



EXPERIMENTELLE DESIGNS

Katharina Spalek
HU Berlin

Linguistischer Methodenworkshop 2016

Was ist das?

- **Allgemein:** Alle Entscheidungen, die vor einem Experiment getroffen werden müssen.
- Experiment: methodisch angelegte Untersuchung zur Gewinnung von Daten
- Speziell:
 - Operationalisierung
 - Manipulation von Variablen
 - Kontrollbedingungen

Operationalisierung

- “Messbarmachung” eines zunächst nur theoretischen Konstrukts
- *Tipp: das Gemessene muss (im weitesten Sinne) etwas Zählbares sein*

Operationalisierung

- Seminar “Versprecher”, WS 2015/ 2016
- Aufgabe: Anhand eines selbstgesammelten Versprecherkorpus die Frage beantworten, ob Phonemvertauschungen (“mit Bleisch und Flut”) häufiger vorkommen, wenn die vertauschten Phoneme einander ähnlich sind.

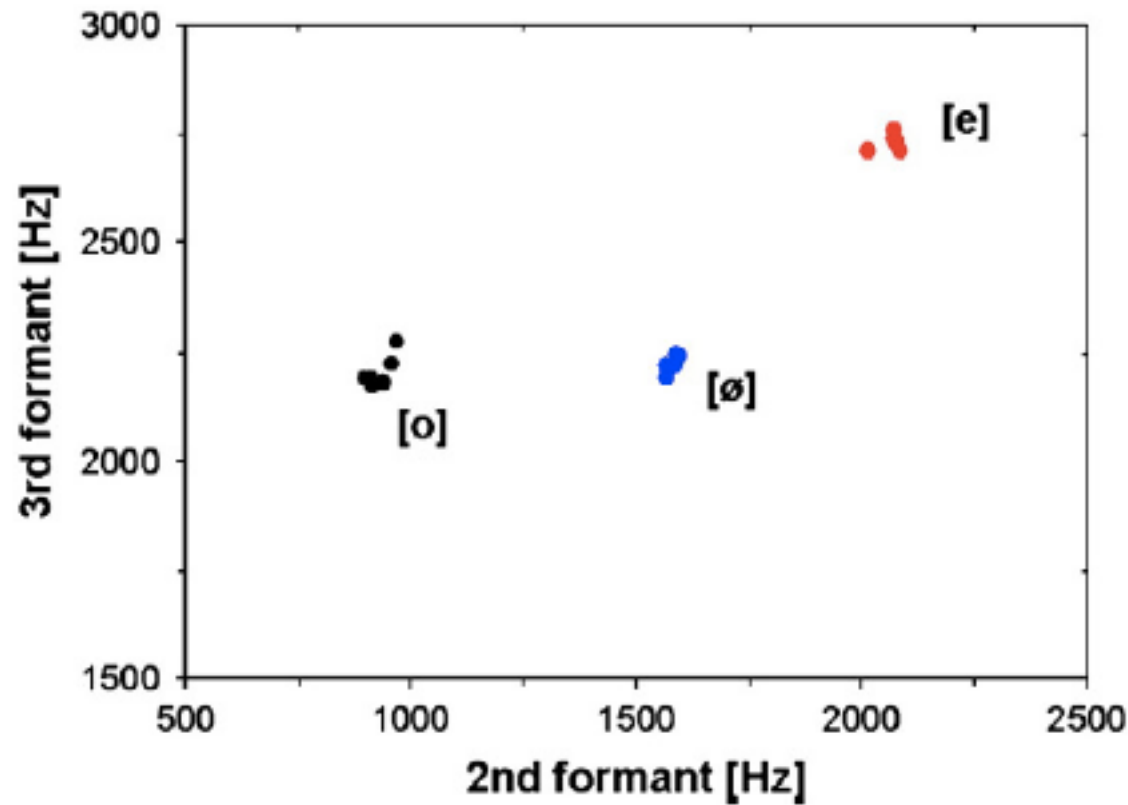
Operationalisierung

- Problem: Wie kann man die Ähnlichkeit zwischen Phonemen messen?
- Lösung 1: Anzahl übereinstimmender Merkmale (Artikulationsart, Artikulationsort, Stimmhaftigkeit) zählen.

Operationalisierung Phonemähnlichkeit

- Problem: Wie kann man die Ähnlichkeit zwischen messen?
- Lösung 2: Lage der Formanten.

Operationalisierung Phonemähnlichkeit



Operationalisierung Phonemähnlichkeit

- Problem: Wie kann man die Ähnlichkeit zwischen messen?
- Lösung 3: Wie leicht verwechselt man zwei Phoneme? (zusätzliche Annahme: Je ähnlicher zwei Stimuli einander sind, umso häufiger werden sie miteinander verwechselt).

Operationalisierung Phonemähnlichkeit

- “confusion matrix”
- Vorgehen: CV und VC Silben wurden aufgenommen und in Rauschen abgespielt. ProbandInnen sollten entweder den Konsonanten oder den Vokal identifizieren.
- Beispielergebnis:

	p	t	k	b	d	g	h
p	85	27	27	3	0	1	6
t	0	145	0	0	0	2	1
k	1	3	139	0	0	0	0

Kleine Variablenkunde

- Unabhängige Variable (UV) -> das, was vom Experimentleiter gezielt manipuliert wird. Üblicherweise hat eine unabhängige Variable mehrere Ausprägungen (mindestens zwei).
- Abhängige Variable (AV) -> das, was man misst und was sich laut Hypothese in Abhängigkeit von der UV ändern sollte.
- Störvariablen (manchmal auch: konfundierende Variablen): Einflussfaktoren, die sowohl die AV als auch die UV beeinflussen könnten und alternative Erklärungen zu den Daten liefern könnten.

