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 <- time_wind_dat %>% group_by(offset) %>% summarize(count = n())


kable(offset_info)
```

offset	count
0	3
200	1
231	1
267	1
300	1
360	5
365	1
367	14
400	1
500	1
1133	1
NA	4

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)
```

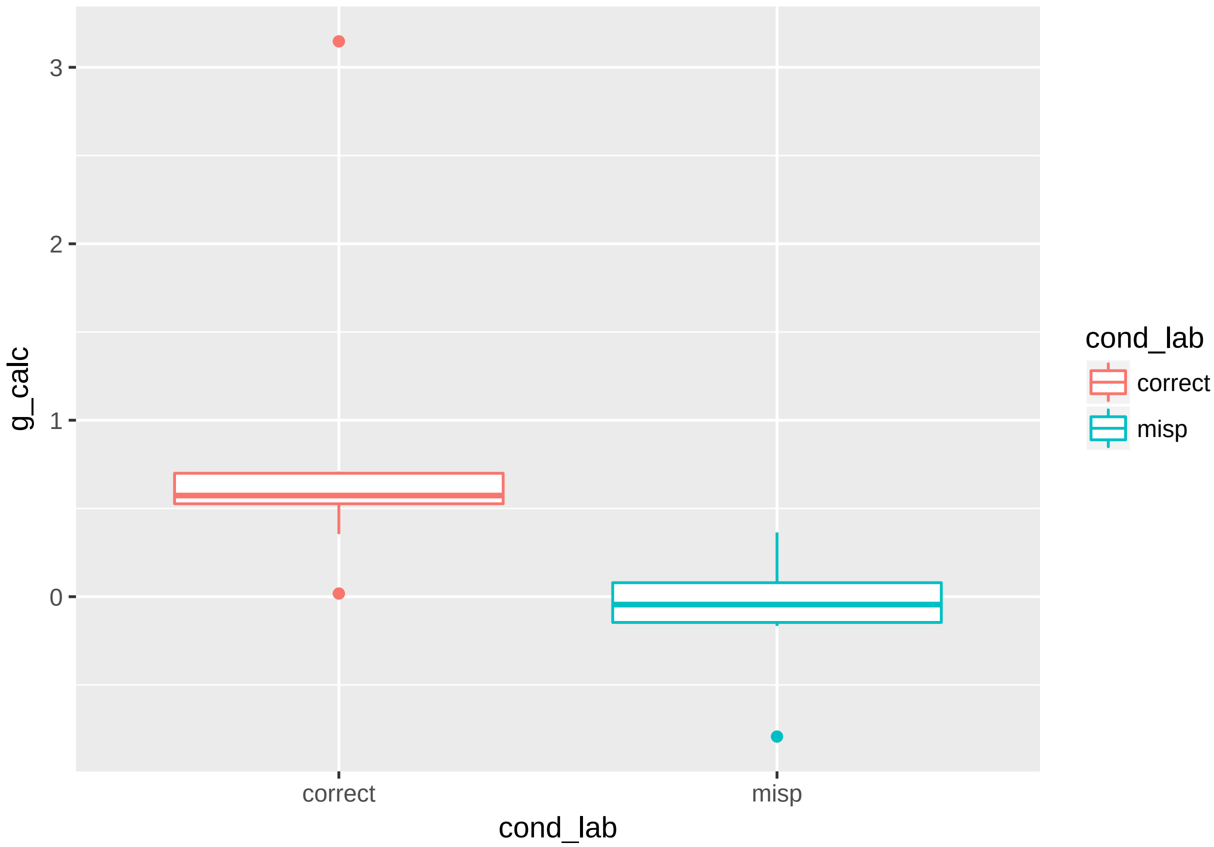
post_nam_dur	count
1.510	2
2.000	45
2.500	18
2.600	4
2.750	4
2.767	1
2.805	4
3.000	13
3.500	6
4.000	6
6.000	1

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
## -111.8857   223.7713   229.7713   237.6755   230.0137
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##              estim    sqrt  fixed
```

```
## tau^2      0.4483  0.6696     no
## rho        0.8886             no
##
## Test for Heterogeneity:
## Q(df = 103) = 625.6267, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.9078   0.1198   7.5784   <.0001   0.6730   1.1426      ***
##
## ---
## 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
## -110.8134   221.6268   229.6268   240.1267   230.0392
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.4458  0.6677     no
## rho        0.8835             no
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 619.1502, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.6778, p-val = 0.4103
##
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.9202  0.1203  7.6515  <.0001   0.6845  1.1559  ***
## age.C      0.0145  0.0176  0.8233  0.4103  -0.0200  0.0490
##
## ---
## 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
## -70.1217  140.2434  146.2434  155.1942  146.4124
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1192  0.3453     no
## rho        0.5924             no
##
## Test for Heterogeneity:
## Q(df = 146) = 462.5143, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.2498   0.0597   4.1835   <.0001   0.1328   0.3668      ***
##
## ---
## 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
## -68.8541  137.7083  145.7083  157.6152  145.9940
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1181  0.3437     no
## rho        0.5830             no
##
## Test for Residual Heterogeneity:
## QE(df = 145) = 449.1871, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 1.7151, p-val = 0.1903
##
```

```
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.2613  0.0599  4.3583  <.0001   0.1438  0.3788  ***
## age.C      0.0149  0.0114  1.3096  0.1903  -0.0074  0.0372
##
## ---
## 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
## -252.9095   505.8189   513.8189   527.8887   513.9829
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1371  0.3703     no
## rho        0.7381             no
##
## Test for Residual Heterogeneity:
## QE(df = 249) = 1088.1411, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 215.7609, p-val < .0001
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## intrcpt      0.2792  0.0652   4.2827  <.0001  0.1514  0.4069  ***
## condition    0.4953  0.0337  14.6888  <.0001  0.4293  0.5614  ***
##
## ---
## 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
## -261.1359   522.2718   528.2718   538.8362   528.3694
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.2069  0.4549     no
## rho        0.8295             no
##
## Test for Residual Heterogeneity:
## QE(df = 250) = 1154.4618, p-val < .0001
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## condition    0.5139  0.0333  15.4186  <.0001  0.4486  0.5793  ***
##
## ---
## 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
## -251.2299   502.4597   514.4597   535.5160   514.8097
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1331  0.3648     no
## rho        0.7254             no
##
## Test for Residual Heterogeneity:
## QE(df = 247) = 1068.3373, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 218.6210, p-val < .0001
##
```

```
## Model Results:
##
##                   estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt             0.2935  0.0648   4.5324  <.0001   0.1666  0.4204  ***
## age.C               0.0171  0.0113   1.5136  0.1301  -0.0051  0.0393
## condition           0.4984  0.0344  14.4930  <.0001   0.4310  0.5658  ***
## age.C:condition     0.0026  0.0076   0.3436  0.7312  -0.0123  0.0175
##
## ---
## 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", "Misp")

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
## -245.9822   491.9645   519.9645   568.6350   521.8395
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1334  0.3653     no
## rho        0.7359             no
##
## Test for Residual Heterogeneity:
## QE(df = 239) = 998.1810, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 225.6133, p-val < .0001
##
## Model Results:
##
##                                          estimate      se     zval
## intrcpt                                    0.2813  0.0713   3.9440
## age.C                                      0.0124  0.0124   1.0028
```

