Varianzanalyse mit Messwiederholungen. (fortgesetzt)

Jonathan Harrington

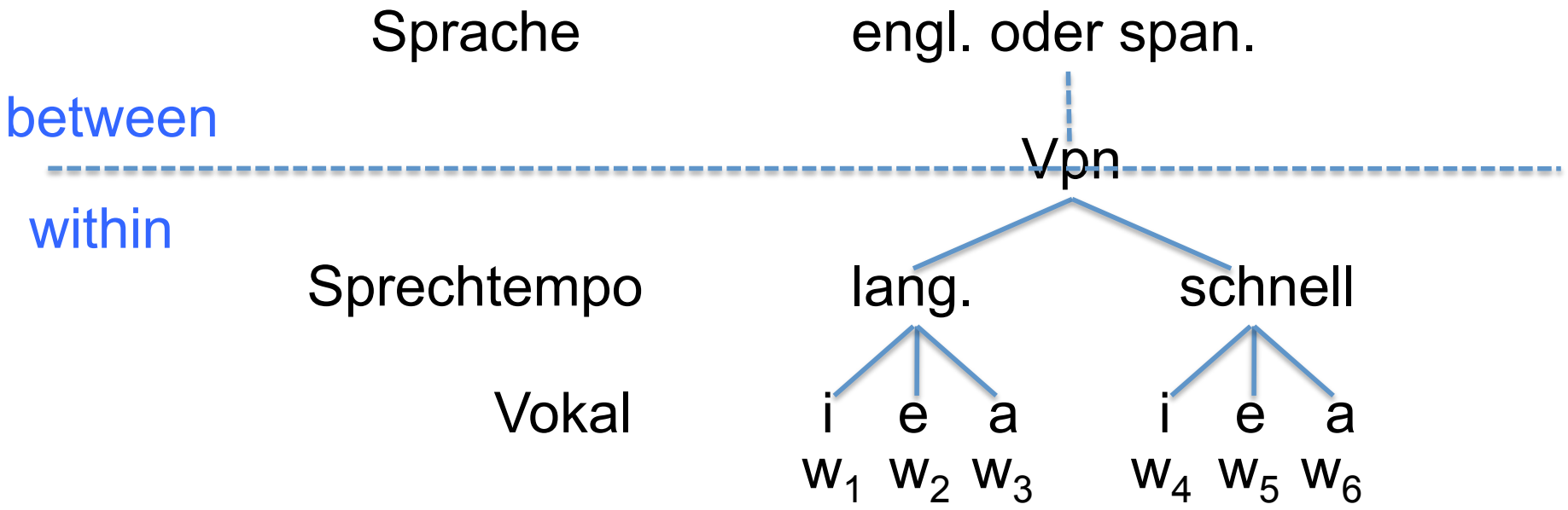
```
pfad = "Verzeichnis wo Sie anovaobjekte gespeichert haben"
attach(paste(pfad, "anovaobjekte", sep="/"))
```

