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. 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	207
proceedings	1	4

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

es_method	n_unique	count
f_two	1	3
group_means_one	18	120
group_means_two	6	45
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	27
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	3
231	4
267	1
300	2
360	25
365	10
367	37
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
2.400	2
2.500	18
2.600	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.

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 = 102; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -105.4985   210.9970   216.9970   224.8423   217.2444
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.3546  0.5955     no
## rho        0.8505             no
##
## Test for Heterogeneity:
## Q(df = 101) = 465.1968, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.9455   0.1067   8.8573   <.0001   0.7362   1.1547      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
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 = 102; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -103.8629   207.7257   215.7257   226.1464   216.1468
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.3393  0.5825     no
## rho        0.8406             no
##
## Test for Residual Heterogeneity:
## QE(df = 100) = 449.1111, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 1.8748, p-val = 0.1709
##
## Model Results:
```

```
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.9670  0.1057  9.1529  <.0001   0.7600  1.1741  ***
## age.C      0.0235  0.0172  1.3692  0.1709  -0.0101  0.0571
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Mispronunciation object identification effect

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

summary(rma_MP)
```

```
##
## Multivariate Meta-Analysis Model (k = 140; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -69.1013  138.2026  144.2026  153.0060  144.3804
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1082  0.3290     no
## rho        0.5446             no
##
## Test for Heterogeneity:
## Q(df = 139) = 409.9230, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##   0.2725   0.0569   4.7885   <.0001   0.1610   0.3841      ***
##
## ---
## 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 = 140; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -67.4238  134.8475  142.8475  154.5565  143.1483
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
```

```
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1046  0.3234     no
## rho        0.5407             no
##
## Test for Residual Heterogeneity:
## QE(df = 138) = 390.4218, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 2.6389, p-val = 0.1043
##
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.2871  0.0568  5.0555  <.0001   0.1758  0.3984  ***
## age.C      0.0178  0.0110  1.6245  0.1043  -0.0037  0.0394
##
## ---
## 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 = 242; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -226.1286   452.2572   460.2572   474.1797   460.4274
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1120  0.3347     no
## rho        0.6716             no
##
## Test for Residual Heterogeneity:
## QE(df = 240) = 875.1198, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 252.2219, p-val < .0001
```

```
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## intrcpt      0.2845  0.0590   4.8220  <.0001  0.1689  0.4001  ***
## condition    0.5478  0.0345  15.8815  <.0001  0.4802  0.6154  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

summary(rma_MPeffect_1)
```

```
##
## Multivariate Meta-Analysis Model (k = 242; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -235.7483   471.4966   477.4966   487.9510   477.5979
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.1858  0.4311     no
## rho        0.8042             no
##
## Test for Residual Heterogeneity:
## QE(df = 241) = 972.2110, p-val < .0001
##
## Model Results:
##
##            estimate      se     zval    pval   ci.lb   ci.ub
## condition    0.5709  0.0340  16.7817  <.0001  0.5042  0.6376  ***
##
## ---
## 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 = 242; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -222.7488   445.4976   457.4976   478.3312   457.8613
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
```

```
## inner factor: collapse   (nlvls = 52)
##
##              estim    sqrt  fixed
## tau^2       0.1026  0.3203     no
## rho         0.6486             no
##
## Test for Residual Heterogeneity:
## QE(df = 238) = 839.5329, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 258.7390, p-val < .0001
##
## Model Results:
##
##                   estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt             0.3035  0.0572   5.3018  <.0001   0.1913  0.4157  ***
## age.C               0.0201  0.0106   1.8963  0.0579  -0.0007  0.0409    .
## condition           0.5575  0.0353  15.8060  <.0001   0.4884  0.6267  ***
## age.C:condition     0.0086  0.0076   1.1301  0.2584  -0.0063  0.0236
##
## ---
## 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 = 242; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -218.4612   436.9224   464.9224   513.0555   466.8759
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##              estim    sqrt  fixed
## tau^2       0.0985  0.3138     no
## rho         0.6357             no
```

```
##
## Test for Residual Heterogeneity:
## QE(df = 230) = 783.2024, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 264.9641, p-val < .0001
##
## Model Results:
##
##                                            estimate      se     zval
## intrcpt                                      0.2886  0.0618   4.6663
## age.C                                        0.0173  0.0116   1.4939
## condition                                    0.5453  0.0376  14.5137
## lang_familyRomanic                           0.2600  0.1672   1.5550
## lang_familySino-Tibetian                    -0.2521  0.2170  -1.1618
## age.C:condition                              0.0087  0.0082   1.0566
## age.C:lang_familyRomanic                     0.0301  0.0327   0.9213
## age.C:lang_familySino-Tibetian              -0.0064  0.0406  -0.1578
## condition:lang_familyRomanic                -0.0111  0.1350  -0.0819
## condition:lang_familySino-Tibetian           0.2219  0.1871   1.1857
## age.C:condition:lang_familyRomanic          -0.0187  0.0294  -0.6356
## age.C:condition:lang_familySino-Tibetian     0.0155  0.0336   0.4604
##                                              pval    ci.lb   ci.ub
## intrcpt                                    <.0001   0.1674  0.4098  ***
## age.C                                      0.1352  -0.0054  0.0400
## condition                                  <.0001   0.4717  0.6189  ***
## lang_familyRomanic                         0.1200  -0.0677  0.5877
## lang_familySino-Tibetian                   0.2453  -0.6774  0.1732
## age.C:condition                            0.2907  -0.0074  0.0248
## age.C:lang_familyRomanic                   0.3569  -0.0340  0.0942
## age.C:lang_familySino-Tibetian             0.8746  -0.0860  0.0731
## condition:lang_familyRomanic               0.9347  -0.2756  0.2535
## condition:lang_familySino-Tibetian         0.2357  -0.1449  0.5887
## age.C:condition:lang_familyRomanic         0.5250  -0.0763  0.0389
## age.C:condition:lang_familySino-Tibetian   0.6452  -0.0504  0.0813
##
## ---
## 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 = 242; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -218.4052   436.8105   456.8105   491.3637   457.7970
##
## Variance Components:
```

```
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.0985  0.3139     no
## rho        0.6431             no
##
## Test for Residual Heterogeneity:
## QE(df = 234) = 824.2765, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 266.6471, p-val < .0001
##
## Model Results:
##
##                                                       estimate      se
## intrcpt                                                 0.3898  0.0697
## age.C                                                   0.0293  0.0130
## condition                                               0.5567  0.0435
## as.factor(object_pair)familiar_novel                   -0.2687  0.1311
## age.C:condition                                         0.0139  0.0094
## age.C:as.factor(object_pair)familiar_novel              0.0006  0.0261
## condition:as.factor(object_pair)familiar_novel          0.0870  0.0911
## age.C:condition:as.factor(object_pair)familiar_novel   -0.0320  0.0199
##                                                          zval    pval
## intrcpt                                                5.5946  <.0001
## age.C                                                  2.2427  0.0249
## condition                                             12.7900  <.0001
## as.factor(object_pair)familiar_novel                  -2.0493  0.0404
## age.C:condition                                        1.4751  0.1402
## age.C:as.factor(object_pair)familiar_novel             0.0230  0.9817
## condition:as.factor(object_pair)familiar_novel         0.9548  0.3397
## age.C:condition:as.factor(object_pair)familiar_novel  -1.6137  0.1066
##                                                         ci.lb    ci.ub
## intrcpt                                                0.2533   0.5264
## age.C                                                  0.0037   0.0548
## condition                                              0.4714   0.6420
## as.factor(object_pair)familiar_novel                  -0.5257  -0.0117
## age.C:condition                                       -0.0046   0.0324
## age.C:as.factor(object_pair)familiar_novel            -0.0505   0.0517
## condition:as.factor(object_pair)familiar_novel        -0.0916   0.2656
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0710   0.0069
##
## 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)

