M21 LDT ERP HC SEMANTIC SENSITIVITY N400 Family Size

Joanna Morris

2025-11-05

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_4A <- read_csv(file.path(dir_path, "fs_m21_ldt_mea_300500_050050_1_AB.csv"))
erp_4B <- read_csv(file.path(dir_path, "fs_m21_ldt_mea_300500_050050_1_BA.csv"))
dmg_lng_vsl <- read_csv(file.path(dir_path, "demo_lang_vsl_pca_hc.csv"))
```

```r
library(dplyr)

erp_4i <- bind_rows(
  erp_4A |> mutate(List = "AB"),
  erp_4B |> mutate(List = "BA")
)
```

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

```r
n400_words_b %>%
  count(family_size, base_freq, Semantic_Sensitivity)
```

```
|| # A tibble: 8 x 4
||   family_size base_freq Semantic_Sensitivity     n
||   <chr>       <chr>     <chr>                <int>
|| 1 Large       High      High Semantic          279
|| 2 Large       High      Low Semantic           261
|| 3 Large       Low       High Semantic          279
|| 4 Large       Low       Low Semantic           261
|| 5 Small       High      High Semantic          279
|| 6 Small       High      Low Semantic           261
|| 7 Small       Low       High Semantic          279
|| 8 Small       Low       Low Semantic           261
```

3.1 Nested ANOVA Model

```r
#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, 58      0.00    .987
|| 2                                family_size   1, 58      2.19    .144
|| 3                                  base_freq   1, 58      1.51    .224
|| 4           Semantic_Sensitivity:family_size   1, 58      0.05    .822
|| 5             Semantic_Sensitivity:base_freq   1, 58      0.45    .507
|| 6                      family_size:base_freq 1, 1498 64.68 ***   <.001
|| 7 Semantic_Sensitivity:family_size:base_freq 1, 1498 12.12 ***   <.001
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```r
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 -4702.4  9436.8
|| family_size in (1 + family_size + base_freq | SubjID)   13 -5008.7 10043.4 612.55  3  < 2.2e-16 ***
|| base_freq in (1 + family_size + base_freq | SubjID)     13 -4831.3  9688.6 257.82  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                    15 -5149.2 10328.4 893.57  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                     |       4.83e-06 | [0.00, 1.00]
|| family_size                              |           0.04 | [0.00, 1.00]
|| base_freq                                |           0.03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size         |       8.81e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:base_freq           |       7.64e-03 | [0.00, 1.00]
|| family_size:base_freq                    |           0.04 | [0.03, 1.00]
|| Semantic_Sensitivity:family_size:base_freq |       8.02e-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: 0.851
||      Marginal R2: 0.009
```

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, 1498	64.68 ***	<.001	0.04
Semantic_Sensitivity:family_size:base_freq	1, 1498	12.12 ***	<.001	8.02e-03

3.3.1 Family Size x Base Frequency Simple Contrasts

```
# Estimated marginal means for the family_size × base frequency interaction
(emm1 <- emmeans(anova_model_n400_words_b, ~ family_size * base_freq))
```

```
||  family_size base_freq emmean    SE   df lower.CL upper.CL
||  Large       High      0.6805 0.370 58.9  -0.0597    1.421
||  Small       High      0.7894 0.444 58.6  -0.0994    1.678
||  Large       Low       0.9669 0.354 59.0   0.2580    1.676
||  Small       Low       0.0461 0.423 58.7  -0.8009    0.893
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get all pairwise contrasts
emm1_contrasts <- contrast(emm1,  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 High - Small High",
          "Large Low - Small Low",
          "Large High - Large Low",
          "Small High - Small Low")
(emm1_contrasts_filtered <- subset(emm1_contrasts, contrast %in% keep))
```

```
||  contrast               estimate    SE   df t.ratio p.value
||  Large High - Small High   -0.109 0.282 64.5  -0.386  0.7005
||  Large High - Large Low    -0.286 0.197 72.5  -1.456  0.1496
||  Small High - Small Low     0.743 0.197 72.5   3.779  0.0003
||  Large Low - Small Low      0.921 0.282 64.5   3.270  0.0017
```

```
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get Confidence Intervals
(emm1_contrasts_filtered_ci <- confint(emm1_contrasts_filtered))
```

```
||  contrast               estimate    SE   df lower.CL upper.CL
||  Large High - Small High   -0.109 0.282 64.5   -0.671    0.454
||  Large High - Large Low    -0.286 0.197 72.5   -0.678    0.106
||  Small High - Small Low     0.743 0.197 72.5    0.351    1.135
||  Large Low - Small Low      0.921 0.282 64.5    0.358    1.483
||
|| 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
effs1 <- eff_size(emm1, sigma = sigma(m1), edf = df.residual(m1))

# Remove the two redundant rows (rows 3 and 4)
(effs1_filtered <- subset(effs1, !contrast %in% c("Large High - Small Low",
                                                  "Small High - Large Low")))
```

