

# Experimental design

Holger Diessel  
holger.diessel@uni-jena.de

# Data types

---

- Nominal data
- Ordinal data
- Interval data

Different data types in linguistic research:

- Nominal data: case, gender, types of RCs
- Ordinal data: ???
- Interval data: length of utterance (measured in msc)

Categorical data: Any variable including a meaningful category  
Categorical data can be nominal or ordinal

# Data types

---

Quantitative data can be continuous or discrete:

- Continuous: time, age
- Discrete: words per sentence, phonemes per word

In theory there is a clear distinction between continuous and discrete data, but in practice the distinction is fuzzy. We often use discrete measures for continuous scales (e.g. in years).

We use statistical measures for continuous data if the data are really continuous or if the data are discrete but include a large number of values.

# Variables

---

- Independent variable
- Dependent variable

Independent variable = explanatory variable, predictor variable

Dependent variable = response variable

- The independent variable must have at least two levels (= conditions)
- The dependent variable must allow for at least two different types of responses

# Example 1

---

Subjects are given two types of constructions and are asked to decide whether the given sentence is grammatical:

- |     |    |                          |                |
|-----|----|--------------------------|----------------|
| (1) | a. | I gave it him.           | Construction 1 |
|     | b. | I gave the book her.     |                |
|     | c. | ...                      |                |
| (2) | a. | I gave it to him .       | Construction 2 |
|     | b. | I gave the note to her . |                |
|     | c. | ...                      |                |

# Example 1

---

IV (two conditions)	DV (forced choice task)
Construction 1 Construction 2	a. grammatical b. ungrammatical

## Example 2

---

Subjects are asked to complete copular sentences with a relative clause. The predicate nominals of the copular clauses belong to three different semantic types: (1) animate/human (2) inanimate/object (3) place.

- (1) This is the man \_\_\_
- (2) This is the ball \_\_\_
- (3) This is the place \_\_\_

## Example 2

---

Subject's responses can be divided into five different types:

- |     |                     |  |
|-----|---------------------|--|
| (1) | This is the man ... | who talked to Jane.<br>who I met.<br>whom I gave the book.<br>to whom she went.<br>whose cat died. |
|-----|---------------------|--|



## Example 2

---

IV	DV
1. This is the man ___	a. SUBJ relative clause
2. This is the ball ___	b. DO relative clause
3. This is the place ___	c. IO relative clause
	d. OBL relative clause
	e. GEN relative clause

## Example 2

---

IV	DV
<ol style="list-style-type: none"><li>1. This is the man ___</li><li>2. This is the thing ___</li><li>3. This is the place ___</li></ol>	<ol style="list-style-type: none"><li>a. SUBJ relative clause</li><li>b. DO relative clause</li><li>c. IO relative clause</li><li>d. OBL relative clause</li><li>e. GEN relative clause</li></ol>
<ol style="list-style-type: none"><li>1. I saw the man ___</li><li>2. I saw the thing ___</li><li>3. I saw the place ___</li></ol>	