summary(rma_DistractorAgeS)
```

```
##
## Multivariate Meta-Analysis Model (k = 173; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -155.6407   311.2813   331.2813   362.3408   332.7099
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 23)
## inner factor: collapse   (nlvls = 38)
##
##             estim    sqrt  fixed
## tau^2      0.1380  0.3715     no
## rho        0.6658             no
##
## Test for Residual Heterogeneity:
## QE(df = 165) = 588.7541, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 194.1267, p-val < .0001
##
## Model Results:
##
##                                                         estimate      se
## intrcpt                                                   0.4156  0.0935
## age.C                                                     0.0045  0.0297
## condition                                                 0.5044  0.0503
## as.factor(object_pair)familiar_novel                     -0.3214  0.1812
## age.C:condition                                           0.0546  0.0211
## age.C:as.factor(object_pair)familiar_novel                0.0406  0.0505
## condition:as.factor(object_pair)familiar_novel            0.1003  0.1134
## age.C:condition:as.factor(object_pair)familiar_novel    -0.0300  0.0355
##                                                            zval    pval
## intrcpt                                                  4.4468  <.0001
## age.C                                                    0.1520  0.8792
## condition                                               10.0314  <.0001
## as.factor(object_pair)familiar_novel                    -1.7735  0.0761
## age.C:condition                                          2.5865  0.0097
## age.C:as.factor(object_pair)familiar_novel               0.8032  0.4219
## condition:as.factor(object_pair)familiar_novel           0.8843  0.3765
```

```
## age.C:condition:as.factor(object_pair)familiar_novel  -0.8459  0.3976
##                                                         ci.lb   ci.ub
## intrcpt                                                0.2324  0.5987  ***
## age.C                                                 -0.0538  0.0628
## condition                                              0.4059  0.6030  ***
## as.factor(object_pair)familiar_novel                  -0.6767  0.0338    .
## age.C:condition                                        0.0132  0.0960   **
## age.C:as.factor(object_pair)familiar_novel            -0.0584  0.1396
## condition:as.factor(object_pair)familiar_novel        -0.1220  0.3225
## age.C:condition:as.factor(object_pair)familiar_novel  -0.0996  0.0396
##
## ---
## 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 = 140; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -59.5608  119.1215  139.1215  167.9496  140.9397
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 32)
## inner factor: collapse   (nlvls = 52)
##
##             estim    sqrt  fixed
## tau^2      0.0927  0.3044     no
## rho        0.5653             no
##
## Test for Residual Heterogeneity:
## QE(df = 132) = 336.3942, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 14.8616, p-val = 0.0378
##
## Model Results:
##
##                                  estimate      se     zval    pval    ci.lb
## intrcpt                            0.0846  0.3639   0.2326  0.8161  -0.6285
## age.C                              0.0185  0.0190   0.9724  0.3309  -0.0188
## distractor_overlapno               0.5276  0.3814   1.3836  0.1665  -0.2198
## distractor_overlapnovel           -0.0163  0.3846  -0.0424  0.9662  -0.7702
## distractor_overlaponset            0.1362  0.3669   0.3711  0.7105  -0.5829
## distractor_overlaponset/medial     0.1559  0.4981   0.3129  0.7543  -0.8204
## age.C:distractor_overlapno         0.0387  0.0279   1.3844  0.1662  -0.0161
## age.C:distractor_overlapnovel      0.0131  0.0301   0.4334  0.6647  -0.0460
```

```
##                                     ci.ub
## intrcpt                            0.7978
## age.C                              0.0558
## distractor_overlapno               1.2751
## distractor_overlapnovel            0.7376
## distractor_overlaponset            0.8552
## distractor_overlaponset/medial     1.1321
## age.C:distractor_overlapno         0.0934
## age.C:distractor_overlapnovel      0.0721
##
## ---
## 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 = 125; method: REML)
##
##    logLik  Deviance       AIC       BIC      AICc
## -58.2248  116.4496  132.4496  154.6826  133.7587
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1089  0.3300     no
## rho        0.4155             no
##
## Test for Residual Heterogeneity:
## QE(df = 119) = 345.6749, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 9.7207, p-val = 0.0835
##
## Model Results:
##
##                               estimate      se     zval    pval    ci.lb
## intrcpt                         0.3130  0.0620   5.0451  <.0001   0.1914
## as.factor(n_feature)2          -0.1722  0.0966  -1.7824  0.0747  -0.3615
## as.factor(n_feature)3          -0.2795  0.1084  -2.5787  0.0099  -0.4919
```

```
## as.factor(n_feature)41640   -0.2711  0.2389  -1.1348  0.2565  -0.7392
## as.factor(n_feature)41641   -0.2555  0.1379  -1.8527  0.0639  -0.5257
## as.factor(n_feature)41672   -0.3130  0.3313  -0.9446  0.3448  -0.9623
##                              ci.ub
## intrcpt                     0.4346  ***
## as.factor(n_feature)2       0.0171    .
## as.factor(n_feature)3      -0.0671   **
## as.factor(n_feature)41640   0.1971
## as.factor(n_feature)41641   0.0148    .
## as.factor(n_feature)41672   0.3364
##
## ---
## 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 = 125; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -58.0985  116.1970  144.1970  182.3805  148.4828
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1138  0.3373     no
## rho        0.3838             no
##
## Test for Residual Heterogeneity:
## QE(df = 113) = 322.0035, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10,11,12):
## QM(df = 11) = 11.9886, p-val = 0.3645
##
## Model Results:
##
##                                   estimate      se     zval    pval
## intrcpt                             0.3205  0.0627   5.1078  <.0001
## as.factor(n_feature)2              -0.1797  0.0979  -1.8361  0.0663
## as.factor(n_feature)3              -0.2697  0.1142  -2.3620  0.0182
## as.factor(n_feature)41640          -0.3160  0.2573  -1.2281  0.2194
## as.factor(n_feature)41641          -0.2722  0.1421  -1.9163  0.0553
## as.factor(n_feature)41672          -0.3205  0.6758  -0.4743  0.6353
## age.C                               0.0159  0.0140   1.1341  0.2568
## as.factor(n_feature)2:age.C         0.0083  0.0196   0.4237  0.6718
## as.factor(n_feature)3:age.C        -0.0045  0.0228  -0.1978  0.8432
## as.factor(n_feature)41640:age.C    -0.0203  0.0522  -0.3894  0.6970
## as.factor(n_feature)41641:age.C    -0.0297  0.0461  -0.6427  0.5204
```

```
## as.factor(n_feature)41672:age.C   -0.0159  1.6306  -0.0097  0.9922
##                                     ci.lb    ci.ub
## intrcpt                            0.1975   0.4435  ***
## as.factor(n_feature)2             -0.3715   0.0121    .
## as.factor(n_feature)3             -0.4934  -0.0459    *
## as.factor(n_feature)41640         -0.8203   0.1883
## as.factor(n_feature)41641         -0.5506   0.0062    .
## as.factor(n_feature)41672         -1.6450   1.0040
## age.C                             -0.0116   0.0434
## as.factor(n_feature)2:age.C       -0.0302   0.0468
## as.factor(n_feature)3:age.C       -0.0492   0.0402
## as.factor(n_feature)41640:age.C  -0.1226   0.0820
## as.factor(n_feature)41641:age.C  -0.1201   0.0608
## as.factor(n_feature)41672:age.C  -3.2118   3.1800
##
## ---
## 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 = 202; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -205.2733   410.5467   424.5467   447.5291   425.1393
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1219  0.3492     no
## rho        0.6342             no
##
## Test for Residual Heterogeneity:
## QE(df = 197) = 777.9053, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5):
## QM(df = 4) = 226.2117, p-val < .0001
##
## Model Results:
```