```
||  contrast               effect.size    SE   df lower.CL upper.CL
||  Large High - Small High     -0.0732 0.189 58.6   -0.452   0.3059
||  Large High - Large Low      -0.1926 0.132 58.9   -0.457   0.0721
||  Small High - Small Low       0.4999 0.132 58.6    0.235   0.7651
||  Large Low - Small Low        0.6193 0.190 58.7    0.240   0.9989
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| sigma used for effect sizes: 1.487
|| 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 $= -0.109; SE = 0.282; t = -0.386; p = 0.7005$. NS.
- At Low base frequency: Large - Small $= 0.921$ $SE = 0.282; t = 3.205; p = 0.0017$. Significant

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 $= -0.286; SE = 0.197; t = -1.456; p = 0.1496$. NS
- Small family_size: High - Low $= 0.743; SE = 0.197; t = 3.779; p = 0.0003$. Significant

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 x Base Frequency Interaction Contrasts

```
#  Interaction contrasts (difference-of-differences)
#    Compare base frequency effect in large vs small family)
contrast(emm1, interaction = "pairwise", by = NULL, adjust = "holm")
```

```
||  family_size_pairwise base_freq_pairwise estimate    SE   df t.ratio p.value
||  Large - Small        High - Low            -1.03 0.128 1498  -8.042  <.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:
||  contrast      estimate    SE   df lower.CL upper.CL
||  Large - Small   -0.109 0.282 64.5   -0.671    0.454
||
|| base_freq = Low:
||  contrast      estimate    SE   df lower.CL upper.CL
||  Large - Small    0.921 0.282 64.5    0.358    1.483
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
confint(contrast(emm1, interaction = c("pairwise", "pairwise")))
```

```
||  family_size_pairwise base_freq_pairwise estimate    SE   df lower.CL upper.CL
||  Large - Small        High - Low            -1.03 0.128 1498    -1.28   -0.779
||
|| 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:

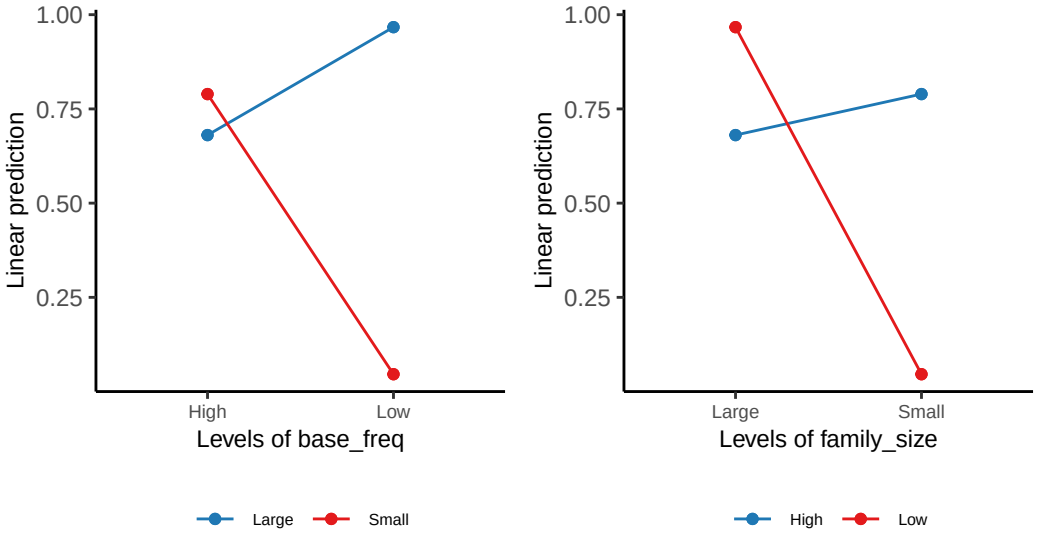
Estimate $= -1.03$; $SE = 0.128$; $t = -8.029$; $p < .0001$

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 Family Size x Base Frequency Interaction Plots

```
p1 <- emmip(anova_model_n400_words_b, family_size ~ base_freq) + my_style
p2 <- emmip(anova_model_n400_words_b, base_freq ~  family_size) + my_style

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



3.3.4 Sensitivity x Family Size x Base Frequency 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       High       0.656 0.514 58.9  -0.3731     1.69
||  Low Semantic         Large       High       0.705 0.532 58.9  -0.3593     1.77
||  High Semantic        Small       High       0.926 0.618 58.6  -0.3100     2.16
||  Low Semantic         Small       High       0.653 0.638 58.6  -0.6248     1.93
||  High Semantic        Large       Low        1.041 0.493 59.0   0.0555     2.03
||  Low Semantic         Large       Low        0.893 0.509 59.0  -0.1265     1.91
||  High Semantic        Small       Low       -0.165 0.588 58.7  -1.3422     1.01
||  Low Semantic         Small       Low        0.257 0.608 58.7  -0.9609     1.47
||
|| 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 High - High Semantic Large Low",
          "High Semantic Small High - High Semantic Small Low",
          "Low Semantic Large High - Low Semantic Large Low",
          "Low Semantic Small High - Low Semantic Small Low",
          "High Semantic Large High - High Semantic Small High",
          "High Semantic Large Low - High Semantic Small Low",
```