Kleine Variablenkunde: “within” oder “between”?

- “Within”: ein Faktor wird in allen Ausprägungen bei derselben Versuchsperson und im selben “Item” getestet.
- “Between”: ein Faktor wird bei einer Gruppe Versuchspersonen in Ausprägung A getestet und bei einer zweiten Gruppe Versuchspersonen in Ausprägung B.

Ein nicht-linguistisches Beispiel

- University of Bristol, Tobacco and Alcohol Research Group, M. Munafó



Design

- Aufgabe: für zwei nahezu identische Gesichter angeben, welches freundlicher ist
- UV: 0, 1, 3 Alkoholeinheiten
- AV: % korrekt (bzw. ein anderes Maß der Diskriminationsfähigkeit)



- Man kann nicht gleichzeitig Alkohol zu sich nehmen und keinen Alkohol zu sich nehmen.
- Ein klarer Fall für between participants, oder?



- Problem 1: Wir wissen, dass Männer und Frauen Alkohol unterschiedlich gut vertragen.
- Lösung 1: Genauso viele Männer wie Frauen in die Alkoholgruppen und in die Kontrollgruppe!

Alkohol	Männer	Frauen
0	10	10
1	10	10
3	10	10

Balanced design



- Problem 2: Wir wissen, dass das Körpergewicht eine Rolle dabei spielt, wie gut Alkohol vertragen wird.
- Lösung 2: Wenn in der ersten Gruppe ein 75 kg schwerer Mann teilgenommen hat, darauf achten, dass der zweiten und dritten Gruppe ein ähnlich schwerer Mann zugewiesen wird.

0 Alkohol	1 Alkohol	3 Alkohol
Mann, 78 kg	Mann, 75 kg	Mann, 80 kg
Frau, 65 kg	Frau, 66 kg	Frau, 62 kg

Pairwise matching



- Problem 3, 4, 5, ...: Wir wissen, dass der Lebensstil, Genetik, Ernährung, Müdigkeit ... eine Rolle dabei spielen, wie gut Alkohol vertragen wird. Außerdem unterscheiden sich Menschen in ihrer Fähigkeit, Gesichtsausdrücke zu deuten.
- Lösung: ?

0 Alkohol	1 Alkohol	3 Alkohol
Katharina, 3.5.	Katharina, 27.4.	Katharina, 10.5.

Doch ein within design!
Optimale Kontrolle über Störvariablen



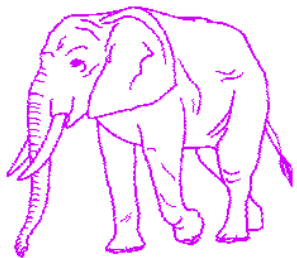
Confound it! Wie gehe ich mit Störvariablen um?

- Eliminieren.
- Gleichhalten.
- In die Analyse miteinbeziehen.

Items

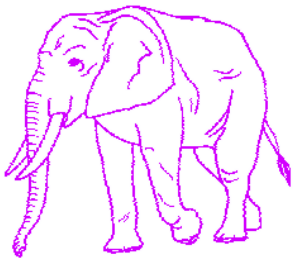
Goldstandard: Vollständiges Kreuzen

- Frage: Spielt bei der Selektion des englischen Artikels in einer Artikel-Adjektiv-Nomen Phrase die Phonologie des Nomens eine Rolle? cf. an apple vs. a tasty apple; a pear vs. an old pear
- AV: Reaktionszeit (Onset) bei der Bildbenennung
- UV: Konsistenz (Ausprägungen: konsistent/ inkonsistent)

