M21 RT Semantic Sensitivity

Joanna Morris

2025-11-11

Setup

Load libraries

1. Set `ggplot2` parameters

Load Files and Format Files

Load Files

```
#DIR <- "csv_files"
df_a <- read_csv("rt_data_hc_A.csv")
df_b <- read_csv( "rt_data_hc_B_fixed.csv")
frq_w <- read_csv("frq_cw.csv")
frq_nw <- read_csv("frq_nw.csv")
dmg <- read_csv("demo_lang_vsl_pca_hc.csv")
```

Format Files

```
# Concatenate datasets
rt <- bind_rows(AB = df_a,
                BA = df_b,
               .id = "List")
rt_dmg<- right_join(dmg, rt, join_by(SubjID == subject_nr)) |>  # Join Participant Demographic and Lang Data
                    mutate(target = tolower(target)) |>
                    filter(correct == 1)

# Divide into Experimental and Filler Items
rt_fill <- rt_dmg |> filter(str_detect(targ_type, "^FILL"))
rt_exp <- rt_dmg |> filter(!str_detect(targ_type, "^FILL"))

# Add logFS to frequency datasets
frq_w <- frq_w |> mutate(Log10FS = log10(FS))
frq_nw <- frq_nw |> mutate(Log10FS = log10(FS))

# Define Factors and Conditions
rt_exp_format <- rt_exp |>
  separate(targ_type, into = c("trial_type", "family_size", "complexity"), sep = "_",
           remove = TRUE, extra = "drop", fill = "right")

# Divide into Words and Nonwords
rt_words <- rt_exp_format |> filter(trial_type == "CW") |> select(- complexity)
rt_nwords <- rt_exp_format |> filter(trial_type == "NW")

# Join Stimulus Frequency Data
rt_words_frq <- left_join(rt_words, frq_w, join_by(target))|>
  select(-cond_trig.y, -word_trig.y) |>
  rename(cond_trig = cond_trig.x, word_trig = word_trig.x)  # remove duplicate columns
rt_nwords_frq <-  left_join(rt_nwords, frq_nw, join_by(target==word)) |>
  select(-cond_trig.y, -word_trig.y) |>
  rename(cond_trig = cond_trig.x, word_trig = word_trig.x)

# Rename BF_Split and FS_Split columns
rt_words_frq <- rt_words_frq |> rename(Base_Frequency = BF_Split, Family_Size = FS_Split) # Rename BF_Split and FS_Split columns
rt_nwords_frq <- rt_nwords_frq |> rename(Base_Frequency = BF_Split, Family_Size = FS_Split)

# Recode factor levels
# rt_words_frq <- rt_words_frq |>
#   mutate(Base_Frequency = case_match(Base_Frequency, "Low" ~ "Low BF", "High" ~ "High BF"),
```

```r
#          Family_Size = case_match(Family_Size, "Small" ~ "Small Family", "Large" ~ "Large Family"))
# rt_nwords_frq <- rt_nwords_frq |> mutate(Base_Frequency = case_match(Base_Frequency, "Low" ~ "Low BF", "High" ~ "High BF"),
#                                          Family_Size = case_match(Family_Size, "Small" ~ "Small Family", "Large" ~ "Large Family"))
#
# rt_words_frq$Semantic_Sensitivity[rt_words_frq$Semantic_Sensitivity == "Low"] <- "Low Sensitivity"
# rt_words_frq$Semantic_Sensitivity[rt_words_frq$Semantic_Sensitivity == "High"] <- "High Sensitivity"
```

Word Data

Use `complete.cases()` to find which rows have missing data in the model-relevant variables:

```r
# Specify only the variables used in the model
model_vars_w <- c("response_time",  "Log10BF","BF", "FS","Family_Size", "Base_Frequency", "Semantic_Sensitivity","SubjID")

# Identify incomplete rows cohort 1
incomplete_cases_words <- rt_words_frq[!complete.cases(rt_words_frq[, model_vars_w]), ]
rt_words_cmpl <- rt_words_frq[complete.cases(rt_words_frq[, model_vars_w]), ]
# View them
print(incomplete_cases_words)
```

```r
# Standardize the predictors
rt_words_cmpl$Log10BF_std <- as.numeric(scale(rt_words_cmpl$Log10BF, center = TRUE, scale = TRUE))
rt_words_cmpl$FS_std <- as.numeric(scale(rt_words_cmpl$FS, center = TRUE, scale = TRUE))
rt_words_cmpl$Log10WF_std <- as.numeric(scale(rt_words_cmpl$Log10WF, center = TRUE, scale = TRUE))
rt_words_cmpl$Log10FS_std <- as.numeric(scale(rt_words_cmpl$Log10FS, center = TRUE, scale = TRUE))
rt_words_cmpl$Dim.2_std <- as.numeric(scale(rt_words_cmpl$Dim.2, center = TRUE, scale = TRUE))
```