```
## condition                                      0.4795  0.0365  13.1408
## lang_familyRomanic                             0.2318  0.1881   1.2323
## lang_familySino-Tibetan                       -0.2461  0.2491  -0.9879
## age.C:condition                                0.0024  0.0082   0.2884
## age.C:lang_familyRomanic                       0.0419  0.0352   1.1900
## age.C:lang_familySino-Tibetan                 -0.0012  0.0461  -0.0264
## condition:lang_familyRomanic                   0.0558  0.1366   0.4083
## condition:lang_familySino-Tibetan              0.2910  0.1877   1.5506
## age.C:condition:lang_familyRomanic            -0.0116  0.0295  -0.3940
## age.C:condition:lang_familySino-Tibetan        0.0215  0.0336   0.6405
##                                                pval    ci.lb   ci.ub
## intrcpt                                      <.0001   0.1415  0.4210  ***
## age.C                                        0.3159  -0.0118  0.0367
## condition                                    <.0001   0.4080  0.5510  ***
## lang_familyRomanic                           0.2178  -0.1369  0.6004
## lang_familySino-Tibetan                      0.3232  -0.7343  0.2421
## age.C:condition                              0.7730  -0.0137  0.0184
## age.C:lang_familyRomanic                     0.2341  -0.0271  0.1108
## age.C:lang_familySino-Tibetan                0.9790  -0.0915  0.0891
## condition:lang_familyRomanic                 0.6831  -0.2120  0.3236
## condition:lang_familySino-Tibetan            0.1210  -0.0768  0.6588
## age.C:condition:lang_familyRomanic           0.6935  -0.0694  0.0462
## age.C:condition:lang_familySino-Tibetan      0.5219  -0.0443  0.0874
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Type of distractor

```r
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
## -247.3148   494.6296   514.6296   549.5602   515.5778
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1357  0.3684     no
## rho        0.7175             no
##
## Test for Residual Heterogeneity:
## QE(df = 243) = 1064.6022, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
```

```
## QM(df = 7) = 224.9573, p-val < .0001
##
## Model Results:
##
##                                                      estimate      se
## intrcpt                                                0.3698  0.0785
## age.C                                                  0.0242  0.0138
## condition                                              0.4666  0.0415
## as.factor(object_pair)familiar_novel                  -0.2541  0.1471
## age.C:condition                                        0.0020  0.0092
## age.C:as.factor(object_pair)familiar_novel             0.0038  0.0288
## condition:as.factor(object_pair)familiar_novel         0.1755  0.0894
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0203  0.0198
##                                                          zval    pval
## intrcpt                                                4.7107  <.0001
## age.C                                                  1.7481  0.0804
## condition                                             11.2325  <.0001
## as.factor(object_pair)familiar_novel                  -1.7273  0.0841
## age.C:condition                                        0.2153  0.8295
## age.C:as.factor(object_pair)familiar_novel             0.1312  0.8956
## condition:as.factor(object_pair)familiar_novel         1.9637  0.0496
## age.C:condition:as.factor(object_pair)familiar_novel  -1.0267  0.3046
##                                                         ci.lb   ci.ub
## intrcpt                                                0.2160  0.5237  ***
## age.C                                                 -0.0029  0.0512    .
## condition                                              0.3852  0.5480  ***
## as.factor(object_pair)familiar_novel                  -0.5425  0.0342    .
## age.C:condition                                       -0.0161  0.0201
## age.C:as.factor(object_pair)familiar_novel            -0.0526  0.0602
## condition:as.factor(object_pair)familiar_novel         0.0003  0.3507    *
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0590  0.0184
##
## ---
## 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)

summary(rma_DistractorAgeS)
```

```
##
## Multivariate Meta-Analysis Model (k = 185; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -181.6354   363.2707   383.2707   415.0322   384.5960
##
## Variance Components:
```

```
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 38)
##
##             estim    sqrt  fixed
## tau^2      0.1852  0.4303     no
## rho        0.7698             no
##
## Test for Residual Heterogeneity:
## QE(df = 177) = 824.6499, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 157.4337, p-val < .0001
##
## Model Results:
##
##                                                      estimate      se
## intrcpt                                                0.4127  0.1035
## age.C                                                 -0.0119  0.0261
## condition                                              0.4086  0.0465
## as.factor(object_pair)familiar_novel                  -0.3230  0.2025
## age.C:condition                                        0.0447  0.0188
## age.C:as.factor(object_pair)familiar_novel             0.0502  0.0538
## condition:as.factor(object_pair)familiar_novel         0.1987  0.1100
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0203  0.0342
##                                                         zval    pval
## intrcpt                                               3.9865  <.0001
## age.C                                                -0.4578  0.6471
## condition                                             8.7811  <.0001
## as.factor(object_pair)familiar_novel                 -1.5949  0.1107
## age.C:condition                                       2.3724  0.0177
## age.C:as.factor(object_pair)familiar_novel            0.9326  0.3510
## condition:as.factor(object_pair)familiar_novel        1.8068  0.0708
## age.C:condition:as.factor(object_pair)familiar_novel -0.5931  0.5531
##                                                        ci.lb   ci.ub
## intrcpt                                               0.2098  0.6156  ***
## age.C                                                -0.0631  0.0392
## condition                                             0.3174  0.4998  ***
## as.factor(object_pair)familiar_novel                 -0.7198  0.0739
## age.C:condition                                       0.0078  0.0816    *
## age.C:as.factor(object_pair)familiar_novel           -0.0553  0.1557
## condition:as.factor(object_pair)familiar_novel       -0.0168  0.4142    .
## age.C:condition:as.factor(object_pair)familiar_novel -0.0873  0.0468
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Distractor Overlap

```
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
## -63.8569  127.7138  147.7138  177.0586  149.4326
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1272  0.3567     no
## rho        0.5803             no
##
## Test for Residual Heterogeneity:
## QE(df = 139) = 426.8044, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 6.1553, p-val = 0.5217
##
## Model Results:
##
##                                 estimate      se     zval    pval    ci.lb
## intrcpt                           0.0983  0.3957   0.2483  0.8039  -0.6772
## age.C                             0.0174  0.0214   0.8130  0.4162  -0.0246
## distractor_overlapno              0.3432  0.4126   0.8319  0.4055  -0.4654
## distractor_overlapnovel          -0.0319  0.4197  -0.0759  0.9395  -0.8544
## distractor_overlaponset           0.1267  0.3979   0.3184  0.7502  -0.6532
## distractor_overlaponset/medial    0.1484  0.5553   0.2672  0.7893  -0.9399
## age.C:distractor_overlapno        0.0132  0.0297   0.4431  0.6577  -0.0451
## age.C:distractor_overlapnovel     0.0142  0.0342   0.4142  0.6787  -0.0529
##                                  ci.ub
## intrcpt                         0.8737
## age.C                           0.0594
## distractor_overlapno            1.1518
## distractor_overlapnovel         0.7907
## distractor_overlaponset         0.9066
## distractor_overlaponset/medial  1.2367
## age.C:distractor_overlapno      0.0714
## age.C:distractor_overlapnovel   0.0812
##
## ---
## 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),
```

