Week 5 Homework

Molly Farry-Thorn

October 3, 2017

1. Run a series of models using a time-invariant nominal covariate. For all models, how does your model change from model to model. What is your final model?

a) where the covariate only predicts the intercept

```
# time invariant covariate that predicts the intecept but not slope
children$Child.Gender <- relevel(children$Child.Gender, ref = "Male")

mod1a <- lmer(Utterances.with.Letters ~ Time + Child.Gender + (1|Subject), data = children)
summary(mod1a)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Utterances.with.Letters ~ Time + Child.Gender + (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5435.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.2033 -0.4072 -0.1776  0.0346 10.3464
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  23.29    4.826
##  Residual             231.59   15.218
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##                     Estimate Std. Error t value
## (Intercept)          1.8813     1.5386   1.223
## Time                 2.6307     0.5163   5.095
## Child.GenderFemale  -1.9533     1.7676  -1.105
##
## Correlation of Fixed Effects:
##             (Intr) Time
## Time        -0.613
## Chld.GndrFm -0.542 -0.002
```

Fixed Effects

Intercept: 1.8813; mean number of utterances with letters for male children at Session 1

Time: 2.63; increase in number of utterances with letters every year

Child.Gender: -1.95; difference between males and females at Session 1, females start lower

b) predicts both intercept and slope

```
# time invariant predictor for the intercept AND slopes
mod1b <- lmer(Utterances.with.Letters ~ Time + Child.Gender + Time*Child.Gender + (1|Subject), data = ch
summary(mod1b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Utterances.with.Letters ~ Time + Child.Gender + Time * Child.Gender +
##     (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5430.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3180 -0.3776 -0.1666  0.0207 10.2906
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  23.32    4.829
##  Residual             230.77   15.191
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##                           Estimate Std. Error t value
## (Intercept)                 0.2871     1.7803   0.161
## Time                        3.5028     0.7121   4.919
## Child.GenderFemale          1.4004     2.5869   0.541
## Time:Child.GenderFemale    -1.8318     1.0320  -1.775
##
## Correlation of Fixed Effects:
##             (Intr) Time   Chl.GF
## Time        -0.731
## Chld.GndrFm -0.688  0.503
## Tm:Chld.GnF  0.504 -0.690 -0.730
```

Fixed Effects

Intercept: 0.29; mean number of utterances with letters for male children at Session 1

Time: 3.50; increase in number of utterances with letters every year, for males

Child.Gender: 1.40; difference between males and females at Session 1, females start higher (different than last model)

Time:Child.Gender: -1.83; difference in the slopes between males and females, the effect of age is smaller in females

c) is rescaled (e.g. centering).

```
# changing dummy coding such that reference group becomes females
children$Child.Gender <- relevel(children$Child.Gender, ref = "Female")
```

```
mod1c <- lmer(Utterances.with.Letters ~ Time + Child.Gender + Time*Child.Gender + (1|Subject), data = cl
summary(mod1c)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Utterances.with.Letters ~ Time + Child.Gender + Time * Child.Gender +
##     (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5430.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3180 -0.3776 -0.1666  0.0207 10.2906
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  23.32    4.829
##  Residual             230.77   15.191
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)              1.688      1.877   0.899
## Time                     1.671      0.747   2.237
## Child.GenderMale        -1.400      2.587  -0.541
## Time:Child.GenderMale    1.832      1.032   1.775
##
## Correlation of Fixed Effects:
##             (Intr) Time   Chl.GM
## Time        -0.730
## Chld.GndrMl -0.725  0.529
## Tm:Chld.GnM  0.528 -0.724 -0.730
```

Fixed Effects

Intercept: 1.69; mean number of utterances with letters for female children at Session 1

Time: 1.67; increase in number of utterances with letters every year, for females

Child.Gender: -1.40; difference between males and females at Session 1, males start lower. This is the same difference we found in mod1b, just in the opposite direction.