```
          "Low Semantic Large High - Low Semantic Small High",
          "Low Semantic Large Low - Low Semantic Small Low",
          "High Semantic Large High - Low Semantic Large High",
          "High Semantic Small High - Low Semantic Small High",
          "High Semantic Large Low - Low Semantic Small Low",
          "High Semantic Small Low - Low Semantic Small Low")

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

```
||  contrast                                        estimate    SE   df t.ratio p.value
||  High Semantic Large High - Low Semantic Large High   -0.0487 0.740 58.9  -0.066  0.9477
||  High Semantic Large High - High Semantic Small High  -0.2697 0.392 64.5  -0.689  0.4936
||  High Semantic Large High - High Semantic Large Low   -0.3850 0.273 72.5  -1.408  0.1634
||  Low Semantic Large High - Low Semantic Small High     0.0520 0.405 64.5   0.128  0.8982
||  Low Semantic Large High - Low Semantic Large Low     -0.1878 0.283 72.5  -0.664  0.5087
||  High Semantic Small High - Low Semantic Small High    0.2729 0.888 58.6   0.307  0.7597
||  High Semantic Small High - High Semantic Small Low    1.0904 0.273 72.5   3.987  0.0002
||  Low Semantic Small High - Low Semantic Small Low      0.3962 0.283 72.5   1.401  0.1654
||  High Semantic Large Low - High Semantic Small Low     1.2058 0.392 64.5   3.079  0.0030
||  High Semantic Large Low - Low Semantic Small Low      0.7845 0.783 73.4   1.002  0.3196
||  Low Semantic Large Low - Low Semantic Small Low       0.6360 0.405 64.5   1.571  0.1211
||  High Semantic Small Low - Low Semantic Small Low     -0.4213 0.846 58.7  -0.498  0.6206
||
|| 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 High - Low Semantic Large High   -0.0487 0.740 58.9   -1.529    1.432
||  High Semantic Large High - High Semantic Small High  -0.2697 0.392 64.5   -1.052    0.513
||  High Semantic Large High - High Semantic Large Low   -0.3850 0.273 72.5   -0.930    0.160
||  Low Semantic Large High - Low Semantic Small High     0.0520 0.405 64.5   -0.757    0.861
||  Low Semantic Large High - Low Semantic Large Low     -0.1878 0.283 72.5   -0.751    0.376
||  High Semantic Small High - Low Semantic Small High    0.2729 0.888 58.6   -1.505    2.051
||  High Semantic Small High - High Semantic Small Low    1.0904 0.273 72.5    0.545    1.635
||  Low Semantic Small High - Low Semantic Small Low      0.3962 0.283 72.5   -0.167    0.960
||  High Semantic Large Low - High Semantic Small Low     1.2058 0.392 64.5    0.424    1.988
||  High Semantic Large Low - Low Semantic Small Low      0.7845 0.783 73.4   -0.775    2.345
||  Low Semantic Large Low - Low Semantic Small Low       0.6360 0.405 64.5   -0.173    1.445
||  High Semantic Small Low - Low Semantic Small Low     -0.4213 0.846 58.7   -2.115    1.273
||
|| 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 High - Low Semantic Large High     -0.0328 0.498 58.9   -1.028    0.963
||  High Semantic Large High - High Semantic Small High    -0.1814 0.263 58.6   -0.708    0.346
||  High Semantic Large High - High Semantic Large Low     -0.2590 0.184 58.9   -0.627    0.109
||  Low Semantic Large High - Low Semantic Small High       0.0350 0.272 58.6   -0.510    0.580
||  Low Semantic Large High - Low Semantic Large Low       -0.1263 0.190 58.9   -0.507    0.254
||  High Semantic Small High - Low Semantic Small High      0.1836 0.597 58.6   -1.012    1.379
||  High Semantic Small High - High Semantic Small Low      0.7334 0.184 58.6    0.365    1.102
||  Low Semantic Small High - Low Semantic Small Low        0.2665 0.190 58.6   -0.114    0.647
||  High Semantic Large Low - High Semantic Small Low       0.8110 0.264 58.7    0.284    1.338
||  High Semantic Large Low - Low Semantic Small Low        0.5276 0.527 58.7   -0.526    1.581
||  Low Semantic Large Low - Low Semantic Small Low         0.4277 0.272 58.7   -0.117    0.973
||  High Semantic Small Low - Low Semantic Small Low       -0.2833 0.569 58.7   -1.423    0.856
||
|| sigma used for effect sizes: 1.487
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

3.3.5 Sensitivity x Family Size x Base Frequency 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 - Small        High - Low           -0.891 0.256 1498  -3.481  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 - Small        High - Low           -0.891 0.256 1498    -1.39   -0.389
||
|| 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:
||  contrast   estimate    SE   df t.ratio p.value
||  Low - High    0.385 0.273 72.5   1.408  0.1634
||
|| Semantic_Sensitivity = Low Semantic, family_size = Large:
||  contrast   estimate    SE   df t.ratio p.value
||  Low - High    0.188 0.283 72.5   0.664  0.5087
||
|| Semantic_Sensitivity = High Semantic, family_size = Small:
||  contrast   estimate    SE   df t.ratio p.value
||  Low - High   -1.090 0.273 72.5  -3.987  0.0002
||
|| Semantic_Sensitivity = Low Semantic, family_size = Small:
||  contrast   estimate    SE   df t.ratio p.value
||  Low - High   -0.396 0.283 72.5  -1.401  0.1654
||
|| 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 - High, Semantic_Sensitivity = High Semantic:
||  contrast1     estimate    SE   df t.ratio p.value
||  Small - Large   -1.475 0.178 1498  -8.287  <.0001
||
|| contrast = Low - High, Semantic_Sensitivity = Low Semantic:
||  contrast1     estimate    SE   df t.ratio p.value
||  Small - Large   -0.584 0.184 1498  -3.173  0.0015
||
|| Degrees-of-freedom method: kenward-roger
```

