M21 LDT ERP HC SEMANTIC SENSITIVITY N250

Joanna Morris

2025-10-17

Contents

Set parameters

Set chunk parameters

Load libraries

Set ggplot parameters

Define standard error of the mean function

1 Load data files

```
dir_path <- "CSV files"

erp_2 <- read_csv(file.path(dir_path, "m21_ldt_mea_200300_050050_1.csv"))
erp_4 <- read_csv(file.path(dir_path, "m21_ldt_mea_300500_050050_1.csv"))
dmg_lng_vsl <- read_csv(file.path(dir_path, "demo_lang_vsl_pca_hc.csv"))
```

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

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

2 Format data files

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`.

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.

3 N250 Word Data

Statistical analysis.

Linear mixed-effects models were fit using the afex::mixed function (method = "KR") to account for both subject-level and electrode-level variability. Each model included random intercepts for participants (SubjID) and electrodes nested within participants (SubjID:chlabel), as well as by-subject random slopes for within-subject factors (Family Size, Complexity, or Base Frequency, depending on the analysis). When a significant interaction was obtained, we probed it using estimated marginal means from the fitted model (emmeans package) to clarify the source of the effect. Because these follow-up contrasts were intended to interpret a significant higher-order interaction rather than to test independent hypotheses, we reported uncorrected p-values (adjust = "none") for interpretive clarity. The robustness of the overall pattern was verified using a Holm correction, which did not change the substantive conclusions.

3.1 Nested ANOVA Model

```
#Fit ANOVA model
anova_model_n250_words_b <- mixed(
    value ~ Semantic_Sensitivity * family_size * base_freq +
    (1 + family_size + base_freq | SubjID) +     # by-subject intercept + slopes
    (1 | SubjID:chlabel),                        # electrode nested within subject
  data   = n250_words_b,
  method = "KR"
)
anova_model_n250_words_b
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * base_freq + (1 +
|| Model:     family_size + base_freq | SubjID) + (1 | SubjID:chlabel)
|| Data: n250_words_b
||                                     Effect      df         F p.value
|| 1                       Semantic_Sensitivity   1, 59      0.68    .411
|| 2                                family_size   1, 59      1.18    .282
|| 3                                  base_freq   1, 59      1.05    .309
|| 4           Semantic_Sensitivity:family_size   1, 59      0.33    .567
|| 5             Semantic_Sensitivity:base_freq   1, 59      0.00    .968
|| 6                      family_size:base_freq 1, 1523 35.00 ***   <.001
|| 7 Semantic_Sensitivity:family_size:base_freq 1, 1523 11.99 ***   <.001
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
m1 <- anova_model_n250_words_b$full_model    # Extract the lmer model
ranova(m1) # Run random effects comparison
```

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| value ~ Semantic_Sensitivity + family_size + base_freq + (1 + family_size + base_freq | SubjID) + (1 | SubjID:chlabel) + Semantic_Sensitivity:f
||                                                    npar  logLik    AIC    LRT Df Pr(>Chisq)
|| <none>                                               16 -4482.9 8997.8
|| family_size in (1 + family_size + base_freq | SubjID)   13 -4797.8 9621.5 629.74  3  < 2.2e-16 ***
|| base_freq in (1 + family_size + base_freq | SubjID)     13 -4712.3 9450.7 458.88  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                   15 -4680.4 9390.8 395.01  1  < 2.2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_n250_words_b, partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------
|| Semantic_Sensitivity                     |           0.01 | [0.00, 1.00]
|| family_size                              |           0.02 | [0.00, 1.00]
|| base_freq                                |           0.02 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size         |       5.58e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:base_freq           |       2.75e-05 | [0.00, 1.00]
|| family_size:base_freq                    |           0.02 | [0.01, 1.00]
|| Semantic_Sensitivity:family_size:base_freq |       7.81e-03 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal(fixed effects only) and Conditional(fixed + random effects) R²
r2(anova_model_n250_words_b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.787
||      Marginal R2: 0.014
```

3.2 Main Effects

Model Fit: Including subject-specific slopes and electrode nesting substantially improves model fit. All partial η^2 values are small ($\leq .02$). This is expected given the large random variance typical of ERP data.

There are no main effects of any single factor.

3.3 Interactions

A robust `family_size × base_freq` interaction that further depends on `Semantic_Sensitivity`. The relationship between `family size` and `base frequency` changes shape depending on whether participants are semantically sensitive or not. Even though the three-way interaction is statistically significant, it explains only a small portion of the variance—typical for ERP component amplitude data.

Effect	df	F	p.value	
family_size:base_freq	1, 1523	35.00 ***	<.001	0.02
Semantic_Sensitivity:family_size:base_freq	1, 1523	11.99 ***	<.001	7.81e-03

3.3.1 `family_size × base_freq` Simple Contrasts