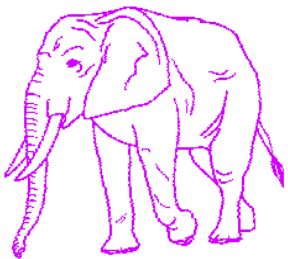
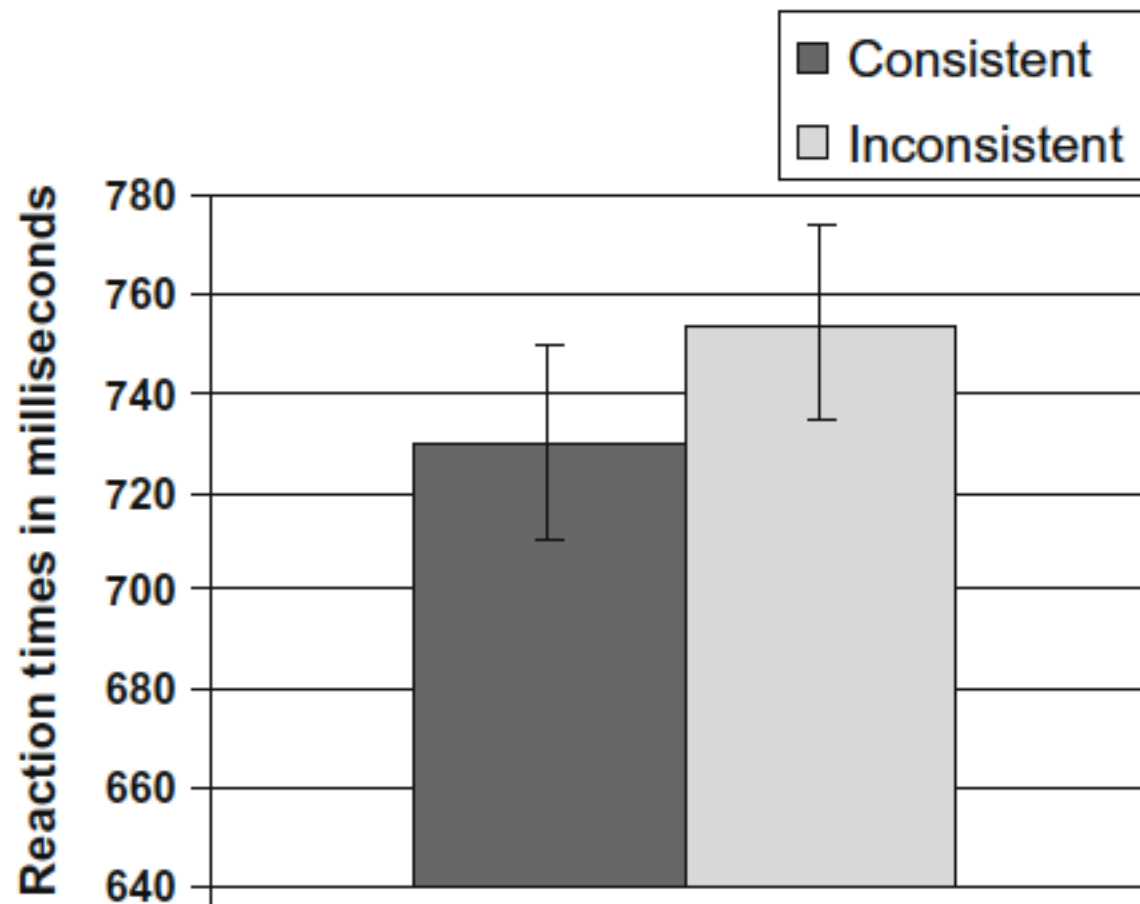


Goldstandard: Vollständiges Kreuzen

Konsistent	Inkonsistent
a purple giraffe	an orange giraffe
an orange elephant	a purple elephant
a purple turtle	an orange turtle
an orange alligator	a purple alligator
a purple leg	an orange leg
an orange eye	a purple eye
...	...



Goldstandard: Vollständiges Kreuzen



Silberstandard ...

- Erst noch einmal Gold:
