

# Bilingualism

## R Markdown

We load the data and remove the cases with NA values.

```
str(D1)

## 'data.frame': 104 obs. of 16 variables:
## $ Name : Factor w/ 104 levels "Albert Denk",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ CoR : Factor w/ 2 levels "0","1": 1 2 2 2 2 2 1 1 2 ...
## $ Hand : Factor w/ 2 levels "0","1": 2 2 2 1 2 2 2 2 2 ...
## $ EO : Factor w/ 2 levels "0","1": 1 1 2 2 2 2 2 1 1 2 ...
## $ List : Factor w/ 2 levels "0","1": 1 1 2 1 1 2 2 2 2 ...
## $ CEF : Ord.factor w/ 3 levels "3"<"4"<"5": 2 3 2 3 2 3 1 3 3 3 ...
## $ SRRC : num 3.8 4 4 4 4 5 3 5 3.5 5 ...
## $ PRE : num 0.33 0.5 1 1 0.78 1 0.83 1 0.88 1 ...
## $ POST1 : num 0.94 1 1 1 0.94 1 1 1 1 1 ...
## $ POST2 : num 1 0.88 1 1 0.88 1 1 0.88 0.88 1 ...
## $ STAY : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
## $ LEAYRS: num 4.2 5 4 3.08 5 5 1.5 7 15 10 ...
## $ HRSD : num 0.25 0.75 0.5 0.125 0 1.5 0 4 1 0.125 ...
## $ RPV : num 5 5 3.5 5 5 2.5 3 2.5 4 2.5 ...
## $ AMGE : num 3.5 4 4.5 4 4.5 5 4.5 5 5 4.5 ...
## $ AMSP : num 4.1 3.67 4.17 4.33 3.33 ...
```

```
L0 <- read.csv2("size.csv",na.strings="#NUM!",sep=";",dec=".",header=TRUE)
L0$Median <- apply(L0[c(3:10,15:21)],1,median,na.rm = TRUE)
LM<-L0 %>% select(Name,MEDIAN.2)
names(LM)[2]<-"time"
#LMM<-L0 %>% select(Name, Median)
LMM<-L0 %>% select(Name, MISMATCH.RTdominantanswerMedian)
#LMM<-L0 %>% select(Name, MEDIAN)#MISMATCH.RTdominantanswerMedian)
names(LMM)[2]<-"time"

SM<-left_join(D1,LM,by="Name")
```

```
## Warning in left_join_impl(x, y, by$x, by$y): joining factors with
## different levels, coercing to character vector
```

```
SMM<-left_join(D1,LMM,by="Name")
```

```
## Warning in left_join_impl(x, y, by$x, by$y): joining factors with
## different levels, coercing to character vector
```

```
S<-rbind(data.frame(SM,type="MATCH"),data.frame(SMM,type="MISMATCH"))
S<-na.omit(S)
```

In total we have 225 data points.

## Multiple Linear Regression

We try first a first order multiple linear regression model

```
S<-S%>%select(-Name)
m<-lm(data=S,time~.)
print(anova(m),signif.stars=TRUE)
```

```
## Analysis of Variance Table
##
## Response: time
##          Df    Sum Sq Mean Sq F value   Pr(>F)
## CoR        1    193909   193909     0.77 0.38050
## Hand        1    412118   412118     1.64 0.20168
## EO          1     70029    70029     0.28 0.59790
## List        1    722139   722139     2.88 0.09162 .
## CEF         2    222770   111385     0.44 0.64212
## SRRC        1     52542    52542     0.21 0.64775
## PRE         1      6535     6535     0.03 0.87195
## POST1       1       318       318     0.00 0.97165
## POST2       1     10381    10381     0.04 0.83902
## STAY        1     75096    75096     0.30 0.58496
## LEAYRS      1     71174    71174     0.28 0.59493
## HRSD        1    505911   505911     2.02 0.15741
## RPV         1    308393   308393     1.23 0.26908
## AMGE        1     18566    18566     0.07 0.78589
## AMSP        1   1614706  1614706     6.44 0.01211 *
## type        1   3062183  3062183    12.21 0.00061 ***
## Residuals 162  40622652   250757
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

and then we include all second order terms:

```
library(MASS)
m2<-lm(data=S,time~.^2)
print(anova(m2),signif.stars=TRUE)
```

```
## Analysis of Variance Table
##
## Response: time
##          Df    Sum Sq Mean Sq F value   Pr(>F)
## CoR        1    193909   193909     1.43 0.23614
## Hand        1    412118   412118     3.03 0.08582 .
## EO          1     70029    70029     0.52 0.47514
## List        1    722139   722139     5.31 0.02400 *
## CEF         2    222770   111385     0.82 0.44460
## SRRC        1     52542    52542     0.39 0.53601
## PRE         1      6535     6535     0.05 0.82704
## POST1       1       318       318     0.00 0.96157
## POST2       1     10381    10381     0.08 0.78303
## STAY        1     75096    75096     0.55 0.45964
## LEAYRS      1     71174    71174     0.52 0.47156
```

## HRSD	1	505911	505911	3.72	0.05756	.
## RPV	1	308393	308393	2.27	0.13627	
## AMGE	1	18566	18566	0.14	0.71274	
## AMSP	1	1614706	1614706	11.88	0.00094	***
## type	1	3062183	3062183	22.53	1.0e-05	***
## CoR:Hand	1	68343	68343	0.50	0.48048	
## CoR:List	1	453552	453552	3.34	0.07180	.
## CoR:CEF	2	733406	366703	2.70	0.07403	.
## CoR:SRRC	1	711153	711153	5.23	0.02506	*
## CoR:PRE	1	16651	16651	0.12	0.72731	
## CoR:POST1	1	216777	216777	1.60	0.21061	
## CoR:POST2	1	1670293	1670293	12.29	0.00078	***
## CoR:STAY	1	71719	71719	0.53	0.46988	
## CoR:LEAYRS	1	111352	111352	0.82	0.36833	
## CoR:HRSD	1	245742	245742	1.81	0.18287	
## CoR:RPV	1	20949	20949	0.15	0.69574	
## CoR:AMGE	1	626630	626630	4.61	0.03509	*
## CoR:AMSP	1	63472	63472	0.47	0.49650	
## CoR:type	1	17003	17003	0.13	0.72457	
## Hand:EO	1	144592	144592	1.06	0.30571	
## Hand:List	1	91829	91829	0.68	0.41374	
## Hand:CEF	2	407254	203627	1.50	0.23028	
## Hand:SRRC	1	94611	94611	0.70	0.40678	
## Hand:PRE	1	41167	41167	0.30	0.58373	
## Hand:POST1	1	348550	348550	2.56	0.11358	
## Hand:POST2	1	43291	43291	0.32	0.57420	
## Hand:STAY	1	4216	4216	0.03	0.86067	
## Hand:type	1	110469	110469	0.81	0.37023	
## EO:List	1	334780	334780	2.46	0.12084	
## EO:CEF	2	311029	155515	1.14	0.32407	
## EO:SRRC	1	26413	26413	0.19	0.66061	
## EO:PRE	1	1950283	1950283	14.35	0.00031	***
## EO:POST1	1	588869	588869	4.33	0.04088	*
## EO:POST2	1	2690	2690	0.02	0.88851	
## EO:STAY	1	522499	522499	3.84	0.05371	.
## EO:LEAYRS	1	263196	263196	1.94	0.16825	
## EO:HRSD	1	15929	15929	0.12	0.73306	
## EO:RPV	1	59301	59301	0.44	0.51096	
## EO:AMGE	1	520016	520016	3.83	0.05427	.
## EO:AMSP	1	2329262	2329262	17.14	9.2e-05	***
## EO:type	1	78451	78451	0.58	0.44982	
## List:CEF	2	766757	383379	2.82	0.06604	.
## List:SRRC	1	1122704	1122704	8.26	0.00530	**
## List:PRE	1	957648	957648	7.05	0.00974	**
## List:POST1	1	5785	5785	0.04	0.83712	
## List:POST2	1	1879642	1879642	13.83	0.00039	***
## List:STAY	1	47016	47016	0.35	0.55822	
## List:LEAYRS	1	72322	72322	0.53	0.46802	
## List:HRSD	1	797	797	0.01	0.93915	
## List:RPV	1	289502	289502	2.13	0.14870	
## List:AMGE	1	694408	694408	5.11	0.02677	*
## List:AMSP	1	596117	596117	4.39	0.03969	*
## List:type	1	624678	624678	4.60	0.03536	*
## CEF:SRRC	2	2578585	1289292	9.49	0.00022	***

