Orientation: L1 vs L2

Data analysis

We have 30 subjects for L1 and 91 for L2 for a total of 1936 observations. After removal of entries with missing data due to wrong answers, we have 1641 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
##   1826.8   1859.2   -907.4   1814.8     1635
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -3.509 -0.656 -0.108  0.540  3.102
##
## Random effects:
##  Groups          Name        Variance Std.Dev.
##  List:Name       (Intercept) 0.0650   0.255
##  List:type:Order (Intercept) 0.0227   0.151
##  Residual                    0.1474   0.384
```

```
## Number of obs: 1641, groups:  List:Name, 121; List:type:Order, 32
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   6.82282    0.04870   140.1
## typeMISMATCH  0.00846    0.05673     0.1
## LangL1       -0.31873    0.05816    -5.5
##
## Correlation of Fixed Effects:
##             (Intr) tMISMA
## typMISMATCH -0.578
## LangL1      -0.309 -0.004
```

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 1825 1852   -907     1815
## lmm    6 1827 1859   -907     1815  0.02      1       0.88
```

```
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 1852 1879   -921     1842
## lmm    6 1827 1859   -907     1815    27      1    2.1e-07 ***
## ---
## 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 1827 1859   -907     1815
## lmm.3  7 1828 1866   -907     1814  0.95      1       0.33
```