```r
# 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
## -60.4794  120.9588  136.9588  159.6491  138.1896
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1229  0.3506     no
## rho        0.4838             no
##
## Test for Residual Heterogeneity:
## QE(df = 126) = 393.2688, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 6.9417, p-val = 0.2250
##
## Model Results:
##
##                              estimate      se     zval    pval    ci.lb
## intrcpt                        0.2767  0.0659   4.1984  <.0001   0.1475
## as.factor(n_feature)2         -0.0889  0.0804  -1.1054  0.2690  -0.2465
## as.factor(n_feature)3         -0.2339  0.1056  -2.2159  0.0267  -0.4409
## as.factor(n_feature)41640     -0.2278  0.2436  -0.9349  0.3498  -0.7053
## as.factor(n_feature)41641     -0.2088  0.1355  -1.5408  0.1234  -0.4743
## as.factor(n_feature)41672     -0.2767  0.3527  -0.7846  0.4327  -0.9680
##                                ci.ub
## intrcpt                       0.4059  ***
## as.factor(n_feature)2         0.0687
## as.factor(n_feature)3        -0.0270    *
## as.factor(n_feature)41640     0.2497
## as.factor(n_feature)41641     0.0568
## as.factor(n_feature)41672     0.4146
##
## ---
## 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
## -60.6887  121.3775  149.3775  188.4023  153.3775
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1305  0.3612     no
## rho        0.4368             no
##
## Test for Residual Heterogeneity:
## QE(df = 120) = 376.5908, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 8.3544, p-val = 0.6812
##
## Model Results:
##
##                                  estimate      se     zval    pval
## intrcpt                            0.2818  0.0668   4.2206  <.0001
## as.factor(n_feature)2             -0.0897  0.0813  -1.1032  0.2699
## as.factor(n_feature)3             -0.2222  0.1107  -2.0076  0.0447
## as.factor(n_feature)41640         -0.2750  0.2681  -1.0256  0.3051
## as.factor(n_feature)41641         -0.2265  0.1409  -1.6074  0.1080
## as.factor(n_feature)41672         -0.2818  0.8856  -0.3182  0.7503
## age.C                              0.0135  0.0149   0.9111  0.3622
## as.factor(n_feature)2:age.C        0.0021  0.0181   0.1153  0.9082
## as.factor(n_feature)3:age.C       -0.0067  0.0226  -0.2964  0.7669
## as.factor(n_feature)41640:age.C   -0.0180  0.0531  -0.3381  0.7353
## as.factor(n_feature)41641:age.C   -0.0311  0.0460  -0.6748  0.4998
## as.factor(n_feature)41672:age.C   -0.0135  1.6520  -0.0082  0.9935
##                                     ci.lb    ci.ub
## intrcpt                            0.1509   0.4127  ***
## as.factor(n_feature)2             -0.2491   0.0697
## as.factor(n_feature)3             -0.4392  -0.0053    *
## as.factor(n_feature)41640         -0.8005   0.2505
## as.factor(n_feature)41641         -0.5026   0.0497
## as.factor(n_feature)41672         -2.0176   1.4540
## age.C                             -0.0156   0.0427
## as.factor(n_feature)2:age.C       -0.0334   0.0376
## as.factor(n_feature)3:age.C       -0.0511   0.0377
## as.factor(n_feature)41640:age.C  -0.1221   0.0862
## as.factor(n_feature)41641:age.C  -0.1213   0.0592
## as.factor(n_feature)41672:age.C  -3.2514   3.2243
##
## ---
## 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
## -234.6537   469.3074   483.3074   506.6025   483.8730
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1530  0.3911     no
## rho        0.6938             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 980.4970, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5):
## QM(df = 4) = 184.5957, p-val < .0001
##
## Model Results:
##
##                       estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt                 0.5966  0.1346   4.4341  <.0001   0.3329   0.8604
## as.factor(n_feature)1  -0.3195  0.1130  -2.8277  0.0047  -0.5409  -0.0980
## as.factor(n_feature)2  -0.2848  0.1290  -2.2078  0.0273  -0.5377  -0.0320
## as.factor(n_feature)3  -0.5037  0.1462  -3.4456  0.0006  -0.7902  -0.2172
## condition               0.1906  0.1062   1.7949  0.0727  -0.0175   0.3987
##
## 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)
```