Time:Child.Gender: 1.83; difference in the slopes between males and females, the effect of age is larger in males. This is the same difference we found in mod1b, just in the opposite direction.

```
anova(mod1a,mod1b)
```

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

```
## Data: children
## Models:
## mod1a: Utterances.with.Letters ~ Time + Child.Gender + (1 | Subject)
## mod1b: Utterances.with.Letters ~ Time + Child.Gender + Time * Child.Gender +
## mod1b:     (1 | Subject)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## mod1a  5 5450.5 5472.9 -2720.3   5440.5
```

```
## mod1b  6 5449.4 5476.2 -2718.7   5437.4 3.1546      1    0.07571 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The likelihood ratio test suggests that simpler model, where covariate only predicts intercept, is preferred.

2. Introduce a time-invariant continuous covariate and run models a-c from #1.

```
# time-invariant continuous covariate that only predicts the intercept
mod2a <- lmer(Utterances.with.Letters ~ Time + SES + (1|Subject), data = children)
summary(mod2a)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Utterances.with.Letters ~ Time + SES + (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5436.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.2214 -0.4072 -0.1713  0.0442 10.3952
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  22.78    4.772
##  Residual             231.60   15.218
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   0.9602     1.2895   0.745
## Time          2.6285     0.5163   5.091
## SES           1.1996     0.8852   1.355
##
## Correlation of Fixed Effects:
##      (Intr) Time
## Time -0.733
## SES   0.000 -0.002
```

Fixed Effects

Intercept: 0.96; mean of number of utterances with letters for children with mean level SES at Session 1

Time: 2.63; increase in number of utterances with letters every year, when SES at mean level

SES: 1.20; increase in number of utterances for every 1 unit increase in SES

```
# time-invariant continuous covariate predicts the intercept AND slopes
mod2b <- lmer(Utterances.with.Letters ~ Time + SES + Time*SES + (1|Subject), data = children)
summary(mod2b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Utterances.with.Letters ~ Time + SES + Time * SES + (1 | Subject)
##    Data: children
```

```
##
## REML criterion at convergence: 5434.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.2554 -0.3922 -0.1626  0.0293 10.4215
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  22.79    4.774
##  Residual             231.38   15.211
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.96469    1.28919   0.748
## Time         2.62507    0.51610   5.086
## SES          0.01034    1.30103   0.008
## Time:SES     0.64840    0.51986   1.247
##
## Correlation of Fixed Effects:
##          (Intr) Time   SES
## Time     -0.733
## SES      -0.002  0.003
## Time:SES  0.003 -0.005 -0.733
```

Fixed Effects

Intercept: 0.96; mean of number of utterances with letters for children with mean level SES at Session 1

Time: 2.63; increase in number of utterances with letters every year, when SES at mean level

SES: 0.01; increase in number of utterances for every 1 unit increase in SES at Session 1

Time:SES: 0.65; the change in the relationship in the time slope for every 1 unit increase in SES

My SES variable is already centered.

