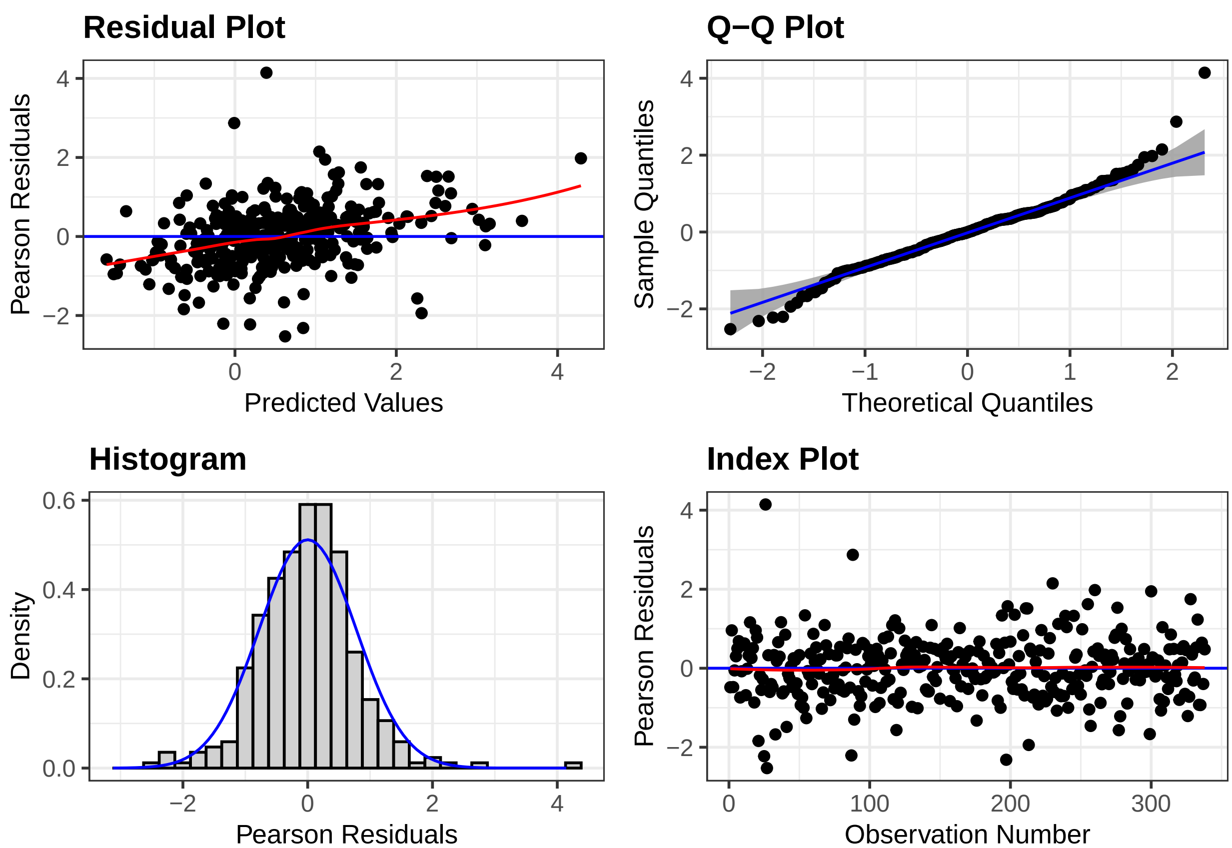
```
emm1_a <- emmeans::emmeans(fit1_a, ~ Years:Voice_Group)
pairs(emm1_a)
```

```
##  contrast                                                          estimate
##  (Years1.73769960282078 Non-Treble) - Years1.73769960282078 Treble    0.502
##     SE  df t.ratio p.value
##  0.188 112   2.673  0.0087
##
## Degrees-of-freedom method: kenward-roger
```

```r
p.fit1_a <- ggResidpanel::resid_panel(fit1_a,
              plots = c("resid", "qq", "hist", "index"),
              smoother = TRUE,
              qqbands = TRUE,
              title.opt = TRUE)
```

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

```r
p.fit1_a
```




```r
summary(fit1_a)
```

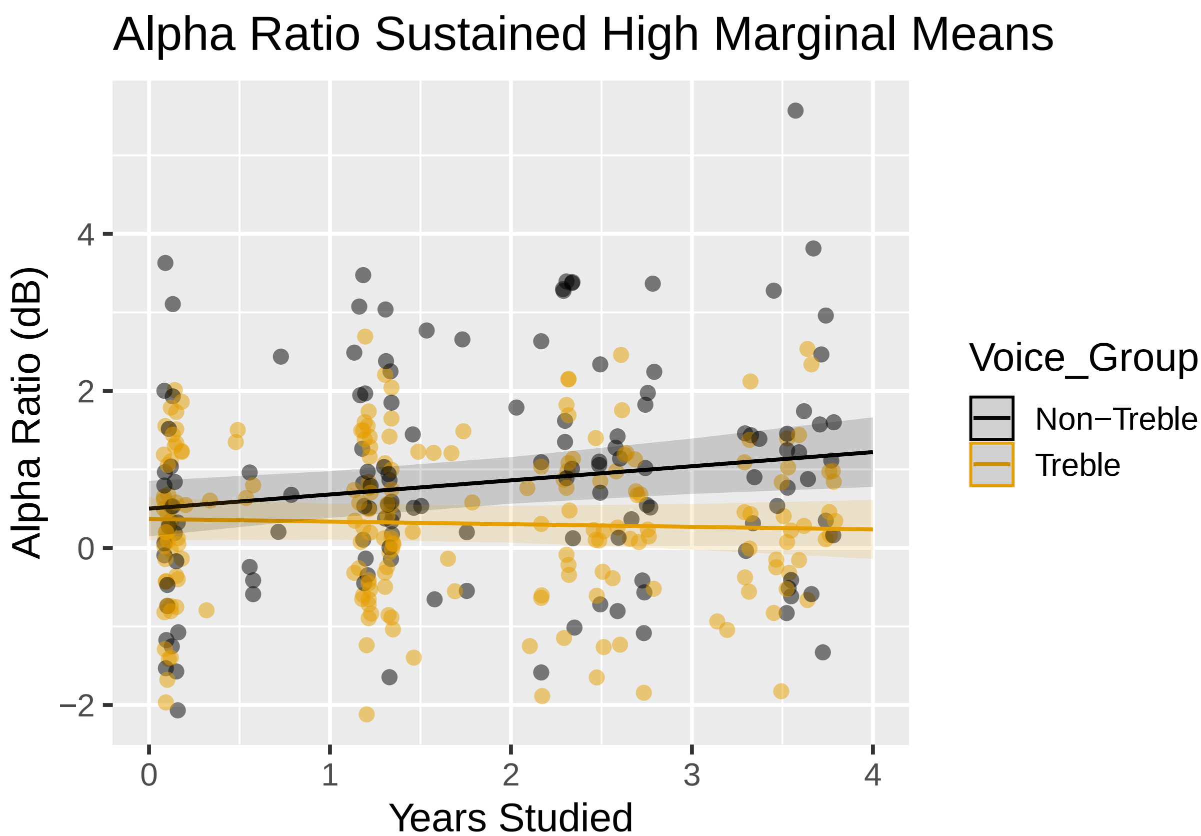
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: alphaRatio ~ Years * Voice_Group + (Years | id)
##    Data: klang
##
## REML criterion at convergence: 934
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5273 -0.4955 -0.0049  0.4556  4.1445
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.83038  0.9113
##           Years       0.07547  0.2747   -0.32
##  Residual             0.41955  0.6477
## Number of obs: 338, groups:  id, 117
##
## Fixed effects:
##                         Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)              0.50089    0.17975 109.60264   2.787  0.00628 **
## Years                    0.17991    0.06842  90.95329   2.630  0.01004 *
## Voice_GroupTreble       -0.13192    0.22750 110.52101  -0.580  0.56319
## Years:Voice_GroupTreble -0.21323    0.08998 101.20560  -2.370  0.01970 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Years  Vc_GrT
## Years       -0.574
## Voc_GrpTrbl -0.790  0.453
## Yrs:Vc_GrpT  0.436 -0.760 -0.576
```

```
fig_a1 <- plot_model(fit1_a, type = "pred", terms = c("Years", "Voice_Group"), show.data= TRUE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
       y = "Alpha Ratio (dB)",
       title = "Alpha Ratio Sustained High Marginal Means")
```

```
## Data points may overlap. Use the 'jitter' argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```
fig_a1
```




```
ggsave("AlphaMarginal_JASA_Final.pdf", width = 6, height = 4, dpi = 300)
```

```
performance::icc(fit1_a)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.679
##   Unadjusted ICC: 0.638
```

```
# INDIVIDUAL SLOPES clean and prep
klang_clean <- klang[!is.na(klang$alphaRatio), ]
klang_clean$fitted <- fitted(fit1_a)

# Ensure factors
klang_clean$id <- as.factor(klang_clean$id)
klang_clean$Voice_Group <- as.factor(klang_clean$Voice_Group)

fig_a1 <- plot_model(fit1_a, type = "pred", terms = c("Years", "Voice_Group"), show.data= FALSE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
        y = "Alpha Ratio (dB)",
        title = "Alpha Ratio Sustained High Individual Slopes")
#fig_h1
#ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)
```

```r
# Plot all subjects, colored by Voice_Group
fig_slopes2 <- fig_a1 + geom_line(data = klang_clean,
            aes(x = Years, y = fitted, group = id, color = Voice_Group),
            alpha = 0.3, linewidth = 0.7, inherit.aes = FALSE) #+

ggsave("AlphaMarginal_JASA_Individual_Final.pdf", width=9.25, height=5.71)
```

CPPS

```r
#Let's make treble the reference group
klang$Voice_Group <- relevel(klang$Voice_Group, ref = "Treble")  # now relevel
```

```r
#Same process for CPPs. _Review interaction term
fit0_c <- lmer(CPPs~Years*Voice_Group + dSPL + (Years | id), data=klang, na.action="na.omit",REML=T)
anova(fit0_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value    Pr(>F)
## Years              16.913  16.913     1  82.761  21.6080  1.25e-05 ***
## Voice_Group       101.239 101.239     1 102.600 129.3435 < 2.2e-16 ***
## dSPL                5.600   5.600     1 277.334   7.1540  0.007924 **
## Years:Voice_Group   0.059   0.059     1  82.461   0.0749  0.785006
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#We can remove the interaction term
fit1_c <- lmer(CPPs~Years + Voice_Group + dSPL + (Years | id), data=klang, na.action="na.omit",REML=T)
anova(fit1_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF   DenDF  F value    Pr(>F)
## Years        17.225  17.225     1  87.469  22.0206 9.899e-06 ***
## Voice_Group 183.010 183.010     1 110.460 233.9611 < 2.2e-16 ***
## dSPL          5.614   5.614     1 278.520   7.1768  0.007824 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Now we can remove dSPL

```r
fit2_c <- lmer(CPPs~Years + Voice_Group + (Years | id), data=klang, na.action="na.omit",REML=T)
anova(fit2_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years        15.353  15.353     1  88.514  19.585 2.733e-05 ***
## Voice_Group 177.456 177.456     1 109.595 226.381 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We have significant effects for Years and Voice_Group

```
margin1 <- ggpredict(fit2_c, c("Years", "Voice_Group"), ci_level=0.95)
margin1
```

