

Die Varianzanalyse II

Jonathan Harrington

```
library(ggplot2)
```

```
library(ez)
```

```
source(file.path(pfadu, "phoc.txt"))
```

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

```
v.df = read.table(file.path(pfadu, "vokal.txt"))
```

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

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

Die Varianzanalyse II

1. Zwei Faktoren
2. Interaktionen zwischen den Faktoren
3. Post-hoc t-tests (wenn Interaktionen vorliegen)
4. Between Faktoren und 'Balanced design'
5. Wiederholungen in within-Stufen
6. Sphericity (mehr als 2 within-Stufen).

1. Zwei Faktoren

Inwiefern wird F2 vom Dialekt und Geschlecht beeinflusst?

head(dg)

names(dg)

with(dg, table(Vpn, interaction(Region, Gen)))

	between/within?	Vpn	A.m	B.m	C.m	A.w	B.w	C.w
Gender	between	S1	1	0	0	0	0	0
		S10	1	0	0	0	0	0
		S11	0	1	0	0	0	0
		S12	0	1	0	0	0	0
Region	between	S13	0	1	0	0	0	0
		S14	0	1	0	0	0	0
		...						

1. Zwei Faktoren

Bei 2 Faktoren, gibt es immer 3 Fragen:

Frage zu Faktor 1

Hat Gender einen Einfluss auf F2?

Frage zu Faktor 2

Hat Region einen Einfluss auf F2?