```

## CEF:PRE      2  657923  328961    2.42 0.09596 .
## CEF:POST1    2  438145  219072    1.61 0.20649
## CEF:POST2    2  822733  411366    3.03 0.05458 .
## CEF:STAY     2  172424   86212    0.63 0.53315
## CEF:LEAYRS   2  464351  232175    1.71 0.18830
## CEF:HRSD     2  208231  104116    0.77 0.46850
## CEF:RPV      1  197012  197012    1.45 0.23246
## CEF:AMGE     1  791578  791578    5.82 0.01831 *
## CEF:AMSP     1   71637   71637    0.53 0.47013
## CEF:type     2  255879  127940    0.94 0.39474
## SRRC:PRE     1 1967958 1967958   14.48 0.00029 ***
## SRRC:POST1   1   22327   22327    0.16 0.68642
## SRRC:POST2   1  230788  230788    1.70 0.19661
## SRRC:STAY    1   87136   87136    0.64 0.42587
## SRRC:LEAYRS  1  187449  187449    1.38 0.24403
## SRRC:HRSD    1   20705   20705    0.15 0.69742
## SRRC:RPV     1   71341   71341    0.52 0.47104
## SRRC:AMGE    1   49342   49342    0.36 0.54866
## SRRC:type    1    295     295    0.00 0.96298
## PRE:type     1   17375   17375    0.13 0.72170
## POST1:type   1  168280  168280    1.24 0.26945
## POST2:type   1   37955   37955    0.28 0.59876
## STAY:type    1    1187    1187    0.01 0.92578
## LEAYRS:type  1   36287   36287    0.27 0.60690
## HRSD:type    1  196854  196854    1.45 0.23264
## RPV:type     1  100917  100917    0.74 0.39165
## AMGE:type    1    8051    8051    0.06 0.80838
## AMSP:type    1  466459  466459    3.43 0.06797 .
## Residuals   73 9920329 135895
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We apply stepwise model search procedure:

```

model.stp<-stepAIC(aov(data=S,time~.^2),scope=list(upper=~.^2,lower=~1),
                  direction="backward",trace=FALSE)
print(model.stp$anova,signif.stars=TRUE)

```

```

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## time ~ (CoR + Hand + EO + List + CEF + SRRC + PRE + POST1 + POST2 +
##       STAY + LEAYRS + HRSD + RPV + AMGE + AMSP + type)^2
##
## Final Model:
## time ~ CoR + Hand + EO + List + CEF + SRRC + PRE + POST1 + POST2 +
##       STAY + LEAYRS + HRSD + RPV + AMGE + AMSP + type + CoR:List +
##       CoR:CEF + CoR:SRRC + CoR:POST1 + CoR:POST2 + CoR:STAY + CoR:HRSD +
##       CoR:RPV + CoR:AMSP + Hand:EO + Hand:List + Hand:SRRC + Hand:PRE +
##       Hand:POST1 + EO:List + EO:CEF + EO:SRRC + EO:PRE + EO:POST1 +
##       EO:STAY + EO:LEAYRS + EO:HRSD + EO:RPV + List:CEF + List:SRRC +
##       List:PRE + List:POST2 + List:STAY + List:LEAYRS + List:HRSD +