- Frage: Wie groß ist der Skopus der phonologischen Planung? Genauer: Wird über das Wort hinaus geplant, bevor die Artikulation beginnt?

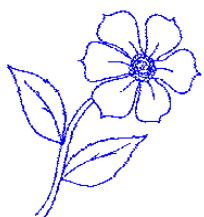
Phonemwiederholung	Kontrolle
green goat	blue goat
blue basket	red basket
red roof	green roof
...	...

Reaktionszeit "green goat" < "blue goat"

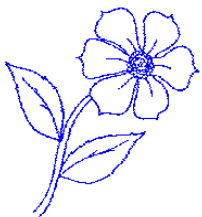
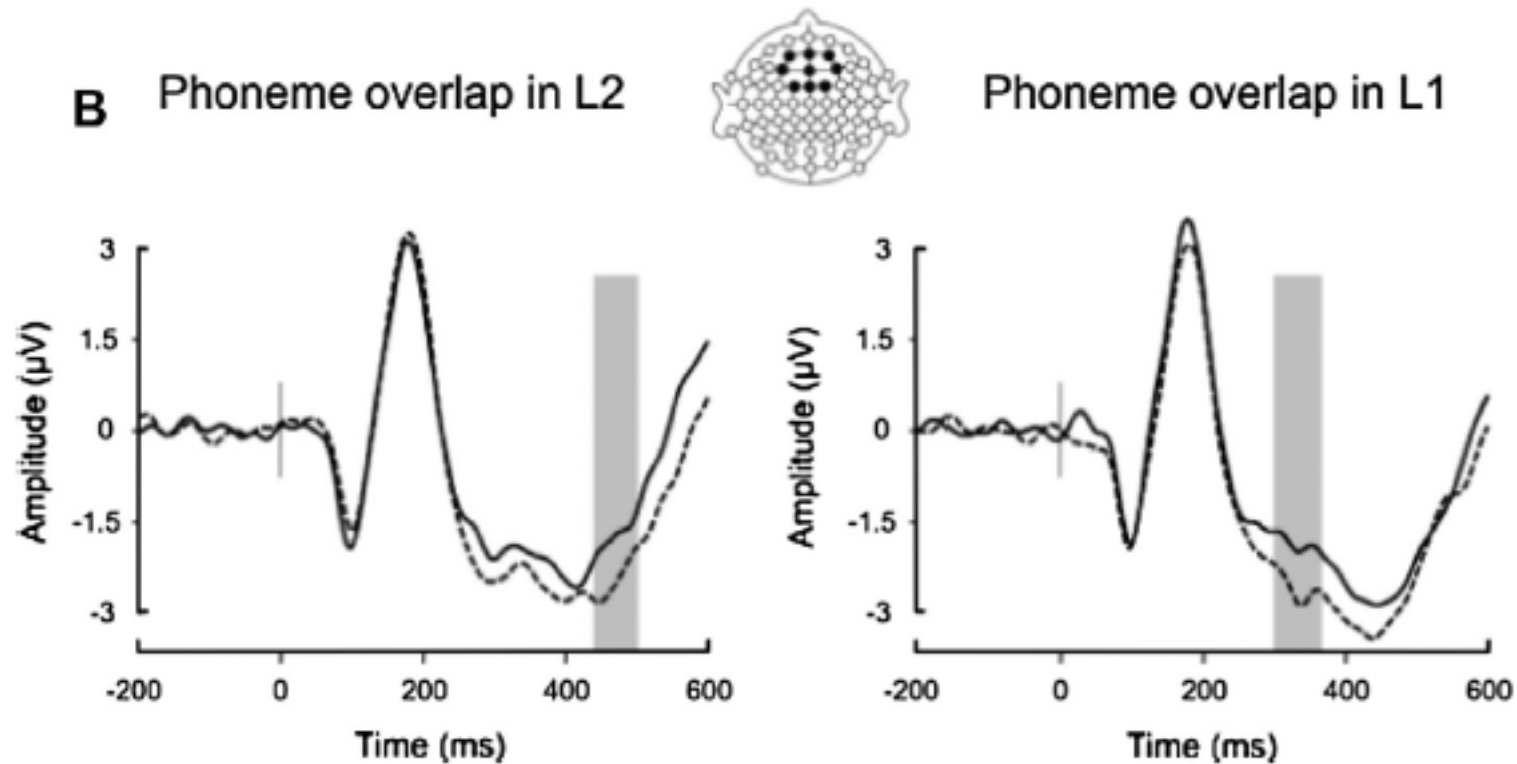
Silberstandard ...