```
## # Predicted values of CPPs
##
## Voice_Group: Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     16.79 | 16.47, 17.11
##     1 |     17.03 | 16.75, 17.30
##     2 |     17.27 | 17.00, 17.54
##     3 |     17.51 | 17.21, 17.81
##     4 |     17.75 | 17.38, 18.11
##
## Voice_Group: Non-Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     20.10 | 19.70, 20.49
##     1 |     20.34 | 19.98, 20.69
##     2 |     20.57 | 20.23, 20.91
##     3 |     20.81 | 20.45, 21.17
##     4 |     21.05 | 20.65, 21.46
##
## Adjusted for:
## * id = 0 (population-level)
```

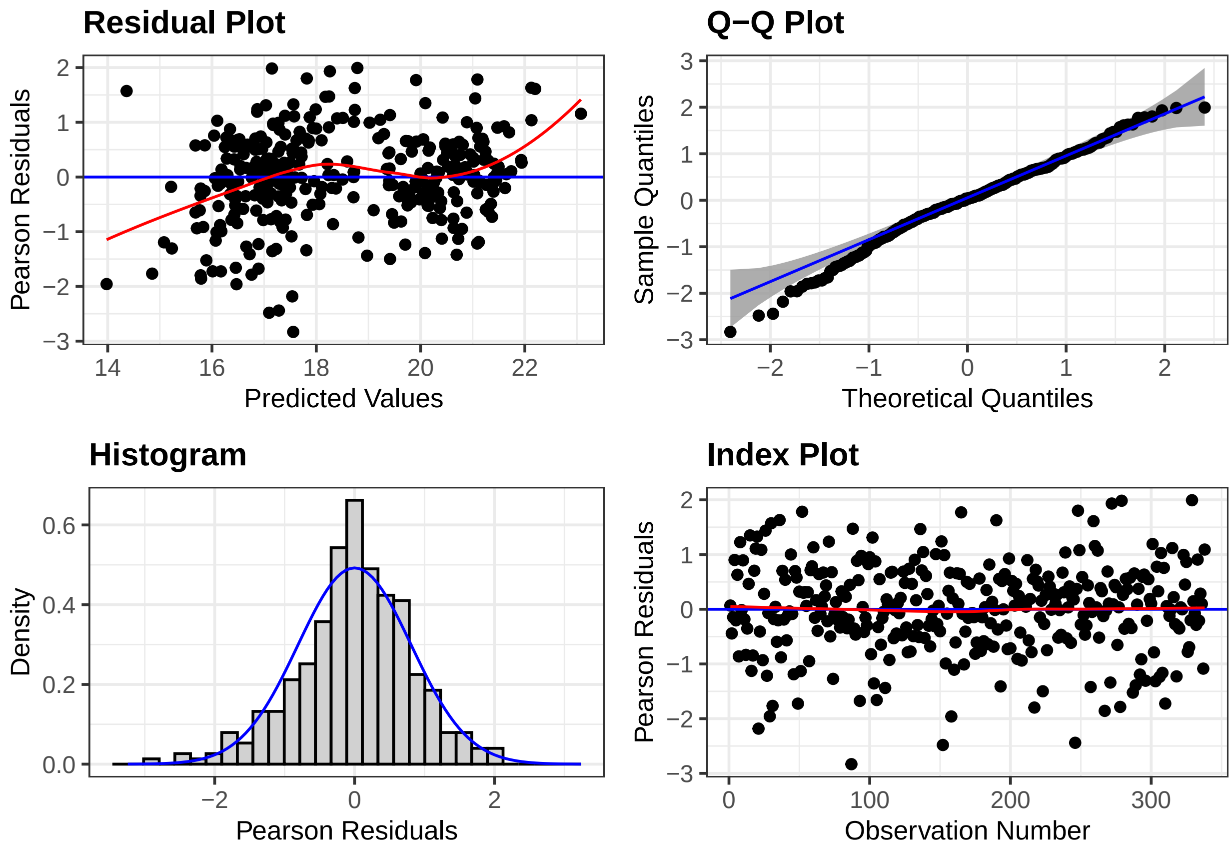
```
emm1_c <- emmeans::emmeans(fit2_c, ~ Years:Voice_Group)
pairs(emm1_c)
```

```
##  contrast                                                          estimate
##  Years1.73769960282078 Treble - (Years1.73769960282078 Non-Treble)    -3.31
##     SE  df t.ratio p.value
##  0.222 111 -14.874  <.0001
##
## Degrees-of-freedom method: kenward-roger
```

```
p.fit2_c <- ggResidpanel::resid_panel(fit2_c,
              plots = c("resid", "qq", "hist", "index"),
              smoother = TRUE,
              qqbands = TRUE,
              title.opt = TRUE)
```

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

```
p.fit2_c
```




```r
summary(fit2_c)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years + Voice_Group + (Years | id)
##    Data: klang
##
## REML criterion at convergence: 1091.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.83155 -0.43820  0.03541  0.54661  1.99323
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 1.33355  1.1548
##           Years       0.07934  0.2817   -0.50
##  Residual             0.78388  0.8854
## Number of obs: 338, groups:  id, 117
##
## Fixed effects:
##                       Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)            16.7899     0.1632 131.1535 102.867  < 2e-16 ***
## Years                   0.2390     0.0540  88.5136   4.426 2.73e-05 ***
## Voice_GroupNon-Treble   3.3066     0.2198 109.5953  15.046  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Years
## Years        -0.560
## Vc_GrpNn-Tr -0.483 -0.049
```

```
fig_c1 <- plot_model(fit2_c, type = "pred", terms = c("Years", "Voice_Group"), show.data= TRUE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
       y = "CPPs (dB)",
       title = "CPPs Sustained High Marginal Means") +
  theme(aspect.ratio = 1)
```

```
## Data points may overlap. Use the `jitter` argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```
#fig_c1
ggsave("CPPsMarginal_JASA_Final.pdf", width=9.25, height=5.71)
```

```
performance::icc(fit2_c)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.588
##   Unadjusted ICC: 0.238
```

Secondary research question: Do female voices became less breathy?

```
fit0_c_f <- lmer(CPPs~Years + dSPL + (Years | id), data=klang_f, na.action="na.omit",REML=T)
```

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

```
anova(fit0_c_f)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##        Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Years 18.4048 18.4048     1 140.07 18.4599 3.226e-05 ***
## dSPL   4.0977  4.0977     1 167.42  4.1099   0.04422 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

dSPL can be removed

```
fit1_c_f <- lmer(CPPs~Years + (Years | id), data=klang_f, na.action="na.omit",REML=T)
```

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

```r
summary(fit1_c_f)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years + (Years | id)
##    Data: klang_f
##
## REML criterion at convergence: 630.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.41490 -0.44183  0.02198  0.62865  1.90051
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 1.621339 1.27332
##           Years       0.006302 0.07938  -1.00
##  Residual             1.007089 1.00354
## Number of obs: 185, groups:  id, 68
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  16.74064    0.20932  67.65117  79.976  < 2e-16 ***
## Years         0.31005    0.07176 138.41230   4.321 2.95e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr)
## Years -0.662
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# INDIVIDUAL SLOPES clean and prep
klang_clean <- klang[!is.na(klang$CPPs), ]
klang_clean$fitted <- fitted(fit2_c)

# Ensure factors
klang_clean$id <- as.factor(klang_clean$id)
klang_clean$Voice_Group <- as.factor(klang_clean$Voice_Group)

fig_c1_ <- plot_model(fit2_c, type = "pred", terms = c("Years", "Voice_Group"), show.data= FALSE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
       y = "CPPs (dB)",
       title = "CPPs Sustained High Individual Slopes")
#fig_h1
#ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)

# Plot all subjects, colored by Voice_Group
fig_slopes3 <- fig_c1_ + geom_line(data = klang_clean,
            aes(x = Years, y = fitted, group = id, color = Voice_Group),
```

```
            alpha = 0.3, linewidth = 0.7, inherit.aes = FALSE) #+

ggsave("CPPsMarginal_JASA_Individual_Final.pdf", width=9.25, height=5.71)
```

Benjamini-Hochberg Correction We take the p-values from all of our initial linear mixed models:

```
p_values_H <- summary(fit0_H)$coefficients[, "Pr(>|t|)"]
p_values_a <- summary(fit0_a)$coefficients[, "Pr(>|t|)"]
p_values_c <- summary(fit0_c)$coefficients[, "Pr(>|t|)"]
p_values_c_f <- summary(fit0_c_f)$coefficients[, "Pr(>|t|)"]

# Organize p-values into a data frame
p_values_df <- data.frame(
  H1H2LTAS = p_values_H,
  AlphaRatio = p_values_a,
  CPPs = p_values_c#,
  #Model4 = p_values_c_f
)

# Combine all p-values into a single vector
all_p_values <- c(p_values_H, p_values_a, p_values_c)

#Save the unadjusted for final comparison:
p_2 <- all_p_values
p_2_f <- p_values_c_f
```

Repertoire: Avezzo a vivere

Now for the repertoire sample "Avezzo a vivere"

Load and Prepare Data

```
klang6 <- read.csv("Klang6_JASA_Final.csv", fileEncoding = "UTF-8")
klang6 <- klang6 %>%
  select(yearDiff, H1H2LTAS, alphaRatio, CPPs, Jahr, Stimmfach, Voice.Type, geschlecht, alter, id, dSPL)
  rename(Years = yearDiff, Gender = geschlecht, Age = alter)

# Create Voice Grouping
klang6 <- klang6 %>%
  mutate(
    Voice_Group = case_when(
      Stimmfach == "Sop/Mezzo/Alt" ~ "Treble",
      Stimmfach == "Ten/Bar/Bass" ~ "Non-Treble"
    ),
    Gender = recode(Gender, "männl." = "Male", "weibl." = "Female")
  )
```

Filtering Data

```
# Step 1: Count total unique students by gender
total_counts <- klang6 %>% group_by(Gender) %>% summarise(Total_Students = n_distinct(id))

# Step 2: Filter students with Years <= 4
klang6_filtered <- klang6 %>% filter(Years <= 4)
filtered_counts <- klang6_filtered %>% group_by(Gender) %>% summarise(Filtered_Students = n_distinct(id

# Step 3: Remove students with only one recording
valid_ids <- klang6_filtered %>% count(id) %>% filter(n > 1) %>% pull(id)
klang6_final <- klang6_filtered %>% filter(id %in% valid_ids)
final_counts <- klang6_final %>% group_by(Gender) %>% summarise(Final_Students = n_distinct(id))

# Print Results
print(total_counts)
```

```
## # A tibble: 2 x 2
##   Gender Total_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               48
```

```
print(filtered_counts)
```

```
## # A tibble: 2 x 2
##   Gender Filtered_Students
##   <chr>              <int>
## 1 Female                68
## 2 Male                  48
```

```
print(final_counts)
```

```
## # A tibble: 2 x 2
##   Gender Final_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               48
```

```
# Update dataset
klang6 <- klang6_final
```

H1H2LTAS Analysis

```
#Let's make treble the reference group
klang6$Voice_Group <- factor(klang6$Voice_Group)  # convert to factor
klang6$Voice_Group <- relevel(klang6$Voice_Group, ref = "Treble")  # now relevel
```

```r
fit0_H <- lmer(H1H2LTAS ~ Years * Voice_Group + dSPL + (Years | id), data = klang6, na.action = "na.omi
anova(fit0_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value  Pr(>F)
## Years               0.688   0.688     1  88.531   1.7478 0.18956
## Voice_Group        96.301  96.301     1 109.362 244.5759 < 2e-16 ***
## dSPL                0.180   0.180     1 212.065   0.4569 0.49981
## Years:Voice_Group   2.331   2.331     1  88.169   5.9202 0.01699 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Remove the dSPL term:

```r
fit1_H <- lmer(H1H2LTAS ~ Years * Voice_Group + (Years | id), data = klang6, na.action = "na.omit", REMI
anova(fit1_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value  Pr(>F)
## Years               0.716   0.716     1  88.242   1.8182 0.18098
## Voice_Group        96.172  96.172     1 109.432 244.0779 < 2e-16 ***
## Years:Voice_Group   2.323   2.323     1  88.242   5.8949 0.01722 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Marginal Means
margin1 <- ggpredict(fit1_H, c("Years", "Voice_Group"), ci_level = 0.95)
print(margin1)
```

```
## # Predicted values of H1H2LTAS
##
## Voice_Group: Treble
##
## Years | Predicted |      95% CI
## --------------------------------
##     0 |      1.16 | 0.79,  1.53
##     1 |      1.34 | 1.03,  1.65
##     2 |      1.52 | 1.23,  1.81
##     3 |      1.70 | 1.38,  2.02
##     4 |      1.88 | 1.49,  2.27
##
## Voice_Group: Non-Treble
##
## Years | Predicted |       95% CI
## ---------------------------------
##     0 |     -3.64 | -4.12, -3.16
##     1 |     -3.70 | -4.10, -3.29
##     2 |     -3.75 | -4.13, -3.36
##     3 |     -3.80 | -4.20, -3.39
##     4 |     -3.85 | -4.32, -3.38
##
## Adjusted for:
## * id = 0 (population-level)
```

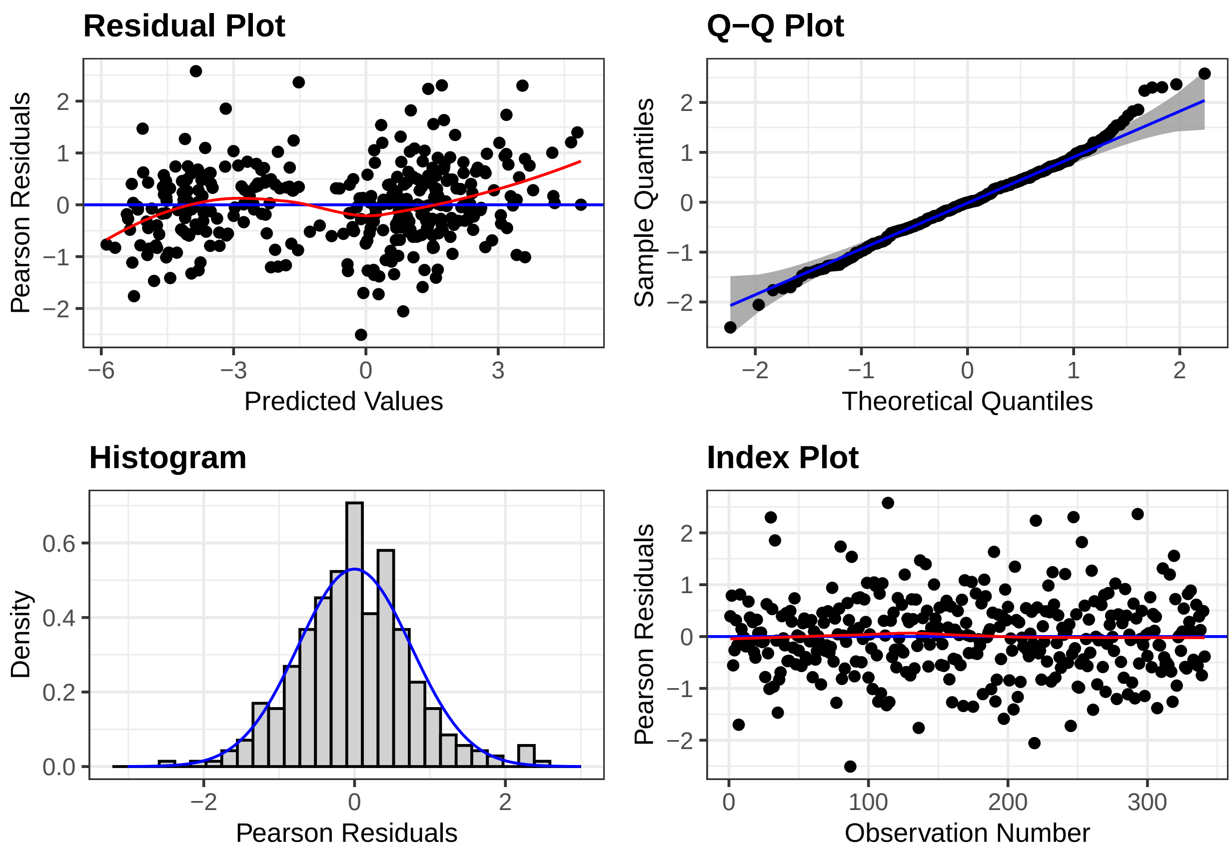
```r
# Pairwise Comparisons
emm0_H <- emmeans(fit1_H, ~ Years:Voice_Group)
pairs(emm0_H)
```