```
# Estimated marginal means for the family_size × base frequency interaction
(emm1a <- emmeans(anova_model_n250_words_b, ~ family_size * base_freq))
```

```
||  family_size  base_freq           emmean    SE   df lower.CL upper.CL
||  Large Family High Base Frequency -0.909 0.279 60.4   -1.467   -0.350
||  Small Family High Base Frequency -0.835 0.349 59.9   -1.534   -0.136
||  Large Family Low Base Frequency  -0.330 0.287 60.3   -0.904    0.245
||  Small Family Low Base Frequency  -0.962 0.341 59.9   -1.644   -0.279
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get all pairswise contrasts
emm1a_contrasts <- contrast(emm1a,  method = "pairwise", by = NULL, adjust = "none")

# Keep only the contrasts you want
# Simple effects of family_size at each level of base_freq
# Simple effects of base_freq at each level of family_size
keep <- c("Large Family High Base Frequency - Small Family High Base Frequency",
          "Large Family Low Base Frequency - Small Family Low Base Frequency",
          "Large Family High Base Frequency - Large Family Low Base Frequency",
          "Small Family High Base Frequency - Small Family Low Base Frequency")
(emm1a_contrasts_filtered <- subset(emm1a_contrasts, contrast %in% keep))
```

```
||  contrast                                                    estimate    SE   df t.ratio p.value
||  Large Family High Base Frequency - Small Family High Base Frequency  -0.0736 0.264 65.5  -0.279  0.7808
||  Large Family High Base Frequency - Large Family Low Base Frequency   -0.5787 0.228 67.9  -2.534  0.0136
||  Small Family High Base Frequency - Small Family Low Base Frequency    0.1266 0.228 67.9   0.554  0.5813
||  Large Family Low Base Frequency - Small Family Low Base Frequency     0.6317 0.264 65.5   2.397  0.0194
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get Confidence Intervals
(emm1a_contrasts_filtered_ci <- confint(emm1a_contrasts_filtered))
```

```
||  contrast                                                    estimate    SE   df lower.CL upper.CL
||  Large Family High Base Frequency - Small Family High Base Frequency  -0.0736 0.264 65.5   -0.600    0.453
||  Large Family High Base Frequency - Large Family Low Base Frequency   -0.5787 0.228 67.9   -1.035   -0.123
||  Small Family High Base Frequency - Small Family Low Base Frequency    0.1266 0.228 67.9   -0.329    0.582
||  Large Family Low Base Frequency - Small Family Low Base Frequency     0.6317 0.264 65.5    0.105    1.158
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get effect sizes
# Get all pairwise effect sizes
effs1a <- eff_size(emm1a, sigma = sigma(m1), edf = df.residual(m1))
```

```
# Remove the two redundant rows (rows 3 and 4)
(effs1a_filtered <- subset(effs1a, !contrast %in% c("Large Family High Base Frequency - Small Family Low Base Frequency",
                                                    "Small Family High Base Frequency - Large Family Low Base Frequency")))
```

```
||  contrast                                                    effect.size    SE   df lower.CL upper.CL
||  Large Family High Base Frequency - Small Family High Base Frequency     -0.0527 0.189 59.9  -0.4302    0.325
||  Large Family High Base Frequency - Large Family Low Base Frequency      -0.4144 0.164 60.3  -0.7418   -0.087
||  Small Family High Base Frequency - Small Family Low Base Frequency       0.0906 0.164 59.9  -0.2366    0.418
||  Large Family Low Base Frequency - Small Family Low Base Frequency        0.4523 0.189 59.9   0.0746    0.830
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| sigma used for effect sizes: 1.397
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

For large-family words, N250 amplitude is more negative when base frequency is high than when it is low. For small-family words, base frequency has little effect. For low-frequency bases, small-family words elicit more negative amplitudes than large-family words.