```
anova(mod2a,mod2b)
```

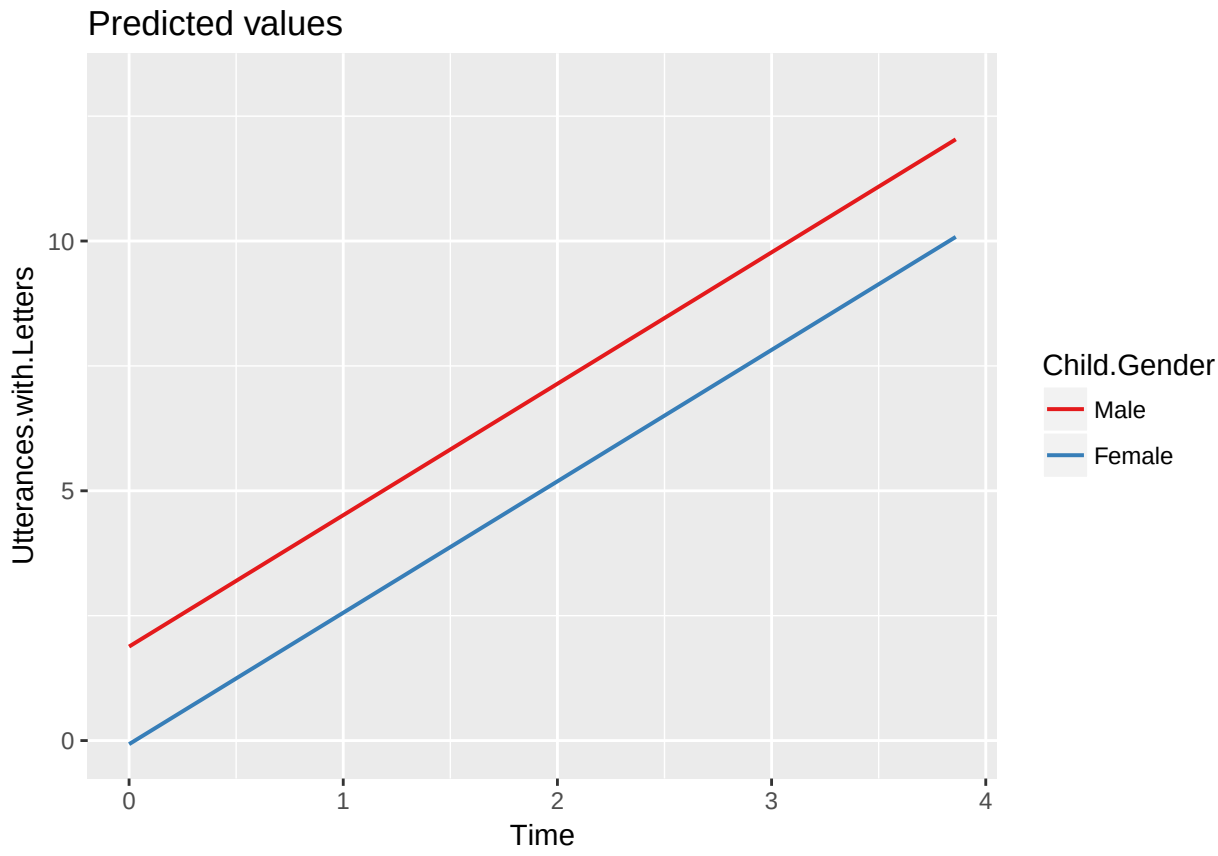
```
## refitting model(s) with ML (instead of REML)
```