```
##  contrast                                                           estimate
##  Years1.72916884264653 Treble - (Years1.72916884264653 Non-Treble)      5.2
##     SE  df t.ratio p.value
##  0.246 112  21.164  <.0001
##
## Degrees-of-freedom method: kenward-roger
```

```r
# Residual Diagnostics
p.fit0_H <- ggResidpanel::resid_panel(fit1_H, plots = c("resid", "qq", "hist", "index"), smoother = TRUE
```

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

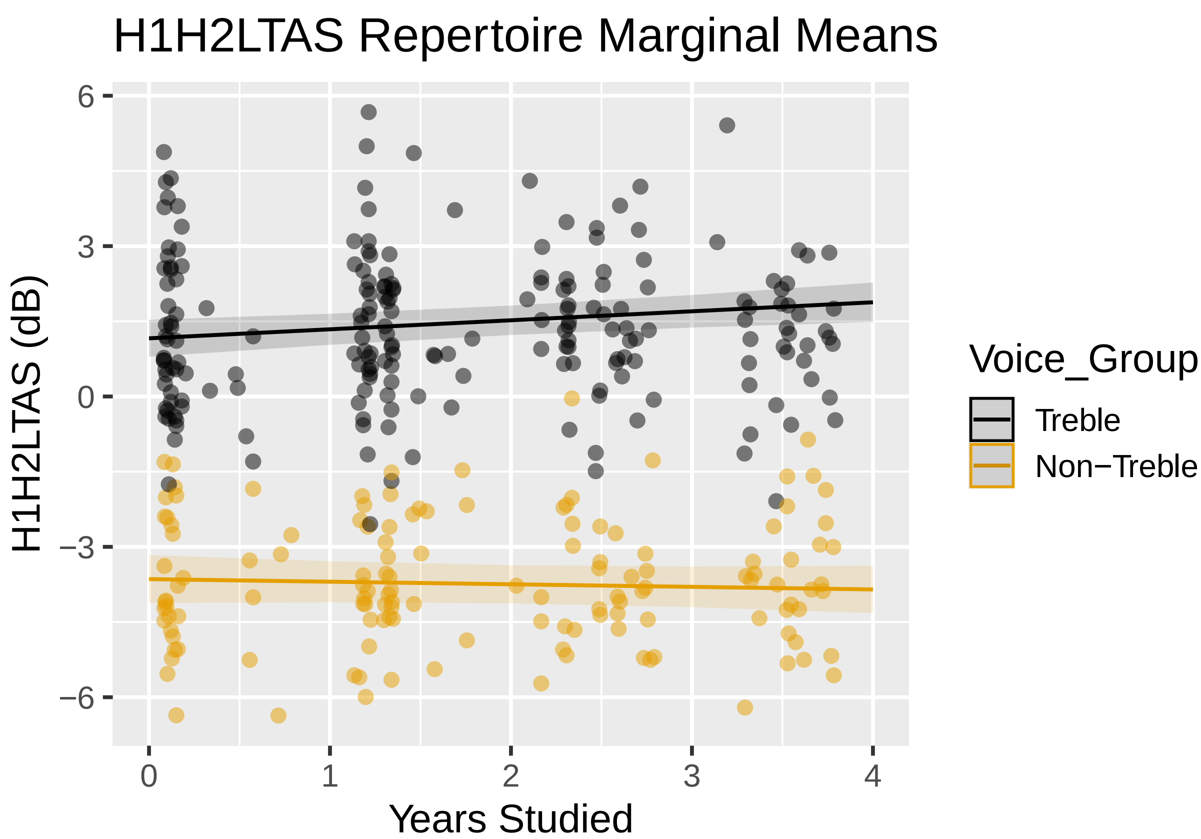
```r
print(p.fit0_H)
```




```r
# Plot Model Predictions
fig_h2 <- plot_model(fit1_H, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE,
                     colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "H1H2LTAS (dB)", title = "H1H2LTAS Repertoire Marginal Means")
```

```
## Data points may overlap. Use the `jitter` argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```r
fig_h2
```




```r
ggsave("H1H2LTASAvezzoMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit1_H)
```

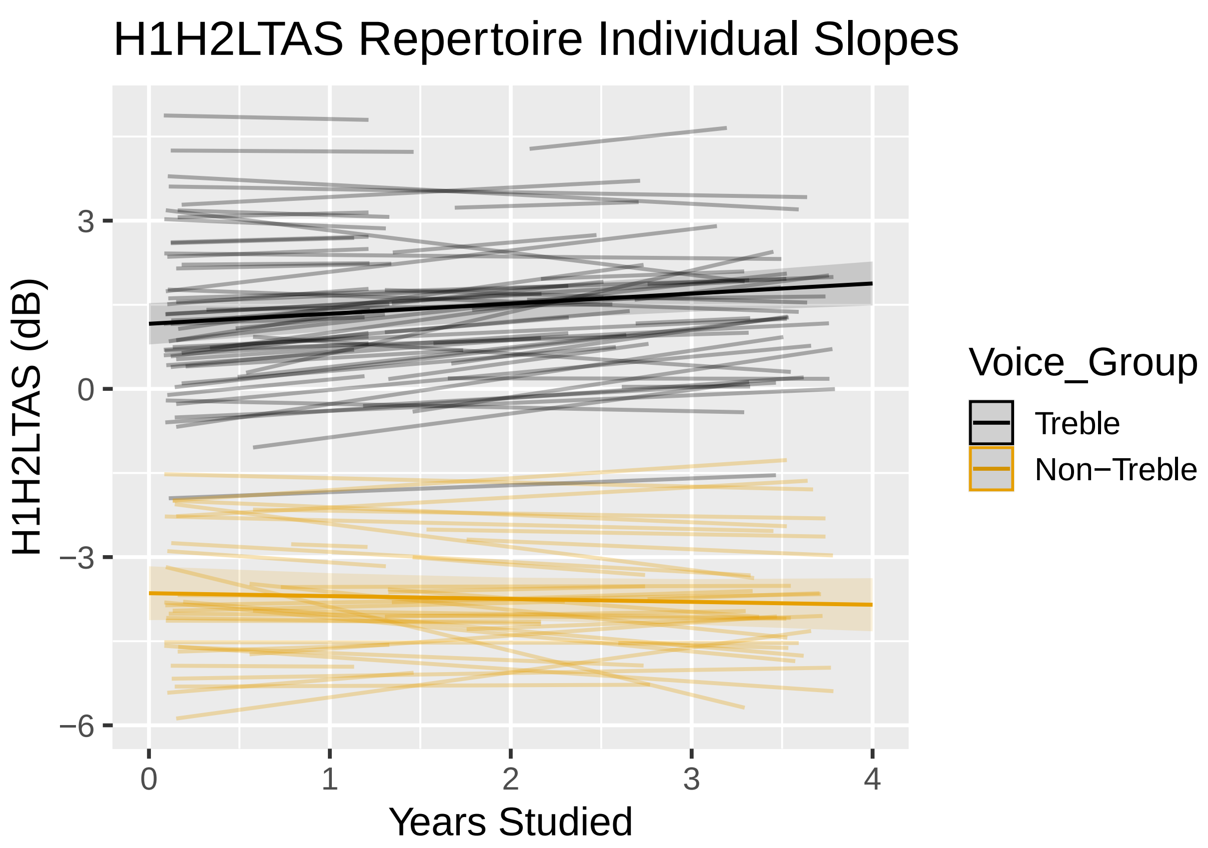
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: H1H2LTAS ~ Years * Voice_Group + (Years | id)
##    Data: klang6
##
## REML criterion at convergence: 1000.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.50827 -0.48095 -0.00819  0.45091  2.57766
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 1.95193  1.3971
##           Years       0.09943  0.3153   -0.53
```

```
##  Residual             0.39402  0.6277
## Number of obs: 341, groups:  id, 116
##
## Fixed effects:
##                              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)                   1.16077    0.18736 112.53006   6.195 9.79e-09 ***
## Years                         0.17967    0.06163 105.39194   2.915  0.00434 **
## Voice_GroupNon-Treble        -4.80476    0.30754 109.43169 -15.623  < 2e-16 ***
## Years:Voice_GroupNon-Treble  -0.23103    0.09515  88.24242  -2.428  0.01722 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Years  V_GN-T
## Years        -0.607
## Vc_GrpNn-Tr -0.609  0.370
## Yrs:Vc_GN-T  0.393 -0.648 -0.605
```

```r
performance::icc(fit1_H)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.801
##   Unadjusted ICC: 0.187
```

```r
# INDIVIDUAL SLOPES clean and prep
klang_clean <- klang6[!is.na(klang6$H1H2LTAS), ]
klang_clean$fitted <- fitted(fit1_H)

# Ensure factors
klang_clean$id <- as.factor(klang_clean$id)
klang_clean$Voice_Group <- as.factor(klang_clean$Voice_Group)

fig_h1 <- plot_model(fit1_H, type = "pred", terms = c("Years", "Voice_Group"), show.data= FALSE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
        y = "H1H2LTAS (dB)",
        title = "H1H2LTAS Repertoire Individual Slopes")
#fig_h1
#ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)

# Plot all subjects, colored by Voice_Group
fig_h1 + geom_line(data = klang_clean,
            aes(x = Years, y = fitted, group = id, color = Voice_Group),
            alpha = 0.3, linewidth = 0.7, inherit.aes = FALSE) #+
```




```r
ggsave("H1H2LTASMarginalRepertoire_JASA_Individual_Final.pdf", width=9.25, height=5.71)
```

Alpha Ratio Analysis

```r
#Let's make treble the reference group
klang6$Voice_Group <- relevel(klang6$Voice_Group, ref = "Non-Treble")  # now relevel
```

```r
fit0_a <- lmer(alphaRatio ~ Years * Voice_Group + dSPL + (Years | id), data = klang6, na.action = "na.om
anova(fit0_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years             0.28331 0.28331     1  92.146  2.3513    0.1286
## Voice_Group       2.57818 2.57818     1 105.892 21.3967 1.061e-05 ***
## dSPL              0.25570 0.25570     1 183.253  2.1221    0.1469
## Years:Voice_Group 0.08646 0.08646     1  91.970  0.7175    0.3992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Remove dSPL
fit1_a <- lmer(alphaRatio ~ Years * Voice_Group + (Years | id), data = klang6, na.action = "na.omit", RE
anova(fit1_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years             0.27418 0.27418     1  91.986  2.2626    0.1360
## Voice_Group       2.54400 2.54400     1 106.962 20.9932 1.253e-05 ***
## Years:Voice_Group 0.08573 0.08573     1  91.986  0.7075    0.4025
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Model Simplification
fit2_a <- update(fit1_a, . ~ . - Years:Voice_Group)
anova(fit2_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##             Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years       0.2294  0.2294     1  97.456  1.8884    0.1725
## Voice_Group 5.0083  5.0083     1 111.200 41.2264 3.446e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fit3_a <- update(fit2_a, . ~ . - Years)
anova(fit3_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##             Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Voice_Group 4.9933  4.9933     1 111.14    41.4 3.234e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Compare Models
anova(fit0_a, fit1_a, fit2_a, fit3_a)
```

```
## refitting model(s) with ML (instead of REML)
```