```r
summary(rma_NFeaturesAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 211; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -232.6365   465.2730   489.2730   528.9127   490.9326
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1581  0.3976     no
## rho        0.7224             no
##
## Test for Residual Heterogeneity:
## QE(df = 201) = 956.3669, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10):
## QM(df = 9) = 190.4816, p-val < .0001
##
## Model Results:
##
##                              estimate      se     zval    pval    ci.lb
## intrcpt                        0.6099  0.1361   4.4828  <.0001   0.3433
## as.factor(n_feature)1         -0.3219  0.1132  -2.8421  0.0045  -0.5438
## as.factor(n_feature)2         -0.2920  0.1293  -2.2593  0.0239  -0.5453
## as.factor(n_feature)3         -0.5182  0.1497  -3.4617  0.0005  -0.8116
## age.C                          0.0842  0.0506   1.6637  0.0962  -0.0150
## condition                      0.1919  0.1063   1.8056  0.0710  -0.0164
## as.factor(n_feature)1:age.C   -0.0691  0.0488  -1.4153  0.1570  -0.1648
## as.factor(n_feature)2:age.C   -0.0486  0.0510  -0.9533  0.3404  -0.1485
## as.factor(n_feature)3:age.C   -0.0535  0.0526  -1.0173  0.3090  -0.1566
## age.C:condition               -0.0648  0.0481  -1.3465  0.1782  -0.1591
##                                 ci.ub
## intrcpt                        0.8766  ***
## as.factor(n_feature)1         -0.0999   **
## as.factor(n_feature)2         -0.0387    *
## as.factor(n_feature)3         -0.2248  ***
## age.C                          0.1835    .
## condition                      0.4003    .
## as.factor(n_feature)1:age.C   0.0266
## as.factor(n_feature)2:age.C   0.0513
## as.factor(n_feature)3:age.C   0.0496
## age.C:condition                0.0295
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MP location

```
# table(db_ET_MP$mispron_location)

db_ET_MPl = 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_MPl,
    random = ~collapse | short_cite)


summary(rma_Location)
```

```
##
## Multivariate Meta-Analysis Model (k = 114; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -57.5043  115.0085  123.0085  133.8825  123.3823
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1502  0.3876     no
## rho        0.5421             no
##
## Test for Residual Heterogeneity:
## QE(df = 112) = 392.6421, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.0419, p-val = 0.8378
##
## Model Results:
##
##                          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                    0.2306  0.0852  2.7063  0.0068   0.0636  0.3977
## mispron_locationmedial     0.0307  0.1498  0.2048  0.8378  -0.2629  0.3243
##
## intrcpt                  **
## mispron_locationmedial
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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


summary(rma_LocationAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 114; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -56.0484  112.0967  124.0967  140.2996  124.9122
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1563  0.3953     no
## rho        0.5238             no
##
## Test for Residual Heterogeneity:
## QE(df = 110) = 386.0990, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 1.2243, p-val = 0.7472
##
## Model Results:
##
##                                 estimate      se    zval    pval    ci.lb
## intrcpt                           0.2296  0.0872  2.6339  0.0084   0.0588
## mispron_locationmedial            0.0832  0.1684  0.4937  0.6215  -0.2470
## age.C                             0.0117  0.0179  0.6531  0.5137  -0.0234
## mispron_locationmedial:age.C      0.0179  0.0337  0.5305  0.5958  -0.0482
##                                  ci.ub
## intrcpt                         0.4005  **
## mispron_locationmedial          0.4133
## age.C                           0.0469
## mispron_locationmedial:age.C  0.0840
##
## ---
## 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)


summary(rma_TypeFeaturesMP)
```

```
##
## Multivariate Meta-Analysis Model (k = 133; method: REML)
##
```

```
##   logLik  Deviance       AIC       BIC      AICc
## -64.0402  128.0804  136.0804  147.5812  136.3979
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1263  0.3553     no
## rho        0.5620             no
##
## Test for Residual Heterogeneity:
## QE(df = 131) = 427.6655, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.1467, p-val = 0.7017
##
## Model Results:
##
##                    estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt              0.2262  0.0729  3.1022  0.0019   0.0833  0.3691  **
## type_featurevowel    0.0338  0.0881  0.3830  0.7017  -0.1390  0.2065
##
## ---
## 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
## -62.8963  125.7927  137.7927  154.9515  138.4812
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1274  0.3570     no
## rho        0.5445             no
##
## Test for Residual Heterogeneity:
## QE(df = 129) = 415.3869, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 1.5441, p-val = 0.6721
##
## Model Results:
```

```
##
##                          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                    0.2283  0.0731  3.1237  0.0018   0.0851  0.3716
## type_featurevowel          0.0439  0.0889  0.4945  0.6210  -0.1302  0.2181
## age.C                      0.0143  0.0147  0.9676  0.3332  -0.0146  0.0431
## type_featurevowel:age.C    0.0008  0.0171  0.0484  0.9614  -0.0327  0.0344
##
## intrcpt                  **
## type_featurevowel
## age.C
## type_featurevowel:age.C
##
## ---
## 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
## -236.8091   473.6183   485.6183   506.0882   486.0054
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1238  0.3519     no
## rho        0.6901             no
##
## Test for Residual Heterogeneity:
## QE(df = 224) = 981.7485, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 154.6077, p-val < .0001
##
## Model Results:
##
##                                 estimate      se     zval    pval
## intrcpt                           0.7114  0.0688  10.3387  <.0001
```

```
## relevel(type_feature, "none")1   -0.4417  0.0423  -10.4486  <.0001
## relevel(type_feature, "none")4   -0.6356  0.1549   -4.1033  <.0001
## relevel(type_feature, "none")5   -0.4680  0.0565   -8.2812  <.0001
##                                   ci.lb    ci.ub
## intrcpt                          0.5765   0.8462  ***
## relevel(type_feature, "none")1  -0.5245  -0.3588  ***
## relevel(type_feature, "none")4  -0.9391  -0.3320  ***
## relevel(type_feature, "none")5  -0.5788  -0.3572  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
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
## -234.9545   469.9090   489.9090   523.8452   490.9616
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1260  0.3549     no
## rho        0.6767             no
##
## Test for Residual Heterogeneity:
## QE(df = 220) = 967.8211, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 158.2894, p-val < .0001
##
## Model Results:
##
##                                       estimate      se      zval    pval
## intrcpt                                 0.7276  0.0702   10.3680  <.0001
## relevel(type_feature, "none")1         -0.4489  0.0427  -10.5083  <.0001
## relevel(type_feature, "none")4         -0.6202  0.1703   -3.6419  0.0003
## relevel(type_feature, "none")5         -0.4874  0.0630   -7.7322  <.0001
## age.C                                   0.0161  0.0124    1.2981  0.1942
## relevel(type_feature, "none")1:age.C    0.0076  0.0104    0.7309  0.4648
## relevel(type_feature, "none")4:age.C    0.0055  0.0311    0.1770  0.8595
## relevel(type_feature, "none")5:age.C   -0.0082  0.0114   -0.7146  0.4748
##                                         ci.lb    ci.ub
## intrcpt                                0.5901   0.8652  ***
## relevel(type_feature, "none")1        -0.5327  -0.3652  ***
## relevel(type_feature, "none")4        -0.9540  -0.2864  ***
## relevel(type_feature, "none")5        -0.6110  -0.3639  ***
```

```
## age.C                                -0.0082   0.0405
## relevel(type_feature, "none")1:age.C  -0.0128   0.0279
## relevel(type_feature, "none")4:age.C  -0.0555   0.0665
## relevel(type_feature, "none")5:age.C  -0.0306   0.0142
##
## ---
## 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
## -226.0585   452.1170   468.1170   494.7400   468.8480
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1293  0.3596     no
## rho        0.5788             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 893.9789, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 158.2471, p-val < .0001
```

```
##
## Model Results:
##
##                                                  estimate      se
## intrcpt                                            0.6597  0.0777
## relevel(type_feature, "none")1                    -0.4135  0.0441
## relevel(type_feature, "none")5                    -0.4830  0.0640
## lang_familyRomanic                                 0.4502  0.1801
## relevel(type_feature, "none")1:lang_familyRomanic  -0.6549  0.2157
## relevel(type_feature, "none")5:lang_familyRomanic   0.0924  0.1490
##                                                      zval    pval
## intrcpt                                            8.4880  <.0001
## relevel(type_feature, "none")1                    -9.3845  <.0001
## relevel(type_feature, "none")5                    -7.5453  <.0001
## lang_familyRomanic                                 2.4991  0.0124
## relevel(type_feature, "none")1:lang_familyRomanic  -3.0359  0.0024
## relevel(type_feature, "none")5:lang_familyRomanic   0.6202  0.5351
##                                                     ci.lb    ci.ub
## intrcpt                                            0.5073   0.8120  ***
## relevel(type_feature, "none")1                    -0.4998  -0.3271  ***
## relevel(type_feature, "none")5                    -0.6084  -0.3575  ***
## lang_familyRomanic                                 0.0971   0.8032    *
## relevel(type_feature, "none")1:lang_familyRomanic  -1.0777  -0.2321   **
## relevel(type_feature, "none")5:lang_familyRomanic  -0.1996   0.3843
##
## ---
## 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)



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
## -226.0585   452.1170   468.1170   494.7400   468.8480
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1293  0.3596     no
## rho        0.5788             no
##
## Test for Residual Heterogeneity:
## QE(df = 206) = 893.9789, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 158.2471, p-val < .0001
##
## Model Results:
##
##                                                  estimate      se
## intrcpt                                            0.6597  0.0777
## relevel(type_feature, "none")1                    -0.4135  0.0441
## relevel(type_feature, "none")5                    -0.4830  0.0640
## lang_familyRomanic                                 0.4502  0.1801
## relevel(type_feature, "none")1:lang_familyRomanic  -0.6549  0.2157
## relevel(type_feature, "none")5:lang_familyRomanic   0.0924  0.1490
##                                                     zval    pval
## intrcpt                                           8.4880  <.0001
## relevel(type_feature, "none")1                   -9.3845  <.0001
## relevel(type_feature, "none")5                   -7.5453  <.0001
## lang_familyRomanic                                2.4991  0.0124
## relevel(type_feature, "none")1:lang_familyRomanic -3.0359  0.0024
## relevel(type_feature, "none")5:lang_familyRomanic  0.6202  0.5351
##                                                    ci.lb   ci.ub
## intrcpt                                           0.5073   0.8120  ***
## relevel(type_feature, "none")1                   -0.4998  -0.3271  ***
## relevel(type_feature, "none")5                   -0.6084  -0.3575  ***
## lang_familyRomanic                                0.0971   0.8032    *
## relevel(type_feature, "none")1:lang_familyRomanic -1.0777  -0.2321   **
## relevel(type_feature, "none")5:lang_familyRomanic -0.1996   0.3843
##
## ---
## 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	count
comprehension	12
none	87
production	5

We have 17 correlations, roughly evenly divided between comprehension and production data. There is reason to believe that production data are different from comprehension data (the former being easier to estimate for parents in the typical questionnaire-based assessment), so we should both analyze this data separately and see whether it makes sense in a joint analysis.

