M21 LDT ERP HC SEMANTIC SENSITIVITY

Joanna Morris

2025-06-25

Contents

Set parameters

Set chunk parameters

Load libraries

Set ggplot parameters

Define standard error of the mean function

1 Load and format data files

```
erp_2 <- read_csv("m21_ldt_mea_200300_050050_1.csv")
erp_4 <- read_csv("m21_ldt_mea_300500_050050_1.csv")
dmg_lng_vsl <- read_csv("demo_lang_vsl_pca_hc.csv")
```

Now we extract `SubjID` from the `ERPset` column

```
# Remove '_LDT_diff_waves' from each string in the ERPset column
# This code first renames the column and then applies the `str_replace` function
# to the newly renamed column.
```

```r
erp_2 <- erp_2 |>
  rename(SubjID = ERPset) |>
  mutate(SubjID = str_replace(SubjID, "_LDT_diff_waves", "")) |>
  mutate(binlabel = str_replace(binlabel, "Critical_", "")) |>
  mutate(binlabel = str_replace(binlabel, "_family", "")) |>
  select(-mlabel)

erp_4 <- erp_4 |>
  rename(SubjID = ERPset) |>
  mutate(SubjID = str_replace(SubjID, "_LDT_diff_waves", "")) |>
  mutate(binlabel = str_replace(binlabel, "Critical_", "")) |>
  mutate(binlabel = str_replace(binlabel, "_family", "")) |>
  select(-mlabel)
```

We then join the ERP data and language into a single data frame

```r
n250 <- erp_2 |>
  left_join(dmg_lng_vsl, by = "SubjID") |>
  select(SubjID, everything())
n400 <- erp_4 |>
  left_join(dmg_lng_vsl, by = "SubjID") |>
  select(SubjID, everything())
```

Divide into word, non-word and difference wave dataframes

Then we do some more formatting and cleanup of the dataframes.We create separate columns, one for each independent variable (anteriority, laterality, morphological family size). To do this we have to use `seperate` function from the `stringr` package. Run `vignette("programming", package = "dplyr")` to see more about `tidy-selection` and `tidy-evaluation`.

```r
# Words
n250_words <- n250_words |>
  separate(binlabel, into = c("trial_type","family_size"), sep = "_", remove = TRUE) |>
  select(-trial_type)
n250_words_b <- n250_words_b |>
  separate(binlabel, into = c("trial_type", "family_size","tmp1", "base_freq", "tmp2"), sep = "_", remove = TRUE) |>
  select(-c(trial_type, tmp1, tmp2))

n400_words <- n400_words |>
  separate(binlabel, into = c("trial_type","family_size"), sep = "_", remove = TRUE) |>
  select(-trial_type)
n400_words_b <- n400_words_b |>
  separate(binlabel, into = c("trial_type", "family_size","tmp1", "base_freq", "tmp2"), sep = "_", remove = TRUE) |>
  select(-c(trial_type, tmp1, tmp2))

# Assuming your data frame is named 'df' and the column is named 'your_column'
n250_words_b$Semantic_Sensitivity[n250_words_b$Semantic_Sensitivity == "Low"] <- "Low Sensitivity"
n250_words_b$Semantic_Sensitivity[n250_words_b$Semantic_Sensitivity == "High"] <- "High Sensitivity"
n250_words_b$base_freq[n250_words_b$base_freq == "Low"] <- "Low Base Frequency"
n250_words_b$base_freq[n250_words_b$base_freq == "High"] <- "High Base Frequency"
n250_words_b$family_size[n250_words_b$family_size == "large"] <- "Large Family"
n250_words_b$family_size[n250_words_b$family_size == "small"] <- "Small Family"

n400_words_b$Semantic_Sensitivity[n400_words_b$Semantic_Sensitivity == "Low"] <- "Low Sensitivity"
n400_words_b$Semantic_Sensitivity[n400_words_b$Semantic_Sensitivity == "High"] <- "High Sensitivity"
n400_words_b$base_freq[n400_words_b$base_freq == "Low"] <- "Low Base Frequency"
n400_words_b$base_freq[n400_words_b$base_freq == "High"] <- "High Base Frequency"
n400_words_b$family_size[n400_words_b$family_size == "large"] <- "Large Family"
n400_words_b$family_size[n400_words_b$family_size == "small"] <- "Small Family"

# Nonwords
n250_nonwords <- n250_nonwords |>
  separate(binlabel, into = c("trial_type", "family_size", "complexity"), sep = "_", remove = TRUE) |>
  select(-trial_type)

n400_nonwords <- n400_nonwords |>
  separate(binlabel, into = c("trial_type", "family_size", "complexity"), sep = "_", remove = TRUE) |>
  select(-trial_type)

# Assuming your data frame is named 'df' and the column is named 'your_column'
n250_nonwords$Semantic_Sensitivity[n250_nonwords$Semantic_Sensitivity == "Low"] <- "Low Sensitivity"
n250_nonwords$Semantic_Sensitivity[n250_nonwords$Semantic_Sensitivity == "High"] <- "High Sensitivity"
n250_nonwords$complexity[n250_nonwords$complexity == "complex"] <- "Complex"
n250_nonwords$complexity[n250_nonwords$complexity == "simple"] <- "Simple"
n250_nonwords$family_size[n250_nonwords$family_size == "large"] <- "Large Family"
n250_nonwords$family_size[n250_nonwords$family_size == "small"] <- "Small Family"

n400_nonwords$Semantic_Sensitivity[n400_nonwords$Semantic_Sensitivity == "Low"] <- "Low Sensitivity"
n400_nonwords$Semantic_Sensitivity[n400_nonwords$Semantic_Sensitivity == "High"] <- "High Sensitivity"
n400_nonwords$complexity[n400_nonwords$complexity == "complex"] <- "Complex"
n400_nonwords$complexity[n400_nonwords$complexity == "simple"] <- "Simple"
n400_nonwords$family_size[n400_nonwords$family_size == "large"] <- "Large Family"
```

```
n400_nonwords$family_size[n400_nonwords$family_size == "small"] <- "Small Family"
```

Now we need to extract just the bins and channels that we intend to analyse. For this analysis we will use 9 channels: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 . We will use the`mutate` function from the `dplyr` package along with the `case_when` function. The `case_when` function is a sequence of two-sided formulas. The left hand side determines which values match this case. The right hand side provides the replacement value.

```
channels_1 <-  c(3, 2, 25, 7, 20, 21, 12, 11, 16)
channels_2 <-  c(3, 2, 29, 8, 23, 24, 14, 13, 19)

# Words
n250_words <- n250_words |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline",grepl("4", chlabel) ~ "Right"))
n250_words$anteriority <- factor(n250_words$anteriority, levels = c("Frontal",  "Central",  "Parietal"))
n250_words$laterality <- factor(n250_words$laterality, levels = c("Left", "Midline", "Right"))

n250_words_b <- n250_words_b |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline",grepl("4", chlabel) ~ "Right"))
n250_words_b$anteriority <- factor(n250_words_b$anteriority, levels = c("Frontal",  "Central",  "Parietal"))
n250_words_b$laterality <- factor(n250_words_b$laterality, levels = c("Left", "Midline", "Right"))

n400_words <- n400_words |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline",grepl("4", chlabel) ~ "Right"))
n400_words$anteriority <- factor(n400_words$anteriority, levels = c("Frontal",  "Central",  "Parietal"))
n400_words$laterality <- factor(n400_words$laterality, levels = c("Left", "Midline", "Right"))

n400_words_b <- n400_words_b |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline",grepl("4", chlabel) ~ "Right"))
n400_words_b$anteriority <- factor(n400_words_b$anteriority, levels = c("Frontal",  "Central",  "Parietal"))
n400_words_b$laterality <- factor(n400_words_b$laterality, levels = c("Left", "Midline", "Right"))

# Nonwords
n250_nonwords <- n250_nonwords |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline", grepl("4", chlabel) ~ "Right"))
n250_nonwords$anteriority <- factor(n250_nonwords$anteriority, levels = c("Frontal",  "Central","Parietal"))
n250_nonwords$laterality <- factor(n250_nonwords$laterality, levels = c("Left", "Midline", "Right"))

n400_nonwords <- n400_nonwords |>
  filter(chindex %in% channels_1) |>
  mutate(anteriority = case_when(grepl("F", chlabel) ~ "Frontal",
                                 grepl("C", chlabel) ~ "Central",
                                 grepl("P", chlabel) ~ "Parietal"),
         laterality = case_when(grepl("3", chlabel) ~ "Left",grepl("z", chlabel) ~ "Midline",
                                grepl("Z", chlabel) ~ "Midline", grepl("4", chlabel) ~ "Right"))
n400_nonwords$anteriority <- factor(n400_nonwords$anteriority, levels = c("Frontal",  "Central","Parietal"))
n400_nonwords$laterality <- factor(n400_nonwords$laterality, levels = c("Left", "Midline", "Right"))
```

