M21 LDT ERP HC SEMANTIC SENSITIVITY N400

Joanna Morris

2025-10-16

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 N400 Word Data

3.1 Nested ANOVA Model

```
#Fit ANOVA model
anova_model_n400_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   = n400_words_b,
  method = "KR"
)
anova_model_n400_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: n400_words_b
||                                     Effect      df         F p.value
|| 1                       Semantic_Sensitivity   1, 59      0.01    .943
|| 2                                family_size   1, 59      1.84    .180
|| 3                                  base_freq   1, 59      1.00    .321
|| 4           Semantic_Sensitivity:family_size   1, 59      0.03    .863
|| 5             Semantic_Sensitivity:base_freq   1, 59      0.33    .567
|| 6                      family_size:base_freq 1, 1523 64.46 ***   <.001
|| 7 Semantic_Sensitivity:family_size:base_freq 1, 1523   8.73 **    .003
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
m1 <- anova_model_n400_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 -4773.9  9579.8
|| family_size in (1 + family_size + base_freq | SubjID)   13 -5065.1 10156.3 582.48  3  < 2.2e-16 ***
|| base_freq in (1 + family_size + base_freq | SubjID)     13 -4908.0  9841.9 268.14  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                    15 -5216.2 10462.5 884.68  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_n400_words_b, partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------------
|| Semantic_Sensitivity                     |       8.82e-05 | [0.00, 1.00]
|| family_size                              |           0.03 | [0.00, 1.00]
|| base_freq                                |           0.02 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size         |       5.10e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:base_freq           |       5.57e-03 | [0.00, 1.00]
|| family_size:base_freq                    |           0.04 | [0.03, 1.00]
|| Semantic_Sensitivity:family_size:base_freq |       5.70e-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_n400_words_b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.846
||      Marginal R2: 0.008
```

3.2 Main Effects

The N400 data for words show no main effects,

3.3 Interactions

There is a robust *family_size × base_freq* interaction, which is further modulated by *Semantic Sensitivity*. This is essentially the same structural pattern as for the N250, but the effect is larger and more reliable in the N400 window.

Effect	df	F	p.value	
family_size:base_freq	1, 1523	64.46 ***	<.001	0.04
Semantic_Sensitivity:family_size:base_freq	1, 1523	8.73 ***	.003	5.70e-03

The family_size × base_freq effect captures a general lexical relationship that applies across participants: e.g., frequency effects differ for large vs. small morphological families. The three-way interaction simply means that this pattern is somewhat different between the sensitivity groups, but the difference between groups is modest.

