M21 LDT ERP HC ORTHOGRAPIC SENSITIVITY N400 Family Size

Joanna Morris

2025-11-05

Contents

Set parameters **1**

1 Load data files **1**

2 Format data files **2**

3 N400 Word Data **2**
3.1 Nested ANOVA Model . . . 2
3.2 Main Effects . . . 3
3.3 Interactions . . . 3
3.3.1 Simple Contrasts . . . 3
3.3.2 Interaction Contrasts . . . 4
3.4 Plots . . . 5

4 N400 Nonword Data **6**
4.1 Compute the ANOVA . . . 6
4.2 Main Effects . . . 6
4.3 Interactions . . . 6
4.3.1 Simple Contrasts . . . 7
4.3.2 Interaction Contrasts . . . 7
4.4 Plots . . . 8

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"))
```

```
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

3.1 Nested ANOVA Model

```
n400_words_b %>%
  count(family_size, base_freq, Orthographic_Sensitivity)
```

```
|| # A tibble: 8 x 4
||   family_size base_freq Orthographic_Sensitivity     n
||   <chr>       <chr>     <chr>                    <int>
|| 1 Large       High      High Orthographic          306
|| 2 Large       High      Low Orthographic           234
|| 3 Large       Low       High Orthographic          306
|| 4 Large       Low       Low Orthographic           234
|| 5 Small       High      High Orthographic          306
|| 6 Small       High      Low Orthographic           234
|| 7 Small       Low       High Orthographic          306
|| 8 Small       Low       Low Orthographic           234
```

```
#Fit ANOVA model
anova_model_n400_words_b <- mixed(
    value ~ Orthographic_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 ~ Orthographic_Sensitivity * family_size * base_freq +
|| Model:     (1 + family_size + base_freq | SubjID) + (1 | SubjID:chlabel)
|| Data: n400_words_b
||                                         Effect      df         F p.value
|| 1                       Orthographic_Sensitivity   1, 58      0.48    .492
|| 2                                    family_size   1, 58      2.39    .128
|| 3                                      base_freq   1, 58      1.26    .265
|| 4           Orthographic_Sensitivity:family_size   1, 58      0.26    .615
|| 5             Orthographic_Sensitivity:base_freq   1, 58      0.81    .373
|| 6                          family_size:base_freq 1, 1498 60.87 ***   <.001
|| 7 Orthographic_Sensitivity:family_size:base_freq 1, 1498    3.99 *    .046
|| ---
|| 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 ~ Orthographic_Sensitivity + family_size + base_freq + (1 + family_size + base_freq | SubjID) + (1 | SubjID:chlabel) + Orthographic_Sensi
||                                                    npar  logLik     AIC    LRT Df Pr(>Chisq)
|| <none>                                               16 -4706.0  9444.1
|| family_size in (1 + family_size + base_freq | SubjID)   13 -5009.6 10045.2 607.13  3  < 2.2e-16 ***
|| base_freq in (1 + family_size + base_freq | SubjID)     13 -4833.0  9692.0 253.94  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                    15 -5150.4 10330.7 888.65  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
|| -------------------------------------------------------------------------------
|| Orthographic_Sensitivity                         |       8.16e-03 | [0.00, 1.00]
|| family_size                                      |           0.04 | [0.00, 1.00]
|| base_freq                                        |           0.02 | [0.00, 1.00]
|| Orthographic_Sensitivity:family_size             |       4.39e-03 | [0.00, 1.00]
|| Orthographic_Sensitivity:base_freq               |           0.01 | [0.00, 1.00]
|| family_size:base_freq                            |           0.04 | [0.02, 1.00]
|| Orthographic_Sensitivity:family_size:base_freq |       2.65e-03 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].
```

```r
# Compute Marginal(fixed effects only) and Conditional(fixed + random effects) R²
r2(anova_model_n400_words_b)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.850
||      Marginal R2: 0.014
```

3.2 Main Effects

No significant main effects

3.3 Interactions

A two-way interaction betweeen Family Size and Base Frequency

Effect	df	F	p.value	
family_size:base_freq	1, 4738.00	87.28 ***	<.001	0.02

3.3.1 Simple Contrasts

```r
# 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       0.653 0.372 58.9  -0.0907    1.397
||  Small       High       0.731 0.445 58.6  -0.1587    1.621
||  Large       Low        0.948 0.357 59.0   0.2332    1.662
||  Small       Low        0.016 0.426 58.7  -0.8362    0.868
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

```r
# Get all pairswise 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  -0.0783 0.284 64.5  -0.276  0.7832
||  Large High - Large Low   -0.2948 0.198 72.7  -1.490  0.1405
||  Small High - Small Low    0.7153 0.198 72.7   3.615  0.0005
||  Large Low - Small Low     0.9318 0.284 64.5   3.286  0.0016
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

