MP MetaAnalysis

Christina Bergmann

Contents


```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following object is masked from 'package:tidyr':
##
##     expand
```

```
## Loading 'metafor' package (version 1.9-9). For an overview
## and introduction to the package please type: help(metafor).

## Loading 'meta' package (version 4.9-0).
## Type 'help(meta)' for a brief overview.

##
## Attaching package: 'meta'

## The following objects are masked from 'package:metafor':
##
##     baujat, forest, funnel, funnel.default, labbe, radial,
##     trimfill
```

Preparation

Read in data and tidy up dataset

Descriptive data

The database contains data from 32 papers consisting of data from 2010 infants. In the tables below, we provide more descriptive information.

The next table shows what type of publications were included in our meta-analysis

publication_status	n_unique	count
dissertation	2	17
gray paper	2	14
paper	27	216
proceedings	1	4

The table below shows based on which data we calculated effect sizes.

es_method	n_unique	count
group_means_one	18	120
group_means_two	7	57
t_one	4	39
t_two	5	35

We also have different ways of comparison of the time-course data, as the next table shows.

within_measure_descriptive	n_unique	count
post-naming compared to pre-naming phase	10	29
post-naming phase compared with chance (=50%)	9	23
post-pre difference score compared with chance (=0)	13	52

Analysis time window

Where possible, we noted the time window for analysis. First, let's look at the offset (in milliseconds) after the start of the word, i.e. the begin of a give analysis window for a naming effect

```
offset_info <- db_ET_correct %>% group_by(offset) %>% summarize(count = n())


kable(offset_info)
```

offset	count
0	7
200	1
231	4
267	1
300	2
360	25
365	10
367	37
400	4
500	2
1133	1
NA	10

Next we look at duration (in seconds) of the post naming window, here,too, we see a lot of heterogeneity.

```
duration_info <- db_ET_correct %>% group_by(post_nam_dur) %>% summarize(count = n())


kable(duration_info)
```

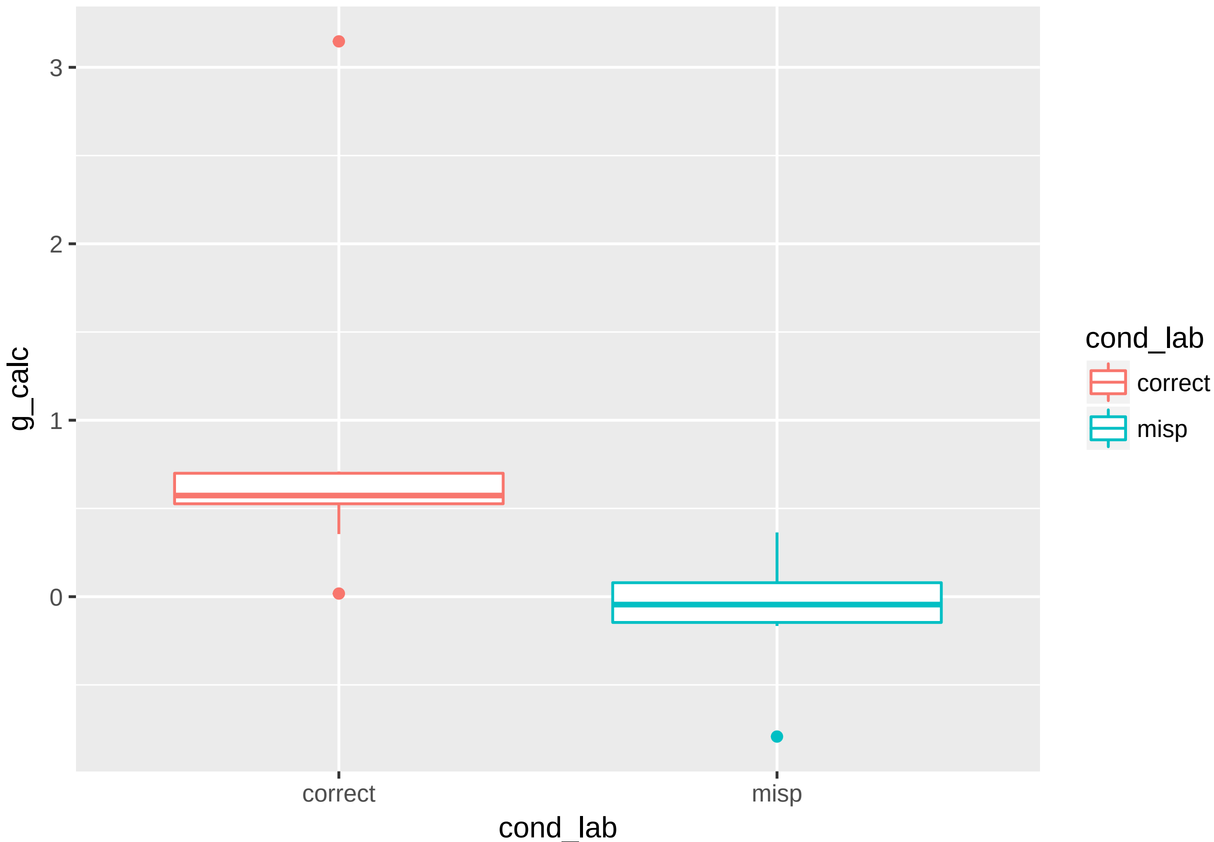
post_nam_dur	count
1.510	2
2.000	45
2.500	18
2.600	4
2.750	4
2.767	1
2.805	3
3.000	14
3.500	6
4.000	6
6.000	1

In summary, we see little consistency in analysis methods of comparable studies looking at naming and mispronunciation effects.

Early Ages

Even the youngest ages in the database (less than 1 year) show mispronunciation sensitivity

```
## [1] Zesiger et al. (2012)       Mani & Plunkett 2010
## [3] Bergelson & Swingley (2017)
## 34 Levels: Altvater-Mackensen (2010) ... Zesiger et al. (2012)
```



Meta-Analysis

Correct object identification effect

```
rma_correct = rma.mv(g_calc, g_var_calc, data = db_ET_correct, random = ~collapse |
    short_cite)

summary(rma_correct)
```

```
##
## Multivariate Meta-Analysis Model (k = 104; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -108.5748   217.1496   223.1496   231.0538   223.3920
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
```

```
## tau^2      0.3496  0.5912     no
## rho        0.8566             no
##
## Test for Heterogeneity:
## Q(df = 103) = 471.7565, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.9477   0.1060   8.9389   <.0001   0.7399   1.1555      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
rma_correct_age = rma.mv(g_calc, g_var_calc, mods = ~age.C, data = db_ET_correct,
    random = ~collapse | short_cite)

summary(rma_correct_age)
```

```
##
## Multivariate Meta-Analysis Model (k = 104; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -106.9830   213.9660   221.9660   232.4659   222.3784
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.3369  0.5805     no
## rho        0.8514             no
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 456.7433, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 1.8570, p-val = 0.1730
##
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.9662  0.1051  9.1932  <.0001   0.7602  1.1722  ***
## age.C      0.0225  0.0165  1.3627  0.1730  -0.0099  0.0549
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Mispronunciation object identification effect

```
rma_MP = rma.mv(g_calc, g_var_calc, data = db_ET_MP, random = ~collapse | short_cite)
```

```r
summary(rma_MP)
```

```
##
## Multivariate Meta-Analysis Model (k = 147; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -68.3155  136.6311  142.6311  151.5819  142.8001
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1059  0.3254     no
## rho        0.5511             no
##
## Test for Heterogeneity:
## Q(df = 146) = 418.1215, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.2726   0.0561   4.8564   <.0001   0.1626   0.3826      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_MP_age = rma.mv(g_calc, g_var_calc, mods = ~age.C, data = db_ET_MP, random = ~collapse |
    short_cite)

summary(rma_MP_age)
```

```
##
## Multivariate Meta-Analysis Model (k = 147; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -66.3311  132.6622  140.6622  152.5692  140.9479
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1005  0.3170     no
## rho        0.5590             no
##
## Test for Residual Heterogeneity:
## QE(df = 145) = 394.0193, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 3.3054, p-val = 0.0691
##
```

```
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.2869  0.0557  5.1550  <.0001   0.1778  0.3960  ***
## age.C      0.0195  0.0107  1.8181  0.0691  -0.0015  0.0405    .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Mispronunciation effect

```
db_ET_correct$condition <- 1
db_ET_MP$condition <- 0

dat <- bind_rows(db_ET_correct, db_ET_MP)

rma_MPeffect <- rma.mv(g_calc, g_var_calc, mods = ~condition, data = dat, random = ~collapse |
    short_cite)

summary(rma_MPeffect)
```

```
##
## Multivariate Meta-Analysis Model (k = 251; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -227.3842   454.7683   462.7683   476.8381   462.9322
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1102  0.3320     no
## rho        0.6750             no
##
## Test for Residual Heterogeneity:
## QE(df = 249) = 889.8779, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 259.5041, p-val < .0001
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## intrcpt      0.2891  0.0583   4.9575  <.0001  0.1748  0.4033  ***
## condition    0.5429  0.0337  16.1091  <.0001  0.4768  0.6089  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
rma_MPeffect_1 <- rma.mv(g_calc, g_var_calc, mods = ~condition - 1, data = dat,
    random = ~collapse | short_cite)
```

```r
summary(rma_MPeffect_1)
```

```
##
## Multivariate Meta-Analysis Model (k = 251; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -237.4011   474.8022   480.8022   491.3666   480.8998
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1875  0.4330     no
## rho        0.8120             no
##
## Test for Residual Heterogeneity:
## QE(df = 250) = 1000.5916, p-val < .0001
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## condition    0.5647  0.0333  16.9659  <.0001  0.4994  0.6299  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_MPeffect_age <- rma.mv(g_calc, g_var_calc, mods = ~age.C * condition, data = dat,
    random = ~collapse | short_cite)

summary(rma_MPeffect_age)
```

```
##
## Multivariate Meta-Analysis Model (k = 251; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -224.1771   448.3541   460.3541   481.4104   460.7041
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1004  0.3168     no
## rho        0.6683             no
##
## Test for Residual Heterogeneity:
## QE(df = 247) = 850.7626, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 265.7722, p-val < .0001
##
```

```
## Model Results:
##
##                    estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt              0.3069  0.0565   5.4279  <.0001   0.1961  0.4177  ***
## age.C                0.0216  0.0103   2.0911  0.0365   0.0014  0.0419    *
## condition            0.5490  0.0344  15.9808  <.0001   0.4816  0.6163  ***
## age.C:condition      0.0056  0.0076   0.7406  0.4589  -0.0093  0.0205
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Language effect

Followup: Per condition (correct or MP) the interaction with age