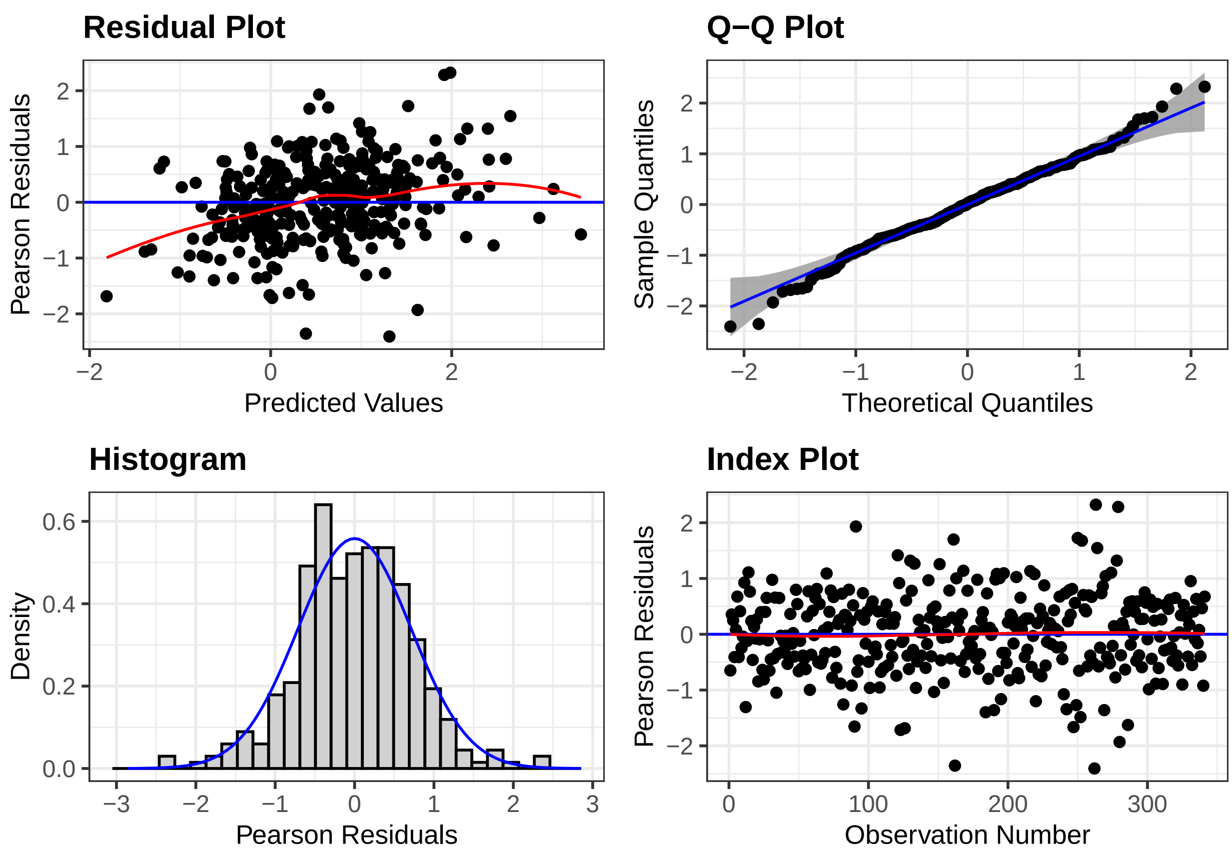
```
## Data: klang6
## Models:
## fit3_a: alphaRatio ~ Voice_Group + (Years | id)
## fit2_a: alphaRatio ~ Years + Voice_Group + (Years | id)
## fit1_a: alphaRatio ~ Years * Voice_Group + (Years | id)
## fit0_a: alphaRatio ~ Years * Voice_Group + dSPL + (Years | id)
##        npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## fit3_a    6 648.56 671.55 -318.28   636.56
## fit2_a    7 648.68 675.50 -317.34   634.68 1.8793  1     0.1704
## fit1_a    8 649.96 680.62 -316.98   633.96 0.7150  1     0.3978
## fit0_a    9 649.86 684.35 -315.93   631.86 2.1030  1     0.1470
```

```
# Residual Diagnostics
p.fit3_a <- ggResidpanel::resid_panel(fit3_a, plots = c("resid", "qq", "hist", "index"), smoother = TRUE
```

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

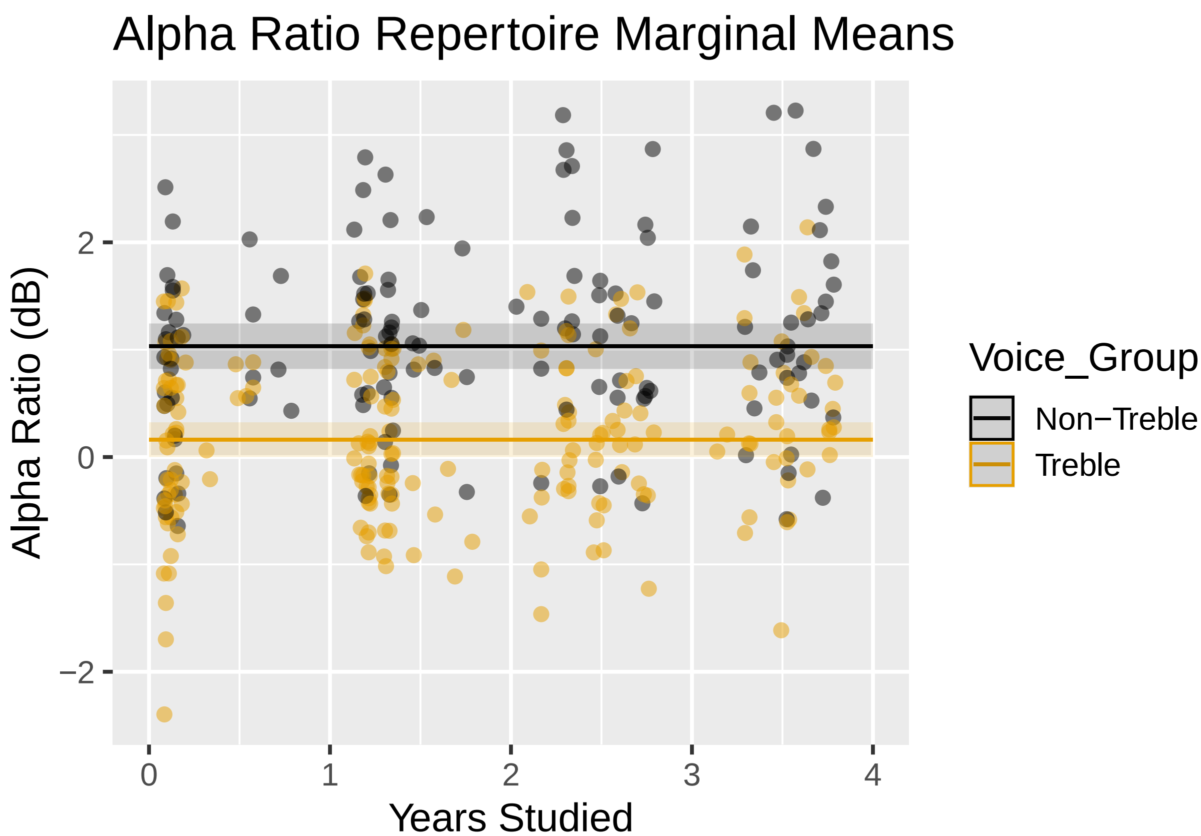
```
## `geom_smooth()` using formula = 'y ~ x'
```

```r
print(p.fit3_a)
```




```r
# Plot Model Predictions
plot_model(fit3_a, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "Alpha Ratio (dB)", title = "Alpha Ratio Repertoire Marginal Means")
```

```
## Data points may overlap. Use the 'jitter' argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```




```r
ggsave("AlphaAvezzoMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit3_a)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: alphaRatio ~ Voice_Group + (Years | id)
##    Data: klang6
##
## REML criterion at convergence: 642.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.40619 -0.45992  0.02018  0.45749  2.32302
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.58328  0.7637
##           Years       0.07575  0.2752   -0.51
##  Residual             0.12061  0.3473
## Number of obs: 341, groups:  id, 116
##
## Fixed effects:
##                      Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)            1.0326     0.1075 108.8496   9.603 3.49e-16 ***
```

```
## Voice_GroupTreble  -0.8708     0.1353 111.1445  -6.434 3.23e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Voc_GrpTrbl -0.795
```

```
performance::icc(fit3_a)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##   the conditional random effect variances.
##   Solution: Respecify fixed structure!
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.829
##   Unadjusted ICC: 0.658
```

CPPs Analysis

```
#Let's make treble the reference group
klang6$Voice_Group <- relevel(klang6$Voice_Group, ref = "Treble")  # now relevel
```

```
fit0_c <- lmer(CPPs ~ Years * Voice_Group + dSPL + (Years | id), data = klang6, na.action = "na.omit", R
anova(fit0_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years              0.0894  0.0894     1  43.412  0.2387   0.62759
## Voice_Group       16.1627 16.1627     1  88.022 43.1476 3.435e-09 ***
## dSPL               0.0001  0.0001     1 254.186  0.0003   0.98719
## Years:Voice_Group  2.5184  2.5184     1  43.105  6.7230   0.01295 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Remove dSPL
fit1_c <- lmer(CPPs ~ Years * Voice_Group + (Years | id), data = klang6, na.action = "na.omit", REML = '
anova(fit1_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Years              0.0888  0.0888     1 43.137  0.2382   0.62796
## Voice_Group       16.0852 16.0852     1 94.628 43.1676 2.711e-09 ***
## Years:Voice_Group  2.5061  2.5061     1 43.137  6.7255   0.01293 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Marginal Means
margin1 <- ggpredict(fit1_c, c("Years", "Voice_Group"), ci_level = 0.95)
print(margin1)
```

```
## # Predicted values of CPPs
##
## Voice_Group: Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     14.45 | 14.18, 14.72
##     1 |     14.57 | 14.35, 14.78
##     2 |     14.68 | 14.49, 14.88
##     3 |     14.80 | 14.58, 15.02
##     4 |     14.91 | 14.63, 15.19
##
## Voice_Group: Non-Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     15.91 | 15.56, 16.25
##     1 |     15.83 | 15.55, 16.11
##     2 |     15.75 | 15.50, 16.00
##     3 |     15.67 | 15.40, 15.94
##     4 |     15.59 | 15.26, 15.91
##
## Adjusted for:
## * id = 0 (population-level)
```

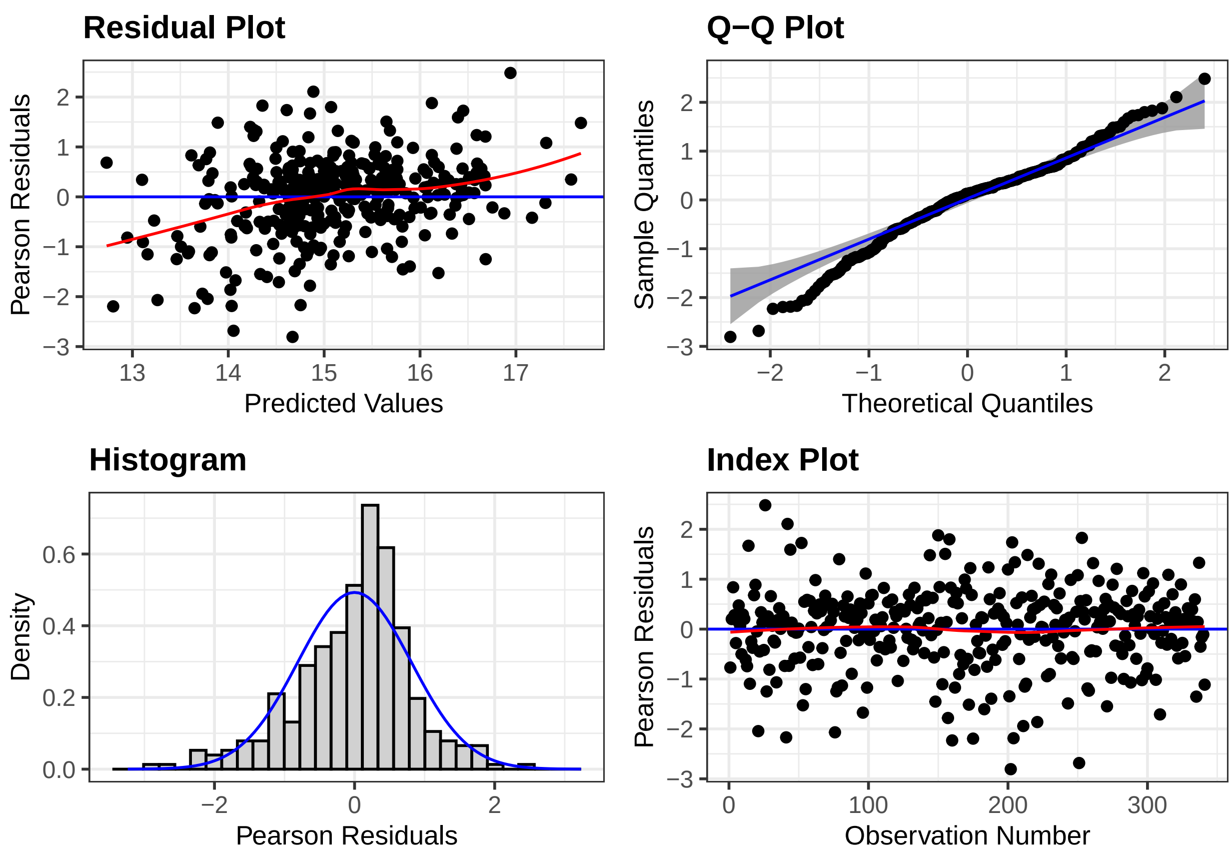
```r
# Pairwise Comparisons
emm1_c <- emmeans(fit1_c, ~ Years:Voice_Group)
pairs(emm1_c)
```