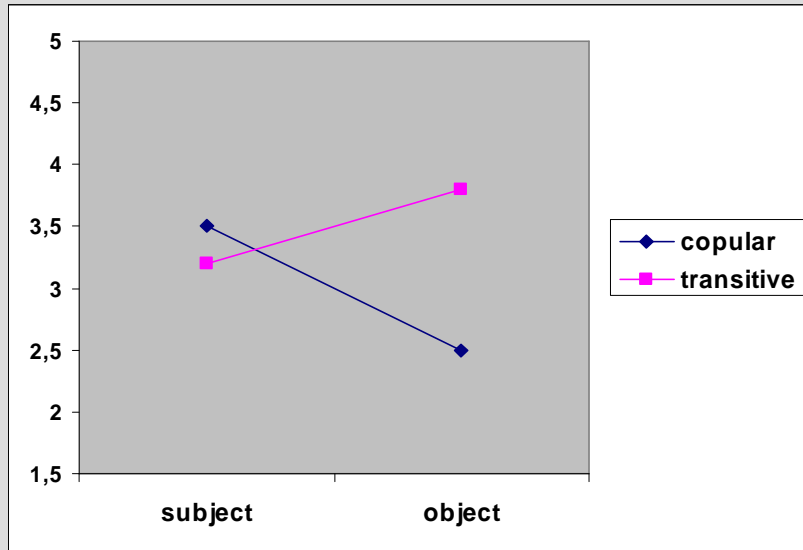
## Example 2

---

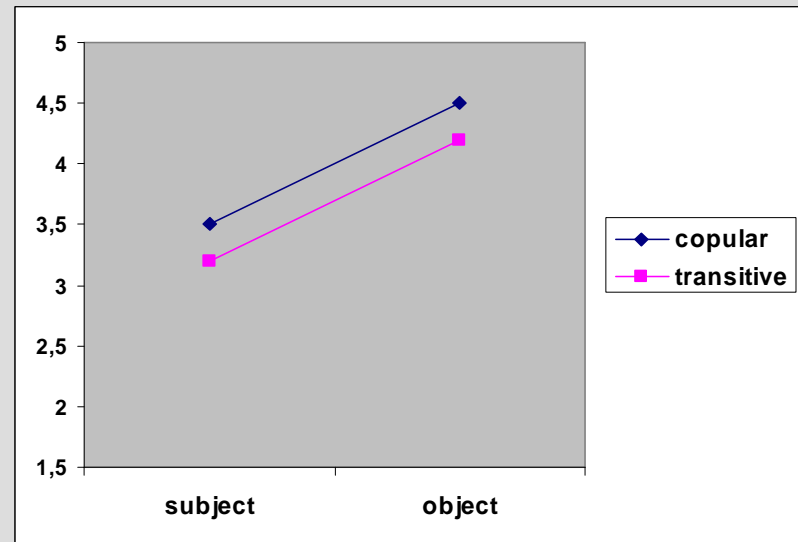
Condition 1 (copular)		Condition 2 (transitive)	
SUBJ	3.5	SUBJ	2.5
DO	3.2	DO	3.8
IO	2.7	IO	3.2
OBL	2.2	OBL	0.5
GEN	0.6	GEN	0.5

# Example 2

---



Interaction



No interaction

# Confounding variable

---

- (1) This is the man who talked to the woman.
- (2) This is the woman who the man talked to.
  
- (3) This is the woman who I talked to.

# Confounding variable

---

Control: Keep the confound constant!

1. Only lexical NPs
2. Equal number of lexical and pronominal NPs in both conditions

# Sampling

---

Sampling frame:

1. First sentence of transcript A
2. Second sentence of transcript A
3. Third sentence of transcript A
4. Forth sentence of transcript A
5. Fifth sentence of transcript A
6. ...
7. 5000th sentence of transcript A

Sample:

200 sentences from transcript A

# Sampling

---

## Sampling techniques

- Simple random sampling
- Stratified random sampling
- Cluster sampling



# Simple random sampling

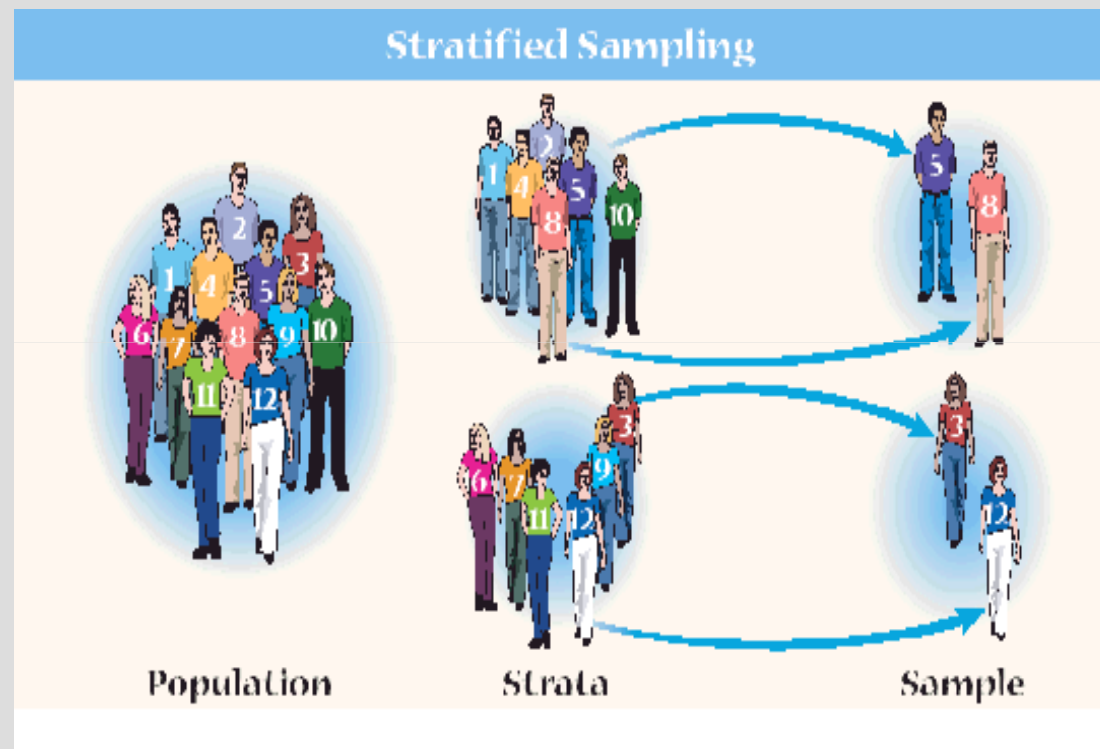
---



Is used when nothing is known about the sample.

# Stratified random sampling

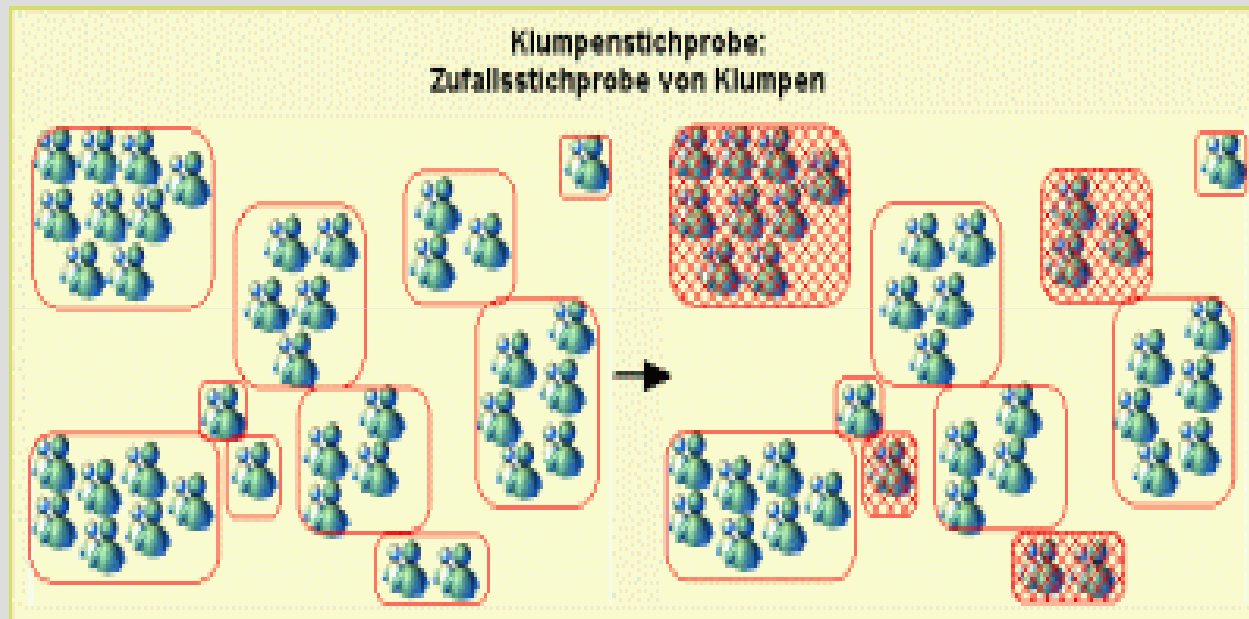
---



Proportional sample vs. disproportional sample

# Cluster sampling

---



# Related and independent design

---

- Within subjects design – related design – repeated measures design
- Between subjects design – unrelated design – independent design

# Related and independent design

---

Advantages of a within subject design:

- Reduction of inter-individual differences
- Fewer subjects

Disadvantages of a within subject design:

- Subjects recognize the purpose of the study.
- Subjects get tired, frustrated, excited.
- Subjects get habituated to the task.

# Differential test and correlational analysis

---

Advantages of a within subject design:

- Correlational analysis (observational statistics)
- Differential test (inferential statistics)

Correlational procedures	Differential tests
Pearson	T-test
Kendell's tau	ANOVA

# Experimental design

---

A child language researcher wants to find out if the acquisition of relative clauses is affected by its syntactic structure. The structure of a relative clause is defined by two features: The syntactic function of the head, i.e. the main clause element that is modified by a relative clause, and the syntactic function of the gap, i.e. the element in the relative clause that is omitted. As can be seen in examples (1) to (4), both head and gap can function as subject or object.