```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")

dat$lang_family = ifelse(dat$native_lang == "American English" | dat$native_lang ==
    "British English" | dat$native_lang == "Dutch" | dat$native_lang == "English" |
    dat$native_lang == "Danish" | dat$native_lang == "Swedish" | dat$native_lang ==
    "German", "Germanic", ifelse(dat$native_lang == "French" | dat$native_lang ==
    "Catalan" | dat$native_lang == "Spanish" | dat$native_lang == "Catalan-Spanish" |
    dat$native_lang == "Swiss French", "Romanic", "Sino-Tibetian"))

rma_lang_interaction <- rma.mv(g_calc, g_var_calc, mods = ~age.C * condition *
    lang_family, data = dat, random = ~collapse | short_cite)
summary(rma_lang_interaction)
```

```
##
## Multivariate Meta-Analysis Model (k = 251; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -219.8260   439.6521   467.6521   516.3226   469.5271
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.0960  0.3098     no
## rho        0.6499             no
##
## Test for Residual Heterogeneity:
## QE(df = 239) = 794.0142, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 272.2265, p-val < .0001
##
## Model Results:
##
##                                          estimate      se     zval
## intrcpt                                    0.2922  0.0608   4.8081
## age.C                                      0.0192  0.0113   1.7016
```

```
## condition                                    0.5366  0.0365  14.7179
## lang_familyRomanic                           0.2628  0.1669   1.5746
## lang_familySino-Tibetian                    -0.2552  0.2170  -1.1762
## age.C:condition                              0.0055  0.0082   0.6666
## age.C:lang_familyRomanic                     0.0290  0.0324   0.8957
## age.C:lang_familySino-Tibetian              -0.0082  0.0403  -0.2028
## condition:lang_familyRomanic                -0.0063  0.1360  -0.0463
## condition:lang_familySino-Tibetian           0.2344  0.1876   1.2493
## age.C:condition:lang_familyRomanic          -0.0161  0.0294  -0.5494
## age.C:condition:lang_familySino-Tibetian     0.0186  0.0336   0.5539
##                                                pval    ci.lb   ci.ub
## intrcpt                                      <.0001   0.1731  0.4113  ***
## age.C                                        0.0888  -0.0029  0.0412    .
## condition                                    <.0001   0.4652  0.6081  ***
## lang_familyRomanic                           0.1154  -0.0643  0.5900
## lang_familySino-Tibetian                     0.2395  -0.6805  0.1701
## age.C:condition                              0.5050  -0.0106  0.0215
## age.C:lang_familyRomanic                     0.3704  -0.0344  0.0924
## age.C:lang_familySino-Tibetian               0.8393  -0.0871  0.0708
## condition:lang_familyRomanic                 0.9631  -0.2729  0.2603
## condition:lang_familySino-Tibetian           0.2116  -0.1333  0.6022
## age.C:condition:lang_familyRomanic           0.5827  -0.0737  0.0414
## age.C:condition:lang_familySino-Tibetian     0.5796  -0.0472  0.0844
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Type of distractor

```
rma_DistractorAge <- rma.mv(g_calc, g_var_calc, mods = ~age.C * condition *
    as.factor(object_pair), data = dat, random = ~collapse | short_cite)

summary(rma_DistractorAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 251; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -220.0721   440.1441   460.1441   495.0748   461.0924
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.0966  0.3108     no
## rho        0.6643             no
##
## Test for Residual Heterogeneity:
## QE(df = 243) = 834.5402, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
```

```
## QM(df = 7) = 273.2542, p-val < .0001
##
## Model Results:
##
##                                                    estimate      se
## intrcpt                                              0.3916  0.0678
## age.C                                                0.0304  0.0125
## condition                                            0.5404  0.0415
## as.factor(object_pair)familiar_novel                -0.2673  0.1283
## age.C:condition                                      0.0091  0.0092
## age.C:as.factor(object_pair)familiar_novel          -0.0009  0.0255
## condition:as.factor(object_pair)familiar_novel       0.1002  0.0892
## age.C:condition:as.factor(object_pair)familiar_novel   -0.0273  0.0198
##                                                       zval    pval
## intrcpt                                             5.7748  <.0001
## age.C                                               2.4368  0.0148
## condition                                          13.0178  <.0001
## as.factor(object_pair)familiar_novel               -2.0844  0.0371
## age.C:condition                                     0.9907  0.3218
## age.C:as.factor(object_pair)familiar_novel         -0.0372  0.9704
## condition:as.factor(object_pair)familiar_novel      1.1232  0.2614
## age.C:condition:as.factor(object_pair)familiar_novel  -1.3812  0.1672
##                                                      ci.lb    ci.ub
## intrcpt                                             0.2587   0.5246
## age.C                                               0.0059   0.0548
## condition                                           0.4590   0.6218
## as.factor(object_pair)familiar_novel               -0.5187  -0.0160
## age.C:condition                                    -0.0089   0.0272
## age.C:as.factor(object_pair)familiar_novel         -0.0510   0.0491
## condition:as.factor(object_pair)familiar_novel     -0.0747   0.2751
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0660   0.0114
##
## intrcpt                                            ***
## age.C                                                *
## condition                                          ***
## as.factor(object_pair)familiar_novel                 *
## age.C:condition
## age.C:as.factor(object_pair)familiar_novel
## condition:as.factor(object_pair)familiar_novel
## age.C:condition:as.factor(object_pair)familiar_novel
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Subset to same age range

```
min_age <- min(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)
max_age <- max(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)

dat_age = dat %>% filter(mean_age_1 > min_age & mean_age_1 < max_age)

rma_DistractorAgeS <- rma.mv(g_calc, g_var_calc, mods = ~age.C * condition *
    as.factor(object_pair), data = dat_age, random = ~collapse | short_cite)
```

```r
summary(rma_DistractorAgeS)
```

```
##
## Multivariate Meta-Analysis Model (k = 185; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -157.7550   315.5101   335.5101   367.2716   336.8354
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 38)
##
##              estim    sqrt  fixed
## tau^2       0.1285  0.3585     no
## rho         0.6668             no
##
## Test for Residual Heterogeneity:
## QE(df = 177) = 601.2211, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 199.5303, p-val < .0001
##
## Model Results:
##
##                                                        estimate      se
## intrcpt                                                  0.4091  0.0869
## age.C                                                    0.0126  0.0251
## condition                                                0.4966  0.0465
## as.factor(object_pair)familiar_novel                    -0.3089  0.1710
## age.C:condition                                          0.0426  0.0188
## age.C:as.factor(object_pair)familiar_novel               0.0325  0.0471
## condition:as.factor(object_pair)familiar_novel           0.1104  0.1098
## age.C:condition:as.factor(object_pair)familiar_novel   -0.0181  0.0342
##                                                           zval    pval
## intrcpt                                                 4.7084  <.0001
## age.C                                                   0.5004  0.6168
## condition                                              10.6791  <.0001
## as.factor(object_pair)familiar_novel                   -1.8066  0.0708
## age.C:condition                                         2.2614  0.0237
## age.C:as.factor(object_pair)familiar_novel              0.6886  0.4911
## condition:as.factor(object_pair)familiar_novel          1.0056  0.3146
## age.C:condition:as.factor(object_pair)familiar_novel  -0.5299  0.5961
##                                                          ci.lb   ci.ub
## intrcpt                                                 0.2388  0.5794  ***
## age.C                                                  -0.0367  0.0619
## condition                                               0.4055  0.5878  ***
## as.factor(object_pair)familiar_novel                   -0.6441  0.0262    .
## age.C:condition                                         0.0057  0.0795    *
## age.C:as.factor(object_pair)familiar_novel             -0.0599  0.1248
## condition:as.factor(object_pair)familiar_novel         -0.1048  0.3257
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0851  0.0489
##
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Distractor Overlap

```r
rma_DistractorOverlap <- rma.mv(g_calc, g_var_calc, mods = ~age.C * distractor_overlap,
    data = db_ET_MP, random = ~collapse | short_cite)

summary(rma_DistractorOverlap)
```

```
##
## Multivariate Meta-Analysis Model (k = 147; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -59.0052  118.0103  138.0103  167.3551  139.7291
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.0926  0.3043     no
## rho        0.5820             no
##
## Test for Residual Heterogeneity:
## QE(df = 139) = 347.9572, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 14.1393, p-val = 0.0488
##
## Model Results:
##
##                                 estimate      se     zval    pval    ci.lb
## intrcpt                           0.0892  0.3625   0.2460  0.8057  -0.6213
## age.C                             0.0182  0.0190   0.9621  0.3360  -0.0189
## distractor_overlapno              0.4669  0.3770   1.2386  0.2155  -0.2719
## distractor_overlapnovel          -0.0179  0.3828  -0.0468  0.9626  -0.7682
## distractor_overlaponset           0.1342  0.3652   0.3675  0.7132  -0.5816
## distractor_overlaponset/medial    0.1547  0.4969   0.3113  0.7555  -0.8192
## age.C:distractor_overlapno        0.0311  0.0264   1.1789  0.2385  -0.0206
## age.C:distractor_overlapnovel     0.0131  0.0300   0.4356  0.6631  -0.0458
##                                  ci.ub
## intrcpt                         0.7996
## age.C                           0.0554
## distractor_overlapno            1.2057
## distractor_overlapnovel         0.7324
## distractor_overlaponset         0.8500
## distractor_overlaponset/medial  1.1286
## age.C:distractor_overlapno      0.0828
## age.C:distractor_overlapnovel   0.0719
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Number of features

```
db_ET_MPf = db_ET_MP %>% filter(n_feature != "1-3" & n_feature != "1-2" & n_feature !=
    "2-3")