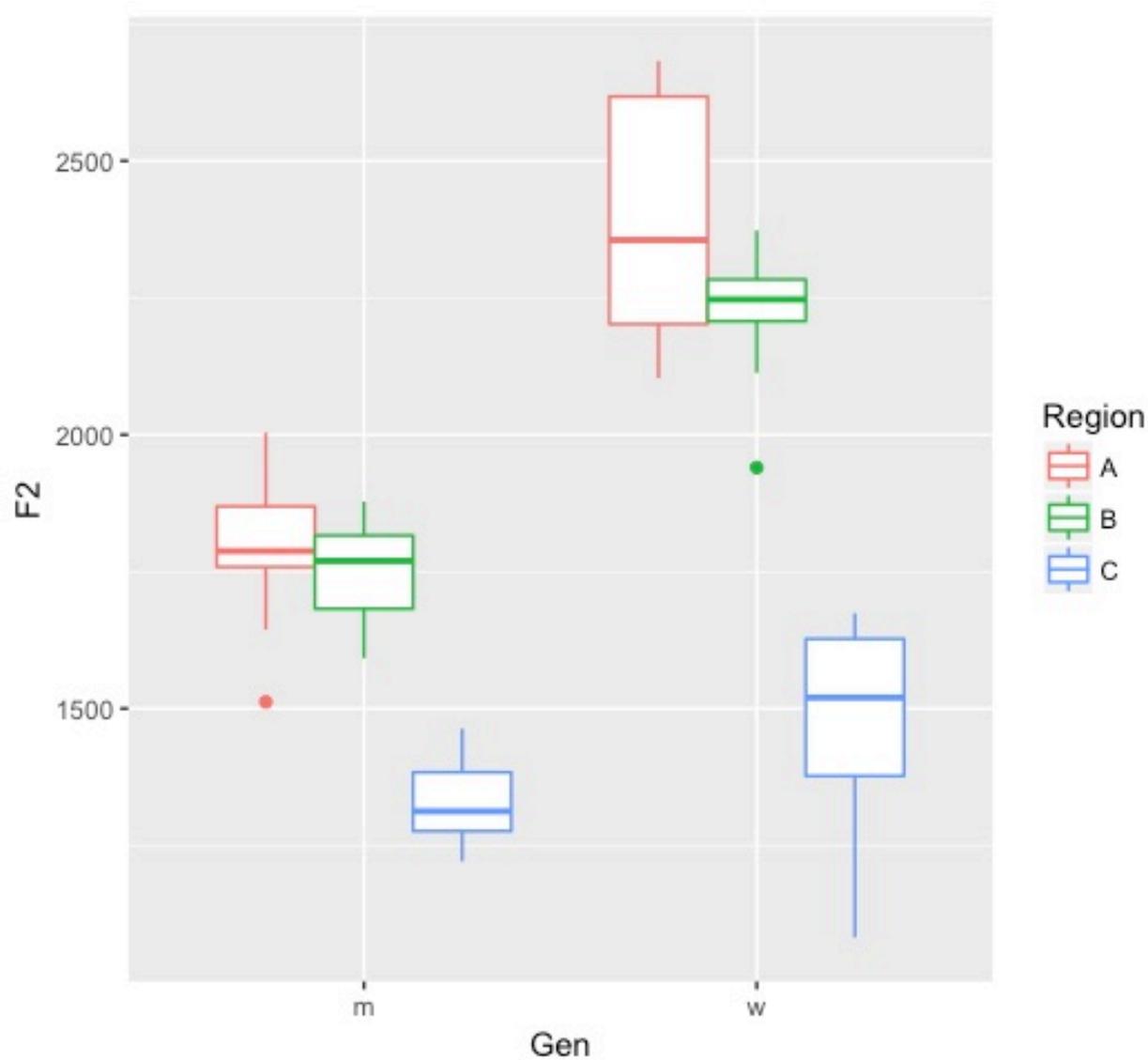
Frage zur Interaktion

Gibt es eine Interaktion zwischen Region und Gender? =

Ist der Unterschied zwischen männlich und weiblich derselbe in allen 3 Regionen?

1 Zwei Faktoren

Hat Region einen Einfluss auf F2? Hat Gender einen Einfluss auf F2?
`ggplot(dg) + aes(y = F2, x = Gen, colour = Region) + geom_boxplot()`



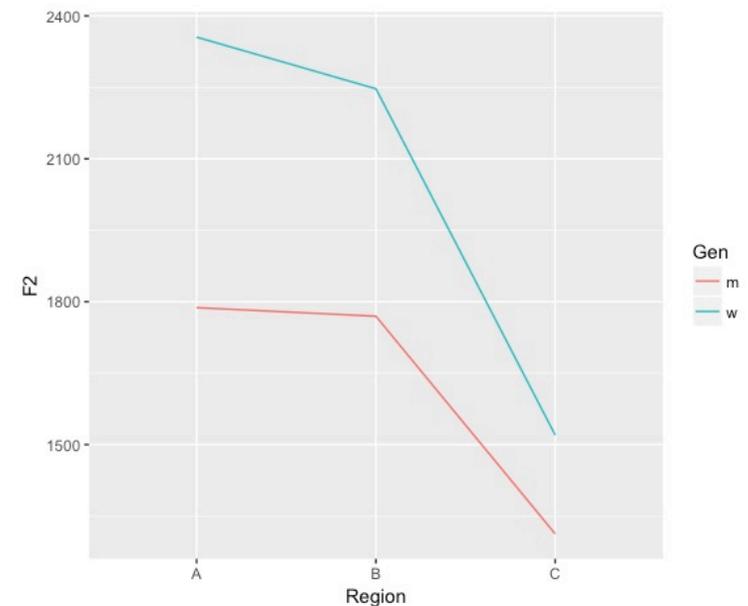
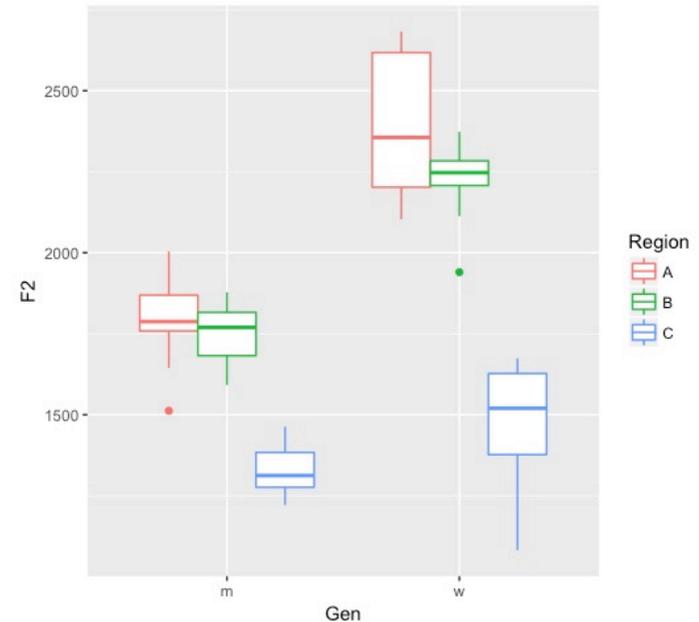
2. Gibt es eine Interaktion zwischen Region und Gender?

