Orientation: L1 vs L2

## Data analysis

We have 30 subjects for L1 and 122 for L2 for a total of 2432 observations.

After removal of entries with missing data due to wrong answers, we have 30 subjects for L1 and 121 subjects for L2 for a total of 1986 observations left.

We set up a categorical variable, Lang, to inform about the speakers L1 and L2 and a another categorical variable, type, to inform about the type of question MATCH and MISMATCH. Each different question is identified by the variable Order that we treat as a categorical variable as well, since we assume no influence of the order in which the question was posed. Finally, we record in the variable List the list from which the question is drawn, that is, list A or list B. The response variable is reaction time measured in milliseconds.

We are not intersted in the effect of the specific questions nor in those of the specific subjects. However, we are are interested in knowing whether these variables are important for the variance of responses. Hence, we specify a mixed effects model with random effects associated to the subjects and to the question order and with fixed effects associated with type and Lang.

In the experiment, we assigned randomly subjects to the two lists. Then we drew 8 questions per each type (matched/mismatched) from each list. Hence question order is nested within type and List and subjects (Name) are nested within List. Within the list nesting subjects and questions are fully crossed. We fitted our model in R using the package lme4 [add reference]. In the syntax of that package the linear mixed model looks as follows: lmer(log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order), data = DLM). We log-transformed the response variable as diagnostic plots (conditional and marginal residuals and qqplot) improved considerably. A Gamma family transformation with log link function yield much worse diagnostic plots. The distribution of reaction times is right skewed and never reaches zero. Hence, logarithm transformation or generalized linear mixed models with Gamma family and log link function are commonly choice.

lmm<- lmer(log(Time) ~ type + Lang + (1 | List:Name)+ (1 | List:type:Order), data = DLM,REML=FALSE)  
#print(anova(lmm),signif.stars=TRUE)  
summary(lmm, transform="log")

## Warning in summary.merMod(lmm, transform = "log"): additional  
## arguments ignored

## Linear mixed model fit by maximum likelihood ['lmerMod']  
## Formula:   
## log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Data: DLM  
##   
## AIC BIC logLik deviance df.resid   
## 2297.8 2331.3 -1142.9 2285.8 1980   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.433 -0.642 -0.107 0.540 3.168   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0642 0.253   
## List:type:Order (Intercept) 0.0219 0.148   
## Residual 0.1554 0.394   
## Number of obs: 1986, groups: List:Name, 151; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 6.8514 0.0457 150.0  
## typeMISMATCH -0.0021 0.0553 0.0  
## LangL1 -0.3420 0.0560 -6.1  
##   
## Correlation of Fixed Effects:  
## (Intr) tMISMA  
## typMISMATCH -0.602   
## LangL1 -0.253 -0.002

We observe that the two random effects, subjects and questions, account for a good portion of the standard deviation of the residuals: 0.2569, 0.1553, 0.3839 respectively. The model is translated by means of a significant intercept. The estimated effects are in the transformed scale.

Note that due to the removal of entries with missing values, our experimental set up is not balanced. Hence, we do not report p-values for F or t statistics as there are no analytical results for null distributions of parameter estimates in complex situations (e.g., unbalanced or partially crossed designs) [cite Bates @Article{, title = {Fitting Linear Mixed-Effects Models Using {lme4}}, author = {Douglas Bates and Martin M{\"a}chler and Ben Bolker and Steve Walker}, journal = {Journal of Statistical Software}, year = {2015}, volume = {67}, number = {1}, pages = {1--48}, doi = {10.18637/jss.v067.i01}, }]. The likelihood ratio test for each individual factors results non significant at a level of 0.05 for the type (0.8815) and for the interaction type:Lang (0.3291). The other fixed and random effects are all significant. The package lmerTest makes availbale an analysis of variance of type III with Satterthwaite approximation for degrees of freedom and a backward elimination of non-significant effects. Both these two procedures confirm our results that type and type:Lang are not significant.

lmm.1 <- update(lmm, .~. - type)  
anova(lmm,lmm.1)

## Data: DLM  
## Models:  
## lmm.1: log(Time) ~ Lang + (1 | List:Name) + (1 | List:type:Order)  
## lmm: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm.1 5 2296 2324 -1143 2286   
## lmm 6 2298 2331 -1143 2286 0 1 0.97

lmm.2 <- update(lmm, .~. - Lang)  
anova(lmm,lmm.2)

## Data: DLM  
## Models:  
## lmm.2: log(Time) ~ type + (1 | List:Name) + (1 | List:type:Order)  
## lmm: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## lmm.2 5 2329 2357 -1160 2319   
## lmm 6 2298 2331 -1143 2286 33.4 1 7.5e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

lmm.3 <- update(lmm, .~. + (type+Lang)^2)  
anova(lmm,lmm.3)

## Data: DLM  
## Models:  
## lmm: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## lmm.3: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order) +   
## lmm.3: type:Lang  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## lmm 6 2298 2331 -1143 2286   
## lmm.3 7 2298 2337 -1142 2284 1.68 1 0.2