2 N250 Word Data

2.1 Compute the ANOVA

```
anova_model_1a <- mixed(
  value ~ Semantic_Sensitivity * family_size * base_freq  +
```

```
    laterality * anteriority  +  # Nuisance variables
    (1 | SubjID),
  data = n250_words_b,
  method = "KR")  # Kenward-Roger approximation for accurate F-tests
# Print ANOVA results
anova_model_1a
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * base_freq + laterality *
|| Model:     anteriority + (1 | SubjID)
|| Data: n250_words_b
||                                        Effect      df         F p.value
|| 1                          Semantic_Sensitivity   1, 59      0.68    .411
|| 2                                   family_size 1, 2121    8.95 **    .003
|| 3                                     base_freq 1, 2121    5.88 *    .015
|| 4                                    laterality 2, 2121      0.30    .744
|| 5                                   anteriority 2, 2121 18.23 ***   <.001
|| 6            Semantic_Sensitivity:family_size 1, 2121      2.51    .113
|| 7              Semantic_Sensitivity:base_freq 1, 2121      0.01    .924
|| 8                       family_size:base_freq 1, 2121 14.30 ***   <.001
|| 9                     laterality:anteriority 4, 2121      0.76    .549
|| 10 Semantic_Sensitivity:family_size:base_freq 1, 2121    4.90 *    .027
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
# Partial Eta Squared
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_1a , partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                 | Eta2 (partial) |       95% CI
|| ------------------------------------------------------------------------------
|| Semantic_Sensitivity                      |           0.01 | [0.00, 1.00]
|| family_size                               |       4.20e-03 | [0.00, 1.00]
|| base_freq                                 |       2.76e-03 | [0.00, 1.00]
|| laterality                                |       2.78e-04 | [0.00, 1.00]
|| anteriority                               |           0.02 | [0.01, 1.00]
|| Semantic_Sensitivity:family_size          |       1.18e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:base_freq            |       4.28e-06 | [0.00, 1.00]
|| family_size:base_freq                     |       6.70e-03 | [0.00, 1.00]
|| laterality:anteriority                    |       1.44e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:base_freq |       2.30e-03 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal (fixed effects) and Conditional (fixed + random effects) R²
r2(anova_model_1a)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.478
||      Marginal R2: 0.024
```

2.2 Significant Effects

Effect	df	F	p.value	
family_size	1, 2121	8.95 **	.003	4.20e-03
base_freq	1, 2121	5.88 *	.015	2.76e-03
family_size:base_freq	1, 2121	14.30 ***	<.001	6.70e-03
Sensitivity:family_size:base_freq	1, 2121	4.90 *	.027	2.30e-03

Main Effects

```
## `family_size` main effect
pairs <- emmeans(anova_model_1a, pairwise ~ family_size, adjust = "bonferroni", pbkrtest.limit = 6480)
(pairs_df <- as.data.frame(pairs$contrasts))
```

```
||  contrast                   estimate         SE   df t.ratio p.value
||  Large Family - Small Family 0.2790251 0.09326883 2121   2.992  0.0028
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
cohensd <- as.data.frame(cohens_d(value ~ family_size, data = n250_words_b))
(family_size_contrasts_df <- bind_cols(pairs_df,cohensd))
```

```
||  contrast                    estimate         SE   df t.ratio p.value
||  Large Family - Small Family 0.2790251 0.09326883 2121   2.992  0.0028
||    Cohens_d   CI      CI_low   CI_high
||  0.09242177 0.95 0.008717338 0.1761052
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
(family_size_means <- as.data.frame(pairs$emmeans))
```

```
||  family_size      emmean        SE    df  lower.CL   upper.CL
||  Large Family -0.6192590 0.2690532 62.71 -1.156968 -0.0815504
||  Small Family -0.8982841 0.2690532 62.71 -1.435993 -0.3605755
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
## `base_freq` main effect
pairs <- emmeans(anova_model_1a, pairwise ~ base_freq, adjust = "bonferroni", pbkrtest.limit = 6480)
(pairs_df <- as.data.frame(pairs$contrasts))
```

```
||  contrast                                  estimate         SE   df t.ratio
||  High Base Frequency - Low Base Frequency -0.2260856 0.09326883 2121  -2.424
||  p.value
||   0.0154
||
|| Results are averaged over the levels of: Semantic_Sensitivity, family_size, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
cohensd <- as.data.frame(cohens_d(value ~ base_freq, data = n250_words_b))
(base_freq_contrasts_df <- bind_cols(pairs_df,cohensd))
```

```
||  contrast                                  estimate         SE   df t.ratio
||  High Base Frequency - Low Base Frequency -0.2260856 0.09326883 2121  -2.424
||  p.value   Cohens_d   CI     CI_low     CI_high
||   0.0154 -0.0755643 0.95 -0.1592348 0.008123396
||
|| Results are averaged over the levels of: Semantic_Sensitivity, family_size, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
(base_freq_means <- as.data.frame(pairs$emmeans))
```

```
||  base_freq               emmean        SE    df  lower.CL   upper.CL
||  High Base Frequency -0.8718144 0.2690532 62.71 -1.409523 -0.3341058
||  Low Base Frequency  -0.6457287 0.2690532 62.71 -1.183437 -0.1080201
||
|| Results are averaged over the levels of: Semantic_Sensitivity, family_size, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

Interactions

```
# `base_freq` x `family_size` interaction

selected_contrasts_famsize <- c("Large Family High Base Frequency - Small Family High Base Frequency",
                                "Large Family Low Base Frequency - Small Family Low Base Frequency")
selected_contrasts_basefreq <- c("Large Family High Base Frequency - Large Family Low Base Frequency",
                                 "Small Family High Base Frequency - Small Family Low Base Frequency")

emmeans_obj <- emmeans(anova_model_1a, pairwise ~ family_size * base_freq, adjust = "bonferroni", pbkrtest.limit = 6480)

# Get selected contrasts and convert the emmGrid object to a dataframe
(contrasts_df <- as.data.frame(emmeans_obj$contrasts))
```

```
||  contrast                                                          estimate
||  Large Family High Base Frequency - Small Family High Base Frequency -0.0736383
||  Large Family High Base Frequency - Large Family Low Base Frequency  -0.5787490
||  Large Family High Base Frequency - Small Family Low Base Frequency   0.0529394
||  Small Family High Base Frequency - Large Family Low Base Frequency  -0.5051107
||  Small Family High Base Frequency - Small Family Low Base Frequency   0.1265777
||  Large Family Low Base Frequency - Small Family Low Base Frequency    0.6316884
||        SE   df t.ratio p.value
||  0.131902 2121  -0.558  1.0000
||  0.131902 2121  -4.388  0.0001
||  0.131902 2121   0.401  1.0000
||  0.131902 2121  -3.829  0.0008
||  0.131902 2121   0.960  1.0000
||  0.131902 2121   4.789  <.0001
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 6 tests
```

```
selected_contrasts_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_famsize, ])
selected_contrasts_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_basefreq,])


cohensd_hi_basefrq <- as.data.frame(cohens_d(value ~ family_size,
                                     data = subset(n250_words_b, base_freq == "High Base Frequency")))
cohensd_lo_basefrq <- as.data.frame(cohens_d(value ~ family_size,
                                      data = subset(n250_words_b, base_freq == "Low Base Frequency")))
cohensd_lrg_fam <- as.data.frame(cohens_d(value ~ base_freq,
                                    data = subset(n250_words_b, family_size == "Large Family")))
cohensd_sml_fam <- as.data.frame(cohens_d(value ~ base_freq,
                                    data = subset(n250_words_b, family_size == "Small Family")))

cohensd_basefrq <- bind_rows(hi_basefrq = cohensd_hi_basefrq,
                             lo_basefrq = cohensd_lo_basefrq,
                             .id = "base_freq")

cohensd_famsize <- bind_rows(lrg_fam = cohensd_lrg_fam,
                             sml_fam = cohensd_sml_fam,
                             .id = "family_size")

(basefreq_contrasts_df <- bind_cols(selected_contrasts_basefrq_df,cohensd_basefrq))
```

```
||  contrast                                                        estimate
||  Large Family High Base Frequency - Large Family Low Base Frequency -0.5787490
||  Small Family High Base Frequency - Small Family Low Base Frequency  0.1265777
||        SE   df t.ratio p.value base_freq     Cohens_d   CI      CI_low
||  0.131902 2121  -4.388  <.0001 hi_basefrq -0.02625453 0.95 -0.14455152
||  0.131902 2121   0.960  0.6747 lo_basefrq  0.21467322 0.95  0.09598588
||    CI_high
||  0.0920544
||  0.3332632
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(sensitivity_contrasts_df <- bind_cols(selected_contrasts_famsize_df,cohensd_famsize))
```