```
##
##                          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt                    0.6713  0.1300   5.1626  <.0001   0.4165   0.9262
## as.factor(n_feature)1     -0.3651  0.1126  -3.2426  0.0012  -0.5858  -0.1444
## as.factor(n_feature)2     -0.5352  0.1376  -3.8897  0.0001  -0.8048  -0.2655
## as.factor(n_feature)3     -0.6180  0.1469  -4.2067  <.0001  -0.9059  -0.3300
## condition                  0.1922  0.1057   1.8184  0.0690  -0.0150   0.3993
##
## 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 = 202; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -200.4790   400.9580   424.9580   464.0479   426.7010
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 27)
## inner factor: collapse   (nlvls = 49)
##
##             estim    sqrt  fixed
## tau^2      0.1235  0.3515     no
## rho        0.7046             no
##
## Test for Residual Heterogeneity:
## QE(df = 192) = 739.9198, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8,9,10):
## QM(df = 9) = 238.4769, p-val < .0001
##
## Model Results:
##
##                                estimate      se     zval    pval    ci.lb
## intrcpt                          0.6923  0.1319   5.2476  <.0001   0.4337
## as.factor(n_feature)1           -0.3695  0.1133  -3.2617  0.0011  -0.5915
## as.factor(n_feature)2           -0.5594  0.1383  -4.0449  <.0001  -0.8304
## as.factor(n_feature)3           -0.6370  0.1512  -4.2134  <.0001  -0.9333
## age.C                            0.0904  0.0500   1.8075  0.0707  -0.0076
## condition                        0.1992  0.1062   1.8765  0.0606  -0.0089
## as.factor(n_feature)1:age.C     -0.0743  0.0484  -1.5350  0.1248  -0.1693
## as.factor(n_feature)2:age.C     -0.0354  0.0511  -0.6925  0.4886  -0.1356
```

```
## as.factor(n_feature)3:age.C   -0.0513  0.0523  -0.9814  0.3264  -0.1538
## age.C:condition               -0.0622  0.0478  -1.3020  0.1929  -0.1558
##                                ci.ub
## intrcpt                       0.9508  ***
## as.factor(n_feature)1        -0.1475   **
## as.factor(n_feature)2        -0.2883  ***
## as.factor(n_feature)3        -0.3407  ***
## age.C                         0.1885    .
## condition                     0.4073    .
## as.factor(n_feature)1:age.C   0.0206
## as.factor(n_feature)2:age.C   0.0648
## as.factor(n_feature)3:age.C   0.0512
## age.C:condition               0.0314
##
## ---
## 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 = 107; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -56.7840  113.5679  121.5679  132.1838  121.9679
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1375  0.3708     no
## rho        0.4900             no
##
## Test for Residual Heterogeneity:
## QE(df = 105) = 343.2173, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.0009, p-val = 0.9760
##
```

```
## Model Results:
##
##                          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                    0.2671  0.0820  3.2580  0.0011   0.1064  0.4278
## mispron_locationmedial     0.0044  0.1453  0.0300  0.9760  -0.2804  0.2891
##
## 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_MP1, random = ~collapse | short_cite)


summary(rma_LocationAge)
```

```
##
## Multivariate Meta-Analysis Model (k = 107; method: REML)
##
##    logLik  Deviance       AIC       BIC      AICc
## -55.1900  110.3799  122.3799  138.1883  123.2549
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 24)
## inner factor: collapse   (nlvls = 41)
##
##             estim    sqrt  fixed
## tau^2      0.1420  0.3768     no
## rho        0.4848             no
##
## Test for Residual Heterogeneity:
## QE(df = 103) = 333.1781, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 1.5638, p-val = 0.6676
##
## Model Results:
##
##                                estimate      se    zval    pval    ci.lb
## intrcpt                          0.2721  0.0842  3.2330  0.0012   0.1072
## mispron_locationmedial           0.0419  0.1604  0.2612  0.7939  -0.2724
## age.C                            0.0159  0.0176  0.9008  0.3677  -0.0187
## mispron_locationmedial:age.C     0.0136  0.0331  0.4104  0.6815  -0.0513
##                                 ci.ub
## intrcpt                        0.4371  **
## mispron_locationmedial         0.3562
## age.C                          0.0504
## mispron_locationmedial:age.C   0.0785
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MP type: Vowel, consonant, or tone?

```
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 = 126; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -63.1824  126.3649  134.3649  145.6460  134.7010
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1141  0.3379     no
## rho        0.5144             no
##
## Test for Residual Heterogeneity:
## QE(df = 124) = 368.5196, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## QM(df = 1) = 0.0214, p-val = 0.8836
##
## Model Results:
##
##                    estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt              0.2609  0.0700  3.7287  0.0002   0.1238  0.3980  ***
## type_featurevowel    0.0127  0.0867  0.1464  0.8836  -0.1573  0.1827
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
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 = 126; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -61.7954  123.5908  135.5908  152.4150  136.3213
```

```
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 26)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.1133  0.3365     no
## rho        0.5031             no
##
## Test for Residual Heterogeneity:
## QE(df = 122) = 353.1397, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 2.0345, p-val = 0.5653
##
## Model Results:
##
##                           estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                     0.2649  0.0698  3.7982  0.0001   0.1282  0.4017
## type_featurevowel           0.0224  0.0870  0.2578  0.7965  -0.1481  0.1930
## age.C                       0.0168  0.0143  1.1804  0.2378  -0.0111  0.0448
## type_featurevowel:age.C     0.0001  0.0168  0.0080  0.9937  -0.0329  0.0331
##
## 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 = 219; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -211.6290   423.2580   435.2580   455.4818   435.6618
##
## Variance Components:
##
```

```
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.0986  0.3140     no
## rho        0.6180             no
##
## Test for Residual Heterogeneity:
## QE(df = 215) = 768.5747, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 188.5001, p-val < .0001
##
## Model Results:
##
##                                  estimate      se      zval    pval
## intrcpt                            0.7817  0.0623   12.5466  <.0001
## relevel(type_feature, "none")1   -0.5192  0.0437  -11.8720  <.0001
## relevel(type_feature, "none")4   -0.6770  0.1536   -4.4091  <.0001
## relevel(type_feature, "none")5   -0.4985  0.0562   -8.8617  <.0001
##                                    ci.lb    ci.ub
## intrcpt                           0.6596   0.9038  ***
## relevel(type_feature, "none")1  -0.6049  -0.4335  ***
## relevel(type_feature, "none")4  -0.9780  -0.3761  ***
## relevel(type_feature, "none")5  -0.6087  -0.3882  ***
##
## ---
## 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 = 219; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -208.6836   417.3672   437.3672   470.8858   438.4672
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 28)
## inner factor: collapse   (nlvls = 46)
##
##             estim    sqrt  fixed
## tau^2      0.0940  0.3065     no
## rho        0.5926             no
##
## Test for Residual Heterogeneity:
## QE(df = 211) = 745.0577, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6,7,8):
## QM(df = 7) = 194.9376, p-val < .0001
```

```
##
## Model Results:
##
##                                        estimate      se      zval    pval
## intrcpt                                 0.8108  0.0623   13.0104  <.0001
## relevel(type_feature, "none")1         -0.5293  0.0441  -11.9902  <.0001
## relevel(type_feature, "none")4         -0.6848  0.1670   -4.1011  <.0001
## relevel(type_feature, "none")5         -0.5304  0.0620   -8.5523  <.0001
## age.C                                   0.0249  0.0115    2.1656  0.0303
## relevel(type_feature, "none")1:age.C    0.0018  0.0102    0.1732  0.8625
## relevel(type_feature, "none")4:age.C   -0.0033  0.0308   -0.1057  0.9158
## relevel(type_feature, "none")5:age.C   -0.0131  0.0114   -1.1487  0.2507
##                                         ci.lb    ci.ub
## intrcpt                                0.6886   0.9329  ***
## relevel(type_feature, "none")1        -0.6158  -0.4428  ***
## relevel(type_feature, "none")4        -1.0121  -0.3575  ***
## relevel(type_feature, "none")5        -0.6520  -0.4088  ***
## age.C                                  0.0024   0.0475    *
## relevel(type_feature, "none")1:age.C  -0.0182   0.0218
## relevel(type_feature, "none")4:age.C  -0.0636   0.0571
## relevel(type_feature, "none")5:age.C  -0.0354   0.0092
##
## ---
## 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 = 203; method: REML)
##
```

```
##    logLik   Deviance        AIC        BIC       AICc
## -202.0398   404.0796   420.0796   446.3452   420.8455
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1028  0.3207     no
## rho        0.5015             no
##
## Test for Residual Heterogeneity:
## QE(df = 197) = 693.7865, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 190.0754, p-val < .0001
##
## Model Results:
##
##                                                   estimate      se
## intrcpt                                             0.7441  0.0703
## relevel(type_feature, "none")1                     -0.4959  0.0458
## relevel(type_feature, "none")5                     -0.5206  0.0636
## lang_familyRomanic                                  0.3791  0.1649
## relevel(type_feature, "none")1:lang_familyRomanic  -0.5443  0.2112
## relevel(type_feature, "none")5:lang_familyRomanic   0.1277  0.1483
##                                                       zval    pval
## intrcpt                                            10.5912  <.0001
## relevel(type_feature, "none")1                    -10.8376  <.0001
## relevel(type_feature, "none")5                     -8.1817  <.0001
## lang_familyRomanic                                  2.2994  0.0215
## relevel(type_feature, "none")1:lang_familyRomanic  -2.5765  0.0100
## relevel(type_feature, "none")5:lang_familyRomanic   0.8614  0.3890
##                                                     ci.lb    ci.ub
## intrcpt                                            0.6064   0.8818  ***
## relevel(type_feature, "none")1                    -0.5856  -0.4062  ***
## relevel(type_feature, "none")5                    -0.6453  -0.3959  ***
## lang_familyRomanic                                 0.0560   0.7023    *
## relevel(type_feature, "none")1:lang_familyRomanic -0.9583  -0.1302   **
## relevel(type_feature, "none")5:lang_familyRomanic -0.1629   0.4184
##
## ---
## 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 ==
```

```r
    "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 = 203; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
## -202.0398   404.0796   420.0796   446.3452   420.8455
##
## Variance Components:
##
## outer factor: short_cite (nlvls = 25)
## inner factor: collapse   (nlvls = 44)
##
##             estim    sqrt  fixed
## tau^2      0.1028  0.3207     no
## rho        0.5015             no
##
## Test for Residual Heterogeneity:
## QE(df = 197) = 693.7865, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## QM(df = 5) = 190.0754, p-val < .0001
##
## Model Results:
##
##                                                   estimate      se
## intrcpt                                             0.7441  0.0703
## relevel(type_feature, "none")1                     -0.4959  0.0458
## relevel(type_feature, "none")5                     -0.5206  0.0636
## lang_familyRomanic                                  0.3791  0.1649
## relevel(type_feature, "none")1:lang_familyRomanic  -0.5443  0.2112
## relevel(type_feature, "none")5:lang_familyRomanic   0.1277  0.1483
##                                                       zval    pval
## intrcpt                                            10.5912  <.0001
## relevel(type_feature, "none")1                    -10.8376  <.0001
## relevel(type_feature, "none")5                     -8.1817  <.0001
## lang_familyRomanic                                  2.2994  0.0215
## relevel(type_feature, "none")1:lang_familyRomanic  -2.5765  0.0100
```

```
## relevel(type_feature, "none")5:lang_familyRomanic    0.8614  0.3890
##                                                      ci.lb    ci.ub
## intrcpt                                             0.6064   0.8818  ***
## relevel(type_feature, "none")1                     -0.5856  -0.4062  ***
## relevel(type_feature, "none")5                     -0.6453  -0.3959  ***
## lang_familyRomanic                                  0.0560   0.7023    *
## relevel(type_feature, "none")1:lang_familyRomanic  -0.9583  -0.1302   **
## relevel(type_feature, "none")5:lang_familyRomanic  -0.1629   0.4184
##
## ---
## 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.