```
## Data: children
## Models:
## mod2a: Utterances.with.Letters ~ Time + SES + (1 | Subject)
## mod2b: Utterances.with.Letters ~ Time + SES + Time * SES + (1 | Subject)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## mod2a  5 5449.9 5472.3 -2719.9   5439.9
## mod2b  6 5450.3 5477.2 -2719.2   5438.3 1.5589      1     0.2118
```

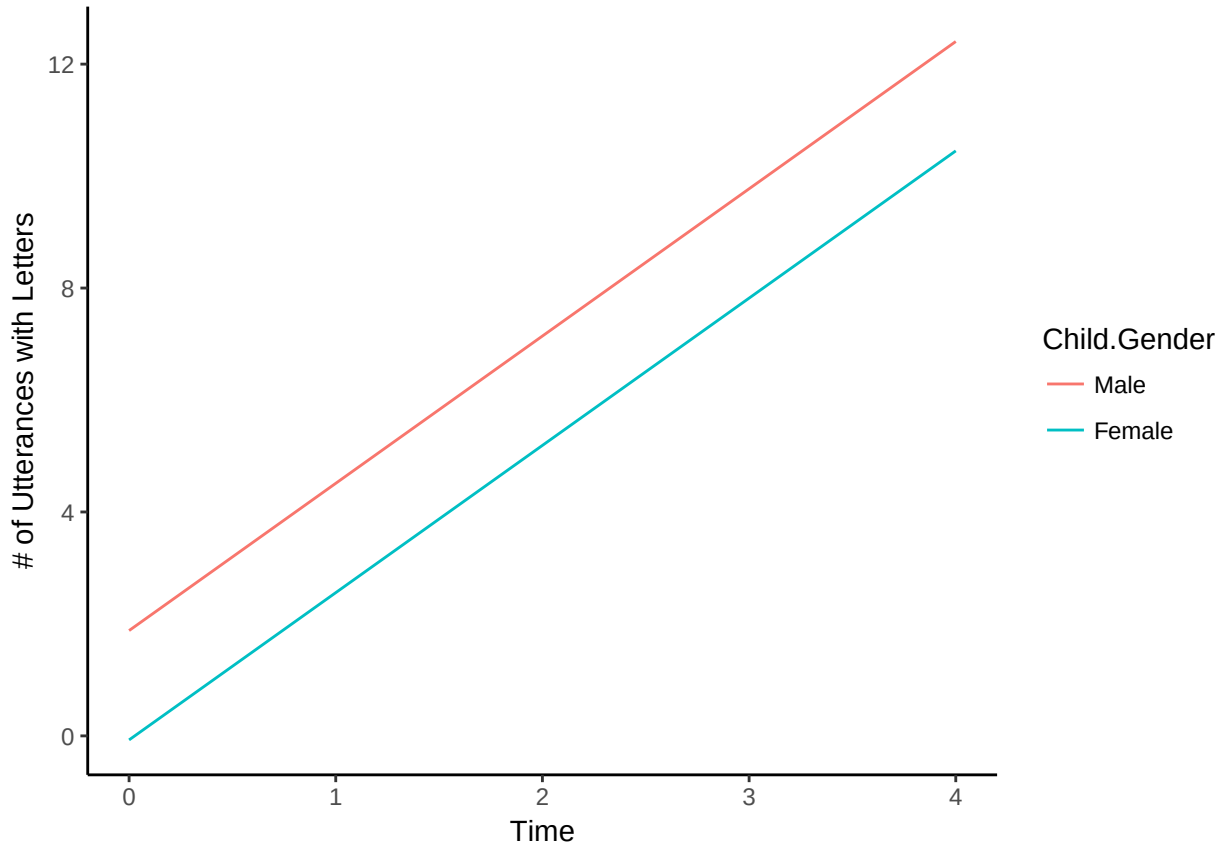
Again, the likelihood ratio test suggests that simpler model, where covariate only predicts intercept, is preferred.

3. Graph both of your final models for the continuous and nominal models above.

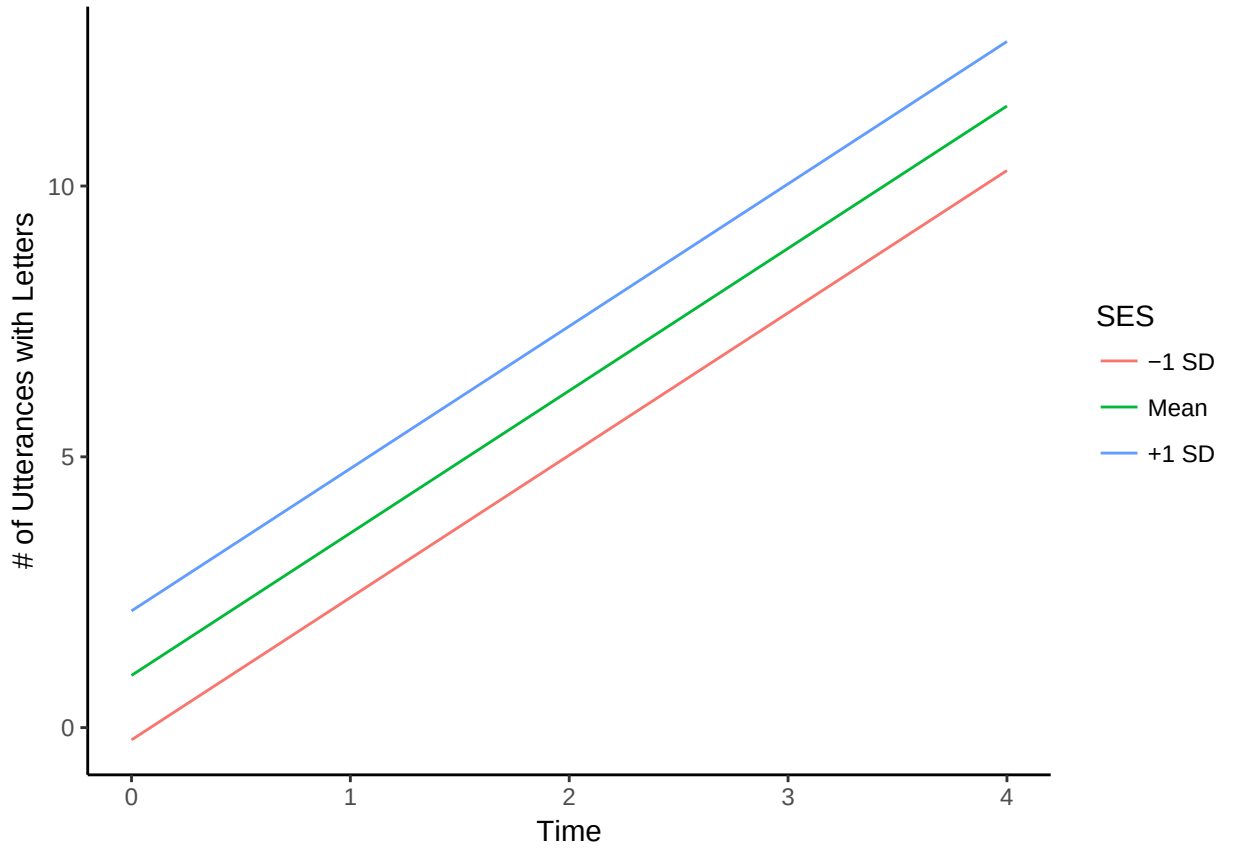
```
# Graphing nominal, where gender only predict the intercept
children$Child.Gender <- relevel(children$Child.Gender, ref = "Male")
sjp.lmer(mod1a, type = "pred.fe", var = c("Time", "Child.Gender"), facet = FALSE, show.scatter = FALSE)
```