```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
```

```r
# 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
```

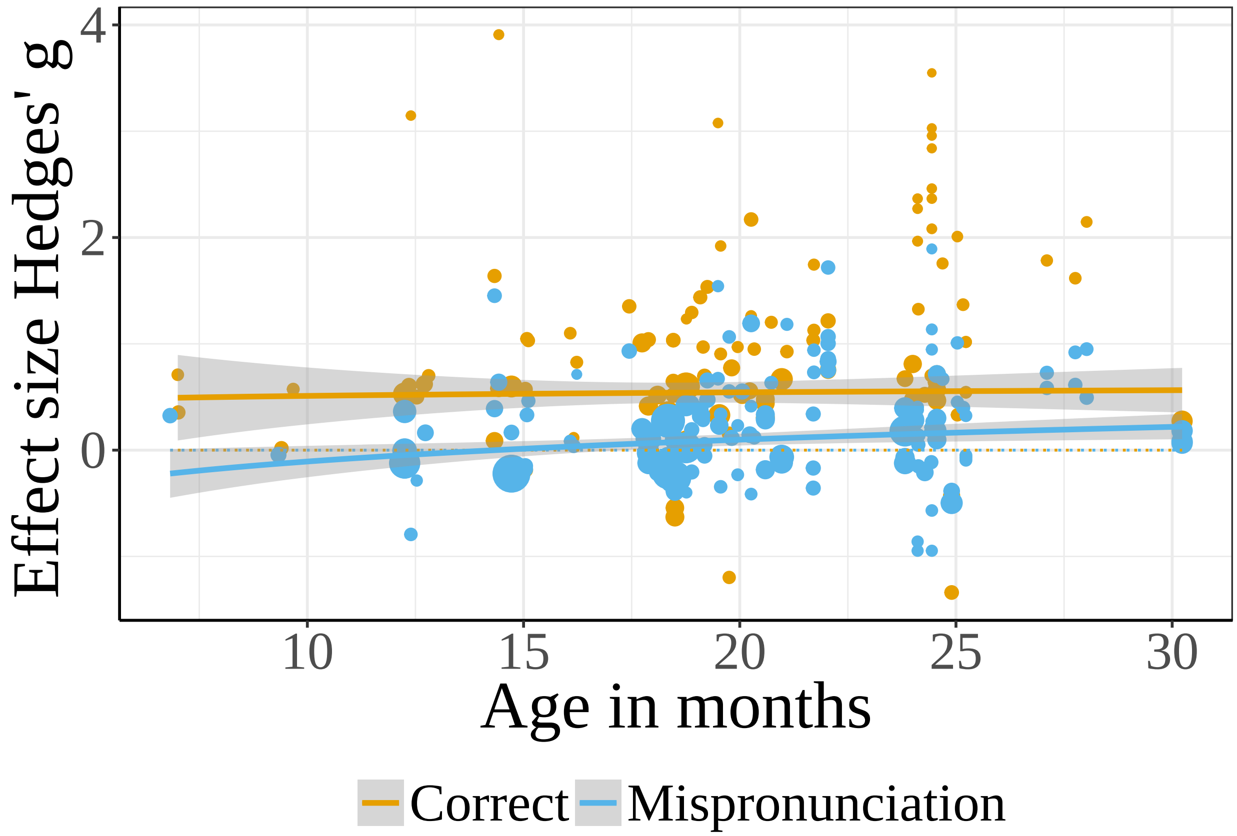
```
##
## Number of studies combined: k = 16
##
##                         COR            95%-CI    z p-value
## Fixed effect model   0.0601 [-0.0331; 0.1533] 1.26  0.2061
## Random effects model 0.0601 [-0.0331; 0.1533] 1.26  0.2061
##
## Quantifying heterogeneity:
## tau^2 = 0; H = 1.00 [1.00; 1.42]; 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

```
#Themes and plot
apatheme=theme_bw()+
  theme(#panel.grid.major=element_blank(),
    #panel.grid.minor=element_blank(),
    #panel.border=element_blank(),
    axis.line=element_line(),
    text=element_text(family='Times', size=25))
```

Mispronunciation Effect by Age (color)

```
dat$condition_label = ifelse(dat$condition == 1, "Correct", "Mispronunciation")

# 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) +
    apatheme + theme(legend.title = element_blank(), legend.position = "bottom") +
    xlab("Age in months") + ylab("Effect size Hedges' g")

p
```




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

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

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

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

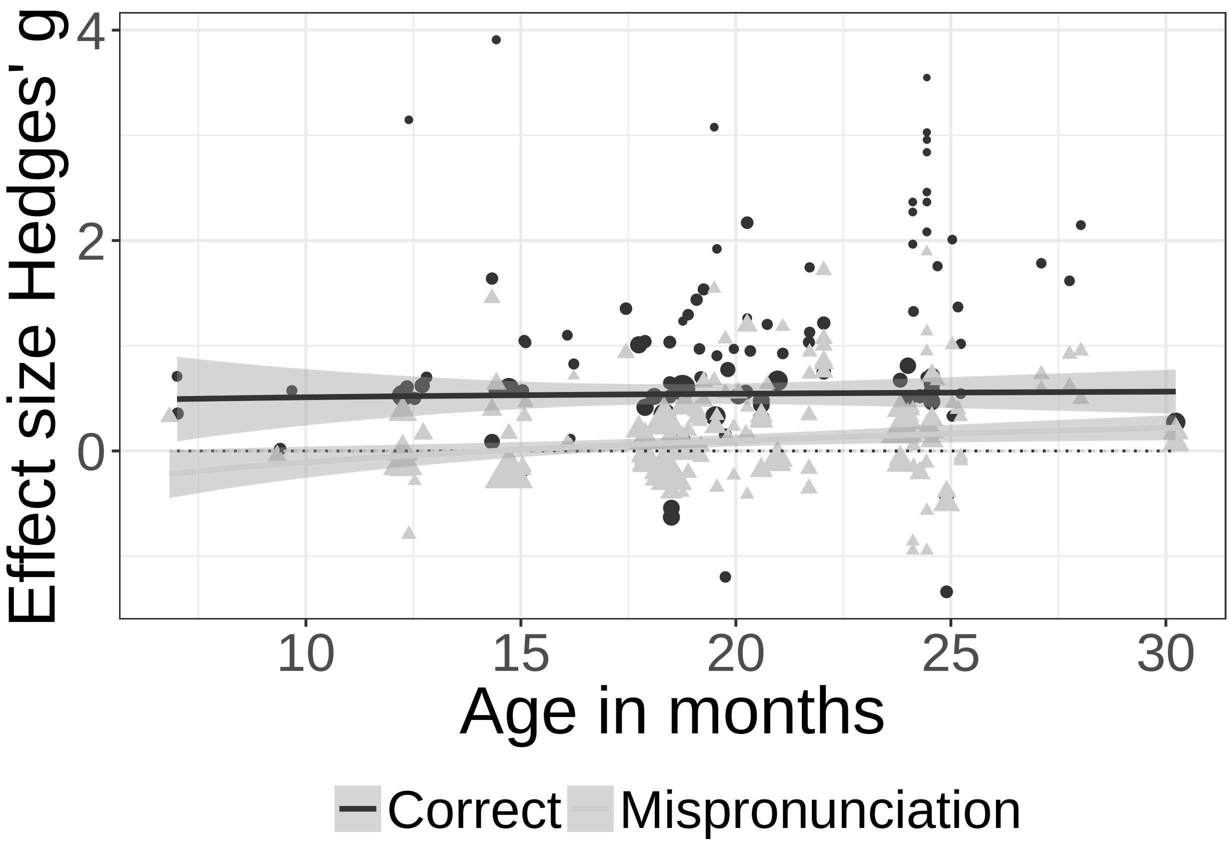
```r
# ggsave('figures/AgeEffect_log.jpg', p,height= 7,width= 6)



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

p

dev.off()
```

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

Mispronunciation Effect by Age (bw)

```r
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") +
```