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

```
|| contrast = Low - High, Semantic_Sensitivity = High Semantic:
||  contrast1     estimate    SE   df lower.CL upper.CL
||  Small - Large   -1.475 0.178 1498   -1.825   -1.126
||
|| contrast = Low - High, Semantic_Sensitivity = Low Semantic:
||  contrast1     estimate    SE   df lower.CL upper.CL
||  Small - Large   -0.584 0.184 1498   -0.945   -0.223
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

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

Semantic Sensitivity	Family Size	High - Low Base Freq	t	p	Interpretation
High Sensitivity	Large Family	-0.3850 µV	-1.408	.1634 (n.s.)	Weak, non-sig. tendency: low-freq words slightly less negative.
Low Sensitivity	Large Family	-0.1878 µV	-0.664	.5087 (n.s.)	Essentially flat.
High Sensitivity	Small Family	1.0904 µV	3.987	**.0002**	More negative N400 for high-freq small-family words.
Low Sensitivity	Small Family	0.3962 µV	1.401	.1654 (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) | 95% CI | p | Interpretation |
| difference in base-freq effect | | | |

|—————————|————————————|———————-|———|———————-| | High Sensitivity | –1.475 µV | [-1.825, -1.126] | <.0001 | the base-freq effect flips between large- and small-family words. | | Low Sensitivity | –0.584 µV | [-0.945, -0.223] | 0.0015 | 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.

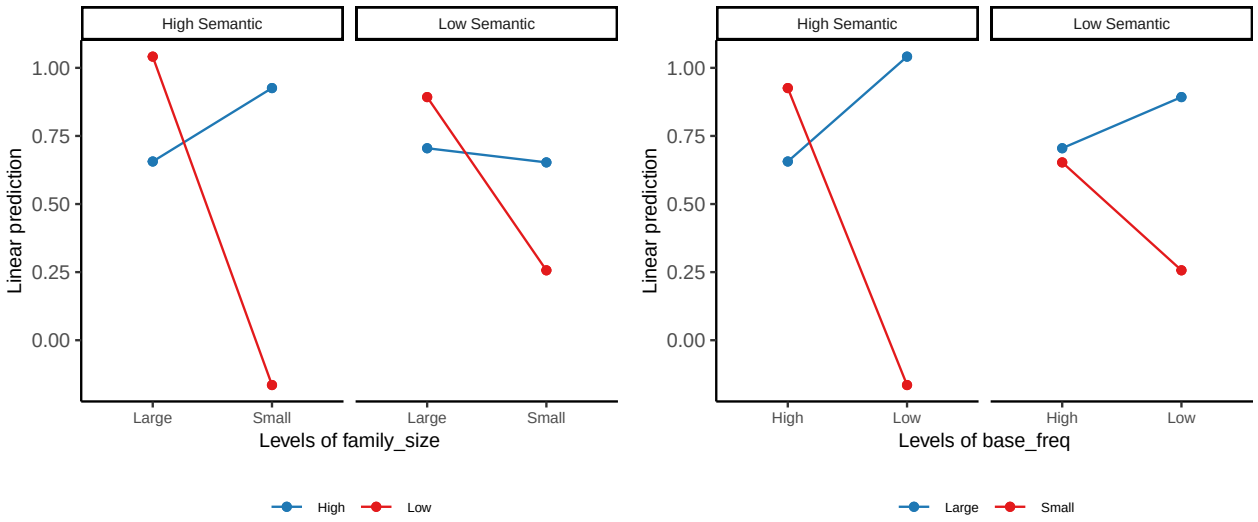
Interpretation: Both groups show that the base-frequency effect differs by family size, but this contrast is about twice as strong in the high-sensitivity group.

(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.

3.3.6 Plots Sensitivity x Family Size x Base Frequency

```
...
p3 <- emmip(anova_model_n400_words_b, base_freq ~ family_size  |  Semantic_Sensitivity) + my_style
p4 <- emmip(anova_model_n400_words_b, family_size ~ base_freq | Semantic_Sensitivity) + my_style
plot_grid(p3, p4,  ncol = 2)
```



4 N400 Nonword Data