- At **High Base Frequency**: `Large` vs. `Small` family → no difference ($p = .781$). Family size doesn't matter when base frequency is high.
- Within **Small Family**: `High` vs. `Low` base frequency → not significant ($p = .581$). Small-family words are unaffected by base frequency.
- At **Low Base Frequency**: Large vs. Small family → significant difference ($p = .0194$). Small-family words yield more negative amplitudes than large-family words, but only when base frequency is low.
- Within **Large Family**: High vs. Low base frequency → significant ($p = .0136$). Large-family words show more negative amplitudes when their base frequency is high.

3.3.2 `family_size` × `base_freq` Interaction Contrasts

```
#  Interaction contrasts (difference-of-differences)
#    Compare base frequency effect in large vs small family)
contrast(emm1a, interaction = "pairwise", by = NULL, adjust = "holm")
```

```
||  family_size_pairwise       base_freq_pairwise                      estimate    SE   df t.ratio p.value
||  Large Family - Small Family High Base Frequency - Low Base Frequency   -0.705 0.119 1523  -5.916  <.0001
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get confidence intervals, for each base frequency effect for each family size and then for interaction effect
confint(contrast(emmeans(m1, ~ family_size | base_freq), "pairwise"))
```

```
|| base_freq = High Base Frequency:
||  contrast                   estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family  -0.0736 0.264 65.5   -0.600    0.453
||
|| base_freq = Low Base Frequency:
||  contrast                   estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family   0.6317 0.264 65.5    0.105    1.158
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
confint(contrast(emm1a, interaction = c("pairwise", "pairwise")))
```

```
||  family_size_pairwise       base_freq_pairwise                      estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family High Base Frequency - Low Base Frequency   -0.705 0.119 1523   -0.939   -0.471
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

3.3.3 `Sensitivity` × `family_size` × `base_freq` Simple Contrasts

```
# Estimated marginal means for the family_size × base_freq interaction
(emm1b <- emmeans(anova_model_n250_words_b, ~ Semantic_Sensitivity * family_size * base_freq))
```

```
||  Semantic_Sensitivity family_size  base_freq           emmean    SE   df lower.CL upper.CL
||  High Semantic        Large Family High Base Frequency -1.309 0.392 60.4   -2.093  -0.5262
||  Low Semantic         Large Family High Base Frequency -0.508 0.398 60.4   -1.304   0.2883
||  High Semantic        Small Family High Base Frequency -0.882 0.490 59.9   -1.862   0.0985
||  Low Semantic         Small Family High Base Frequency -0.788 0.498 59.9   -1.785   0.2080
||  High Semantic        Large Family Low Base Frequency  -0.515 0.403 60.3   -1.321   0.2904
||  Low Semantic         Large Family Low Base Frequency  -0.144 0.410 60.3   -0.963   0.6747
||  High Semantic        Small Family Low Base Frequency  -1.206 0.479 59.9   -2.163  -0.2483
||  Low Semantic         Small Family Low Base Frequency  -0.717 0.487 59.9   -1.691   0.2558
```

```
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get all pairswise contrasts
emm1b_contrasts <- contrast(emm1b,  method = "pairwise", by = NULL, adjust = "none")

# Keep only the contrasts you want
# Simple effects of family_size at each level of base_freq
# Simple effects of base_freq at each level of family_size
keep1b <- c("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",
          "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",
          "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",
          "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",
          "High Semantic Large Family High Base Frequency - Low Semantic Large Family High Base Frequency",
          "High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency",
          "High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency",
          "High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency")

(emm1b_contrasts_filtered <- subset(emm1b_contrasts, contrast %in% keep1b))
```

```
||  contrast                                                                                    estimate    SE   df t.ratio p.value
||  High Semantic Large Family High Base Frequency - Low Semantic Large Family High Base Frequency   -0.8016 0.558 60.4  -1.435  0.1563
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency  -0.4278 0.370 65.5  -1.157  0.2513
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency   -0.7941 0.320 67.9  -2.479  0.0157
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency     0.2805 0.376 65.5   0.747  0.4580
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency     -0.3634 0.326 67.9  -1.116  0.2684
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency   -0.0932 0.699 59.9  -0.133  0.8943
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency    0.3241 0.320 67.9   1.012  0.3153
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency     -0.0710 0.326 67.9  -0.218  0.8282
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency     0.6904 0.370 65.5   1.868  0.0663
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency      0.2020 0.632 81.4   0.320  0.7499
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency       0.5730 0.376 65.5   1.525  0.1321
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency     -0.4883 0.682 59.9  -0.716  0.4771
||
|| Degrees-of-freedom method: kenward-roger
```