```

```
## List:RPV + List:AMGE + List:AMSP + List:type + CEF:POST1 +
## CEF:POST2 + CEF:LEAYRS + CEF:HRSD + CEF:RPV + CEF:AMGE +
## CEF:AMSP + CEF:type + SRRC:PRE + SRRC:POST2 + SRRC:STAY +
## SRRC:LEAYRS + SRRC:RPV + SRRC:AMGE + SRRC:type + POST2:type +
## STAY:type + HRSD:type + AMSP:type
##
##
```

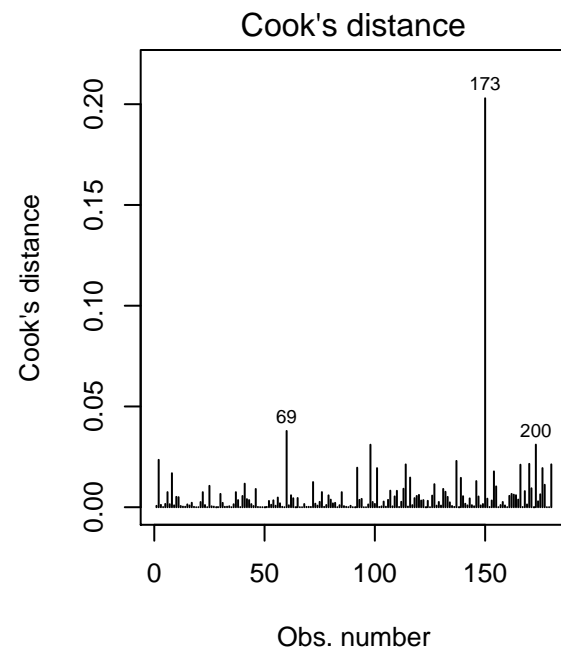
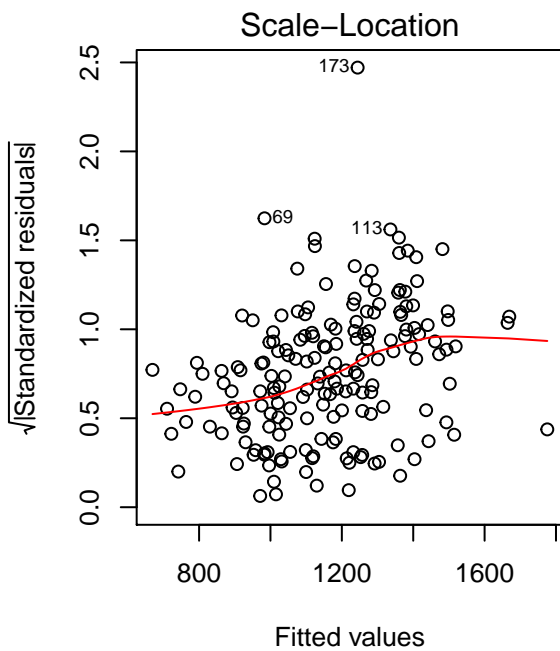
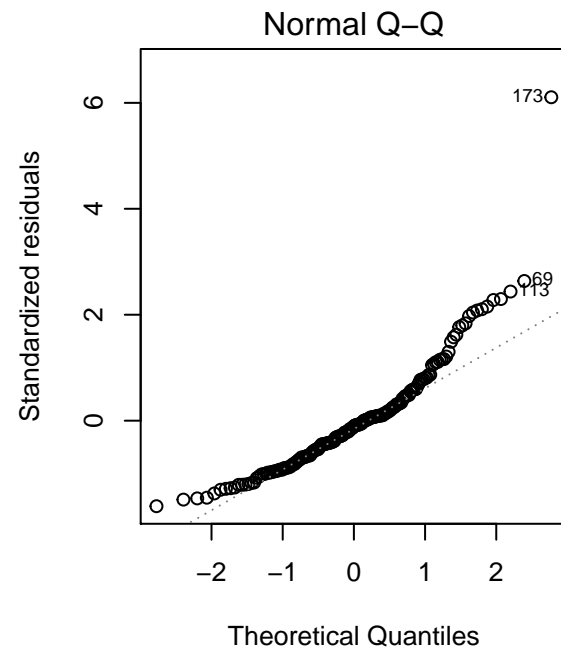
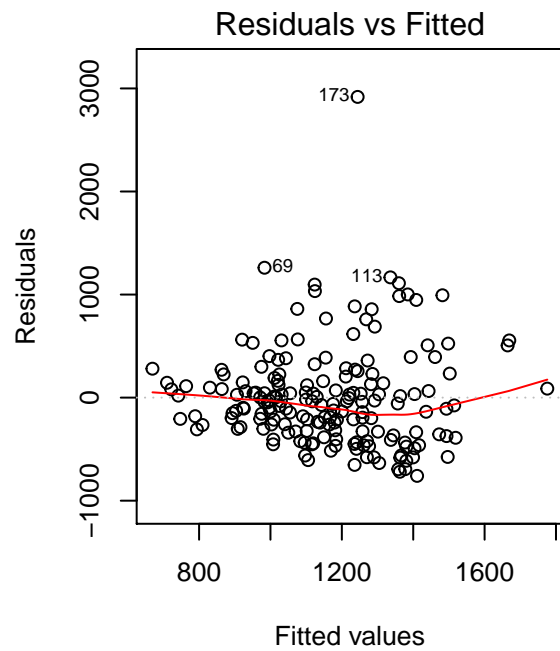
	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
## 1				73	9920329	2179.1
## 2	- AMGE:AMSP	0	0.0000e+00	73	9920329	2179.1
## 3	- RPV:AMSP	0	0.0000e+00	73	9920329	2179.1
## 4	- RPV:AMGE	0	0.0000e+00	73	9920329	2179.1
## 5	- HRSD:AMSP	0	0.0000e+00	73	9920329	2179.1
## 6	- HRSD:AMGE	0	0.0000e+00	73	9920329	2179.1
## 7	- HRSD:RPV	0	0.0000e+00	73	9920329	2179.1
## 8	- LEAYRS:AMSP	0	0.0000e+00	73	9920329	2179.1
## 9	- LEAYRS:AMGE	0	0.0000e+00	73	9920329	2179.1
## 10	- LEAYRS:RPV	0	0.0000e+00	73	9920329	2179.1
## 11	- LEAYRS:HRSD	0	0.0000e+00	73	9920329	2179.1
## 12	- STAY:AMSP	0	0.0000e+00	73	9920329	2179.1
## 13	- STAY:AMGE	0	0.0000e+00	73	9920329	2179.1
## 14	- STAY:RPV	0	0.0000e+00	73	9920329	2179.1
## 15	- STAY:HRSD	0	0.0000e+00	73	9920329	2179.1
## 16	- STAY:LEAYRS	0	0.0000e+00	73	9920329	2179.1
## 17	- POST2:AMSP	0	0.0000e+00	73	9920329	2179.1
## 18	- POST2:AMGE	0	0.0000e+00	73	9920329	2179.1
