

M21 LDT ERP HC ORTHOGRAPHIC SENSITIVITY N250 Base Frequency

Joanna Morris

2025-11-05

Contents

Set parameters	1
1 Load data files	1
2 Format data files	1
3 N250 Nonword Data	2
3.1 Compute the ANOVA	2
3.2 Main Effects and Interactions	3
3.2.1 Interaction Contrasts	3
3.3 Plots	4

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_2A <- read_csv(file.path(dir_path, "bf_m21_ldt_mea_200300_050050_1_AB.csv"))
erp_2B <- read_csv(file.path(dir_path, "bf_m21_ldt_mea_200300_050050_1_BA.csv"))

dmg_lng_vsl <- read_csv(file.path(dir_path, "demo_lang_vsl_pca_hc.csv"))

library(dplyr)

erp_2i <- bind_rows(
  erp_2A |> mutate(List = "AB"),
  erp_2B |> 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 `separate` 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 `thematate` function from the `dplyr` package along with the `case_when` function. The `case_when` function is a sequence of two-sided formulas. The left hand side determines which values match this case. The right hand side provides the replacement value.

3 N250 Nonword Data

```
n250_nonwords %>%
  count(Base_Frequency, Complexity, Orthographic_Sensitivity)
```

3.1 Compute the ANOVA

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

```
anova_model_n250_nonwords

|| Mixed Model Anova Table (Type 3 tests, KR-method)
||
|| Model: value ~ Orthographic_Sensitivity * Base_Frequency * Complexity +
|| Model: (1 + Base_Frequency + Complexity | SubjID) + (1 | SubjID:chlabel)
|| Data: n250_nonwords
||
|| Effect df F p.value
|| 1 Orthographic_Sensitivity 1, 58 0.13 .715
|| 2 Base_Frequency 1, 58 0.04 .851
|| 3 Complexity 1, 58 1.16 .286
|| 4 Orthographic_Sensitivity:Base_Frequency 1, 58 1.30 .260
|| 5 Orthographic_Sensitivity:Complexity 1, 58 1.70 .197
|| 6 Base_Frequency:Complexity 1, 1498 1.49 .222
|| 7 Orthographic_Sensitivity:Base_Frequency:Complexity 1, 1498 0.35 .555
|| ---
|| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| value ~ Orthographic_Sensitivity + Base_Frequency + Complexity + (1 + Base_Frequency + Complexity | SubjID) + (1 | SubjID:chlabel) + Orthographic_Sensitivity:Base_Frequency:Complexity
|| npar logLik AIC LRT Df Pr(>Chisq)
|| <none> 16 -6240.7 12513
|| Base_Frequency in (1 + Base_Frequency + Complexity | SubjID) 13 -6447.7 12921 414.06 3 < 2.2e-16 ***
|| Complexity in (1 + Base_Frequency + Complexity | SubjID) 13 -6486.1 12998 490.87 3 < 2.2e-16 ***
|| (1 | SubjID:chlabel) 15 -6362.9 12756 244.42 1 < 2.2e-16 ***
|| ---
|| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter | Eta2 (partial) | 95% CI
|| -----|-----|-----
|| Orthographic_Sensitivity | 2.32e-03 | [0.00, 1.00]
|| Base_Frequency | 6.17e-04 | [0.00, 1.00]
|| Complexity | 0.02 | [0.00, 1.00]
|| Orthographic_Sensitivity:Base_Frequency | 0.02 | [0.00, 1.00]
|| Orthographic_Sensitivity:Complexity | 0.03 | [0.00, 1.00]
|| Base_Frequency:Complexity | 9.94e-04 | [0.00, 1.00]
|| Orthographic_Sensitivity:Base_Frequency:Complexity | 2.32e-04 | [0.00, 1.00]
||
|| - One-sided CIs: upper bound fixed at [1.00].

# Compute Marginal(fixed effects only) and Conditional(fixed + random effects) R^2
r2(anova_model_n250_nonwords)
```

```
|| # R2 for Mixed Models
||
|| Conditional R2: 0.704
|| Marginal R2: 0.009
```