```
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

```
# 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")
```

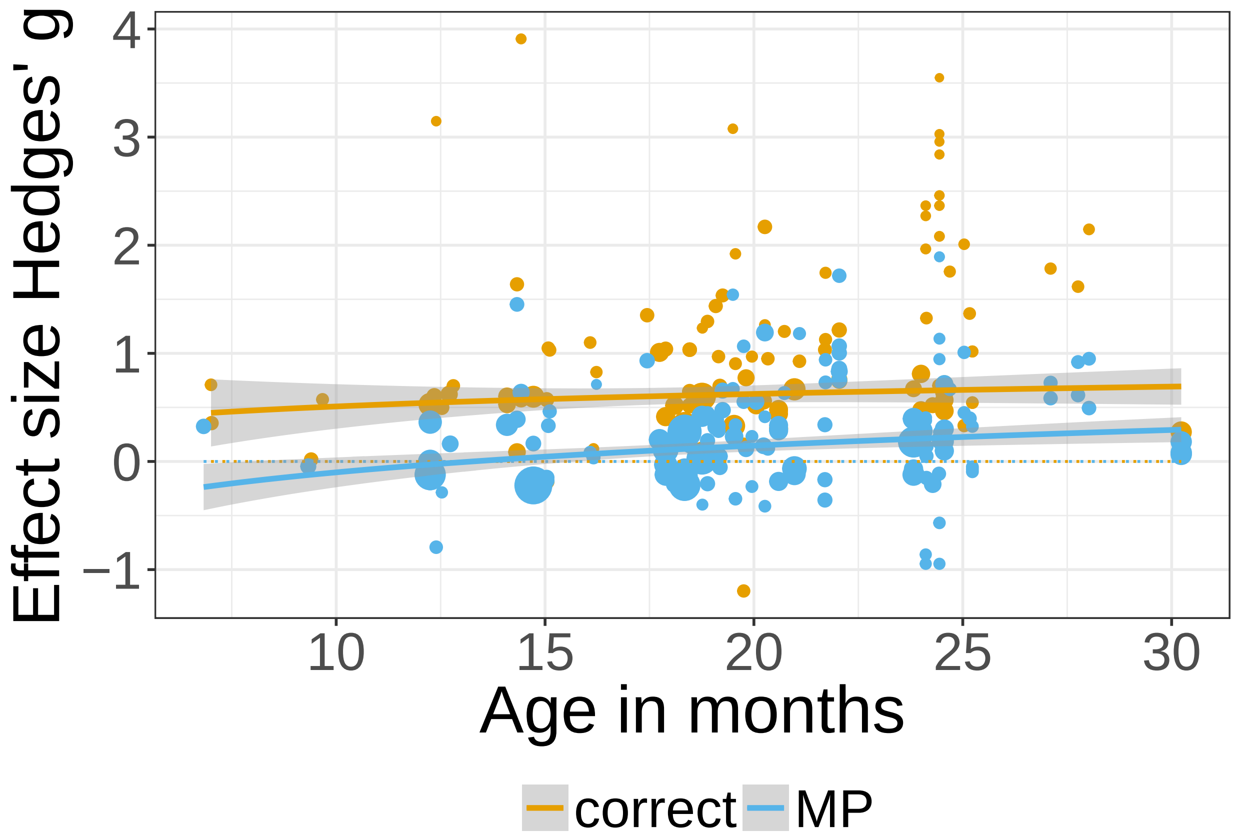
```
##                                      COR            95%-CI %W(fixed)
## Zesiger et al. (2012)            -0.0090 [-0.4268; 0.4088]       5.0
## Zesiger et al. (2012)            -0.1720 [-0.5775; 0.2335]       5.3
## Mani, Coleman, & Plunkett (2008)  0.0700 [-0.1861; 0.3261]      13.2
## Mani & 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

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
```