Anova

```r
anova_model_words <- mixed(
  response_time ~ Base_Frequency * Family_Size * Semantic_Sensitivity +
    (1 + Base_Frequency + Family_Size | SubjID) +
    (1 | STRING),
  data = rt_words_cmpl,
  method = "S")
anova_model_words
```

```
|| Mixed Model Anova Table (Type 3 tests, S-method)
||
|| Model: response_time ~ Base_Frequency * Family_Size * Semantic_Sensitivity +
|| Model:     (1 + Base_Frequency + Family_Size | SubjID) + (1 | STRING)
|| Data: rt_words_cmpl
||                                            Effect         df        F p.value
|| 1                                  Base_Frequency   1, 93.30 10.16 **    .002
|| 2                                     Family_Size   1, 88.08  9.06 **    .003
|| 3                            Semantic_Sensitivity   1, 64.88     0.00    .990
|| 4                      Base_Frequency:Family_Size   1, 92.22     1.01    .318
|| 5             Base_Frequency:Semantic_Sensitivity  1, 673.76     0.37    .543
|| 6                Family_Size:Semantic_Sensitivity   1, 58.44     0.31    .579
|| 7 Base_Frequency:Family_Size:Semantic_Sensitivity 1, 5626.31     1.03    .310
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```r
m1 <- anova_model_words$full_model    # Extract the lmer model
ranova(m1) # formally test whether adding each random effect improves fit
```

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| response_time ~ Base_Frequency + Family_Size + Semantic_Sensitivity + (1 + Base_Frequency + Family_Size | SubjID) + (1 | STRING) + Base_Frequen
||                                                               npar logLik   AIC     LRT Df Pr(>Chisq)
|| <none>                                                          16 -35808 71649
|| Base_Frequency in (1 + Base_Frequency + Family_Size | SubjID)   13 -35810 71646   2.849  3     0.4155
|| Family_Size in (1 + Base_Frequency + Family_Size | SubjID)      13 -35809 71643   0.223  3     0.9738
|| (1 | STRING)                                                    15 -35898 71826 179.026  1     <2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                       | Eta2 (partial) |       95% CI
|| ------------------------------------------------------------------------------
|| Base_Frequency                                  |           0.10 | [0.02, 1.00]
|| Family_Size                                     |           0.09 | [0.02, 1.00]
|| Semantic_Sensitivity                            |       2.50e-06 | [0.00, 1.00]
```

```
|| Base_Frequency:Family_Size                      |           0.01 | [0.00, 1.00]
|| Base_Frequency:Semantic_Sensitivity             |       5.51e-04 | [0.00, 1.00]
|| Family_Size:Semantic_Sensitivity                |       5.30e-03 | [0.00, 1.00]
|| Base_Frequency:Family_Size:Semantic_Sensitivity |       1.83e-04 | [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_words)
```

```
|| Random effect variances not available. Returned R2 does not account for random effects.
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: NA
||      Marginal R2: 0.017
```

Main Effects

Effect	df	F	p.value
Base_Frequency	1, 93.73	10.08 **	.002
Family_Size	1, 92.52	9.00 **	.003

```r
emmeans(anova_model_words, ~ Family_Size)
```

Means

```
||  Family_Size emmean   SE  df asymp.LCL asymp.UCL
||  Large          602 10.0 Inf       582       621
||  Small          620 10.1 Inf       600       640
||
|| Results are averaged over the levels of: Base_Frequency, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```r
emmeans(anova_model_words, ~ Base_Frequency)
```

```
||  Base_Frequency emmean    SE  df asymp.LCL asymp.UCL
||  High              601  9.83 Inf       582       620
||  Low               621 10.30 Inf       600       641
||
|| Results are averaged over the levels of: Family_Size, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

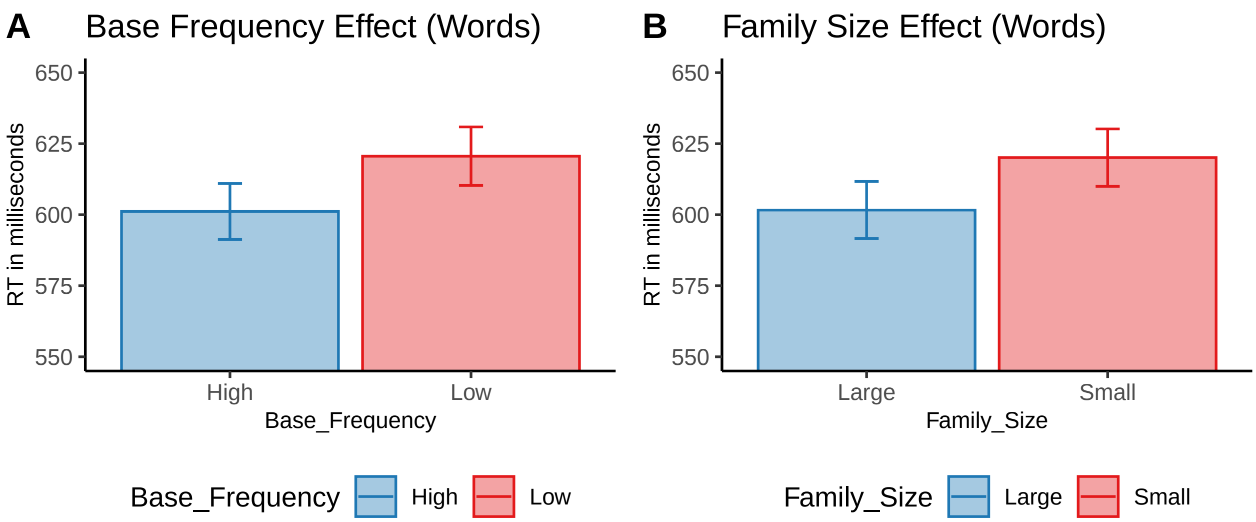
```r
emmeans(anova_model_words, ~ Semantic_Sensitivity)
```

```
||  Semantic_Sensitivity emmean   SE  df asymp.LCL asymp.UCL
||  High                    611 13.2 Inf       585       637
||  Low                     611 13.4 Inf       585       637
||
|| Results are averaged over the levels of: Base_Frequency, Family_Size
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

Plots

