m21_202303

Joanna Morris

2023-03-18

Compute PCA

This script computes separate ANOVAs for simple and complex non-words.

Following Andrews and Lo (2013) this script computes a PCA for our spelling and vocabulary measures. Because the standardised spelling and vocabulary scores were correlated, to facilitate interpretation, two orthogonal measures of individual differences were derived from a principal components analysis. Analysis based on this tutorial

```
library(readr)
library(dplyr)
library(datawizard)
sv_202303 <- read_csv("m21_spell_vocab_raw.csv")
sv_202303.na <- na.omit(sv_202303)
sv_202303.na <- mutate(sv_202303.na,
                       z_ART = standardise(ART_correct),
                       z_vocab = standardise(vocab_correct),
                       z_spell = standardise(spell_correct))

cor.test(sv_202303.na$z_vocab, sv_202303.na$z_spell)
```

```

    Pearson's product-moment correlation

data:  sv_202303.na$z_vocab and sv_202303.na$z_spell
t = 1.9352, df = 61, p-value = 0.05761
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.007724833  0.460807138
sample estimates:
      cor
0.2405005
```

By default, the function `PCA()` in `FactoMineR`, standardizes the data automatically during the PCA; so you don't need do this transformation before the PCA.

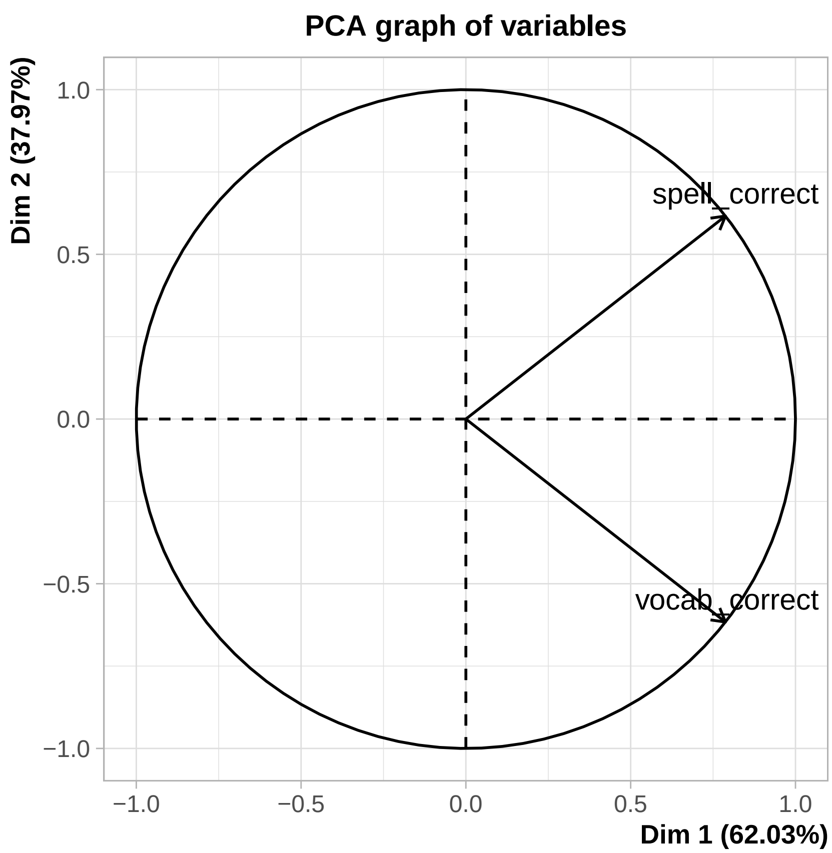
- `X`: a data frame. Rows are individuals and columns are numeric variables
- `scale.unit`: a logical value. If TRUE, the data are scaled to unit variance before the analysis. This standardization to the same scale avoids some variables to become dominant just because of their large measurement units. It makes variable comparable.

- `ncp`: number of dimensions kept in the final results.
- `graph`: a logical value. If TRUE a graph is displayed.

The plot shows the relationships between all variables. It can be interpreted as follow:

- Positively correlated variables are grouped together.
- Negatively correlated variables are positioned on opposite sides of the plot origin (opposed quadrants).
- The distance between variables and the origin measures the quality of the variables on the factor map. Variables that are away from the origin are well represented on the factor map.

```
library(FactoMineR)
library(factoextra)

res.pca <- PCA(sv_202303.na[,3:4], scale.unit = TRUE, ncp = 2, graph = FALSE)
plot(res.pca, choix = "varcor", graph.type = c("ggplot"))
```



The eigenvalues measure the amount of variation retained by each principal component. Eigenvalues are large for the first PCs and small for the subsequent PCs. That is, the first PCs corresponds to the directions with the maximum amount of variation in the data set.