- Frage: Wird die Phonologie in der L1 aktiviert, wenn ein Sprecher seine L2 spricht?

	L1	L2
Wiederholung	red slide green brain blue flower	red roof green goat blue basket
Kontrolle	blue slide red brain green flower	blue roof red goat green basket



Silberstandard ...



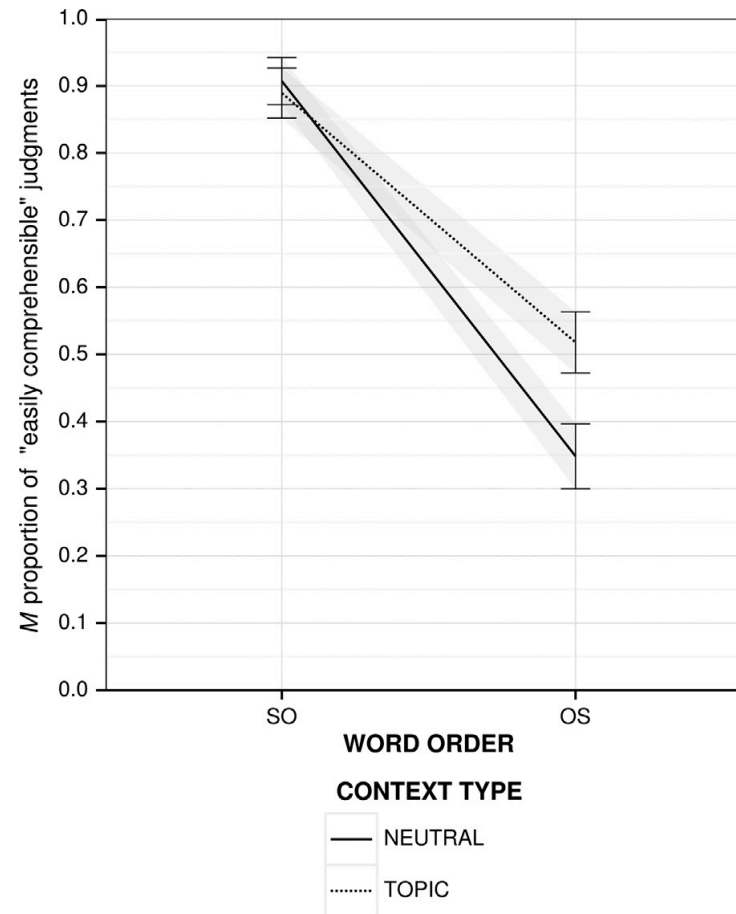
Noch ein silbernes Beispiel

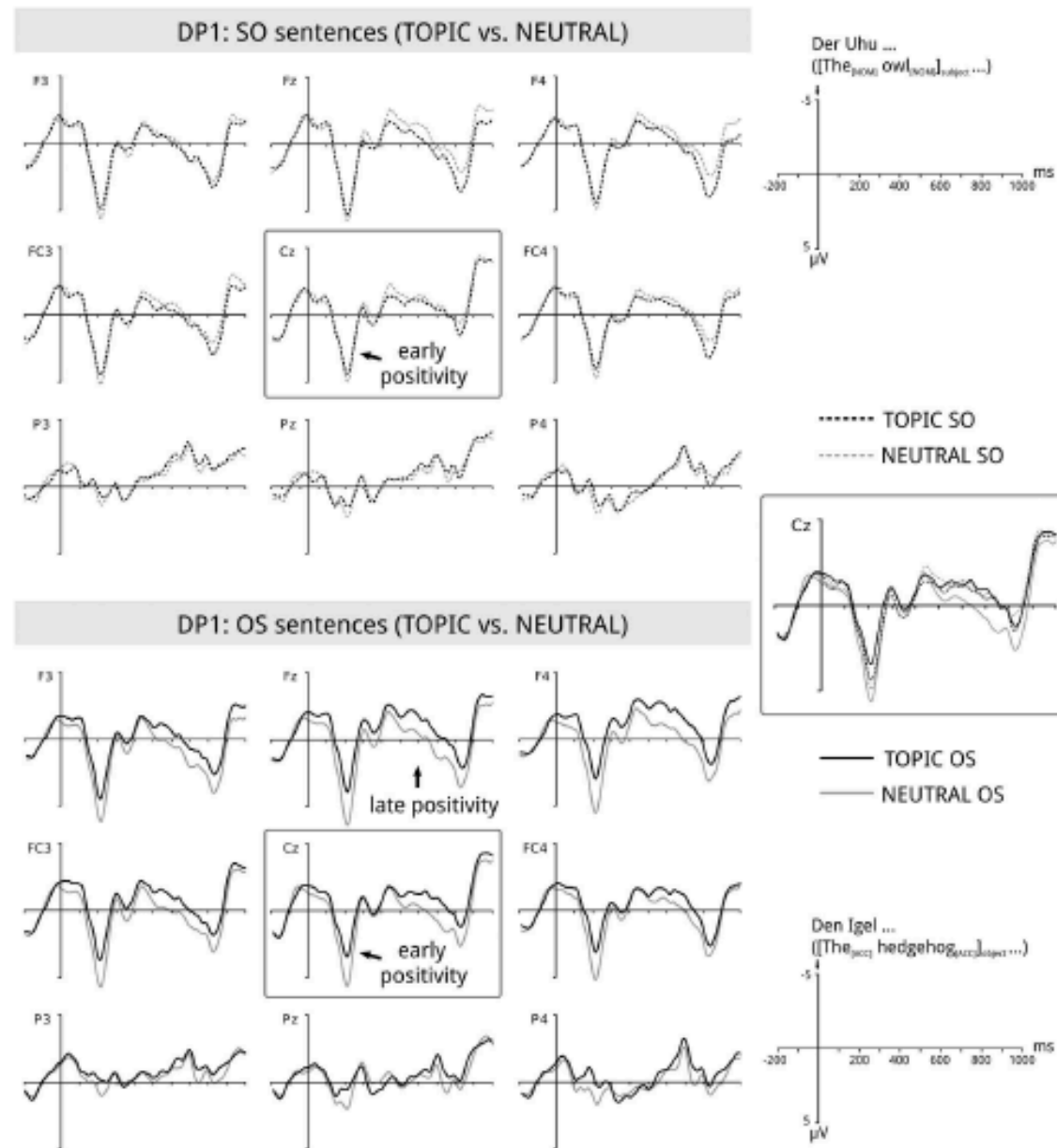
- Frage: Beeinflusst Topikinformation die bekannten Verarbeitungsschwierigkeiten für OVS-Sätze?

	Topik	Neutral
SO	Was ist mit dem Hamster? Der Hamster wäscht den Käfer ...	Was ist denn genau los? Der Hamster wäscht den Käfer ...
OS	Was ist mit dem Hamster? Den Hamster wäscht der Käfer ...	Was ist denn genau los? Den Hamster wäscht der Käfer ...