```
||  contrast                                                         estimate
||  Large Family High Base Frequency - Small Family High Base Frequency -0.0736383
||  Large Family Low Base Frequency - Small Family Low Base Frequency    0.6316884
||        SE   df t.ratio p.value family_size    Cohens_d   CI     CI_low
||  0.131902 2121  -0.558  1.0000 lrg_fam     -0.20058496 0.95 -0.3191349
||  0.131902 2121   4.789  <.0001 sml_fam      0.04230879 0.95 -0.0760120
||      CI_high
||  -0.08194398
||   0.16061027
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(famsize.basefreq_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  family_size  base_freq              emmean        SE    df   lower.CL
||  Large Family High Base Frequency -0.9086335 0.2770183 70.46 -1.4610662
||  Small Family High Base Frequency -0.8349952 0.2770183 70.46 -1.3874279
||  Large Family Low Base Frequency  -0.3298845 0.2770183 70.46 -0.8823172
||  Small Family Low Base Frequency  -0.9615729 0.2770183 70.46 -1.5140056
||    upper.CL
||  -0.3562009
||  -0.2825626
||   0.2225481
||  -0.4091403
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Semantic Sensitivity x Family Size x Base Frequency

selected_contrasts_hisem_famsize <- c("High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency",
                                      "High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency")
selected_contrasts_losem_famsize <- c("Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency",
                                      "Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency")
selected_contrasts_hisem_basefrq <- c("High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency",
                                      "High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency")
selected_contrasts_losem_basefrq <- c("Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency",
                                      "Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency")
```

```r
emmeans_obj <- emmeans(anova_model_1a,  pairwise ~ Semantic_Sensitivity * family_size * base_freq, adjust = "bonferroni",  pbkrtest.limit = 6480)

contrasts_df <- as.data.frame(emmeans_obj$contrasts)
selected_contrasts_hisem_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_hisem_famsize, ])
selected_contrasts_losem_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_losem_famsize, ])
selected_contrasts_hisem_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_hisem_basefrq, ])
selected_contrasts_losem_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_losem_basefrq, ])

cohensd_1 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n250_words_b, Semantic_Sensitivity == "High Semantic" & family_size == "Large
cohensd_2 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n250_words_b, Semantic_Sensitivity == "High Semantic" & family_size == "Small

cohensd_3 <- as.data.frame(cohens_d(value ~ base_freq,  data = subset(n250_words_b, Semantic_Sensitivity == "Low Semantic" & family_size == "Large
cohensd_4 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n250_words_b, Semantic_Sensitivity == "Low Semantic" & family_size == "Small

cohensd_5 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n250_words_b, Semantic_Sensitivity == "High Semantic" & base_freq == "High
cohensd_6 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n250_words_b, Semantic_Sensitivity == "High Semantic" & base_freq == "Low

cohensd_7 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n250_words_b, Semantic_Sensitivity == "Low Semantic" & base_freq == "High
cohensd_8 <- as.data.frame(cohens_d(value ~ family_size, data = subset(n250_words_b, Semantic_Sensitivity == "Low Semantic" & base_freq == "Low Ba

cohensd_hisem_famsize_df<- bind_rows(cohensd_1, cohensd_2)
cohensd_losem_famsize_df <- bind_rows(cohensd_3, cohensd_4)
cohensd_hisem_basefrq_df <- bind_rows(cohensd_5, cohensd_6)
cohensd_losem_basefrq_df <- bind_rows(cohensd_7, cohensd_8)


(hisem_famsize_df <- bind_cols(selected_contrasts_hisem_famsize_df, cohensd_hisem_famsize_df))
```

```
||  contrast
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value   Cohens_d   CI     CI_low
||  -0.7940609 0.1850023 2121  -4.292  <.0001 -0.2695937 0.95 -0.4361718
||   0.3241147 0.1850023 2121   1.752  0.1599  0.1094160 0.95 -0.0567014
||     CI_high
||  -0.1027753
||   0.2754352
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```r
(losem_famsize_df <- bind_cols(selected_contrasts_losem_famsize_df, cohensd_losem_famsize_df))
```

```
||  contrast
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value    Cohens_d   CI     CI_low
||  -0.3634370 0.1880604 2121  -1.933  0.1068 -0.12847553 0.95 -0.2972773
||  -0.0709593 0.1880604 2121  -0.377  1.0000 -0.02238617 0.95 -0.1910679
||    CI_high
||  0.0404454
||  0.1463164
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```r
(hisem_basefrq_df <- bind_cols(selected_contrasts_hisem_basefrq_df, cohensd_hisem_basefrq_df) )
```

```
||  contrast
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value   Cohens_d   CI      CI_low
||  -0.4278136 0.1850023 2121  -2.312  0.0417 -0.1453696 0.95 -0.31146790
||   0.6903620 0.1850023 2121   3.732  0.0004  0.2328631 0.95  0.06625212
||    CI_high
||  0.0208591
||  0.3992661
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```r
(basefrq_df <- bind_cols(selected_contrasts_losem_basefrq_df, cohensd_losem_basefrq_df))
```

```
||  contrast
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency
||   estimate        SE   df t.ratio p.value   Cohens_d   CI      CI_low   CI_high
||  0.2805370 0.1880604 2121   1.492  0.2718 0.09084545 0.95 -0.07797083 0.2595774
```

```
||  0.5730148 0.1880604 2121   3.047  0.0047 0.19638433 0.95  0.02719923 0.3653878
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(sensitivity.familysize.basefreq_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  Semantic_Sensitivity family_size  base_freq               emmean        SE
||  High Semantic        Large Family High Base Frequency -1.3094337 0.3885386
||  Low Semantic         Large Family High Base Frequency -0.5078333 0.3949612
||  High Semantic        Small Family High Base Frequency -0.8816201 0.3885386
||  Low Semantic         Small Family High Base Frequency -0.7883704 0.3949612
||  High Semantic        Large Family Low Base Frequency  -0.5153728 0.3885386
||  Low Semantic         Large Family Low Base Frequency  -0.1443963 0.3949612
||  High Semantic        Small Family Low Base Frequency  -1.2057348 0.3885386
||  Low Semantic         Small Family Low Base Frequency  -0.7174111 0.3949612
||     df   lower.CL   upper.CL
||  70.46 -2.0842612 -0.5346061
||  70.46 -1.2954688  0.2798021
||  70.46 -1.6564476 -0.1067925
||  70.46 -1.5760059 -0.0007349
||  70.46 -1.2902003  0.2594548
||  70.46 -0.9320318  0.6432392
||  70.46 -1.9805623 -0.4309072
||  70.46 -1.5050466  0.0702244
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