# rma_NFeatures <- rma.mv(g_calc, g_var_calc, mods = ~as.ordered(n_feature),
# data = db_ET_MP, random = ~collapse | short_cite)
rma_NFeatures <- rma.mv(g_calc, g_var_calc, mods = ~as.factor(n_feature), data = db_ET_MPf,
    random = ~collapse | short_cite)


summary(rma_NFeatures)
```

```
##
## Multivariate Meta-Analysis Model (k = 132; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -57.1099  114.2198  130.2198  152.9101  131.4506
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1057  0.3252     no
## rho        0.4295             no
##
## Test for Residual Heterogeneity:
## QE(df = 126) = 354.2893, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 10.5852, p-val = 0.0603
##
## Model Results:
##
##                              estimate      se     zval    pval    ci.lb
## intrcpt                        0.3147  0.0610   5.1612  <.0001   0.1952
## as.factor(n_feature)2         -0.1600  0.0798  -2.0054  0.0449  -0.3163
## as.factor(n_feature)3         -0.2741  0.1050  -2.6094  0.0091  -0.4800
## as.factor(n_feature)41640     -0.2666  0.2354  -1.1324  0.2575  -0.7280
## as.factor(n_feature)41641     -0.2497  0.1350  -1.8490  0.0645  -0.5144
## as.factor(n_feature)41672     -0.3147  0.3289  -0.9569  0.3386  -0.9593
##                                ci.ub
## intrcpt                       0.4342  ***
## as.factor(n_feature)2        -0.0036    *
## as.factor(n_feature)3        -0.0682   **
## as.factor(n_feature)41640     0.1948
## as.factor(n_feature)41641     0.0150    .
## as.factor(n_feature)41672     0.3299
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_NFeaturesAge <- rma.mv(g_calc, g_var_calc, mods = ~as.factor(n_feature) *
    age.C, data = db_ET_MPf, random = ~collapse | short_cite)


summary(rma_NFeaturesAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 132; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -56.7024  113.4047  141.4047  180.4296  145.4047
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1087  0.3297     no
## rho        0.4092             no
##
## Test for Residual Heterogeneity:
## QE(df = 120) = 326.6211, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 13.4855, p-val = 0.2628
##
## Model Results:
##
##                                   estimate      se     zval    pval
## intrcpt                             0.3209  0.0614   5.2262  <.0001
## as.factor(n_feature)2              -0.1655  0.0804  -2.0575  0.0396
## as.factor(n_feature)3              -0.2639  0.1100  -2.3981  0.0165
## as.factor(n_feature)41640          -0.3105  0.2536  -1.2244  0.2208
## as.factor(n_feature)41641          -0.2689  0.1404  -1.9149  0.0555
## as.factor(n_feature)41672          -0.3209  0.8487  -0.3781  0.7054
## age.C                               0.0171  0.0138   1.2374  0.2159
## as.factor(n_feature)2:age.C         0.0111  0.0180   0.6151  0.5385
## as.factor(n_feature)3:age.C        -0.0035  0.0225  -0.1574  0.8749
## as.factor(n_feature)41640:age.C    -0.0215  0.0508  -0.4234  0.6720
## as.factor(n_feature)41641:age.C    -0.0283  0.0459  -0.6161  0.5379
## as.factor(n_feature)41672:age.C    -0.0171  1.5926  -0.0107  0.9914
##                                     ci.lb    ci.ub
## intrcpt                            0.2005   0.4412  ***
## as.factor(n_feature)2             -0.3232  -0.0078    *
## as.factor(n_feature)3             -0.4795  -0.0482    *
## as.factor(n_feature)41640         -0.8074   0.1865
## as.factor(n_feature)41641         -0.5440   0.0063    .
## as.factor(n_feature)41672         -1.9843   1.3426
## age.C                             -0.0100   0.0442
## as.factor(n_feature)2:age.C       -0.0242   0.0463
## as.factor(n_feature)3:age.C       -0.0477   0.0406
## as.factor(n_feature)41640:age.C   -0.1212   0.0781
## as.factor(n_feature)41641:age.C   -0.1183   0.0617
```

```
## as.factor(n_feature)41672:age.C  -3.1386   3.1044
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction with condition

```r
dat_f <- subset(dat, n_feature == "0" | n_feature == "1" | n_feature == "2" |
    n_feature == "3")

# rma_NFeatures <- rma.mv(g_calc, g_var_calc, mods = ~as.ordered(n_feature),
# data = db_ET_MP, random = ~collapse | short_cite)
rma_NFeatures <- rma.mv(g_calc, g_var_calc, mods = ~as.factor(n_feature) * condition,
    data = dat_f, random = ~collapse | short_cite)


summary(rma_NFeatures)
```

```
##
## Multivariate Meta-Analysis Model (k = 211; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -207.2555   414.5109   428.5109   451.8061   429.0766
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1190  0.3449     no
## rho        0.6371             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 793.1423, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5):
## QM(df = 4) = 232.2411, p-val < .0001
##
## Model Results:
##
##                         estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt                   0.6673  0.1294   5.1569  <.0001   0.4137   0.9209
## as.factor(n_feature)1    -0.3567  0.1123  -3.1769  0.0015  -0.5767  -0.1366
## as.factor(n_feature)2    -0.4795  0.1284  -3.7343  0.0002  -0.7311  -0.2278
## as.factor(n_feature)3    -0.5946  0.1456  -4.0837  <.0001  -0.8800  -0.3092
## condition                 0.1919  0.1056   1.8168  0.0692  -0.0151   0.3989
##
## intrcpt                 ***
## as.factor(n_feature)1    **
## as.factor(n_feature)2   ***
## as.factor(n_feature)3   ***
## condition                 .
```

```
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_NFeaturesAge <- rma.mv(g_calc, g_var_calc, mods = ~as.factor(n_feature) *
    age.C * condition, data = dat_f, random = ~collapse | short_cite)


summary(rma_NFeaturesAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 211; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -202.9240   405.8481   429.8481   469.4878   431.5077
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1191  0.3451     no
## rho        0.7152             no
##
## Test for Residual Heterogeneity:
## QE(df = 201) = 751.7844, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10):
## QM(df = 9) = 243.7505, p-val < .0001
##
## Model Results:
##
##                              estimate      se     zval    pval    ci.lb
## intrcpt                        0.6883  0.1307   5.2677  <.0001   0.4322
## as.factor(n_feature)1         -0.3637  0.1125  -3.2315  0.0012  -0.5843
## as.factor(n_feature)2         -0.4960  0.1286  -3.8579  0.0001  -0.7480
## as.factor(n_feature)3         -0.6127  0.1490  -4.1113  <.0001  -0.9047
## age.C                          0.0887  0.0498   1.7795  0.0752  -0.0090
## condition                      0.1913  0.1057   1.8087  0.0705  -0.0160
## as.factor(n_feature)1:age.C   -0.0708  0.0483  -1.4645  0.1431  -0.1655
## as.factor(n_feature)2:age.C   -0.0377  0.0505  -0.7451  0.4562  -0.1367
## as.factor(n_feature)3:age.C   -0.0512  0.0522  -0.9808  0.3267  -0.1534
## age.C:condition               -0.0626  0.0477  -1.3128  0.1893  -0.1560
##                                 ci.ub
## intrcpt                        0.9444  ***
## as.factor(n_feature)1         -0.1431   **
## as.factor(n_feature)2         -0.2440  ***
## as.factor(n_feature)3         -0.3206  ***
## age.C                          0.1863    .
## condition                      0.3985    .
## as.factor(n_feature)1:age.C    0.0240
## as.factor(n_feature)2:age.C    0.0614
## as.factor(n_feature)3:age.C    0.0511
## age.C:condition                0.0308
```

```
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MP location

```r
# table(db_ET_MP$mispron_location)

db_ET_MP1 = db_ET_MP %>% filter(mispron_location == "onset" | mispron_location ==
    "medial")

# rma_NFeatures <- rma.mv(g_calc, g_var_calc, mods = ~as.ordered(n_feature),
# data = db_ET_MP, random = ~collapse | short_cite)
rma_Location <- rma.mv(g_calc, g_var_calc, mods = ~mispron_location, data = db_ET_MP1,
    random = ~collapse | short_cite)


summary(rma_Location)
```

```
##
## Multivariate Meta-Analysis Model (k = 114; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -56.0730  112.1460  120.1460  131.0200  120.5199
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1333  0.3651     no
## rho        0.5013             no
##
## Test for Residual Heterogeneity:
## QE(df = 112) = 351.3582, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.0003, p-val = 0.9854
##
## Model Results:
##
##                         estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                   0.2677  0.0802  3.3384  0.0008   0.1106  0.4249
## mispron_locationmedial    0.0026  0.1430  0.0183  0.9854  -0.2777  0.2829
##
## intrcpt                 ***
## mispron_locationmedial
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_LocationAge <- rma.mv(g_calc, g_var_calc, mods = ~mispron_location * age.C,
    data = db_ET_MP1, random = ~collapse | short_cite)


summary(rma_LocationAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 114; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -54.2638  108.5275  120.5275  136.7304  121.3431
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1355  0.3681     no
## rho        0.5078             no
##
## Test for Residual Heterogeneity:
## QE(df = 110) = 337.5012, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 2.0758, p-val = 0.5568
##
## Model Results:
##
##                                 estimate      se    zval    pval    ci.lb
## intrcpt                           0.2713  0.0816  3.3251  0.0009   0.1114
## mispron_locationmedial            0.0427  0.1586  0.2692  0.7878  -0.2681
## age.C                             0.0195  0.0170  1.1449  0.2523  -0.0139
## mispron_locationmedial:age.C      0.0107  0.0320  0.3331  0.7390  -0.0521
##                                  ci.ub
## intrcpt                         0.4313  ***
## mispron_locationmedial          0.3535
## age.C                           0.0528
## mispron_locationmedial:age.C    0.0735
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MP type: Vowel, consonant, or tone?

```r
db_MP_type <- subset(db_ET_MP, type_feature == "consonant" | type_feature ==
    "vowel")




rma_TypeFeaturesMP <- rma.mv(g_calc, g_var_calc, mods = ~type_feature, data = db_MP_type,
    random = ~collapse | short_cite)
```

```r
summary(rma_TypeFeaturesMP)
```

```
##
## Multivariate Meta-Analysis Model (k = 133; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -62.4062  124.8124  132.8124  144.3131  133.1298
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1111  0.3333     no
## rho        0.5216             no
##
## Test for Residual Heterogeneity:
## QE(df = 131) = 375.4320, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.0180, p-val = 0.8932
##
## Model Results:
##
##                    estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt              0.2615  0.0687  3.8070  0.0001   0.1269  0.3961  ***
## type_featurevowel    0.0116  0.0861  0.1343  0.8932  -0.1572  0.1804
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
rma_TypeFeaturesMPAge <- rma.mv(g_calc, g_var_calc, mods = ~type_feature * age.C,
    data = db_MP_type, random = ~collapse | short_cite)


summary(rma_TypeFeaturesMPAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 133; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -60.7240  121.4480  133.4480  150.6069  134.1365
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1083  0.3291     no
## rho        0.5198             no
##
```

```
## Test for Residual Heterogeneity:
## QE(df = 129) = 356.5950, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 2.6683, p-val = 0.4456
##
## Model Results:
##
##                          estimate      se     zval    pval    ci.lb
## intrcpt                    0.2625  0.0682   3.8497  0.0001   0.1288
## type_featurevowel          0.0244  0.0864   0.2819  0.7780  -0.1450
## age.C                      0.0198  0.0140   1.4153  0.1570  -0.0076
## type_featurevowel:age.C   -0.0018  0.0166  -0.1099  0.9125  -0.0344
##                           ci.ub
## intrcpt                  0.3961  ***
## type_featurevowel        0.1937
## age.C                    0.0471
## type_featurevowel:age.C  0.0308
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction with condition

```
dat_type <- subset(dat, type_feature == "consonant" | type_feature == "vowel" |
    type_feature == "tone")
dat_type$type_feature <- as.factor(ifelse(dat_type$condition == 1, "none", dat_type$type_feature))


rma_TypeFeatures <- rma.mv(g_calc, g_var_calc, mods = ~relevel(type_feature,
    "none") * condition, data = dat_type, random = ~collapse | short_cite)


summary(rma_TypeFeatures)
```

```
##
## Multivariate Meta-Analysis Model (k = 228; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -212.7856   425.5712   437.5712   458.0411   437.9583
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.0966  0.3108     no
## rho        0.6207             no
##
## Test for Residual Heterogeneity:
## QE(df = 224) = 782.1529, p-val < .0001
##
```