3.3.1 family_size × base_freq Simple Contrasts

```
# Estimated marginal means for the family_size × base frequency interaction
(emm1a <- emmeans(anova_model_n400_words_b, ~ family_size * base_freq))
```

```
||  family_size  base_freq           emmean    SE   df lower.CL upper.CL
||  Large Family High Base Frequency  0.680 0.366 59.9  -0.0516     1.41
||  Small Family High Base Frequency  0.833 0.432 59.6  -0.0314     1.70
||  Large Family Low Base Frequency   1.004 0.348 60.0   0.3085     1.70
||  Small Family Low Base Frequency   0.137 0.406 59.7  -0.6759     0.95
||
|| 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   -0.153 0.271 66.1  -0.564  0.5744
||  Large Family High Base Frequency - Large Family Low Base Frequency    -0.324 0.197 73.5  -1.647  0.1039
||  Small Family High Base Frequency - Small Family Low Base Frequency     0.696 0.197 73.5   3.542  0.0007
||  Large Family Low Base Frequency - Small Family Low Base Frequency      0.867 0.271 66.1   3.205  0.0021
||
|| 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   -0.153 0.271 66.1   -0.693    0.388
||  Large Family High Base Frequency - Large Family Low Base Frequency    -0.324 0.197 73.5   -0.715    0.068
||  Small Family High Base Frequency - Small Family Low Base Frequency     0.696 0.197 73.5    0.305    1.088
||  Large Family Low Base Frequency - Small Family Low Base Frequency      0.867 0.271 66.1    0.327    1.407
||
|| 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      -0.103 0.182 59.6   -0.466   0.2612
||  Large Family High Base Frequency - Large Family Low Base Frequency       -0.218 0.132 59.9   -0.482   0.0468
||  Small Family High Base Frequency - Small Family Low Base Frequency        0.468 0.132 59.6    0.203   0.7325
||  Large Family Low Base Frequency - Small Family Low Base Frequency         0.583 0.182 59.7    0.219   0.9470
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| sigma used for effect sizes: 1.488
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

For large-family words, n400 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 - Small $= -0.153; SE = 0.271; t = -0.564; p = 0.5744$. This difference is not statistically significant ($p > .05$).
- At Low base frequency: Large - Small $= 0.867$ $SE = 0.271; t = 3.205; p = 0.0021$. This difference is statistically significant ($p = 0.0023$ after adjustment.

When base frequency is *low*, large vs small family_size differ significantly in predicted N400; when base frequency is *high*, they do not differ.

Next, contrasting High vs Low base_freq within each family_size:

- Large family_size: High - Low $= -0.324; SE = 0.197; t = -1.647; p = 0.1039$. This is not significant.
- Small family_size: High - Low $= 0.696; SE = 0.197; t = 3.542; p = 0.0007$. Significant difference: base_freq level matters when family_size is small.

When family_size is *small*, high vs low base frequency yields a significant difference; when family_size is *large*, the difference is not strong.

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    -1.02 0.127 1523  -8.029  <.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   -0.153 0.271 66.1   -0.693    0.388
||
|| base_freq = Low Base Frequency:
||  contrast                   estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family    0.867 0.271 66.1    0.327    1.407
||
|| 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    -1.02 0.127 1523    -1.27   -0.771
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

The final contrast tests whether the difference between Large vs Small family_size is itself different between High vs Low base_freq:

Estimate $= -1.02$; $SE = 0.127$; $t = -8.029$; $p < .0001$

That is, the slope or effect of `family_size` depends strongly on the level of `base_freq` (consistent with your ANOVA). Put differently: the family size difference (Large - Small) is much more positive in the low base frequency condition than it is in the high base frequency condition. That difference of differences is highly significant.

3.3.3 Sensitivity × family_size × base_freq Simple Contrasts

```
# Estimated marginal means for the family_size × base_freq interaction
(emm1b <- emmeans(anova_model_n400_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  0.6374 0.513 59.9  -0.3893     1.66
||  Low Semantic         Large Family High Base Frequency  0.7234 0.522 59.9  -0.3202     1.77
||  High Semantic        Small Family High Base Frequency  0.9321 0.606 59.6  -0.2805     2.14
||  Low Semantic         Small Family High Base Frequency  0.7341 0.616 59.6  -0.4985     1.97
||  High Semantic        Large Family Low Base Frequency   1.0418 0.488 60.0   0.0662     2.02
||  Low Semantic         Large Family Low Base Frequency   0.9664 0.496 60.0  -0.0253     1.96
||  High Semantic        Small Family Low Base Frequency  -0.0588 0.570 59.7  -1.1987     1.08
||  Low Semantic         Small Family Low Base Frequency   0.3325 0.579 59.7  -0.8263     1.49
||
|| 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   -0.0860 0.732 59.9  -0.118  0.9068
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency  -0.2947 0.380 66.1  -0.777  0.4402
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency   -0.4044 0.276 73.5  -1.467  0.1467
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency    -0.0107 0.386 66.1  -0.028  0.9780
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency     -0.2430 0.280 73.5  -0.867  0.3887
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency    0.1980 0.864 59.6   0.229  0.8195
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency    0.9909 0.276 73.5   3.594  0.0006
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency      0.4016 0.280 73.5   1.433  0.1561
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency     1.1006 0.380 66.1   2.900  0.0051
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency      0.7093 0.757 74.9   0.937  0.3519
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency       0.6339 0.386 66.1   1.643  0.1051
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency     -0.3913 0.813 59.7  -0.482  0.6319
||
|| 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   -0.0860 0.732 59.9   -1.550    1.378
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency  -0.2947 0.380 66.1   -1.052    0.463
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency   -0.4044 0.276 73.5   -0.954    0.145
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency    -0.0107 0.386 66.1   -0.781    0.760
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency     -0.2430 0.280 73.5   -0.801    0.315
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency    0.1980 0.864 59.6   -1.531    1.927
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency    0.9909 0.276 73.5    0.442    1.540
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency      0.4016 0.280 73.5   -0.157    0.960
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency     1.1006 0.380 66.1    0.343    1.858
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency      0.7093 0.757 74.9   -0.799    2.218
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency       0.6339 0.386 66.1   -0.136    1.404
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency     -0.3913 0.813 59.7   -2.017    1.234
||
|| 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     -0.05781 0.492 59.9  -1.0416   0.9260
||  High Semantic Large Family High Base Frequency - High Semantic Small Family High Base Frequency    -0.19806 0.255 59.6  -0.7083   0.3122
||  High Semantic Large Family High Base Frequency - High Semantic Large Family Low Base Frequency     -0.27178 0.185 59.9  -0.6425   0.0989
||  Low Semantic Large Family High Base Frequency - Low Semantic Small Family High Base Frequency      -0.00717 0.259 59.6  -0.5258   0.5115
||  Low Semantic Large Family High Base Frequency - Low Semantic Large Family Low Base Frequency       -0.16329 0.188 59.9  -0.5401   0.2135
||  High Semantic Small Family High Base Frequency - Low Semantic Small Family High Base Frequency      0.13309 0.581 59.6  -1.0289   1.2950
||  High Semantic Small Family High Base Frequency - High Semantic Small Family Low Base Frequency      0.66589 0.185 59.6   0.2948   1.0370
||  Low Semantic Small Family High Base Frequency - Low Semantic Small Family Low Base Frequency        0.26988 0.188 59.6  -0.1070   0.6467
||  High Semantic Large Family Low Base Frequency - High Semantic Small Family Low Base Frequency       0.73960 0.255 59.7   0.2290   1.2502
||  High Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency        0.47668 0.509 59.7  -0.5414   1.4948
||  Low Semantic Large Family Low Base Frequency - Low Semantic Small Family Low Base Frequency         0.42600 0.259 59.7  -0.0928   0.9448
||  High Semantic Small Family Low Base Frequency - Low Semantic Small Family Low Base Frequency      -0.26292 0.546 59.7  -1.3553   0.8294
||
|| sigma used for effect sizes: 1.488
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