2.3 Model Comparisons

```
# Family Size
reduced_model <- update(anova_model_1a,
                        . ~ . - family_size - Semantic_Sensitivity:family_size - family_size:base_freq - Semantic_Sensitivity:family_size:base_fre
anova(anova_model_1a, reduced_model)
```

```
|| Data: data
|| Models:
|| reduced_model: value ~ Semantic_Sensitivity + base_freq + laterality + anteriority + Semantic_Sensitivity:base_freq + laterality:anteriority +
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                npar    AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model    14 9957.5 10037 -4964.8   9929.5
|| anova_model_1a   18 9952.1 10055 -4958.0   9916.1 13.442  4   0.009305 **
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Base Frequency
reduced_model <- update(anova_model_1a,
                        . ~ . - base_freq - Semantic_Sensitivity:base_freq - base_freq:family_size - Semantic_Sensitivity:base_freq:family_size)
anova(anova_model_1a, reduced_model)
```

```
|| Data: data
|| Models:
|| reduced_model: value ~ Semantic_Sensitivity + family_size + laterality + anteriority + Semantic_Sensitivity:family_size + laterality:anteriorit
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                npar    AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model    14 9952.2 10032 -4962.1   9924.2
|| anova_model_1a   18 9952.1 10055 -4958.0   9916.1 8.0901  4    0.08833 .
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

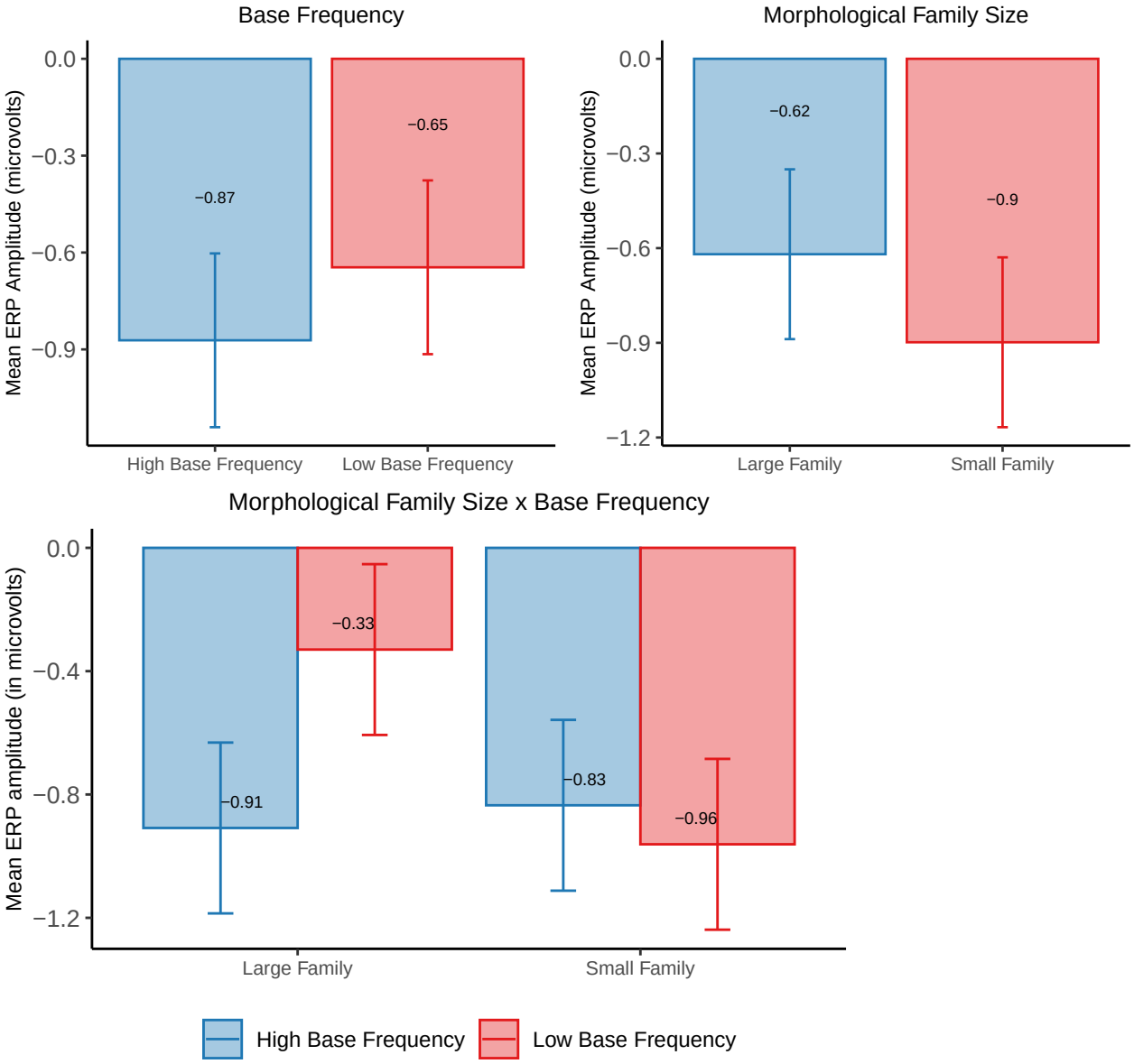
```
# Family Size x Base Frequency
reduced_model_int <- update(anova_model_1a,
  . ~ . - family_size:base_freq - Semantic_Sensitivity:family_size:base_freq)
anova(anova_model_1a, reduced_model_int)
```

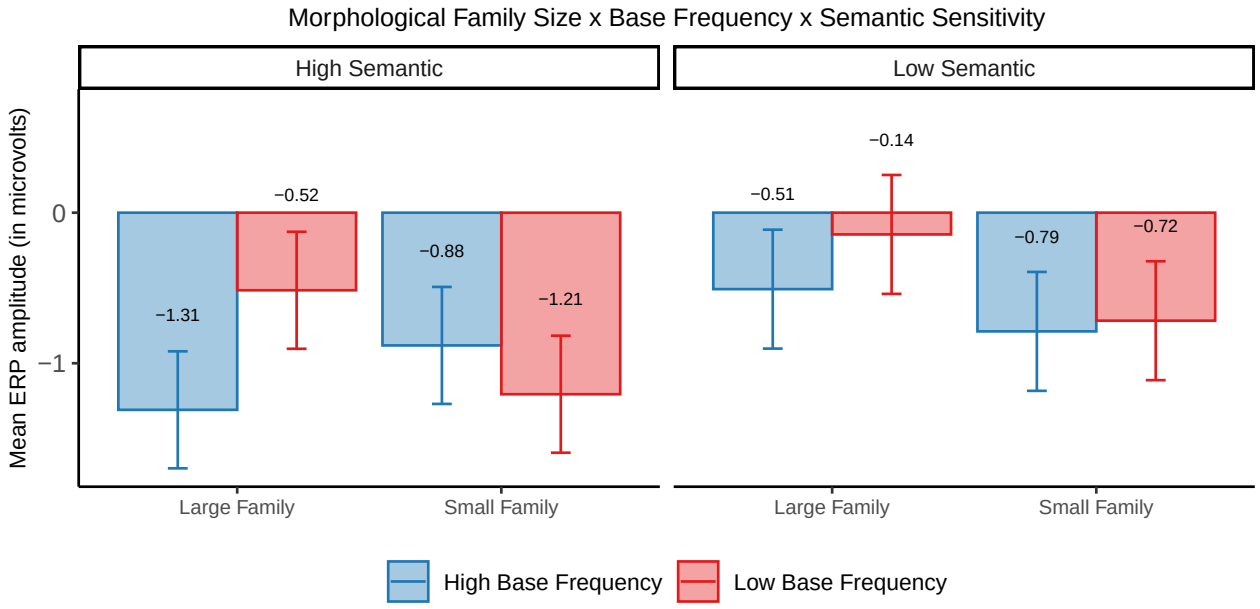
```
|| Data: data
|| Models:
|| reduced_model_int: value ~ Semantic_Sensitivity + family_size + base_freq + laterality + anteriority + Semantic_Sensitivity:family_size + Seman
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                    npar    AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model_int   16 9958.9 10050 -4963.4   9926.9
|| anova_model_1a      18 9952.1 10055 -4958.0   9916.1 10.817  2   0.004478 **
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Semantic Sensitivity x Family Size x Base Frequency
reduced_model_int <- update(anova_model_1a,
  . ~ .  - Semantic_Sensitivity:family_size:base_freq)
anova(anova_model_1a, reduced_model_int)
```

```
|| Data: data
|| Models:
|| reduced_model_int: value ~ Semantic_Sensitivity + family_size + base_freq + laterality + anteriority + Semantic_Sensitivity:family_size + Seman
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                   npar    AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model_int   17 9950.7 10048 -4958.3   9916.7
|| anova_model_1a      18 9952.1 10055 -4958.0   9916.1 0.6018  1     0.4379
```

2.4 Plots





3 N250 Nonword Data

3.1 Compute the ANOVA

```
# Fit the ANOVA/mixed model
anova_model_1b <- mixed(
  value ~ Semantic_Sensitivity * family_size * complexity +
    laterality * anteriority  +  # Nuisance variables
    (1 | SubjID),
  data = n250_nonwords,
  method = "KR"  # Kenward-Roger approximation for accurate F-tests
)

# Print ANOVA results
anova_model_1b
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * complexity + laterality *
|| Model:     anteriority + (1 | SubjID)
|| Data: n250_nonwords
||                                         Effect      df         F p.value
|| 1                           Semantic_Sensitivity   1, 59      0.13    .717
|| 2                                    family_size 1, 2121      0.57    .450
|| 3                                     complexity 1, 2121      0.03    .860
|| 4                                     laterality 2, 2121      0.51    .598
|| 5                                    anteriority 2, 2121 35.73 ***   <.001
|| 6             Semantic_Sensitivity:family_size 1, 2121      2.01    .157
|| 7              Semantic_Sensitivity:complexity 1, 2121 11.16 ***   <.001
|| 8                       family_size:complexity 1, 2121      1.13    .288
|| 9                       laterality:anteriority 4, 2121      0.81    .518
|| 10 Semantic_Sensitivity:family_size:complexity 1, 2121      0.50    .480
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
# Partial Eta Squared
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_1b , partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                | Eta2 (partial) |       95% CI
|| ---------------------------------------------------------------------------
|| Semantic_Sensitivity                     |       2.24e-03 | [0.00, 1.00]
|| family_size                              |       2.69e-04 | [0.00, 1.00]
|| complexity                               |       1.47e-05 | [0.00, 1.00]
|| laterality                               |       4.85e-04 | [0.00, 1.00]
|| anteriority                              |           0.03 | [0.02, 1.00]
|| Semantic_Sensitivity:family_size         |       9.46e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity          |       5.23e-03 | [0.00, 1.00]
```

```
|| family_size:complexity                     |       5.31e-04 | [0.00, 1.00]
|| laterality:anteriority                     |       1.53e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |       2.36e-04 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal (fixed effects) and Conditional (fixed + random effects) R²
r2(anova_model_1b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.408
||      Marginal R2: 0.025
```

```
# Fit the ANOVA/mixed model
anova_model_1b <- mixed(
  value ~ Semantic_Sensitivity * family_size * complexity +
    laterality * anteriority  +  # Nuisance variables
    (1 | SubjID),
  data = n250_nonwords,
  method = "KR"  # Kenward-Roger approximation for accurate F-tests
)