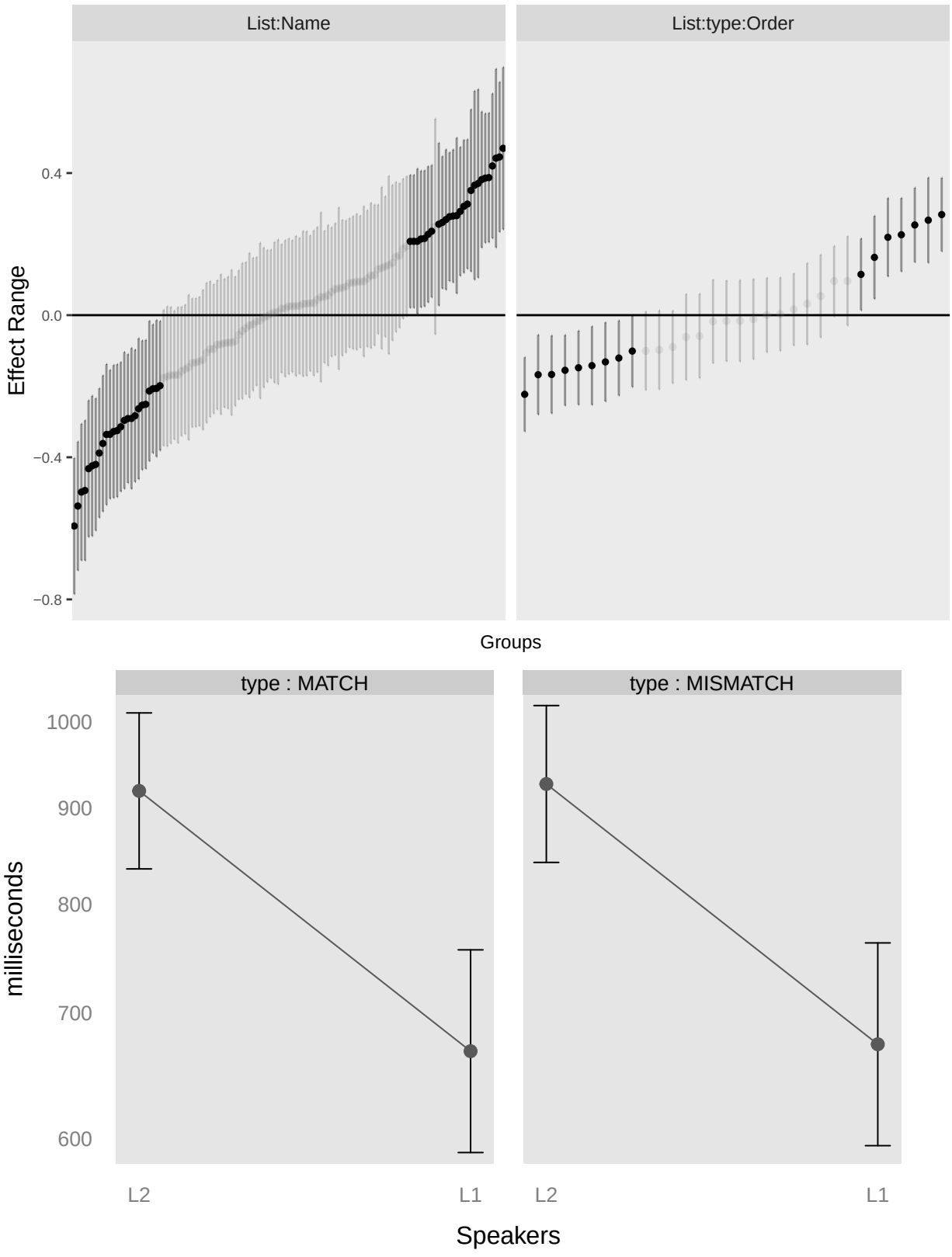
```
lmm.4 <- update(lmm, .~. - (1|List:Name))
anova(lmm,lmm.4)
```

```
## 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 2171 2198  -1080     2161
## lmm    6 1827 1859   -907     1815   346      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 1986 2013   -988     1976
## lmm    6 1827 1859   -907     1815   161      1     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

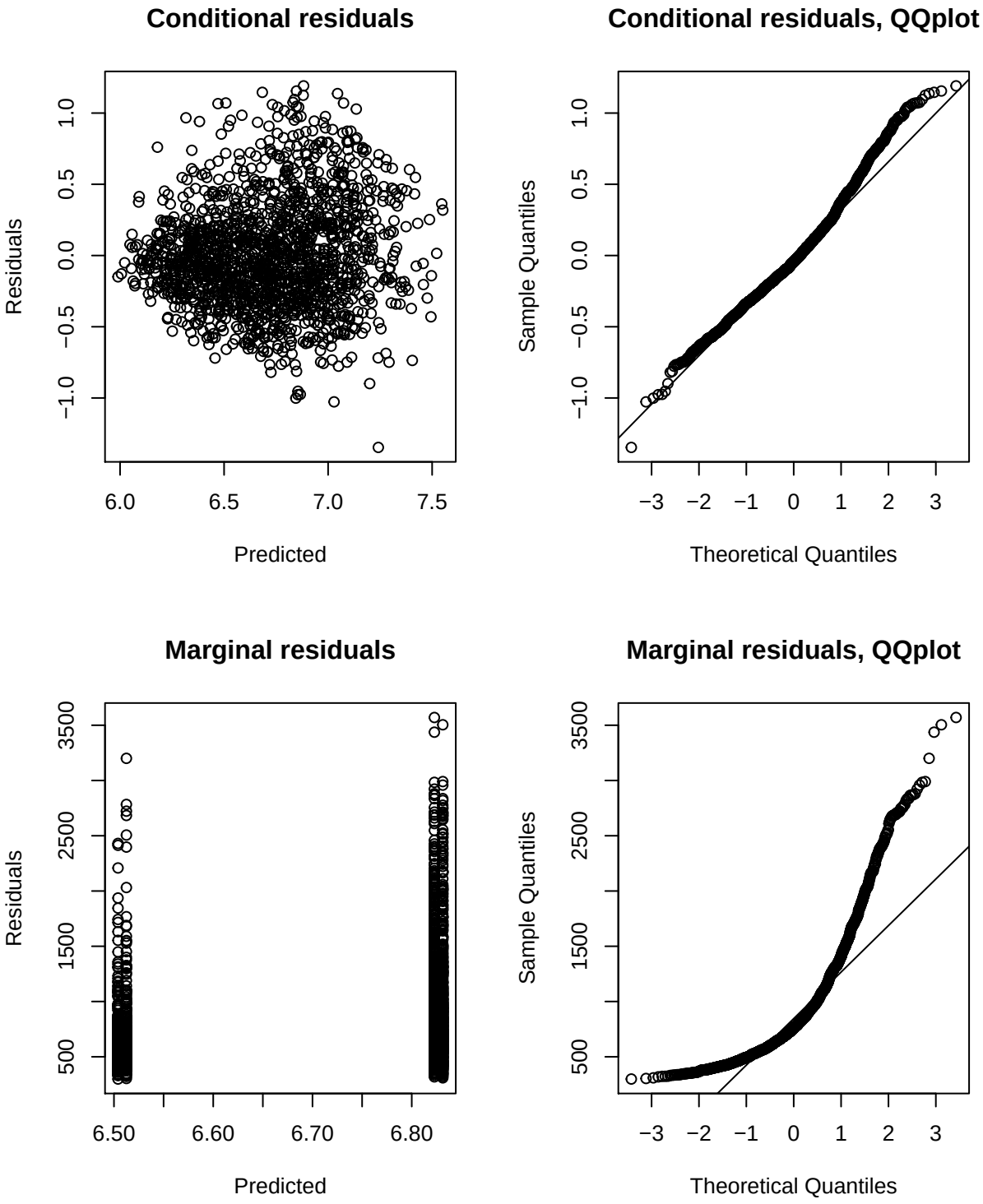
We plot the random and the fixed effects together with their 95% confidence levels in the Figure. The confidence levels are obtained by the estimated standard deviation and the normal distribution. However, other analysis with boostrapped confidence intervals confirmed the same conclusions.



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   0.00    0.00     1  31.7    0.02    0.88
## Lang   4.43    4.43     1 116.1   30.04 2.5e-07 ***
## ---
## 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
##   1826.8   1859.2   -907.4   1814.8     1635
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -3.509 -0.656 -0.108  0.540  3.102
##
## Random effects:
##  Groups          Name        Variance Std.Dev.
##  List:Name       (Intercept) 0.0650   0.255
##  List:type:Order (Intercept) 0.0227   0.151
##  Residual                    0.1474   0.384
## Number of obs: 1641, groups:  List:Name, 121; List:type:Order, 32
##
## Fixed effects:
##               Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)    6.82282    0.04870  63.00000  140.10  < 2e-16 ***
## typeMISMATCH   0.00846    0.05673  31.70000    0.15     0.88
## LangL1        -0.31873    0.05816 116.10000   -5.48  2.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) tMISMA
## typMISMATCH -0.578
## LangL1      -0.309 -0.004
```