(a) Base-frequency effect within each Family Size × Sensitivity cell

Semantic Sensitivity	Family Size	Low – High Base Freq	t	p	Interpretation
High Sensitivity	Large Family	+0.40 μV	1.47	.15 (n.s.)	Weak, non-sig. tendency: low-freq words slightly less negative.
Low Sensitivity	Large Family	+0.24 μV	0.87	.39 (n.s.)	Essentially flat.
High Sensitivity	Small Family	–0.99 μV	–3.59	**.0006**	More negative N400 for high-freq small-family words.
Low Sensitivity	Small Family	–0.40 μV	–1.43	.16 (n.s.)	Modest, non-sig. version of same trend.

Pattern: Only high-sensitivity readers show a pronounced base-frequency effect—and only for small-family words, where high-frequency bases elicit larger (more negative) N400s.

(b) Difference of those frequency effects across family size (within each group)

Semantic Sensitivity	(Small – Large Family) difference in base-freq effect	95% CI	p	Interpretation
High Sensitivity	–1.40 μV	[–1.74, –1.05]	< .001	the base-freq effect flips between large- and small-family words.
Low Sensitivity	–0.65 μV	[–1.00, –0.29]	.0004	Same pattern but weaker: a smaller differential between family sizes.

Semantic Sensitivity (Small – Large Family) difference in base-freq effect 95 % CI p Interpretation High Sensitivity –1.40 μV [–1.74, –1.05] < .001 Very large difference: the base-freq effect flips between large- and small-family words.
Low Sensitivity –0.65 μV [–1.00, –0.29] .0004 Same pattern but weaker: a smaller differential between family sizes.

Interpretation: Both groups show that the base-frequency effect differs by family size, but this contrast is about twice as strong in the high-sensitivity group.