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

```
||  contrast               estimate    SE   df lower.CL upper.CL
||  Large High - Small High  -0.0783 0.284 64.5   -0.645   0.4881
||  Large High - Large Low   -0.2948 0.198 72.7   -0.689   0.0995
||  Small High - Small Low    0.7153 0.198 72.7    0.321   1.1096
||  Large Low - Small Low     0.9318 0.284 64.5    0.365   1.4982
||
```

```
|| Results are averaged over the levels of: Orthographic_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% keep))
```

```
||  contrast               effect.size    SE   df lower.CL upper.CL
||  Large High - Small High     -0.0525 0.190 58.6   -0.433   0.3281
||  Large High - Large Low      -0.1978 0.133 58.9   -0.463   0.0679
||  Small High - Small Low       0.4798 0.133 58.6    0.214   0.7457
||  Large Low - Small Low        0.6250 0.190 58.7    0.244   1.0061
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| sigma used for effect sizes: 1.491
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

For low-frequency bases, small-family words elicit more negative amplitudes than large-family words. When base frequency is high, family size has no effect.

- Low base frequency: Large - Small $= 0.9381$ $SE = 0.250; z = 3.747; p = 0.0002$. This difference is statistically significant.
- High base frequency: Large - Small $= 0.0395$ $SE = 0.272; z = 0.158; p = 0.8748$. This difference is NS.

For small-family words, low base frequency bases elicit more negative responses than high base frequency bases; when family_size is large, the difference is marginal.

- Small family_size: High - Low $= 0.5495; SE = 0.188; z = 2.926; p = 0.0034$. Significant difference
- Large family_size: High - Low $= -0.3491; SE = 0.188; z = -1.859; p = 0.0631$. This is a trend but not significant.

3.3.2 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            -1.01 0.129 1498  -7.802  <.0001
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
```

```
# Get confidence intervals, for the 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  -0.0783 0.284 64.5   -0.645    0.488
||
|| base_freq = Low:
||  contrast      estimate    SE   df lower.CL upper.CL
||  Large - Small   0.9318 0.284 64.5    0.365    1.498
||
|| Results are averaged over the levels of: Orthographic_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            -1.01 0.129 1498    -1.26   -0.756
||
|| Results are averaged over the levels of: Orthographic_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:

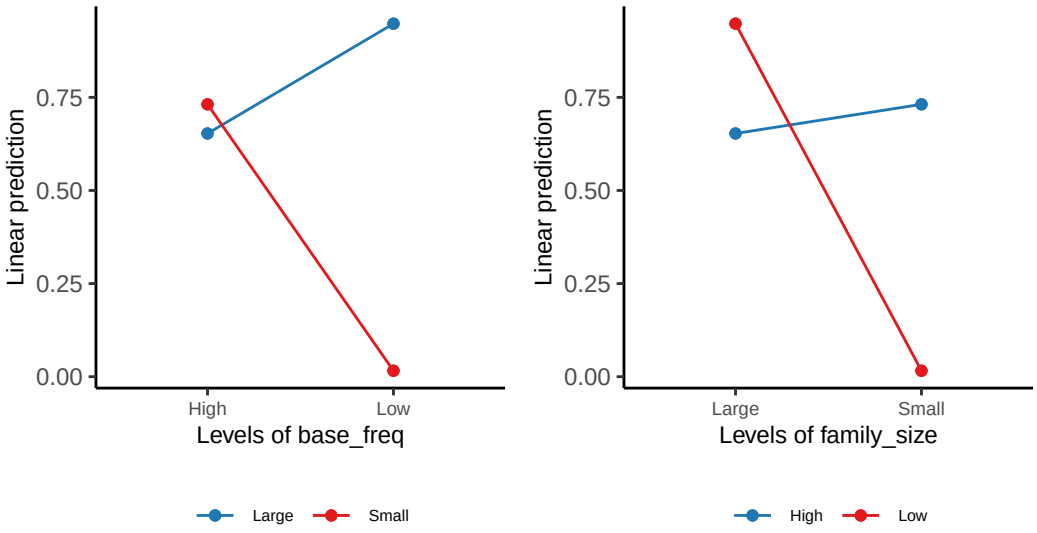
Estimate $= -0.899$; $SE = 0.0962$; $z = -9.342$; $p < .0001$

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.

3.4 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)
```