1. Das Problem mit mehreren Werten pro Zelle
2. Post-hoc Tests

Befehle: anova3.txt

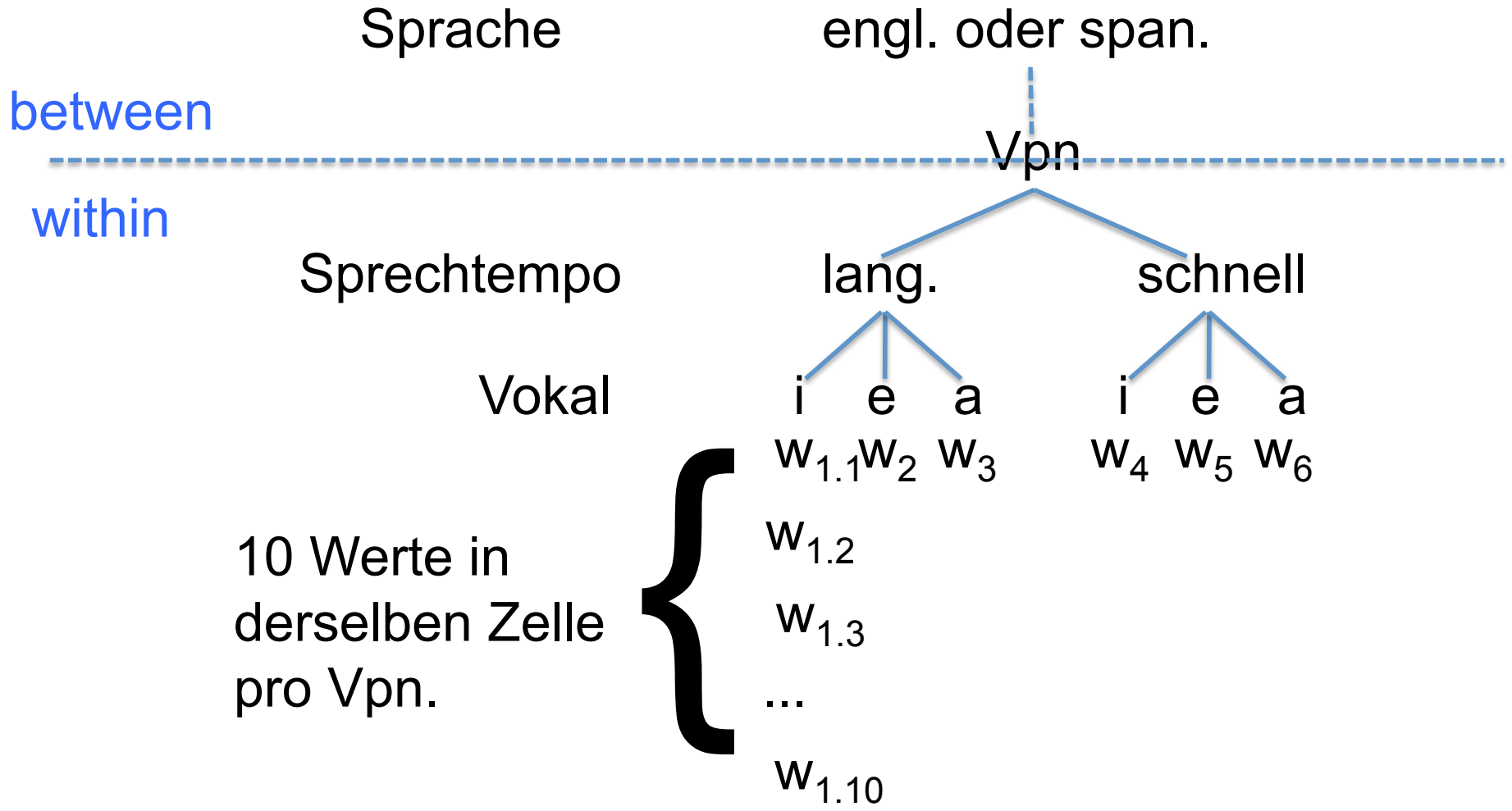
1. Wiederholungen in derselben Zelle

In allen bislang untersuchten ANOVAs gab es **einen Wert pro Vpn. pro Zelle**. z.B. 2 Faktoren mit 3 und 2 Stufen, dann 6 Werte pro Vpn, also einen Wert pro Stufen-Kombination pro Vpn.

