Mixed Models

Jonathan Harrington

```
library(ez)
library(lme4)
library(multcomp)
source(file.path(pfadu, "phoc.txt"))
```

Mixed Models und die Varianzanalyse

Mixed Models bieten eine flexiblere Alternative zur Varianzanalyse

- Keine within/between Trennung
- Keine Notwendigkeit für ein 'balanced' design
- Die Stufen müssen nicht pro Vpn. vollständig sein
- Es muss nicht über Wiederholungen in der selben Stufe gemittelt werden
- Keine Greenhouse-Geißer Korrektur
- Die Variabilität mehrerer Faktoren kann ausgeklammert werden (in ANOVA nur eines: meistens der Sprecher)
- Eine Mischung aus unabhängigen numerischen und kategorialen Faktoren ist möglich. z.B. Haben f0 (numerisch) und Dialekt (kategorial) einen Einfluss auf die Dauer?
- In R: die selbe Funktion/Syntax egal ob die abhängige Variable kontinuierlich (MM) oder kategorial (GLMM) ist.

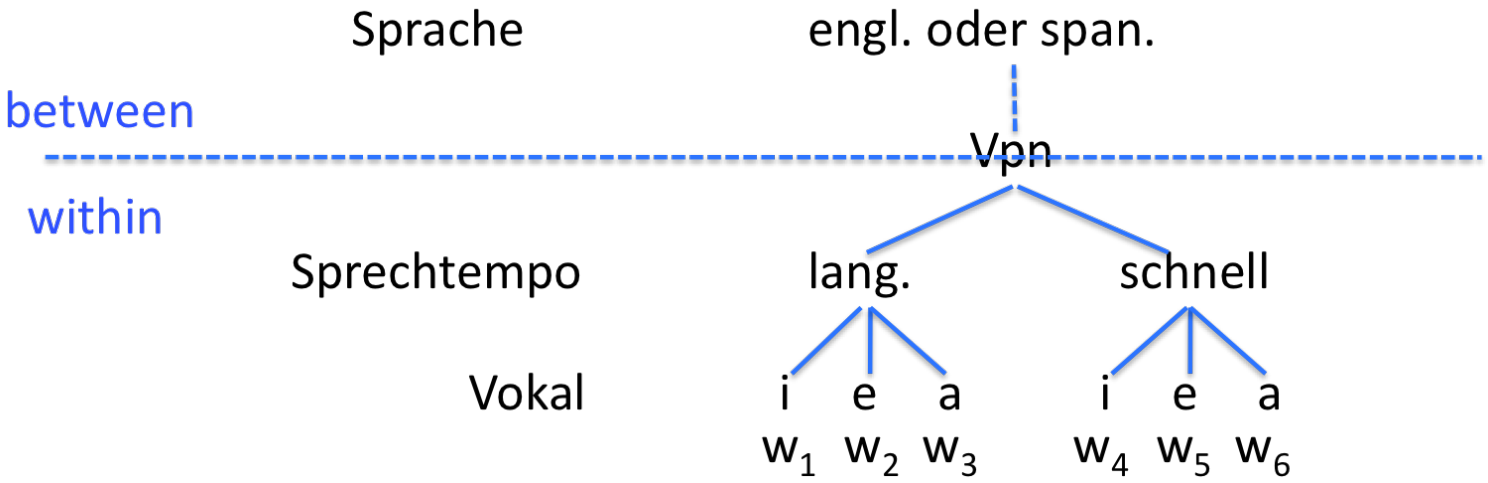
Faktoren in einem MM

Im MM-Verfahren wird prinzipiell zwischen 2 verschiedenen Sorten von Faktoren differenziert

- Fixed = Faktoren die vorhanden sind, unabhängig von dem experimentellen Design (z.B. Geschlecht, Sprache, Alter). Meistens die Faktoren, die geprüft werden sollen.
- Random: Faktoren, die randomisierte Stichproben aus einer Bevölkerung enthalten (z.B. Versuchspersonen, Wörter).