```
# Get Confidence Intervals
(emm1b_contrasts_filtered_ci <- confint(emm1b_contrasts_filtered))
```

```
||  contrast                                                                                    estimate    SE   df lower.CL upper.CL
||  High Semantic Large Family High Base Frequency - Low Semantic Large Family High Base Frequency   -0.8016 0.558 60.4  -1.9185    0.315
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency  -0.4278 0.370 65.5  -1.1659    0.310
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency   -0.7941 0.320 67.9  -1.4334   -0.155
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency     0.2805 0.376 65.5  -0.4698    1.031
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency     -0.3634 0.326 67.9  -1.0133    0.286
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency   -0.0932 0.699 59.9  -1.4909    1.304
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency    0.3241 0.320 67.9  -0.3152    0.963
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency     -0.0710 0.326 67.9  -0.7208    0.579
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency     0.6904 0.370 65.5  -0.0478    1.428
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency      0.2020 0.632 81.4  -1.0547    1.459
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency       0.5730 0.376 65.5  -0.1773    1.323
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency     -0.4883 0.682 59.9  -1.8535    0.877
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get effect sizes
# Get all pairwise effect sizes
effs1b <- eff_size(emm1b, sigma = sigma(m1), edf = df.residual(m1))

# Remove the  redundant rows
(effs1b_filtered <- subset(effs1b, contrast %in% keep1b))
```

```
||  contrast                                                                                    effect.size    SE   df lower.CL upper.CL
||  High Semantic Large Family High Base Frequency - Low Semantic Large Family High Base Frequency      -0.5740 0.400 60.4  -1.3739    0.226
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency     -0.3063 0.265 59.9  -0.8359    0.223
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency      -0.5686 0.230 60.3  -1.0278   -0.109
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency        0.2009 0.269 59.9  -0.3374    0.739
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency       -0.2602 0.233 60.3  -0.7267    0.206
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency     -0.0668 0.500 59.9  -1.0676    0.934
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency      0.2321 0.229 59.9  -0.2269    0.691
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency       -0.0508 0.233 59.9  -0.5173    0.416
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency       0.4943 0.265 59.9  -0.0354    1.024
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency        0.1447 0.452 59.9  -0.7601    1.049
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency         0.4103 0.269 59.9  -0.1281    0.949
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency       -0.3497 0.489 59.9  -1.3273    0.628
||
```

```
|| sigma used for effect sizes: 1.397
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

3.3.4 Sensitivity × family_size × base_freq Interaction Contrasts

The interaction contrast tests whether the difference in the *base frequency* effect for large vs small families differs across semantic sensitivity?