```
##  contrast                                                          estimate
##  Years1.72916884264653 Treble - (Years1.72916884264653 Non-Treble)    -1.12
##     SE  df t.ratio p.value
##  0.162 111  -6.892  <.0001
##
## Degrees-of-freedom method: kenward-roger
```

```r
# Residual Diagnostics
p.fit1_c <- ggResidpanel::resid_panel(fit1_c, plots = c("resid", "qq", "hist", "index"), smoother = TRUE
```

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

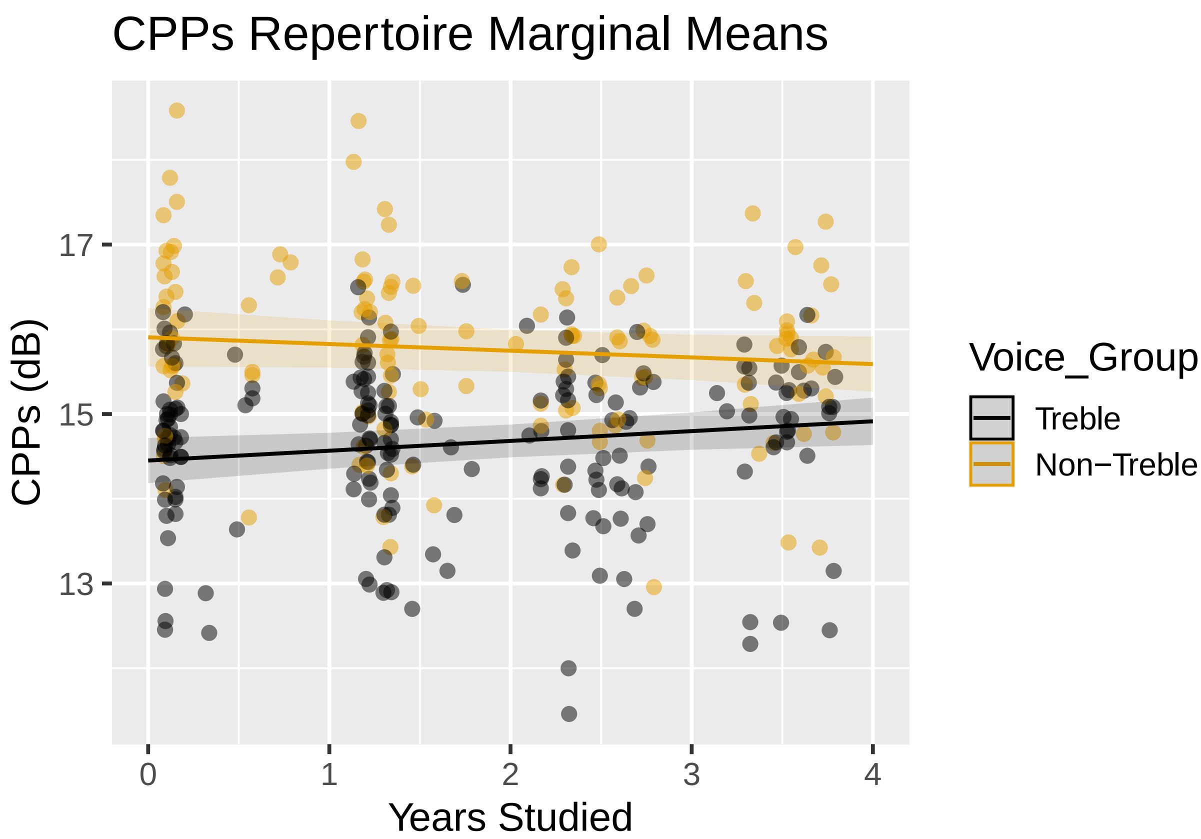
```r
print(p.fit1_c)
```




```
# Plot Model Predictions
fig_c2 <- plot_model(fit1_c, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE,
                     colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "CPPs (dB)", title = "CPPs Repertoire Marginal Means")
```

```
## Data points may overlap. Use the `jitter` argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```
fig_c2
```




```r
ggsave("CPPsAvezzoMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit1_c)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years * Voice_Group + (Years | id)
##    Data: klang6
##
## REML criterion at convergence: 863
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8067 -0.4262  0.1258  0.4816  2.4822
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.82440  0.9080
##           Years       0.03708  0.1926   -0.62
##  Residual             0.37262  0.6104
## Number of obs: 341, groups:  id, 116
##
## Fixed effects:
##                              Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)                  14.45110    0.13534 97.65697 106.773  < 2e-16 ***
```

```
## Years                         0.11569    0.04895 53.68252   2.363   0.0218 *
## Voice_GroupNon-Treble        1.45414    0.22132 94.62833   6.570 2.71e-09 ***
## Years:Voice_GroupNon-Treble -0.19474    0.07509 43.13713  -2.593   0.0129 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Years  V_GN-T
## Years       -0.687
## Vc_GrpNn-Tr -0.612  0.420
## Yrs:Vc_GN-T  0.448 -0.652 -0.689
```

```r
performance::icc(fit1_c)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.622
##   Unadjusted ICC: 0.471
```

```r
fit0_c_f <- lmer(CPPs ~ Years  + dSPL + (Years | id), data = klang6 %>% filter(Gender == "Female"), REML
anova(fit0_c_f)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##        Sum Sq Mean Sq NumDF   DenDF F value  Pr(>F)
## Years 1.93554 1.93554     1  25.734  4.9403 0.03524 *
## dSPL  0.01308 0.01308     1 119.527  0.0334 0.85531
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Remove dSPL
fit1_c_f <- lmer(CPPs ~ Years + (Years | id), data = klang6 %>% filter(Gender == "Female"), REML = TRUE)
summary(fit1_c_f)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years + (Years | id)
##    Data: klang6 %>% filter(Gender == "Female")
##
## REML criterion at convergence: 478.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7188 -0.4879  0.1385  0.5166  1.8067
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.72887  0.8537
##           Years       0.02663  0.1632   -0.51
##  Residual             0.39050  0.6249
## Number of obs: 187, groups:  id, 68
##
## Fixed effects:
```

```
##              Estimate Std. Error       df t value Pr(>|t|)
## (Intercept) 14.42295    0.13783 52.24399  104.64   <2e-16 ***
## Years        0.11214    0.05074 25.90354    2.21   0.0361 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr)
## Years -0.660
```

```r
p_values_H <- summary(fit0_H)$coefficients[, "Pr(>|t|)"]
p_values_a <- summary(fit0_a)$coefficients[, "Pr(>|t|)"]
p_values_c <- summary(fit0_c)$coefficients[, "Pr(>|t|)"]
p_values_c_f <- summary(fit0_c_f)$coefficients[, "Pr(>|t|)"]

all_p_values <- c(p_values_H, p_values_a, p_values_c)#, p_values_c_f)
p_6 <- all_p_values
p_6_f <- p_values_c_f
```

Medium Sustained /a/

Load and Prepare Data

```r
# Load data
klang1 <- read.csv("Klang1_JASA_Final.csv", fileEncoding = "UTF-8")
klang1 <- klang1[c("yearDiff", "H1H2LTAS", "alphaRatio", "CPPs", "Jahr", "Stimmfach", "Voice.Type", "ges
names(klang1)[names(klang1) == "geschlecht"] <- "Gender"
names(klang1)[names(klang1) == "yearDiff"] <- "Years"
names(klang1)[names(klang1) == "alter"] <- "Age"
klang1$Voice_Group[klang1$Stimmfach == "Sop/Mezzo/Alt"] <- "Treble"
klang1$Voice_Group[klang1$Stimmfach == "Ten/Bar/Bass"] <- "Non-Treble"
klang1$Gender[klang1$Gender == "männl."] <- "Male"
klang1$Gender[klang1$Gender == "weibl."] <- "Female"
```

Filtering Data

```r
# Step 1: Count total unique male and female students
total_counts <- klang1 %>%
  group_by(Gender) %>%
  summarise(Total_Students = n_distinct(id))

# Step 2: Filter to only include students with Years <= 4
klang1_filtered <- klang1 %>%
  filter(Years <= 4)

# Count unique male and female students after filtering
filtered_counts <- klang1_filtered %>%
  group_by(Gender) %>%
```

```r
  summarise(Filtered_Students = n_distinct(id))

# Step 3: Remove students with only one recording
id_counts <- klang1_filtered %>%
  count(id)  # Count occurrences of each student (id)

valid_ids <- id_counts %>%
  filter(n > 1) %>%
  pull(id)  # Get list of students who appear more than once

klang1_final <- klang1_filtered %>%
  filter(id %in% valid_ids)  # Keep only students with multiple recordings

# Count unique male and female students after final filtering
final_counts <- klang1_final %>%
  group_by(Gender) %>%
  summarise(Final_Students = n_distinct(id))

# Print results
print(total_counts)      # Total students by gender
```

```
## # A tibble: 2 x 2
##   Gender Total_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               49
```

```r
print(filtered_counts)   # After Years <= 4 restriction
```

```
## # A tibble: 2 x 2
##   Gender Filtered_Students
##   <chr>              <int>
## 1 Female                68
## 2 Male                  49
```

```r
print(final_counts)      # After removing single-recording students
```

```
## # A tibble: 2 x 2
##   Gender Final_Students
##   <chr>           <int>
## 1 Female             68
## 2 Male               49
```

```r
# Update the klang1 dataframe to keep only the final filtered version
klang1 <- klang1_final
```

H1H2LTAS

```r
#Let's make treble the reference group
klang1$Voice_Group <- factor(klang1$Voice_Group)  # convert to factor
klang1$Voice_Group <- relevel(klang1$Voice_Group, ref = "Treble")  # now relevel
```

```r
fit0_H <- lmer(H1H2LTAS ~ Years * Voice_Group + dSPL + (Years | id), data = klang1, REML = TRUE)
anova(fit0_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF  F value  Pr(>F)
## Years                0.00    0.00     1  81.801   0.0008 0.97713
## Voice_Group        565.62  565.62     1 110.933 436.6740 < 2e-16 ***
## dSPL                 5.15    5.15     1 269.429   3.9767 0.04714 *
## Years:Voice_Group    5.94    5.94     1  82.022   4.5868 0.03519 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#After significance correction, not significant: Remove dSPL
fit1_H <- lmer(H1H2LTAS ~ Years * Voice_Group + (Years | id), data = klang1, REML = TRUE)
anova(fit1_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF   DenDF F value  Pr(>F)
## Years                0.00    0.00     1  81.743   0.000 0.99496
## Voice_Group        582.93  582.93     1 110.522 431.735 < 2e-16 ***
## Years:Voice_Group    6.06    6.06     1  81.743   4.486 0.03721 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#After significance correction, not significant: Remove interaction
fit2_H <- update(fit1_H, . ~ . - Years:Voice_Group)
anova(fit2_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF   DenDF  F value Pr(>F)
## Years           0.1     0.1     1  89.965   0.0723 0.7886
## Voice_Group 1027.6  1027.6     1 107.458 761.8817 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Remove years
fit3_H <- update(fit2_H, . ~ . - Years)
anova(fit3_H)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Voice_Group 1033.8  1033.8     1 107.07  764.95 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