Vergleich: MM und ANOVA

Die Kieferposition wurde in 3 Vokalen /i, e, a/ und jeweils zu 2 Sprechtempi (langsam, schnell) gemessen. Die Messungen sind von 8 mit Muttersprache spanisch, 8 mit Muttersprache englisch aufgenommen worden.



Mixed model

Fixed: Sprache, Sprechtempo, Vokal — soll geprüft werden

Random: Sprecher — soll ausgeklammert werden

Mixed model (MM)

In einem MM wird ein Response (abhängige Variable) aus einer Kombinationen von gewichteten Faktoren eingeschätzt.

Lineares Modell, Minimierung vom Abstand zwischen tatsächlichen und eingeschätzten Werten – sehr ähnlich wie Regression*. Aber zusätzlich die Möglichkeit, Random-Faktors zu deklarieren.

*Das Verfahren um dies zu tun, ist aber nicht least-squares wie in Regression sondern maximum-likelihood.

Vergleich: MM und ANOVA

```
soa = read.table(file.path(pfadu, "soa.txt"))
```

F1 von /a:/ wurde in 3 verschiedenen Wörtern gemessen (*Bart*, *Pfad*, *Bart*). Die Wörter wurden von 3 Vpn produziert sowohl phrasenmedial als auch phrasenfinal (Faktor Pos). Inwiefern wird F1 von Phrasenposition beeinflusst? (N.B. F1 variiert sehr stark wegen Kontext, also von Wort zu Wort).

(a) Wir wollen die Sprechervariation ausklammern

(b) Wir wollen aber *auch* die Wortvariation ausklammern (die unterschiedlichen F1-Werte in *Bart* vs. *Pfad* usw. interessieren uns nicht).

(a) und (b) gleichzeitig ausklammern mit einer einzigen ANOVA geht nicht.

MM

Fixed: Phrasenposition (Pos) Random: Sprecher (Vpn), Wort (W)

Unterschiede in der Neigung, wenn sich die Stufen von Pos unterscheiden

Mixed model (MM)

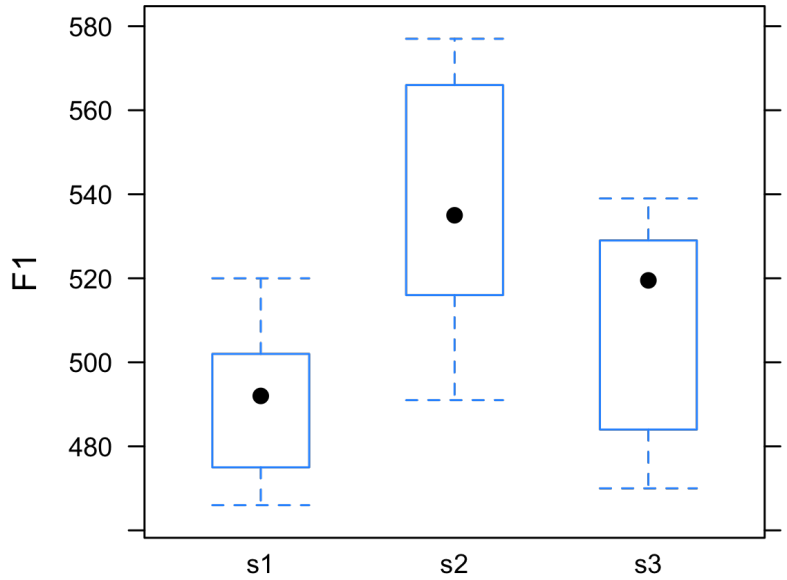
Das Intercept steht im Verhältnis zum Mittelwert aller Beobachtungen

x ist ein Faktor-Code* (0 oder 1 für die 2 Stufen von Pos)

