Foundations of Empirical Methods

Torsten Kai Jachmann
jachmann@coli.uni-saarland.de

http://www.coli.uni-saarland.de/~jachmann/teaching/FEM.html

Fr 8:30–10:00

UNIVERSITÄT
DES
SAARLANDES

About the course

- Once weekly every Friday 8:30 – 10:00 am
- Slides (with comments/annotations) will be uploaded beforehand
- We use Microsoft Teams for the lectures as well as communication and file distribution

While Teams is used as our main platform, you can also write me an email to jachmann@coli.uni-saarland.de

Course requirements

- Group project (starting 15. January 2021)
  - 15% of the final grade
- The final exam is scheduled for Friday 12.02.2021 – 8:30 am
  - On site, Building C7 4, Conference room
  - 85% of the final grade

Your final grade will be evaluated based on two parts:

1) a group project worth 15% of the final grade
2) the (written) final exam worth 85% of the final grade

Overview

- The role of empirical data in cognitive and linguistic research
- Psycholinguistic methods (offline and online)
- Corpus studies
- Experimental design
- Basic statistics
- Interpretation of results

- Why do we need empirical data?

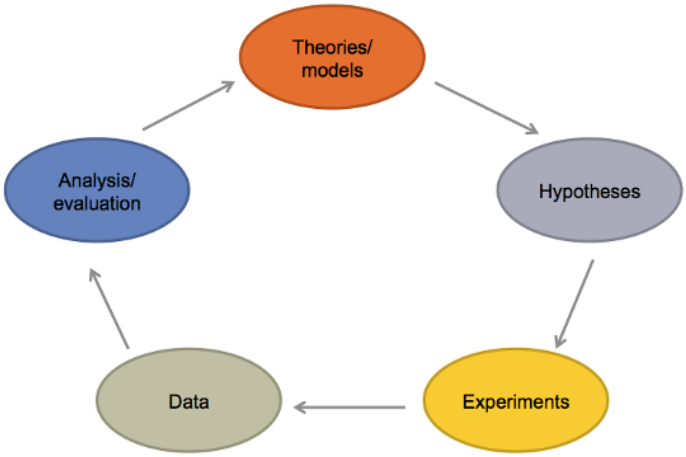
- What methods of data collection exist?
- How do they differ?
- What question can we (potentially) answer with a given method?

- How do we prepare an experiment to answer our research questions?
- What do we have to consider?

- Once we have collected data, how do we analyse them?

- Once we analysed them, how do we interpret the results?
- What do the results tell us?

Research Cycle



- This is the idealized research cycle
- We might start with a general theory or model of some phenomenon. A good theory or model should lead to hypotheses that are empirically testable.
- To test a particular hypothesis, we need to design an appropriate experiment or series of experiments, which will generate some data.
- We then need to analyze the data and see whether it supports the hypothesis or not. In some cases, the results of our experiment may lead us to revise the original theory or model.
- This week we will focus on the Experiment – Analysis part of the cycle

Deductive reasoning starts with a theory/hypothesis and makes predictions about what observations should be found if the theory were correct.
Inductive reasoning makes broad generalizations from specific observations – from many observations, discern a pattern, make a generalization, and infer an explanation or a theory. Scientists use it to form hypotheses to be tested.
Abductive reasoning starts with incomplete set of observations and proceeds to likeliest possible explanation, which often entails making an educated guess. Useful for forming hypotheses to be tested, often used by doctors to make diagnosis based on test results and by jurors to make decisions based on evidence presented.

Observation

- Language provides choices for speakers on how to encode their message
  - Choice of words
  - Choice of syntactic structures
  - Choice of sentence arrangement in discourse

Many questions that we try to answer in science originate from observations that we can make in everyday life.
This is an example of an observation we can make in the utilization of language.

When we talk to people and try to convey a certain massage, we are always encountering a multitude of options:
Imagine you want to talk about the displayed object.

You could choose to say “the ball”, but you could also say “the basket ball” or even “the round thing”.

If you choose the latter, you could also instead say “the thing that is round”.

Now, we could expand on this ”Do you see the basket ball? I was playing with it yesterday.”