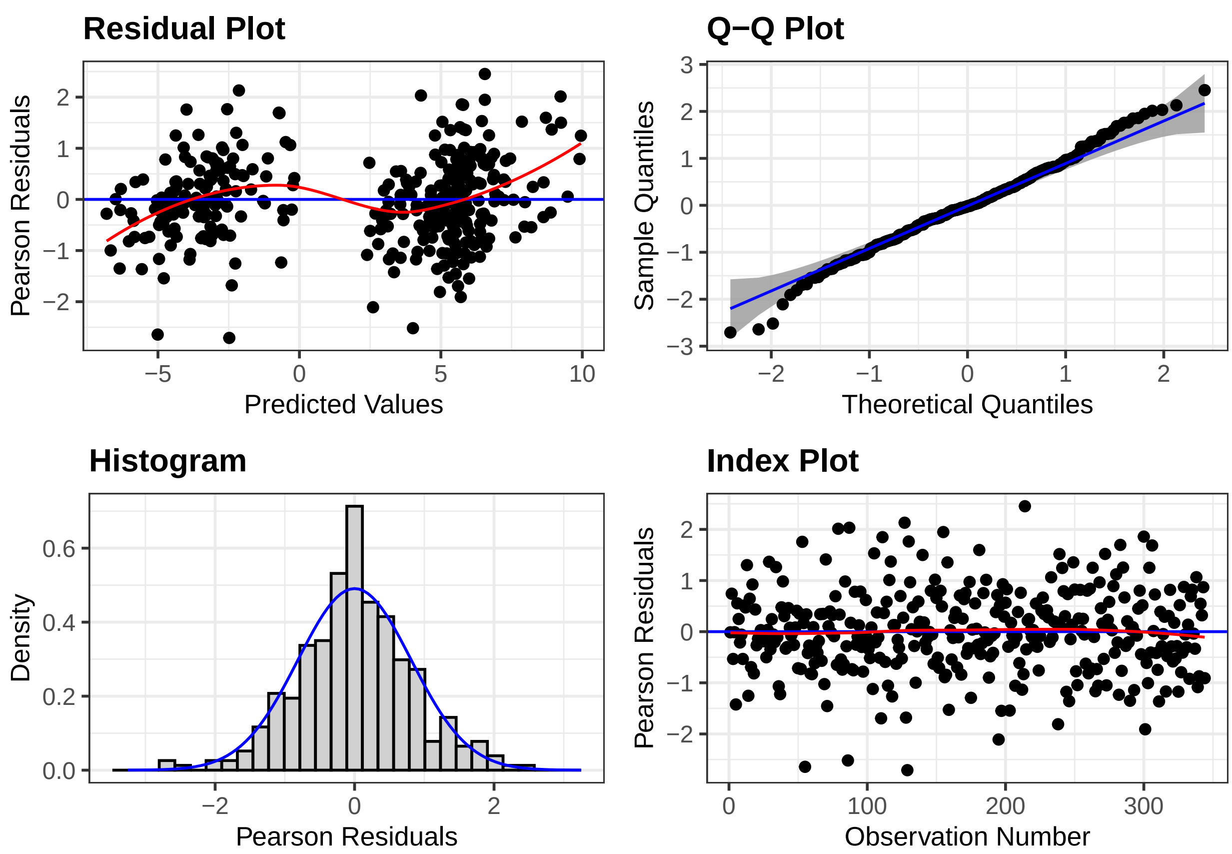
```r
anova(fit0_H, fit1_H, fit2_H, fit3_H, digits = 10, test = "LRT")
```

```
## refitting model(s) with ML (instead of REML)

## Data: klang1
## Models:
## fit3_H: H1H2LTAS ~ Voice_Group + (Years | id)
## fit2_H: H1H2LTAS ~ Years + Voice_Group + (Years | id)
## fit1_H: H1H2LTAS ~ Years * Voice_Group + (Years | id)
## fit0_H: H1H2LTAS ~ Years * Voice_Group + dSPL + (Years | id)
##        npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## fit3_H    6 1328.7 1351.8 -658.36   1316.7
## fit2_H    7 1330.7 1357.5 -658.33   1316.7 0.0674  1    0.79512
## fit1_H    8 1328.3 1359.0 -656.14   1312.3 4.3759  1    0.03645 *
## fit0_H    9 1326.6 1361.1 -654.29   1308.6 3.7125  1    0.05401 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
p.fit3_H <- resid_panel(fit3_H, plots = c("resid", "qq", "hist", "index"), smoother = TRUE, qqbands = T
```

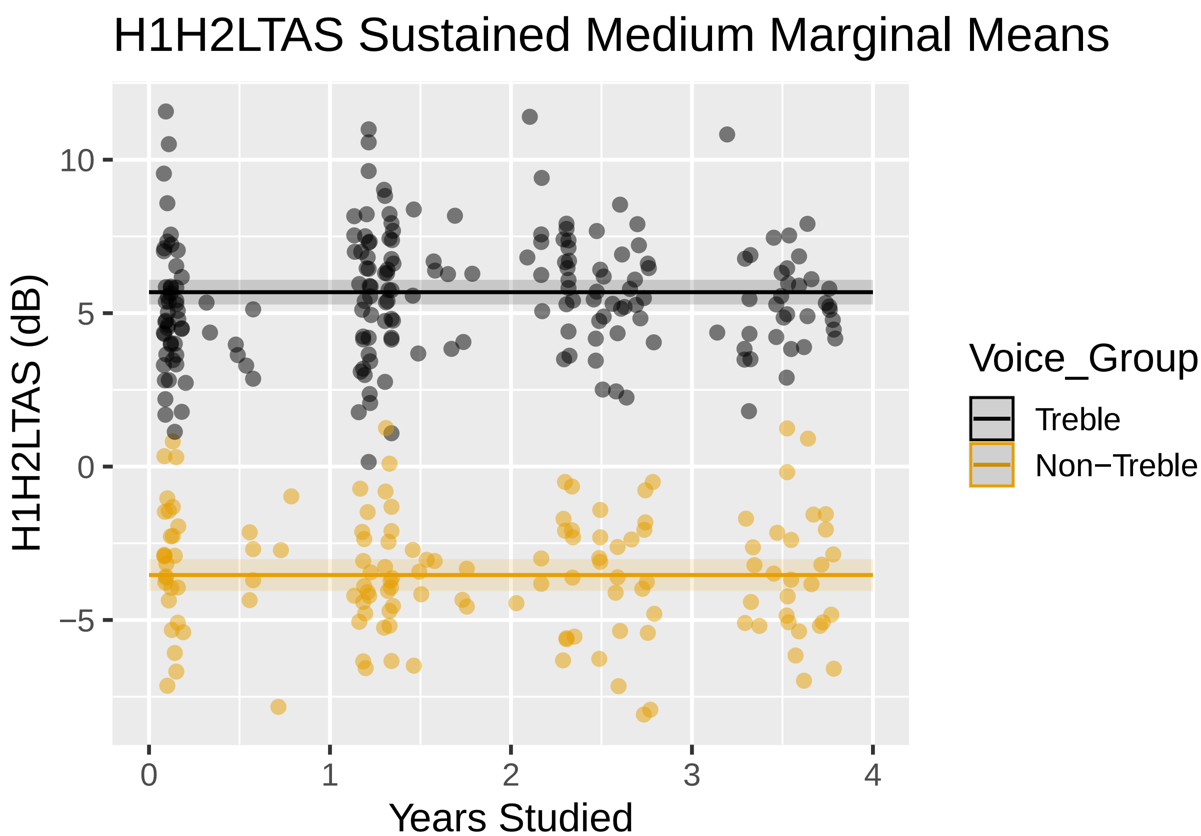
```
## `geom_smooth()` using formula = 'y ~ x'

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

```r
p.fit3_H
```



```r
plot_model(fit3_H, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE, colors = palette
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "H1H2LTAS (dB)", title = "H1H2LTAS Sustained Medium Marginal Means")
```

```
## Data points may overlap. Use the ‘jitter‘ argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```




```r
ggsave("H1H2LTASMediumMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit3_H)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite’s method [
## lmerModLmerTest]
## Formula: H1H2LTAS ~ Voice_Group + (Years | id)
##    Data: klang1
##
## REML criterion at convergence: 1318.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.70808 -0.50892 -0.01983  0.48192  2.45272
##
## Random effects:
```

```
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 3.17398  1.782
##           Years       0.09732  0.312    -0.49
##  Residual             1.35141  1.162
## Number of obs: 344, groups:  id, 117
##
## Fixed effects:
##                        Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)              5.6833     0.2046 113.6419   27.77   <2e-16 ***
## Voice_GroupNon-Treble   -9.2175     0.3333 107.0672  -27.66   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Vc_GrpNn-Tr -0.614
```

```
performance::icc(fit3_H)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##   the conditional random effect variances.
##   Solution: Respecify fixed structure!

## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.701
##   Unadjusted ICC: 0.127
```

Alpha-Ratio

```
#Let's make treble the reference group
klang1$Voice_Group <- relevel(klang1$Voice_Group, ref = "Non-Treble")  # now relevel
```

```
fit0_a <- lmer(alphaRatio ~ Years * Voice_Group + dSPL + (Years | id), data = klang1, REML = TRUE)
anova(fit0_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                   Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years             0.2510  0.2510     1  99.244  0.6124    0.4357
## Voice_Group      10.6419 10.6419     1 104.623 25.9627 1.552e-06 ***
## dSPL              0.2316  0.2316     1 250.295  0.5651    0.4529
## Years:Voice_Group 0.0966  0.0966     1  99.420  0.2357    0.6284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Remove dSPL
fit1_a <- lmer(alphaRatio ~ Years * Voice_Group + (Years | id), data = klang1, REML = TRUE)
anova(fit1_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                     Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
## Years               0.2547  0.2547     1  99.507  0.6233    0.4317
## Voice_Group        10.4299 10.4299     1 104.081 25.5249 1.874e-06 ***
## Years:Voice_Group   0.0905  0.0905     1  99.507  0.2215    0.6389
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fit2_a <- update(fit1_a, . ~ . - Years:Voice_Group)
anova(fit2_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## Years        0.2194  0.2194     1 104.72  0.5369    0.4654
## Voice_Group 21.9513 21.9513     1 105.36 53.7237 4.906e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fit3_a <- update(fit2_a, . ~ . - Years)
anova(fit3_a)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##             Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
## Voice_Group 22.122  22.122     1 105.2  54.266 4.112e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(fit0_a, fit1_a, fit2_a, fit3_a, digits = 10)
```

```
## refitting model(s) with ML (instead of REML)
```

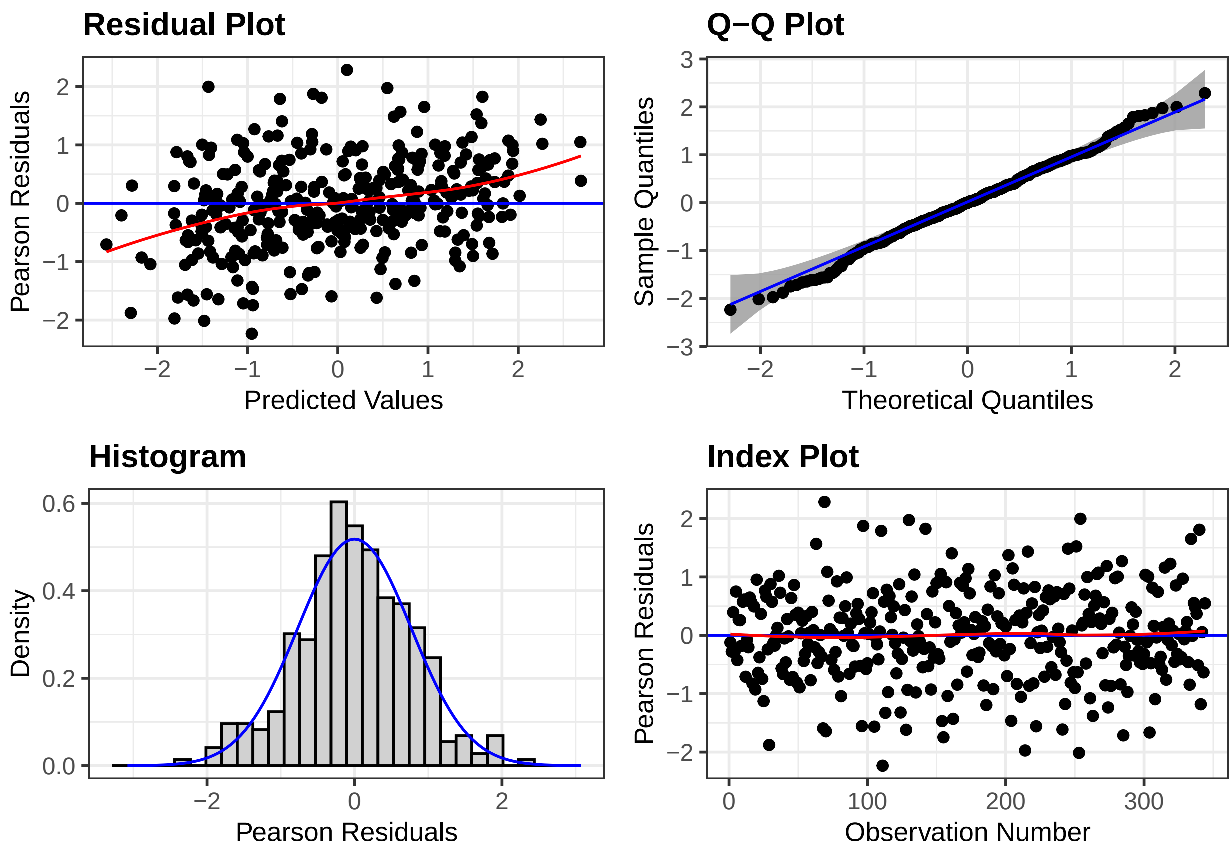
```
## Data: klang1
## Models:
## fit3_a: alphaRatio ~ Voice_Group + (Years | id)
## fit2_a: alphaRatio ~ Years + Voice_Group + (Years | id)
## fit1_a: alphaRatio ~ Years * Voice_Group + (Years | id)
## fit0_a: alphaRatio ~ Years * Voice_Group + dSPL + (Years | id)
##        npar    AIC     BIC  logLik deviance  Chisq Df Pr(>Chisq)
## fit3_a    6 962.68  985.72 -475.34   950.68
## fit2_a    7 964.14  991.03 -475.07   950.14 0.5354  1     0.4643
## fit1_a    8 965.92  996.65 -474.96   949.92 0.2232  1     0.6366
## fit0_a    9 967.35 1001.92 -474.68   949.35 0.5694  1     0.4505
```

```
p.fit3_a <- resid_panel(fit3_a, plots = c("resid", "qq", "hist", "index"), smoother = TRUE, qqbands = TI
```