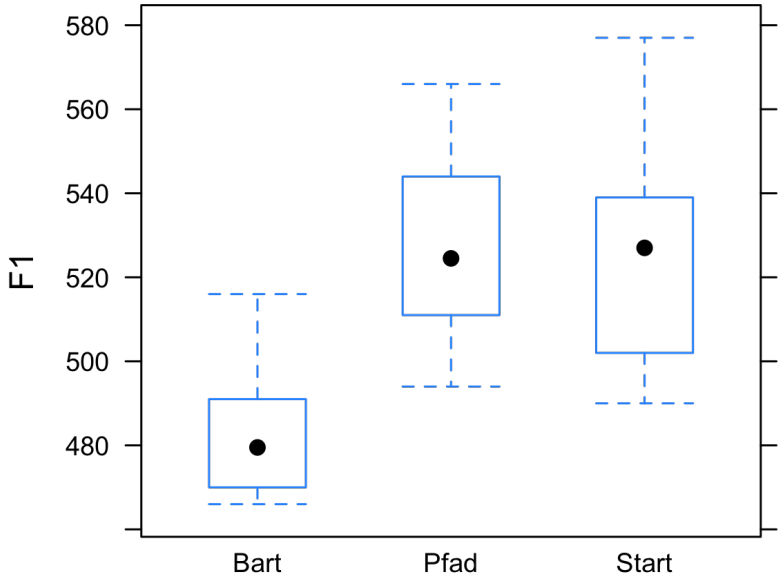
Eingeschätzte Werte

$$\hat{y} = bx + k + e_{Vpn} + e_{W}$$

By-subject intercept
Sprechervariation ausklammern



By-item intercept
Wortvariation ausklammern



*siehe: http://www.ats.ucla.edu/stat/r/library/contrast_coding.htm

Die eingeschätzen Werte

fitted(o)

MM in R

$$\hat{y} = bx + k + e_{Vpn} + e_{W}$$

o = lmer(F1 ~ Pos + (1 | Vpn) + (1 | W), data = soa)

Fixed-Koeffiziente

fixef(o)

```
(Intercept)    Posmedial
  522.11111    -18.88889
```

Random-Koeffiziente

ranef(o)

```
$Vpn
    (Intercept)
s1   -20.557646
s2    22.948070
s3    -2.390424

$W
      (Intercept)
Bart    -27.94979
Pfad     14.13553
Start    13.81427
```

z.B. für den 4en Wert

tatsächlich

```
  Vpn    W     Pos  F1
4  s1 Bart medial 475
```

eingeschätzt durch MM

-18.88889* 1 + 522.11111 -20.557646 -27.94979

454.7148

contrasts(soa$Pos)

```
       final      0
       medial     1
```

das gleiche fitted(o)[4]

```
454.7148
```

MM und die Prüfstatistik (ein Fixed Factor)

o = lmer(F1 ~ Pos + (1 | Vpn) + (1 | W), data = soa)

anova(o)

```
Analysis of Variance Table
    Df Sum Sq Mean Sq F value
Pos  1 1605.6  1605.6   11.69
```

Je höher der F-Wert umso wahrscheinlicher, dass Pos einen signifikanten Einfluss auf F1 hat. Um dies zu prüfen, das Modell noch einmal ohne den Fixed-Factor berechnen, und dann die beiden Modelle mit einem χ^2-Test vergleichen:

Modell ohne fixed-Faktoren*

ohne = lmer(F1 ~ 1 + (1 | Vpn) + (1 | W), data = soa)

oder: das ursprüngliche Modell ohne Pos:

ohne = update(o, ~ . -Pos)

* Siehe http://lme4.r-forge.r-project.org/slides/2011-03-16-Amsterdam/2Longitudinal.pdf, S. 29

Ein Fixed-Factor

o = lmer(F1 ~ Pos + (1 | Vpn) + (1 | W), data = soa)

ohne = update(o, ~ . -Pos)

anova(o, ohne)

```
Data: soa
Models:
ohne: F1 ~ (1 | Vpn) + (1 | W)
o: F1 ~ Pos + (1 | Vpn) + (1 | W)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
ohne   4 171.23 174.79 -81.614
o      5 164.35 168.81 -77.176 8.8758      1    0.00289 **
```

F1 wurde signifikant von der Phrasenposition beeinflusst

($\chi^2[1] = 8.9$, $p < 0.01$)