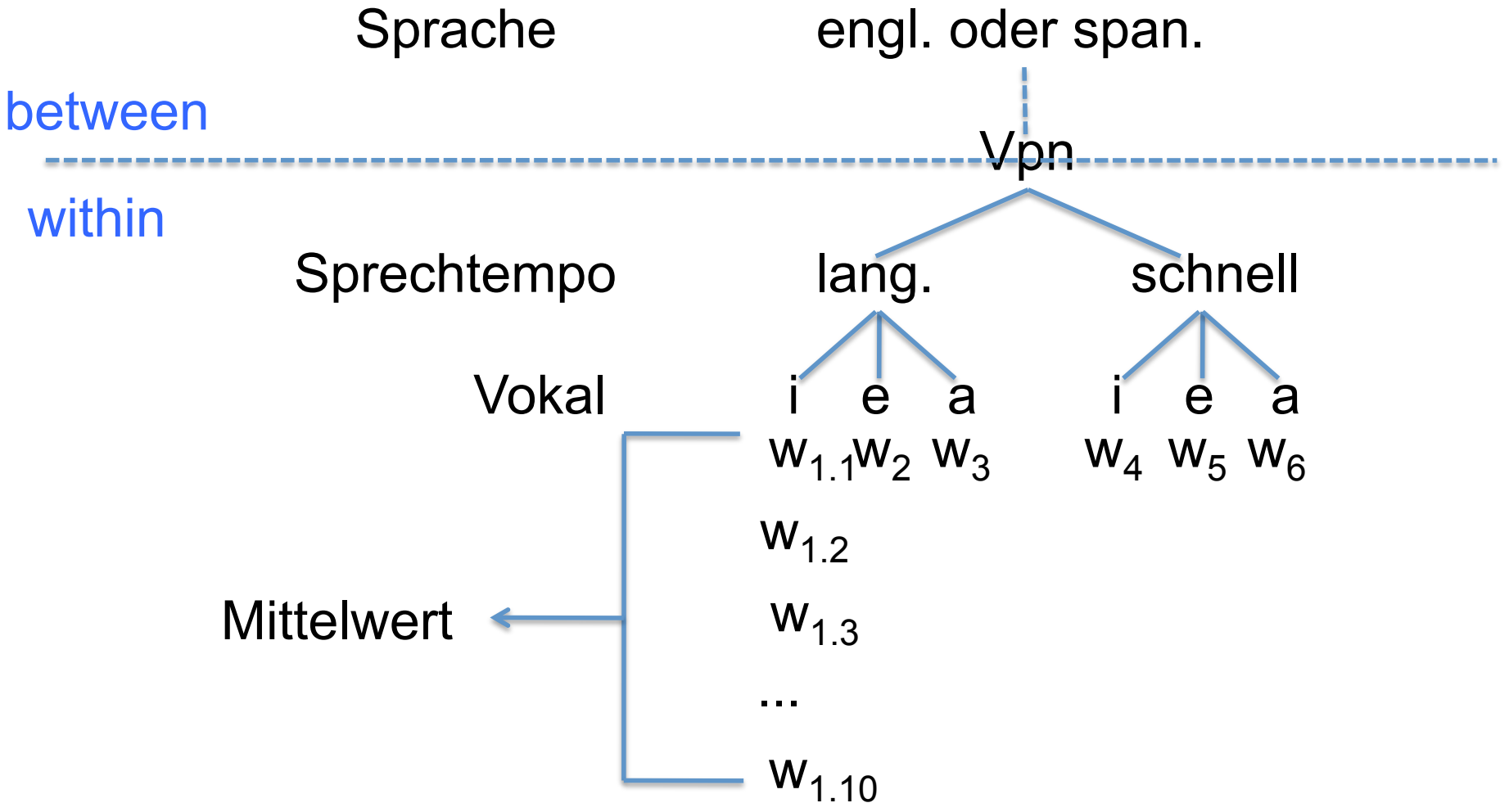


Wiederholungen in derselben Zelle

Jedoch haben die meisten phonetischen Untersuchungen **mehrere Werte pro Zelle**. zB. jede Vpn. erzeugte 'hid', 'head', 'had' zu einer langsamen und schnellen Sprechgeschwindigkeit **jeweils 10 Mal.**



Wiederholungen innerhalb der Zelle in einem ANOVA sind nicht zulässig und müssen gemittelt werden – damit wir pro Vpn. **einen within-subjects Wert pro Kombination der within-subjects Stufen** haben (6 Mittelwerte pro Vpn. in diesem Beispiel).



Wiederholungen in derselben Zelle

In einer Untersuchung zur /u/-Frontierung im Standardenglischen wurde von 12 Sprecherinnen (6 alt, 6 jung) F2 zum zeitlichen Mittelpunkt in drei verschiedenen /u/-Wörtern erhoben (*used*, *swoop*, *who'd*). Jedes Wort ist von jeder Vpn. 10 Mal erzeugt worden. Ist /u/ in den jungen Vpn. frontierter? (bis zu 60 Werte pro Vpn).