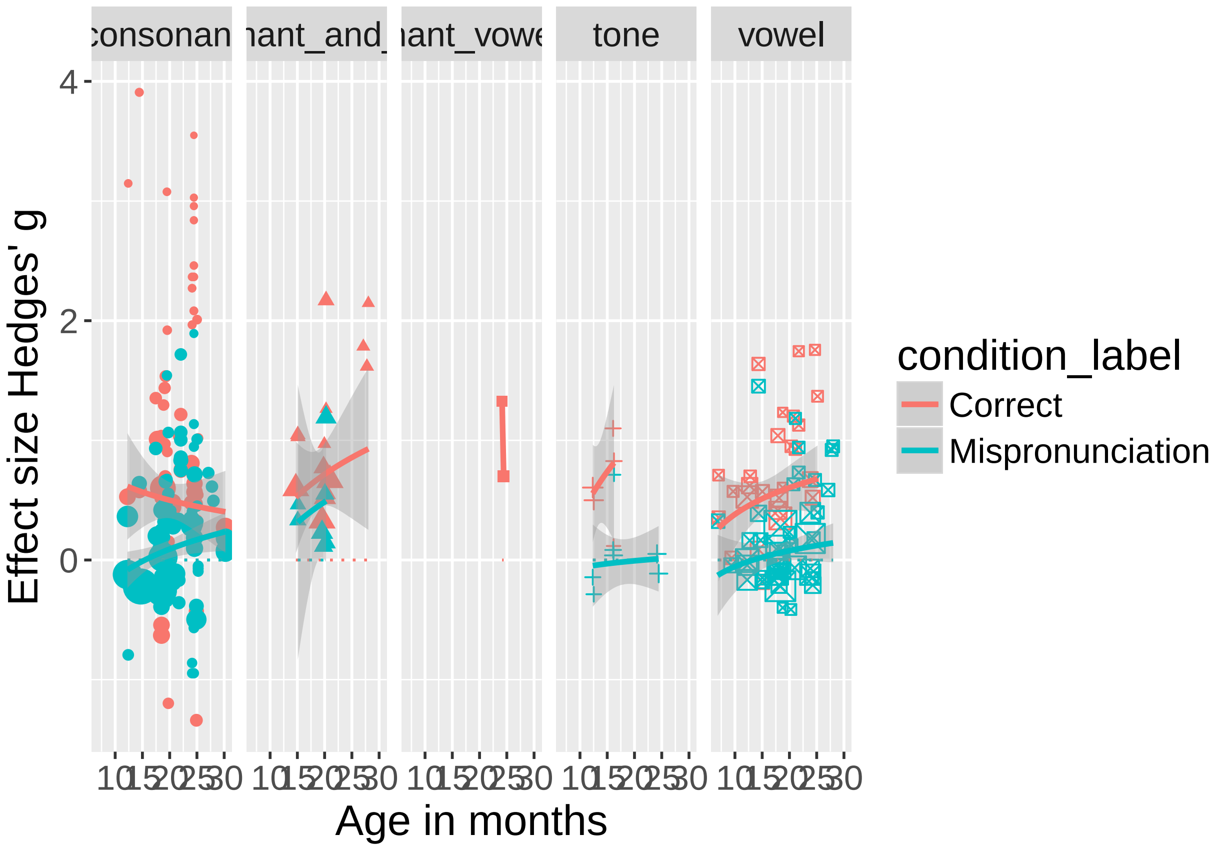
```
    ylab("Effect size Hedges' g")
p
```




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

MP type: Consonant, Vowel, or Tone?

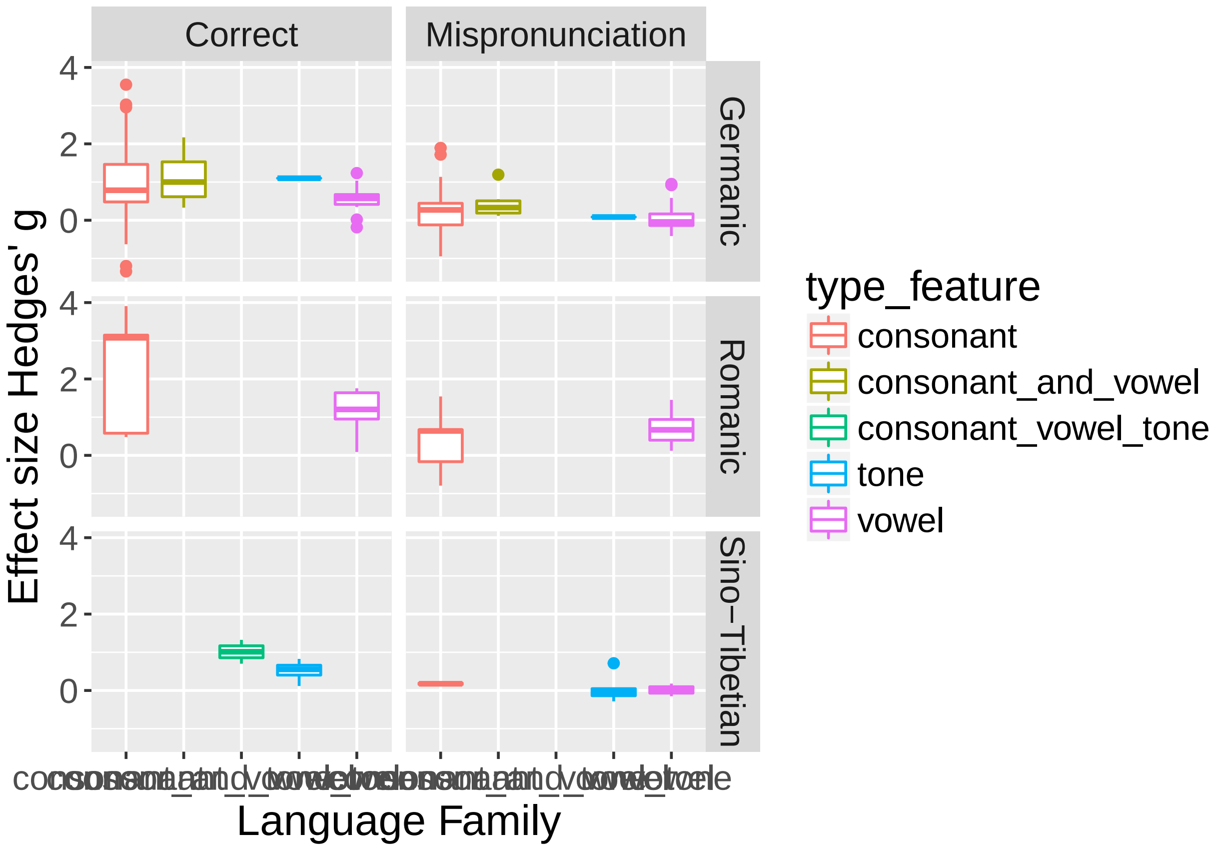
```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = type_feature), show.legend = FALSE) + facet_grid(. ~ type_feature) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g))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




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

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