Mehr zu Random Factors

Man kann mit demselben Verfahren prüfen, ob beide Random-Factors wirklich benötigt werden (N.B.: im MM muss mindestens ein Random-Factor vorkommen).

```
o = lmer(F1 ~ Pos + (1 | Vpn) + (1 | W), data = soa)
```

benötigen wir Wort als RF?

```
ow = update(o, ~ . -(1|W))
anova(o, ow)
```

Ja.

```
Chisq Chi Df Pr(>Chisq)
16.105      1  5.992e-05 ***
```

benötigen wir Sprecher als RF?

```
ovpn = update(o, ~ . -(1|Vpn))
anova(o, ovpn)
```

Ja.

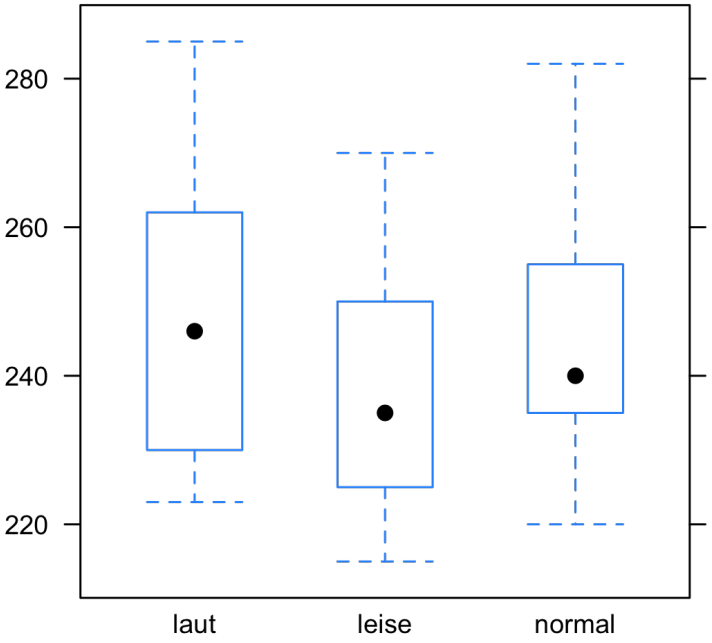
```
 Chisq Chi Df Pr(>Chisq)
14.387      1  0.0001488 ***
```

Ein Fixed-Factor mit mehreren Stufen

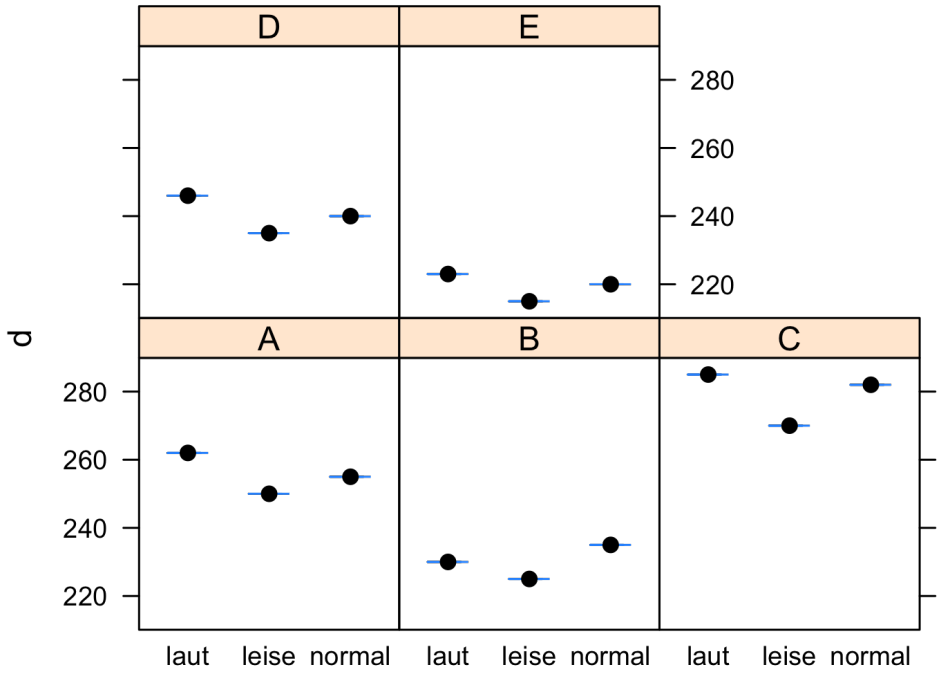
Die folgenden Daten zeigen die Wortdauer von fünf Sprechern wenn sie leise, normal, und laut sprechen. Hat die Lautstärke einen Einfluss auf die Dauer?

```
amp = read.table(file.path(pfadu, "amplitude.txt"))
head(amp)
```

```
bwplot(d ~ Amplitude, data = amp)
```




```
bwplot(d ~ Amplitude|Vpn , data = amp)
```