Bedeutung: ist der Unterschied zwischen männlich und weiblich ähnlich in den 3 Regionen?

Wenn ja, müsste der Abstand zwischen den m-w Medianen ähnlich sein

d.h. diese Linien müssten mehr oder weniger parallel zu einander sein:

```
dg.m = aggregate(F2 ~ Gen * Region,  
median, data = dg)
```



```
ggplot(dg.m) + aes(y = F2, x = Region, group=Gen, colour = Gen) + geom_line()
```

2. Zwei Faktoren und Interaktionen

ezANOVA(dg, .(F2), .(Vpn), between =.(Region, Gen))

Effect	DFn	DFd	F	p	p<.05	ges
1	Region	2	54	119.63719	1.439560e-20	* 0.8158721
2	Gen	1	54	106.14696	2.353977e-14	* 0.6628097
3	Region:Gen	2	54	12.08336	4.602985e-05	* 0.3091690

F2 wurde signifikant von der Region ($F[2,54] = 119.6$, $p < 0.001$) und von Geschlecht ($F[1,54] = 106.1$, $p < 0.001$) beeinflusst und es gab eine signifikante Interaktion zwischen diesen Faktoren ($F[2,54] = 12.1$, $p < 0.001$).

3. post-hoc t-tests

Wenn eine Interaktion vorliegt, sollte durch t-tests geprüft werden, ob sich alle Paare von Stufen-Kombinationen in der abhängigen Variable (hier F2) unterscheiden.

Paare wie: A-männlich vs. B-männlich, A-männlich vs. A-weiblich usw...

Die Anzahl dieser Tests:

Region: 3 Stufen. Geschlecht: 2 Stufen = $3 \times 2 = 6$ Stufen.

Alle Paare davon:

$6!/(4! \times 2!) = 15$ Testpaare

$\text{factorial}(6)/(\text{factorial}(4) * \text{factorial}(2))$

3. post-hoc t-Tests

Data-Frame

Versuchspersonen



```
phoc(dg, .(F2), .(Vpn), .(Region, Gen))
```



Abhängige Variable

Alle Faktoren, die
post-hoc geprüft
werden sollen
(egal ob 'within'
oder 'between')

\$res: die Ergebnisse der t-tests

\$name: die Testpaare

\$paired: ob ein gepaarter oder ungepaarter t-test durchgeführt wurde

\$bonf: Anzahl der durchgeführten Tests

3. post-hoc t-tests und Bonferroni-Korrektur

`prob-adj` ist die sogenannte Bonferroni-Korrektur

```
$res
          t          df      prob-adj
A:m-B:m    0.8313356  15.22192  1.0000000e+00
A:m-C:m    8.7155048  13.98591  7.531888e-06
A:m-A:w   -7.1586378  15.68960  3.814827e-05
...
```

Bonferroni-Korrektur: Der Wahrscheinlichkeitswert der individuellen t-tests wird mit der Anzahl der theoretisch möglichen Testkombinationen (15 in diesem Fall) multipliziert.

Der Grund: Je mehr post-hoc Tests durchgeführt werden, um so wahrscheinlicher ist es, dass ein von den vielen Tests per Zufall signifikant sein wird. Die Bonferroni-Korrektur ist eine Maßnahme dagegen.

3. post-hoc t-tests und Auswahl

Nicht alle t-tests werden benötigt sondern eher nur Vergleiche zwischen Stufen von **einem** Faktor, wenn die Stufen aller anderen Faktoren **konstant** sind.

1. Unterscheiden sich die Regionen desselben Geschlechts?
(Region variiert, Geschlecht ist konstant).

A vs B in Männern

A vs B in Frauen

A vs C in Männern

A vs C in Frauen

B vs C in Männern

B vs C in Frauen

2. Unterscheiden sich Männer und Frauen in derselben Region?
(Geschlecht variiert, Region ist konstant)

m vs. w in A

m vs. w in B

m vs. w in C

Aber *nicht* wenn beide Faktoren variieren.

m-A vs. w-B, m-C vs w-A usw.