We examine the eigenvalues to determine the number of principal components to be considered

```
(eig.val <- get_eigenvalue(res.pca))
```

```
      eigenvalue variance.percent cumulative.variance.percent
Dim.1  1.2405005         62.02503                    62.02503
Dim.2  0.7594995         37.97497                   100.00000
```

The quality of representation of the variables on factor map is called cos2 (square cosine, squared coordinates). A high cos2 indicates a good representation of the variable on the principal component. In this case the variable is positioned close to the circumference of the correlation circle. A low cos2 indicates that the variable is not perfectly represented by the PCs. In this case the variable is close to the center of the circle. For a given variable, the sum of the cos2 on all the principal components is equal to one. If a variable is perfectly represented by only two principal components (Dim.1 & Dim.2), the sum of the cos2 on these two PCs is equal to one. In this case the variables will be positioned on the circle of correlations.

```
res.pca$var$cos2
```

```
                 Dim.1     Dim.2
vocab_correct 0.6202503 0.3797497
spell_correct 0.6202503 0.3797497
```

The contributions of variables in accounting for the variability in a given principal component are expressed in percentage. Variables that are correlated with PC1 (i.e., Dim.1) and PC2 (i.e., Dim.2) are the most important in explaining the variability in the data set. Variables that do not correlated with any PC or correlated with the last dimensions are variables with low contribution and might be removed to simplify the overall analysis.

```
res.pca$var$contrib
```

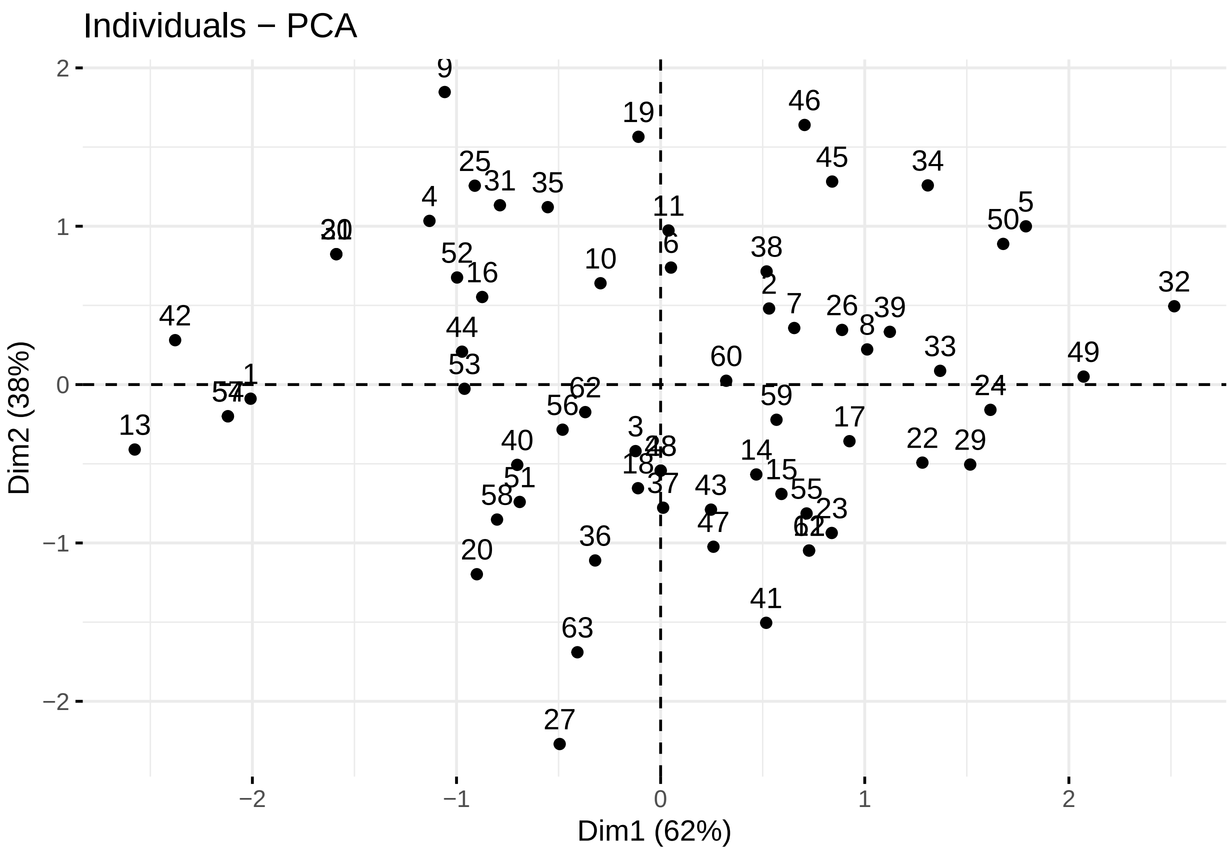
```
              Dim.1 Dim.2
vocab_correct    50    50
spell_correct    50    50
```

```
(res.desc <- dimdesc(res.pca, axes = c(1,2), proba = 0.05))
```

```
$Dim.1

Link between the variable and the continuous variables (R-square)
=================================================================================
              correlation      p.value
vocab_correct   0.7875597 1.913196e-14
spell_correct   0.7875597 1.913196e-14

$Dim.2

Link between the variable and the continuous variables (R-square)
=================================================================================
              correlation      p.value
spell_correct   0.6162384 7.592937e-08
vocab_correct  -0.6162384 7.592937e-08
```