1. Mixed model

a = lmer(d ~ Amplitude + (1|Vpn), data = amp)

2. Mixed model ohne fixed factor

ohne = update(a, ~ . -Amplitude)

3. Vergleich mit und ohne

anova(a, ohne)

```
Data: amp
Models:
a2: d ~ 1 + (1 | Vpn)
a: d ~ Amplitude + (1 | Vpn)
   Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
a2  3 119.84 121.97 -56.921
a   5 107.04 110.58 -48.522 16.798      2   0.000225 ***
```

Die Dauer wurde signifikant von der Amplitude beeinflusst ($\chi^2[2] = 16.8$, $p < 0.001$)

Post-hoc Tests (wenn der Fixed-Factor mehr als 2 Stufen hat)

```
a = lmer(d ~ Amplitude + (1|Vpn), data = amp)
summary(glht(a, linfct = mcp(Amplitude = "Tukey")))
```

```
Linear Hypotheses:
                     Estimate Std. Error z value Pr(>|z|)
leise - laut == 0     -10.200      1.793  -5.690   <1e-04 ***
normal - laut == 0     -2.800      1.793  -1.562   0.2623
normal - leise == 0     7.400      1.793   4.128   0.0001 ***
```

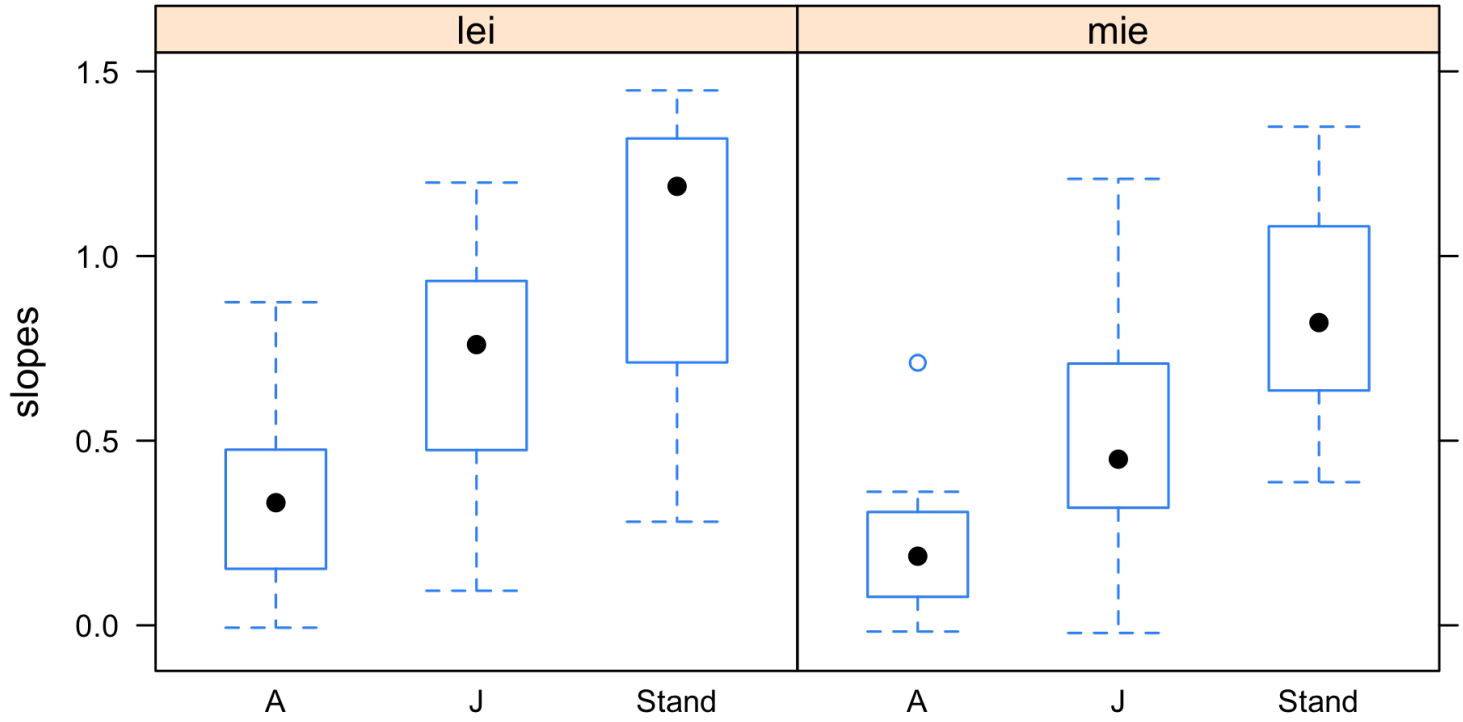
Post-hoc Tukey-Tests zeigten signifikante Dauer-Unterschiede zwischen leise und laut ($z = 5.7$, $p < 0.001$) und zwischen normal und leise ($z = 4.1$, $p < 0.001$), jedoch nicht zwischen normal und laut.

