Varianzanalyse mit Messwiederholungen (fortgesetzt)

Jonathan Harrington

Befehle: anova3.txt

path = "Verzeichnis wo Sie anova1 gespeichert haben"
attach(paste(path, "anova1", sep="/"))

1. Greenhouse-Geisser Korrektur
2. Mehr zu post-hoc Tests
3. Das Problem mit mehreren Werten pro Zelle

Greenhouse-Geisser Korrektur

In einem RM-ANOVA wird angenommen, dass die Ebenen von jedem within-subject Faktor ähnliche Varianzen haben.

Wenn die Varianzen der Ebenen unähnlich sind, kann trotzdem ein RM-ANOVA durchgeführt werden, aber dann **müssen die Freiheitsgrade mit einem Korrekturfaktor multipliziert werden**.

Der Korrekturfaktor ist der **Greenhouse-Geisser Epsilon (ε)**

ε variiert zwischen 1 (keine Korrektur notwendig) und 1/(k-1), wo *k* die Anzahl der Ebenen ist.

Der Korrekturfaktor muss **nur bei einem Within-Subjects-Faktor mit 3 oder mehr Ebenen** angewandt werden.

Greenhouse-Geisser Korrektur

In Sussman et al (1997) sind die Neigungen von Lokusgleichungen in initialen (CV), medialen (VCV), und finalen (VC) Silben für C = /b, d, g/ und jeweils in 10 Vpn. (5 männlich, 5 weiblich) berechnet worden.

Inwiefern werden die Neigungen vom Konsonant, Silbenposition und Geschlecht beeinflusst?

Faktor	between/within?	Wieviele Ebenen?	ε Korrektur?
Konsonant	within	3	ja
Silbenposition	within	3	ja
Geschlecht	between	2	nein

ANOVA mit Messwiederholungen: between und within

Die Daten: lok names(lok) attach(lok)

Faktor	Ebenen	between/within
Kons	b, d, g	within
P	initial,medial,final	within
G	m, w	between
Spr	Die Sprecher...	-

RM-ANOVA ausführen

```
lok.aov    = aov(slopes ~ G * Kons * P + Error(Spr/(Kons * P)))

summary(lok.aov)
```

```
Error: Spr
          Df   Sum Sq  Mean Sq F value Pr(>F)
G          1 0.005760 0.005760  0.3105 0.5926
Residuals  8 0.148422 0.018553
```

between

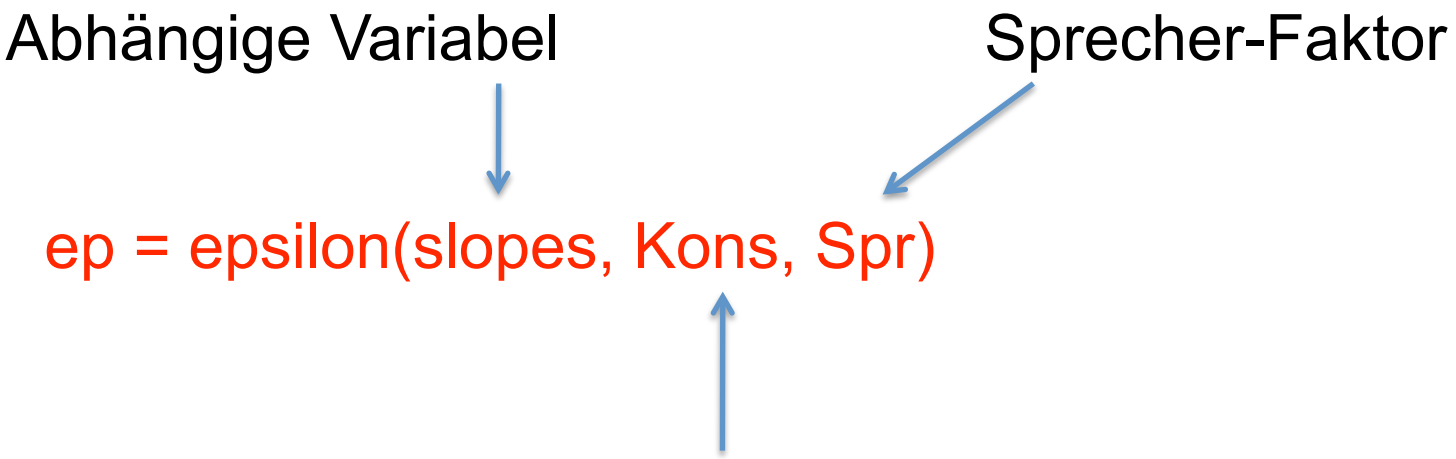
```
Error: Spr:Kons
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons       2 2.32113 1.16056 70.8838 1.119e-08 ***
G:Kons     2 0.03971 0.01985  1.2126    0.3233
Residuals 16 0.26196 0.01637
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Spr:P
          Df   Sum Sq  Mean Sq F value  Pr(>F)
P          2 0.053816 0.026908  3.4211 0.05795 .
G:P        2 0.030807 0.015403  1.9584 0.17346
Residuals 16 0.125844 0.007865
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Spr:Kons:P
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons:P     4 0.62695 0.15674 19.3407 3.557e-08 ***
G:Kons:P   4 0.09279 0.02320  2.8624   0.03908 *
Residuals 32 0.25933 0.00810
```