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

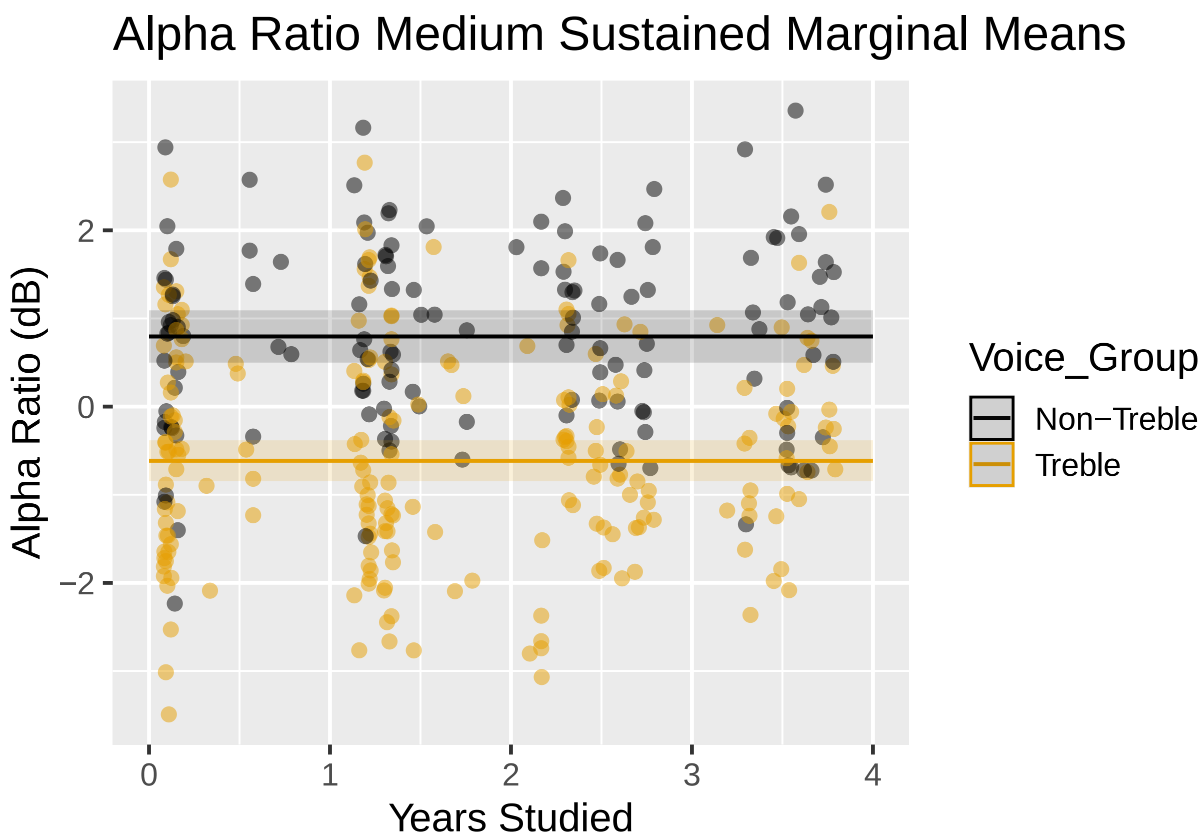
```
## 'geom_smooth()' using formula = 'y ~ x'
```

```
p.fit3_a
```




```
plot_model(fit3_a, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE, colors = palette
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "Alpha Ratio (dB)", title = "Alpha Ratio Medium Sustained Marginal Mean
```

```
## Data points may overlap. Use the ‘jitter‘ argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```




```r
ggsave("AlphaMediumMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit3_a)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: alphaRatio ~ Voice_Group + (Years | id)
##    Data: klang1
##
## REML criterion at convergence: 955.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.23334 -0.46638  0.00131  0.50417  2.28475
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 1.27060  1.1272
##           Years       0.09595  0.3098   -0.60
##  Residual             0.40766  0.6385
## Number of obs: 344, groups:  id, 117
##
## Fixed effects:
##                     Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)           0.7954     0.1509 100.8358   5.273 7.69e-07 ***
```

```
## Voice_GroupTreble  -1.4099     0.1914 105.2015  -7.367 4.11e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr)
## Voc_GrpTrbl -0.788
```

```
performance::icc(fit3_a)
```

```
## Warning: Random slopes not present as fixed effects. This artificially inflates
##   the conditional random effect variances.
##   Solution: Respecify fixed structure!
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.757
##   Unadjusted ICC: 0.589
```

CPPs

```
#Let's make treble the reference group
klang1$Voice_Group <- relevel(klang1$Voice_Group, ref = "Treble")  # now relevel
```

```
fit0_c <- lmer(CPPs ~ Years * Voice_Group + dSPL + (Years | id), data = klang1, REML = TRUE)
```

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

```
anova(fit0_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF  DenDF  F value  Pr(>F)
## Years               0.494   0.494     1 154.73   0.4586 0.49931
## Voice_Group       255.599 255.599     1 146.44 237.1433 < 2e-16 ***
## dSPL                0.170   0.170     1 317.66   0.1578 0.69147
## Years:Voice_Group   7.283   7.283     1 154.87   6.7569 0.01024 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Remove dSPL
fit1_c <- lmer(CPPs ~ Years * Voice_Group + (Years | id), data = klang1, REML = TRUE)
```

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

```
### Not enough variance in individual slopes, change to random intercepts:
fit1_c <- lmer(CPPs ~ Years * Voice_Group + (1 | id), data = klang1, REML = TRUE)
anova(fit1_c)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##                    Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
## Years               0.830   0.830     1 272.01   0.7497  0.387322
## Voice_Group       200.470 200.470     1 254.78 180.9897 < 2.2e-16 ***
## Years:Voice_Group   8.808   8.808     1 272.01   7.9523  0.005156 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
margin1 <- ggpredict(fit1_c, c("Years", "Voice_Group"), ci_level = 0.95)
print(margin1)
```

```
## # Predicted values of CPPs
##
## Voice_Group: Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     14.16 | 13.83, 14.48
##     1 |     14.35 | 14.09, 14.61
##     2 |     14.54 | 14.29, 14.79
##     3 |     14.73 | 14.43, 15.04
##     4 |     14.92 | 14.52, 15.33
##
## Voice_Group: Non-Treble
##
## Years | Predicted |       95% CI
## --------------------------------
##     0 |     17.78 | 17.36, 18.20
##     1 |     17.68 | 17.34, 18.01
##     2 |     17.58 | 17.27, 17.89
##     3 |     17.48 | 17.12, 17.84
##     4 |     17.38 | 16.92, 17.83
##
## Adjusted for:
## * id = 0 (population-level)
```

```r
emm0_c <- emmeans(fit1_c, ~ Years:Voice_Group)
pairs(emm0_c)
```

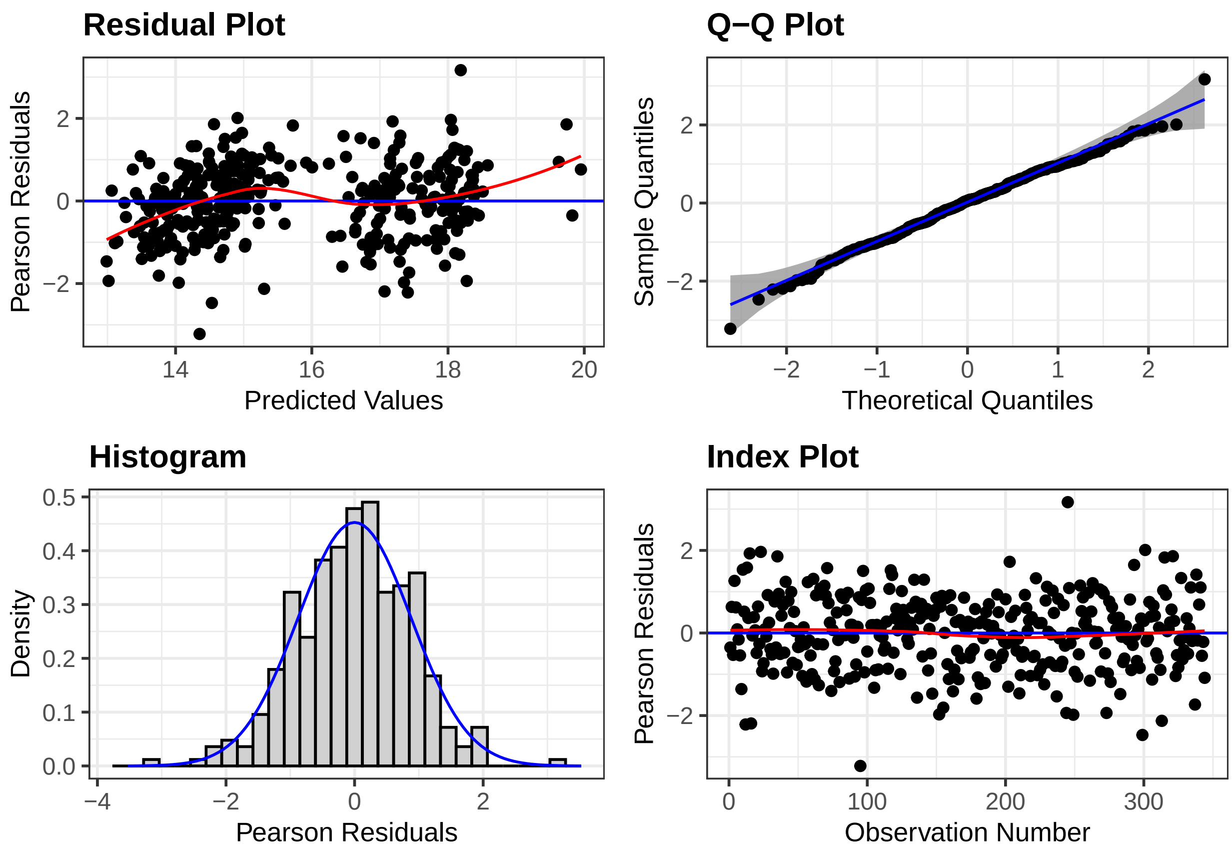
```
##  contrast                                                          estimate
##  Years1.72300892003823 Treble - (Years1.72300892003823 Non-Treble)    -3.12
##     SE  df t.ratio p.value
##  0.201 111 -15.485  <.0001
##
## Degrees-of-freedom method: kenward-roger
```

```r
p.fit1_c <- resid_panel(fit1_c, plots = c("resid", "qq", "hist", "index"), smoother = TRUE, qqbands = TR
```

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

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

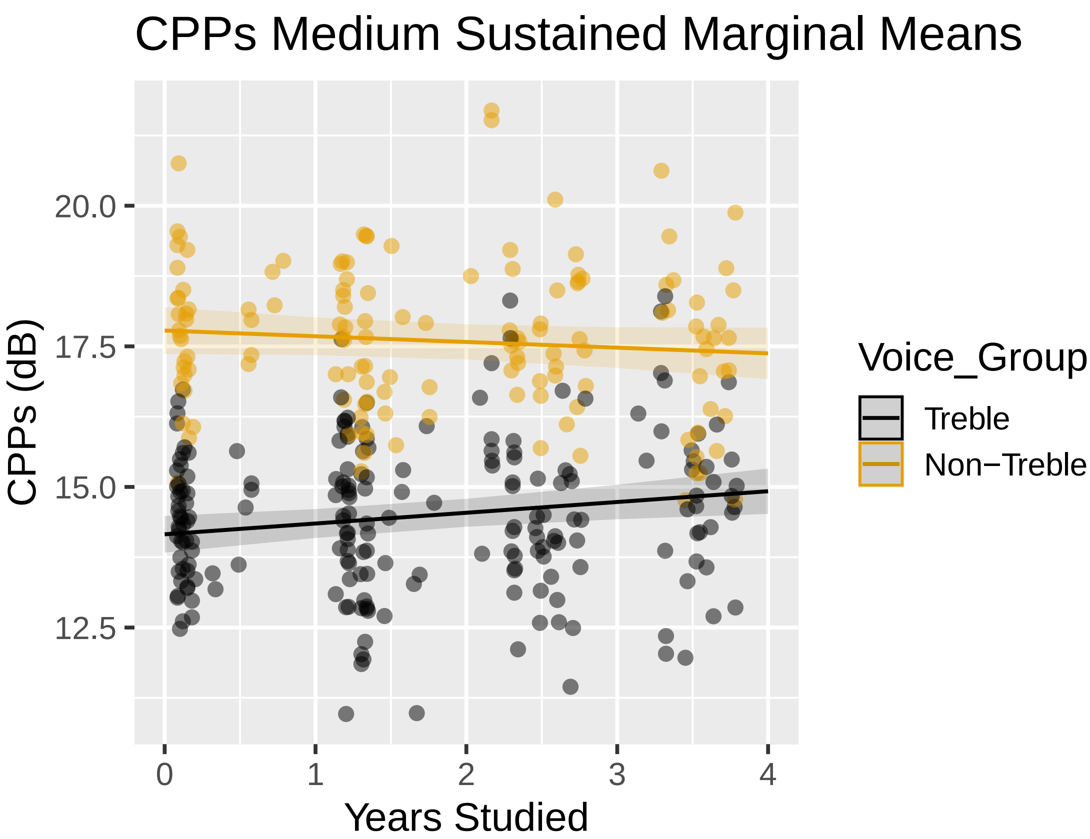
```
p.fit1_c
```