Faktor	within/between	wieviele Stufen?
Word	within	3
Alter	between	2

Wieviele Werte pro Vpn. dürfen in der ANOVA vorkommen? 3

Wieviele Werte insgesamt in der ANOVA wird es geben? 36

Wiederholungen in derselben Zelle: Beispiel

form.ssb Trackdatei, F1 und F2 englischer /u/ Vokale

age.ssb Alter: jung oder alt

word.ssb Wort: swoop, used, who'd

spk.ssb Sprecher: 12 Sprecherinnen (6 jung, 6 alt)

F2 zum zeitlichen Mittelpunkt

F2ssb = dcut(form.ssb[,2], .5, prop=T)

Wiederholungen in derselben Zelle: Beispiel

Anzahl der Wort-Wiederholungen pro Sprecher

table(word.ssb, spk.ssb)

word.ssb	arkn	elwi	frwa	gisa	jach	jeny	kapo	mapr	nata	rohi	rusy	shle
swoop	10	9	10	10	10	10	10	10	10	10	10	10
used	10	10	10	10	10	10	10	10	10	10	10	10
who'd	10	10	10	10	10	10	10	10	10	10	10	10

Die Funktion anova.mean() mittelt über die 10 Werte pro Vpn. pro Stufen-Kombinationen und bereitet alles für den RM-ANOVA vor.

alle Faktoren

F2m = anova.mean(F2ssb, spk.ssb, age.ssb, word.ssb)

abhängige Variable

Wiederholungen in derselben Zelle: Beispiel

F2m

```
    m          X1    X2     X3
1  10.527359 arkn   alt swoop
2  14.186585 arkn   alt  used
3  10.326474 arkn   alt who'd
4   8.662981 elwi   alt swoop
5  14.100450 elwi   alt  used
6   9.002776 elwi   alt who'd
7   7.495192 frwa   alt swoop
8  10.166607 frwa   alt  used
...
```


Mittelwert über die 10 Wiederholungen von *used*, Sprecherin elwi

F2m ist ein Data-Frame mit den erwünschten 36 Zeilen und mit 3 Werten pro Vpn.

Man kann/soll die Faktoren-Namen umbennen:

names(F2m) = c("F2", "Vpn", "Alter", "Wort")

Wiederholungen in derselben Zelle: Beispiel

```
names(F2m)
[1] "F2"    "Vpn"   "Alter" "Wort"
code =          c("d", "s", "b", "w")
ssb.t =         Anova.prepare(F2m, code)
Alter = factor(ssb.t$b)
ssb.lm =        lm(ssb.t$d ~ Alter)
ssb.ana =       Anova(ssb.lm, idata=ssb.t$w, idesign=~Wort)

ssb.ana
```

```
Type II Repeated Measures MANOVA Tests: Pillai test statistic
            Df test stat approx F num Df den Df     Pr(>F)
Alter        1     0.598   14.877      1     10  0.003175 **
Wort         1     0.912   46.652      2      9 1.777e-05 ***
Alter:Wort   1     0.548    5.449      2      9  0.028142 *
```

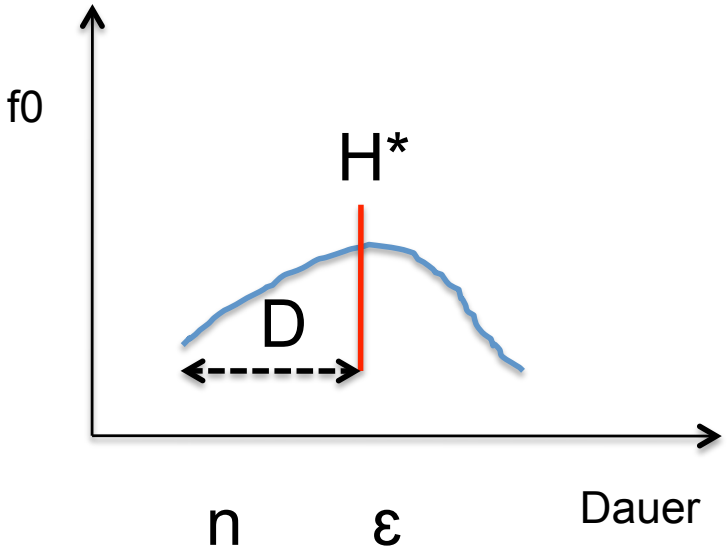
```
Type II Repeated Measures MANOVA Tests: Pillai test statistic
           Df test stat approx F num Df den Df    Pr(>F)
Alter       1     0.598   14.877      1     10  0.003175 **
Wort        1     0.912   46.652      2      9 1.777e-05 ***
Alter:Wort  1     0.548    5.449      2      9  0.028142 *
```

Alter hatte einen signifikanten Einfluss auf F2 ($F[1, 10] = 14.9$, $p < 0.01$) und es gab eine signifikante Interaktion zwischen Alter und Wort ($F[2, 9] = 5.5$, $p < 0.05$).