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   -0.751 0.254 1523  -2.955  0.0032
||
|| 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   -0.751 0.254 1523    -1.25   -0.252
||
|| 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    0.404 0.276 73.5   1.467  0.1467
||
|| Semantic_Sensitivity = Low Semantic, family_size = Large Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency    0.243 0.280 73.5   0.867  0.3887
||
|| Semantic_Sensitivity = High Semantic, family_size = Small Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency   -0.991 0.276 73.5  -3.594  0.0006
||
|| Semantic_Sensitivity = Low Semantic, family_size = Small Family:
||  contrast                                estimate    SE   df t.ratio p.value
||  Low Base Frequency - High Base Frequency   -0.402 0.280 73.5  -1.433  0.1561
||
|| 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   -1.395 0.178 1523  -7.831  <.0001
||
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = Low Semantic:
||  contrast1                  estimate    SE   df t.ratio p.value
||  Small Family - Large Family   -0.645 0.181 1523  -3.559  0.0004
||
|| 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   -1.395 0.178 1523    -1.74   -1.046
||
|| contrast = Low Base Frequency - High Base Frequency, Semantic_Sensitivity = Low Semantic:
||  contrast1                  estimate    SE   df lower.CL upper.CL
||  Small Family - Large Family   -0.645 0.181 1523    -1.00   -0.289
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

(High – Low Sensitivity) × (Large – Small Family) × (High – Low Base Freq) = –0.75 μV [–1.25, –0.25], $p = .003$.

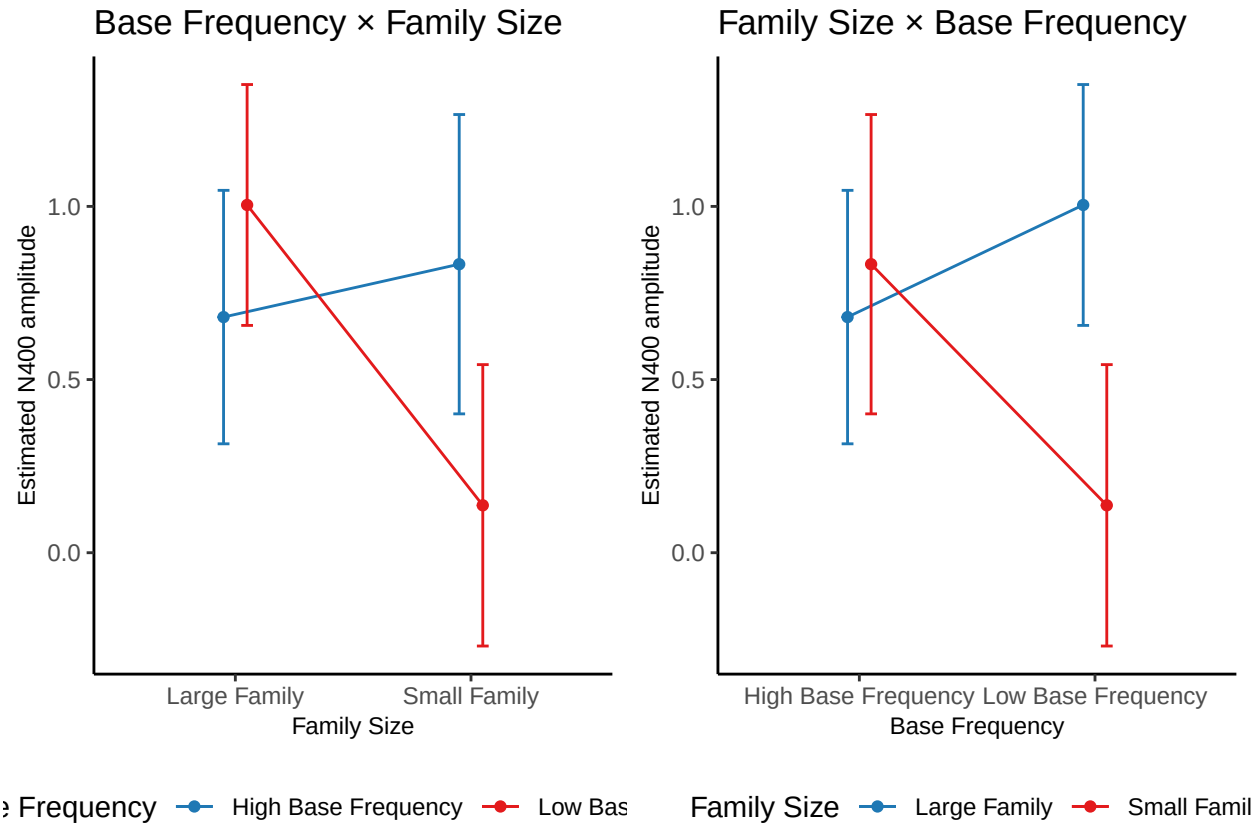
Meaning: The difference between family-size patterns across base-frequency conditions is 0.75 μV larger in the high-sensitivity group than in the low-sensitivity group—confirming the N400 interaction found in the omnibus ANOVA.