within

Greenhouse-Geisser Korrektur, Faktor Kons



ep

```
Greenhouse-Geisser epsilon     Huynh-Feldt epsilon
                 0.9309326                1.1640752
```

Anwendung von Greenhouse-Geisser Korrektur

```
Error: Spr:Kons
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons       2 2.32113 1.16056 70.8838 1.119e-08 ***
G:Kons     2 0.03971 0.01985  1.2126    0.3233
Residuals 16 0.26196 0.01637
```

ersetzt

Die Freiheitsgrade und daher der p-Wert wird mit ε gewichtet:

```
2 * (1 - pf(70.8838, 2*ep [1], 16*ep [1]))
[1] 6.617335e-08
```

Der Einfluss der Artikulationsstelle auf die Neigungen war signifikant (F(1.86, 14.89) = 70.88, p < 0.001, Greenhouse-Geisser korrigiert).

Greenhouse-Geisser Korrigierung für Faktor 'Position'

```
Error: Spr:P
          Df   Sum Sq  Mean Sq F value  Pr(>F)
P          2 0.053816 0.026908  3.4211 0.05795 .
G:P        2 0.030807 0.015403  1.9584 0.17346
Residuals 16 0.125844 0.007865
```

pos.ep = epsilon(slopes, P, Spr)

```
2 * (1 - pf(3.4211, 2*pos.ep [1], 16*pos.ep [1]))
```

[1] 0.1378011

Position hatte keinen signifikanten Einfluss
auf die Neigungen.

post-hoc tests

Post-hoc tests müssen durchgeführt werden, wenn Interaktionen vorliegen, weil das heißt: man kann **einen signifikanten Haupteffekt nicht uneingeschränkt annehmen.**

Konkret für diese Ergebnisse heißt das: Konsonant hat deutlich einen Einfluss auf die Neigungen, aber vielleicht nicht für alle 3 Positionen und beide Geschlechter

```
Error: Spr
          Df   Sum Sq  Mean Sq F value Pr(>F)
G          1 0.005760 0.005760  0.3105 0.5926
```

between

---

within

```
Error: Spr:Kons
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons       2 2.32113 1.16056 70.8838 1.119e-08 ***
G:Kons     2 0.03971 0.01985  1.2126    0.3233

Error: Spr:P
          Df  Sum Sq Mean Sq F value  Pr(>F)
P          2 0.053816 0.026908  3.4211 0.05795 .
G:P        2 0.030807 0.015403  1.9584 0.17346
Residuals 16 0.125844 0.007865

Error: Spr:Kons:P
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons:P     4 0.62695 0.15674 19.3407 3.557e-08 ***
G:Kons:P   4 0.09279 0.02320  2.8624   0.03908 *
```

post-hoc tests

In diesem Fall müssen wir post-hoc Tests auf **Kombinationen von allen 3 Faktoren** anwenden, da es **zwischen den 3 Faktoren eine signifikante Interaktion gegeben hat**.

```
Error: Spr:Kons:P
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons:P     4 0.62695 0.15674 19.3407 3.557e-08 ***
G:Kons:P   4 0.09279 0.02320  2.8624   0.03908 *
```

Wenn nur die Interaktion Kons:P signifikant gewesen wäre, dann reicht es, nur Kons:P abzuprüfen, also:

b.initial vs. d.initial, b.initial vs. g.initial, d.initial vs. g.initial, b.medial vs. d.medial usw.

und wir hätten annehmen dürfen, dass alle signifikanten Ergebnisse für beide Geschlechter gültig sind...

post-hoc tests

Wenn wir (wie in diesem Fall) Kons für Interaktionen prüfen wollen, dann soll dies geschehen bei **gleichbleibenden Ebenen der anderen Faktoren.**

zB	nicht
b-initial-M vs. d-initial-M	b-initial-M vs. g-medial-W
d-final-W vs. g-final-W	d-final-W vs. g-initial-W
usw.	

weil nur die Tests links (aber nicht rechts) die Antwort auf die Frage vermittelen: ist Kons signifikant in allen 3 Positionen und für beide Geschlechter?