1. Unterscheiden sich die Regionen im selben Geschlecht (Region variiert, Geschlecht ist konstant)?

	t	df	prob-adj
A:m-B:m	0.8313356	15.22192	1.000000e+00
A:m-C:m	8.7155048	13.98591	7.531888e-06
A:m-A:w	-7.1586378	15.68960	3.814827e-05
A:m-B:w	-7.0876370	17.28901	2.482025e-05
A:m-C:w	4.1291502	16.66330	1.092264e-02
B:m-C:m	10.6837180	17.65040	5.898958e-08
B:m-A:w	-8.5319197	12.11771	2.708776e-05
B:m-B:w	-9.8137671	16.97522	3.098841e-07
B:m-C:w	3.9943383	12.84208	2.345945e-02
C:m-A:w	-14.3108625	11.38030	1.881869e-07
C:m-B:w	-19.4274325	15.79614	2.840451e-11
C:m-C:w	-2.1074735	11.95523	8.530528e-01
A:w-B:w	2.2029457	13.88744	6.749777e-01
A:w-C:w	9.8529861	17.77397	1.896196e-07
B:w-C:w	10.2391336	14.86067	5.992353e-07

2. Unterscheiden sich Männer und Frauen derselben Region? (Geschlecht variiert, Region ist konstant)?

Die anderen werden meistens nicht benötigt

3. post-hoc t-tests und Auswahl

Die benötigten Tests können mit `phsel()` ausgesucht werden

```
vok.ph = phoc(dg, .(F2), .(Vpn), .(Region, Gen))
```

```
phsel(vok.ph$res, 1)
```

```
phsel(vok.ph$res, 2)
```

oder

```
phsel(vok.ph$res)
```

gibt die post-hoc t-Tests für
Region (mit Gender konstant)

gibt die post-hoc t-Tests für
Gender (mit Region konstant)

3. post-hoc t-tests

ersichtlicher wenn auf z.B. 3 Zahlen aufgerundet:

`p1 = phsel(vok.ph$res)`

`p2 = phsel(vok.ph$res, 2)`

`round(p1, 3)`

`round(p2, 3)`

	t	df	prob-adj
A:m-B:m	0.831	15.222	1.000
A:m-C:m	8.716	13.986	0.000
B:m-C:m	10.684	17.650	0.000
A:w-B:w	2.203	13.887	0.675
A:w-C:w	9.853	17.774	0.000
B:w-C:w	10.239	14.861	0.000

	t	df	prob-adj
A:m-A:w	-7.159	15.690	0.000
B:m-B:w	-9.814	16.975	0.000
C:m-C:w	-2.107	11.955	0.853

Post-hoc Bonferroni-adjusted t-tests zeigten signifikante **F2-Unterschiede zwischen A vs C** ($p < 0.001$) und **zwischen B vs C** ($p < 0.001$) **jedoch nicht zwischen A vs. B**. F2 von Männern und Frauen unterschieden sich signifikant für **Regionen A** ($p < 0.001$) und **B** ($p < 0.001$), jedoch nicht für C.