$$[[(A_1 - A_2) \text{ in } B_1] - [(A_1 - A_2) \text{ in } B_2] \text{ in Condition } C_1] - [[(A_1 - A_2) in B_1] - [(A_1 - A_2) in B_2] \text{in Condition } C_2]$$

```
#  Interaction contrasts (difference-of-differences)
#    Compare base_freq effect in large vs small family)
contrast(emm1b, interaction = "pairwise", by = NULL, adjust = "holm")
```

```
||  Semantic_Sensitivity_pairwise family_size_pairwise       base_freq_pairwise                      estimate    SE   df t.ratio p.value
||  High Semantic - Low Semantic  Large Family - Small Family High Base Frequency - Low Base Frequency   -0.826 0.238 1523  -3.463  0.0005
||
|| Degrees-of-freedom method: kenward-roger
```

```
confint(contrast(emm1b, interaction = c("pairwise", "pairwise")))
```

```
||  Semantic_Sensitivity_pairwise family_size_pairwise       base_freq_pairwise                      estimate    SE   df lower.CL upper.CL
||  High Semantic - Low Semantic  Large Family - Small Family High Base Frequency - Low Base Frequency   -0.826 0.238 1523    -1.29   -0.358
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Compute the A1 - A2 difference within each combination of B × C
(base_freq_diff <- contrast(emm1b, method = "revpairwise",
                            by = c("Semantic_sensitivity", "family_size"),
                            simple = "base_freq"))
```

```
|| Semantic_Sensitivity = High Semantic, family_size = Large Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency    0.794 0.320 67.9   2.479  0.0157
||
|| Semantic_Sensitivity = Low Semantic, family_size = Large Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency    0.363 0.326 67.9   1.116  0.2684
||
|| Semantic_Sensitivity = High Semantic, family_size = Small Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency   -0.324 0.320 67.9  -1.012  0.3153
||
|| Semantic_Sensitivity = Low Semantic, family_size = Small Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency    0.071 0.326 67.9   0.218  0.8282
||
|| Degrees-of-freedom method: kenward-roger
```

```
# Compute how that A-effect changes across the levels of B, separately for each level of C
(family_size_base_freq_int_within_sensitivity <- contrast(base_freq_diff,
                                                          method = "revpairwise",
                                                          by = "Semantic_sensitivity", simple = "family_size"))
```

```
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = High Semantic:
||  contrast1                   estimate    SE   df t.ratio p.value
||  Small Family - Large Family   -1.118 0.167 1523  -6.687  <.0001
||
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = Low Semantic:
||  contrast1                   estimate    SE   df t.ratio p.value
||  Small Family - Large Family   -0.292 0.170 1523  -1.721  0.0855
||
|| Degrees-of-freedom method: kenward-roger
```

```
# Get confidence intervals
confint(family_size_base_freq_int_within_sensitivity)
```

```
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = High Semantic:
||  contrast1                   estimate    SE   df lower.CL upper.CL
||  Small Family - Large Family   -1.118 0.167 1523   -1.446  -0.7902
||
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = Low Semantic:
||  contrast1                   estimate    SE   df lower.CL upper.CL
||  Small Family - Large Family   -0.292 0.170 1523   -0.626   0.0409
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

Compute the effect of Base Frequency (Low - High) within each *Semantic Sensitivity* × *Family Size* combination.

High Sensitivity- Large Family: `Low Base Frequency - High Base Frequency = -0.515 - (-1.309)  = 0.794`

High Sensitivity- Small Family: `Low Base Frequency - High Base Frequency = -1.206 - (-0.882) = -0.324`

Low Sensitivity - Large Family: `Low Base Frequency - High Base Frequency = -0.144 - (-0.508) =  0.364`

Low Sensitivity - Small Family: `Low Base Frequency - High Base Frequency = -0.717 - (-0.788) =  0.071`

Compute the difference of differences: compare how the effect of base frequency differs across sensitivity groups:`(High Sensitivity base frequency effect) - (Low Sensitivity base frequency effect)`

For Large Family:

```
High:  0.794
Low:   0.364
Difference: 0.794 - 0.364 = 0.43
```

For Small Family:

```
High: -0.324
Low:  +0.071
Difference: -0.324 - 0.071 = -0.395
```

This is a reversal of the *base frequency effect* between High and Low sensitivity participants for Small Family nonwords — and that's the core of your significant 3-way interaction.

Now take the difference of these differences (Small - Large): `-0.395 - 0.43 = -0.825`. That's the interaction contrast estimate: '-0.826, $SE = 0.238$, $t(1523) = -3.463$, $p = 0.0005$

Only High-Sensitivity participants responding to words from Large morphological familes show a clear base-frequency effect. When responding to low-base frequency words they show a clear family size effect.

In high-semantic-sensitivity readers, the N250 shows a selective interaction: Low vs. high base frequency matters only for large-family words. This dependency disappears in low-sensitivity readers.

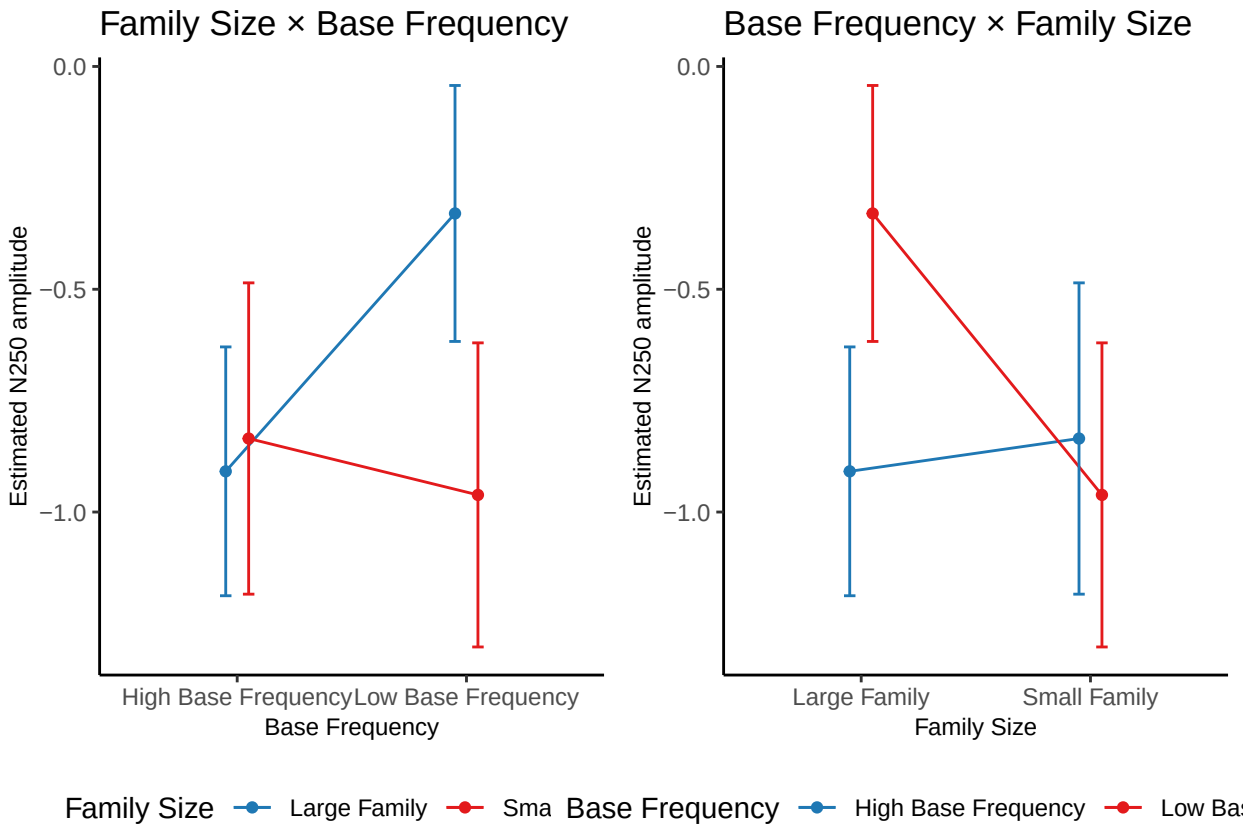
3.4 Plots

3.4.1 family_size × base_freq

```r
emm1a_df <- as.data.frame(emm1a)
p1 <- ggplot(emm1a_df,
       aes(x = base_freq, y = emmean,
           color = family_size, group = family_size)) +
  geom_line(position = position_dodge(0.2)) +
  geom_point(position = position_dodge(0.2)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = 0.1, position = position_dodge(0.2)) +
  labs(x = "Base Frequency", y = "Estimated N250 amplitude",
       color = "Family Size",
       title = "Family Size × Base Frequency") +
  scale_color_custom() +
  scale_fill_custom()


p2 <- ggplot(emm1a_df,
       aes(x = family_size, y = emmean,
           color = base_freq, group = base_freq)) +
  geom_line(position = position_dodge(0.2)) +
  geom_point(position = position_dodge(0.2)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = 0.1, position = position_dodge(0.2)) +
  labs(x = "Family Size", y = "Estimated N250 amplitude",
       color = "Base Frequency",
       title = "Base Frequency × Family Size") +
  scale_color_custom() +
  scale_fill_custom()