# Print ANOVA results
anova_model_1b
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * complexity + laterality *
|| Model:     anteriority + (1 | SubjID)
|| Data: n250_nonwords
||                                       Effect      df         F p.value
|| 1                          Semantic_Sensitivity   1, 59      0.13    .717
|| 2                                   family_size 1, 2121      0.57    .450
|| 3                                    complexity 1, 2121      0.03    .860
|| 4                                    laterality 2, 2121      0.51    .598
|| 5                                  anteriority 2, 2121 35.73 ***   <.001
|| 6             Semantic_Sensitivity:family_size 1, 2121      2.01    .157
|| 7              Semantic_Sensitivity:complexity 1, 2121 11.16 ***   <.001
|| 8                       family_size:complexity 1, 2121      1.13    .288
|| 9                       laterality:anteriority 4, 2121      0.81    .518
|| 10 Semantic_Sensitivity:family_size:complexity 1, 2121      0.50    .480
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
# Partial Eta Squared
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_1b , partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                  | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------------
|| Semantic_Sensitivity                       |       2.24e-03 | [0.00, 1.00]
|| family_size                                |       2.69e-04 | [0.00, 1.00]
|| complexity                                 |       1.47e-05 | [0.00, 1.00]
|| laterality                                 |       4.85e-04 | [0.00, 1.00]
|| anteriority                                |           0.03 | [0.02, 1.00]
|| Semantic_Sensitivity:family_size           |       9.46e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity            |       5.23e-03 | [0.00, 1.00]
|| family_size:complexity                     |       5.31e-04 | [0.00, 1.00]
|| laterality:anteriority                     |       1.53e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |       2.36e-04 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal (fixed effects) and Conditional (fixed + random effects) R²
r2(anova_model_1b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.408
||      Marginal R2: 0.025
```

3.2 Effects

Effect	df	F	p.value	
Semantic_Sensitivity:complexity	1, 2121	11.16 **	<.001	5.23e-03

Interactions

```r
# `Semantic_Sensitivity` x `Complexity` interaction
selected_contrasts<- c("High Semantic Complex - High Semantic Simple",
                       "Low Semantic Complex - Low Semantic Simple")

emmeans_obj <- emmeans(anova_model_1b, pairwise ~ Semantic_Sensitivity * complexity, adjust = "bonferroni", pbkrtest.limit = 6480)

# Get selected contrasts and convert the emmGrid object to a dataframe
contrasts_df <- as.data.frame(emmeans_obj$contrasts)
selected_contrasts_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts, ])

cohensd_hisem <- as.data.frame(cohens_d(value ~ complexity, data = subset(n250_nonwords, Semantic_Sensitivity == "High Semantic")))
cohensd_losem <- as.data.frame(cohens_d(value ~ complexity,  data = subset(n250_nonwords, Semantic_Sensitivity == "Low Semantic")))

cohensd_df <- bind_rows(cohensd_hisem,cohensd_losem, .id = "sensitivity" )

(sensitivity.complexity_contrasts_df <- bind_cols(selected_contrasts_df,cohensd_df))
```

```
||  contrast                                       estimate        SE   df t.ratio
||  High Semantic Complex - High Semantic Simple -0.2985376 0.1323422 2121  -2.256
||  Low Semantic Complex - Low Semantic Simple    0.3318593 0.1345299 2121   2.467
||  p.value sensitivity   Cohens_d   CI      CI_low    CI_high
||   0.0484 1           -0.1036559 0.95 -0.22105155 0.01378619
||   0.0274 2            0.1181303 0.95 -0.00128082 0.23748671
||
|| Results are averaged over the levels of: family_size, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

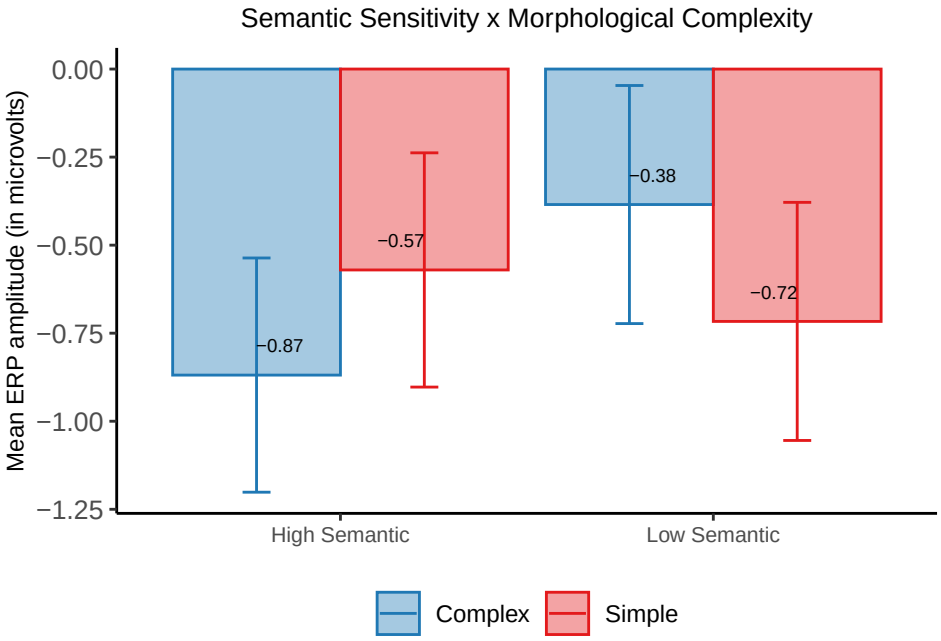
```r
(sensitivity.complexity_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  Semantic_Sensitivity complexity     emmean        SE    df  lower.CL
||  High Semantic        Complex    -0.8689803 0.3325424 63.96 -1.533318
||  Low Semantic         Complex    -0.3847019 0.3380393 63.96 -1.060021
||  High Semantic        Simple     -0.5704427 0.3325424 63.96 -1.234780
||  Low Semantic         Simple     -0.7165611 0.3380393 63.96 -1.391880
||     upper.CL
||  -0.20464292
||   0.29061704
||   0.09389471
||  -0.04124222
||
|| Results are averaged over the levels of: family_size, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

3.3 Model Comparisons

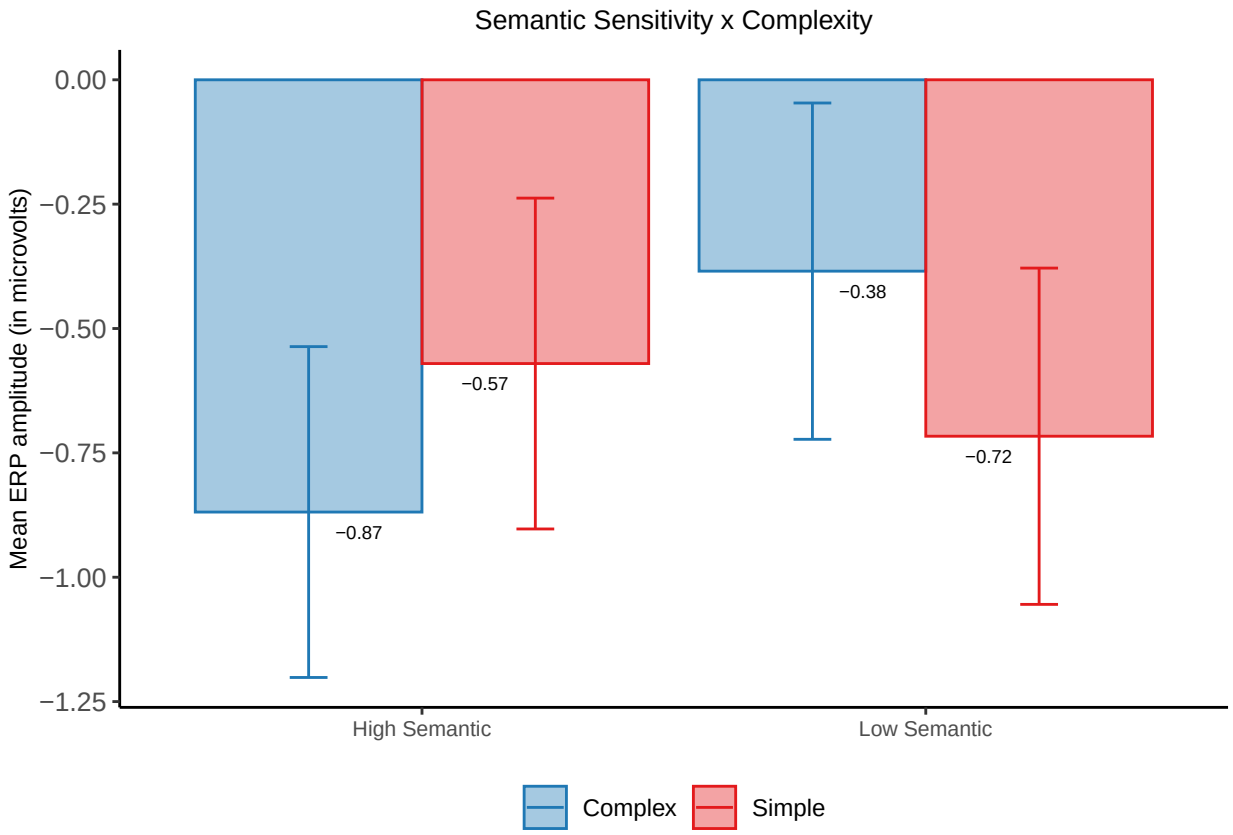
```r
# Reduced model: remove Semantic_Sensitivity:complexity interaction and all higher-order terms that include it
reduced_model <- update(anova_model_1b,
  . ~ . - Semantic_Sensitivity:complexity - Semantic_Sensitivity:family_size:complexity
)
# Model comparison
anova(anova_model_1b, reduced_model)
```