```
||  Base_Frequency   emmean        SE  df asymp.LCL asymp.UCL
||  High           601.1459  9.830243 Inf   581.879  620.4128
||  Low            620.6046 10.303517 Inf   600.410  640.7991
||
|| Results are averaged over the levels of: Family_Size, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```
||  Family_Size   emmean       SE  df asymp.LCL asymp.UCL
||  Large       601.6482 10.04806 Inf  581.9543  621.3420
||  Small       620.1023 10.09854 Inf  600.3095  639.8951
||
|| Results are averaged over the levels of: Base_Frequency, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```


A
Base Frequency Effect (Words)
RT in milliseconds
650
625
600
575
550
High
Low
Base_Frequency
Base_Frequency
High
Low
B
Family Size Effect (Words)
RT in milliseconds
650
625
600
575
550
Large
Small
Family_Size
Family_Size
Large
Small

Non-word Data

Use `complete.cases()` to find which rows had missing data in the model-relevant variables:

```
# Specify only the variables used in the model
# model_vars <- c("response_time", "Dim.2","SubjID")
model_vars_nw <- c("response_time","Complexity", "Family_Size", "Base_Frequency",
                   "SubjID", "ItemID", "Semantic_Sensitivity")

# Identify incomplete rows
incomplete_cases_nwords <- rt_nwords[!complete.cases(rt_nwords_frq[, model_vars_nw]), ]
rt_nwords_cmpl <- rt_nwords_frq[complete.cases(rt_nwords_frq[, model_vars_nw]), ]
# View them
print(incomplete_cases_nwords)

# str(rt_nwords_1_cmpl)
```

Standardize the predictors

```
rt_nwords_cmpl$LogBF_std <- as.numeric(scale(rt_nwords_cmpl$LogBF, center = TRUE, scale = TRUE))
rt_nwords_cmpl$FS_std <- as.numeric(scale(rt_nwords_cmpl$FS, center = TRUE, scale = TRUE))
rt_nwords_cmpl$BF_std <- as.numeric(scale(rt_nwords_cmpl$BF, center = TRUE, scale = TRUE))
rt_nwords_cmpl$Dim.2_std <- as.numeric(scale(rt_nwords_cmpl$Dim.2, center = TRUE, scale = TRUE))
```

Anova Family Size

```
anova_model_nwords_fs <- mixed(
  response_time ~ Complexity * Family_Size * Semantic_Sensitivity +
    (1 + Complexity + Family_Size | SubjID) +
    (1 | ItemID),
  data = rt_nwords_cmpl,
  method = "S")
anova_model_nwords_fs
```

```
|| Mixed Model Anova Table (Type 3 tests, S-method)
||
|| Model: response_time ~ Complexity * Family_Size * Semantic_Sensitivity +
|| Model:     (1 + Complexity + Family_Size | SubjID) + (1 | ItemID)
|| Data: rt_nwords_cmpl
||                                       Effect         df         F p.value
|| 1                                 Complexity   1, 62.65 88.28 ***   <.001
|| 2                                Family_Size   1, 94.52      0.91    .344
|| 3                       Semantic_Sensitivity   1, 63.40      0.00    .957
|| 4                     Complexity:Family_Size 1, 4471.37      0.46    .498
|| 5            Complexity:Semantic_Sensitivity   1, 59.60      0.14    .710
|| 6           Family_Size:Semantic_Sensitivity   1, 67.28      0.09    .760
|| 7 Complexity:Family_Size:Semantic_Sensitivity 1, 4387.63    4.73 *    .030
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

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

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| response_time ~ Complexity + Family_Size + Semantic_Sensitivity + (1 + Complexity + Family_Size | SubjID) + (1 | ItemID) + Complexity:Family_Si
||                                                          npar logLik   AIC     LRT Df Pr(>Chisq)
|| <none>                                                     16 -28031 56094
|| Complexity in (1 + Complexity + Family_Size | SubjID)     13 -28033 56092   3.489  3     0.3222
|| Family_Size in (1 + Complexity + Family_Size | SubjID)    13 -28032 56090   2.111  3     0.5497
|| (1 | ItemID)                                               15 -28103 56236 143.964  1     <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_nwords_fs, partial = TRUE)
```