4. Balanced design und Between Faktoren.

Zwei Bedingungen für die Durchführung der Varianzanalyse
Between

Die selbe Anzahl pro Stufen-
Kombination

Dialekt	BY	Alter	
		jung	alt
	BY	10	10
	SH	10	10

ezANOVA() gibt eine
Warnmeldung

Dialekt	BY	Alter	
		jung	alt
	BY	4	11
	SH	6	3

5. Wiederholungen in within Stufen.

Within
Ein Wert pro Stufe pro Vpn

Anzahl der Werte

	i	e	a
S1	1	1	1
S2	1	1	1
S3	1	1	1
Sn	1	1	1

geht nicht

Sn	0	1	1
----	---	---	---

muss gemittelt werden

Sn	4	4	4
----	---	---	---

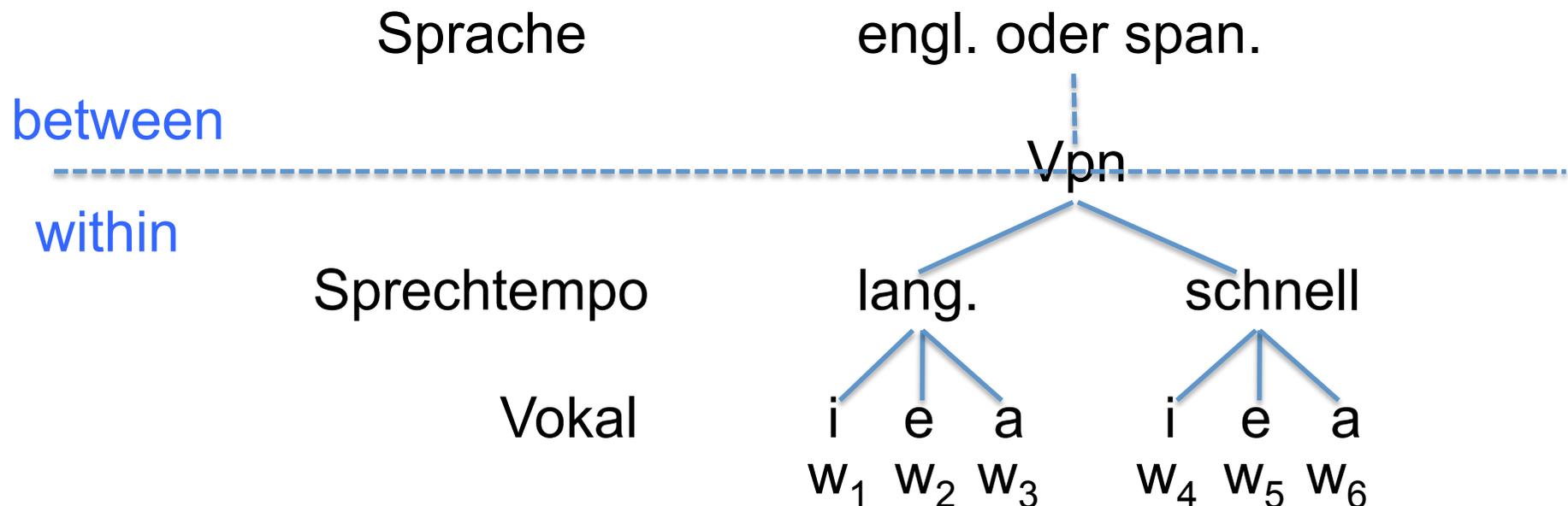
(nächste Folie)

5. Wiederholungen in within-Stufen

Wenn es n within-Stufen gibt, dann müssen es n Werte pro Vpn sein, einen Wert pro within-Stufe z.B:

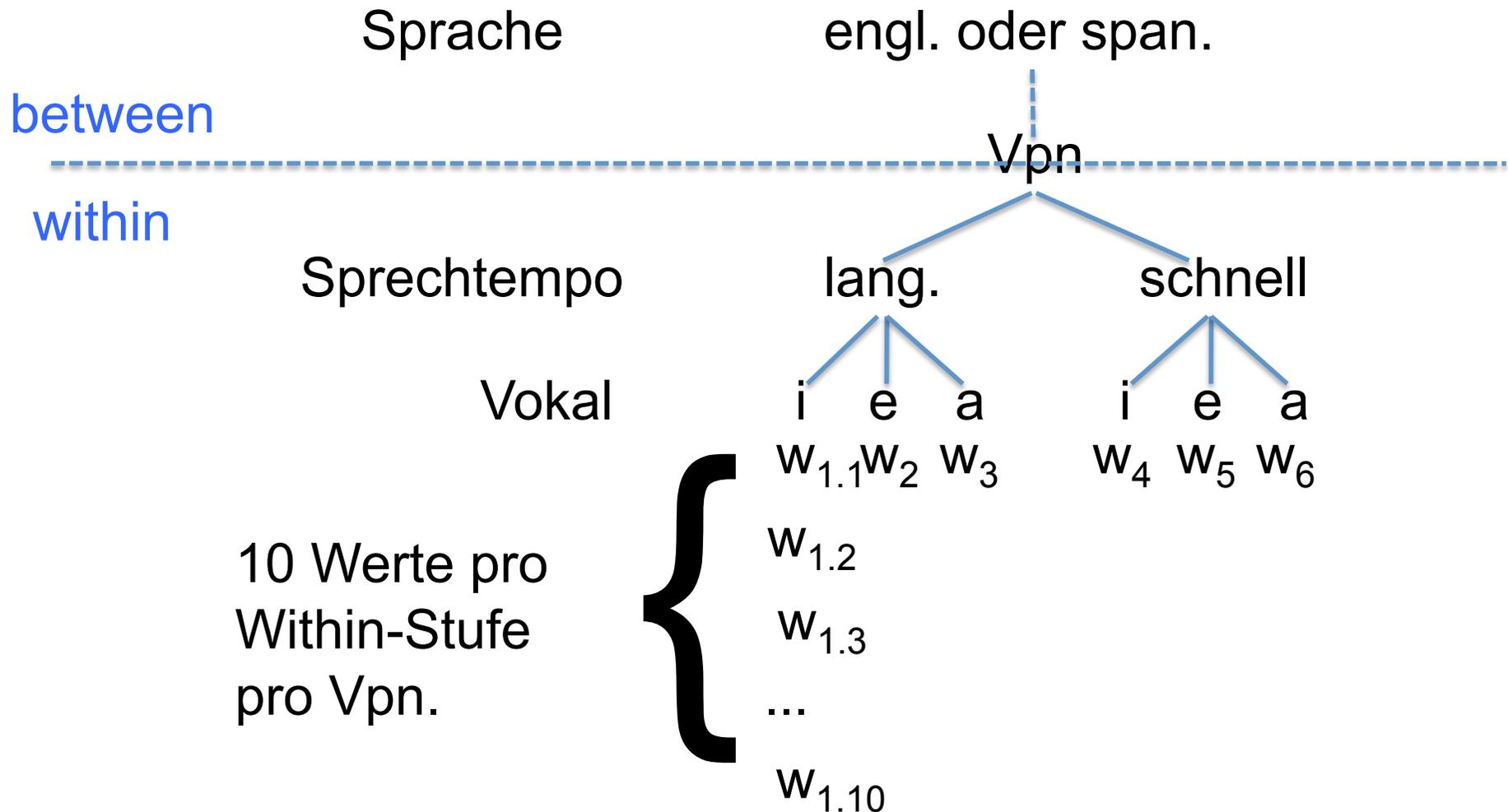
Englische und spanische Vpn produzierten /i, e, a/ zu 2 Sprechgeschwindigkeiten

Within: Vokal (3 Stufen) und Sprechgeschwindigkeit (2 Stufen)
Daher: $3 \times 2 = 6$ within-Werte pro Vpn
(ein Wert pro within-Stufe pro Vpn).