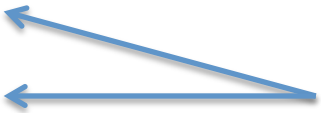
post-Hoc Tukey test für alle 3 Faktoren

```
Error: Spr:Kons:P
          Df  Sum Sq Mean Sq F value    Pr(>F)
Kons:P     4 0.62695 0.15674 19.3407 3.557e-08 ***
G:Kons:P   4 0.09279 0.02320  2.8624   0.03908 *
Residuals 32 0.25933 0.00810
```

tk = Tukey.rm(slopes, 0.00810, 32, G, Kons, P)

tk hat 153 Einträge (!), von denen wir nur 18 brauchen (auf der nächsten Seite). In anova3.txt sind die Befehle, um diese 18 aus dem Vektor tk zu suchen.

M.b.initial-M.d.initial	0.00000621
M.b.initial-M.g.initial	0.00000091
F.b.initial-F.d.initial	0.00001645
F.b.initial-F.g.initial	0.00000000
M.d.initial-M.g.initial	0.99999887
F.d.initial-F.g.initial	0.21804338
M.b.medial-M.d.medial	0.00001009
M.b.medial-M.g.medial	0.00000002
F.b.medial-F.d.medial	0.00009657
F.b.medial-F.g.medial	0.00000010
M.d.medial-M.g.medial	0.66535745
F.d.medial-F.g.medial	0.52660795
M.b.final-M.d.final	0.00092722
M.b.final-M.g.final	0.09245728
F.b.final-F.d.final	0.99969205
F.b.final-F.g.final	0.99992148
M.d.final-M.g.final	0.92870692
F.d.final-F.g.final	0.84702874



Nur für dieses Paar gibt es unterschiedliche Ergebnisse bei M vs. F

/b/ vs. /d/: in allen Positionen sig. außer finalem /b/ vs. finalem /d/ für Frauen.

/b/ vs. /g/: in initialer und medialer Position sig.