## 19	- POST2:RPV	0	0.0000e+00	73	9920329	2179.1
## 20	- POST2:HRSD	0	0.0000e+00	73	9920329	2179.1
## 21	- POST2:LEAYRS	0	0.0000e+00	73	9920329	2179.1
## 22	- POST2:STAY	0	0.0000e+00	73	9920329	2179.1
## 23	- POST1:AMSP	0	0.0000e+00	73	9920329	2179.1
## 24	- POST1:AMGE	0	0.0000e+00	73	9920329	2179.1
## 25	- POST1:RPV	0	0.0000e+00	73	9920329	2179.1
## 26	- POST1:HRSD	0	0.0000e+00	73	9920329	2179.1
## 27	- POST1:LEAYRS	0	0.0000e+00	73	9920329	2179.1
## 28	- POST1:STAY	0	0.0000e+00	73	9920329	2179.1
## 29	- POST1:POST2	0	0.0000e+00	73	9920329	2179.1
## 30	- PRE:AMSP	0	0.0000e+00	73	9920329	2179.1
## 31	- PRE:AMGE	0	0.0000e+00	73	9920329	2179.1
## 32	- PRE:RPV	0	0.0000e+00	73	9920329	2179.1
## 33	- PRE:HRSD	0	0.0000e+00	73	9920329	2179.1
## 34	- PRE:LEAYRS	0	0.0000e+00	73	9920329	2179.1
## 35	- PRE:STAY	0	0.0000e+00	73	9920329	2179.1
## 36	- PRE:POST2	0	0.0000e+00	73	9920329	2179.1
## 37	- PRE:POST1	0	0.0000e+00	73	9920329	2179.1
## 38	- SRRC:AMSP	0	0.0000e+00	73	9920329	2179.1
## 39	- EO:AMSP	0	1.0803e-07	73	9920329	2179.1
## 40	- EO:AMGE	0	3.4086e-07	73	9920329	2179.1
## 41	- Hand:AMSP	0	0.0000e+00	73	9920329	2179.1
## 42	- Hand:AMGE	0	0.0000e+00	73	9920329	2179.1
## 43	- Hand:RPV	0	0.0000e+00	73	9920329	2179.1
## 44	- Hand:HRSD	0	0.0000e+00	73	9920329	2179.1
## 45	- Hand:LEAYRS	0	0.0000e+00	73	9920329	2179.1
## 46	- Hand:STAY	0	3.5390e-08	73	9920329	2179.1

