

Einige Einschränkungen in der Durchführung einer Varianzanalyse

Jonathan Harrington

ANOVA und 'balanced design'

40 Vpn. 20 aus BY (davon 10 alt, 10 jung), 20 aus SH
(davon 10 alt, 10 jung) produzierten /i/, /e/, /a/

Between

Dieselbe Anzahl und
mindestens 5 pro Stufe

Dialekt		Alter	
		jung	alt
BY	10	10	10
	SH	10	10

geht meistens nicht

Dialekt		Alter	
		jung	alt
BY	2	2	2
	SH	2	2

Dialekt		Alter	
		jung	alt
BY	4	11	11
	SH	6	3

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
----	---	---	---

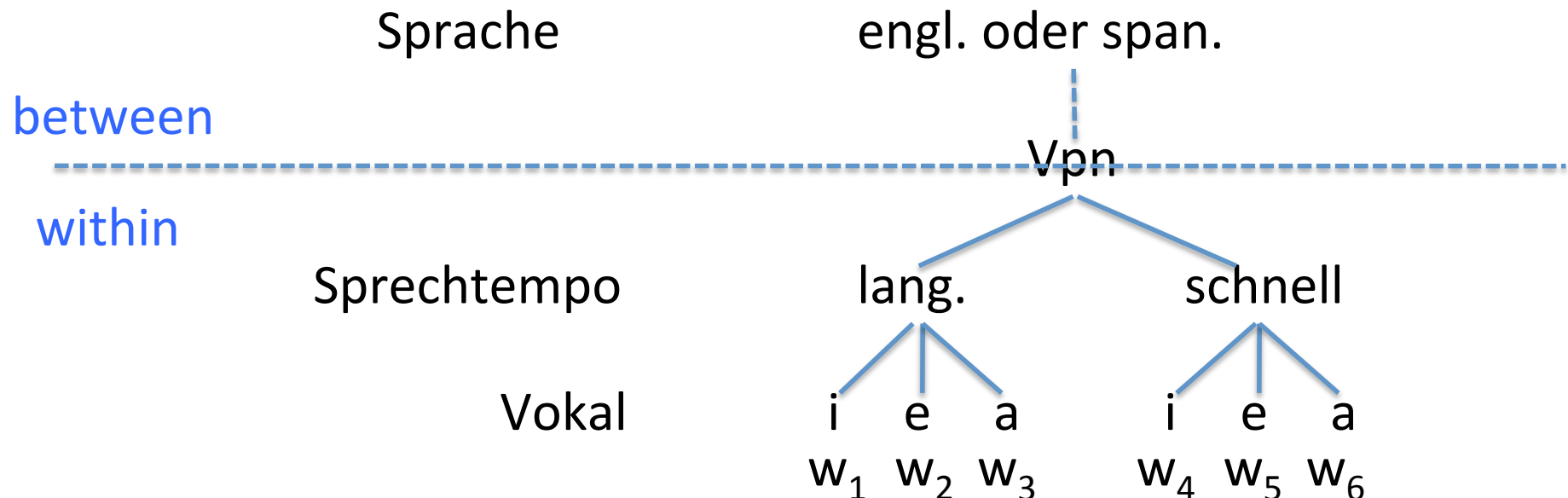
(nächste Folie)