Wir brauchen den Wort-Effekt nicht zu berichten, weil das uns nicht interessiert – **war nicht Bestandteil der Fragestellung**: unterscheiden sich alt und jung in F2?.

RM-(M)anova und Interaktionen

Die Dauer, *D*, (ms) wurde gemessen zwischen dem Silbenonset und dem H* Tonakzent in äußerungsinitialen Silben (zB *nächstes*) und -finalen Silben (*demnächst*) jeweils von 10 Vpn., 5 aus Bayern (B) und 5 aus Schleswig-Holstein (SH).



Fragestellung(en): Inwiefern wird die Dauer vom Dialekt und/oder der Position beeinflusst?

Die Daten: dr names(dr)

RM-(M)anova und Interaktionen

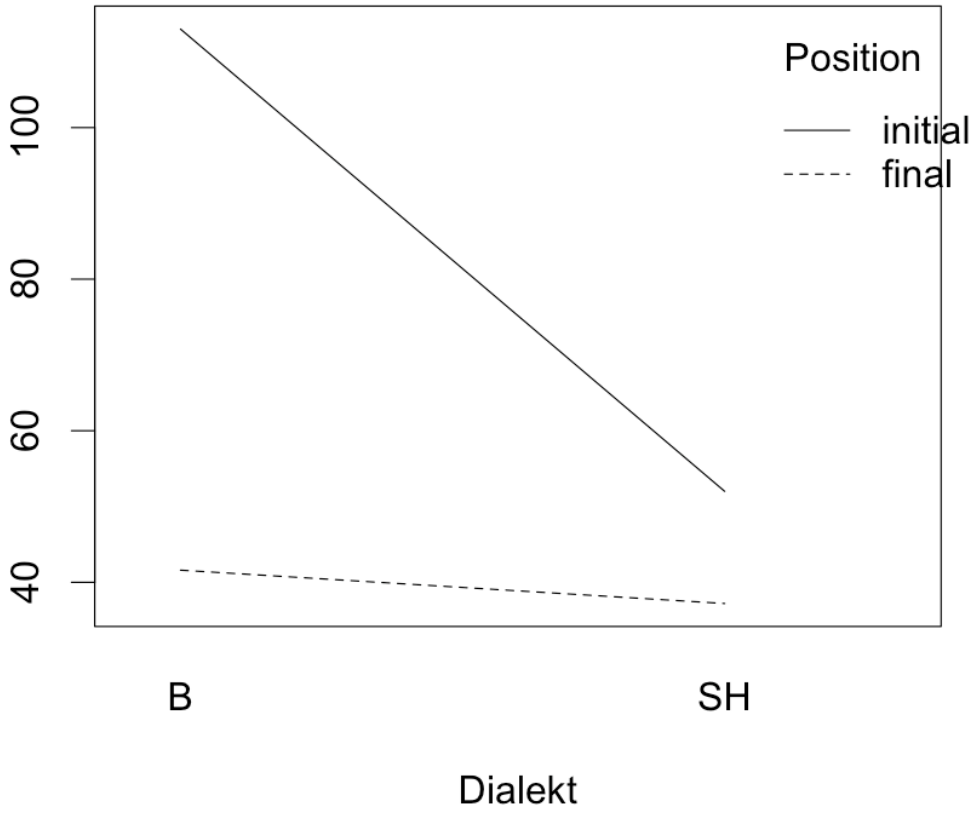
Fragestellung(en): Inwiefern wird die Dauer von der Position und/oder Dialekt beeinflusst?

```
dr.t = Anova.prepare(dr, c("d", "b", "s", "w"))
Dialekt = factor(dr.t$b)
dr.lm = lm(dr.t$d ~ Dialekt)
Anova(dr.lm, idata=dr.t$w, idesign=~Position)
```

```
Type II Repeated Measures MANOVA Tests: Pillai test statistic
              Df test stat approx F num Df den Df    Pr(>F)
Dial           1     0.581   11.081      1      8 0.0104034 *
Position       1     0.925   98.547      1      8 8.965e-06 ***
Dial:Position  1     0.842   42.488      1      8 0.0001845 ***
```

```
Type II Repeated Measures MANOVA Tests: Pillai test statistic
              Df test stat approx F num Df den Df    Pr(>F)
Dial           1     0.581   11.081      1      8 0.0104034 *
Position       1     0.925   98.547      1      8 8.965e-06 ***
Dial:Position  1     0.842   42.488      1      8 0.0001845 ***
```

