Colour experiment

```
library(ggplot2)
library(lme4)
```

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

```
##
## Attaching package: 'lme4'
```

```
## The following object is masked from 'package:stats':
##
##     sigma
```

Load data

```
variants = read.csv('../data/processedData/variants_summary.csv', stringsAsFactors = F)
```

Transform some variables.

```
variants$Teach = variants$Teach >0

variants$check.any = variants$check>0

variants$freq_week_1.logcenter = log(variants$freq_week_1 + 1)
variants$freq_week_1.logcenter =
  variants$freq_week_1.logcenter - mean(variants$freq_week_1.logcenter)

# cut TryMarking into two categories
variants$TryMarked.cat = cut(variants$TryMarked,
                             c(-Inf,3,Inf),
                             labels = c("Low",'High'))

variants$averageLength_week_1.logcenter = log(variants$averageLength_week_1)
variants$averageLength_week_1.logcenter =
  variants$averageLength_week_1.logcenter -
    mean(variants$averageLength_week_1.logcenter)
```

LMER models

```
m0 = lmer(freq_week_4_withinColour ~
            1
          + (1 | colourName),
          data=variants)
m1 = lmer(freq_week_4_withinColour ~
            (indexical) +
            + (1 | colourName),
          data=variants)
m2 = lmer(freq_week_4_withinColour ~
            (indexical) +
```

```
            (Teach)
          + (1 | colourName),
          data=variants)

m3 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach) + (TryMarked.cat)
          + (1 | colourName),
          data=variants)

m4 = lmer(freq_week_4_withinColour ~
            (indexical) +
            Teach * TryMarked.cat
          + (1 | colourName),
          data=variants)

m5 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter
          + (1 | colourName),
          data=variants)

m6 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter
          + (1 | colourName),
          data=variants)

m7 = lmer(freq_week_4_withinColour ~
            (indexical) +
            (Teach * TryMarked.cat) +
            freq_week_1.logcenter +
            averageLength_week_1.logcenter+
            check.any
          + (1 | colourName),
          data=variants)
```

Results

```
anova(m0,m1,m2,m3,m4,m5, m6,m7)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: variants
## Models:
## m0: freq_week_4_withinColour ~ 1 + (1 | colourName)
## m1: freq_week_4_withinColour ~ (indexical) + +(1 | colourName)
## m2: freq_week_4_withinColour ~ (indexical) + (Teach) + (1 | colourName)
## m3: freq_week_4_withinColour ~ (indexical) + (Teach) + (TryMarked.cat) +
```

```
## m3:     (1 | colourName)
## m4: freq_week_4_withinColour ~ (indexical) + Teach * TryMarked.cat +
## m4:     (1 | colourName)
## m5: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m5:     freq_week_1.logcenter + (1 | colourName)
## m6: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m6:     freq_week_1.logcenter + averageLength_week_1.logcenter +
## m6:     (1 | colourName)
## m7: freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
## m7:     freq_week_1.logcenter + averageLength_week_1.logcenter +
## m7:     check.any + (1 | colourName)
##    Df      AIC     BIC logLik deviance   Chisq Chi Df Pr(>Chisq)
## m0  3  -64.284 -56.499 35.142  -70.284
## m1  5  -73.910 -60.934 41.955  -83.910 13.6255      2   0.001100 **
## m2  6  -71.963 -56.392 41.982  -83.963  0.0535      1   0.817080
## m3  7  -80.246 -62.080 47.123  -94.246 10.2829      1   0.001343 **
## m4  8  -81.624 -60.863 48.812  -97.624  3.3782      1   0.066064 .
## m5  9  -99.584 -76.228 58.792 -117.584 19.9599      1  7.908e-06 ***
## m6 10 -101.675 -75.724 60.838 -121.675  4.0911      1   0.043110 *
## m7 11  -99.785 -71.239 60.893 -121.785  0.1101      1   0.739992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(m7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## freq_week_4_withinColour ~ (indexical) + (Teach * TryMarked.cat) +
##     freq_week_1.logcenter + averageLength_week_1.logcenter +
##     check.any + (1 | colourName)
##    Data: variants
##
## REML criterion at convergence: -78.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5660 -0.4890  0.0082  0.2626  4.5416
##
## Random effects:
##  Groups     Name        Variance  Std.Dev.
##  colourName (Intercept) 0.0007442 0.02728
##  Residual               0.0183026 0.13529
## Number of obs: 99, groups:  colourName, 6
##
## Fixed effects:
##                                Estimate Std. Error t value
## (Intercept)                     0.06554    0.02627   2.495
## indexicalYes                   -0.01389    0.05296  -0.262
## indexicalYes-body               0.07613    0.04213   1.807
## TeachTRUE                      -0.05830    0.04775  -1.221
## TryMarked.catHigh               0.10522    0.08222   1.280
## freq_week_1.logcenter           0.16692    0.03663   4.556
## averageLength_week_1.logcenter -0.03214    0.01687  -1.906
## check.anyTRUE                  -0.01084    0.03301  -0.328
## TeachTRUE:TryMarked.catHigh    -0.17122    0.09450  -1.812
```

```
##
## Correlation of Fixed Effects:
##             (Intr) indxcY indxY- TcTRUE TryM.H fr__1. aL__1. c.TRUE
## indexicalYs -0.234
## indxclYs-bd -0.377  0.139
## TeachTRUE   -0.300  0.094  0.063
## TryMrkd.ctH -0.204 -0.186 -0.082  0.126
## frq_wk_1.lg  0.415  0.091 -0.223 -0.200 -0.465
## avrgLng__1.  0.053  0.067  0.029  0.128  0.015 -0.046
## chck.nyTRUE -0.543  0.098  0.193  0.133  0.064 -0.362 -0.180
## TcTRUE:TM.H  0.090  0.085 -0.003 -0.457 -0.651  0.017 -0.073  0.009
```

Summary

There was a significant main effect of frequency in week 1 (beta = 0.17 , std.err = 0.037 , Wald t = 4.6 ; log likelihood difference = 10 , df = 1 , Chi Squared = 19.96 , p = 7.9e-06).

There was a significant main effect of indexicality (beta = -0.014 , std.err = 0.053 , Wald t = -0.26 ; log likelihood difference = 6.8 , df = 2 , Chi Squared = 13.63 , p = 0.0011).

There was a significant main effect of sign length (beta = -0.032 , std.err = 0.017 , Wald t = -1.9 ; log likelihood difference = 2 , df = 1 , Chi Squared = 4.09 , p = 0.043).

```
#sjp.lmer(m5,'fe')
```