```
fixed.frame <- data.frame(expand.grid(Time = seq(0,4,1), Child.Gender = c("Male", "Female"))) %>%
  mutate(pred = predict(mod1a, newdata = ., re.form = NA))

nom <- ggplot(aes(x = Time, y = pred, color = Child.Gender), data = fixed.frame) +
  geom_line() +
  labs(x = "Time", y = "# of Utterances with Letters") +
  theme_classic()
nom
```




```
# Graphing continuous, where SES only predicts the intercept
fixed.frame2 <- letters %>%
  summarise(mean = mean(SES, na.rm = T), sd = sd(SES, na.rm = T))
fixed.frame2 <- data.frame(
  expand.grid(
    Time = seq(0,4,1),
    SES = c(fixed.frame2$mean - fixed.frame2$sd,
            fixed.frame2$mean,
            fixed.frame2$mean + fixed.frame2$sd))) %>%
  mutate(pred = predict(mod2a, newdata = ., re.form = NA))

fixed.frame2$SES <- as.factor(fixed.frame2$SES)
levels(fixed.frame2$SES) <- c("-1 SD","Mean", "+1 SD")

con <- ggplot(aes(x = Time, y = pred, color = SES), data = fixed.frame2) +
  geom_line() +
  labs(x = "Time", y = "# of Utterances with Letters") +
  theme_classic()
con
```



4. Calculate confidence intervals around your estimates for your final models

```
#Confidence intervals around nominal model
mod1a.ci <- confint(mod1a, level = .95, oldNames = F, method = "boot", nsim = 1000)
```

```
## Computing bootstrap confidence intervals ...
```

```
broom::tidy(mod1a.ci)
```

```
##                .rownames    X2.5..   X97.5..
## 1 sd_(Intercept)|Subject  2.804677  6.489865
## 2                  sigma 14.372709 16.122581
## 3            (Intercept) -1.066993  4.943337
## 4                   Time  1.561244  3.650354
## 5     Child.GenderFemale -5.358063  1.505485
```

```
#Confidence intervals around continuous model
mod2a.ci <- confint(mod2a, level = .95, oldNames = F, method = "boot", nsim = 1000)
```

```
## Computing bootstrap confidence intervals ...
```

```
## Warning in optwrap(object@optinfo$optimizer, ff, x0, lower = lower, control
## = control$optCtrl, : convergence code 3 from bobyqa: bobyqa -- a trust
## region step failed to reduce q
```

```
## Warning in optwrap(object@optinfo$optimizer, ff, x0, lower = lower, control
## = control$optCtrl, : convergence code 3 from bobyqa: bobyqa -- a trust
## region step failed to reduce q