with(dr, interaction.plot(Dialekt, Position, D))



Fragestellung(en): Inwiefern wird die Dauer von der Position und/oder Dialekt beeinflusst?

Interpretation

B und SH unterscheiden sich in initialer, nicht in finaler Position

Die Unterschiede zwischen initialer und finaler Position sind für B, nicht für SH signifikant

RM-(M)anovas, Interaktionen, und post-hoc Tests

Für einen RM-(M)anova kann **ein post-hoc t-test mit Bonferroni Korrektur** angewandt werden.

Je mehr Tests wir post-hoc anwenden, um so wahrscheinlicher ist es, dass wir Signifikanzen per Zufall bekommen werden. Der Tukey und Bonferroni-adjusted Tests sind Maßnahmen dagegen.

Bonferroni-Korrektur: Der Wahrscheinlichkeitswert der inviduellen Tests wird mit der **Anzahl der theoretisch möglichen Testkombinationen** multipliziert.

Post-hoc t-test mit Bonferroni Korrektur

1. t-tests aller Stufen-Kombinationen durchführen: als **g**epaart mit denselben Between-Stufen, sonst ungepaart.

SH-initial mit SH-final	**g**	SH-final mit B-initial
SH-initial mit B-initial		SH-final mit B-final
SH-initial mit B-final		B-initial mit B-final **g**

2. Bonferroni Korrektur: den Wahrscheinlichkeitswert eines t-tests mit der Anzahl der Tests multiplizieren

zB wenn SH-initial vs SH-final p = 0.035, Bonferroni-Korrektur: 0.035 * 6 = 0.21 (weil es 6 mögliche Testpaare gibt).

3. Auswahl: nur die Test-Kombinationen, die sich in **einer** Stufe unterscheiden

(Zur Info): wieviele Tests?

Fur *n* Stufen gibt es n!/(n-2)!2! mögliche Kombinationen.

zB
Dialekt * Position * Geschlecht war signifikant.
Dialekt = Hessen, Bayern, S-H
Geschlecht = M, W
Position = initial, medial, final

Wir haben 3 x 2 x 3 = 18 Stufen-Kombinationen

Das gibt 18!/16!2! = 18 x 17/2 = 153 t-Tests.

Bonferroni Korrektur: Die Wahrscheinlichkeiten mit 153 multiplizieren.

Post-hoc t-test mit Bonferroni Korrektur

posthoc(Fak1, Fak2, Fak3, ...Fakn, code = factorcode, d = depvariable)

Fak1, Fak2, Fak3...Fakn sind die gekreuzten Faktoren, die posthoc geprüft werden sollen

```
dr.t = Anova.prepare(dr, c("d", "b", "s", "w"))
```

Die gekreuzten Faktoren, die posthoc geprüft werden sollen

```
Dialekt = dr.t$b
Position = dr.t$w
```

```
dr.p = posthoc(Dialekt, Position, code = c("b", "w"), d = dr.t$d)
```

NB. immer **code =** etwas und **d =** etwas

dr.p$stats

listet nur die Test-Kombinationen, die sich in **einer** Stufe unterscheiden (daher *nicht* SH.final-B.initial usw.)

	stat	df	Bonferroni p
SH.final-B.final	-0.4666613	7.999611	1.000000000
SH.final-SH.initial	-2.5709017	4.000000	0.371518380
B.final-B.initial	-10.9833157	4.000000	0.002342832
SH.initial-B.initial	-5.1226150	6.475584	0.010372660

dr.p$bonf

Bonferroni-Multiplikator

6

dr.p$paired

Wurde ein gepaarter t-Test durchgeführt?