```
|| Data: data
|| Models:
|| reduced_model: value ~ Semantic_Sensitivity + family_size + complexity + laterality + anteriority + Semantic_Sensitivity:family_size + family_s
|| anova_model_1b: value ~ Semantic_Sensitivity * family_size * complexity + laterality * anteriority + (1 | SubjID)
||                npar    AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model    16 9984.9 10076 -4976.4   9952.9
|| anova_model_1b   18 9985.8 10088 -4974.9   9949.8 3.0973  2     0.2125
```



3.4 Plots

```
...
p2 <- sensitivity.complexity_means |>
  ggplot(aes(x = Semantic_Sensitivity,
             y = emmean,
             fill = complexity,
             colour = complexity)) +
  geom_col(alpha = .4, position = position_dodge(.9)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = .15,
                position = position_dodge(.9)) +
  labs(y="Mean ERP amplitude (in microvolts)") +
  geom_text(aes(label = round(emmean, digits = 2)),
                colour = "black",size = 2.5, vjust = 2, position = position_dodge(.5)) +
  # coord_cartesian(ylim = c(-.25, 1.75)) +
  scale_color_custom() +
  scale_fill_custom() +
  labs(title = "Semantic Sensitivity x Complexity") +
  theme( plot.title = element_text(size = 10, hjust = 0.5),
         legend.title = element_blank(),
         axis.title.x = element_blank(),
         axis.text.x = element_text(size = 8))
p2
```



4 N400 Word Data

4.1 Compute the ANOVA

```
anova_model_2a <- mixed(
  value ~ Semantic_Sensitivity * family_size * base_freq  +
    laterality * anteriority  +  # Nuisance variables
    (1 | SubjID),
  data = n400_words_b,
  method = "KR")  # Kenward-Roger approximation for accurate F-tests
# Print ANOVA results
anova_model_2a
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * base_freq + laterality *
|| Model:     anteriority + (1 | SubjID)
|| Data: n400_words_b
||                                       Effect      df          F p.value
|| 1                        Semantic_Sensitivity   1, 59       0.01    .943
|| 2                                 family_size 1, 2121  11.30 ***   <.001
|| 3                                   base_freq 1, 2121     3.07 +    .080
|| 4                                  laterality 2, 2121    4.76 **    .009
|| 5                                 anteriority 2, 2121 104.51 ***   <.001
|| 6            Semantic_Sensitivity:family_size 1, 2121       0.18    .668
|| 7              Semantic_Sensitivity:base_freq 1, 2121       1.01    .314
|| 8                       family_size:base_freq 1, 2121  23.02 ***   <.001
|| 9                     laterality:anteriority 4, 2121       0.78    .541
|| 10 Semantic_Sensitivity:family_size:base_freq 1, 2121      3.12 +    .078
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
# Partial Eta Squared
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_2a , partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                | Eta2 (partial) |       95% CI
```

```
|| ------------------------------------------------------------------------
|| Semantic_Sensitivity                       |       8.82e-05 | [0.00, 1.00]
|| family_size                                |       5.30e-03 | [0.00, 1.00]
|| base_freq                                  |       1.45e-03 | [0.00, 1.00]
|| laterality                                 |       4.47e-03 | [0.00, 1.00]
|| anteriority                                |          0.09 | [0.07, 1.00]
|| Semantic_Sensitivity:family_size           |       8.69e-05 | [0.00, 1.00]
|| Semantic_Sensitivity:base_freq             |       4.77e-04 | [0.00, 1.00]
|| family_size:base_freq                      |          0.01 | [0.00, 1.00]
|| laterality:anteriority                     |       1.46e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:base_freq |       1.47e-03 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal (fixed effects) and Conditional (fixed + random effects) R²
r2(anova_model_2a)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.569
||      Marginal R2: 0.052
```

4.2 Effects

Effect	df	F	p.value	
family_size	1, 2121	11.63 **	.003	5.45e-03
family_size:base_freq	1, 2121	23.02 ***	<.001	0.01
Sensitivity:family_size:base_freq	1, 2121	3.12 +	.078	1.47e-03

Main Effects

```
## `family_size` main effect
emmeans_obj <- emmeans(anova_model_2a, pairwise ~ family_size, adjust = "bonferroni", pbkrtest.limit = 6480)
(contrasts_df <- as.data.frame(emmeans_obj$contrasts))
```

```
||  contrast                    estimate        SE   df t.ratio p.value
||  Large Family - Small Family 0.3572735 0.1062998 2121   3.361  0.0008
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
cohensd_df <- as.data.frame(cohens_d(value ~ family_size, data = n400_words_b))
(family_size_contrasts_df <- bind_cols(contrasts_df,cohensd_df))
```

```
||  contrast                    estimate        SE   df t.ratio p.value
||  Large Family - Small Family 0.3572735 0.1062998 2121   3.361  0.0008
||    Cohens_d   CI     CI_low  CI_high
||  0.09544102 0.95 0.01173328 0.179127
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
```

```
(family_size_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  family_size     emmean        SE   df   lower.CL upper.CL
||  Large Family 0.8422566 0.3571103 61.7  0.1283347 1.556178
||  Small Family 0.4849831 0.3571103 61.7 -0.2289388 1.198905
||
|| Results are averaged over the levels of: Semantic_Sensitivity, base_freq, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