Zwei Fixed-Factors: keine Interaktion

```
param = read.table(file.path(pfadu, "param.txt"))
```

Die Daten zeigen Neigungen (slopes) für 3 Sprecher-Gruppen (Group) und zwei Kontinua (Cont). Inwiefern werden die Neigungen von der Gruppe und/oder Kontinuum beeinflusst?

```
bwplot(slopes ~ Group | Cont, data = param)
```



Zwei Fixed-Factors: keine Interaktion

```
o = lmer(slopes ~ Group * Cont + (1|Vpn), data = param)
```

```
anova(o)
```

```
Analysis of Variance Table
            Df  Sum Sq Mean Sq F value
Group        2 1.25371 0.62686 22.4981
Cont         1 0.61108 0.61108 21.9320
Group:Cont   2 0.00658 0.00329  0.1181
```

Wahrscheinlich keine Interaktion...

Interaktion sig?

```
o2 = lmer(slopes ~ Group + Cont + (1|Vpn), data = param)
```

oder äquivalent

```
o2 = update(o, ~ . -Group:Cont)
```

```
anova(o, o2)
```

```
    Df    AIC    BIC  logLik Chisq Chi Df Pr(>Chisq)
o2   6 16.719 32.105 -2.3593
o    8 20.445 40.959 -2.2223 0.274      2     0.8719
```

Keine signifikante Interaktion

Faktors prüfen

o2 = lmer(slopes ~ Group + Cont + (1|Vpn), data = param)

Faktor Group testen

o3 = update(o2, ~ . -Group)

anova(o2, o3)

```
o3: slopes ~ Cont + (1 | Vpn)
o2: slopes ~ Group + Cont + (1 | Vpn)
   Df    AIC    BIC   logLik  Chisq Chi Df Pr(>Chisq)
o3  4 45.967 56.225 -18.9836
o2  6 16.719 32.105  -2.3593 33.249      2  6.028e-08 ***
```

Faktor Cont testen

o4 = update(o2, ~ . -Cont)

anova(o2, o4)

```
o4: slopes ~ Group + (1 | Vpn)
o2: slopes ~ Group + Cont + (1 | Vpn)
   Df    AIC    BIC   logLik  Chisq Chi Df Pr(>Chisq)
o4  5 33.720 46.542 -11.8602
o2  6 16.719 32.105  -2.3593 19.002      1  1.306e-05 ***
```