```r
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
```

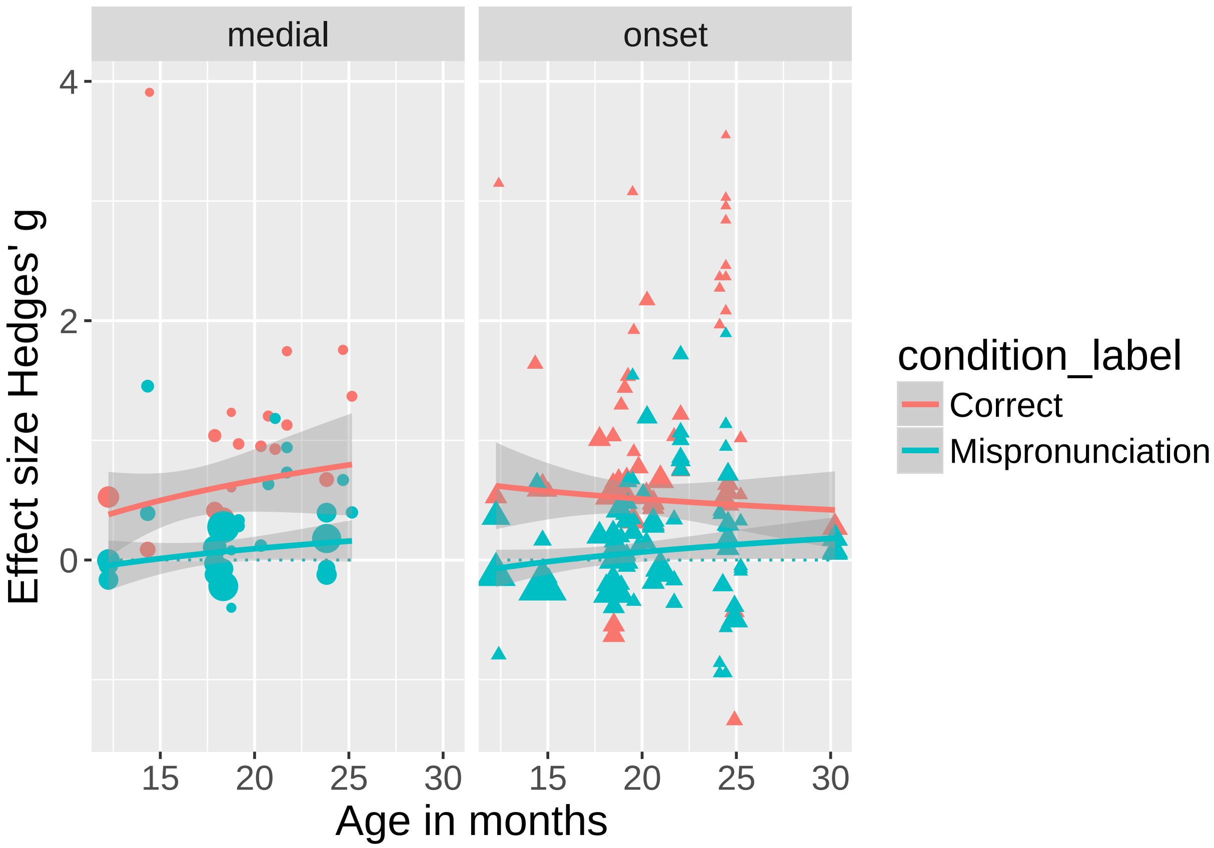
```
## Error: A continuous variable can not be mapped to shape
```

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

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

Position of Mispronunciation

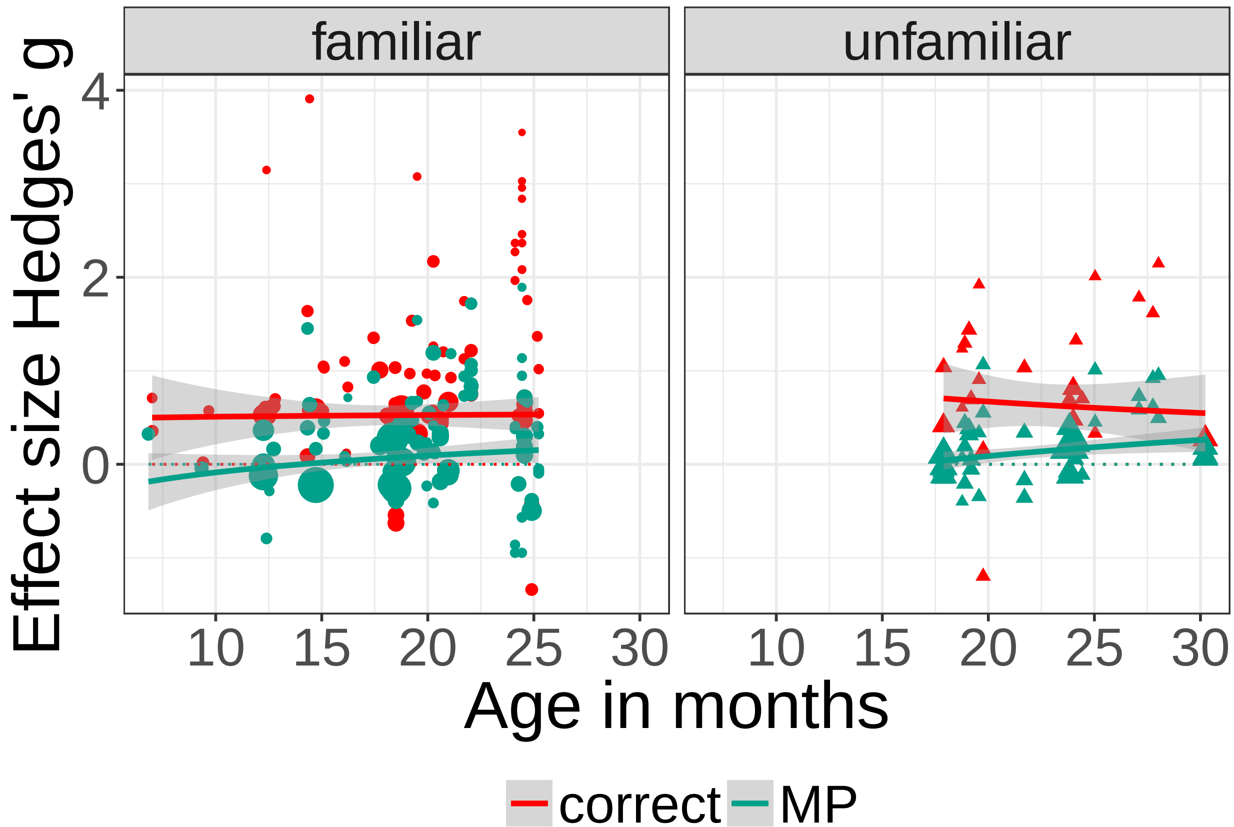
```r
dat.p <- subset(dat, mispron_location == "onset" | mispron_location == "medial" |
    mispron_location == "offset")

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

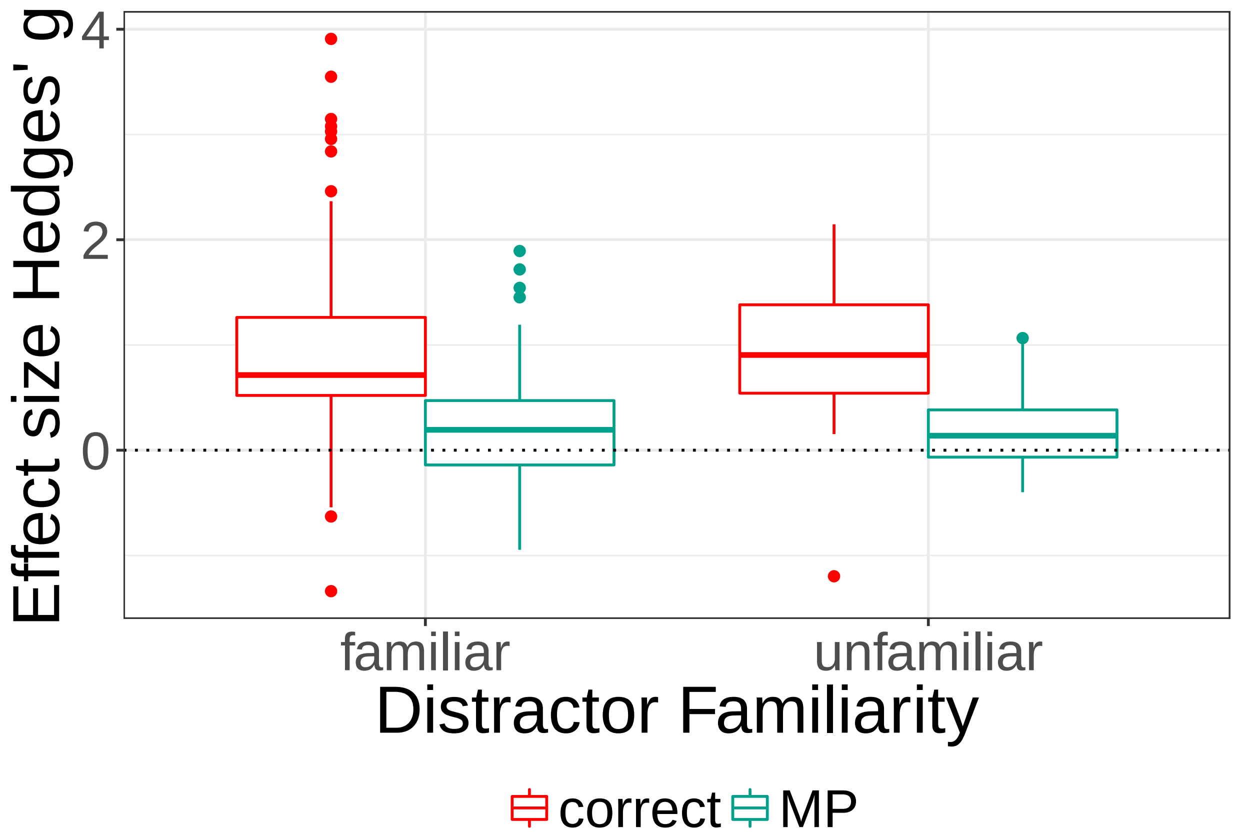
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