3.4 Plots

3.4.1 family_size × base_freq

```
emm1a_df <- as.data.frame(emm1a)
p1<- 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 N400 amplitude",
       color = "Base Frequency",
       title = "Base Frequency × Family Size") +
  scale_color_custom() +
  scale_fill_custom()

p2 <- 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)) +
```

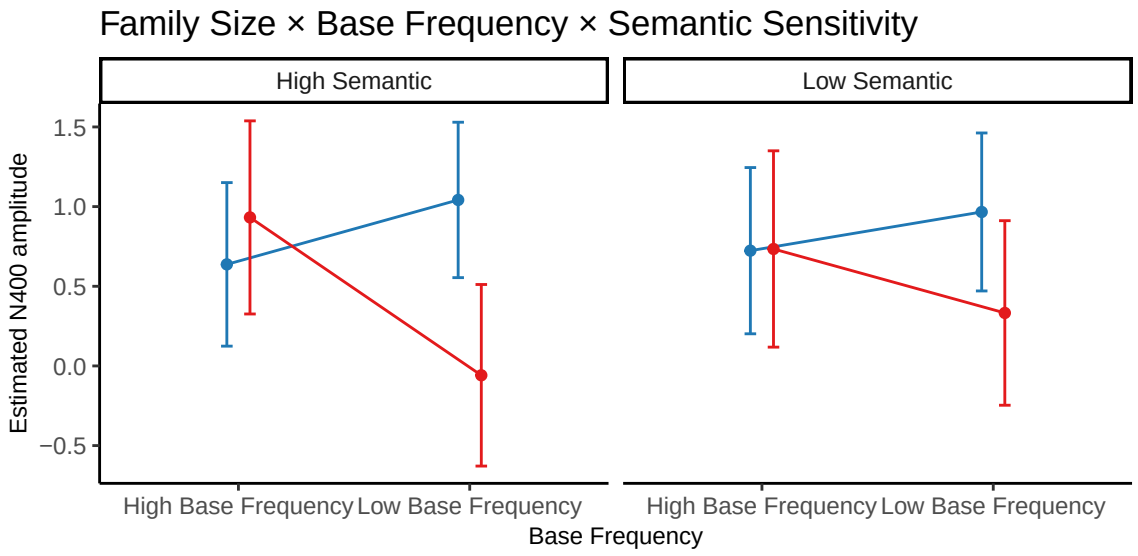
```
  geom_errorbar(aes(ymin = emmean - SE, ymax = emmean + SE),
                width = 0.1, position = position_dodge(0.2)) +
  labs(x = "Base Frequency", y = "Estimated N400 amplitude",
       color = "Family Size",
       title = "Family Size × Base Frequency") +
  scale_color_custom() +
  scale_fill_custom()

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

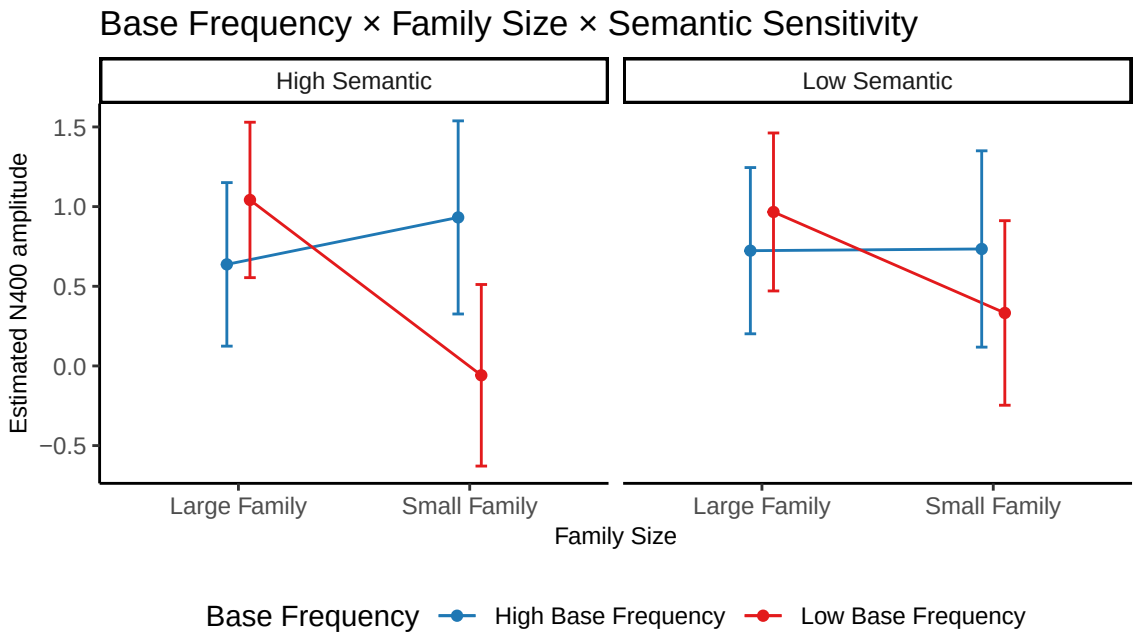


3.4.2 Sensitivity × family_size × base_freq