Zwei Fixed-Factors (keine Interaktion)

Group

anova(o2, o3)

```
Chisq Chi Df Pr(>Chisq)
33.249      2  6.028e-08 ***
```

Cont

anova(o3, o4)

```
Chisq Chi Df Pr(>Chisq)
19.002      1  1.306e-05 ***
```

Slopes wurde signifikant von Group ($\chi^2[2] = 33.3$, $p < 0.001$) und von Cont ($\chi^2[1] = 19.0$, $p < 0.001$) beeinflusst, und es gab keine signifikante Interaktion zwischen diesen Faktoren.

anova(o, o2)

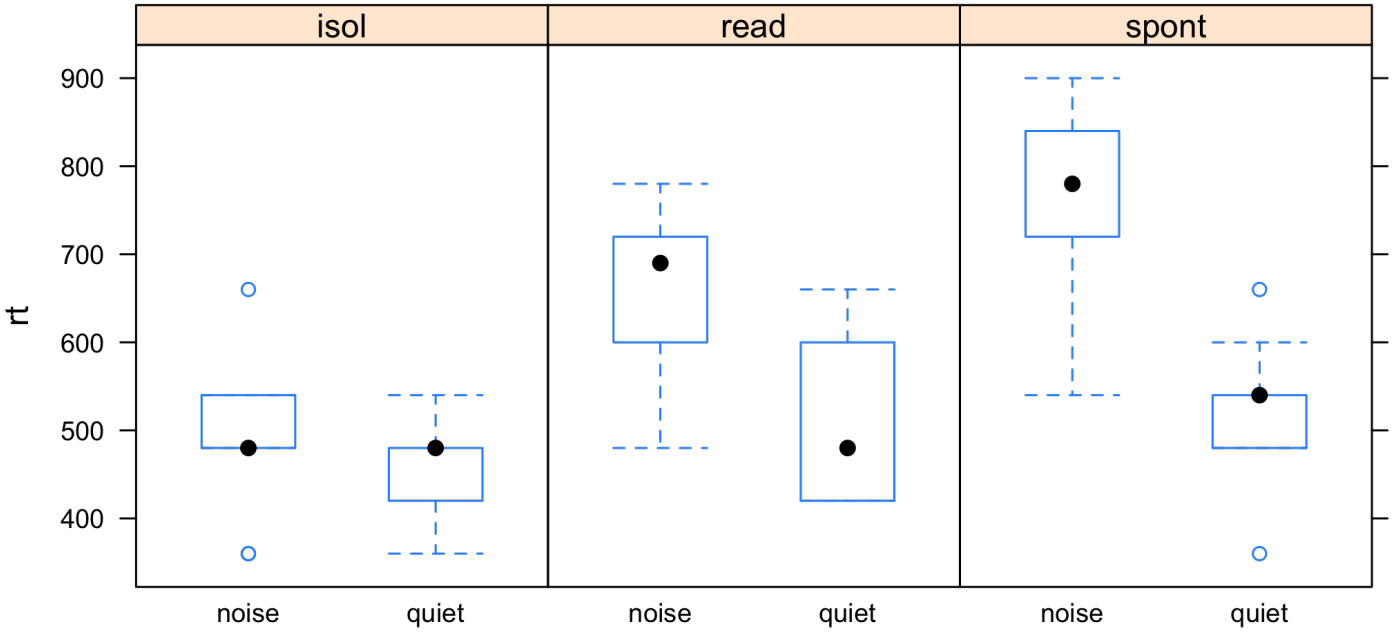
```
    Df    AIC    BIC  logLik Chisq Chi Df Pr(>Chisq)
o2   6 16.719 32.105 -2.3593
o    8 20.445 40.959 -2.2223 0.274      2     0.8719
```

Zwei Fixed-Factors mit Interaktion

```
noise = read.table(file.path(pfadu, "noise.txt"))
```