Question

- How do we decide on a specific encoding?
  - Why do we choose a certain word?
  - What leads us to utter one syntactic structure over another?
  - How do we construct and expand the discurse?

When we encounter certain phenomena in language, we can ask question about them:

What led us to say ”basket ball” rather than just “ball”, or the otherway around?
Why would we say ”the thing that is round”, rather than ”the round thig”?
Are those choices random, or can we pinpoint some sort of systematicity?

Theory

- The goal of communication is to convey a certain message
  - The speaker wants to express the message
  - The speaker wants to be understood by the listener

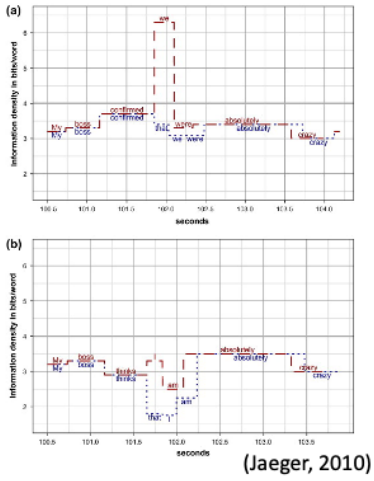
➢ Speakers choose the encoding that optimizes those two goals

If we consider that the primary goal of communication is to convey a certain message, we could argue that the choice of encoding should be in line with this goal. We could thus consider some subgoals:

1) The speaker wants to express the intended message, possibly in a way such that it is as effortles as possible while being complete
2) The speaker wants to be understood by the listener

Assumption

- Communication through a noisy channel has a limited bandwidth (Shannon, 1948)
- Information transfer is optimized by distributing information uniformly close to that channels capacity (Genzel & Charniak, 2002)
- Speakers distribute information uniformly across the signal without under- or overutilizing the channel (Jaeger, 2010)



(Jaeger, 2010)

Oftentimes certain assumptions are (/have to be) made under a specific model (or theory)
Such assumptions can also be taken from previous work and the literature to extent ones theory

For example, the Uniform Information Density (UID) hypothesis by Jaeger (2010) states that speakers aim to utilize the communicative (noisy) channel in a way such that their produced signal (speech) does not contain sudden strong peaks or plunges in the signal. Thus, speakers aim to distribute the information equalkly across the signal.
As you see, this theory utilizes Shannon's noisy channel (1948) as well as theories formulated by Genzel & Charniak (2002)

The example sentences can be understood as follows:
The verb "confirm" in example (a) can either take a direkt object ("My boss confirmed the order") or introduce a sub-clause. The following word then provides information about the specific syntac tic structure we actually encountered. While "that" is optional in the sentence, it is adding an extra element to the signal that leads to the same amount of information being distributed across 2 elements ("that we") rathewr

than a single element (“we”).
The verb “think” in example (b) on the other hand is usually introducing a sub-clause. Thus, the information conveyed by “I” is lower as of the “we” in the previous example. In this case, adding “that” is distributiong this lower amount of information across two words, which leads to a plunge in the distribution of informativity.

Theory to Hypothesis

- Theories have to be tested in order to
  a. confirm them
  b. refute them
- More concrete hypotheses are formulated to gather evidence for a theory

Models and theories need to be tested in order to be able to confirm their validity. As such, it is necessary to formulate more concrete hypotheses that can be tested. Such hypotheses are usually formulated in a way such that a confirmation of any given hypothesis would be in support of the theory or model that we want to confirm.