plot_grid(p1, p2, ncol = 2)
```



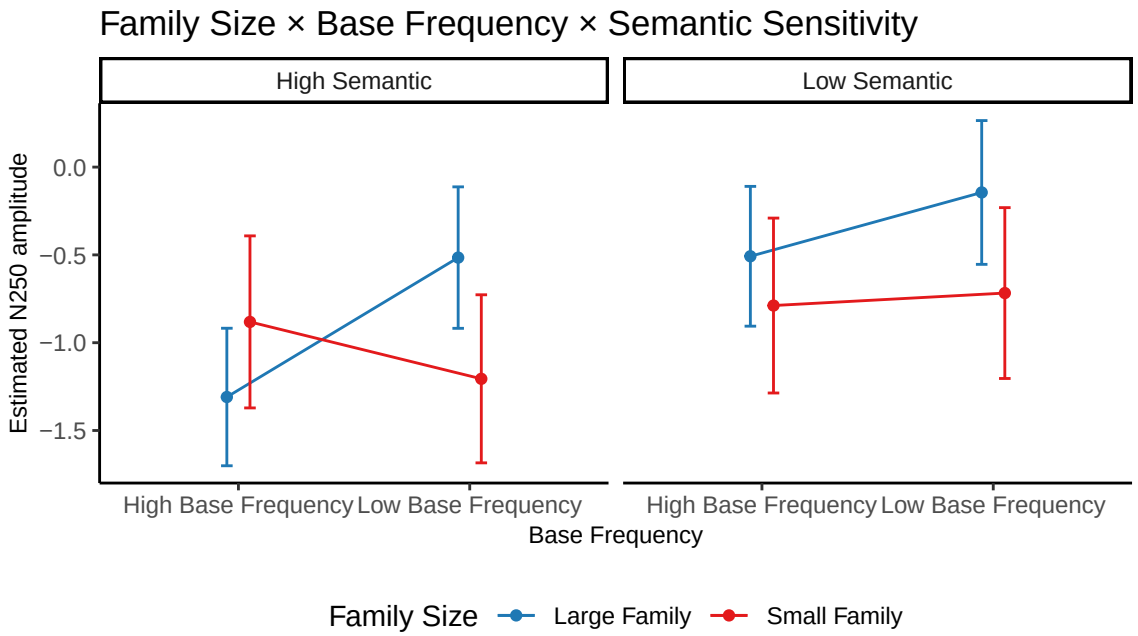
3.4.2 Sensitivity × family_size × base_freq

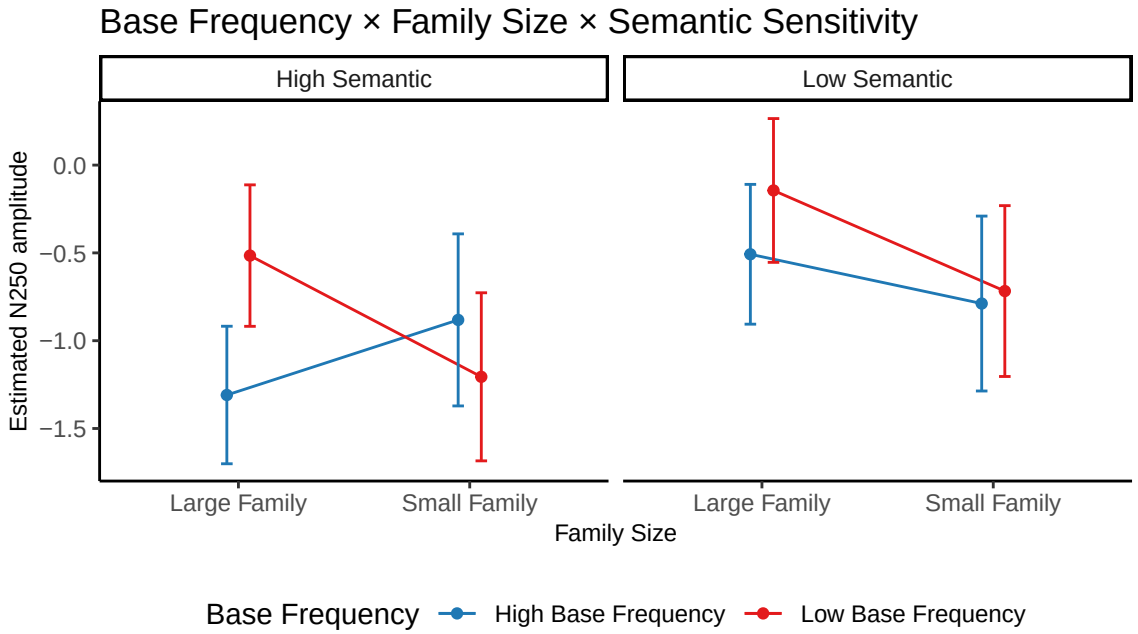
```
# Plot the interaction
library(ggplot2)

emm1b_df <- as.data.frame(emm1b)
p3 <- ggplot(emm1b_df,
       aes(x = base_freq, y = emmean,
           color = family_size, group = family_size)) +
  facet_wrap(~ Semantic_Sensitivity) +
  geom_line(position = position_dodge(0.2)) +
  geom_point(position = position_dodge(0.2)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = 0.1, position = position_dodge(0.2)) +
  labs(x = "Base Frequency", y = "Estimated N250 amplitude",
       color = "Family Size",
       title =  "Family Size × Base Frequency × Semantic Sensitivity") +
  scale_color_custom() +
  scale_fill_custom()