```
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 195.9748, p-val < .0001
##
## Model Results:
##
##                                  estimate      se      zval    pval
## intrcpt                            0.7819  0.0615   12.7207  <.0001
## relevel(type_feature, "none")1   -0.5130  0.0421  -12.1823  <.0001
## relevel(type_feature, "none")4   -0.6762  0.1531   -4.4178  <.0001
## relevel(type_feature, "none")5   -0.4976  0.0561   -8.8688  <.0001
##                                    ci.lb    ci.ub
## intrcpt                           0.6615   0.9024  ***
## relevel(type_feature, "none")1  -0.5955  -0.4305  ***
## relevel(type_feature, "none")4  -0.9762  -0.3762  ***
## relevel(type_feature, "none")5  -0.6076  -0.3877  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
rma_TypeFeaturesAge <- rma.mv(g_calc, g_var_calc, mods = ~relevel(type_feature,
    "none") * age.C * condition, data = dat_type, random = ~collapse | short_cite)


summary(rma_TypeFeaturesAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 228; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -209.4438   418.8876   438.8876   472.8238   439.9402
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.0919  0.3031     no
## rho        0.6087             no
##
## Test for Residual Heterogeneity:
## QE(df = 220) = 755.7262, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 203.2440, p-val < .0001
##
## Model Results:
##
##                                        estimate      se      zval    pval
## intrcpt                                  0.8054  0.0614   13.1267  <.0001
## relevel(type_feature, "none")1         -0.5233  0.0425  -12.3184  <.0001
## relevel(type_feature, "none")4         -0.6758  0.1681   -4.0203  <.0001
## relevel(type_feature, "none")5         -0.5266  0.0623   -8.4565  <.0001
## age.C                                    0.0233  0.0113    2.0582  0.0396
## relevel(type_feature, "none")1:age.C    0.0063  0.0103    0.6140  0.5392
```

```
## relevel(type_feature, "none")4:age.C   -0.0017  0.0308   -0.0545  0.9565
## relevel(type_feature, "none")5:age.C   -0.0115  0.0113   -1.0161  0.3096
##                                         ci.lb    ci.ub
## intrcpt                                0.6851   0.9256  ***
## relevel(type_feature, "none")1        -0.6066  -0.4401  ***
## relevel(type_feature, "none")4        -1.0053  -0.3463  ***
## relevel(type_feature, "none")5        -0.6486  -0.4045  ***
## age.C                                  0.0011   0.0456    *
## relevel(type_feature, "none")1:age.C  -0.0138   0.0264
## relevel(type_feature, "none")4:age.C  -0.0620   0.0586
## relevel(type_feature, "none")5:age.C  -0.0338   0.0107
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction with language

```r
# dat_type <- subset(dat, type_feature == 'consonant' | type_feature ==
# 'vowel' | type_feature == 'tone')

dat_type <- subset(dat, type_feature == "consonant" | type_feature == "vowel")

dat_type$type_feature <- as.factor(ifelse(dat_type$condition == 1, "none", dat_type$type_feature))

dat_type$lang_family = ifelse(dat_type$native_lang == "American English" | dat_type$native_lang ==
    "British English" | dat_type$native_lang == "Dutch" | dat_type$native_lang ==
    "Danish" | dat_type$native_lang == "Swedish" | dat_type$native_lang == "English" |
    dat_type$native_lang == "German", "Germanic", ifelse(dat_type$native_lang ==
    "French" | dat_type$native_lang == "Catalan" | dat_type$native_lang == "Spanish" |
    dat_type$native_lang == "Catalan-Spanish" | dat_type$native_lang == "Swiss French",
    "Romanic", "Sino-Tibetian"))

dat_type_sub <- subset(dat_type, lang_family != "Sino-Tibetian")

rma_TypeFeatures_Lang <- rma.mv(g_calc, g_var_calc, mods = ~relevel(type_feature,
    "none") * lang_family, data = dat_type_sub, random = ~collapse | short_cite)


summary(rma_TypeFeatures_Lang)
```

```
##
## Multivariate Meta-Analysis Model (k = 212; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -203.2441   406.4882   422.4882   449.1112   423.2192
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1004  0.3168     no
```

```
## rho        0.5036             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 707.7561, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 197.5595, p-val < .0001
##
## Model Results:
##
##                                                     estimate      se
## intrcpt                                               0.7456  0.0690
## relevel(type_feature, "none")1                       -0.4909  0.0439
## relevel(type_feature, "none")5                       -0.5205  0.0635
## lang_familyRomanic                                    0.3774  0.1634
## relevel(type_feature, "none")1:lang_familyRomanic   -0.5456  0.2103
## relevel(type_feature, "none")5:lang_familyRomanic    0.1273  0.1481
##                                                         zval    pval
## intrcpt                                              10.8008  <.0001
## relevel(type_feature, "none")1                      -11.1728  <.0001
## relevel(type_feature, "none")5                       -8.2018  <.0001
## lang_familyRomanic                                    2.3100  0.0209
## relevel(type_feature, "none")1:lang_familyRomanic   -2.5940  0.0095
## relevel(type_feature, "none")5:lang_familyRomanic    0.8591  0.3903
##                                                       ci.lb    ci.ub
## intrcpt                                              0.6103   0.8809  ***
## relevel(type_feature, "none")1                      -0.5770  -0.4048  ***
## relevel(type_feature, "none")5                      -0.6448  -0.3961  ***
## lang_familyRomanic                                   0.0572   0.6976    *
## relevel(type_feature, "none")1:lang_familyRomanic  -0.9579  -0.1334   **
## relevel(type_feature, "none")5:lang_familyRomanic  -0.1631   0.4176
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction with condition and language

```
dat_type <- subset(dat, type_feature == "consonant" | type_feature == "vowel")  # |
# type_feature == 'tone')
dat_type$type_feature <- as.factor(ifelse(dat_type$condition == 1, "none", dat_type$type_feature))

dat_type$lang_family = ifelse(dat_type$native_lang == "American English" | dat_type$native_lang ==
    "British English" | dat_type$native_lang == "Dutch" | dat_type$native_lang ==
    "Danish" | dat_type$native_lang == "Swedish" | dat_type$native_lang == "English" |
    dat_type$native_lang == "German", "Germanic", ifelse(dat_type$native_lang ==
    "French" | dat_type$native_lang == "Catalan" | dat_type$native_lang == "Spanish" |
    dat_type$native_lang == "Catalan-Spanish" | dat_type$native_lang == "Swiss French",
    "Romanic", "Sino-Tibetian"))

dat_type_sub <- subset(dat_type, lang_family != "Sino-Tibetian")
dat_type_sub$lang_family <- as.factor(dat_type_sub$lang_family)
```

```r
rma_TypeFeatures_Lang <- rma.mv(g_calc, g_var_calc, mods = ~relevel(type_feature,
    "none") * lang_family * condition, data = dat_type_sub, random = ~collapse |
    short_cite)


summary(rma_TypeFeatures_Lang)
```

```
##
## Multivariate Meta-Analysis Model (k = 212; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -203.2441   406.4882   422.4882   449.1112   423.2192
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1004  0.3168     no
## rho        0.5036             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 707.7561, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 197.5595, p-val < .0001
##
## Model Results:
##
##                                                     estimate      se
## intrcpt                                               0.7456  0.0690
## relevel(type_feature, "none")1                       -0.4909  0.0439
## relevel(type_feature, "none")5                       -0.5205  0.0635
## lang_familyRomanic                                    0.3774  0.1634
## relevel(type_feature, "none")1:lang_familyRomanic   -0.5456  0.2103
## relevel(type_feature, "none")5:lang_familyRomanic    0.1273  0.1481
##                                                         zval    pval
## intrcpt                                              10.8008  <.0001
## relevel(type_feature, "none")1                      -11.1728  <.0001
## relevel(type_feature, "none")5                       -8.2018  <.0001
## lang_familyRomanic                                    2.3100  0.0209
## relevel(type_feature, "none")1:lang_familyRomanic   -2.5940  0.0095
## relevel(type_feature, "none")5:lang_familyRomanic    0.8591  0.3903
##                                                       ci.lb    ci.ub
## intrcpt                                              0.6103   0.8809  ***
## relevel(type_feature, "none")1                      -0.5770  -0.4048  ***
## relevel(type_feature, "none")5                      -0.6448  -0.3961  ***
## lang_familyRomanic                                   0.0572   0.6976    *
## relevel(type_feature, "none")1:lang_familyRomanic  -0.9579  -0.1334   **
## relevel(type_feature, "none")5:lang_familyRomanic  -0.1631   0.4176
##
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Correlation MP effect and vocabulary

First, let's take a look at which vocabulary data we have available.

```r
vocab_info <- db_ET_correct %>% mutate(has_vocab = ifelse(!is.na(r_comprehension),
    "comprehension", ifelse(!is.na(r_production), "production", "none"))) %>%
    group_by(has_vocab) %>% summarize(count = n())

kable(vocab_info)
```

has_vocab
comprehension
none
production

We have 17 corre lations, roughly evenly divided between comprehension and production data. There is reason to believe

```r
# we're relying on the library meta function metacor
compr <- subset(db_ET_correct, !is.na(db_ET_correct$r_comprehension) & r_comprehension >
    -1)

metacor(cor = r_comprehension, n = n_1, studlab = short_cite, data = compr,
    sm = "COR")
```

```
##                                        COR            95%-CI %W(fixed)
## Zesiger et al. (2012)               0.0610 [-0.3553; 0.4773]       5.8
## Zesiger et al. (2012)              -0.1590 [-0.5663; 0.2483]       6.1
## Mani, Coleman, & Plunkett (2008)   0.0300 [-0.2271; 0.2871]      15.2
## Swingley & Aslin (2000)             0.1050 [-0.1564; 0.3664]      14.7
## Mani & Plunkett 2007               -0.1700 [-0.5234; 0.1834]       8.0
## Mani & Plunkett 2007               -0.1700 [-0.5175; 0.1775]       8.3
## Swingley & Aslin (2002)             0.1410 [-0.2432; 0.5252]       6.8
## Swingley & Aslin (2002)             0.1410 [-0.2596; 0.5416]       6.3
## Swingley 2003                       0.3400 [ 0.0470; 0.6330]      11.7
## Swingley 2003                       0.0600 [-0.3472; 0.4672]       6.1
## H\xbfjen et al.                     0.2220 [-0.2591; 0.7031]       4.3
## H\xbfjen et al.                     0.4820 [ 0.0935; 0.8705]       6.7
##                                   %W(random)
## Zesiger et al. (2012)                     6.2
## Zesiger et al. (2012)                     6.5
## Mani, Coleman, & Plunkett (2008)         13.7
## Swingley & Aslin (2000)                  13.4
## Mani & Plunkett 2007                      8.3
## Mani & Plunkett 2007                      8.5
## Swingley & Aslin (2002)                   7.2
## Swingley & Aslin (2002)                   6.7
## Swingley 2003                            11.2
## Swingley 2003                             6.5
## H\xbfjen et al.                           4.8
## H\xbfjen et al.                           7.0
##
## Number of studies combined: k = 12
```

```
##
##                          COR            95%-CI    z p-value
## Fixed effect model   0.0897 [-0.0105; 0.1900] 1.75  0.0795
## Random effects model 0.0893 [-0.0212; 0.1999] 1.58  0.1132
##
## Quantifying heterogeneity:
## tau^2 = 0.0060; H = 1.09 [1.00; 1.50]; I^2 = 15.7% [0.0%; 55.4%]
##
## Test of heterogeneity:
##      Q d.f. p-value
##  13.05   11  0.2899
##
## Details on meta-analytical method:
## - Inverse variance method
## - DerSimonian-Laird estimator for tau^2
## - Untransformed correlations
```

```
# we're relying on the library meta function metacor
prodr <- subset(db_ET_correct, !is.na(db_ET_correct$r_production) & r_production <
    1)

metacor(cor = r_production, n = n_1, studlab = short_cite, data = prodr, sm = "COR")
```