```
n400_nonwords %>%
  count(family_size, complexity, Semantic_Sensitivity)
```

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, 58      0.00    .956
|| 2                                  family_size   1, 58      0.01    .931
|| 3                                   complexity   1, 58    4.15 *    .046
|| 4              Semantic_Sensitivity:family_size   1, 58      0.06    .811
|| 5               Semantic_Sensitivity:complexity   1, 58      0.01    .904
|| 6                     family_size:complexity 1, 1498   8.08 **    .005
|| 7 Semantic_Sensitivity:family_size:complexity 1, 1498 34.43 ***   <.001
|| ---
|| 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 -4944.9  9921.9
|| family_size in (1 + family_size + complexity | SubjID)   13 -5193.2 10412.4  496.56  3  < 2.2e-16 ***
|| complexity in (1 + family_size + complexity | SubjID)    13 -5273.0 10572.0  656.15  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                     15 -5468.1 10966.1 1046.23  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                      |       5.26e-05 | [0.00, 1.00]
|| family_size                               |       1.31e-04 | [0.00, 1.00]
|| complexity                                |           0.07 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size          |       9.90e-04 | [0.00, 1.00]
|| Semantic_Sensitivity:complexity           |       2.53e-04 | [0.00, 1.00]
|| family_size:complexity                    |       5.36e-03 | [0.00, 1.00]
|| Semantic_Sensitivity:family_size:complexity |          0.02 | [0.01, 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: 0.856
||      Marginal R2: 0.008
```

4.2 Main Effects and Interactions

Effect	df	F	p.value	eta-sqrd
complexity	1, 58	4.15 *	.046	0.07

Reponses to Simple words are more negative than those to Complex words.

Effect	df	F	p.value	eta-sqrd
family_size:complexity	1, 1498	8.08 **	.005	5.36e-03
Semantic_Sensitivity:family_size:complexity	1, 1498	34.43 ***	<.001	0.02

4.2.1 Family Size x Complexity 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
||  Small       Simple     -0.4824 0.412 58.8   -1.308    0.343
||  Large       Simple     -0.6574 0.461 58.7   -1.580    0.265
||  Small       Complex    -0.0484 0.431 58.8   -0.911    0.814
||  Large       Complex     0.1735 0.431 58.8   -0.689    1.036
||
|| 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("Small Simple - Large Simple",
          "Small Complex - Large Complex",
          "Small Simple - Small Complex",
          "Large Simple - Large Complex")
(emm2_contrasts_filtered <- subset(emm2_contrasts, contrast %in% keep2))
```

```
||  contrast                    estimate    SE df t.ratio p.value
||  Small Simple - Large Simple     0.175 0.278 66   0.629  0.5316
||  Small Simple - Small Complex   -0.434 0.318 64  -1.364  0.1774
||  Large Simple - Large Complex   -0.831 0.318 64  -2.611  0.0112
||  Small Complex - Large Complex  -0.222 0.278 66  -0.798  0.4279
||
|| 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
||  Small Simple - Large Simple     0.175 0.278 66   -0.380    0.730
||  Small Simple - Small Complex   -0.434 0.318 64   -1.070    0.202
||  Large Simple - Large Complex   -0.831 0.318 64   -1.467   -0.195
||  Small Complex - Large Complex  -0.222 0.278 66   -0.777    0.333
||
|| 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
||  Small Simple - Large Simple       0.108 0.172 58.7   -0.235    0.451
||  Small Simple - Small Complex     -0.268 0.196 58.8   -0.660    0.125
||  Large Simple - Large Complex     -0.512 0.196 58.7   -0.905   -0.120
||  Small Complex - Large Complex    -0.137 0.172 58.8   -0.480    0.206
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| sigma used for effect sizes: 1.622
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