# Experimental design

---

- |   |    |
|---|----|
| (1) Peter saw the man <u>who talked to Sally last night</u> . | OS |
| (2) Jack noticed the man <u>who Sally met yesterday</u> .     | OO |
| (3) The man <u>who talked to Sally last night</u> saw Peter.  | SS |
| (4) The man <u>who Sally met yesterday</u> noticed Jack.      | SO |



# Experimental design

---

IV	DV
1. SUBJ-relative 2. OBJ-relative	Number of correct responses
1. SUBJ head 2. OBJ head	

# Experimental design

---

schematic token set		HEAD	
		animate	inanimate
REL PRO	SUBJ	a (s+s)	b (s+b)
	OBJ	c (b+s)	d (b+b)

token set 1		HEAD	
		subject	object
REL	SUBJ	The man who met Sally talked to me.	I talked to the man who met Sally.
PRO	OBJ	The man who Sally met talked to me.	I talked to the man who Sally met.

token set 2		HEAD	
		subject	object
REL	SUBJ	The girl who saw Jack kissed Bill.	Bill saw the girl who kissed Jack.
PRO	OBJ	The girl who Jack saw kissed Bill.	Bill saw the girl who Jack kissed.

token set 3		HEAD	
		subject	object
REL	SUBJ	The cat that chased the dog scared the horse.	The cat chased the dog that scared the horse.
PRO	OBJ	The cat that the dog chased scared the horse.	The cat chased the dog that the horse scared.

token set 4		HEAD	
		subject	object
REL	SUBJ	The car that hit the bus pumped into the van.	The car hit the bus that pumped into the van.
PRO	OBJ	The car that the bus hit pumped into the van.	The car hit the bus that pumped into the van.

# (=number)	blocks	REL PRO	HEAD	experimental condition	stimulus
1	1	SUBJ	subj	a	The man who met Sally talked to me.
2	1	SUBJ	obj	b	I talked to the man who met Sally.
3	1	OBJ	subj	c	The man who Sally met talked to me.
4	1	OBJ	obj	d	I talked to the man who Sally met.
5	2	SUBJ	subj	a	The girl who saw Jack kissed Bill.
6	2	SUBJ	obj	b	Bill saw the girl who kissed Jack.
7	2	OBJ	subj	c	The girl who Jack saw kissed Bill.
8	2	OBJ	obj	d	Bill saw the girl who Jack kissed.
9	3	SUBJ	subj	a	The cat that chased the dog scared the horse.
10	3	SUBJ	obj	b	The cat chased the dog that scared the horse.
11	3	OBJ	subj	c	The cat that the dog chased scared the horse.
12	3	OBJ	obj	d	The cat chased the dog that the horse scared.
13	4	SUBJ	subj	a	The car that hit the bus pumped into the van.
14	4	SUBJ	obj	b	The car hit the bus that pumped into the van.
15	4	OBJ	subj	c	The car that the bus hit pumped into the van.
16	4	OBJ	obj	d	The car hit the bus that pumped into the van.

# Ordering effects

---

- Random ordering
- Insertion of fillers (= distractors)

Counterbalancing:

ABBA -> AB + BA

Problem: effect is non-linear

Between subject balancing:

Group A: AB

Group B: BA

Problem: effect is non-symmetrical

# Ordering effects

---

Complete counterbalancing:

1. ABC
2. ACB
3. BAC
4. CAB
5. BCA
6. CBA

# (=number)	blocks	REL PRO	HEAD	experimental condition	stimulus	randomizer
2	1	SUBJ	obj	b	I talked to the man who met Sally.	0,307790921
42	1	filler 26	-	-	[...]	0,410340229
4	1	OBJ	obj	d	I talked to the man who Sally met.	0,420995715
22	1	filler 6	-	-	[...]	0,439751365
17	1	filler 1	-	-	[...]	0,46188501
18	1	filler 2	-	-	[...]	0,589339167
21	1	filler 5	-	-	[...]	0,612003047
19	1	filler 3	-	-	[...]	0,705094748
41	1	filler 25	-	-	[...]	0,755874665
20	1	filler 4	-	-	[...]	0,769874943
3	1	OBJ	subj	c	The man who Sally met talked to me.	0,856969895
1	1	SUBJ	subj	a	The man who met Sally talked to me.	0,908738867
28	2	filler 12	-	-	[...]	0,074022153