The fviz_pca_ind() is used to produce the graph of individuals.

```
fviz_pca_ind(res.pca)
```




```
sv_202303.na<-bind_cols(sv_202303.na,res.pca$ind$coord)
```

Loads RT data and join to PCA dataset

```
cw_frq <- read_csv("CW_frq.csv")
nw_frq <- read_csv("NW_frq.csv")


CW_rt <- read_csv("CW_rt_2.csv")
CW_rt$cw_target <- NULL
CW_rt <- rename(CW_rt, cw_target = target_lower)

NW_rt <- read_csv("NW_rt_2.csv")
NW_rt$nw_target <- NULL
NW_rt <- rename(NW_rt, nw_target = target_lower)

cw_rt_pca <- inner_join(sv_202303.na, CW_rt, by = "SubjID")  #join subject PCA data
nw_rt_pca <- inner_join(sv_202303.na, NW_rt, by = "SubjID")

cw <- left_join(cw_rt_pca, cw_frq, by = c("cw_target")) #join word frequency data
nw <- left_join(nw_rt_pca, nw_frq, by = c("nw_target"))
```

Divide participants based on median split of Dim2. Higher values on this factor indicate that spelling scores were relatively higher than vocabulary,

```
cw.median <- median(cw$Dim.2)
cw <- cw |>
    mutate(lang_type = case_when(
      Dim.2 <= cw.median ~ "Semantic",
      Dim.2 > cw.median ~ "Orthographic"
    ))

nw.median <- median(nw$Dim.2)
nw <- nw |>
    mutate(lang_type = case_when(
      Dim.2 <= nw.median ~ "Semantic",
      Dim.2 > nw.median ~ "Orthographic"
    ))
```

```
cols <- c( "cw_famsize", "lang_type")  # recode ind variable columns as factors
cw <- cw |> mutate_at(cols, factor)
cw$cw_famsize <- recode_factor(cw$cw_famsize, S = "Small", L = "Large")


cols <- c( "cw_famsize", "lang_type")
cw <- cw |> mutate_at(cols, factor)
cw$cw_famsize <- recode_factor(cw$cw_famsize, S = "Small", L = "Large")


cols <- c( "nw_famsize", "lang_type", "complexity")
nw <- nw |> mutate_at(cols, factor)
nw$nw_famsize <- recode_factor(nw$nw_famsize, S = "Small", L = "Large")
nw$complexity <- recode_factor(nw$complexity, SIMP = "Simple", COMP = "Complex")


nw_smpl <- filter(nw, complexity == "Simple")
nw_smpl$complexity <- NULL

nw_cplx <- filter(nw, complexity == "Complex")
nw_cplx$complexity <- NULL

rm(CW_rt)  #remove original rt file after joining neuropsych data
rm(NW_rt)
```

Removes rts for errors (column rt.err) and then imputes missing values with the mean for the dataset (column "rt.err.imp") then creates a new column with inverse RTs

```
library(tidyr)
cw <- cw |> mutate(rt.err = response_time * correct)   # convert error rts to 0
cw <- cw |> mutate(rt.err = na_if(rt.err, 0))          # convert 0 rts to NA
cw.mean <- mean(cw$rt.err, na.rm = TRUE)  # get mean rt excluding errors
cw <- cw |> mutate(rt.err.imp = ifelse(is.na(rt.err),
                                       cw.mean,
                                       rt.err))  # replace missing values with mean
cw <- cw |> mutate(inv.rt = 1/rt.err.imp)  # creates new column with inverse RTs


nw_smpl <- nw_smpl |> mutate(rt.err = response_time * correct)  # convert error rts to 0
```

```
nw_smpl <- nw_smpl |> mutate(rt.err = na_if(rt.err, 0))         # convert 0 rts to NA
nw_smpl.mean <- mean(nw_smpl$rt.err, na.rm = TRUE)  # get mean rt excluding errors
nw_smpl <- nw_smpl |> mutate(rt.err.imp = ifelse(is.na(rt.err),
                                                 nw_smpl.mean,
                                                 rt.err))  # replace missing values with mean
nw_smpl <- nw_smpl |> mutate(inv.rt = 1/rt.err.imp)  # creates new column with inverse RTs

nw_cplx <- nw_cplx |> mutate(rt.err = response_time * correct)  # convert error rts to 0
nw_cplx <- nw_cplx |> mutate(rt.err = na_if(rt.err, 0))         # convert 0 rts to NA
nw_cplx.mean <- mean(nw_cplx$rt.err, na.rm = TRUE)  # get mean rt excluding errors
nw_cplx <- nw_cplx |> mutate(rt.err.imp = ifelse(is.na(rt.err),
                                                 nw_smpl.mean,
                                                 rt.err))  # replace missing values with mean
nw_cplx <- nw_cplx |> mutate(inv.rt = 1/rt.err.imp)  # creates new column with inverse RTs
```