```
fig_c3 <- plot_model(fit1_c, type = "pred", terms = c("Years", "Voice_Group"), show.data = TRUE, colors
  theme_gray(base_size = 15) +
  labs(x = "Years Studied", y = "CPPs (dB)", title = "CPPs Medium Sustained Marginal Means") +
  theme(aspect.ratio = 1)
```

```
## Data points may overlap. Use the ‘jitter‘ argument to add some amount of
##   random variation to the location of data points and avoid overplotting.
```

```
fig_c3
```




```r
ggsave("CPPsMediumMarginal_JASA_Final.pdf", width = 9.25, height = 5.71)
```

```r
summary(fit1_c)
```

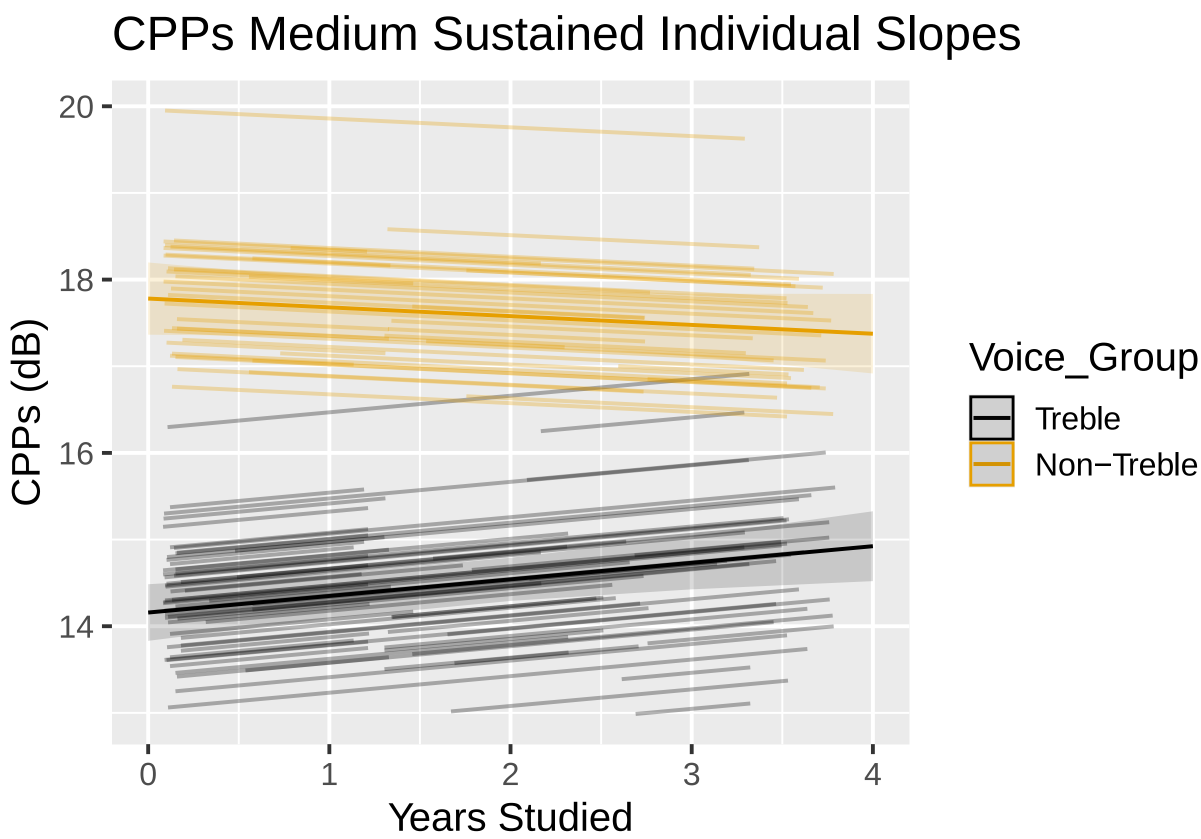
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years * Voice_Group + (1 | id)
##    Data: klang1
##
## REML criterion at convergence: 1140.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2204 -0.5703  0.0605  0.6205  3.1681
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept) 0.7138   0.8448
##  Residual             1.1076   1.0524
## Number of obs: 344, groups:  id, 117
##
## Fixed effects:
##                                Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)                    14.15794    0.16617 257.26516  85.202  < 2e-16 ***
## Years                           0.19138    0.06851 288.37866   2.793  0.00557 **
```

```
## Voice_GroupNon-Treble         3.62317    0.26932 254.78457  13.453  < 2e-16 ***
## Years:Voice_GroupNon-Treble  -0.29284    0.10384 272.00623  -2.820  0.00516 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Years  V_GN-T
## Years       -0.666
## Vc_GrpNn-Tr -0.617  0.411
## Yrs:Vc_GN-T  0.440 -0.660 -0.664
```

```
performance::icc(fit1_c)
```

```
## # Intraclass Correlation Coefficient
##
##     Adjusted ICC: 0.392
##   Unadjusted ICC: 0.170
```

```
# INDIVIDUAL SLOPES clean and prep
klang_clean <- klang1[!is.na(klang1$H1H2LTAS), ]
klang_clean$fitted <- fitted(fit1_c)

# Ensure factors
klang_clean$id <- as.factor(klang_clean$id)
klang_clean$Voice_Group <- as.factor(klang_clean$Voice_Group)

fig_h1 <- plot_model(fit1_c, type = "pred", terms = c("Years", "Voice_Group"), show.data= FALSE,
           colors = palette.colors(palette = "Okabe-Ito")) +
  theme_gray(base_size=15) +
  labs(x = "Years Studied",
        y = "CPPs (dB)",
        title = "CPPs Medium Sustained Individual Slopes")
#fig_h1
#ggsave("H1H2LTASMarginal_JASA_Final.pdf", width=9.25, height=5.71)

# Plot all subjects, colored by Voice_Group
fig_h1 + geom_line(data = klang_clean,
            aes(x = Years, y = fitted, group = id, color = Voice_Group),
            alpha = 0.3, linewidth = 0.7, inherit.aes = FALSE) #+
```




```
ggsave("CPPsMarginalMedium_JASA_Individual_Final.pdf", width=9.25, height=5.71)
```

```
fit0_c_f <- lmer(CPPs ~ Years + dSPL + (Years | id), data = klang1 %>% filter(Gender == "Female"), REML
```

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

```
anova(fit0_c_f)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##       Sum Sq Mean Sq NumDF   DenDF F value   Pr(>F)
## Years 6.8403  6.8403     1  69.312  7.1048 0.009557 **
## dSPL  0.1988  0.1988     1 147.864  0.2065 0.650167
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Remove dSPL
fit1_c_f <- lmer(CPPs ~ Years + (Years | id), data = klang1 %>% filter(Gender == "Female"), REML = TRUE
summary(fit1_c_f)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: CPPs ~ Years + (Years | id)
##    Data: klang1 %>% filter(Gender == "Female")
```

```
##
## REML criterion at convergence: 606.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5721 -0.5373  0.0340  0.6275  1.9780
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.24519  0.4952
##           Years       0.05445  0.2333   0.86
##  Residual             0.95370  0.9766
## Number of obs: 188, groups:  id, 68
##
## Fixed effects:
##             Estimate Std. Error       df t value Pr(>|t|)
## (Intercept) 14.11851    0.14341 42.26271  98.447   <2e-16 ***
## Years        0.20497    0.07603 36.65604   2.696   0.0105 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr)
## Years -0.554
```

```
p_values_H <- summary(fit0_H)$coefficients[, "Pr(>|t|)"]
p_values_a <- summary(fit0_a)$coefficients[, "Pr(>|t|)"]
p_values_c <- summary(fit0_c)$coefficients[, "Pr(>|t|)"]

all_p_values <- c(p_values_H, p_values_a, p_values_c)
p_1 <- all_p_values
p_1_f <- summary(fit0_c_f)$coefficients[, "Pr(>|t|)"]
```

```
# Stack plots with superscripts and a caption
combined_plot <- (fig_c1 + labs(tag = "A")) /
                 (fig_c3 + labs(tag = "B")) /
                 (fig_c2 + labs(tag = "C")) +
  plot_annotation(
    tag_levels = "A"
    #title = "Figure 1"#,
    #caption = "CPPs marginal means for sustained high, repertoire, and medium sustained contexts."
  ) +
  plot_layout(heights = c(1, 1, 1))  # Equal heights

# Save the combined plot with fixed dimensions
#combined_plot
ggsave("combined_CPPs_JASA_test_Final.pdf", plot = combined_plot, width = 6, height = 12, dpi = 300)
```

```
# Stack plots with superscripts and a caption
combined_plot <- (fig_h1 + labs(tag = "A")) /
                 (fig_h2 + labs(tag = "B")) +
  plot_annotation(
    tag_levels = "A"
```

```
    #title = "Figure 1"#,
    #caption = "CPPs marginal means for sustained high, repertoire, and medium sustained contexts."
  ) +
  plot_layout(heights = c(1, 1))  # Equal heights

# Save the combined plot with fixed dimensions
#combined_plot
ggsave("combined_H1H2LTAS_JASA_test_Final.pdf", plot = combined_plot, width = 6, height = 8, dpi = 300)
```

```
# Stack plots with superscripts and a caption
combined_plot <- (fig_slopes1 + labs(tag = "A")) /
                 (fig_slopes2 + labs(tag = "B")) /
                 (fig_slopes3 + labs(tag = "C")) +
  plot_annotation(
    tag_levels = "A"
    #title = "Figure 1"#,
    #caption = "CPPs marginal means for sustained high, repertoire, and medium sustained contexts."
  ) +
  plot_layout(heights = c(1, 1, 1))  # Equal heights

# Save the combined plot with fixed dimensions
#combined_plot
ggsave("combined_Slopes.pdf", plot = combined_plot, width = 6, height = 8, dpi = 300)
```

P-Value Adjustments

```
total_p_values <- c(p_2, p_6, p_1, p_2_f, p_6_f, p_1_f)

# Apply the Benjamini-Hochberg correction using p.adjust()
adjusted_p_values_total <- p.adjust(total_p_values, method = "BH")

#Final DFs
p_values_BH_2 <- data.frame(
  H1H2LTAS = adjusted_p_values_total[1:5],
  AlphaRatio = adjusted_p_values_total[6:10],
  CPPs = adjusted_p_values_total[11:15]#,
)

p_values_BH_6 <- data.frame(
  H1H2LTAS = adjusted_p_values_total[16:20],
  AlphaRatio = adjusted_p_values_total[21:25],
  CPPs = adjusted_p_values_total[26:30]#,
)

p_values_BH_1 <- data.frame(
  H1H2LTAS = adjusted_p_values_total[31:35],
  AlphaRatio = adjusted_p_values_total[36:40],
  CPPs = adjusted_p_values_total[41:45]#,
)
```

```r
p_values_BH_f <- data.frame(
  Intercept = adjusted_p_values_total[c(46, 49, 52)],
  Years = adjusted_p_values_total[c(47, 50, 53)],
  dSPL = adjusted_p_values_total[c(48, 51, 54)]
)
rownames(p_values_BH_f) <- c("High", "Rep", "Med")

# Display results with formatted tables
cat("### High Frequency Sustained Phonation\n")
```

```
## ### High Frequency Sustained Phonation
```

```r
kable(p_values_BH_2, digits = 3)
```

	H1H2LTAS	AlphaRatio	CPPs
(Intercept)	0.000	0.082	0.000
Years	0.024	0.020	0.003
Voice_GroupNon-Treble	0.000	0.637	0.000
dSPL	0.280	0.455	0.019
Years:Voice_GroupNon-Treble	0.019	0.029	0.815

```r
cat("\n### Medium Frequency Sustained Phonation\n")
```

```
##
## ### Medium Frequency Sustained Phonation
```

```r
kable(p_values_BH_1, digits = 3)
```

	H1H2LTAS	AlphaRatio	CPPs
(Intercept)	0.000	0.017	0.000
Years	0.141	0.490	0.025
Voice_GroupNon-Treble	0.000	0.000	0.000
dSPL	0.069	0.532	0.732
Years:Voice_GroupNon-Treble	0.054	0.693	0.020

```r
cat("\n### Repertoire\n")
```

```
##
## ### Repertoire
```

```r
kable(p_values_BH_6, digits = 3)
```

	H1H2LTAS	AlphaRatio	CPPs
(Intercept)	0.000	0.000	0.000

	H1H2LTAS	AlphaRatio	CPPs
Years	0.012	0.172	0.036
Voice_GroupNon-Treble	0.000	0.000	0.000
dSPL	0.574	0.193	0.987
Years:Voice_GroupNon-Treble	0.029	0.490	0.024

```
cat("\n### Female CPPs\n")
```

```
##
## ### Female CPPs
```

```
kable(p_values_BH_f, digits = 3)
```

	Intercept	Years	dSPL
High	0	0.000	0.066
Rep	0	0.054	0.871
Med	0	0.020	0.702