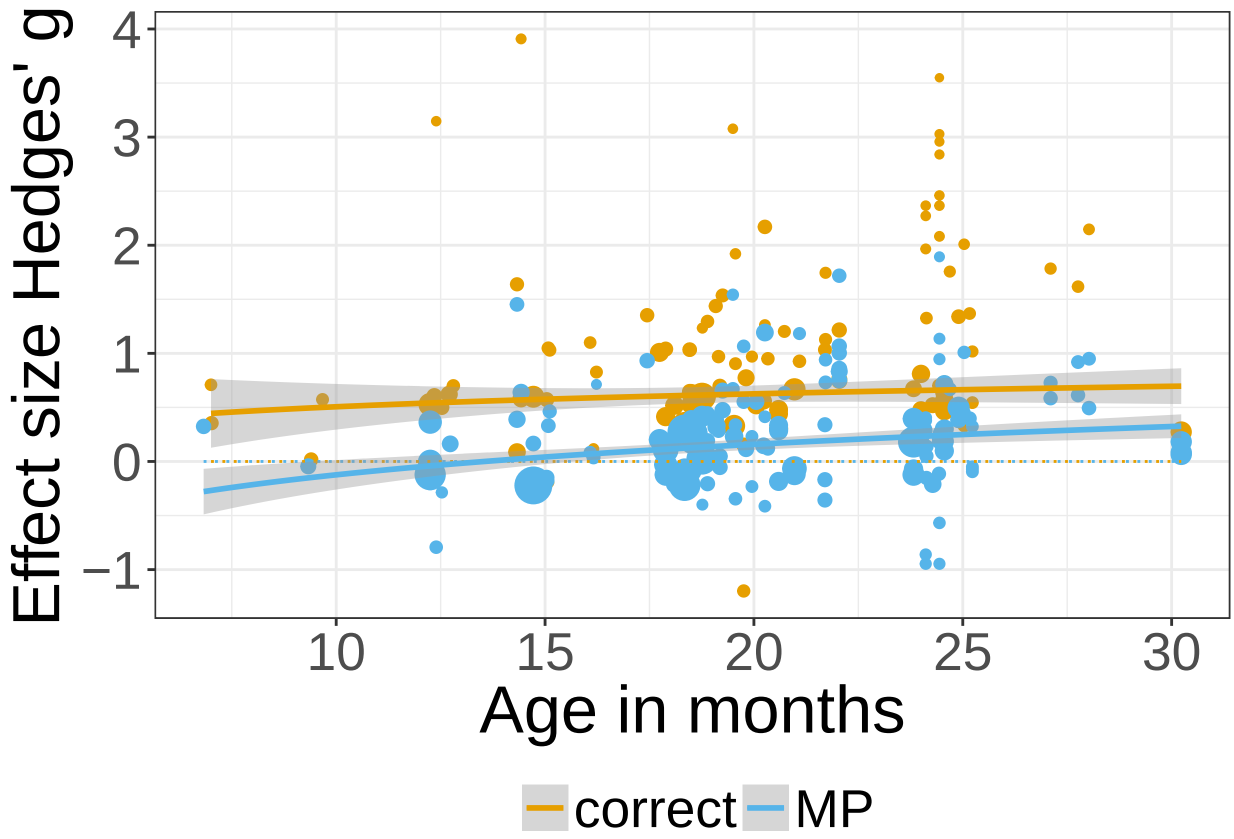
```
##                                         COR            95%-CI %W(fixed)
## Zesiger et al. (2012)               -0.0090 [-0.4268; 0.4088]       5.0
## Zesiger et al. (2012)               -0.1720 [-0.5775; 0.2335]       5.3
## Mani, Coleman, & Plunkett (2008)     0.0700 [-0.1861; 0.3261]      13.2
## Mani & Plunkett 2007                -0.1100 [-0.4696; 0.2496]       6.7
## Mani & Plunkett 2007                -0.1100 [-0.4635; 0.2435]       6.9
## Swingley & Aslin (2002)              0.1820 [-0.1970; 0.5610]       6.0
## Swingley & Aslin (2002)              0.1820 [-0.2131; 0.5771]       5.6
## Swingley 2003                        0.1800 [-0.1406; 0.5006]       8.4
## Swingley 2003                        0.0700 [-0.3367; 0.4767]       5.2
## Ramon-Casas et al. 2009              0.0980 [-0.3068; 0.5028]       5.3
## Ramon-Casas et al. 2009             -0.1470 [-0.5468; 0.2528]       5.4
## Ramon-Casas et al. 2009             -0.2300 [-0.6171; 0.1571]       5.8
## Ramon-Casas et al. 2009              0.2400 [-0.1451; 0.6251]       5.9
## Ramon-Casas et al. 2009              0.4350 [ 0.1037; 0.7663]       7.9
## H\xbfjen et al.                      0.2220 [-0.2591; 0.7031]       3.7
## H\xbfjen et al.                     -0.1480 [-0.6430; 0.3470]       3.5
##                                     %W(random)
## Zesiger et al. (2012)                      5.0
## Zesiger et al. (2012)                      5.3
## Mani, Coleman, & Plunkett (2008)          13.2
## Mani & Plunkett 2007                       6.7
## Mani & Plunkett 2007                       6.9
## Swingley & Aslin (2002)                    6.0
## Swingley & Aslin (2002)                    5.6
## Swingley 2003                              8.4
## Swingley 2003                              5.2
## Ramon-Casas et al. 2009                    5.3
## Ramon-Casas et al. 2009                    5.4
## Ramon-Casas et al. 2009                    5.8
## Ramon-Casas et al. 2009                    5.9
## Ramon-Casas et al. 2009                    7.9
```

```
## H\xbfjen et al.                         3.7
## H\xbfjen et al.                         3.5
##
## Number of studies combined: k = 16
##
##                           COR            95%-CI    z p-value
## Fixed effect model   0.0601 [-0.0331; 0.1533] 1.26  0.2061
## Random effects model 0.0601 [-0.0331; 0.1533] 1.26  0.2061
##
## Quantifying heterogeneity:
## tau^2 = 0; H = 1.00 [1.00; 1.42]; I^2 = 0.0% [0.0%; 50.7%]
##
## Test of heterogeneity:
##      Q d.f. p-value
##  14.51   15  0.4870
##
## Details on meta-analytical method:
## - Inverse variance method
## - DerSimonian-Laird estimator for tau^2
## - Untransformed correlations
```

Plotting

Mispronunciation Effect by Age (color)

```
# Color Blind palette:
cbPalette <- c("#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00",
    "#CC79A7")

p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    show.legend = FALSE) + geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm",
    formula = y ~ log(x), aes(weight = weights_g)) + scale_colour_manual(values = cbPalette) +
    theme_bw() + theme(text = element_text(size = 25), legend.title = element_blank(),
    legend.position = "bottom") + xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
min(dat$mean_age_1/30.44)
```

```
## [1] 6.826544
```

```
max(dat$mean_age_1/30.44)
```

```
## [1] 30.22996
```

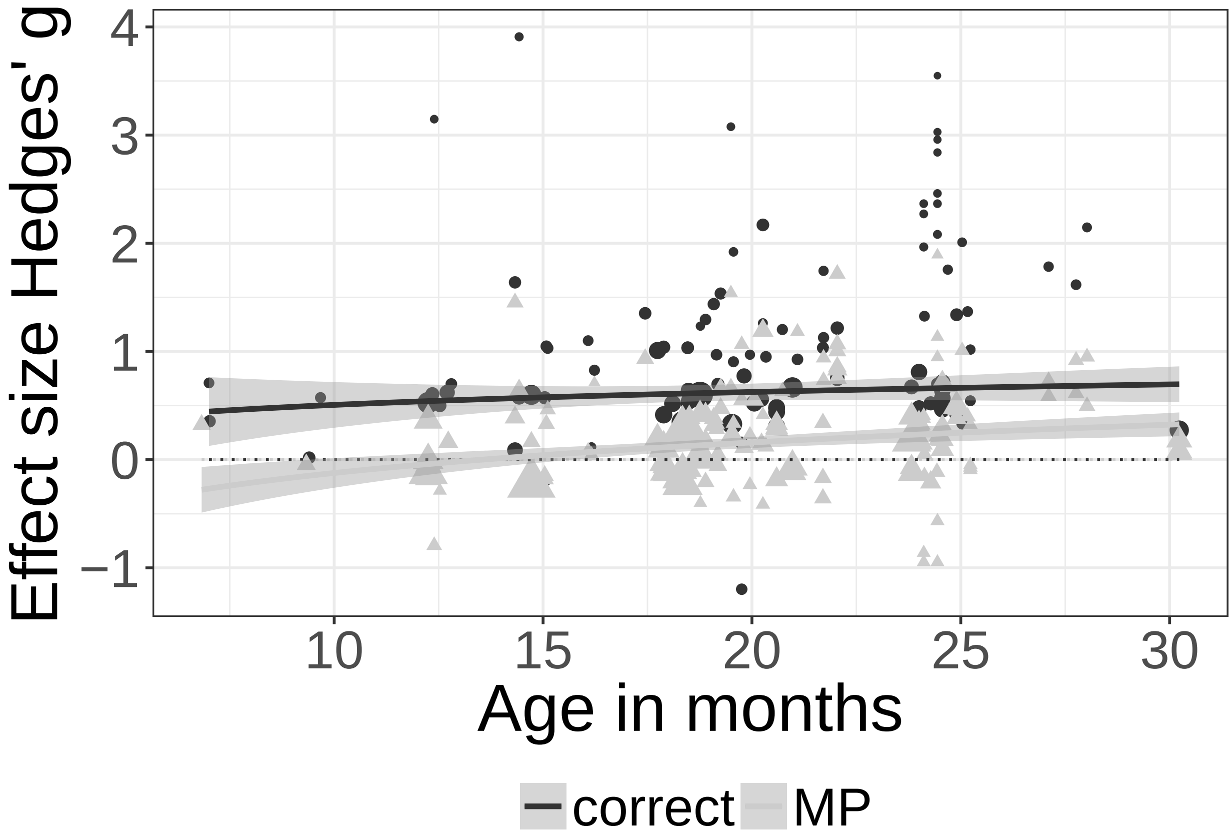
```
# ggsave('figures/AgeEffect_log.jpg', p,height= 7,width= 6)



jpeg(filename = "figures/AgeEffect_log.jpg", width = 600, height = 400, units = "px")

p

dev.off()
```

```
## pdf
##   2
```