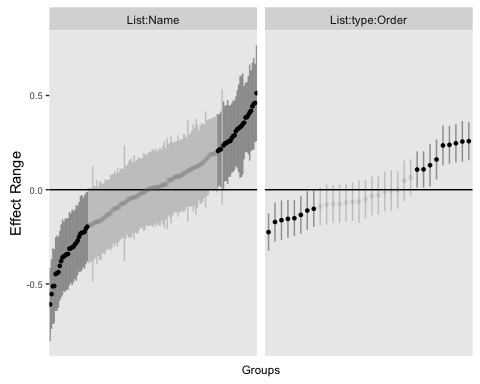
lmm.4 <- update(lmm, .~. - (1|List:Name))  
anova(lmm,lmm.4)

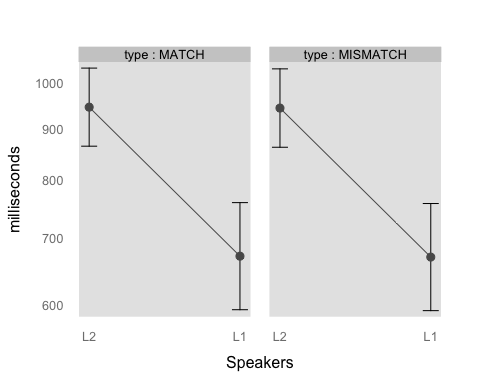
## Data: DLM  
## Models:  
## lmm.4: log(Time) ~ type + Lang + (1 | List:type:Order)  
## lmm: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## lmm.4 5 2680 2708 -1335 2670   
## lmm 6 2298 2331 -1143 2286 384 1 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

lmm.5 <- update(lmm, .~. - (1|List:type:Order))  
anova(lmm,lmm.5)

## Data: DLM  
## Models:  
## lmm.5: log(Time) ~ type + Lang + (1 | List:Name)  
## lmm: log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## lmm.5 5 2474 2502 -1232 2464   
## lmm 6 2298 2331 -1143 2286 178 1 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

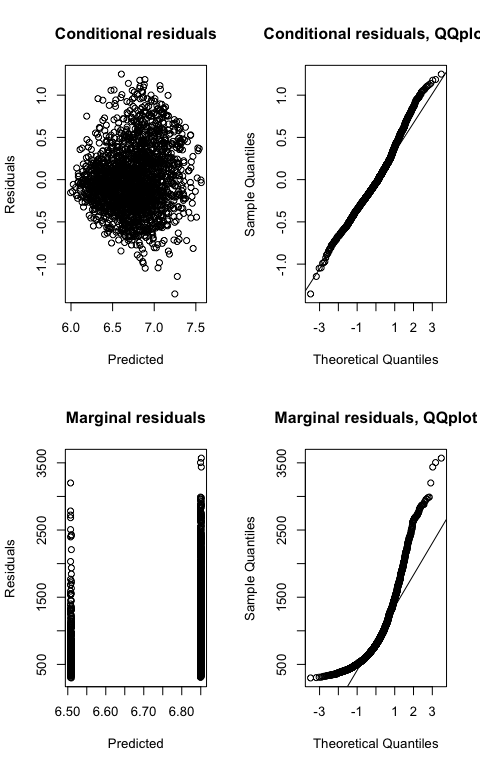
We plot the random and the fixed effects together with their 95% confidence levels in the Figure. The confidence levels are obtained by the estimated standard deviation and the normal distribution. However, other analysis with boostrapped confidence intervals confirmed the same conclusions.





From the figure we can observe that the random effects cause often a significant effect by shifting the intercept by a quantity larger than zero. For the fixed effects we transformed back the effects in the linear time scale of milliseconds. It is evident that L1 speakers have a shorter reaction time and that the type does not have a significant impact.

## Diagnostic plots



The joint qqplot looks normal. The marginal looks less nice.

## Tables with p-values (Satterwhite approximation)

It could be possible to include a syntetic ANOVA table like this:

require(lmerTest)  
lmm<- lmer(log(Time) ~ type + Lang + (1 | List:Name)+ (1 | List:type:Order), data = DLM,REML=FALSE)  
anova(lmm)

## Analysis of Variance Table of type III with Satterthwaite   
## approximation for degrees of freedom  
## Sum Sq Mean Sq NumDF DenDF F.value Pr(>F)   
## type 0.0 0.0 1 31.5 0.0 0.97   
## Lang 5.8 5.8 1 141.2 37.3 9.3e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

however, as explained above it is controversial. It is also possible to have t-test and p-values like this:

summary(lmm)

## Linear mixed model fit by maximum likelihood t-tests use  
## Satterthwaite approximations to degrees of freedom [lmerMod]  
## Formula:   
## log(Time) ~ type + Lang + (1 | List:Name) + (1 | List:type:Order)  
## Data: DLM  
##   
## AIC BIC logLik deviance df.resid   
## 2297.8 2331.3 -1142.9 2285.8 1980   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.433 -0.642 -0.107 0.540 3.168   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## List:Name (Intercept) 0.0642 0.253   
## List:type:Order (Intercept) 0.0219 0.148   
## Residual 0.1554 0.394   
## Number of obs: 1986, groups: List:Name, 151; List:type:Order, 32  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 6.8514 0.0457 56.3000 149.96 < 2e-16 \*\*\*  
## typeMISMATCH -0.0021 0.0553 31.5000 -0.04 0.97   
## LangL1 -0.3420 0.0560 141.2000 -6.11 9.3e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) tMISMA  
## typMISMATCH -0.602   
## LangL1 -0.253 -0.002