Wiederholungen in within-Stufen

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

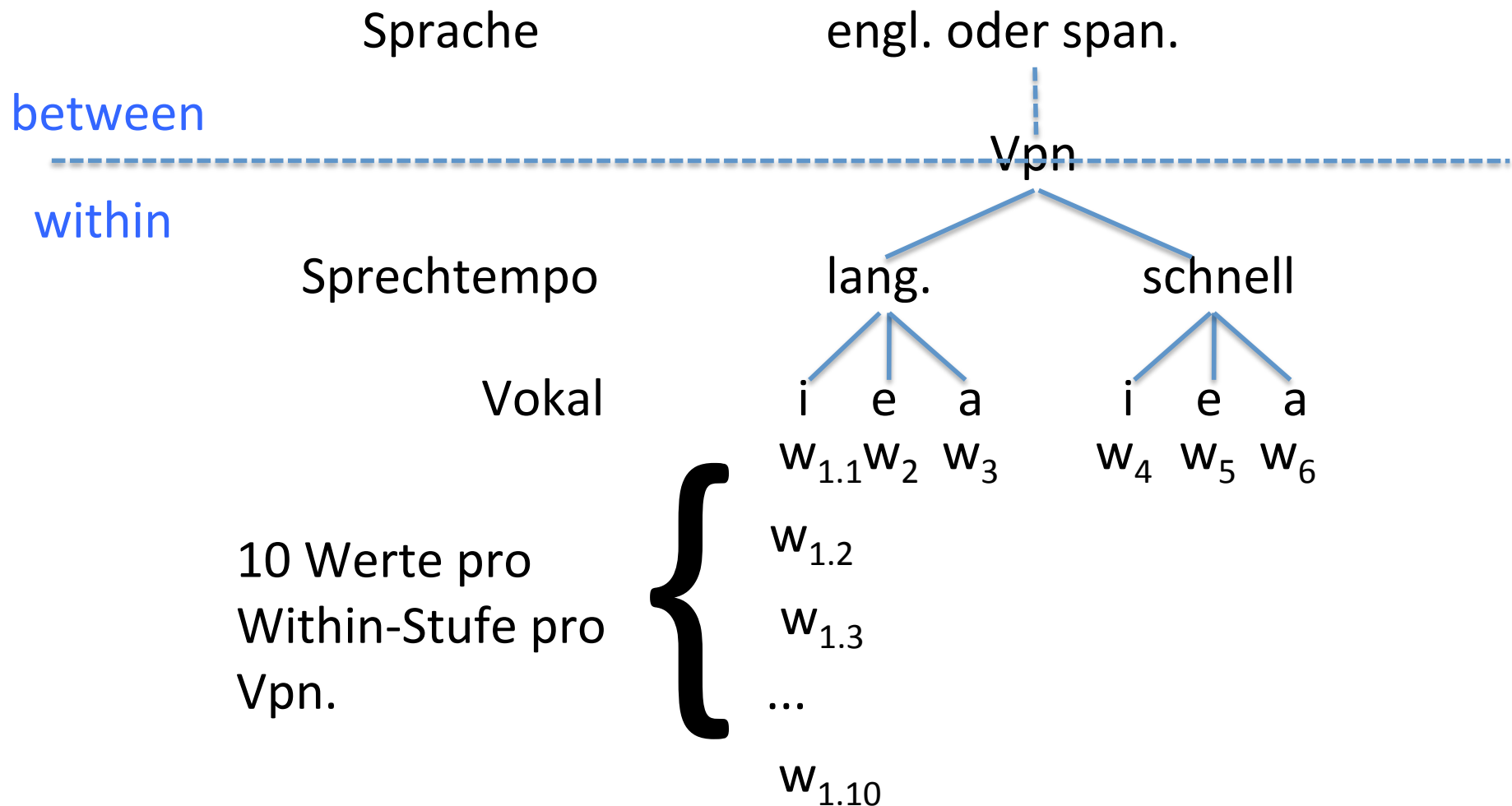
Englische und spanische Vpn produzierten /i, e, a/ zu 2 Sprechgeschwindigkeiten