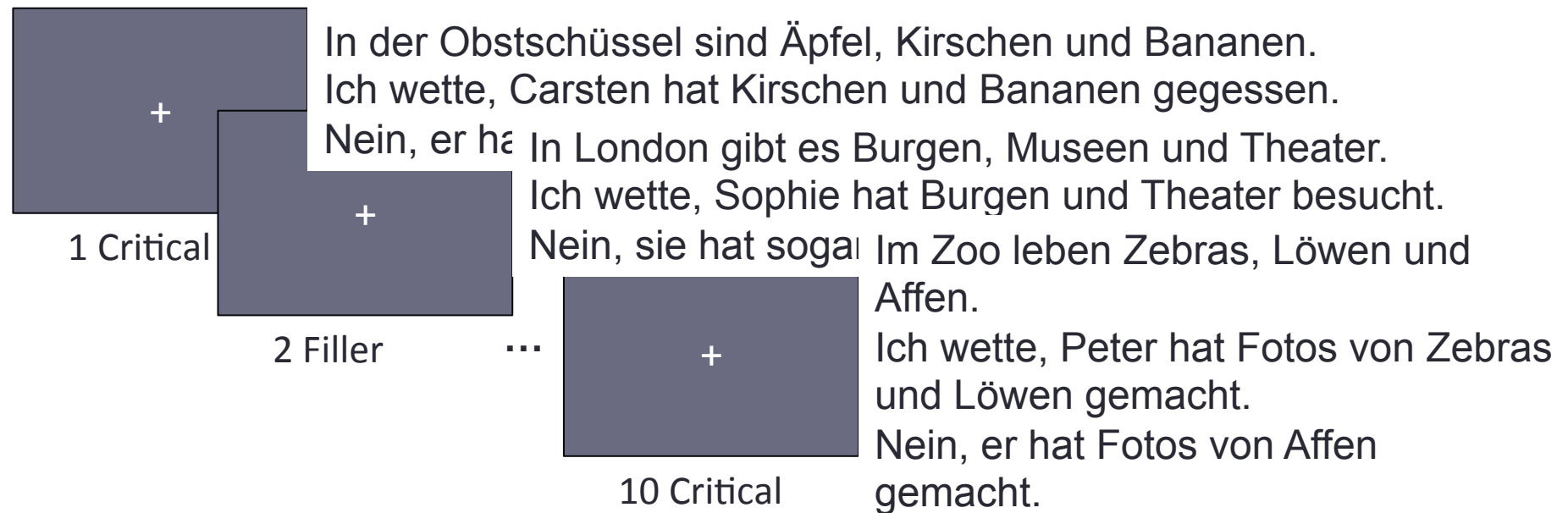
Noch ein “silbernes” Beispiel





Was macht man, wenn eine Itemwiederholung nicht gut ist?

- Frage: Verbessert die Anwesenheit einer Fokuspartikel die Erinnerung an Alternativen?

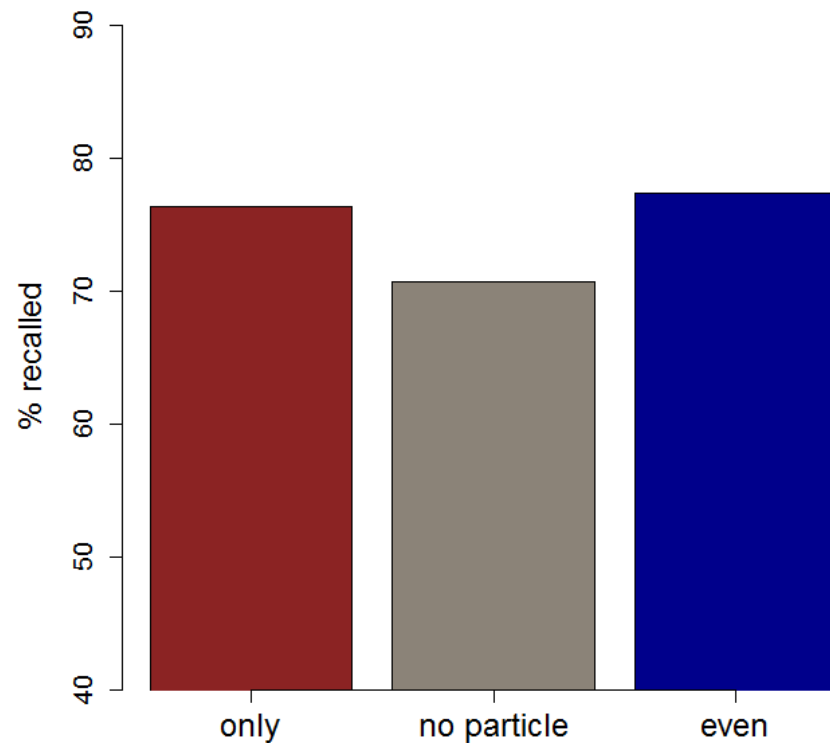


Lateinisches Kreuz

- Design:
- AV: Anzahl erinnerter Alternativen.
- UV: Fokuspartikel (Ausprägungen: keine, “nur”, “sogar”)
- Problem: Wenn within-participants, within-items und vollständiges Kreuzen der Goldstandard ist, dann hört jede Versuchsperson jeden Dialog 3x, mit einem winzigen Unterschied im dritten Satz -> Übersättigung, Deckeneffekt ...
- Lösung: lateinisches Kreuz

Lateinisches Kreuz

	VP1 (= Liste 1)	VP2 (= Liste 2)	VP3 (= Liste 3)
nur	Obtschüssel	Instrumente	Zoo
		Obtschüssel	Instrumente
		Zoo	Obtschüssel



Noch ein paar Bemerkungen zum “sauberen Design”












- Im Katalog sind Hemden, Hosen und Jacken. Ich wette, Matthias hat sich Hemden und Hosen gekauft. Nein, er hat sich _/nur/sogar Jacken gekauft.
- In der Obstschüssel liegen Pfirsiche, Kirschen und Bananen. Ich wette, Carsten hat Kirschen und Bananen gegessen. Nein, er hat _/nur/ sogar Pfirsiche gegessen.
- Im Getränkemarkt gibt es Wasser, Cola und Saft. Ich wette, Angelika hat Saft und Wasser gekauft. Nein, sie hat _/nur/sogar Cola gekauft.





Baselines

- Man benötigt etwas, womit man vergleichen kann. Meist vergleicht man die Bedingungen untereinander, manchmal ist jedoch eine zusätzliche Baseline nützlich.