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                   | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------------
|| Complexity                                  |           0.58 | [0.45, 1.00]
|| Family_Size                                 |       9.49e-03 | [0.00, 1.00]
|| Semantic_Sensitivity                        |       4.73e-05 | [0.00, 1.00]
|| Complexity:Family_Size                      |       1.03e-04 | [0.00, 1.00]
|| Complexity:Semantic_Sensitivity             |       2.33e-03 | [0.00, 1.00]
|| Family_Size:Semantic_Sensitivity            |       1.39e-03 | [0.00, 1.00]
|| Complexity:Family_Size:Semantic_Sensitivity |       1.08e-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_nwords_fs)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.462
||      Marginal R2: 0.016
```

Main Effects

Effect	df	F	p.value
Complexity	1, 62.25	87.32 ***	<.001

```r
emmeans(anova_model_nwords_fs, ~ Complexity)
```

Main Effects Means

```
||  Complexity emmean   SE  df asymp.LCL asymp.UCL
||  Complex       734 11.5 Inf       712       757
||  Simple        700 11.3 Inf       678       722
||
|| Results are averaged over the levels of: Family_Size, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

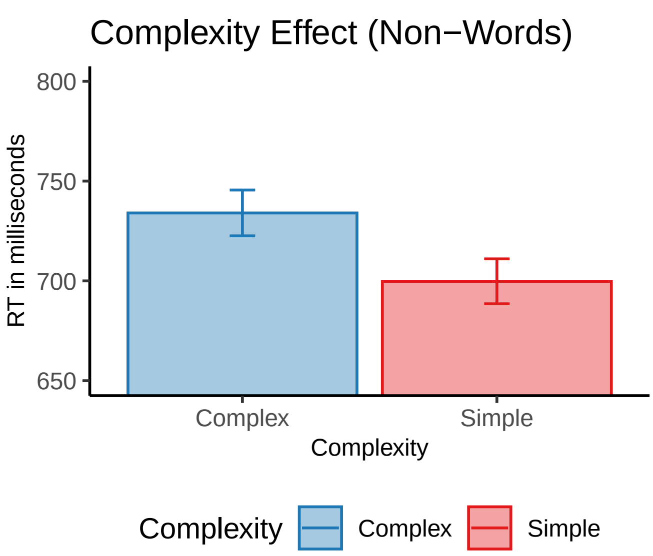
```r
emmeans(anova_model_nwords_fs, ~ Family_Size)
```

```
||  Family_Size emmean   SE  df asymp.LCL asymp.UCL
||  Large          720 11.6 Inf       697       743
||  Small          714 11.8 Inf       691       737
||
|| Results are averaged over the levels of: Complexity, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```r
emmeans(anova_model_nwords_fs, ~ Semantic_Sensitivity)
```

```
||  Semantic_Sensitivity emmean   SE  df asymp.LCL asymp.UCL
||  High                    716 15.5 Inf       686       747
||  Low                     717 15.8 Inf       687       748
||
|| Results are averaged over the levels of: Complexity, Family_Size
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

Main Effects Plots ...



Interaction Effects

Effect	df	F	p.value
Complexity:Family__Size:Semantic__Sensitivity	1, 198.28	4.67 *	.032

Simple Contrasts

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

This gives you: 4 contrasts: one for each `Family Size × Complexity` combination. Each shows whether `High` vs `Low Semantic Sensitivity` differs significantly

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?

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

```
||  Semantic_Sensitivity Family_Size Complexity emmean   SE  df asymp.LCL asymp.UCL
||  High                 Large       Complex       742 16.2 Inf       710       774
||  Low                  Large       Complex       734 16.7 Inf       702       767
||  High                 Small       Complex       726 16.5 Inf       694       759
||  Low                  Small       Complex       734 16.9 Inf       701       767
||  High                 Large       Simple        698 15.9 Inf       667       729
||  Low                  Large       Simple        706 16.2 Inf       674       738
||  High                 Small       Simple        699 16.2 Inf       668       731
||  Low                  Small       Simple        696 16.5 Inf       664       729
||
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```
# Get all pairswise 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 Large Complex - High Large Simple",
          "High Small Complex - High Small Simple",
          "Low Large Complex - Low Large Simple",
          "Low Small Complex - Low Small Simple",
          "High Large Complex - High Small Complex",
          "High Large Simple - High Small Simple",
          "Low Large Complex - Low Small Complex",
          "Low Large Simple - Low Small Simple",
          "High Large Complex - Low Large Complex",
          "High Small Complex - Low Small Complex",
          "High Large Simple - Low Small Simple",
          "High Small Simple - Low Small Simple")

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

```
||  contrast                                estimate    SE  df z.ratio p.value
||  High Large Complex - Low Large Complex     7.862 22.60 Inf   0.348  0.7282
||  High Large Complex - High Small Complex   15.914  8.51 Inf   1.870  0.0615
||  High Large Complex - High Large Simple    44.380  6.50 Inf   6.828  <.0001
||  Low Large Complex - Low Small Complex      0.495  9.10 Inf   0.054  0.9566
||  Low Large Complex - Low Large Simple      28.375  7.16 Inf   3.965  0.0001
||  High Small Complex - Low Small Complex    -7.557 23.00 Inf  -0.328  0.7427
||  High Small Complex - High Small Simple    26.839  6.34 Inf   4.233  <.0001
||  Low Small Complex - Low Small Simple      37.468  6.90 Inf   5.429  <.0001
||  High Large Simple - High Small Simple     -1.627  7.95 Inf  -0.205  0.8378
||  High Large Simple - Low Small Simple       1.444 22.90 Inf   0.063  0.9497
||  Low Large Simple - Low Small Simple        9.587  8.30 Inf   1.155  0.2481
||  High Small Simple - Low Small Simple       3.071 22.50 Inf   0.137  0.8912
||
|| Degrees-of-freedom method: asymptotic
```