FALSE TRUE TRUE FALSE

Fragestellung(en): Inwiefern wird die Dauer von der Position und/oder Dialekt beeinflusst?

	stat	df	Bonferroni p
SH.final-B.final	-0.4666613	7.999611	1.000000000
SH.final-SH.initial	-2.5709017	4.000000	0.371518380
B.final-B.initial	-10.9833157	4.000000	0.002342832
SH.initial-B.initial	-5.1226150	6.475584	0.010372660

Post-hoc t-Tests mit Bonferroni-Korrektur zeigten signifikante Unterschiede zwischen Bayern und Schleswig-Holstein in initialer ($p < 0.05$) jedoch nicht in finaler Position. Die Unterschiede zwischen initialer und finaler Position waren nur für Bayern ($p < 0.01$) jedoch nicht für Schleswig-Holstein signifikant.

In einer Untersuchung zur /u/-Frontierung im Standardenglischen wurde von 12 Sprecherinnen (6 alt, 6 jung) F2 zum zeitlichen Mittelpunkt in drei verschiedenen /u/-Wörtern erhoben (*used, swoop, who'd*). Jedes Wort ist von jeder Vpn. 10 Mal erzeugt worden.

Die Fragestellung: Ist F2 höher (/u/ frontierter) für die junge im Vergleich zur alten Gruppe?

Die Fragestellung: Ist F2 höher (/u/ frontierter) für die junge im Vergleich zur alten Gruppe?

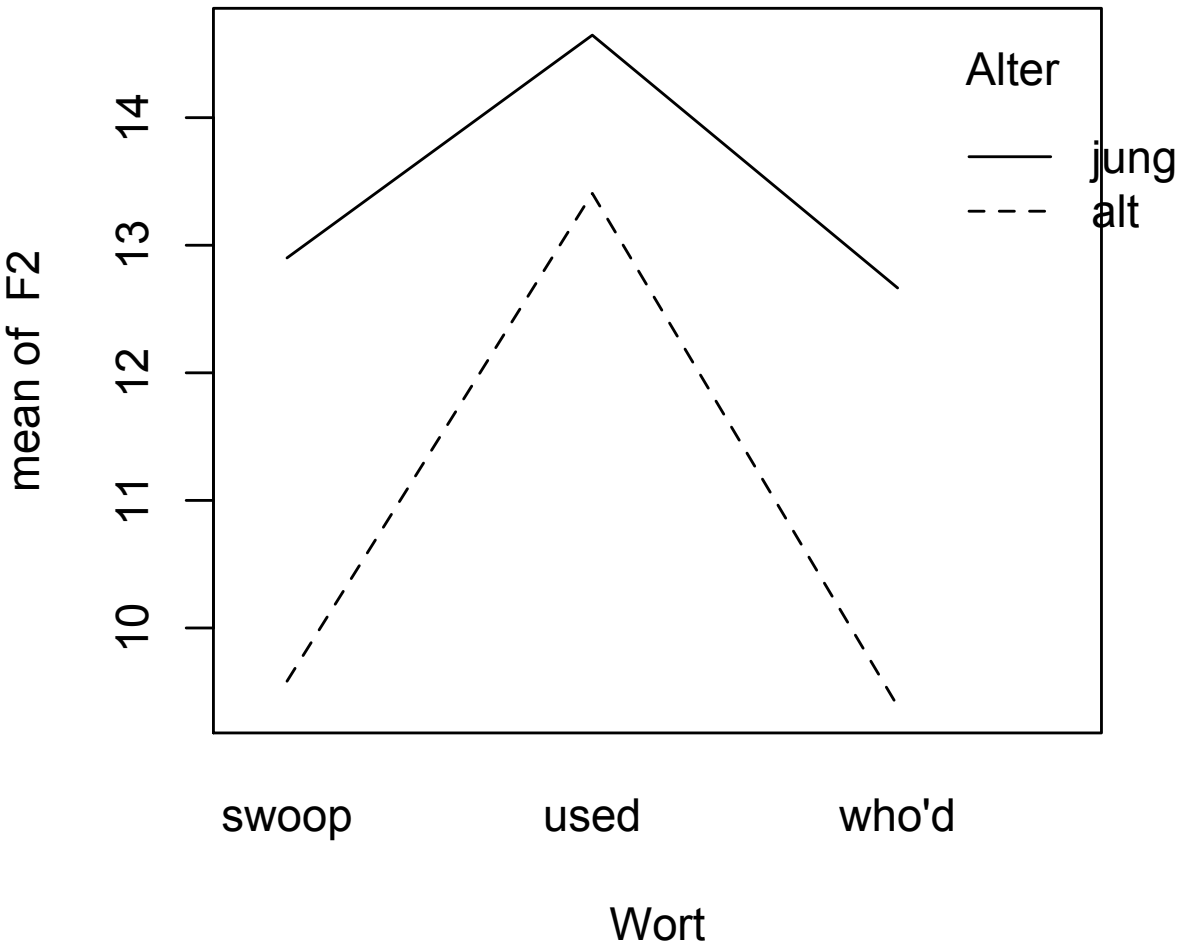
```
code = c("d", "s", "b", "w")
ssb.t = Anova.prepare(F2m, code)
Alter = factor(ssb.t$b)
ssb.lm = lm(ssb.t$d ~ Alter)
ssb.ana = Anova(ssb.lm, idata=ssb.t$w, idesign= ~Wort)
ssb.ana
```

```
Type II Repeated Measures MANOVA Tests: Pillai test statistic
           Df test stat approx F num Df den Df    Pr(>F)
Alter       1     0.598   14.877      1     10  0.003175 **
Wort        1     0.912   46.652      2      9 1.777e-05 ***
Alter:Wort  1     0.548    5.449      2      9  0.028142 *
```