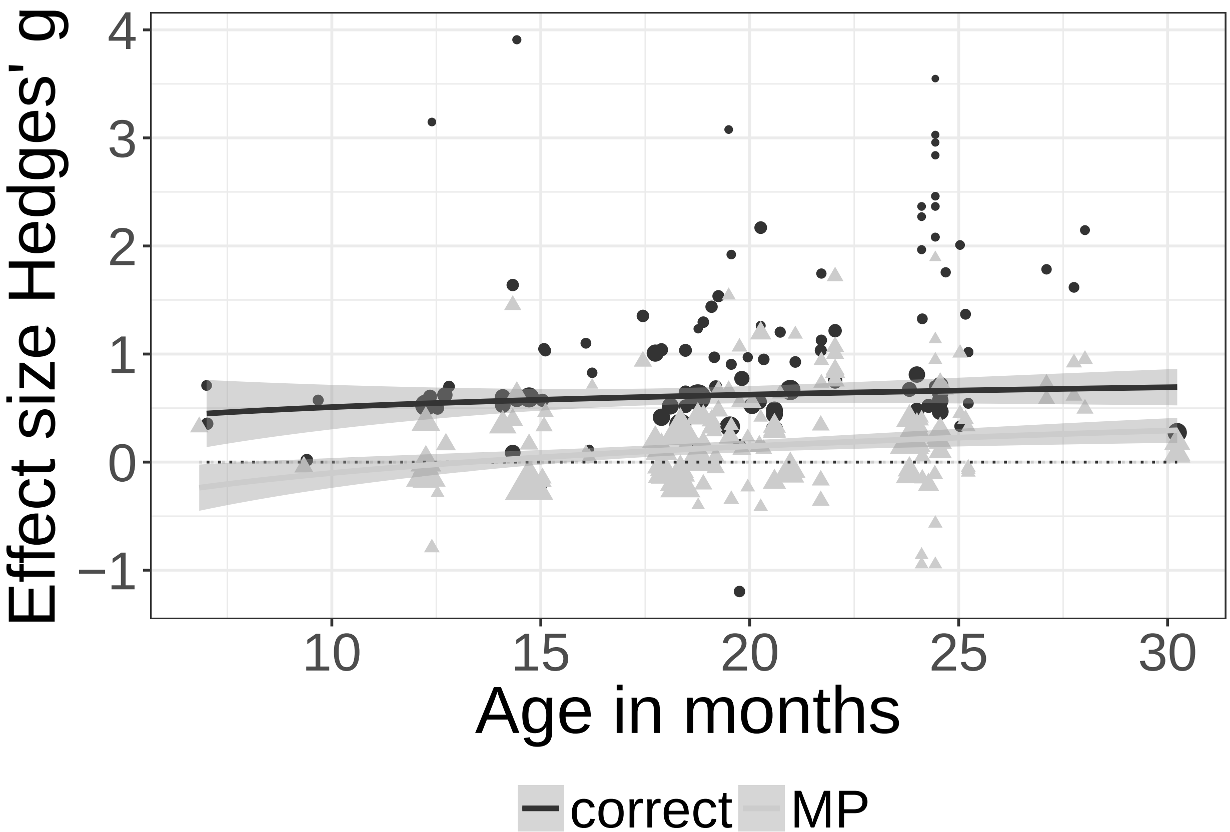
```
# ggsave('figures/AgeEffect_log.jpg', p,height= 7,width= 6)



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") +
```

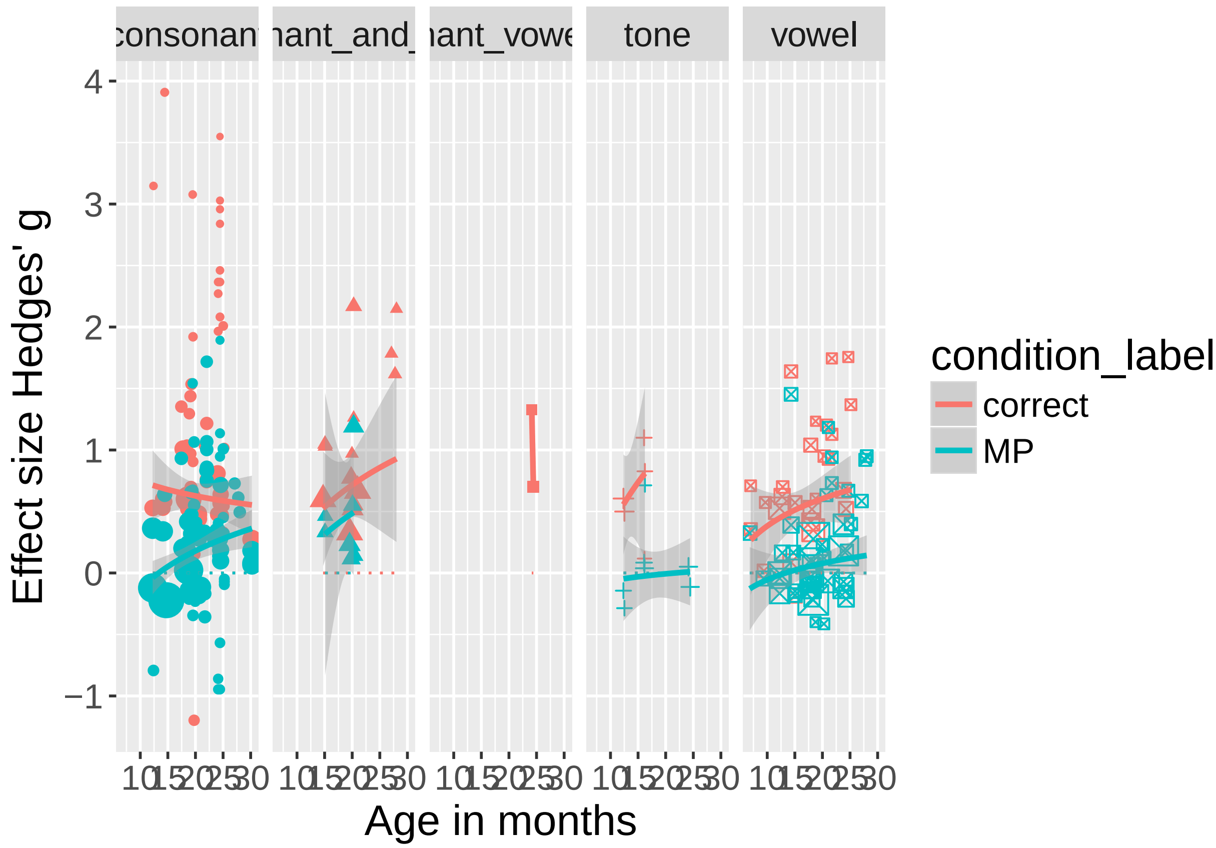
```
    ylab("Effect size Hedges' g")
p
```




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

MP type: Consonant, Vowel, or Tone?