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

```
||  contrast                                estimate    SE  df asymp.LCL asymp.UCL
||  High Large Complex - Low Large Complex     7.862 22.60 Inf   -36.469      52.2
||  High Large Complex - High Small Complex   15.914  8.51 Inf    -0.768      32.6
||  High Large Complex - High Large Simple    44.380  6.50 Inf    31.642      57.1
||  Low Large Complex - Low Small Complex      0.495  9.10 Inf   -17.338      18.3
||  Low Large Complex - Low Large Simple      28.375  7.16 Inf    14.348      42.4
```

```
|| High Small Complex - Low Small Complex    -7.557 23.00 Inf   -52.684      37.6
|| High Small Complex - High Small Simple    26.839  6.34 Inf    14.412      39.3
|| Low Small Complex - Low Small Simple      37.468  6.90 Inf    23.942      51.0
|| High Large Simple - High Small Simple     -1.627  7.95 Inf   -17.202      13.9
|| High Large Simple - Low Small Simple       1.444 22.90 Inf   -43.452      46.3
|| Low Large Simple - Low Small Simple        9.587  8.30 Inf    -6.683      25.9
|| High Small Simple - Low Small Simple       3.071 22.50 Inf   -40.939      47.1
||
|| Degrees-of-freedom method: asymptotic
|| 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 asymp.LCL asymp.UCL
||  High Large Complex - Low Large Complex      0.07728 0.2220 Inf  -0.35849     0.513
||  High Large Complex - High Small Complex     0.15643 0.0837 Inf  -0.00758     0.320
||  High Large Complex - High Large Simple      0.43625 0.0640 Inf   0.31072     0.562
||  Low Large Complex - Low Small Complex       0.00486 0.0894 Inf  -0.17044     0.180
||  Low Large Complex - Low Large Simple        0.27893 0.0704 Inf   0.14092     0.417
||  High Small Complex - Low Small Complex     -0.07429 0.2260 Inf  -0.51789     0.369
||  High Small Complex - High Small Simple      0.26383 0.0624 Inf   0.14155     0.386
||  Low Small Complex - Low Small Simple        0.36831 0.0679 Inf   0.23514     0.501
||  High Large Simple - High Small Simple      -0.01599 0.0781 Inf  -0.16910     0.137
||  High Large Simple - Low Small Simple        0.01420 0.2250 Inf  -0.42713     0.456
||  Low Large Simple - Low Small Simple         0.09424 0.0816 Inf  -0.06571     0.254
||  High Small Simple - Low Small Simple        0.03019 0.2210 Inf  -0.40243     0.463
||
|| sigma used for effect sizes: 101.7
|| Degrees-of-freedom method: inherited from asymptotic when re-gridding
|| Confidence level used: 0.95
```

Interaction Contrasts

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

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

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

```
||  Semantic_Sensitivity_pairwise Family_Size_pairwise Complexity_pairwise estimate   SE  df z.ratio p.value
||  High - Low                    Large - Small        Complex - Simple        26.6 12.2 Inf   2.175  0.0296
||
|| Degrees-of-freedom method: asymptotic
```

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

```
||  Semantic_Sensitivity_pairwise Family_Size_pairwise Complexity_pairwise estimate   SE  df asymp.LCL asymp.UCL
||  High - Low                    Large - Small        Complex - Simple        26.6 12.2 Inf      2.63      50.6
||
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```
# Compute the A1 - A2 difference within each combination of B × C
(complexity_diff <- contrast(emm2, method = "revpairwise",
                            by = c("Semantic_Sensitivity", "Family_Size"),
                            simple = "Complexity"))
```

```
|| Semantic_Sensitivity = High, Family_Size = Large:
||  contrast         estimate   SE  df z.ratio p.value
||  Simple - Complex    -44.4 6.50 Inf  -6.828  <.0001
||
|| Semantic_Sensitivity = Low, Family_Size = Large:
||  contrast         estimate   SE  df z.ratio p.value
||  Simple - Complex    -28.4 7.16 Inf  -3.965  0.0001
||
|| Semantic_Sensitivity = High, Family_Size = Small:
||  contrast         estimate   SE  df z.ratio p.value
||  Simple - Complex    -26.8 6.34 Inf  -4.233  <.0001
||
|| Semantic_Sensitivity = Low, Family_Size = Small:
||  contrast         estimate   SE  df z.ratio p.value
```

```
||  Simple - Complex    -37.5 6.90 Inf  -5.429  <.0001
||
|| Degrees-of-freedom method: asymptotic
```

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

```
|| contrast = Simple - Complex, Semantic_Sensitivity = High:
||  contrast1     estimate   SE  df z.ratio p.value
||  Small - Large    17.54 8.29 Inf   2.116  0.0344
||
|| contrast = Simple - Complex, Semantic_Sensitivity = Low:
||  contrast1     estimate   SE  df z.ratio p.value
||  Small - Large    -9.09 9.15 Inf  -0.993  0.3206
||
|| Degrees-of-freedom method: asymptotic
```