## 47	- CoR:EO	0	0.0000e+00	73	9920329	2179.1
## 48	- CEF:STAY	2	3.5553e+03	75	9923885	2175.1
## 49	- CEF:SRRC	2	1.4853e+04	77	9938738	2171.4
## 50	- CEF:PRE	2	2.3450e+04	79	9962188	2167.8
## 51	- Hand:POST2	1	2.1903e+03	80	9964379	2165.9
## 52	- List:POST1	1	9.2679e+03	81	9973646	2164.0
## 53	- PRE:type	1	1.5147e+04	82	9988794	2162.3
## 54	- CoR:Hand	1	1.8482e+04	83	10007276	2160.7
## 55	- EO:POST2	1	2.5133e+04	84	10032409	2159.1
## 56	- SRRC:HRSD	1	1.4139e+04	85	10046548	2157.4
## 57	- EO:type	1	2.9474e+04	86	10076022	2155.9
## 58	- Hand:type	1	2.2460e+04	87	10098482	2154.3
## 59	- CoR:LEAYRS	1	4.3594e+04	88	10142076	2153.1
## 60	- SRRC:POST1	1	4.4720e+04	89	10186796	2151.9
## 61	- POST1:type	1	4.8044e+04	90	10234840	2150.7
## 62	- RPV:type	1	5.6933e+04	91	10291773	2149.7
## 63	- LEAYRS:type	1	5.7023e+04	92	10348796	2148.7
## 64	- AMGE:type	1	6.4067e+04	93	10412862	2147.8
## 65	- CoR:type	1	4.9806e+04	94	10462668	2146.7
## 66	- CoR:PRE	1	6.5757e+04	95	10528426	2145.8
## 67	- CoR:AMGE	1	6.4960e+04	96	10593386	2144.9
## 68	- Hand:CEF	2	2.2012e+05	98	10813508	2144.6

```
#m2b<-lm(data=D1,)
#print(anova(m2b),signif.stars = TRUE)
#anova(m2b,m2)
```

## Diagnostic plots

```
par(mfrow=c(2,2))
plot(m,which=1:4)
```



From the qqplot we see that the point 1071 is an outlier: With hindsight we already remove it above. Afterwards the qqplot looks perfectly normal.