Determines how much missing data there is. Creates new dataframe with just the non-missing data

```
cw_missing_data<- filter(cw, is.na(cw$rt.err))
(xtab.missing.data <- xtabs(~cw_famsize+lang_type, data=cw_missing_data))
```

```
          lang_type
cw_famsize Orthographic Semantic
     Small          206      176
     Large          224      164
```

```
nw.smpl_missing_data<- filter(nw_smpl, is.na(nw_smpl$rt.err))
(xtab.missing.data <- xtabs(~nw_famsize+lang_type, data=nw.smpl_missing_data))
```

```
          lang_type
nw_famsize Orthographic Semantic
     Small          194      225
     Large          195      277
```

```
nw.cplx_missing_data<- filter(nw_cplx, is.na(nw_cplx$rt.err))
(xtab.missing.data <- xtabs(~nw_famsize+lang_type, data=nw.cplx_missing_data))
```

```
          lang_type
nw_famsize Orthographic Semantic
     Small          215      237
     Large          233      264
```

With RT as dependent variable

```
library(ez)
library(car)
(m.cw <- ezANOVA(cw,
        dv = rt.err.imp,
        wid = SubjID,
        within = cw_famsize,
        between = lang_type))
```

```
Warning: Converting "SubjID" to factor for ANOVA.

Warning: Data is unbalanced (unequal N per group). Make sure you specified a
well-considered value for the type argument to ezANOVA().

Warning: Collapsing data to cell means. *IF* the requested effects are a subset
of the full design, you must use the "within_full" argument, else results may be
inaccurate.

$ANOVA
                Effect DFn DFd         F            p p<.05          ges
2            lang_type   1  58  3.020685 8.751568e-02       0.0480026898
3           cw_famsize   1  58 29.284656 1.242722e-06     * 0.0158162335
4 lang_type:cw_famsize   1  58  0.811432 3.714225e-01       0.0004450876
```

```
(m.nw_smpl <- ezANOVA(nw_smpl,
        dv = rt.err.imp,
        wid = SubjID,
        within = .(nw_famsize),
        between = lang_type))
```

```
Warning: Converting "SubjID" to factor for ANOVA.

Warning: Data is unbalanced (unequal N per group). Make sure you specified a
well-considered value for the type argument to ezANOVA().

Warning: Collapsing data to cell means. *IF* the requested effects are a subset
of the full design, you must use the "within_full" argument, else results may be
inaccurate.

$ANOVA
                Effect DFn DFd         F          p p<.05         ges
2            lang_type   1  58 5.5423736 0.02196919     * 0.080556561
3           nw_famsize   1  58 2.1315506 0.14969037       0.003045769
4 lang_type:nw_famsize   1  58 0.8958419 0.34782581       0.001282331
```

```
(m.nw_cplx <- ezANOVA(nw_cplx,
        dv = rt.err.imp,
        wid = SubjID,
        within = .(nw_famsize),
        between = lang_type))
```

```
Warning: Converting "SubjID" to factor for ANOVA.

Warning: Data is unbalanced (unequal N per group). Make sure you specified a
well-considered value for the type argument to ezANOVA().

Warning: Collapsing data to cell means. *IF* the requested effects are a subset
of the full design, you must use the "within_full" argument, else results may be
inaccurate.
```

```
$ANOVA
               Effect DFn DFd          F         p p<.05          ges
2           lang_type   1  58 0.60386305 0.4402659       9.747770e-03
3          nw_famsize   1  58 0.57081666 0.4529920       5.363524e-04
4 lang_type:nw_famsize   1  58 0.01075519 0.9177595       1.011114e-05
```

Get condition means

```r
#Define standard error of the mean function

sem <- function(x) sd(x)/sqrt(length(x))

(cw.cond.means <- cw |>
   group_by(cw_famsize, lang_type) |>
   summarise(mean = mean(rt.err.imp),
             se = sem(rt.err.imp),
             num_stim = n()))
```

```
# A tibble: 4 x 5
# Groups:   cw_famsize [2]
  cw_famsize lang_type     mean    se num_stim
  <fct>      <fct>        <dbl> <dbl>    <int>
1 Small      Orthographic  597.  3.15     1450
2 Small      Semantic      628.  3.28     1506
3 Large      Orthographic  584.  3.17     1450
4 Large      Semantic      609.  3.14     1515
```

```r
(nw_smpl.cond.means <- nw_smpl |>
    group_by(nw_famsize, lang_type) |>
    summarise(mean = mean(rt.err.imp),
              se = sem(rt.err.imp),
              num_stim = n()))
```

```
# A tibble: 4 x 5
# Groups:   nw_famsize [2]
  nw_famsize lang_type     mean    se num_stim
  <fct>      <fct>        <dbl> <dbl>    <int>
1 Small      Orthographic  688.  4.06      725
2 Small      Semantic      716.  4.06      755
3 Large      Orthographic  690.  4.06      725
4 Large      Semantic      725.  3.58      757
```

```r
(nw_cplx.cond.means <- nw_cplx |>
    group_by(nw_famsize, lang_type) |>
    summarise(mean = mean(rt.err.imp),
              se = sem(rt.err.imp),
              num_stim = n()))
```