4.2.2 Family Size x Complexity Interaction Contrasts

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
||  Small - Large        Simple - Complex       0.397 0.14 1498   2.842  0.0045
||
|| 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"))
```

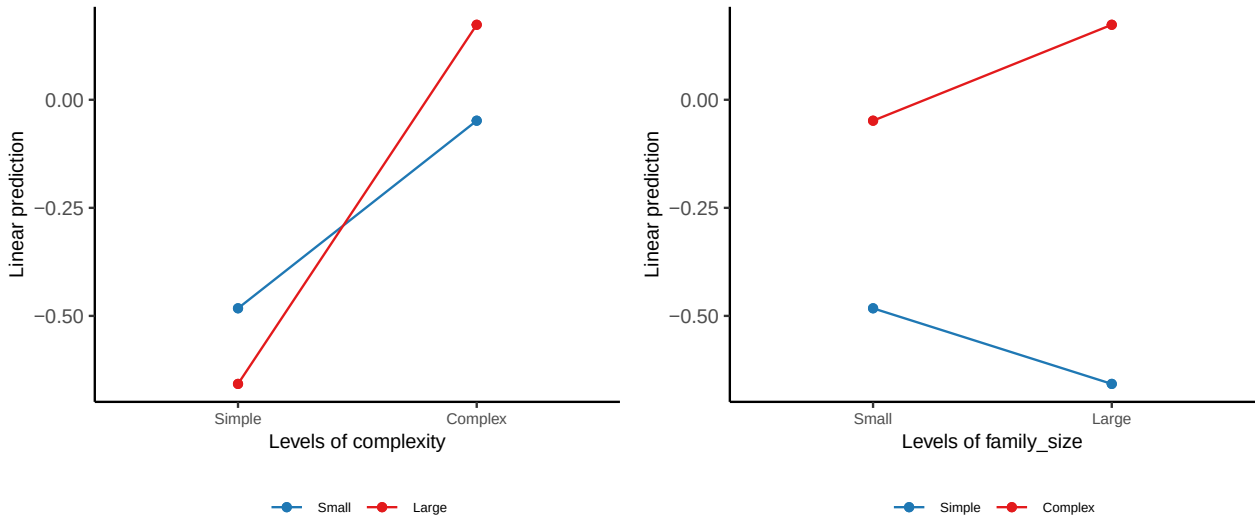
```
|| complexity = Simple:
||  contrast      estimate    SE df lower.CL upper.CL
||  Small - Large    0.175 0.278 66   -0.380    0.730
||
|| complexity = Complex:
||  contrast      estimate    SE df lower.CL upper.CL
||  Small - Large   -0.222 0.278 66   -0.777    0.333
||
|| Results are averaged over 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
||  Small - Large        Simple - Complex       0.397 0.14 1498    0.123    0.671
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

4.2.3 Family Size x Complexity Plots

```
p5 <- emmip(anova_model_n400_nonwords, family_size ~ complexity) + my_style
p6 <- emmip(anova_model_n400_nonwords, complexity ~  family_size) + my_style
plot_grid(p5, p6, ncol = 2)
```