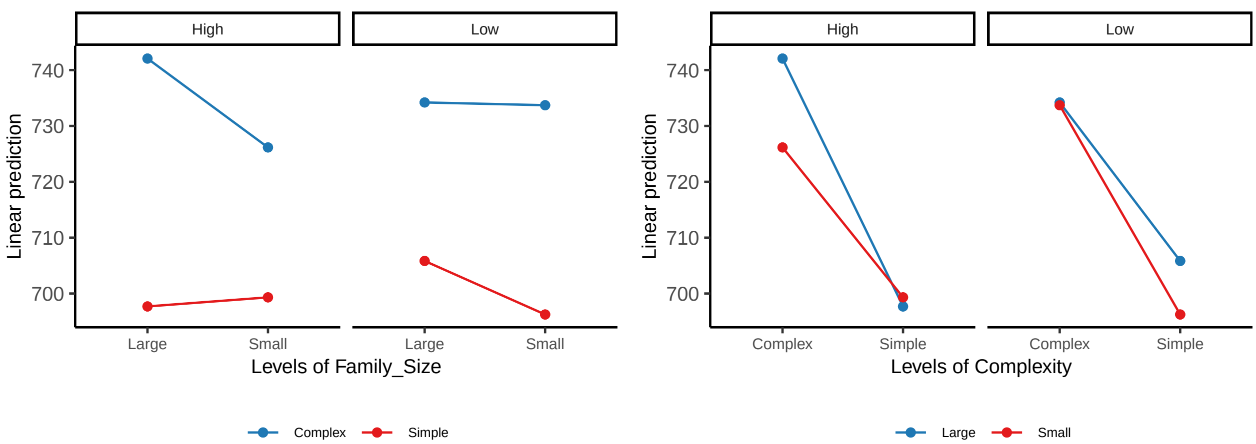
```r
# Get confidence intervals
confint(family_size_complexity_int_within_sensitivity)
```

```
|| contrast = Simple - Complex, Semantic_Sensitivity = High:
||  contrast1     estimate   SE  df asymp.LCL asymp.UCL
||  Small - Large    17.54 8.29 Inf      1.29     33.79
||
|| contrast = Simple - Complex, Semantic_Sensitivity = Low:
||  contrast1     estimate   SE  df asymp.LCL asymp.UCL
||  Small - Large    -9.09 9.15 Inf    -27.03      8.85
||
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

Interaction Plots

```r
p4 <- emmip(anova_model_nwords_fs, Complexity ~ Family_Size | Semantic_Sensitivity) + my_style
p5 <- emmip(anova_model_nwords_fs, Family_Size ~ Complexity | Semantic_Sensitivity) + my_style

plot_grid(p4, p5, ncol = 2)
```



Responses were slower to *Complex* than *Simple* nonwords in every condition. The *Complexity effect* (`Complex - Simple`) varied with both *Family Size* and *Semantic Sensitivity*.

The Complexity effect (slower responses for complex vs. simple nonwords) is robust across all groups.

However, its magnitude varies:

- Among **high-sensitivity participants**, the effect is larger for large families ($\approx$ 45 ms) than small families ($\approx$ 27 ms).
- Among **low-sensitivity participants**, the pattern reverses slightly ($\approx$ 29 ms vs. 38 ms).

The difference in the Complexity $\times$ Family Size interaction between high- and low-sensitivity participants is about 27 ms.

- **High-sensitivity participants** showed a stronger complexity effect for large-family nonwords than for small-family ones.
- **Low-sensitivity participants** showed the opposite or no difference.

This indicates that semantic sensitivity modulates how morphological family size influences the cost of morphological complexity in nonword processing.

All groups show reliable complexity effects (complex slower than simple). Only one cross-condition difference is significant: High-sensitivity participants respond faster to complex nonwords from small families than to complex nonwords from large-families.

Summary interpretation (for Results section):

Response times to morphologically complex nonwords were significantly slower than to simple nonwords, indicating greater processing cost for complexity. While overall family size and semantic sensitivity did not produce main effects, there was a significant `Complexity × Family Size × Semantic Sensitivity` interaction ($p = .028$).

Follow-up contrasts showed that for participants with high semantic sensitivity, the complexity effect was larger for large-family nonwords (≈ 45 ms) than for small-family nonwords (≈ 27 ms). In contrast, participants with low semantic sensitivity showed little difference or the reverse pattern. This suggests that individuals with greater semantic knowledge are more sensitive to morphological family size cues when processing novel morphological structures, showing amplified complexity costs when nonwords resemble rich morphological families.

Anova Base Frequency

```r
anova_model_nwords_bf <- mixed(
  response_time ~ Complexity * Base_Frequency * Semantic_Sensitivity +
    (1 + Base_Frequency + Complexity | SubjID) +
    (1 | ItemID),
  data = rt_nwords_cmpl,
  method = "S")
anova_model_nwords_bf
```