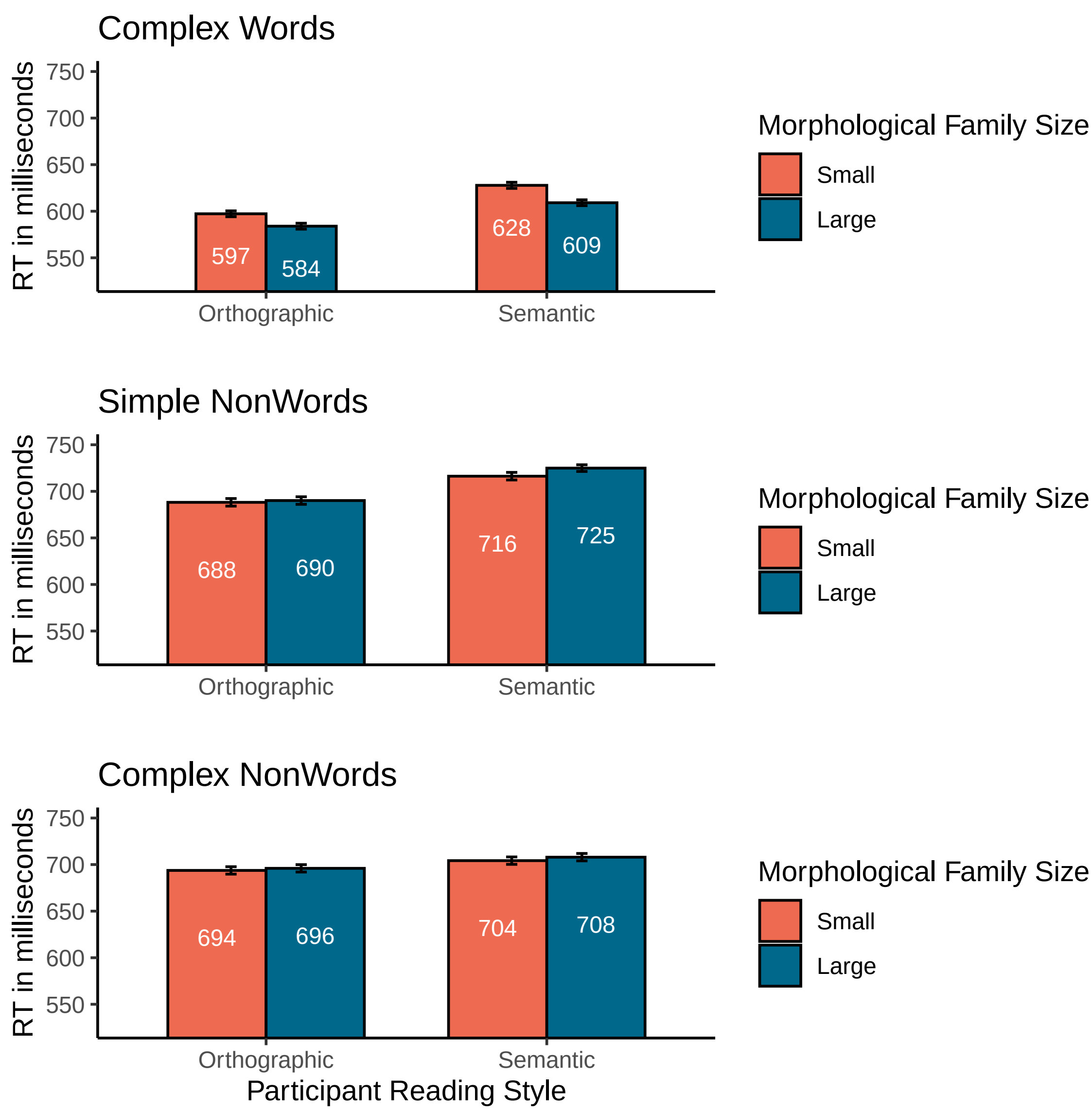
```
# A tibble: 4 x 5
# Groups:   nw_famsize [2]
  nw_famsize lang_type     mean    se num_stim
  <fct>      <fct>        <dbl> <dbl>    <int>
```

```
1 Small      Orthographic  694.  3.99      725
2 Small      Semantic      704.  3.97      761
3 Large      Orthographic  696.  3.96      725
4 Large      Semantic      708.  3.97      762
```