p4 <- ggplot(emm1b_df,
       aes(x = family_size, y = emmean,
           color = base_freq, group = base_freq)) +
  facet_wrap(~ Semantic_Sensitivity) +
  geom_line(position = position_dodge(0.2)) +
  geom_point(position = position_dodge(0.2)) +
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = 0.1, position = position_dodge(0.2)) +
  labs(x = "Family Size", y = "Estimated N250 amplitude",
       color = "Base Frequency",
       title = "Base Frequency × Family Size × Semantic Sensitivity") +
  scale_color_custom() +
  scale_fill_custom()

plot_grid(p3, p4, nrow = 2)
```


Family Size × Base Frequency × Semantic Sensitivity
High Semantic
Low Semantic
Estimated N250 amplitude
0.0
−0.5
−1.0
−1.5
High Base Frequency
Low Base Frequency
Base Frequency
Family Size
Large Family
Small Family


Base Frequency × Family Size × Semantic Sensitivity
High Semantic
Low Semantic
Estimated N250 amplitude
0.0
−0.5
−1.0
−1.5
Large Family
Small Family
Family Size
Base Frequency
High Base Frequency
Low Base Frequency

4 N250 Nonword Data

4.1 Compute the ANOVA

```
anova_model_n250_nonwords <- mixed(
    value ~ Semantic_Sensitivity * family_size * complexity +
    (1 + family_size + complexity | SubjID) +     # by-subject intercept + slopes
    (1 | SubjID:chlabel),                         # electrode nested within subject
  data   = n250_nonwords,
  method = "KR"
)
anova_model_n250_nonwords
```

```
|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Semantic_Sensitivity * family_size * complexity + (1 +
|| Model:     family_size + complexity | SubjID) + (1 | SubjID:chlabel)
|| Data: n250_nonwords
||                                       Effect      df      F p.value
|| 1                         Semantic_Sensitivity   1, 59   0.13    .717
|| 2                                  family_size   1, 59   0.11    .742
|| 3                                   complexity   1, 59   0.00    .951
|| 4             Semantic_Sensitivity:family_size   1, 59   0.38    .538
|| 5              Semantic_Sensitivity:complexity   1, 59   1.34    .252
|| 6                        family_size:complexity 1, 1523 2.75 +    .097
|| 7 Semantic_Sensitivity:family_size:complexity 1, 1523   1.22    .269
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
m2 <- anova_model_n250_nonwords$full_model    # Extract the lmer model
ranova(m2)    # Run random effects comparison
```

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| value ~ Semantic_Sensitivity + family_size + complexity + (1 + family_size + complexity | SubjID) + (1 | SubjID:chlabel) + Semantic_Sensitivity
||                                                    npar  logLik    AIC    LRT Df Pr(>Chisq)
|| <none>                                               16 -4507.9 9047.8
|| family_size in (1 + family_size + complexity | SubjID)   13 -4721.4 9468.8 427.00  3  < 2.2e-16 ***
|| complexity in (1 + family_size + complexity | SubjID)    13 -4849.7 9725.5 683.67  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                     15 -4708.4 9446.8 400.96  1  < 2.2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Extract effect sizes from your ANOVA model
eta_squared(anova_model_n250_nonwords, partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------------
|| Semantic_Sensitivity                     |       2.24e-03 | [0.00, 1.00]
|| family_size                              |       1.84e-03 | [0.00, 1.00]
|| complexity                               |       6.33e-05 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size         |       6.46e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity          |          0.02 | [0.00, 1.00]
|| family_size:complexity                   |       1.81e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |       8.01e-04 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```
# Compute Marginal(fixed effects only) and Conditional(fixed + random effects) R²
r2(anova_model_n250_nonwords)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.758
||      Marginal R2: 0.005
```

4.2 Main Effects and Interactions

The random structure is well supported; most variance lies in subject- and electrode-specific fluctuations. None of the effects reach conventional significance. Nearly all systematic variance is due to subject/electrode differences, not the fixed experimental factors. There is no evidence that Semantic Sensitivity modulates N250 responses to non-words. Semantically sensitive and insensitive participants behave alike for non-word stimuli. Family size and morphological complexity do not affect the N250 when there is no real morphological or semantic content to activate. There is at most a weak family_size × complexity trend ($p = .097$)); perhaps more "complex" pseudowords elicit slightly different early orthographic responses, but not reliably.