```
# 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 N400 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 N400 amplitude",
```

```
      color = "Base Frequency",
      title = "Base Frequency × Family Size × Semantic Sensitivity") +
  scale_color_custom() +
  scale_fill_custom()

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





4 N400 Nonword Data

4.1 Compute the ANOVA

```
anova_model_n400_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   = n400_nonwords,
  method = "KR"
)
anova_model_n400_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: n400_nonwords
||                                       Effect      df      F p.value
|| 1                         Semantic_Sensitivity   1, 59   0.00    .998
|| 2                                  family_size   1, 59   0.02    .900
|| 3                                   complexity   1, 59   0.01    .935
|| 4            Semantic_Sensitivity:family_size   1, 59   0.06    .806
|| 5             Semantic_Sensitivity:complexity   1, 59   0.42    .520
|| 6                       family_size:complexity 1, 1523 5.34 *    .021
|| 7 Semantic_Sensitivity:family_size:complexity 1, 1523   0.23    .628
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```
m2 <- anova_model_n400_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 -5056.5 10145
|| family_size in (1 + family_size + complexity | SubjID)   13 -5322.9 10672  532.84  3  < 2.2e-16 ***
|| complexity in (1 + family_size + complexity | SubjID)    13 -5420.6 10867  728.17  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                     15 -5562.0 11154 1011.10  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_n400_nonwords, partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                 | Eta2 (partial) |       95% CI
|| ------------------------------------------------------------------------------
|| Semantic_Sensitivity                      |       6.06e-08 | [0.00, 1.00]
|| family_size                               |       2.68e-04 | [0.00, 1.00]
|| complexity                                |       1.12e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size          |       1.03e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity           |       7.04e-03 | [0.00, 1.00]
|| family_size:complexity                    |       3.49e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |       1.54e-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_n400_nonwords)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.852
||      Marginal R2: 0.001
```

The model's random structure is strongly supported: variability across subjects and electrodes dominates. Nearly all explainable variance arises from subject and electrode differences, not from the fixed predictors.

4.2 Main Effects

No evidence that semantic sensitivity modulates N400 responses to nonwords.No main effects of morphological "family size" or "complexity."

4.3 Interactions

A small *family_size* × *complexity* interaction may reflect weak structural sensitivity, but this effect is marginal and not easily interpretable. The absence of group effects suggests that N400 activity to pseudowords reflects form-level processing only, not meaningful lexical–semantic integration.

4.3.1 Simple Contrasts

(a) Effect of family_size within each level of complexity. Tests whether "*large vs. small family*" differs for simple and complex items separately. This helps you see where the interaction is coming from — e.g., if the family size effect flips between complexity levels.