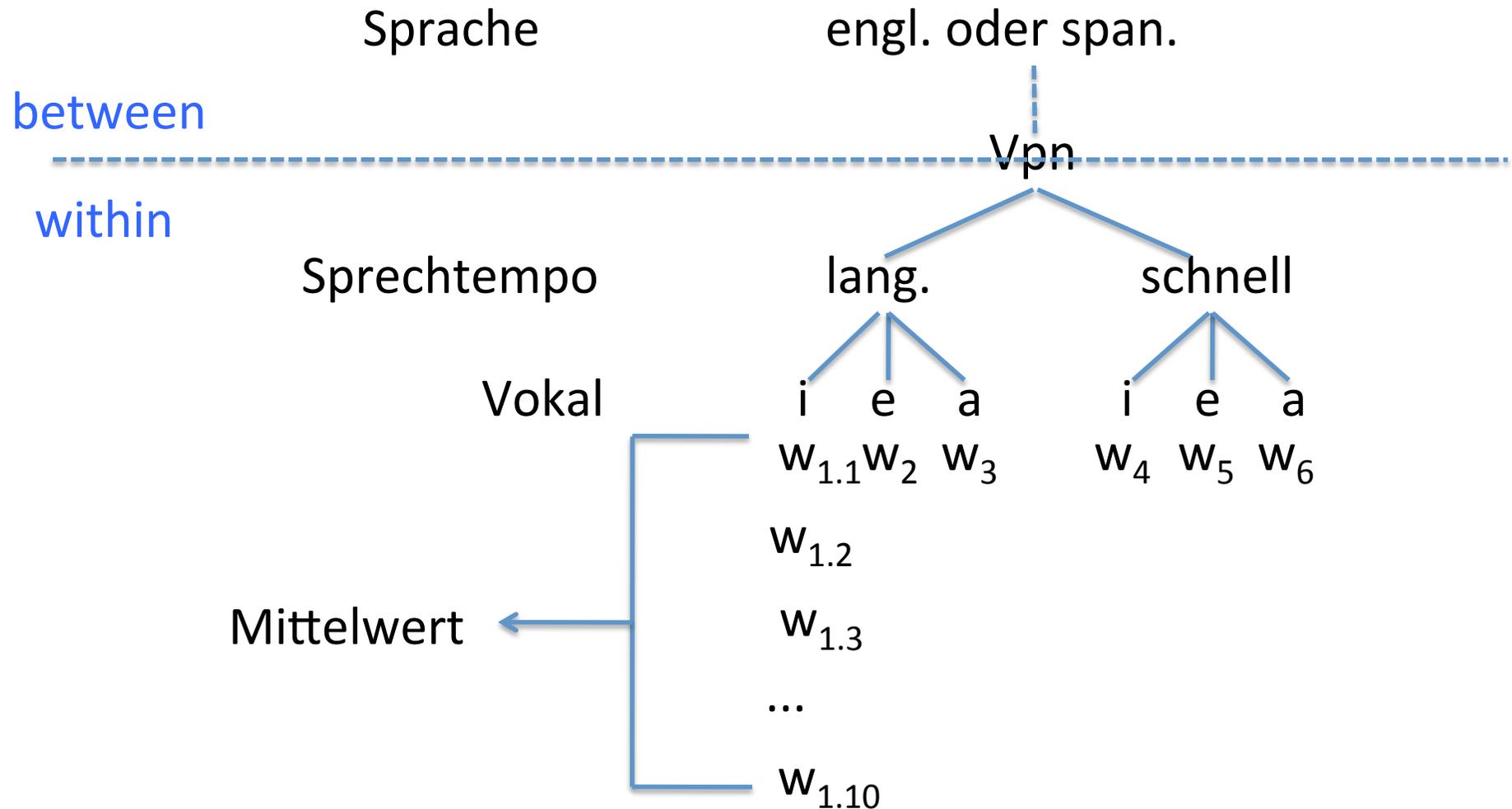


5. Wiederholungen in within-Stufen

Jedoch haben die meisten phonetischen Untersuchungen mehrere Werte pro within-Stufe. zB. jede Vpn. erzeugte /i, e, a/ zu einer langsamen und schnellen Sprechgeschwindigkeit **jeweils 10 Mal**.



Wiederholungen in derselben within-Stufe sind in einem ANOVA nicht zulässig und müssen gemittelt werden – damit wir pro Vpn. einen Wert pro within-Stufe haben (6 **Mittelwerte** pro Vpn. in diesem Beispiel).



5. Wiederholungen in within-Stufen

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

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. Inwiefern wird F2 vom Alter und Wort beeinflusst?

Faktor	within/between	wieviele Stufen?
Wort	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**

5. Wiederholungen in within-Stufen

1. Anzahl der Wort-Wiederholungen pro Sprecher prüfen

```
with(ssb, table(Vpn, interaction(Wort, Alter)))
```

Vpn	swoop.alt	used.alt	who'd.alt	swoop.jung	used.jung	who'd.jung
arkn	10	10	10	0	0	0
elwi	9	10	10	0	0	0
frwa	10	10	10	0	0	0
gisa	10	10	10	0	0	0
jach	0	0	0	10	10	10
jeny	0	0	0	10	10	10
kapo	0	0	0	10	10	10
mapr	10	10	10	0	0	0
nata	10	10	10	0	0	0
rohi	0	0	0	10	10	10
rusy	0	0	0	10	10	10
shle	0	0	0	10	10	10

5. Wiederholungen in within-Stufen

2. Über die Wort-Wiederholungen mit `aggregate()` mitteln
abhängige Variable alle anderen Variablen

```
ssbm = aggregate(F2 ~ Wort * Alter * Vpn, mean, data = ssb)
```

```
dim(ssbm)
```

```
[1] 36 4
```

```
head(ssbm)
```

```
      Wort Alter  Vpn      F2  
1 swoop   alt  arkn 10.527359
```

```
with(ssbm, table(Vpn, interaction(Wort, Alter)))
```

```
Vpn      swoop.alt used.alt who'd.alt swoop.jung used.jung who'd.jung  
arkn          1         1         1          0         0         0  
elwi          1         1         1          0         0         0  
frwa          1         1         1          0         0         0  
...
```