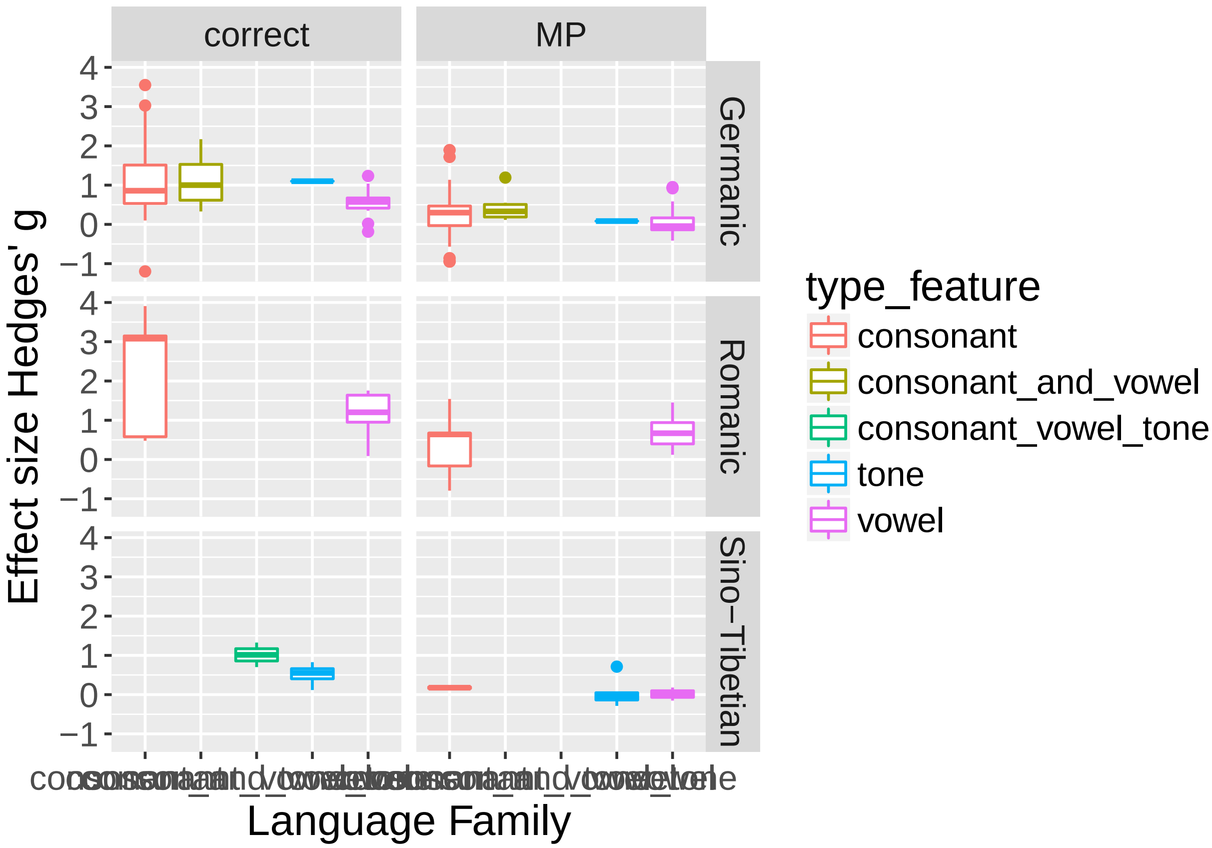
```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = type_feature), show.legend = FALSE) + facet_grid(. ~ type_feature) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight = weights_g))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




```
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
```

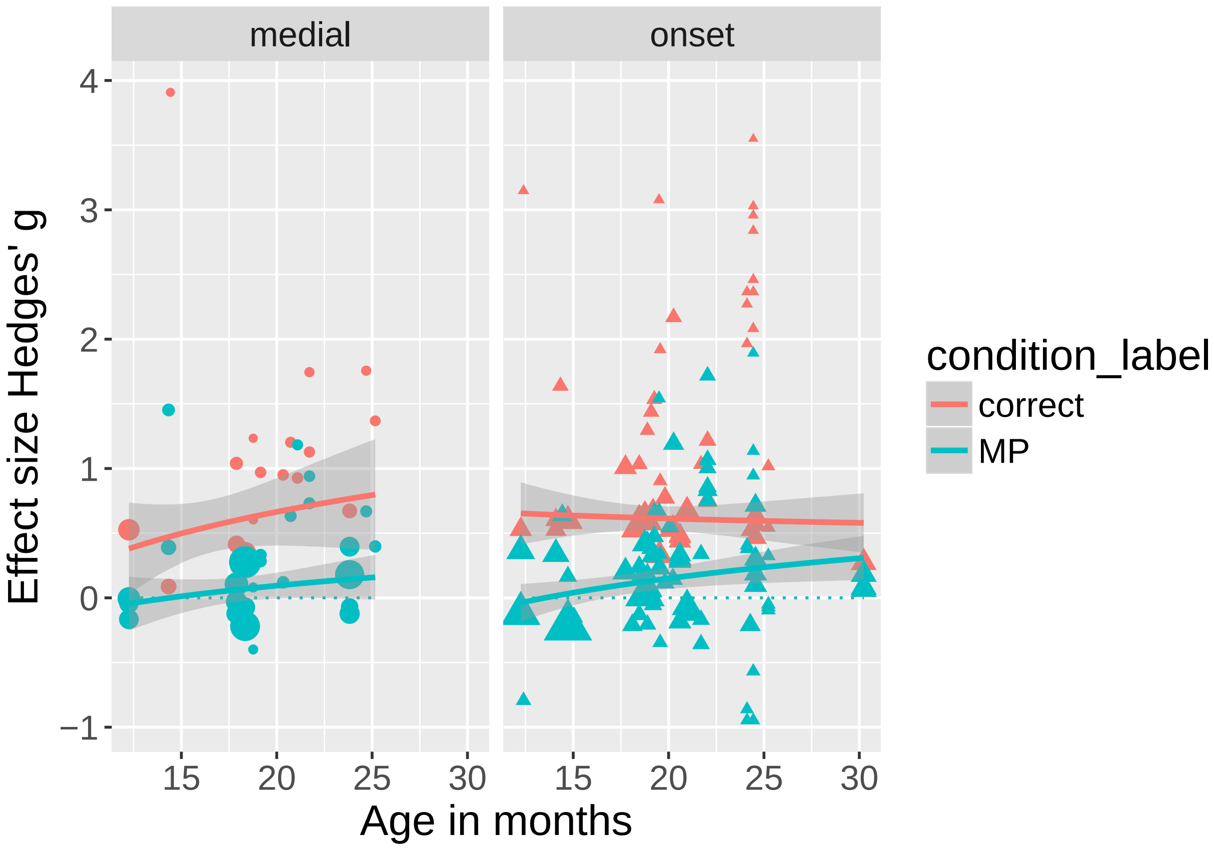
```
## Error: A continuous variable can not be mapped to shape
```

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

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

Position of Mispronunciation

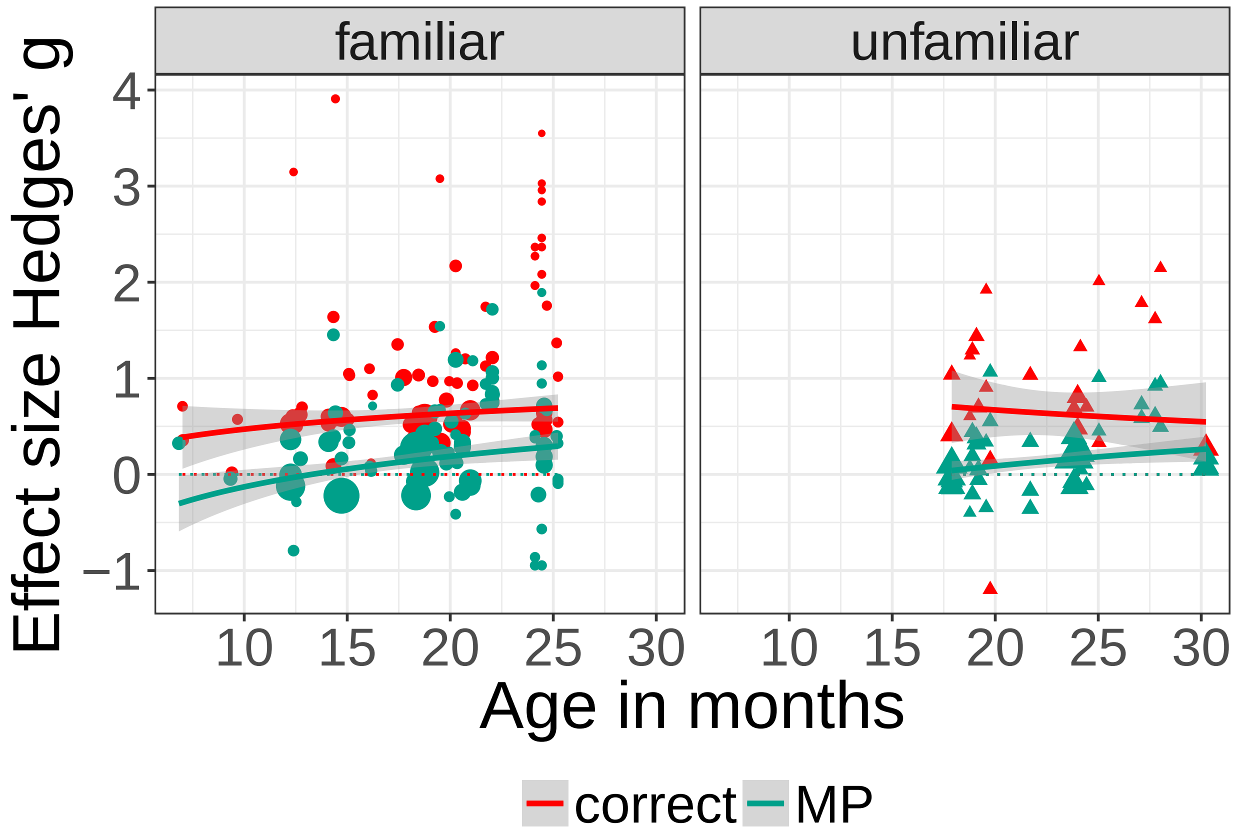
```r
dat.p <- subset(dat, mispron_location == "onset" | mispron_location == "medial" |
    mispron_location == "offset")