Alter hatte einen signifikanten Einfluss auf F2 ($F[1, 10] = 14.9$, $p < 0.01$) und es gab eine signifikante Interaktion zwischen Alter und Wort ($F[2, 9] = 5.5$, $p < 0.05$).

Die Fragestellung: Ist F2 höher (/u/ frontierter) für die junge im Vergleich zur alten Gruppe?

with(F2m, interaction.plot(Wort, Alter, F2))



Die Fragestellung: Ist F2 höher (/u/ frontierter) für die junge im Vergleich zur alten Gruppe?

```
Wort = ssb.t$w
Alter = ssb.t$b

ssb.p = posthoc(Wort, Alter, code=c("w", "b"), d=ssb.t$d)

ssb.p$stats
```

	stat	df	Bonferroni p
swoop.alt-used.alt	-7.382146	5.000000	0.01075660
swoop.alt-who'd.alt	0.956723	5.000000	1.00000000
swoop.alt-swoop.jung	-4.275313	9.555319	0.02700452
used.alt-who'd.alt	7.973837	5.000000	0.00750801
used.alt-used.jung	-1.785802	5.428486	1.00000000
who'd.alt-who'd.jung	-4.316846	7.924107	0.03921836
swoop.jung-used.jung	-4.604262	5.000000	0.08726669
swoop.jung-who'd.jung	1.010658	5.000000	1.00000000
used.jung-who'd.jung	6.458623	5.000000	0.01986783

Die Fragestellung: Ist F2 höher (/u/ frontierter) für die junge im Vergleich zur alten Gruppe?

Alter hatte einen signifikanten Einfluss auf F2 ($F(1, 10)=14.9$, $p < 0.01$) und es gab eine signifikante Interaktion zwischen Alter und Wort ($F(2, 9) = 5.5$, $p < 0.05$).

	stat	df	Bonferroni p
swoop.alt-used.alt	-7.382146	5.000000	0.01075660
swoop.alt-who'd.alt	0.956723	5.000000	1.00000000
swoop.alt-swoop.jung	-4.275313	9.555319	0.02700452
used.alt-who'd.alt	7.973837	5.000000	0.00750801
used.alt-used.jung	-1.785802	5.428486	1.00000000
who'd.alt-who'd.jung	-4.316846	7.924107	0.03921836
swoop.jung-used.jung	-4.604262	5.000000	0.08726669
swoop.jung-who'd.jung	1.010658	5.000000	1.00000000
used.jung-who'd.jung	6.458623	5.000000	0.01986783

Post-hoc t-Tests mit Bonferroni-Korrektur zeigten

signifikante altersbedingte Unterschiede in *who'd* ($p < 0.05$) und *swoop* ($p < 0.05$), jedoch nicht in *used*.