3. Abbildung

```
ggplot(ssbm) + aes(y = F2, x = Alter, colour = Wort) + geom_boxplot()
```

4. Anova wie üblich durchführen

```
ezANOVA(ssbm, .(F2), .(Vpn), .(Wort), between = .(Alter))
```

6. Sphericity-Korrektur

Sphericity ist die Annahme, dass die Unterschiede zwischen den Stufen eines within-Faktors **dieselbe Varianz haben**.

Wenn Sphericity nicht gegeben ist, **werden die Wahrscheinlichkeiten durch Änderungen in den Freiheitsgraden nach oben gesetzt**.

Dieses Problem kommt nur dann vor, **wenn ein within-Faktor mehr als 2 Stufen hat**.

Man soll grundsätzlich immer für Sphericity korrigieren, wenn Sphericity-Korrektur in der Ausgabe von ezANOVA() erscheint.

6. Sphericity-Korrektur

\$ANOVA

	Effect	DFn	DFd	F	p	p<.05	ges
2	Alter	1	10	14.876957	3.175409e-03	*	0.5519903
3	Wort	2	20	78.505534	3.390750e-10	*	0.5742513
4	Alter:Wort	2	20	9.890888	1.031474e-03	*	0.1452519

\$`Mauchly's Test for Sphericity`

	Effect	W	p	p<.05
3	Wort	0.5423826	0.06373468	
4	Alter:Wort	0.5423826	0.06373468	

\$`Sphericity Corrections`

	Effect	GGe	p[GG]	p[GG]<.05	HFe	p[HF]	p[HF]<.05
3	Wort	0.6860511	1.340736e-07	*	0.7587667	3.342362e-08	*
4	Alter:Wort	0.6860511	4.370590e-03	*	0.7587667	3.120999e-03	*

1. Die **betroffenen Freiheitsgrade** werden mit dem **Greenhouse-Geisser-Epsilon** multipliziert, wenn er unter 0.75 liegt¹, sonst mit dem **Huynh-Feldt-Epsilon**: sollte in diesem letzten Fall der H-F-Epsilon > 1 sein, dann einfach **die ursprünglichen Freiheitsgrade** nehmen d.h. **keine Korrektur einsetzen**.

Wort: $F[2,20] \rightarrow F[2 * 0.6860511, 20 * 0.6860511] = F[1.4, 13.7]$

Alter × Wort Interaktion: $F[2,20] \rightarrow F[1.4, 13.7]$

1. Nach Girden (1992) *ANOVA: Repeated Measures*. Sage, Ca.

6. Sphericity-Korrektur

```
$ANOVA
```

	Effect	DFn	DFd	F	p	p<.05	ges
2	Alter	1	10	14.876957	3.175409e-03	*	0.5519903
3	Wort	2	20	78.505534	3.390750e-10	*	0.5742513
4	Alter:Wort	2	20	9.890888	1.031474e-03	*	0.1452519

```
$`Sphericity Corrections`
```

	Effect	GGe	p[GG]	p[GG]<.05	HFe	p[HF]	p[HF]<.05
3	Wort	0.6860511	1.340736e-07	*	0.7587667	3.342362e-08	*
4	Alter:Wort	0.6860511	4.370590e-03	*	0.7587667	3.120999e-03	*

2. Die neuen damit verbundenen Wahrscheinlichkeiten sind $p[GG]$ (wenn mit GGe multipliziert wurde) sonst $p[HF]$.

Das sind die Wahrscheinlichkeiten mit den korrigierten Freiheitsgraden
z.B. $1 - pf(9.8908882, 2 * 0.6860511, 20 * 0.6860511)$

```
[1] 0.004370589
```

Alter ($F[1,10] = 14.9$, $p < 0.001$), Wort ($F[1.4, 13.7] = 78.5$, $p < 0.001$) sowie die Interaktion Wort und Alter ($F[1.4, 13.7] = 9.9$, $p < 0.01$) hatten einen signifikanten Einfluss auf F2.