Mispronunciation Effect by Age (bw)

```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = condition_label, color = condition_label), show.legend = FALSE) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g)) + scale_color_grey() + theme_bw() + theme(text = element_text(size
    legend.title = element_blank(), legend.position = "bottom") + xlab("Age in months") +
```

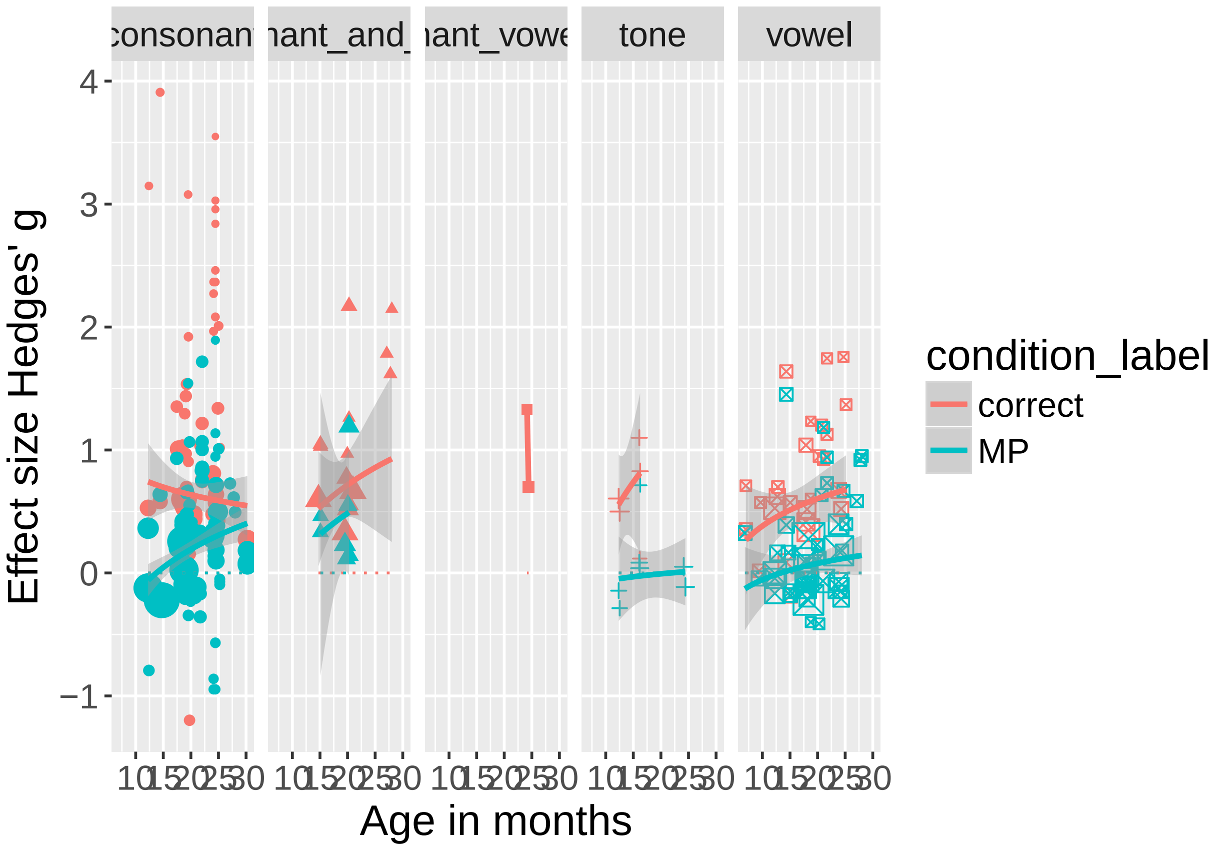
```
    ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_BW.jpg", p, height = 3, width = 6)
```

MP type: Consonant, Vowel, or Tone?

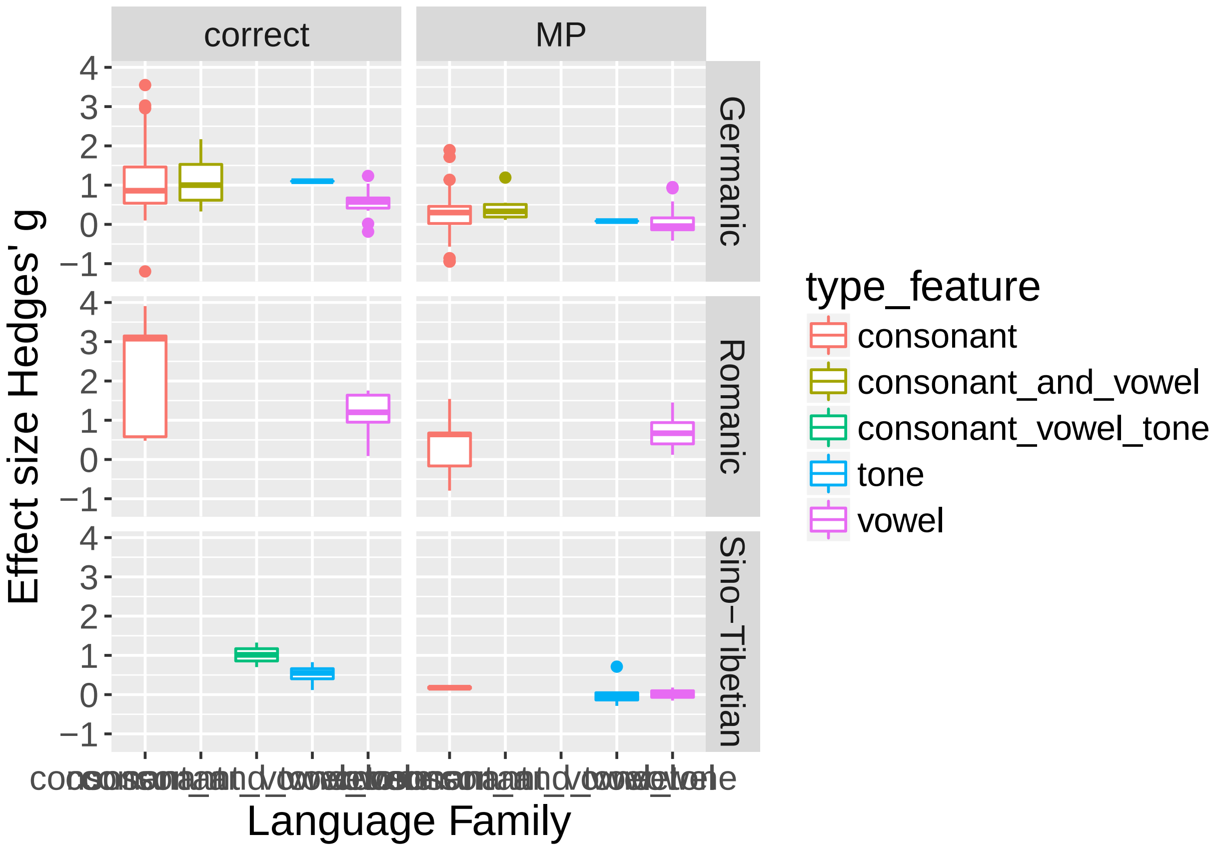
```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weight
    shape = type_feature), show.legend = FALSE) + facet_grid(. ~ type_feature) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g)) + theme(text = element_text(size = 16)) +
    xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_CV.jpg", p)
```

Language Family by MP type: Consonant, Vowel, or Tone?

```
p <- ggplot(dat, aes(type_feature, g_calc, color = type_feature)) + geom_boxplot() +
    facet_grid(lang_family ~ condition_label) + # geom_line(y= 0, linetype='dotted') + geom_smooth(meth
# y ~ log(x), aes(weight=weights_g)) +
theme(text = element_text(size = 16)) + xlab("Language Family") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/LangFamily_CV.jpg", p)
```

Number of Features

```
# dat_f <- subset(dat, n_feature == '0' | n_feature == '1' | n_feature ==
# '2' | n_feature == '3')

p <- ggplot(dat_f, aes(mean_age_1/30.44, g_calc, color = n_feature)) + geom_point(aes(size = weights_g,
    shape = n_feature), show.legend = FALSE) + # facet_grid(.~type_feature)+
geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g)) + theme(text = element_text(size = 16)) +
    xlab("Age in months") + ylab("Effect size Hedges' g")
p
```

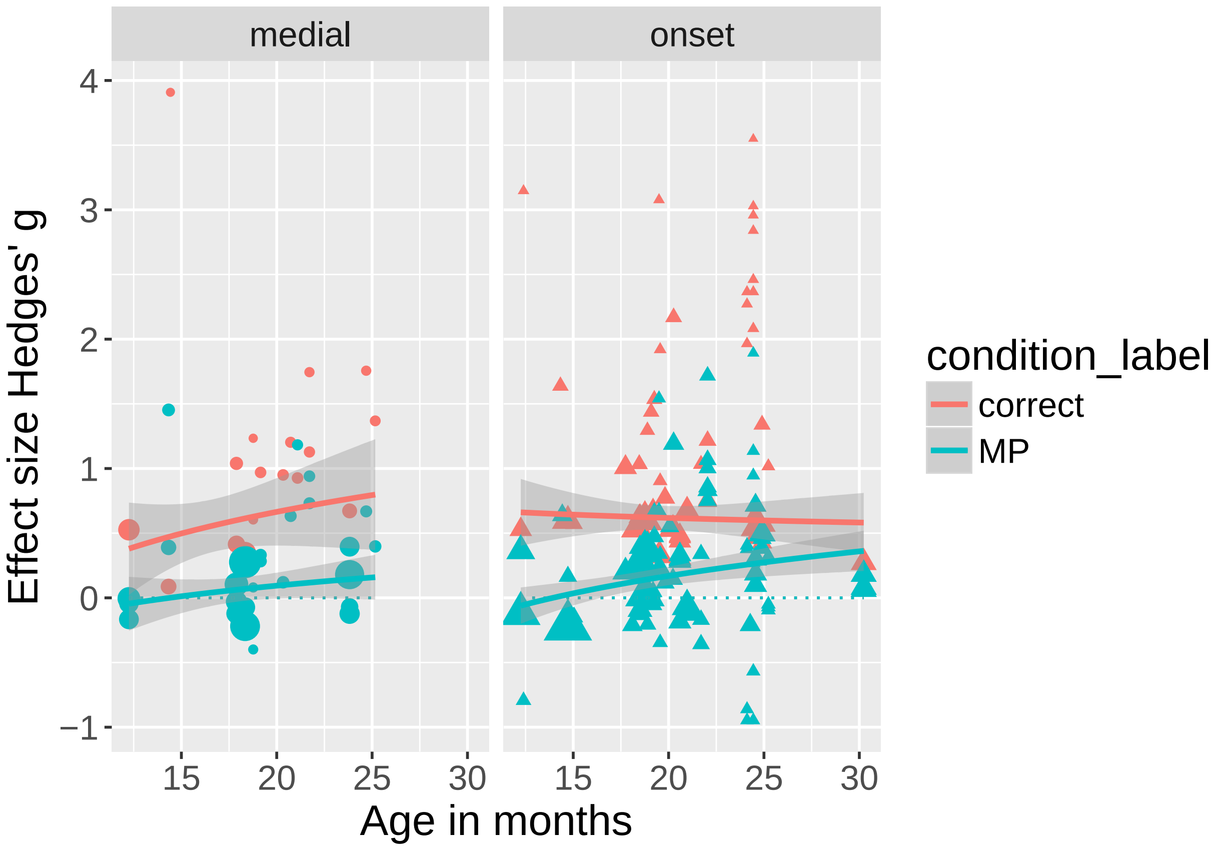
```
## Error: A continuous variable can not be mapped to shape
```

```r
ggsave("figures/AgeEffect_log_feat.jpg", p)
```

```
## Error: A continuous variable can not be mapped to shape
```