Barplots

```
library(gridExtra)
p1 <-  cw.cond.means %>% ggplot(aes(x=lang_type,
                                   y=mean,
                                   fill = cw_famsize,
                                   ymin = mean - se,
                                   ymax = mean + se)) +
  coord_cartesian(xlim = NULL,
                 ylim = c(525, 750),
                 expand = TRUE,
                 default = FALSE,
                 clip = "on") +
  geom_col(position = "dodge", width = 0.5, color = "black")  +
  ylab("RT in milliseconds")  +
  xlab("")  +
  ggtitle("Complex Words") +
  scale_fill_manual(values = c("coral2", "deepskyblue4"))+
  geom_errorbar(width = .08, position = position_dodge(0.5)) +
  theme_classic() +
   geom_text(aes(label = round(mean, digits = 0)),
            colour = "white",
            size = 3,
            vjust = 3,
            position = position_dodge(.5))+
  guides(fill=guide_legend(title="Morphological Family Size"))

p2 <-  nw_smpl.cond.means %>% ggplot(aes(x=lang_type,
                                        y=mean,
                                        fill = nw_famsize,
                                        ymin = mean - se,
                                        ymax = mean + se)) +
  coord_cartesian(xlim = NULL, ylim = c(525, 750),
                 expand = TRUE,
                 default = FALSE,
                 clip = "on") +
  geom_col(position = "dodge", width = .7, color = "black")  +
  xlab("")  +
  ylab("RT in milliseconds")  +
  ggtitle("Simple NonWords") +
  scale_fill_manual(values = c("coral2", "deepskyblue4"))+
  geom_errorbar(width = .08, position = position_dodge(0.5)) +
  theme_classic() +
  geom_text(aes(label = round(mean, digits = 0)),
            colour = "white",
            size = 3,
            vjust = 4.5,
           position = position_dodge(.7)) +
  guides(fill=guide_legend(title="Morphological Family Size"))
```

```
p3 <-  nw_cplx.cond.means %>% ggplot(aes(x=lang_type, y=mean, fill = nw_famsize, ymin = mean - se, ymax
  coord_cartesian(xlim = NULL, ylim = c(525, 750), expand = TRUE, default = FALSE,clip = "on") +
  geom_col(position = "dodge", width = .7, color = "black")  +
  xlab("Participant Reading Style")  +
  ylab("RT in milliseconds")  +
  ggtitle("Complex NonWords") +
  scale_fill_manual(values = c("coral2", "deepskyblue4"))+
  geom_errorbar(width = .08, position = position_dodge(0.5)) + theme_classic() +
   geom_text(aes(label = round(mean, digits = 0)),colour = "white", size = 3, vjust = 4.5, position = po
  guides(fill=guide_legend(title="Morphological Family Size"))

grid.arrange(p1, p2, p3)
```