3.2 Main Effects and Interactions

Effect	df	F	p	Eta2 (partial)
Base_Frequency:Complexity	1, 4738.00	8.33 **	.004	1.76e-03

```
# Estimated marginal means for the Base_Frequency x Complexity interaction
(emm2 <- emmeans(anova_model_n250_nonwords, ~ Base_Frequency * Complexity))

|| Base_Frequency Complexity emmean SE df lower.CL upper.CL
|| High Complex -1.29 0.665 59.5 -2.62 0.0367
|| Low Complex -1.01 0.676 59.4 -2.37 0.3398
|| High Simple -1.72 0.575 60.0 -2.87 -0.5737
|| Low Simple -1.81 0.606 59.8 -3.02 -0.5935
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95

# Get all pairwise contrasts
emm2_contrasts <- contrast(emm2, method = "pairwise", by = NULL, adjust = "none")
# emm2_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 Complex - Low Complex",
           "High Simple - Low Simple",
           "High Complex - High Simple",
           "Low Complex - Low Simple")

(emm2_contrasts_filtered <- subset(emm2_contrasts, contrast %in% keep2))

|| contrast estimate SE df t.ratio p.value
|| High Complex - Low Complex -0.2801 0.546 67.5 -0.513 0.6098
|| High Complex - High Simple 0.4313 0.588 66.1 0.734 0.4655
|| Low Complex - Low Simple 0.7925 0.588 66.1 1.349 0.1820
|| High Simple - Low Simple 0.0811 0.546 67.5 0.148 0.8825
||
|| 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
|| High Complex - Low Complex -0.2801 0.546 67.5 -1.370 0.81
|| High Complex - High Simple 0.4313 0.588 66.1 -0.742 1.60
|| Low Complex - Low Simple 0.7925 0.588 66.1 -0.381 1.97
|| High Simple - Low Simple 0.0811 0.546 67.5 -1.009 1.17
||
|| 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 redundant rows
(effs2_filtered <- subset(effs2, contrast %in% keep2))

|| contrast effect.size SE df lower.CL upper.CL
|| High Complex - Low Complex -0.0822 0.160 59.4 -0.403 0.239
|| High Complex - High Simple 0.1266 0.173 59.5 -0.218 0.472
|| Low Complex - Low Simple 0.2327 0.173 59.4 -0.113 0.578
|| High Simple - Low Simple 0.0238 0.160 59.8 -0.297 0.345
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| sigma used for effect sizes: 3.406
|| Degrees-of-freedom method: inherited from kenward-roger when re-gridding
|| Confidence level used: 0.95
```

3.2.1 Interaction Contrasts

```
# Interaction contrasts (difference-of-differences)

contrast(emm2, interaction = "pairwise", by = NULL, adjust = "holm")

|| Base_Frequency_pairwise Complexity_pairwise estimate SE df t.ratio p.value
```

```

|| High - Low          Complex - Simple      -0.361 0.296 1498  -1.221  0.2223
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
confint(contrast(emm2, interaction = c("pairwise", "pairwise")))

|| Base_Frequency_pairwise Complexity_pairwise estimate   SE   df lower.CL upper.CL
|| High - Low          Complex - Simple      -0.361 0.296 1498  -0.941  0.219
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95
# Get confidence intervals, for each base frequency effect for each family size and then for interaction effect
confint(contrast(emmeans(m2, ~ Base_Frequency | Complexity), "pairwise"))

|| Complexity = Complex:
|| contrast estimate SE df lower.CL upper.CL
|| High - Low -0.2801 0.546 67.5 -1.37 0.81
||
|| Complexity = Simple:
|| contrast estimate SE df lower.CL upper.CL
|| High - Low 0.0811 0.546 67.5 -1.01 1.17
||
|| Results are averaged over the levels of: Orthographic_Sensitivity
|| Degrees-of-freedom method: kenward-roger
|| Confidence level used: 0.95

```

3.3 Plots

```

p1 <- emmip(anova_model_n250_nonwords, Base_Frequency ~ Complexity) + my_style
p2 <- emmip(anova_model_n250_nonwords, Complexity ~ Base_Frequency) + my_style
plot_grid(p1, p2, ncol = 2)

```