Position of Mispronunciation

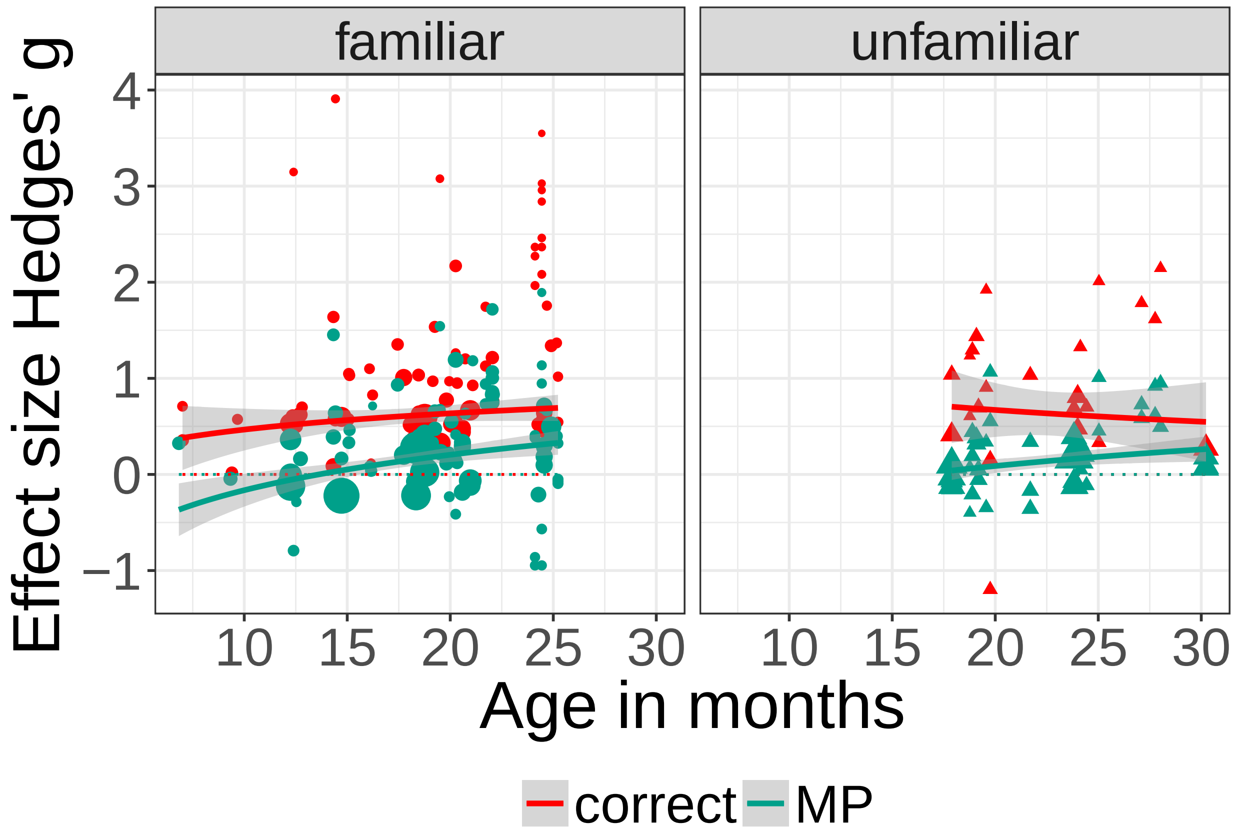
```r
dat.p <- subset(dat, mispron_location == "onset" | mispron_location == "medial" |
    mispron_location == "offset")

p <- ggplot(dat.p, aes(mean_age_1/30.44, g_calc, color = condition_label)) +
    geom_point(aes(size = weights_g, shape = mispron_location), show.legend = FALSE) +
    facet_grid(. ~ mispron_location) + geom_line(y = 0, linetype = "dotted") +
    geom_smooth(method = "lm", formula = y ~ log(x), aes(weight = weights_g)) +
    theme(text = element_text(size = 16)) + xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_position.jpg", p)
```

Distractor Familiarity

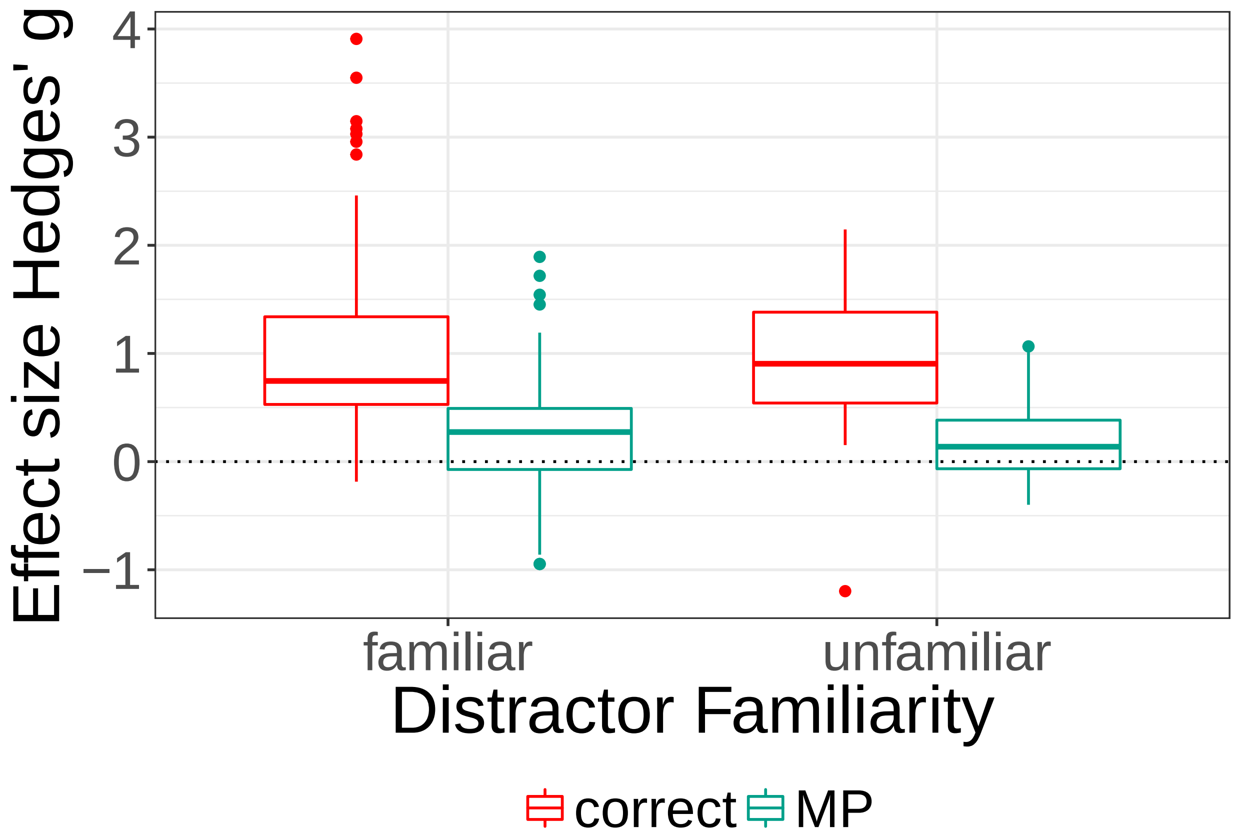
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = dist_code), show.legend = FALSE) + facet_grid(. ~ dist_code) + geom_line(y = 0,
    linetype = "dotted") + geom_smooth(method = "lm", formula = y ~ log(x),
    aes(weight = weights_g)) + scale_color_manual(values = wes_palette(name = "Darjeeling")) +
    theme_bw() + theme(text = element_text(size = 25), legend.title = element_blank(),
    legend.position = "bottom") + xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_distractor_fam.jpg", p)
```

Distractor Familiarity (w/o age)

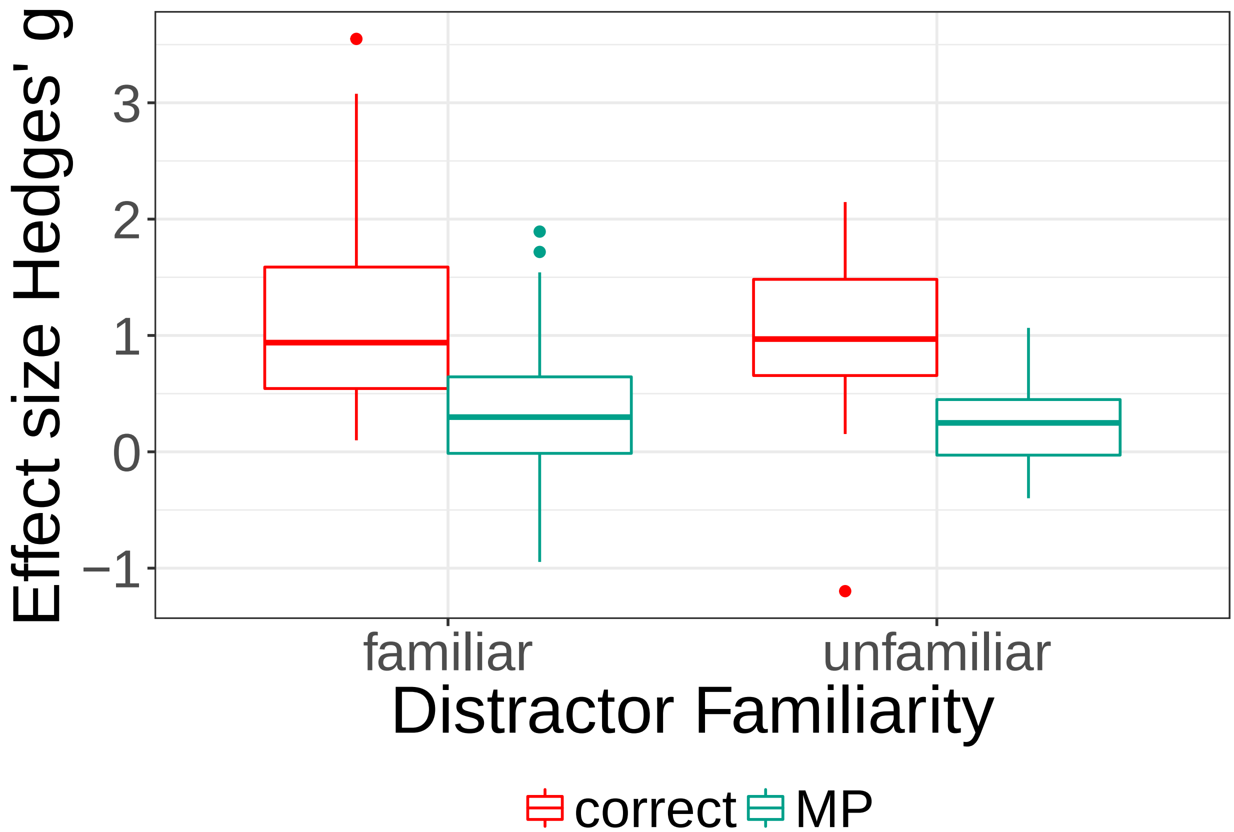
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

p <- ggplot(dat, aes(dist_code, g_calc, color = condition_label)) + geom_boxplot() +
    # geom_smooth(method = 'lm', formula = y ~ log(x), aes(weight=weights_g)) +
scale_color_manual(values = wes_palette(name = "Darjeeling")) + theme_bw() +
    theme(text = element_text(size = 25), legend.title = element_blank(), legend.position = "bottom") +
    xlab("Distractor Familiarity") + geom_hline(yintercept = 0, linetype = "dotted") +
    ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/Distractor_fam_log.jpg", p)
```