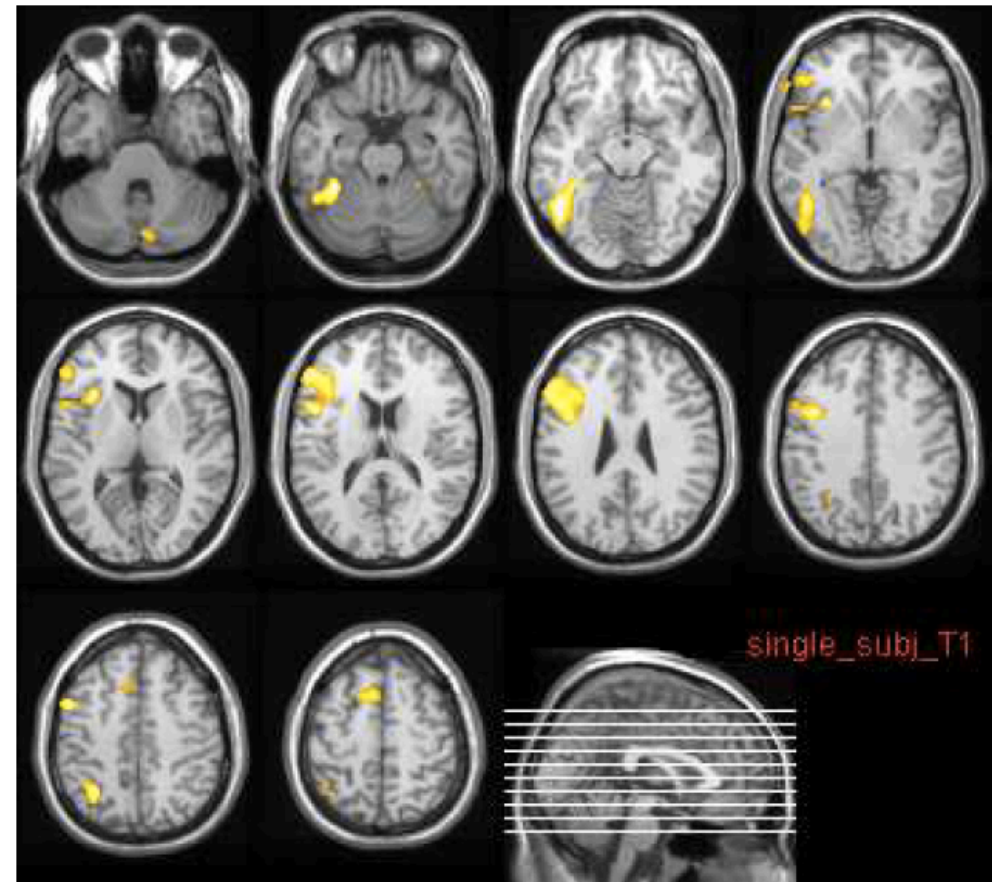
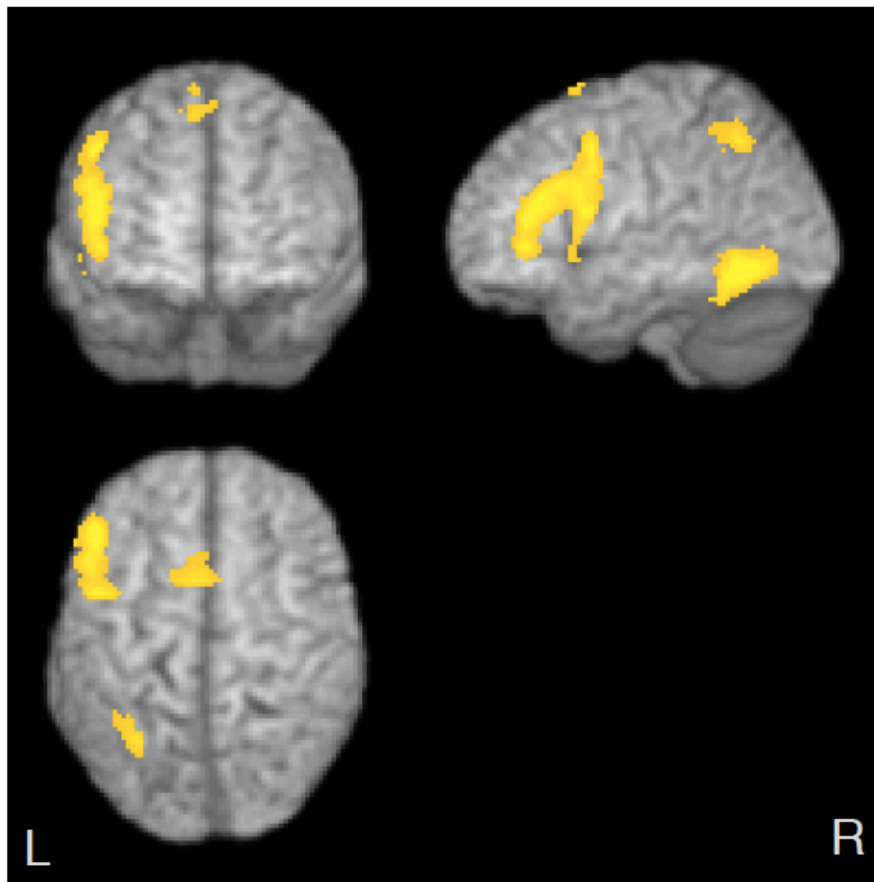
Baselines, Bsp. 1

- Frage: Wie und wo zeigt sich der semantische Interferenzeffekt bei der Sprachproduktion im Gehirn?

Task	Object Naming	Color Naming	Baseline
Condition	same  apple	same  elephant	yes  Idoahs
	related  peach	related  rhinoceros	no  oally
	unrelated  table	unrelated  jar	
	catch trial  sslardt	catch trial  skmayu	catch trial  #####

Trial-type		Response
object (unrelated)	*  table	“apple”
color (related)	*  peach	“red”
catch trial	*  ssiardt	--
baseline	*  ldoahs	“yes”

Kontrast [alle experimentellen Durchgänge] – Baseline ergab (nur) bedeutungsverarbeitende Hirnareale -> Vorteil für darauf folgende Analysen der experimentellen Faktoren (Aufgabe [Objekt vs. Farbe] * Verwandtschaft [identisch, semantisch, nicht verwandt]).