(b) Effect of complexity within each level of family_size. Tests whether "*complex vs. simple*" differs within large and small families.

```
# Estimated marginal means for the family_size × complexity interaction
(emm2 <- emmeans(anova_model_n400_nonwords, ~ family_size * complexity))
```

```
||  family_size  complexity  emmean    SE   df lower.CL upper.CL
||  Large Family Complex    -0.2846 0.416 59.9   -1.116    0.547
||  Small Family Complex    -0.0867 0.398 59.9   -0.883    0.710
||  Large Family Simple     -0.1487 0.469 59.7   -1.087    0.789
||  Small Family Simple     -0.2762 0.454 59.7   -1.184    0.632
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get all pairswise contrasts
emm2_contrasts <- contrast(emm2,  method = "pairwise", by = NULL, adjust = "none")

# Keep only the contrasts you want
# Simple effects of family_size at each level of complexity
# Simple effects of complexity at each level of family_size
keep2 <- c("Large Family Complex - Small Family Complex",
          "Large Family Simple - Small Family Simple",
          "Large Family Complex - Large Family Simple",
          "Small Family Complex - Small Family Simple")
(emm2_contrasts_filtered <- subset(emm2_contrasts, contrast %in% keep2))
```

```
||  contrast                                 estimate    SE   df t.ratio p.value
||  Large Family Complex - Small Family Complex   -0.198 0.289 66.7  -0.686  0.4953
||  Large Family Complex - Large Family Simple    -0.136 0.336 64.5  -0.405  0.6872
||  Small Family Complex - Small Family Simple     0.189 0.336 64.5   0.564  0.5750
||  Large Family Simple - Small Family Simple      0.127 0.289 66.7   0.442  0.6601
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get Confidence Intervals
(emm2_contrasts_filtered_ci <- confint(emm2_contrasts_filtered))
```

```
||  contrast                                 estimate    SE   df lower.CL upper.CL
||  Large Family Complex - Small Family Complex   -0.198 0.289 66.7   -0.774    0.378
||  Large Family Complex - Large Family Simple    -0.136 0.336 64.5   -0.807    0.535
||  Small Family Complex - Small Family Simple     0.189 0.336 64.5   -0.482    0.861
||  Large Family Simple - Small Family Simple      0.127 0.289 66.7   -0.449    0.704
||
|| 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
effs2 <- eff_size(emm2, sigma = sigma(m2), edf = df.residual(m2))

# Remove the two redundant rows (rows 3 and 4)
(effs2_filtered <- subset(effs2, contrast %in% keep2))
```

```
||  contrast                                 effect.size    SE   df lower.CL upper.CL
||  Large Family Complex - Small Family Complex     -0.1200 0.175 59.9   -0.470    0.230
||  Large Family Complex - Large Family Simple      -0.0824 0.204 59.7   -0.490    0.325
||  Small Family Complex - Small Family Simple       0.1148 0.204 59.7   -0.293    0.522
||  Large Family Simple - Small Family Simple        0.0773 0.175 59.7   -0.273    0.427
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| sigma used for effect sizes: 1.65
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

4.3.2 Interaction Contrasts

If simple effects aren't significant, try looking at interaction contrasts, which test differences in the differences. You're now asking: Does the effect of Sensitivity change more in some complexity/family combinations than others?

The interaction contrast tests:

Is the difference in the effect of A across levels of B different at Complex vs. Simple levels?

Mathematically:

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

```
#  Interaction contrasts (difference-of-differences)
#    Compare complexity effect in large vs small family)
contrast(emm2, interaction = "pairwise", by = NULL, adjust = "holm")
```

```
||  family_size_pairwise        complexity_pairwise estimate    SE   df t.ratio p.value
||  Large Family - Small Family Complex - Simple      -0.325 0.141 1523  -2.310  0.0210
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get confidence intervals, for each complexity effect for each family size and then for interaction effect
confint(contrast(emmeans(m2, ~ family_size | complexity), "pairwise"))
```

```
|| complexity = Complex:
||  contrast                    estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family   -0.198 0.289 66.7   -0.774    0.378
||
|| complexity = Simple:
||  contrast                    estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family    0.127 0.289 66.7   -0.449    0.704
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