## Warning in optwrap(object@optinfo$optimizer, ff, x0, lower = lower, control
## = control$optCtrl, : convergence code 3 from bobyqa: bobyqa -- a trust
## region step failed to reduce q
```

```
broom::tidy(mod2a.ci)
```

```
##                .rownames     X2.5..   X97.5..
## 1 sd_(Intercept)|Subject  2.9151774  6.478857
## 2                  sigma 14.3322945 16.067017
## 3            (Intercept) -1.6916625  3.323114
## 4                   Time  1.6409310  3.632219
## 5                    SES -0.6134671  2.929835
```

5. Include both types of covariates in a single model. How does your interpretation of parameters change?

```
mod3 <- lmer(Utterances.with.Letters ~ Time*Child.Gender*SES + (1|Subject), data = children)
summary(mod3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Utterances.with.Letters ~ Time * Child.Gender * SES + (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5414.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.4280 -0.3531 -0.1579  0.0137 10.4621
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  22.7     4.764
##  Residual             229.3    15.144
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                    0.4702     1.7944   0.262
## Time                           3.2063     0.7205   4.450
## Child.GenderFemale             1.3477     2.6002   0.518
## SES                           -1.2529     2.1517  -0.582
## Time:Child.GenderFemale       -1.5853     1.0406  -1.523
## Time:SES                       2.0292     0.8641   2.348
## Child.GenderFemale:SES         2.1404     2.7117   0.789
## Time:Child.GenderFemale:SES   -2.3718     1.0856  -2.185
##
## Correlation of Fixed Effects:
```

```
##             (Intr) Time   Chl.GF SES    Tm:C.GF Tm:SES C.GF:S
## Time        -0.733
## Chld.GndrFm -0.690  0.506
## SES         -0.164  0.122  0.113
## Tm:Chld.GnF  0.507 -0.692 -0.732 -0.084
## Time:SES     0.122 -0.171 -0.084 -0.733  0.119
## Chld.GF:SES  0.130 -0.097 -0.033 -0.793  0.026   0.582
## Tm:C.GF:SES -0.097  0.136  0.026  0.584 -0.039  -0.796 -0.733
```

Fixed Effects

Intercept: 0.47; mean of number of utterances with letters for males with mean level SES at Session 1

Time: 3.21; increase in number of utterances with letters every year, for males and when SES at mean level

Child.Gender: 1.35; the difference in the number of utterances with letters between males and females when SES at mean level

SES: -1.25; decrease, for males, in number of utterances for every 1 unit increase in SES at Session 1

Session:Child.Gender: -1.59; the difference between the slopes for males and females when SES at mean level

Time:SES: 2.03; the change in the slope of Time, for males, for every 1 unit increase in SES

Child.Gender:SES: 2.14; the difference between the relationship of SES and the number of utterances with letters for males and females at Session 1

Time:Child.Gender:SES: -2.37; the difference between the interaction of SES and Time for males and females

6. If you have one available, introduce a time-varying covariate.

```
mod4 <- lmer(Utterances.with.Letters ~ Time + Utterances + (1|Subject), data = children)
```

```
## Warning: Some predictor variables are on very different scales: consider
## rescaling
```

```
summary(mod4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Utterances.with.Letters ~ Time + Utterances + (1 | Subject)
##    Data: children
##
## REML criterion at convergence: 5436.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3610 -0.3866 -0.1699  0.0560 10.6000
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)  14.58    3.818
##  Residual             230.71   15.189
## Number of obs: 652, groups:  Subject, 55
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) -0.1856687  1.2589451  -0.147
```

```
## Time         0.5827555  0.7103248   0.820
## Utterances   0.0033089  0.0007916   4.180
##
## Correlation of Fixed Effects:
##            (Intr) Time
## Time       -0.393
## Utterances -0.219 -0.688
## fit warnings:
## Some predictor variables are on very different scales: consider rescaling
```

Fixed Effects

Intercept: -0.19; mean of number of utterances with letters for children at Session 1

Time: 0.58; increase in number of utterances with letters every year

Utterances: 0.0033; increase in number of utterances for every 1 unit increase number of utterances