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

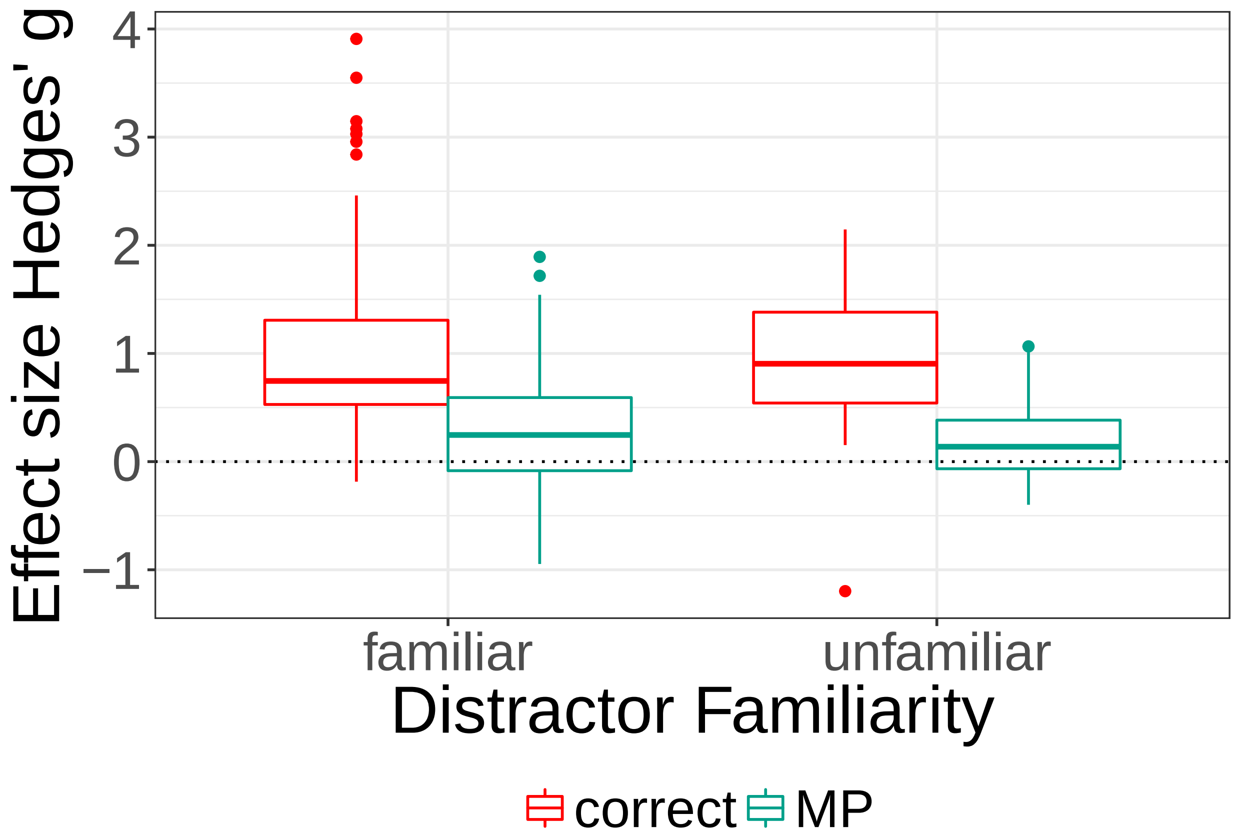
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

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)