```
|| Mixed Model Anova Table (Type 3 tests, S-method)
||
|| Model: response_time ~ Complexity * Base_Frequency * Semantic_Sensitivity +
|| Model:     (1 + Base_Frequency + Complexity | SubjID) + (1 | ItemID)
|| Data: rt_nwords_cmpl
||                                          Effect         df         F p.value
|| 1                                    Complexity   1, 62.89 90.74 ***   <.001
|| 2                                Base_Frequency   1, 90.71  11.45 **    .001
|| 3                          Semantic_Sensitivity   1, 63.36      0.00    .970
|| 4                     Complexity:Base_Frequency 1, 4491.15    3.56 +    .059
|| 5               Complexity:Semantic_Sensitivity   1, 59.79      0.23    .633
|| 6           Base_Frequency:Semantic_Sensitivity   1, 67.26      1.10    .297
|| 7 Complexity:Base_Frequency:Semantic_Sensitivity 1, 4394.20      2.69    .101
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

```r
m3 <- anova_model_nwords_bf$full_model    # Extract the lmer model
ranova(m3) # Run random effects comparison
```

```
|| ANOVA-like table for random-effects: Single term deletions
||
|| Model:
|| response_time ~ Complexity + Base_Frequency + Semantic_Sensitivity + (1 + Base_Frequency + Complexity | SubjID) + (1 | ItemID) + Complexity:Bas
||                                                            npar logLik   AIC     LRT Df Pr(>Chisq)
|| <none>                                                       16 -28024 56080
|| Base_Frequency in (1 + Base_Frequency + Complexity | SubjID)   13 -28026 56078   3.402  3     0.3336
|| Complexity in (1 + Base_Frequency + Complexity | SubjID)       13 -28026 56078   3.937  3     0.2683
|| (1 | ItemID)                                                 15 -28086 56201 122.935  1     <2e-16 ***
|| ---
|| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
|| # Effect Size for ANOVA (Type III)
||
|| Parameter                                      | Eta2 (partial) |       95% CI
|| -----------------------------------------------------------------------------
|| Complexity                                     |           0.59 | [0.46, 1.00]
|| Base_Frequency                                 |           0.11 | [0.03, 1.00]
|| Semantic_Sensitivity                           |       2.20e-05 | [0.00, 1.00]
|| Complexity:Base_Frequency                      |       7.93e-04 | [0.00, 1.00]
|| Complexity:Semantic_Sensitivity                |       3.84e-03 | [0.00, 1.00]
|| Base_Frequency:Semantic_Sensitivity            |           0.02 | [0.00, 1.00]
|| Complexity:Base_Frequency:Semantic_Sensitivity |       6.13e-04 | [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_nwords_bf)
```

```
|| # R2 for Mixed Models
||
||   Conditional R2: 0.462
||      Marginal R2: 0.021
```

Main Effects

Effect	df	F	p.value
Complexity	1, 61.44	88.55 ***	<.001
Base_Frequency	1, 91.73	11.22 **	.001

Participants responded more slowly to complex nonwords and to low-frequency-base nonwords.

```r
emmeans(anova_model_nwords_bf, ~ Complexity)
```

Means

```
||  Complexity emmean   SE  df asymp.LCL asymp.UCL
||  Complex       734 11.4 Inf       712       757
||  Simple        699 11.2 Inf       677       721
||
|| Results are averaged over the levels of: Base_Frequency, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```r
emmeans(anova_model_nwords_bf, ~ Base_Frequency)
```

```
||  Base_Frequency emmean   SE  df asymp.LCL asymp.UCL
||  High              727 11.3 Inf       705       749
||  Low               707 11.9 Inf       683       730
||
|| Results are averaged over the levels of: Complexity, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```r
emmeans(anova_model_nwords_bf, ~ Semantic_Sensitivity)
```

```
||  Semantic_Sensitivity emmean   SE  df asymp.LCL asymp.UCL
||  High                    716 15.4 Inf       686       747
||  Low                     717 15.7 Inf       686       748
||
|| Results are averaged over the levels of: Complexity, Base_Frequency
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

Interaction Effects: `Complexity x Base_Frequency`

Effect	df	F	p.value
Complexity:Base_Frequency	1, 545.66	3.14 *	077

Simple Contrasts

```r
# Estimated marginal means for the family_size × base frequency interaction
(emm1 <- emmeans(anova_model_nwords_bf, ~ Complexity * Base_Frequency))
```

```
||  Complexity Base_Frequency emmean   SE  df asymp.LCL asymp.UCL
||  Complex    High              747 11.7 Inf       724       770
||  Simple     High              707 11.3 Inf       684       729
||  Complex    Low               721 12.2 Inf       697       745
||  Simple     Low               692 12.0 Inf       669       716
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```r
# Get all pairswise contrasts
emm1_contrasts <- contrast(emm1,  method = "pairwise", by = NULL, adjust = "none")
emm1_contrasts
```

```
||  contrast                   estimate   SE  df z.ratio p.value
||  Complex High - Simple High     40.7 5.00 Inf   8.144  <.0001
||  Complex High - Complex Low     26.4 7.04 Inf   3.751  0.0002
||  Complex High - Simple Low      55.3 7.26 Inf   7.615  <.0001
||  Simple High - Complex Low     -14.3 6.88 Inf  -2.083  0.0372
||  Simple High - Simple Low       14.6 6.59 Inf   2.209  0.0272
||  Complex Low - Simple Low       28.9 4.63 Inf   6.242  <.0001
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
```