Baselines, Bsp. 2

- Frage: Wird bei L2-Sprechern in der L2-Produktion das Genus der Muttersprache aktiviert?



Baselines, Bsp. 2

	Noun phrase
<i>Cognates</i>	
Gender-compatible	1178 (177)
Gender-incompatible	1261 (281)
<i>Non-cognates</i>	
Gender-compatible	1193 (178)
Gender-incompatible	1281 (203)

Baselines, Bsp. 2

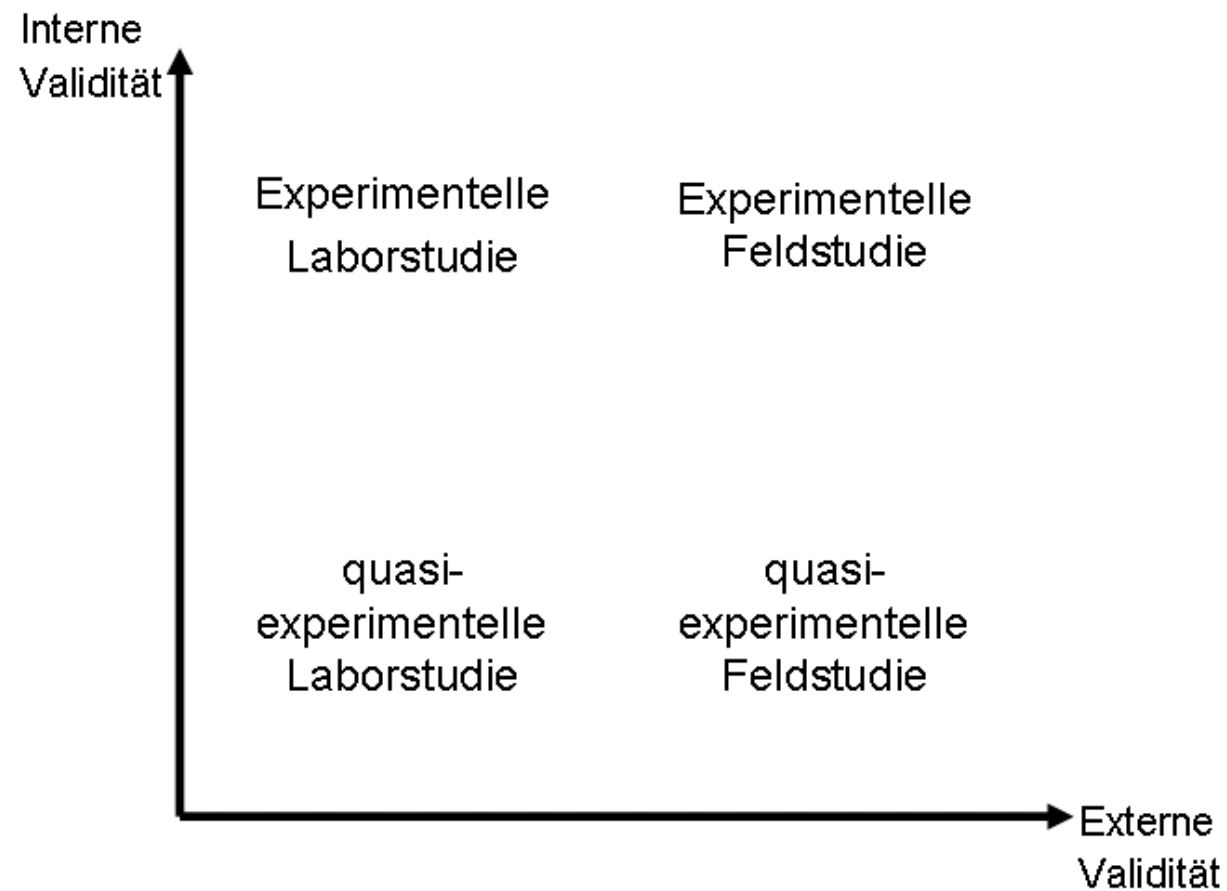
	RTs		
	Noun phrase	Bare noun	Phrase Type effect ^a
<i>Cognates</i>			
Gender-compatible	1178 (177)	1002 (115)	176* [57]
Gender-incompatible	1261 (281)	1044 (126)	217* [121]
<i>Non-cognates</i>			
Gender-compatible	1193 (178)	1091 (130)	102* [64]
Gender-incompatible	1281 (203)	1089 (126)	192* [82]

Fazit

- Für ein aussagekräftiges, interpretierbares und “sauberes” Experiment muss man
 - Störvariablen eliminieren oder konstant halten
 - am besten within-participants und within-items testen
 - balanzieren
 - vollständig kreuzen
 - wo nötig, Baselines verwenden

Fazit

- Aber wird das dann nicht völlig unnatürlich?



Language and Cognition (2015), Page 1 of 28. doi:10.1017/langcog.2015.12
© UK Cognitive Linguistics Association, 2015

Converging evidence for the relevance of alternative
sets: data from NPs with focus sensitive particles in
German

KATHARINA SPALEK*
Humboldt-Universität zu Berlin

AND

AMIR ZELDES*
Georgetown University