Distractor Familiarity (w/o age, subset to age range)

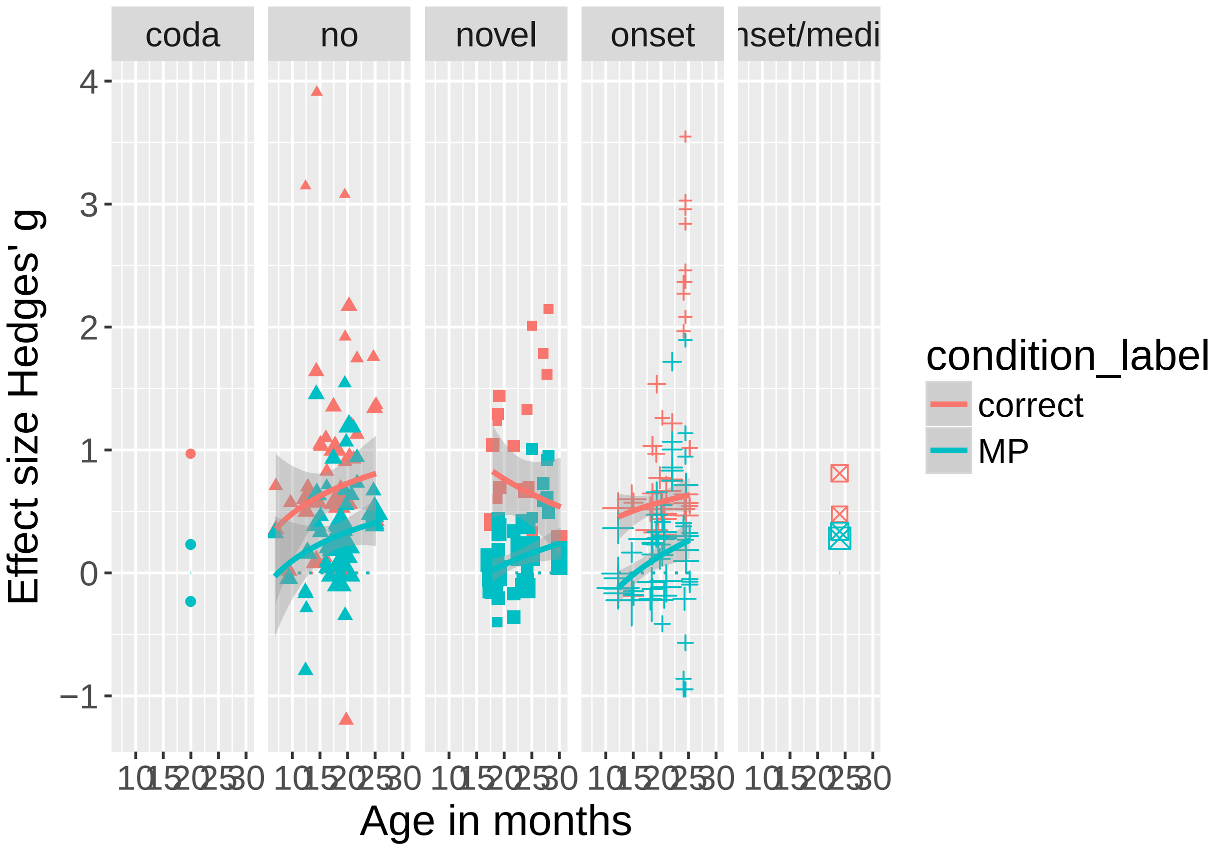
```
min_age <- min(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)
max_age <- max(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)

dat_age = dat %>% filter(mean_age_1 > min_age & mean_age_1 < max_age)

dat_age$condition_label = ifelse(dat_age$condition == 1, "correct", "MP")
dat_age$dist_code <- ifelse(dat_age$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

p <- ggplot(dat_age, aes(dist_code, g_calc, color = condition_label)) + geom_boxplot() +
    # geom_line(y= 0, linetype='dotted') + geom_smooth(method = 'lm', formula =
# y ~ log(x), aes(weight=weights_g)) +
scale_color_manual(values = wes_palette(name = "Darjeeling")) + theme_bw() +
    theme(text = element_text(size = 25), legend.title = element_blank(), legend.position = "bottom") +
    xlab("Distractor Familiarity") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeMatch_Distractor_fam_log.jpg", p)
```

Overlap between distractor and target

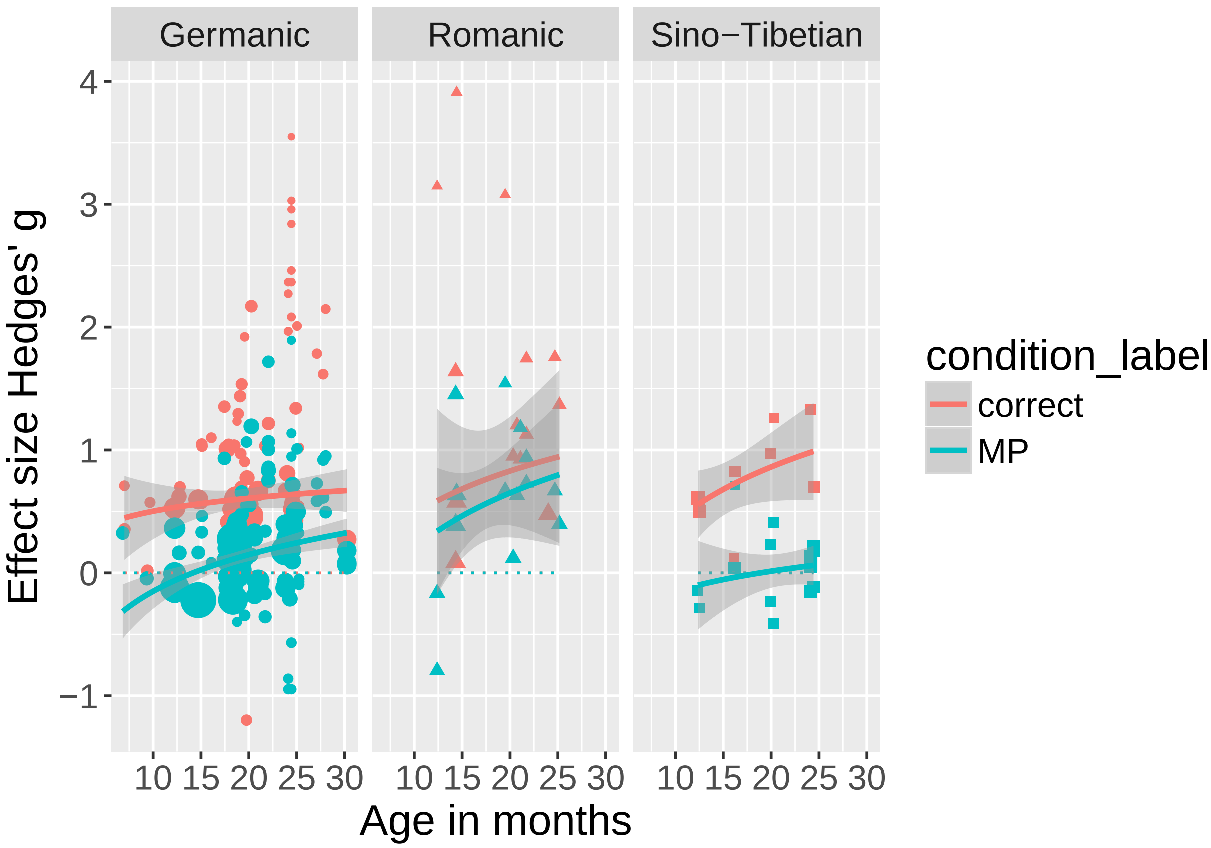
```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = distractor_overlap), show.legend = FALSE) + facet_grid(. ~ distractor_overlap) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g)) + theme(text = element_text(size = 16)) +
    xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_distractor_overlap.jpg", p)
```

Language Family

```
dat$lang_family = ifelse(dat$native_lang == "American English" | dat$native_lang ==
    "British English" | dat$native_lang == "Dutch" | dat$native_lang == "English" |
    dat$native_lang == "German", "Germanic", ifelse(dat$native_lang == "French" |
    dat$native_lang == "Catalan" | dat$native_lang == "Spanish" | dat$native_lang ==
    "Catalan-Spanish" | dat$native_lang == "Swiss French", "Romanic", "Sino-Tibetian"))

p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = lang_family), show.legend = FALSE) + facet_grid(. ~ lang_family) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g)) + theme(text = element_text(size = 16)) +
    xlab("Age in months") + ylab("Effect size Hedges' g")
p
```




```
ggsave("figures/AgeEffect_log_language.jpg", p)
```