4.2.4 Semantic Sensitivity x Family Size x Complexity Simple Contrasts

Compare High vs Low Semantic Sensitivity within each combination of Family Size and Complexity

```
# Estimated marginal means for the family_size × complexity interaction
(emm3 <- emmeans(anova_model_n400_nonwords, ~ Semantic_Sensitivity * family_size * complexity))
```

```
||  Semantic_Sensitivity family_size complexity   emmean    SE   df lower.CL upper.CL
||  High Semantic        Small       Simple     -0.67970 0.574 58.8   -1.827    0.468
||  Low Semantic         Small       Simple     -0.28513 0.593 58.8   -1.472    0.902
||  High Semantic        Large       Simple     -0.38040 0.641 58.7   -1.663    0.903
||  Low Semantic         Large       Simple     -0.93432 0.663 58.7   -2.261    0.392
||  High Semantic        Small       Complex     0.12639 0.599 58.8   -1.073    1.326
||  Low Semantic         Small       Complex    -0.22320 0.620 58.8   -1.463    1.017
||  High Semantic        Large       Complex     0.00323 0.600 58.8   -1.197    1.203
||  Low Semantic         Large       Complex     0.34384 0.620 58.8   -0.897    1.584
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get all pairswise contrasts
emm3_contrasts <- contrast(emm3,  method = "pairwise", by = NULL, adjust = "none")
# emm3_contrasts

# 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("High Semantic Large Simple - High Semantic Large Complex",
           "High Semantic Small Simple - High Semantic Small Complex",
           "Low Semantic Large Simple - Low Semantic Large Complex",
           "Low Semantic Small Simple - Low Semantic Small Complex",
           "High Semantic Small Simple - High Semantic Large Simple",
           "High Semantic Small Complex - High Semantic Large Complex",
           "Low Semantic Small Simple - High Semantic Large Simple",
           "Low Semantic Small Complex - Low Semantic Large Complex",
           "High Semantic Large Simple - Low Semantic Large Simple",
           "High Semantic Large Complex - Low Semantic Large Complex",
           "High Semantic Small Simple - Low Semantic Small Simple",
           "High Semantic Small Complex - Low Semantic Small Complex")

(emm3_contrasts_filtered <- subset(emm3_contrasts, contrast %in% keep2))
```

```
||  contrast                                                estimate    SE   df t.ratio p.value
||  High Semantic Small Simple - Low Semantic Small Simple    -0.3946 0.825 58.8  -0.478  0.6342
||  High Semantic Small Simple - High Semantic Large Simple   -0.2993 0.387 66.0  -0.774  0.4418
||  High Semantic Small Simple - High Semantic Small Complex  -0.8061 0.442 64.0  -1.822  0.0731
||  Low Semantic Small Simple - High Semantic Large Simple     0.0953 0.873 70.6   0.109  0.9134
||  Low Semantic Small Simple - Low Semantic Small Complex    -0.0619 0.457 64.0  -0.135  0.8927
||  High Semantic Large Simple - Low Semantic Large Simple     0.5539 0.922 58.7   0.601  0.5504
||  High Semantic Large Simple - High Semantic Large Complex  -0.3836 0.442 64.0  -0.867  0.3891
||  Low Semantic Large Simple - Low Semantic Large Complex    -1.2782 0.457 64.0  -2.794  0.0069
||  High Semantic Small Complex - Low Semantic Small Complex   0.3496 0.862 58.8   0.406  0.6865
||  High Semantic Small Complex - High Semantic Large Complex  0.1232 0.387 66.0   0.318  0.7512
||  Low Semantic Small Complex - Low Semantic Large Complex   -0.5670 0.400 66.0  -1.418  0.1609
||  High Semantic Large Complex - Low Semantic Large Complex  -0.3406 0.862 58.8  -0.395  0.6943
||
|| Degrees-of-freedom method: kenward-roger
```

```
# Get Confidence Intervals
(emm3_contrasts_filtered_ci <- confint(emm3_contrasts_filtered))
```

```
||  contrast                                                estimate    SE   df lower.CL upper.CL
||  High Semantic Small Simple - Low Semantic Small Simple    -0.3946 0.825 58.8   -2.045   1.2563
||  High Semantic Small Simple - High Semantic Large Simple   -0.2993 0.387 66.0   -1.072   0.4730
||  High Semantic Small Simple - High Semantic Small Complex  -0.8061 0.442 64.0   -1.690   0.0778
||  Low Semantic Small Simple - High Semantic Large Simple     0.0953 0.873 70.6   -1.646   1.8368
||  Low Semantic Small Simple - Low Semantic Small Complex    -0.0619 0.457 64.0   -0.976   0.8519
||  High Semantic Large Simple - Low Semantic Large Simple     0.5539 0.922 58.7   -1.292   2.3994
||  High Semantic Large Simple - High Semantic Large Complex  -0.3836 0.442 64.0   -1.268   0.5003
||  Low Semantic Large Simple - Low Semantic Large Complex    -1.2782 0.457 64.0   -2.192  -0.3643
||  High Semantic Small Complex - Low Semantic Small Complex   0.3496 0.862 58.8   -1.375   2.0744
||  High Semantic Small Complex - High Semantic Large Complex  0.1232 0.387 66.0   -0.649   0.8955
||  Low Semantic Small Complex - Low Semantic Large Complex   -0.5670 0.400 66.0   -1.366   0.2314
||  High Semantic Large Complex - Low Semantic Large Complex  -0.3406 0.862 58.8   -2.066   1.3851
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Get effect sizes
# Get all pairwise effect sizes
effs3 <- eff_size(emm3, sigma = sigma(m2), edf = df.residual(m2))

# Remove the  redundant rows
(effs3_filtered <- subset(effs3, contrast %in% keep2))
```

```
||  contrast                                                effect.size    SE   df lower.CL upper.CL
||  High Semantic Small Simple - Low Semantic Small Simple       -0.2433 0.509 58.8   -1.261    0.775
||  High Semantic Small Simple - High Semantic Large Simple      -0.1846 0.239 58.7   -0.662    0.293
||  High Semantic Small Simple - High Semantic Small Complex     -0.4971 0.273 58.8   -1.043    0.049
```

```
||  Low Semantic Small Simple - High Semantic Large Simple         0.0588 0.539 58.7   -1.019    1.137
||  Low Semantic Small Simple - Low Semantic Small Complex        -0.0382 0.282 58.8   -0.603    0.526
||  High Semantic Large Simple - Low Semantic Large Simple         0.3416 0.569 58.7   -0.797    1.480
||  High Semantic Large Simple - High Semantic Large Complex      -0.2366 0.273 58.7   -0.783    0.310
||  Low Semantic Large Simple - Low Semantic Large Complex        -0.7883 0.282 58.7   -1.353   -0.224
||  High Semantic Small Complex - Low Semantic Small Complex       0.2156 0.532 58.8   -0.848    1.279
||  High Semantic Small Complex - High Semantic Large Complex      0.0760 0.239 58.8   -0.401    0.553
||  Low Semantic Small Complex - Low Semantic Large Complex       -0.3497 0.247 58.8   -0.843    0.144
||  High Semantic Large Complex - Low Semantic Large Complex      -0.2101 0.532 58.8   -1.274    0.854
||
|| sigma used for effect sizes: 1.622
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

4.2.5 Semantic Sensitivity x Family Size x Complexity Interaction Contrasts

The interaction contrast tests whether the difference in the complexity effect for large vs small families differs across sensitivity?

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

```
#  Interaction contrasts (difference-of-differences)
#    Compare complexity effect in large vs small family)
contrast(emm3, interaction = "pairwise", by = NULL, adjust = "holm")
```

```
||  Semantic_Sensitivity_pairwise family_size_pairwise complexity_pairwise estimate    SE   df t.ratio p.value
||  High Semantic - Low Semantic  Small - Large        Simple - Complex       -1.64 0.279 1498  -5.868  <.0001
||
|| Degrees-of-freedom method: kenward-roger
```

```
confint(contrast(emm3, interaction = c("pairwise", "pairwise")))
```

```
||  Semantic_Sensitivity_pairwise family_size_pairwise complexity_pairwise estimate    SE   df lower.CL upper.CL
||  High Semantic - Low Semantic  Small - Large        Simple - Complex       -1.64 0.279 1498    -2.19    -1.09
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```
# Compute the A1 - A2 difference within each combination of B × C
(complexity_diff <- contrast(emm3, method = "revpairwise",
                             by = c("Semantic_sensitivity", "family_size"),
                             simple = "complexity"))
```