LME

Models

```r
library(lme4)

# COMPLEX WORDS

cw_null.model = lmer(rt.err.imp ~ 1 + (1|SubjID) + (1|cw_target),
                      data= cw, REML=FALSE)
summary(cw_null.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ 1 + (1 | SubjID) + (1 | cw_target)
   Data: cw

     AIC      BIC   logLik deviance df.resid
 72112.9  72139.7 -36052.5  72104.9     5917

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2785 -0.6813 -0.1465  0.4764  4.9695

Random effects:
 Groups    Name        Variance Std.Dev.
 cw_target (Intercept)   576.8   24.02
 SubjID    (Intercept)  3982.7   63.11
 Residual              10713.6  103.51
Number of obs: 5921, groups:  cw_target, 100; SubjID, 60

Fixed effects:
            Estimate Std. Error t value
(Intercept)  605.232      8.601   70.37
```

```r
# Main effects models with random intercepts
cw_main.model = lmer(rt.err.imp ~ lang_type + cw_famsize + (1|SubjID) + (1|cw_target),
                      data= cw, REML=FALSE)
summary(cw_main.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type + cw_famsize + (1 | SubjID) + (1 | cw_target)
   Data: cw

     AIC      BIC   logLik deviance df.resid
 72104.9  72145.1 -36046.5  72092.9     5915

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2878 -0.6784 -0.1493  0.4726  4.9400

Random effects:
```

```
 Groups    Name        Variance Std.Dev.
 cw_target (Intercept)   511.3   22.61
 SubjID    (Intercept)  3781.6   61.50
 Residual              10713.5  103.51
Number of obs: 5921, groups:  cw_target, 100; SubjID, 60

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        598.646     12.089  49.522
lang_typeSemantic   28.388     16.116   1.762
cw_famsizeLarge    -16.129      5.262  -3.065

Correlation of Fixed Effects:
            (Intr) lng_tS
lng_typSmnt -0.688
cw_famszLrg -0.218  0.000
```

```
# Interaction effects models with random intercepts
cw_inter.model = lmer(rt.err.imp ~ lang_type * cw_famsize + (1|SubjID) + (1|cw_target),
                      data= cw, REML=FALSE)
summary(cw_inter.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type * cw_famsize + (1 | SubjID) + (1 | cw_target)
   Data: cw

     AIC      BIC   logLik deviance df.resid
 72105.9  72152.7 -36045.9  72091.9     5914

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3011 -0.6791 -0.1469  0.4727  4.9275

Random effects:
 Groups    Name        Variance Std.Dev.
 cw_target (Intercept)   511.4   22.61
 SubjID    (Intercept)  3781.8   61.50
 Residual              10711.6  103.50
Number of obs: 5921, groups:  cw_target, 100; SubjID, 60

Fixed effects:
                                  Estimate Std. Error t value
(Intercept)                        597.234     12.166  49.089
lang_typeSemantic                   31.164     16.340   1.907
cw_famsizeLarge                    -13.306      5.936  -2.242
lang_typeSemantic:cw_famsizeLarge   -5.535      5.382  -1.028

Correlation of Fixed Effects:
            (Intr) lng_tS cw_fmL
lng_typSmnt -0.693
cw_famszLrg -0.244  0.076
lng_typS:_L  0.113 -0.165 -0.463
```

```r
# SIMPLE NONWORDS

nw.smpl_null.model = lmer(rt.err.imp ~ 1 + (1|SubjID) + (1|nw_target),
                          data= nw_smpl,
                          REML=FALSE)
summary(nw.smpl_null.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ 1 + (1 | SubjID) + (1 | nw_target)
   Data: nw_smpl

     AIC      BIC   logLik deviance df.resid
 35537.9  35561.8 -17764.9  35529.9     2958

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3363 -0.6270 -0.0966  0.5739  4.0367

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  474.6   21.78
 SubjID    (Intercept) 2575.7   50.75
 Residual              8689.4   93.22
Number of obs: 2962, groups:  nw_target, 100; SubjID, 60

Fixed effects:
            Estimate Std. Error t value
(Intercept)  705.123      7.117   99.08
```

```r
# Main effects models with random intercepts
nw.smpl_main.model = lmer(rt.err.imp ~ lang_type + nw_famsize + (1|SubjID) + (1|nw_target),
                          data= nw_smpl, REML=FALSE)
summary(nw.smpl_main.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
   Data: nw_smpl

     AIC      BIC   logLik deviance df.resid
 35535.4  35571.3 -17761.7  35523.4     2956

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3148 -0.6312 -0.0916  0.5719  4.0305

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  465.9   21.58
 SubjID    (Intercept) 2336.3   48.34
 Residual              8689.5   93.22
Number of obs: 2962, groups:  nw_target, 100; SubjID, 60

Fixed effects:
```

```
                  Estimate Std. Error t value
(Intercept)        686.303      9.943  69.026
lang_typeSemantic   30.926     12.956   2.387
nw_famsizeLarge      5.728      5.513   1.039

Correlation of Fixed Effects:
            (Intr) lng_tS
lng_typSmnt -0.672       
nw_famszLrg -0.277  0.000
```

```r
# Interaction effects models with random intercepts
nw.smpl_inter.model = lmer(rt.err.imp ~ lang_type * nw_famsize + (1|SubjID) + (1|nw_target),
                           data= nw_smpl, REML=FALSE)
summary(nw.smpl_inter.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type * nw_famsize + (1 | SubjID) + (1 | nw_target)
   Data: nw_smpl

     AIC      BIC   logLik deviance df.resid 
 35536.2  35578.1 -17761.1  35522.2     2955 

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-3.3368 -0.6306 -0.0937  0.5708  4.0514 

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  466.1   21.59   
 SubjID    (Intercept) 2337.5   48.35   
 Residual              8685.7   93.20   
Number of obs: 2962, groups:  nw_target, 100; SubjID, 60

Fixed effects:
                                  Estimate Std. Error t value
(Intercept)                        688.219     10.098  68.153
lang_typeSemantic                   27.166     13.407   2.026
nw_famsizeLarge                      1.895      6.533   0.290
lang_typeSemantic:nw_famsizeLarge    7.502      6.860   1.094

Correlation of Fixed Effects:
            (Intr) lng_tS nw_fmL
lng_typSmnt -0.684              
nw_famszLrg -0.323  0.137       
lng_typS:_L  0.174 -0.256 -0.536
```

```r
# COMPLEX NONWORDS

nw.cplx_null.model = lmer(rt.err.imp ~ 1 + (1|SubjID) + (1|nw_target),
                          data= nw_cplx, REML=FALSE)
summary(nw.cplx_null.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
```

```
Formula: rt.err.imp ~ 1 + (1 | SubjID) + (1 | nw_target)
   Data: nw_cplx

     AIC      BIC   logLik deviance df.resid
 35610.7  35634.7 -17801.3  35602.7     2969

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2418 -0.6006 -0.1426  0.5543  3.4087

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  379.9   19.49
 SubjID    (Intercept) 2810.5   53.01
 Residual              8537.8   92.40
Number of obs: 2973, groups:  nw_target, 99; SubjID, 60

Fixed effects:
            Estimate Std. Error t value
(Intercept)   700.41       7.32   95.69
```

```r
# Main effects models with random intercepts
nw.cplx_main.model = lmer(rt.err.imp ~ lang_type + nw_famsize + (1|SubjID) + (1|nw_target),
                          data= nw_cplx, REML=FALSE)
summary(nw.cplx_main.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
   Data: nw_cplx

     AIC      BIC   logLik deviance df.resid
 35613.7  35649.7 -17800.9  35601.7     2967

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2312 -0.5997 -0.1458  0.5547  3.4188

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  378     19.44
 SubjID    (Intercept) 2777     52.70
 Residual              8538     92.40
Number of obs: 2973, groups:  nw_target, 99; SubjID, 60

Fixed effects:
                  Estimate Std. Error t value
(Intercept)        693.036     10.598  65.394
lang_typeSemantic   11.342     14.033   0.808
nw_famsizeLarge      3.003      5.179   0.580

Correlation of Fixed Effects:
            (Intr) lng_tS
lng_typSmnt -0.684
nw_famszLrg -0.246  0.000
```

```r
# Interaction effects models with random intercepts
nw.cplx_inter.model = lmer(rt.err.imp ~ lang_type * nw_famsize  + (1|SubjID) + (1|nw_target),
                           data= nw_cplx, REML=FALSE)
summary(nw.cplx_inter.model)
```