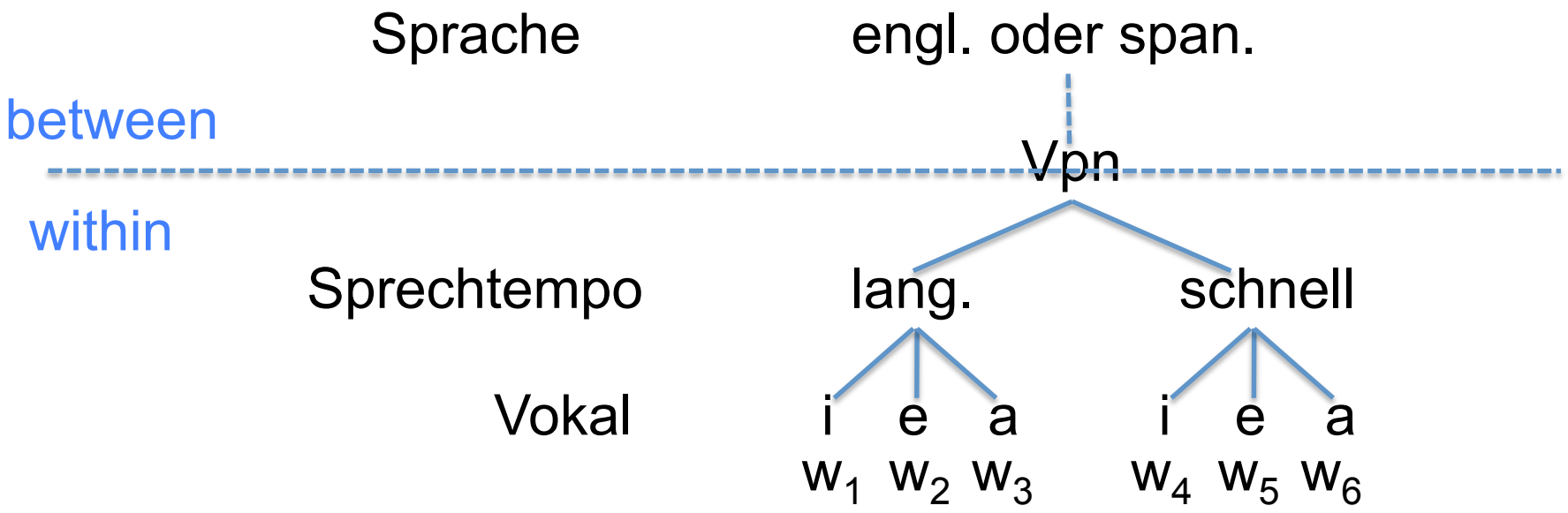
/d/ vs. /g/:nicht signifikant

Zusammenfassung der Ergebnisse

Der Einfluss der Artikulationsstelle auf die Neigungen war signifikant ($F(1.86, 14.89) = 70.88$, $p < 0.001$, Greenhouse-Geisser korrigiert). Der Einfluss der Position war nicht signifikant. Die Interaktionen Konsonant x Position ($F(4, 32) = 19.34$, $p < 0.001$) sowie Konsonant x Position x Geschlecht ($F(4, 32) = 2.86$, $p < 0.05$) waren beide signifikant. Post-hoc Tukey tests zeigten signifikante ($p < 0.001$) Unterschiede zwischen /b/ und /d/ in allen Positionen außer für finalen /b/ vs. finalen /d/ für Frauen. Die Unterschiede zwischen /b/ und /g/ waren in initialer und medialer Position signifikant ($p < 0.001$). Der /d-g/ Unterschied war nicht signifikant.