```
|| Semantic_Sensitivity = High Semantic, family_size = Small:
||  contrast         estimate    SE df t.ratio p.value
||  Complex - Simple   0.8061 0.442 64   1.822  0.0731
||
|| Semantic_Sensitivity = Low Semantic, family_size = Small:
||  contrast         estimate    SE df t.ratio p.value
||  Complex - Simple   0.0619 0.457 64   0.135  0.8927
||
|| Semantic_Sensitivity = High Semantic, family_size = Large:
||  contrast         estimate    SE df t.ratio p.value
||  Complex - Simple   0.3836 0.442 64   0.867  0.3891
||
|| Semantic_Sensitivity = Low Semantic, family_size = Large:
||  contrast         estimate    SE df t.ratio p.value
||  Complex - Simple   1.2782 0.457 64   2.794  0.0069
||
|| Degrees-of-freedom method: kenward-roger
```

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

```
|| contrast = Complex - Simple, Semantic_Sensitivity = High Semantic:
||  contrast1     estimate    SE   df t.ratio p.value
||  Large - Small   -0.422 0.194 1498  -2.176  0.0297
||
|| contrast = Complex - Simple, Semantic_Sensitivity = Low Semantic:
||  contrast1     estimate    SE   df t.ratio p.value
||  Large - Small    1.216 0.201 1498   6.059  <.0001
||
|| Degrees-of-freedom method: kenward-roger
```

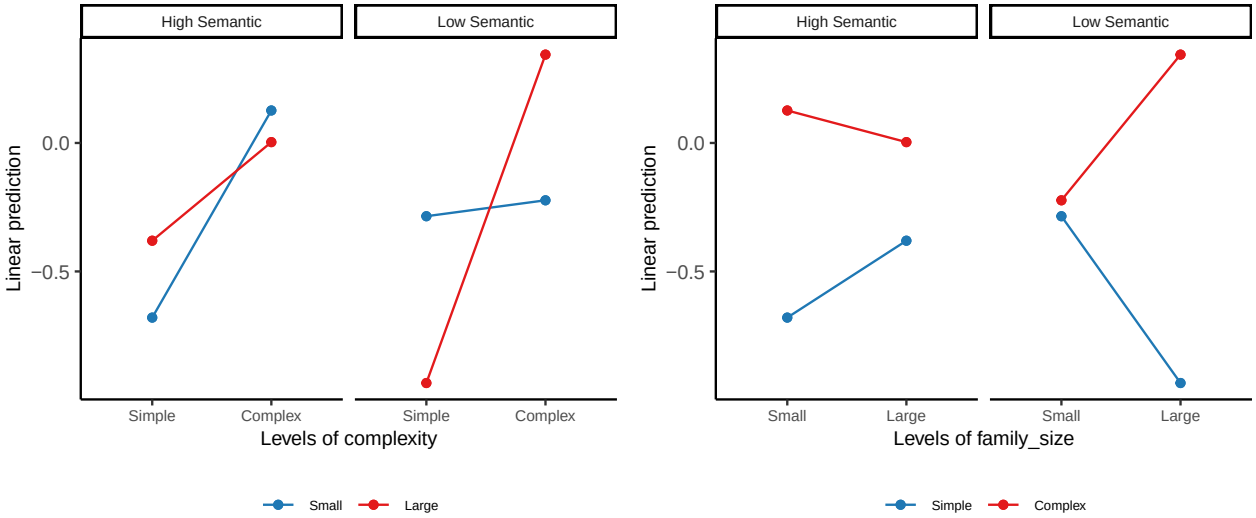
```
# Get confidence intervals
confint(familysize_complexity_int_within_sensitivity)
```

```
|| contrast = Complex - Simple, Semantic_Sensitivity = High Semantic:
||  contrast1     estimate    SE   df lower.CL upper.CL
||  Large - Small   -0.422 0.194 1498   -0.803  -0.0416
```

```
||
|| contrast = Complex - Simple, Semantic_Sensitivity = Low Semantic:
||  contrast1     estimate    SE   df lower.CL upper.CL
||  Large - Small    1.216 0.201 1498    0.822   1.6100
||
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

4.2.6 Semantic Sensitivity x Family Size x Complexity Plots

```
p7 <- emmip(anova_model_n400_nonwords, family_size ~   complexity|   Semantic_Sensitivity)  + my_style
p8 <- emmip(anova_model_n400_nonwords, complexity ~ family_size  | Semantic_Sensitivity)  + my_style
plot_grid(p7, p8, ncol = 2)
```



- High-semantic participants: show a greater complexity effect for small families than for large families.
- Low-semantic participants: show the opposite pattern—a stronger complexity effect for large-family items and no effect for small ones.

Summary (N400) - High-semantic participants: familiarity (large families) neutralizes complexity effects. - Low-semantic participants: familiarity amplifies complexity effects.