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

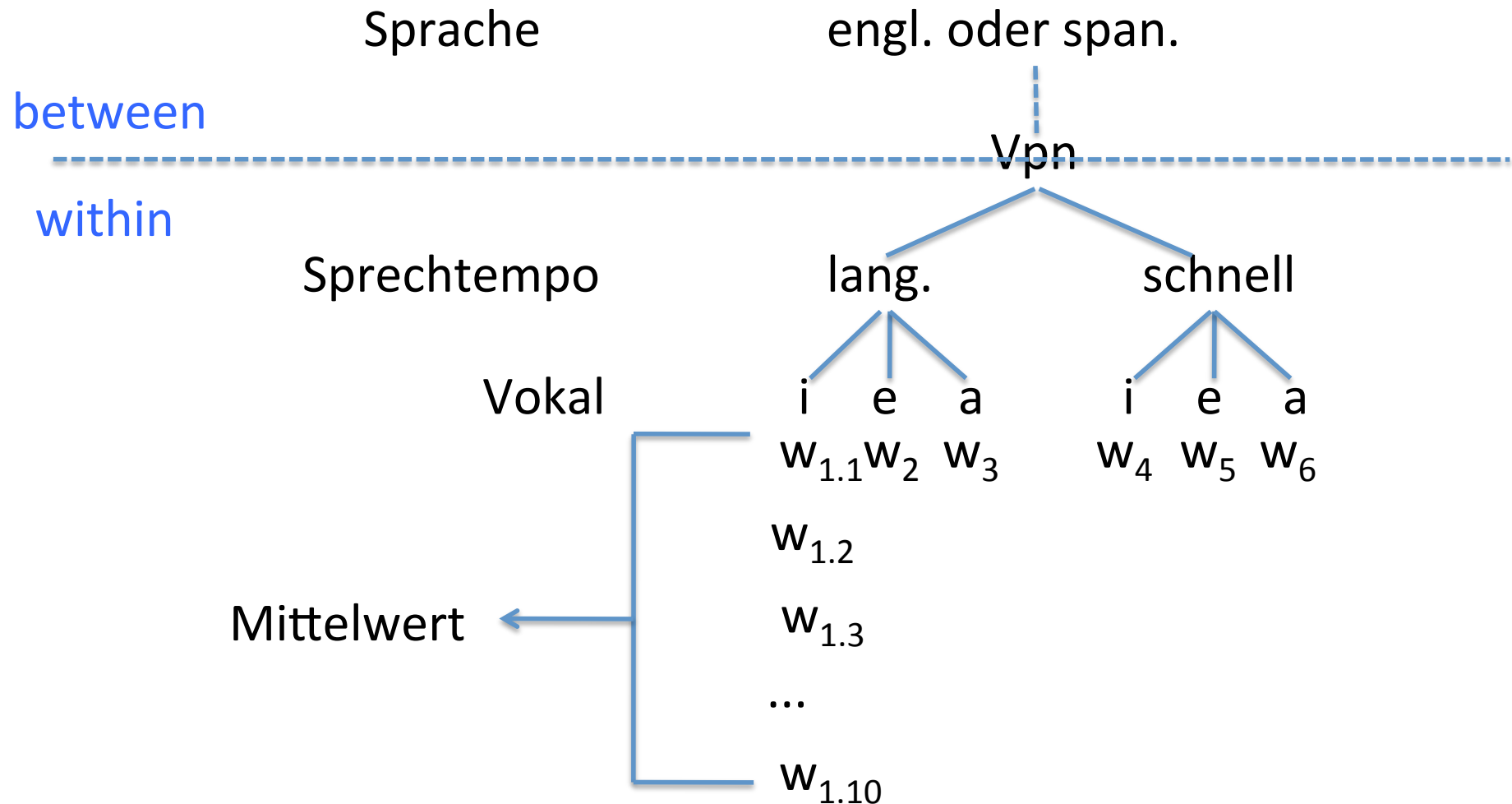


Wiederholungen in within-Stufen

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).



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

Wiederholungen in derselben Zelle

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

Wiederholungen in derselben Zelle

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

```
ssbm = with(ssb, aggregate(F2, list(Wort, Alter, Vpn), mean))
```

```
dim(ssbm)
```

```
[1] 36 4
```

```
head(ssbm)
```

```
Group.1 Group.2 Group.3      x1  
swoop    alt    arkn 10.527359
```

3. Neue Namen vergeben

```
names(ssbm) = c("Wort", "Alter", "Vpn", "F2")
```

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

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

4. Anova wie üblich durchführen

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

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 tritt nur auf 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.

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 liegt, 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 $\times$ Wort Interaktion: $F[2,20] \rightarrow F[1.4, 13.7]$

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 verbunden Wahrscheinlichkeiten sind **p[GG]** (wenn mit GGe multipliziert wurde) sonst **p[HF]**.

Das sind die Wahrscheinlichkeiten mit den korrigierten Freiheitsgraden

z.B. $1 - \text{pf}(9.8908882, 2 * 0.6860511, 20 * 0.6860511)$

[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 ($F[1.4, 13.7] = 9.9$, **$p < 0.001$**) hatten einen signifikanten Einfluss auf F2.