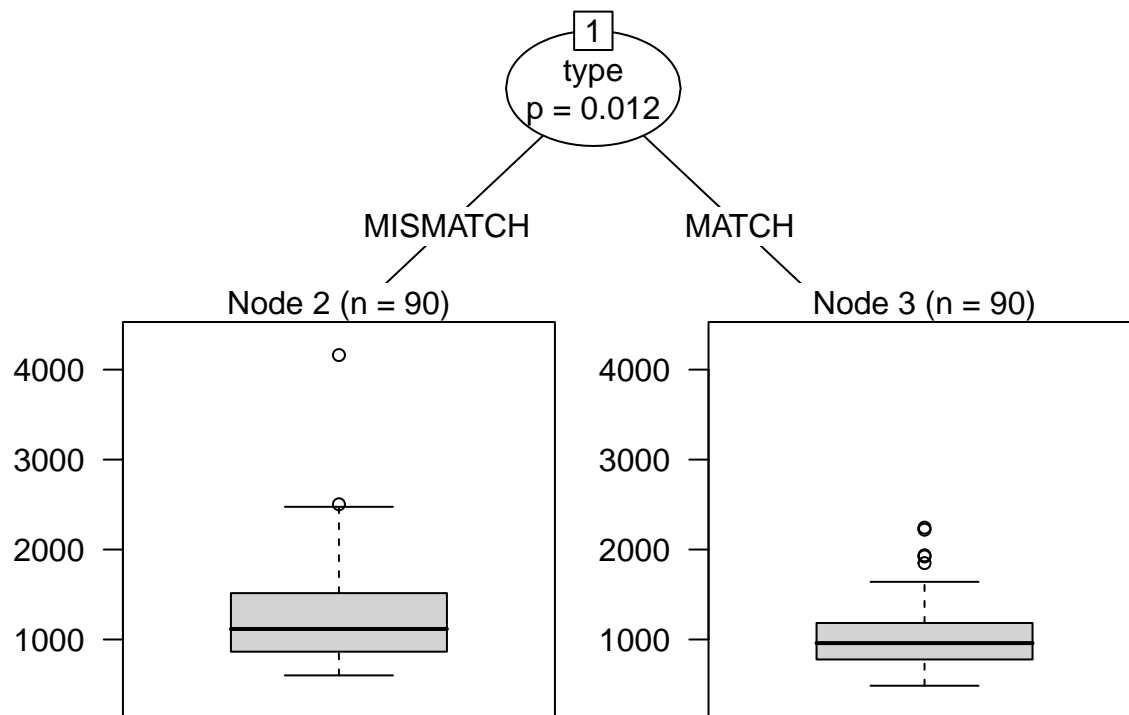
## Regression tree

Another type of analysis by means of regression tree:

```
ct <- ctree(data=S,time~.^2)
ct
```

```
##
## Conditional inference tree with 2 terminal nodes
##
## Response: time
## Inputs: CoR, Hand, EO, List, CEF, SRRC, PRE, POST1, POST2, STAY, LEAYRS, HRSD, RPV, AMGE, AMSP, type
## Number of observations: 180
##
## 1) type == {MISMATCH}; criterion = 0.988, statistic = 11.427
## 2)* weights = 90
## 1) type == {MATCH}
## 3)* weights = 90
```

```
plot(ct)
```



From the plot we understand that regression trees find only two factors POST1 and Hand and their interaction as relevant explain differences in the response time.