Interactions

```
# `base_freq` x `family_size` interaction
selected_contrasts_famsize <- c("Large Family High Base Frequency - Small Family High Base Frequency",
                                "Large Family Low Base Frequency - Small Family Low Base Frequency")
selected_contrasts_basefreq <- c("Large Family High Base Frequency - Large Family Low Base Frequency",
                                 "Small Family High Base Frequency - Small Family Low Base Frequency")

emmeans_obj <- emmeans(anova_model_2a, pairwise ~ family_size * base_freq, adjust = "bonferroni", pbkrtest.limit = 6480)

contrasts_df <- as.data.frame(emmeans_obj$contrasts)
selected_contrasts_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_famsize, ])
selected_contrasts_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_basefreq,])

cohensd_hi_basefrq <- as.data.frame(cohens_d(value ~ family_size,data = subset(n400_words_b, base_freq == "High Base Frequency")))
cohensd_lo_basefrq <- as.data.frame(cohens_d(value ~ family_size,data = subset(n400_words_b, base_freq == "Low Base Frequency")))
cohensd_lrg_fam <- as.data.frame(cohens_d(value ~ base_freq,data = subset(n400_words_b, family_size == "Large Family")))
cohensd_sml_fam <- as.data.frame(cohens_d(value ~ base_freq,data = subset(n400_words_b, family_size == "Small Family")))
```

```
cohensd_basefrq <- bind_rows(hi_basefrq = cohensd_hi_basefrq,lo_basefrq = cohensd_lo_basefrq,.id = "base_freq")
cohensd_famsize <- bind_rows(lrg_fam = cohensd_lrg_fam,sml_fam = cohensd_sml_fam,.id = "family_size")

(basefreq_contrasts_df <- bind_cols(selected_contrasts_basefrq_df,cohensd_basefrq))
```

```
||  contrast                                                    estimate
||  Large Family High Base Frequency - Large Family Low Base Frequency -0.3237041
||  Small Family High Base Frequency - Small Family Low Base Frequency  0.6962480
||         SE   df t.ratio p.value base_freq    Cohens_d   CI     CI_low   CI_high
||  0.1503306 2121  -2.153  0.0628 hi_basefrq -0.0407526 0.95 -0.1590535 0.0775669
||  0.1503306 2121   4.631  <.0001 lo_basefrq  0.2366583 0.95  0.1178928 0.3553166
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(sensitivity_contrasts_df <- bind_cols(selected_contrasts_famsize_df,cohensd_famsize))
```

```
||  contrast                                                     estimate
||  Large Family High Base Frequency - Small Family High Base Frequency -0.1527026
||  Large Family Low Base Frequency - Small Family Low Base Frequency    0.8672495
||         SE   df t.ratio p.value family_size   Cohens_d   CI      CI_low
||  0.1503306 2121  -1.016  0.6197 lrg_fam     -0.0915178 0.95 -0.20985682
||  0.1503306 2121   5.769  <.0001 sml_fam      0.1786143 0.95  0.06003985
||     CI_high
||  0.02686293
||  0.29710762
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(famsize.basefreq_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  family_size  base_freq             emmean        SE    df   lower.CL
||  Large Family High Base Frequency 0.6804045 0.3649351 67.28 -0.0479526
||  Small Family High Base Frequency 0.8331071 0.3649351 67.28  0.1047499
||  Large Family Low Base Frequency  1.0041086 0.3649351 67.28  0.2757514
||  Small Family Low Base Frequency  0.1368591 0.3649351 67.28 -0.5914981
||   upper.CL
||  1.4087617
||  1.5614643
||  1.7324658
||  0.8652163
||
|| Results are averaged over the levels of: Semantic_Sensitivity, laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Semantic Sensitivity x Family Size x Base Frequency

selected_contrasts_hisem_famsize <- c("High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency",
                                      "High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency")
selected_contrasts_losem_famsize <- c("Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency",
                                      "Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency")
selected_contrasts_hisem_basefrq <- c("High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency",
                                      "High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency")
selected_contrasts_losem_basefrq <- c("Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency",
                                      "Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency")

emmeans_obj <- emmeans(anova_model_2a,  pairwise ~ Semantic_Sensitivity * family_size * base_freq, adjust = "bonferroni",  pbkrtest.limit = 6480)

contrasts_df <- as.data.frame(emmeans_obj$contrasts)
selected_contrasts_hisem_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_hisem_famsize, ])
selected_contrasts_losem_famsize_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_losem_famsize, ])
selected_contrasts_hisem_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_hisem_basefrq, ])
selected_contrasts_losem_basefrq_df <- as.data.frame(emmeans_obj$contrasts[contrasts_df$contrast %in% selected_contrasts_losem_basefrq, ])

cohensd_1 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n400_words_b, Semantic_Sensitivity == "High Semantic" & family_size == "Large
cohensd_2 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n400_words_b, Semantic_Sensitivity == "High Semantic" & family_size == "Small

cohensd_3 <- as.data.frame(cohens_d(value ~ base_freq,  data = subset(n400_words_b, Semantic_Sensitivity == "Low Semantic" & family_size == "Large
cohensd_4 <- as.data.frame(cohens_d(value ~ base_freq, data = subset(n400_words_b, Semantic_Sensitivity == "Low Semantic" & family_size == "Small

cohensd_5 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n400_words_b, Semantic_Sensitivity == "High Semantic" & base_freq == "High
cohensd_6 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n400_words_b, Semantic_Sensitivity == "High Semantic" & base_freq == "Low

cohensd_7 <- as.data.frame(cohens_d(value ~ family_size,  data = subset(n400_words_b, Semantic_Sensitivity == "Low Semantic" & base_freq == "High
cohensd_8 <- as.data.frame(cohens_d(value ~ family_size, data = subset(n400_words_b, Semantic_Sensitivity == "Low Semantic" & base_freq == "Low Ba

cohensd_hisem_famsize_df<- bind_rows(cohensd_1, cohensd_2)
```

```
cohensd_losem_famsize_df <- bind_rows(cohensd_3, cohensd_4)
cohensd_hisem_basefrq_df <- bind_rows(cohensd_5, cohensd_6)
cohensd_losem_basefrq_df <- bind_rows(cohensd_7, cohensd_8)


(hisem_famsize_df <- bind_cols(selected_contrasts_hisem_famsize_df, cohensd_hisem_famsize_df))
```

```
||  contrast
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value   Cohens_d   CI      CI_low
||  -0.4044229 0.2108498 2121  -1.918  0.1105 -0.1122388 0.95 -0.27826319
||   0.9908961 0.2108498 2121   4.700  <.0001  0.2530799 0.95  0.08635814
||    CI_high
||  0.0538864
||  0.4195758
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(losem_famsize_df <- bind_cols(selected_contrasts_losem_famsize_df, cohensd_losem_famsize_df))
```

```
||  contrast
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value    Cohens_d   CI      CI_low
||  -0.2429852 0.2143351 2121  -1.134  0.5141 -0.06935518 0.95 -0.23806070
||   0.4016000 0.2143351 2121   1.874  0.1222  0.10202763 0.95 -0.06681663
||     CI_high
||  0.09941474
||  0.27077720
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(hisem_basefrq_df <- bind_cols(selected_contrasts_hisem_basefrq_df, cohensd_hisem_basefrq_df))
```

```
||  contrast
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value    Cohens_d   CI     CI_low
||  -0.2947348 0.2108498 2121  -1.398  0.3246 -0.07816681 0.95 -0.2441390
||   1.1005842 0.2108498 2121   5.220  <.0001  0.29314070 0.95  0.1261748
||    CI_high
||  0.0878757
||  0.4598457
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(basefrq_df <- bind_cols(selected_contrasts_losem_basefrq_df, cohensd_losem_basefrq_df))
```

```
||  contrast
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency
||    estimate        SE   df t.ratio p.value    Cohens_d   CI      CI_low
||  -0.0106704 0.2143351 2121  -0.050  1.0000 -0.00277579 0.95 -0.17146143
||   0.6339148 0.2143351 2121   2.958  0.0063  0.17587667 0.95  0.00678155
||    CI_high
||  0.1659124
||  0.3448090
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| P value adjustment: bonferroni method for 2 tests
```

```
(sensitivity.familysize.basefreq_means <- as.data.frame(emmeans_obj$emmeans))
```

```
||  Semantic_Sensitivity family_size  base_freq              emmean        SE
||  High Semantic        Large Family High Base Frequency  0.6373943 0.5118484
||  Low Semantic         Large Family High Base Frequency  0.7234148 0.5203093
||  High Semantic        Small Family High Base Frequency  0.9321290 0.5118484
||  Low Semantic         Small Family High Base Frequency  0.7340852 0.5203093
||  High Semantic        Large Family Low Base Frequency   1.0418172 0.5118484
||  Low Semantic         Large Family Low Base Frequency   0.9664000 0.5203093
||  High Semantic        Small Family Low Base Frequency  -0.0587670 0.5118484
||  Low Semantic         Small Family Low Base Frequency   0.3324852 0.5203093
||     df   lower.CL  upper.CL
||  67.28 -0.3841804 1.6589689
```

```
||  67.28 -0.3150465 1.7618761
||  67.28 -0.0894456 1.9537037
||  67.28 -0.3043761 1.7725465
||  67.28  0.0202426 2.0633919
||  67.28 -0.0720613 2.0048613
||  67.28 -1.0803417 0.9628076
||  67.28 -0.7059761 1.3709465
||
|| Results are averaged over the levels of: laterality, anteriority
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

4.3 Model Comparisons

```
# Family Size
reduced_model <- update(anova_model_2a,
                        . ~ . - family_size - Semantic_Sensitivity:family_size - family_size:base_freq - Semantic_Sensitivity:family_size:base_fre
anova(anova_model_1a, reduced_model)
```

```
|| Data: data
|| Models:
|| reduced_model: value ~ Semantic_Sensitivity + base_freq + laterality + anteriority + Semantic_Sensitivity:base_freq + laterality:anteriority +
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                npar     AIC   BIC  logLik deviance Chisq Df Pr(>Chisq)
|| reduced_model    14 10554.3 10634 -5263.1  10526.3
|| anova_model_1a   18  9952.1 10055 -4958.0   9916.1 610.2  4  < 2.2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Family Size x Base Frequency
reduced_model_int <- update(anova_model_2a,
  . ~ . - family_size:base_freq - Semantic_Sensitivity:family_size:base_freq)
anova(anova_model_1a, reduced_model_int)
```

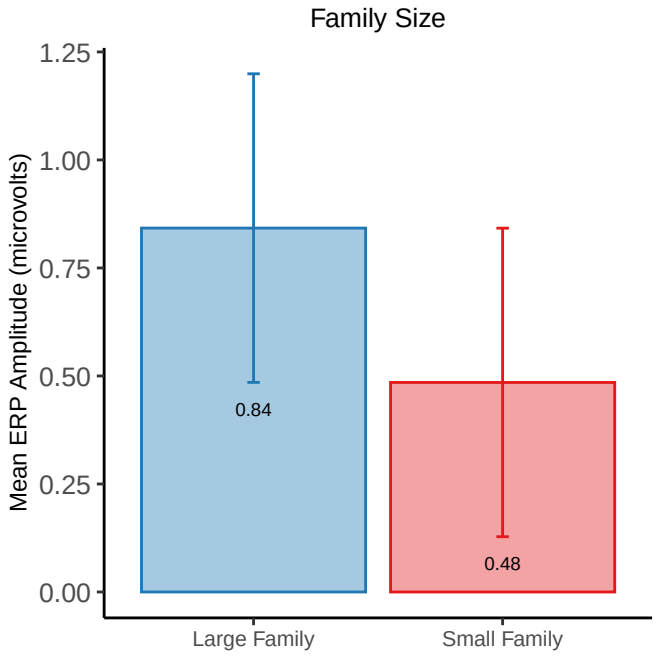
```
|| Data: data
|| Models:
|| reduced_model_int: value ~ Semantic_Sensitivity + family_size + base_freq + laterality + anteriority + Semantic_Sensitivity:family_size + Seman
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                    npar     AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model_int   16 10554.9 10646 -5261.5  10522.9
|| anova_model_1a      18  9952.1 10055 -4958.0   9916.1 606.86  2  < 2.2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