```r
# Keep only the contrasts you want
# Simple effects of Complexity at each level of Base_Frequency
# Simple effects of Base_Frequency at each level of Complexity
keep <- c("Complex High - Simple High",
          "Complex Low - Simple Low",
          "Complex High - Complex Low",
          "Simple High - Simple Low")
(emm1_contrasts_filtered <- subset(emm1_contrasts, contrast %in% keep))
```

```
||  contrast                    estimate   SE  df z.ratio p.value
||  Complex High - Simple High      40.7 5.00 Inf   8.144  <.0001
||  Complex High - Complex Low      26.4 7.04 Inf   3.751  0.0002
||  Simple High - Simple Low        14.6 6.59 Inf   2.209  0.0272
||  Complex Low - Simple Low        28.9 4.63 Inf   6.242  <.0001
||
|| Results are averaged over the levels of: Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
```

A marginal Complexity × Base Frequency interaction (p = .048) suggests that the complexity cost was smaller for nonwords derived from low-frequency bases.

Complexity	Base Frequency	Mean RT (ms)	Interpretation
Complex	High	748	slowest
Simple	High	707	41 ms faster
Complex	Low	721	28 ms slower than Simple Low
Simple	Low	692	fastest

Both complexity and base frequency affect RTs additively, but their combination reveals that high-frequency bases magnify the complexity cost.

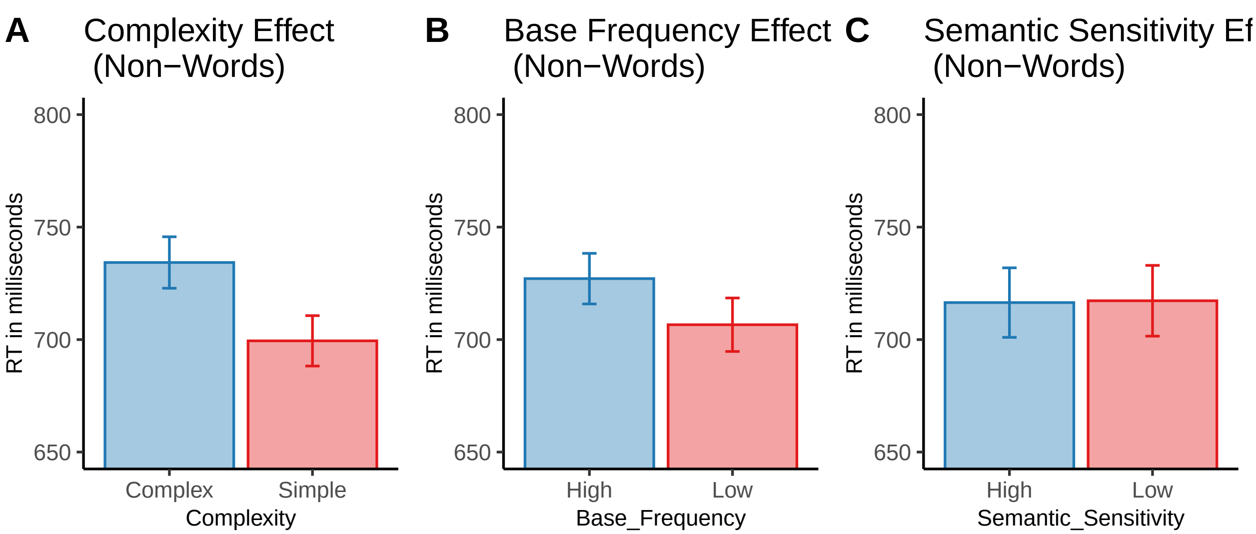
- The complexity effect (`Complex - Simple`) is larger for *high-frequency* bases (41 ms) than for *low-frequency* ones (29 ms).
- The base-frequency advantage (`High - Low`) is stronger for *complex* items (27 ms) than for *simple* ones (15 ms).

Main Effects Plots

```
||  Complexity   emmean       SE  df asymp.LCL asymp.UCL
||  Complex    734.2518 11.43605 Inf  711.8376  756.6661
||  Simple     699.4390 11.19614 Inf  677.4949  721.3830
||
|| Results are averaged over the levels of: Base_Frequency, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```
||  Base_Frequency   emmean       SE  df asymp.LCL asymp.UCL
||  High           727.0885 11.25292 Inf  705.0332  749.1438
||  Low            706.6023 11.88104 Inf  683.3159  729.8888
||
|| Results are averaged over the levels of: Complexity, Semantic_Sensitivity
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```

```
||  Semantic_Sensitivity   emmean       SE  df asymp.LCL asymp.UCL
||  High                 716.4391 15.44970 Inf  686.1582  746.7199
||  Low                  717.2518 15.72849 Inf  686.4245  748.0790
||
|| Results are averaged over the levels of: Complexity, Base_Frequency
|| Degrees-of-freedom method: asymptotic
|| Confidence level used: 0.95
```



Interaction Plots

```
p9 <- emmip(anova_model_nwords_bf, Complexity ~ Base_Frequency) + my_style
p10 <- emmip(anova_model_nwords_bf, Base_Frequency ~ Complexity) + my_style

plot_grid(p9, p10, ncol = 2)
```