Uniform Information Density Hypothesis

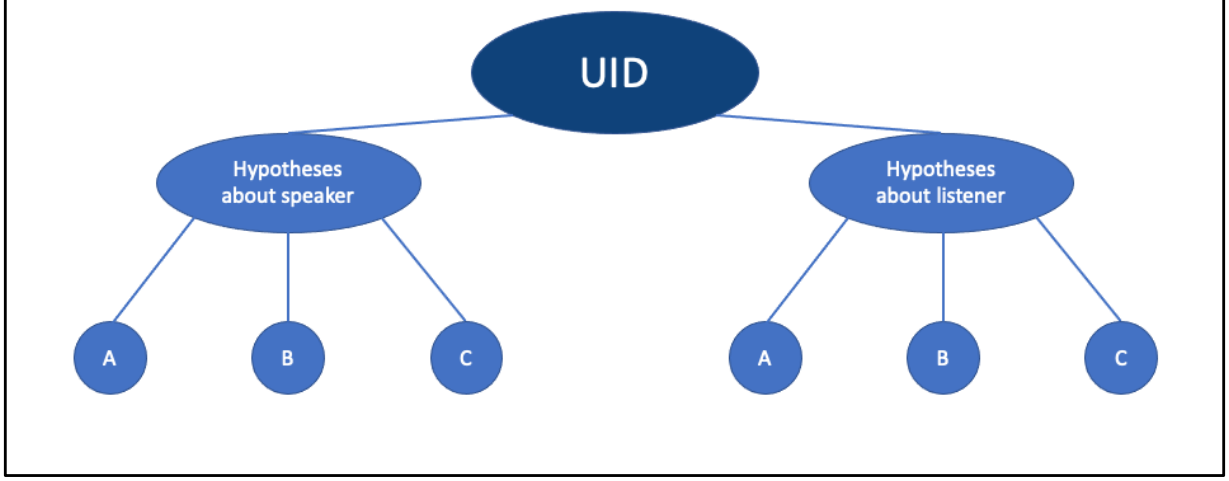
Within the bounds defined by grammar, speakers prefer utterances that distribute information uniformly across the signal (information density). Where speakers have a choice between several variants to encode their message, they prefer the variant with more uniform information density (ceteris paribus).

(Jaeger, 2010)

This is the UID hypothesis as formulated by Jaeger (2010).

As you can see, this hypothesis is somewhat broad in its formulation.
In order to properly test a hypothesis, evidences in support of it (or against it) have to be collected.
As such, it is usual to formulate sub-hypotheses that can be approached by running experiments.

Hypothesis to Hypotheses



Starting from the main hypothesis, we can derive sub-hypotheses which, again, could lead to their own sub-hypotheses.

Considering the UID hypothesis, one can come up with hypotheses concerning the speakers behaviour that would be required if the UID was in fact correct.
At the same time, the ideas behind UID also carry some level of audience design. Thus, listeners should also show certain reactions when exposed to certain structures.
(E.g., is it really harder for listeners to hear unevenly distributed information?)

In order to test the more fine grained hypothesis, experiments have to be designed that can potentially answer the underlying research questions.

When a theory is on a large scale and offers various hypotheses and grounds for testing, a project can be introduced in which multiple researchers cooperate to answer the same question from different angles.

Hypotheses to Experiment

- Formulation of research questions
- Creating an experiment designed to answer that question

Speakers distribute information uniformly across the signal without under- or overutilizing the channel

➢ Do speakers adjust their behaviour according the content of information… ?

When testing for the validity of a given hypothesis, one should formulate research question which we then aim to answer with an experiment designed exactly for this purpose.

If our hypothesis is that speakers distribute information uniformly across the signal without under- or overutilizing the channel, we can reformulate this as a question such as “do speakers adjust their behaviour in order not to over- or underutilize the channel?”

While this seems trivial, it isn’t always. Especially if you consider what you need to show in order to be able to answer that question with yes.
Ideally, you can not only ask the question, but also think of a way to find proof for the answer.

Experimental design

- Aiming to answer specific research questions
- Choose from a variety of experiment types tailored to the question
  - ”Offline” Studies
    - Self-Paced Reading (SPR)
    - …
  - “Online” Studies
    - Event-Related Potentials (ERP)
    - …

An experiment is usually designed to answer specific research questions.

There is a variety of experiment types that can be chosen. Thus, it is important to select the right tool for the right task.

Next week we will talk about the basic structure, design and problems of experiments and their creation.

Next week

- Experimental design
- In preparation, participate in this 10 minute example experiment

https://expt.pcibex.net/ibexexps/tokaja/FEM-Experiment/experiment.html

- Take note of
  - perceived oddities
  - your thoughts on what this experiment could be about

Next week we will talk about the basic structure, design and problems of experiments and their creation.

Please follow the provided link and participate in the example experiment.
Note down anything you find odd and think about what could be tested in this experiment.