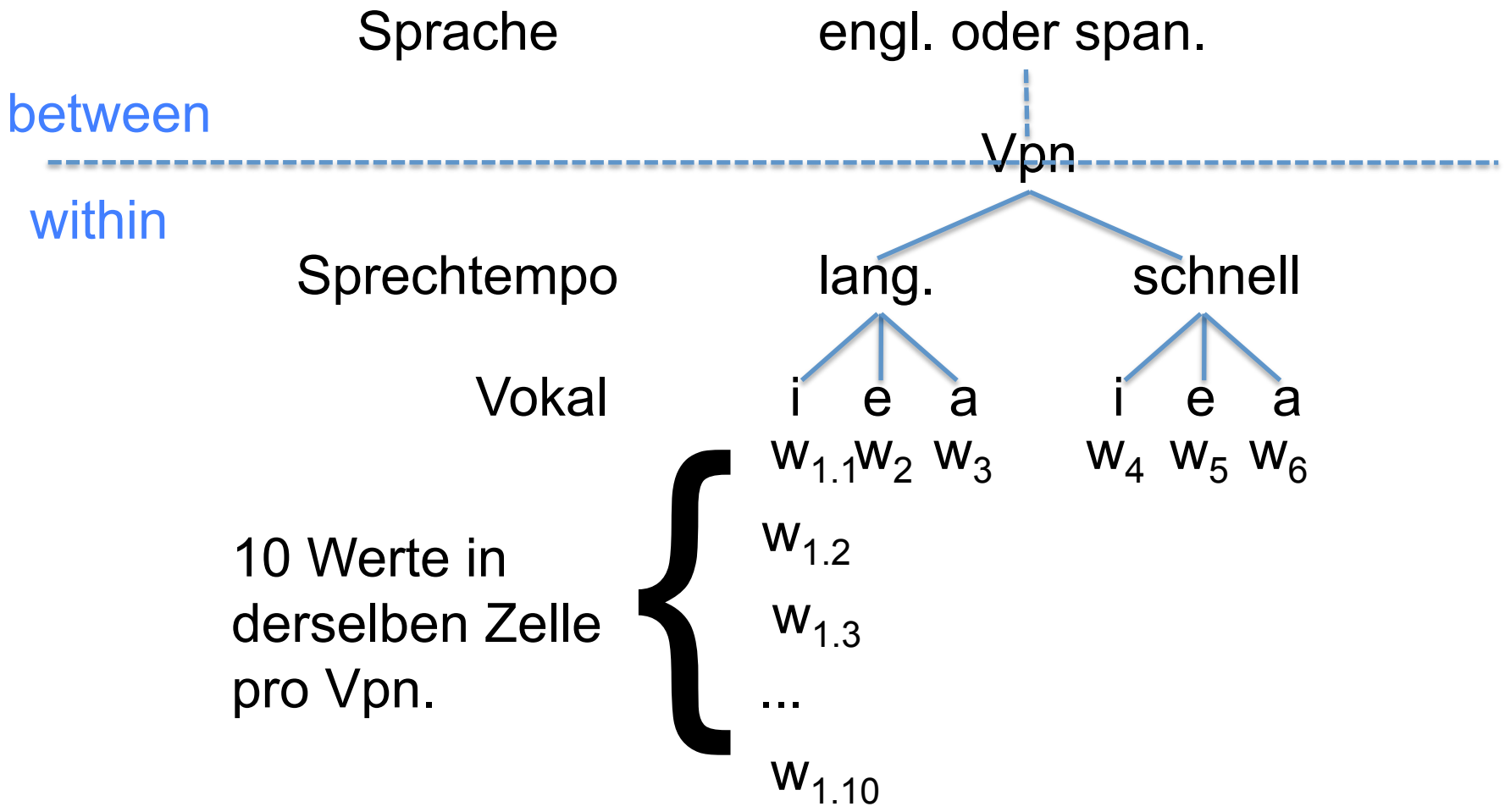
3. Wiederholungen in derselben Zelle

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

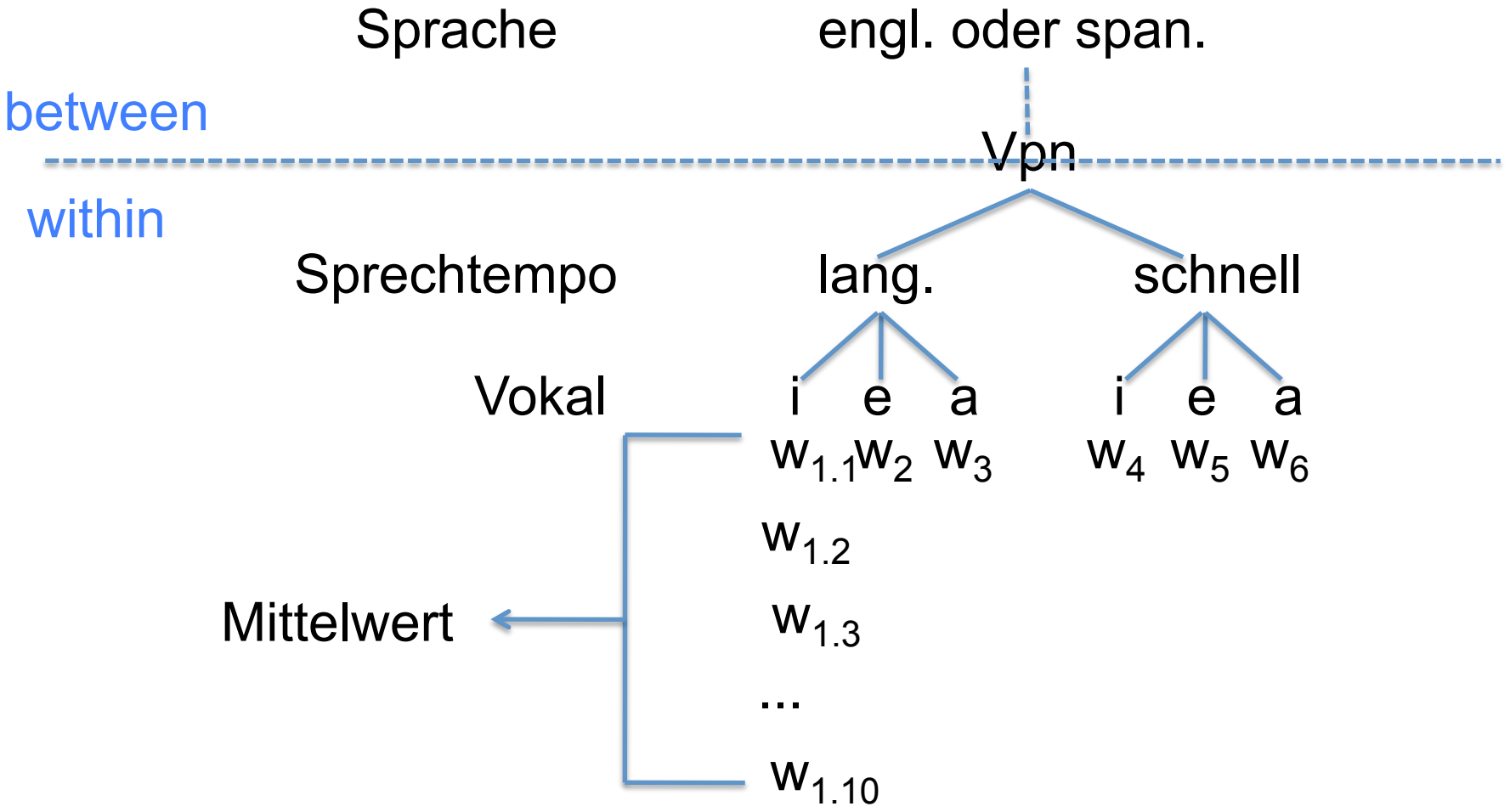


3. Wiederholungen in derselben Zelle

In fast allen phonetischen Untersuchungen gibt es jedoch **mehrere Werte pro Zelle**. zB. jede Vpn. erzeugte 'hid', 'head', 'had' zu einer langsamen und schnellen Sprechgeschwindigkeit **jeweils 10 Mal.**



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



3. 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 Ebenen?
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.prepare() mittelt über die 10 Werte pro Vpn. pro Ebenen-Kombinationen und bereitet alles für den RM-ANOVA vor.

alle Faktoren

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

abhängige Variabel

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 die Faktoren-Namen umbennen:

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

Wiederholungen in derselben Zelle: Beispiel

```
attach(F2m)
names(F2m)
[1] "F2"    "Vpn"   "Alter" "Wort"
```

ANOVA-Befehl?

```
ssb.aov =     aov(F2 ~ Alter * Wort + Error(Vpn/Wort))
```

```
summary(ssb.aov)
```

```
Error: Vpn
          Df Sum Sq Mean Sq F value   Pr(>F)
Alter      1 61.395  61.395  14.877 0.003175 **
Residuals 10 41.268   4.127
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
1

Error: Vpn:Wort
           Df Sum Sq Mean Sq F value    Pr(>F)
Wort        2 67.210  33.605 78.5055 3.391e-10 ***
Alter:Wort  2  8.468   4.234  9.8909  0.001031 **
Residuals  20  8.561   0.428
---
```

Hat Alter einen Einfluss auf Wort?