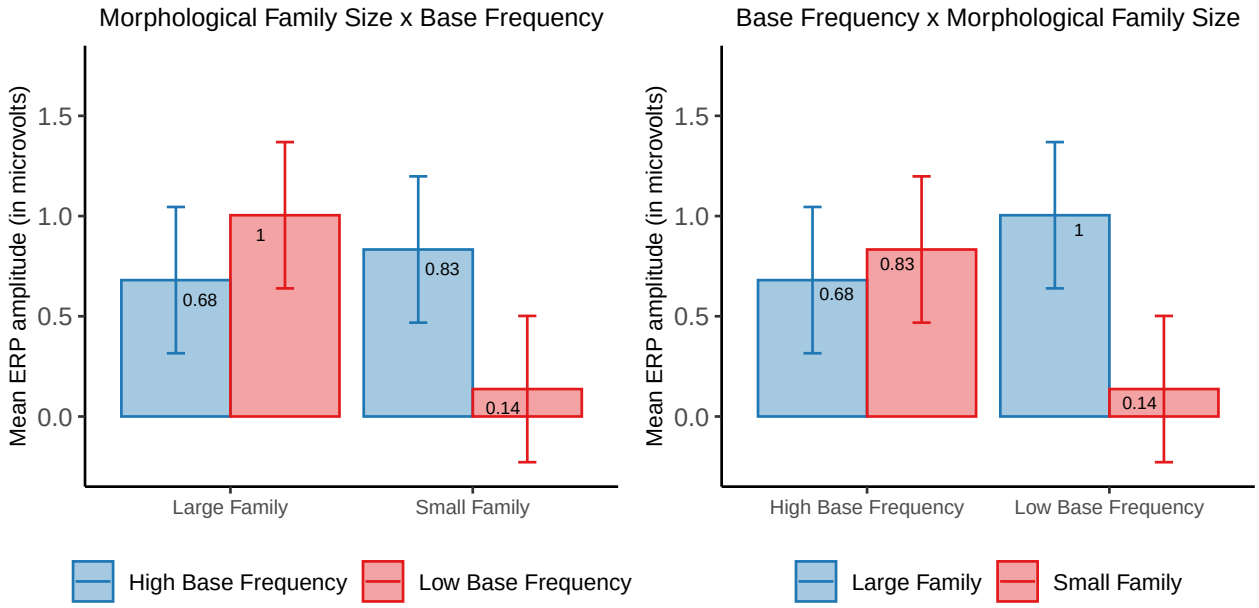
```
# Semantic Sensitivity x Family Size x Base Frequency
reduced_model_int <- update(anova_model_2a,
  . ~ . - Semantic_Sensitivity:family_size:base_freq)
anova(anova_model_1a, reduced_model_int)
```

```
|| Data: data
|| Models:
|| reduced_model_int: value ~ Semantic_Sensitivity + family_size + base_freq + laterality + anteriority + Semantic_Sensitivity:family_size + Seman
|| anova_model_1a: value ~ Semantic_Sensitivity * family_size * base_freq + laterality * anteriority + (1 | SubjID)
||                    npar     AIC   BIC  logLik deviance  Chisq Df Pr(>Chisq)
|| reduced_model_int   17 10537.8 10635 -5251.9  10503.8
|| anova_model_1a      18  9952.1 10055 -4958.0   9916.1 587.73  1  < 2.2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.4 Plots

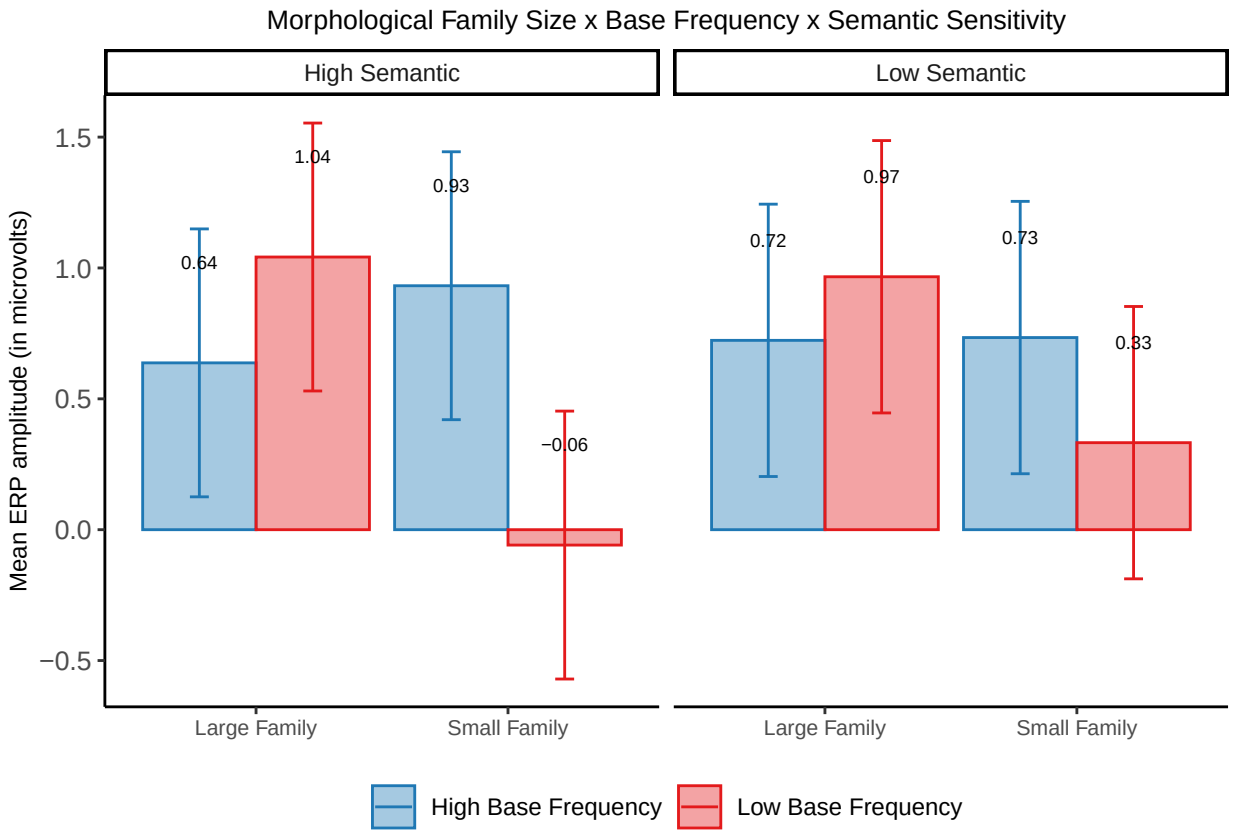
...






```
p4 <- sensitivity.familysize.basefreq_means|>
  ggplot(aes(x = family_size,
             y = emmean,
             fill = base_freq,
             colour = base_freq)) +
  facet_grid( ~ Semantic_Sensitivity) +
  geom_col(alpha = .4, position = position_dodge(.9)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = .15,
                position = position_dodge(.9)) +
  labs(y="Mean ERP amplitude (in microvolts)") +
  geom_text(aes(label = round(emmean, digits = 2)),
                colour = "black",size = 2.5, vjust = -7, position = position_dodge(.9)) +
 # coord_cartesian(ylim = c(-1.7, .7)) +
  scale_color_custom() +
  scale_fill_custom() +
  labs(title = "Morphological Family Size x Base Frequency x Semantic Sensitivity") +
  theme( plot.title = element_text(size = 10, hjust = 0.5),
         legend.title = element_blank(),
         axis.title.x = element_blank(),
         axis.text.x = element_text(size = 8))
```

```
p4
```



5 N400 Nonword Data

5.1 Compute the ANOVA

```
# Fit the ANOVA/mixed model
anova_model_2b <- mixed(
  value ~ Semantic_Sensitivity * family_size * complexity +
    laterality * anteriority  +  # Nuisance variables
    (1 | SubjID),
  data = n400_nonwords,
  method = "KR"  # Kenward-Roger approximation for accurate F-tests
)

# Print ANOVA results
anova_model_2b
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * complexity + laterality *
|| Model:     anteriority + (1 | SubjID)
|| Data: n400_nonwords
||                                         Effect      df          F p.value
|| 1                          Semantic_Sensitivity   1, 59       0.00    .998
|| 2                                   family_size 1, 2121       0.08    .781
|| 3                                    complexity 1, 2121       0.04    .832
|| 4                                    laterality 2, 2121     4.21 *    .015
|| 5                                   anteriority 2, 2121 140.67 ***   <.001
|| 6             Semantic_Sensitivity:family_size 1, 2121       0.30    .584
|| 7              Semantic_Sensitivity:complexity 1, 2121     2.83 +    .093
|| 8                        family_size:complexity 1, 2121       1.66    .198
|| 9                       laterality:anteriority 4, 2121       0.40    .812
|| 10 Semantic_Sensitivity:family_size:complexity 1, 2121       0.07    .787
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```r
# Partial Eta Squared
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_2b , partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                 | Eta2 (partial) |       95% CI
|| ------------------------------------------------------------------------------
|| Semantic_Sensitivity                      |       6.06e-08 | [0.00, 1.00]
|| family_size                               |       3.66e-05 | [0.00, 1.00]
|| complexity                                |       2.11e-05 | [0.00, 1.00]
|| laterality                                |       3.96e-03 | [0.00, 1.00]
|| anteriority                               |           0.12 | [0.10, 1.00]
|| Semantic_Sensitivity:family_size          |       1.41e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity           |       1.33e-03 | [0.00, 1.00]
|| family_size:complexity                    |       7.81e-04 | [0.00, 1.00]
|| laterality:anteriority                    |       7.45e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |       3.43e-05 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```r
# Compute Marginal (fixed effects) and Conditional (fixed + random effects) R²
r2(anova_model_2b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.522
||      Marginal R2: 0.065
```

5.2 Effects

No Significant Effects