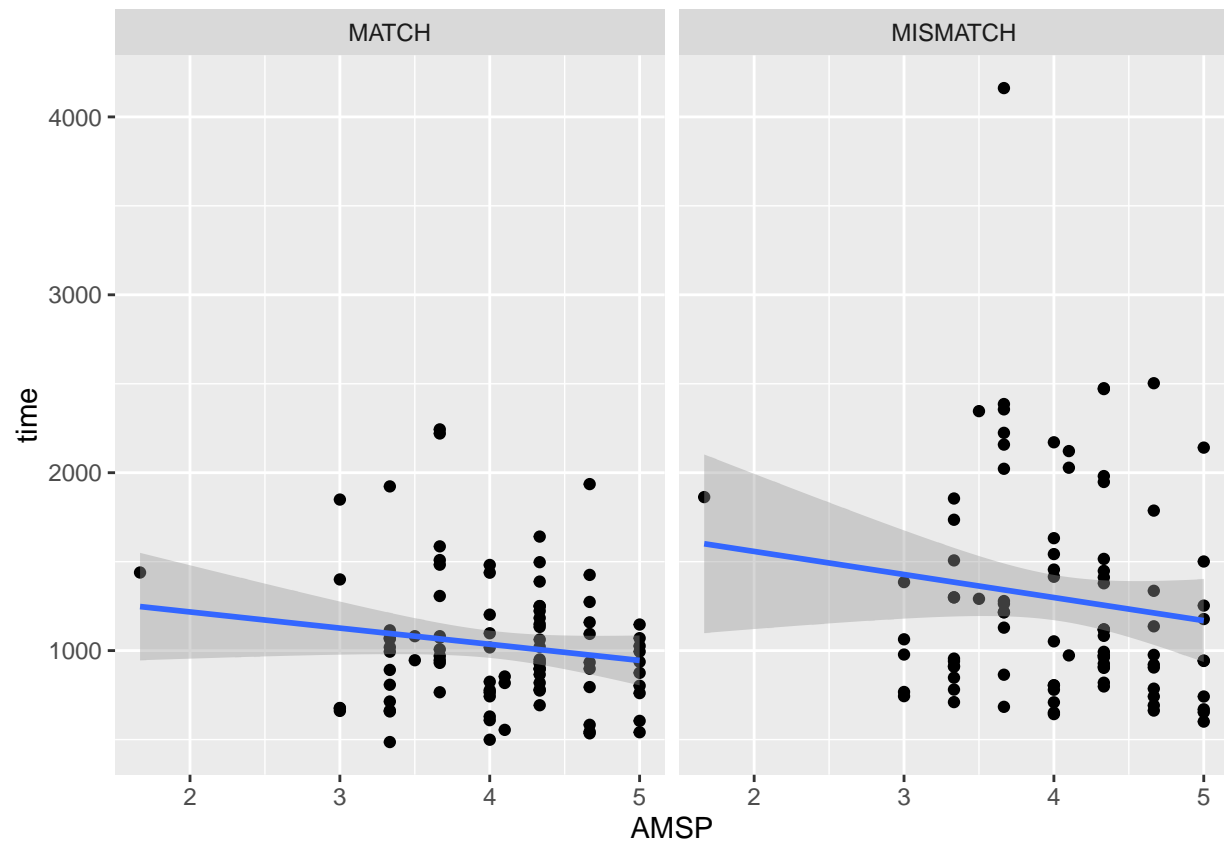
## Data visualization

We visualize the influence of the two factors that have been found explaining the most the response time. We separate the plots in match and mismatch cases, although this factor was found not to be important.

```
q<-ggplot(data=S,aes(y=time,x=AMSP))
q<-q+facet_grid(.~type)
q<-q+geom_point()
q<-q+geom_smooth(method=lm)
```



```
q<-q+theme(legend.position="top")
print(q)
```



We shall conclude that the response time depends on the hand and on the general proficiency in the verbs as assessed by the POST1 test. Right handed subjects with good performance in the POST1 test respond faster.