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: rt.err.imp ~ lang_type * nw_famsize + (1 | SubjID) + (1 | nw_target)
   Data: nw_cplx

     AIC      BIC   logLik deviance df.resid
 35615.7  35657.6 -17800.8  35601.7     2966

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2354 -0.5991 -0.1472  0.5543  3.4150

Random effects:
 Groups    Name        Variance Std.Dev.
 nw_target (Intercept)  378.2   19.45
 SubjID    (Intercept) 2776.8   52.70
 Residual              8537.5   92.40
Number of obs: 2973, groups:  nw_target, 99; SubjID, 60

Fixed effects:
                                  Estimate Std. Error t value
(Intercept)                        693.402     10.740  64.565
lang_typeSemantic                   10.628     14.437   0.736
nw_famsizeLarge                      2.271      6.241   0.364
lang_typeSemantic:nw_famsizeLarge    1.428      6.788   0.210

Correlation of Fixed Effects:
            (Intr) lng_tS nw_fmL
lng_typSmnt -0.694
nw_famszLrg -0.292  0.131
lng_typS:_L  0.162 -0.235 -0.558
```

Model Comparisons

```r
anova(cw_null.model,cw_main.model)
```

```
Data: cw
Models:
cw_null.model: rt.err.imp ~ 1 + (1 | SubjID) + (1 | cw_target)
cw_main.model: rt.err.imp ~ lang_type + cw_famsize + (1 | SubjID) + (1 | cw_target)
              npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
cw_null.model    4 72113 72140 -36052    72105
cw_main.model    6 72105 72145 -36046    72093 11.999  2   0.002479 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(cw_main.model,cw_inter.model)
```

```
Data: cw
Models:
cw_main.model: rt.err.imp ~ lang_type + cw_famsize + (1 | SubjID) + (1 | cw_target)
cw_inter.model: rt.err.imp ~ lang_type * cw_famsize + (1 | SubjID) + (1 | cw_target)
               npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
cw_main.model     6 72105 72145 -36046    72093
cw_inter.model    7 72106 72153 -36046    72092 1.0575  1     0.3038
```

```
anova(nw.smpl_null.model,nw.smpl_main.model)
```

```
Data: nw_smpl
Models:
nw.smpl_null.model: rt.err.imp ~ 1 + (1 | SubjID) + (1 | nw_target)
nw.smpl_main.model: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
                   npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
nw.smpl_null.model    4 35538 35562 -17765    35530
nw.smpl_main.model    6 35535 35571 -17762    35523 6.5175  2    0.03844 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(nw.smpl_main.model,nw.smpl_inter.model)
```

```
Data: nw_smpl
Models:
nw.smpl_main.model: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
nw.smpl_inter.model: rt.err.imp ~ lang_type * nw_famsize + (1 | SubjID) + (1 | nw_target)
                    npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
nw.smpl_main.model     6 35535 35571 -17762    35523
nw.smpl_inter.model    7 35536 35578 -17761    35522 1.1958  1     0.2742
```

```
anova(nw.cplx_null.model,nw.cplx_main.model)
```

```
Data: nw_cplx
Models:
nw.cplx_null.model: rt.err.imp ~ 1 + (1 | SubjID) + (1 | nw_target)
nw.cplx_main.model: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
                   npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
nw.cplx_null.model    4 35611 35635 -17801    35603
nw.cplx_main.model    6 35614 35650 -17801    35602 0.9854  2      0.611
```

```
anova(nw.cplx_main.model,nw.cplx_inter.model)
```

```
Data: nw_cplx
Models:
nw.cplx_main.model: rt.err.imp ~ lang_type + nw_famsize + (1 | SubjID) + (1 | nw_target)
nw.cplx_inter.model: rt.err.imp ~ lang_type * nw_famsize + (1 | SubjID) + (1 | nw_target)
                    npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
nw.cplx_main.model     6 35614 35650 -17801    35602
nw.cplx_inter.model    7 35616 35658 -17801    35602 0.0443  1     0.8333
```