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)

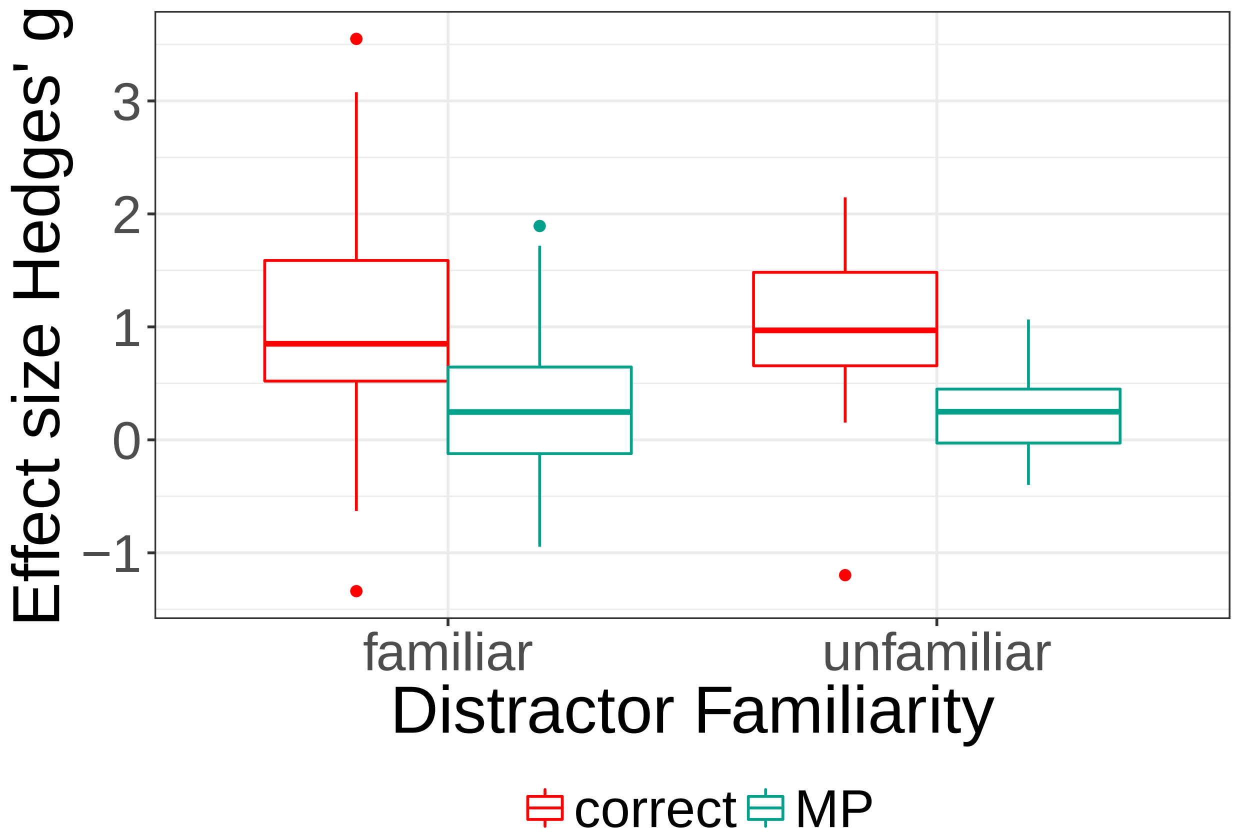
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

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)
```

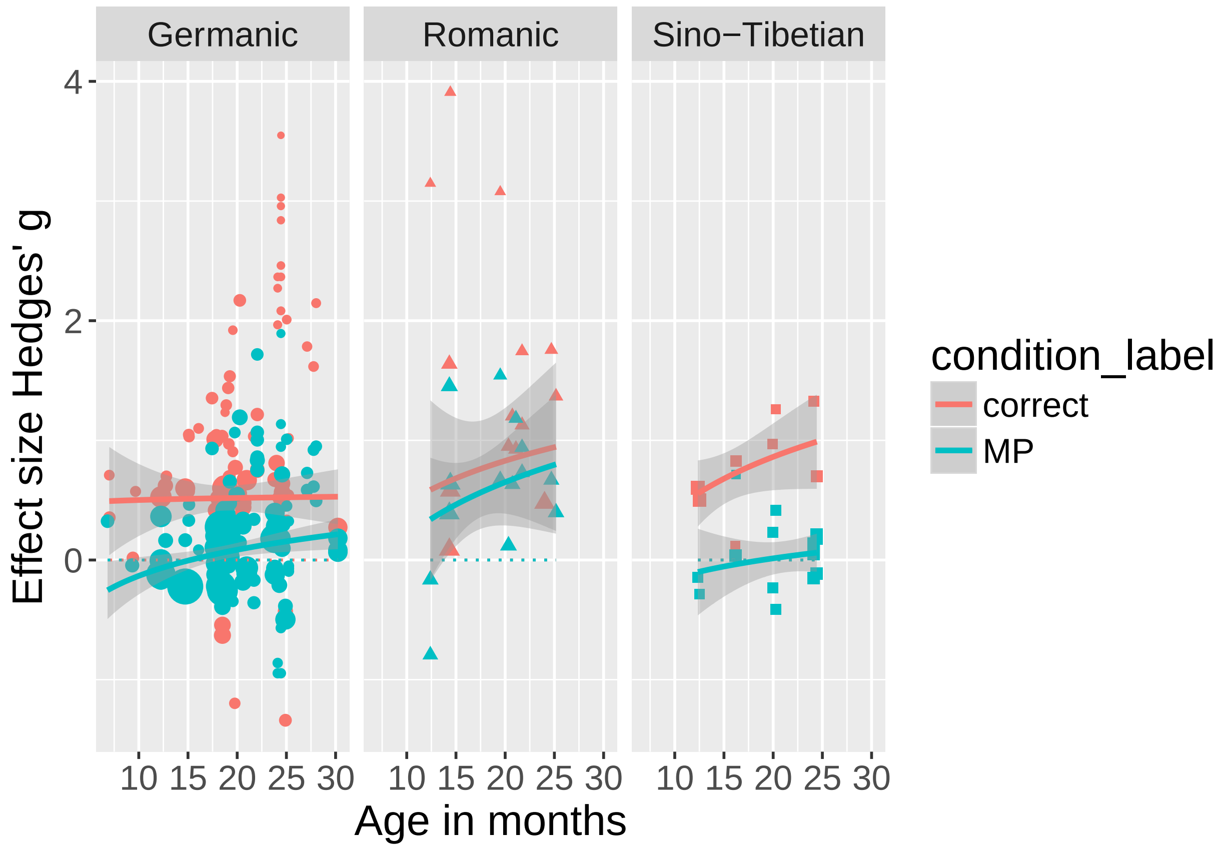
Overlap between distractor and target

```
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)
```