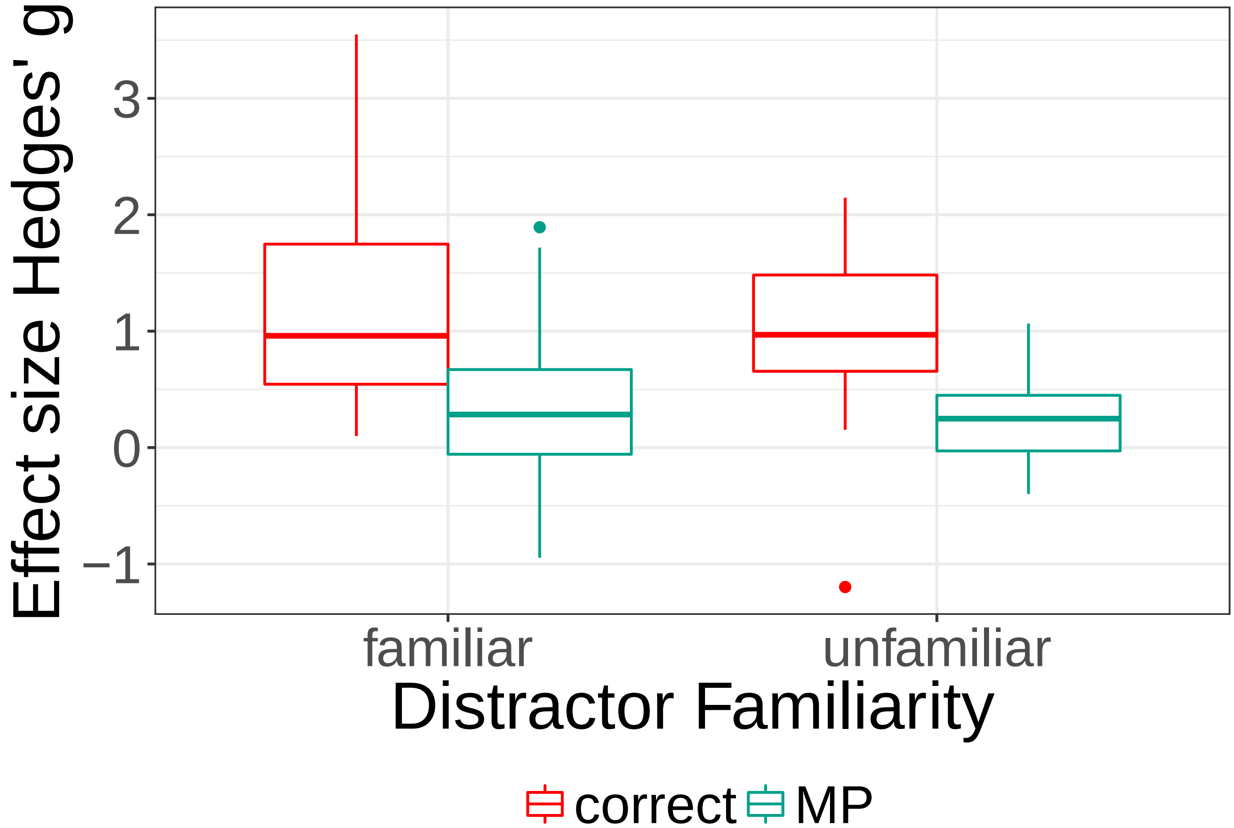
```
dat$condition_label = ifelse(dat$condition == 1, "correct", "MP")
dat$dist_code <- ifelse(dat$object_pair == "familiar_familiar", "familiar",
    "unfamiliar")

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)

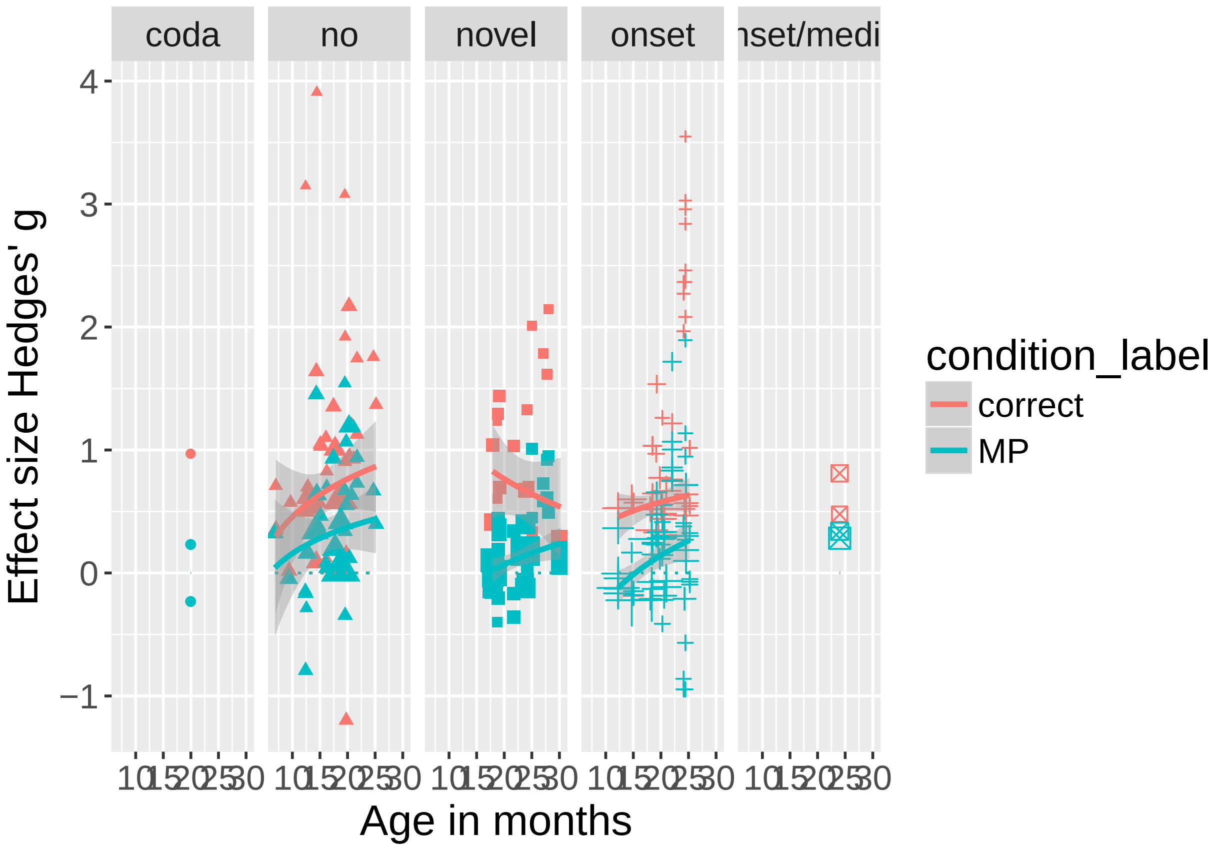
```
min_age <- min(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)
max_age <- max(dat[dat$object_pair == "familiar_novel", ]$mean_age_1)

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

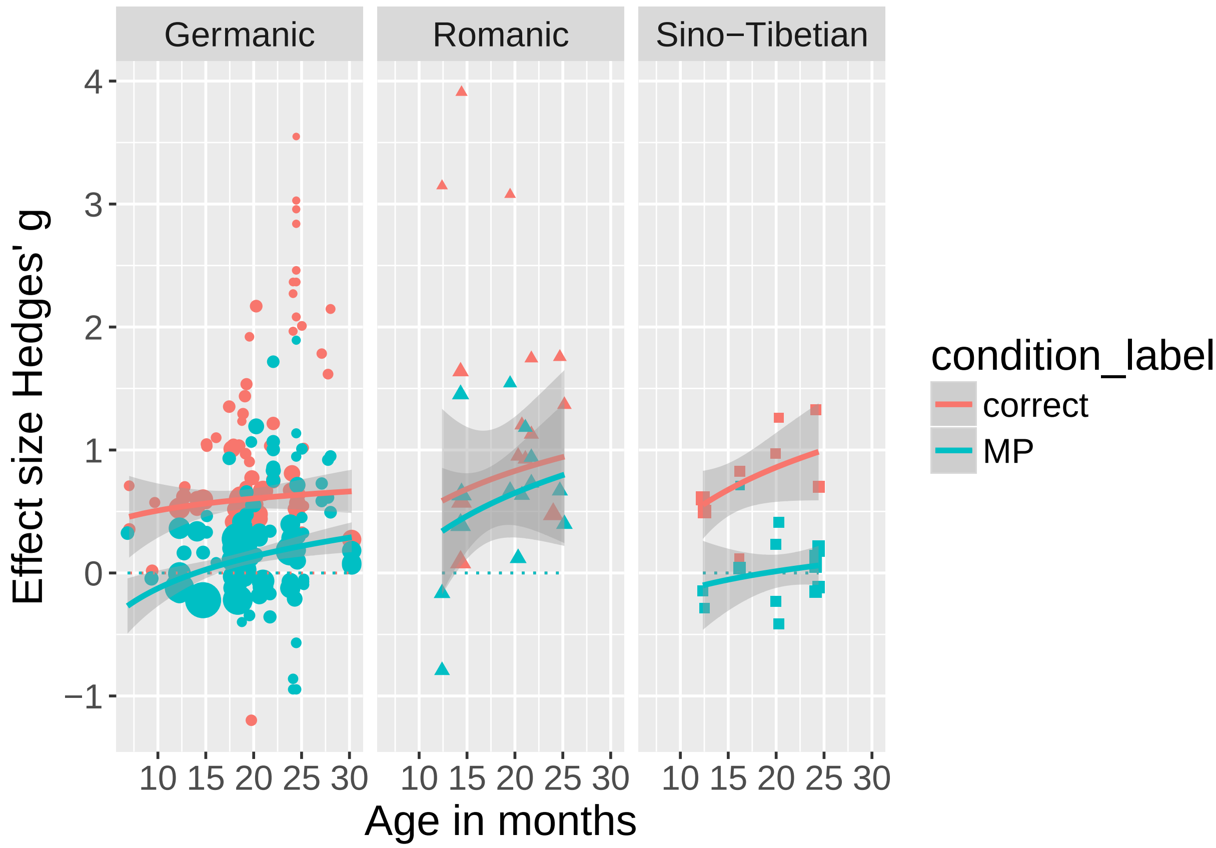
```
p <- ggplot(dat, aes(mean_age_1/30.44, g_calc, color = condition_label)) + geom_point(aes(size = weights
    shape = distractor_overlap), show.legend = FALSE) + facet_grid(. ~ distractor_overlap) +
    geom_line(y = 0, linetype = "dotted") + geom_smooth(method = "lm", formula = y ~
    log(x), aes(weight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




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

Language Family

```r
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
```




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