Reaktionszeiten wurden von Versuchsperson erhoben unter zwei Bedingungen: mit und ohne Lärm über Kopfhörer (Faktor Noise) und in isolierten Wörtern, in der gelesenen Sprache, und in der Spontansprache (Faktor Type). Inwiefern wurden die Reaktionszeiten durch Noise und Type beeinflusst?

```
bwplot(rt ~ Noise | Type, data = noise)
```



Zwei Fixed-Faktors mit Interaktion*

```
o = lmer(rt ~ Type * Noise + (1|Subj), data = noise)
```

```
anova(o)
```

```
Analysis of Variance Table
           Df Sum Sq Mean Sq F value
Type        2 289920  144960  40.492
Noise       1 285660  285660  79.793
Type:Noise  2 105120   52560  14.682
```

Wahrscheinlich liegt eine Interaktion vor. Dafür prüfen:

Modell ohne Interaktion

```
o2 = update(o, ~ . -Type:Noise)
```

Modell mit und ohne Interaktion vergleichen

```
anova(o, o2)
```

```
 Chisq Chi Df Pr(>Chisq)
25.123      2  3.505e-06 ***
```

Es gibt eine signifikante Interaktion zwischen Type und Noise ($\chi^2[2] = 25.1$, $p < 0.001$)

* Siehe http://lme4.r-forge.r-project.org/slides/2011-03-16-Amsterdam/2Longitudinal.pdf, S. 31

Zwei Fixed-Factors mit Interaktion

Wenn eine Interaktion vorliegt, dann die Faktoren miteinander kombinieren

```
beide = with(noise, interaction(Type, Noise))
```

MM damit berechnen

```
b = lmer(rt ~ beide + (1|Subj), data = noise)
```

Post-hoc Tukey-Tests anwenden

```
p = summary(glht(b, linfct = mcp(beide = "Tukey")))
```

Stufen-Kombinationen für Faktor 1 (Type variiert)

```
round(phsel(p), 3)
```

Stufen-Kombinationen für Faktor 2 (Noise variiert)

```
round(phsel(p, 2), 3)
```

Zwei Fixed-Factors mit Interaktion

round(phsel(p), 3)

	z value	Adjusted p values
read.noise - isol.noise	6.278	0.000
spont.noise - isol.noise	10.090	0.000
spont.noise - read.noise	3.812	0.002
read.quiet - isol.quiet	1.794	0.470
spont.quiet - isol.quiet	2.467	0.134
spont.quiet - read.quiet	0.673	0.985

Type variiert

round(phsel(p, 2), 3)

	z value	Adjusted p values
isol.quiet - isol.noise	-1.121	0.873
read.quiet - read.noise	-5.606	0.000
spont.quiet - spont.noise	-8.745	0.000

Noise variiert

Post-hoc Tukey Tests zeigten signifikante Unterschiede zwischen 'noise' und 'quiet' in gelesener ($p < 0.001$) und in spontaner ($p < 0.001$) Sprache jedoch nicht in isolierten Wörtern. Es gab Unterschiede zwischen allen drei Spechstilkombinationen aber nur in noise (read vs. isolated: $p < 0.001$; spont. vs isol: $p < 0.001$; spont. vs read: $p < 0.01$), aber nicht in quiet.

Konsistent mit ANOVA: siehe
http://www.phonetik.uni-muenchen.de/~jmh/lehre/sem/ss12/statistik/anova2ant.pdf
Aufgabe 4

Mehr als ein Random Faktor

Der Data-Frame `asp` enthält Werte der Aspirationsdauer von silbeninitialem /t/ und /k/ aus gelesenen Sätzen in dem Kielcorpus. Diese Dauern sind für 55 Versuchspersonen und 287 Wort-Typen (items) erhoben worden. (Die Versuchspersonen produzierten nicht alle dieselben items). Inwiefern wird die Aspirationsdauer von der Artikulationsstelle (/k/, /t/) oder von der Silbenbetonung ("betont", "unbetont") beeinflusst?

```
asp = read.table(file.path(pfadu, "asp.txt"))
```