(Wir brauchen den Wort-Effekt nicht zu berichten, weil das uns nicht interessiert – war nicht Bestandteil der Fragestellung ob sich *used* vs. *swoop* vs. *who'd* unterscheiden).

Wiederholungen in derselben Zelle: Beispiel

post-hoc Tukey Test

```
Error: Vpn
          Df Sum Sq Mean Sq F value   Pr(>F)
Alter      1 61.395  61.395  14.877 0.003175 **
Residuals 10 41.268   4.127
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '
1

Error: Vpn:Wort
           Df Sum Sq Mean Sq F value    Pr(>F)
Wort        2 67.210  33.605 78.5055 3.391e-10 ***
Alter:Wort  2  8.468   4.234  9.8909  0.001031 **
Residuals  20  8.561   0.428
---
```

ssb.tukey = Tukey.rm(F2, 0.428, 20, Alter, Wort)

Wiederholungen in derselben Zelle: Beispiel

ssb.tukey

```
 [,1]
alt.swoop-alt.used     0.00000004
alt.swoop-alt.who'd    0.99485548
alt.swoop-jung.swoop   0.00000036
alt.swoop-jung.used    0.00000000
alt.swoop-jung.who'd   0.00000116
alt.used-alt.who'd     0.00000002
alt.used-jung.swoop    0.76048611
alt.used-jung.used     0.03741249
alt.used-jung.who'd    0.39631771
alt.who'd-jung.swoop   0.00000014
alt.who'd-jung.used    0.00000000
alt.who'd-jung.who'd   0.00000044
jung.swoop-jung.used   0.00197170
jung.swoop-jung.who'd 0.98793755
jung.used-jung.who'd   0.00048910
```

Alter hatte einen signifikanten Einfluss auf F2 ($F(1, 10)=14.88$, $p < 0.01$) und es gab eine signifikante Interaktion zwischen Alter und Wort ($F(2, 20) = 78.51$, $p < 0.01$). Post-hoc Tukey Tests zeigten signifikante Unterschiede ($p < 0.05$) zwischen den Altersgruppen für alle 3 Wortkategorien