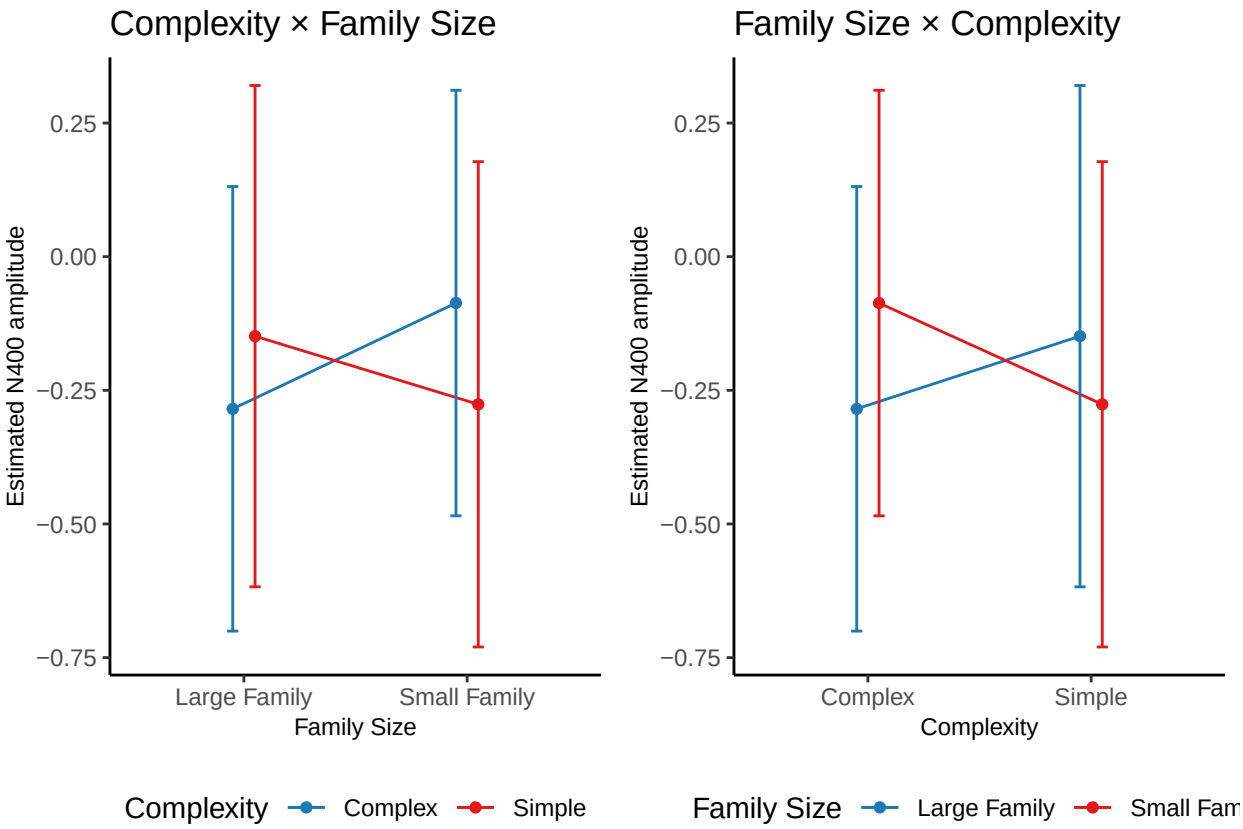
```
confint(contrast(emm2, interaction = c("pairwise", "pairwise")))
```

```
||  family_size_pairwise        complexity_pairwise estimate    SE   df lower.CL upper.CL
||  Large Family - Small Family Complex - Simple      -0.325 0.141 1523   -0.602  -0.0491
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

4.4 Plots

4.4.1 family_size × base_freq

```
# Plot the interaction
library(ggplot2)

emm2_df <- as.data.frame(emm2)

p3<- ggplot(emm2_df,
       aes(x = family_size, y = emmean,
           color = complexity, group = complexity)) +
  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 N400 amplitude",
       color = "Complexity",
       title = "Complexity × Family Size") +
  scale_color_custom() +
  scale_fill_custom()

p4 <- ggplot(emm2_df,
       aes(x = complexity, 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 = "Complexity", y = "Estimated N400 amplitude",
       color = "Family Size",
       title = "Family Size × Complexity") +
  scale_color_custom() +
  scale_fill_custom()

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


Complexity × Family Size
Estimated N400 amplitude
0.25
0.00
−0.25
−0.50
−0.75
Large Family
Small Family
Family Size
Complexity
Complex
Simple
Family Size × Complexity
Estimated N400 amplitude
0.25
0.00
−0.25
−0.50
−0.75
Complex
Simple
Complexity
Family Size
Large Family
Small Fam