4 N400 Nonword Data

4.1 Compute the ANOVA

```
# n400_nonwords %>%
#   count(family_size, complexity, Orthographic_Sensitivity)

anova_model_n400_nonwords <- mixed(
    value ~ Orthographic_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 ~ Orthographic_Sensitivity * family_size * complexity +
|| Model:     (1 + family_size + complexity | SubjID) + (1 | SubjID:chlabel)
|| Data: n400_nonwords
||                                          Effect      df       F p.value
|| 1                        Orthographic_Sensitivity   1, 58    0.00    .975
|| 2                                     family_size   1, 58    0.03    .853
|| 3                                      complexity   1, 58  4.64 *    .035
|| 4            Orthographic_Sensitivity:family_size   1, 58    0.47    .498
|| 5             Orthographic_Sensitivity:complexity   1, 58    0.85    .362
|| 6                          family_size:complexity 1, 1498 7.28 **    .007
|| 7 Orthographic_Sensitivity:family_size:complexity 1, 1498    0.59    .442
|| ---
|| 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 ~ Orthographic_Sensitivity + family_size + complexity + (1 + family_size + complexity | SubjID) + (1 | SubjID:chlabel) + Orthographic_Sen
||                                                    npar  logLik     AIC     LRT Df Pr(>Chisq)
|| <none>                                               16 -4960.8  9953.6
|| family_size in (1 + family_size + complexity | SubjID)   13 -5201.9 10429.8  482.18  3  < 2.2e-16 ***
|| complexity in (1 + family_size + complexity | SubjID)    13 -5277.9 10581.9  634.32  3  < 2.2e-16 ***
|| (1 | SubjID:chlabel)                                     15 -5473.1 10976.1 1024.54  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
|| -------------------------------------------------------------------------------
|| Orthographic_Sensitivity                         |       1.64e-05 | [0.00, 1.00]
|| family_size                                      |       5.95e-04 | [0.00, 1.00]
|| complexity                                       |           0.07 | [0.00, 1.00]
|| Orthographic_Sensitivity:family_size             |       7.97e-03 | [0.00, 1.00]
|| Orthographic_Sensitivity:complexity              |           0.01 | [0.00, 1.00]
|| family_size:complexity                           |       4.84e-03 | [0.00, 1.00]
|| Orthographic_Sensitivity:family_size:complexity |       3.94e-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: 0.853
||      Marginal R2: 0.008
```

4.2 Main Effects

Main effect of Complexity. Simple words elicit more negative responses than complex ones.

4.3 Interactions

Effec	df	F	p.value	
complexity	1, 58	4.64 *	.035	0.07
family_size:complexity	2, 4736	7.28 **	<.001	7.67e-03

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
||  Small       Simple     -0.5151 0.415 58.9   -1.346    0.315
||  Large       Simple     -0.6569 0.465 58.7   -1.588    0.274
||  Small       Complex    -0.0379 0.435 58.8   -0.908    0.832
||  Large       Complex     0.2045 0.434 58.8   -0.663    1.072
||
|| Results are averaged over the levels of: Orthographic_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 - Small Complex",
           "Large Simple - Large Complex",
           "Small Simple - Large Simple",
           "Small Complex - Large Complex")
(emm2_contrasts_filtered <- subset(emm2_contrasts, contrast %in% keep2))
```

```
||  contrast                     estimate    SE   df t.ratio p.value
||  Small Simple - Large Simple     0.142 0.280 66.3   0.507  0.6140
||  Small Simple - Small Complex   -0.477 0.319 64.2  -1.497  0.1394
||  Large Simple - Large Complex   -0.861 0.319 64.2  -2.701  0.0088
||  Small Complex - Large Complex  -0.242 0.280 66.3  -0.866  0.3896
||
|| Results are averaged over the levels of: Orthographic_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     0.142 0.280 66.3   -0.417    0.700
||  Small Simple - Small Complex   -0.477 0.319 64.2   -1.114    0.160
||  Large Simple - Large Complex   -0.861 0.319 64.2   -1.498   -0.224
||  Small Complex - Large Complex  -0.242 0.280 66.3   -0.801    0.316
||
|| Results are averaged over the levels of: Orthographic_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       0.0865 0.171 58.7   -0.255   0.4280
||  Small Simple - Small Complex     -0.2910 0.195 58.8   -0.680   0.0982
||  Large Simple - Large Complex     -0.5253 0.195 58.7   -0.915  -0.1360
||  Small Complex - Large Complex    -0.1478 0.171 58.8   -0.489   0.1937
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| sigma used for effect sizes: 1.64
|| 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
||  Small - Large        Simple - Complex       0.384 0.142 1498   2.698  0.0071
||
|| Results are averaged over the levels of: Orthographic_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, ~ complexity | family_size), "pairwise"))
```

```
|| family_size = Small:
||  contrast         estimate    SE   df lower.CL upper.CL
||  Simple - Complex   -0.477 0.319 64.2    -1.11    0.160
||
|| family_size = Large:
||  contrast         estimate    SE   df lower.CL upper.CL
||  Simple - Complex   -0.861 0.319 64.2    -1.50   -0.224
||
|| Results are averaged over the levels of: Orthographic_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       0.384 0.142 1498    0.105    0.663
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
```

4.4 Plots

```
p3 <- emmip(anova_model_n400_nonwords, complexity ~ family_size)  + my_style
p4 <- emmip(anova_model_n400_nonwords, family_size ~ complexity)  + my_style
plot_grid(p3, p4, ncol = 3)
```