How (not) to do an experiment: Evidence from a self-paced reading study

Chiara Tschirner

Potsdam University

November 29, 2019

Table of Contents

Finding a research topic

The process of finding a research topic for an experiment is similar to other research projects.
However, the following issues need to kept in mind to avoid an over-ambitious project:

- Each stage of the experiment takes time
- What can be done to shorten time is to add onto existing experiments, instead of creating one from scratch
- Everything depends on the method, especially the time frame, but also anything between design and analysis

Pre-registration (aspredicted.org) is a good idea because it ensures that the plan is sound and feasible, but important details are not over-looked.

Background

- In basic terms: verbs have argument structure and the brain has to figure this structure out during processing
- Languages have different ways of showing what the arguments are, for example case in German
- Another way is through animacy, which has an effect in German even though it is not marked specifically
- Frisch & Schlesewsky (2001): contrast in animacy important for processing
- Czypionka et al. (2017): object animacy effects

→ what about subject animacy and contrast?

The idea

Proposal

A self-paced reading experiment answering the question whether a contrast affects processing and whether it is independent of subject type

- Self-paced reading because it allows an insight into processing (without the hassle of an EEG)
- Contrast seems to be important, but it might not be just any contrast, but the specific constellation of animate subject with an inanimate object
- For this reason, all four combinations of animacy with subject and object have to be tested (see (1) for visualisation)

Designing the experiment

The basic idea of an experiment is to measure the effects of a certain manipulation compared to a baseline.

- Minimal pairs are desirable because they show where exactly the effect comes from
- Several factors play into the choice of type and number of items, such as the method and the participants
- It is a good idea to check for length and frequency effects
- Other materials such as fillers and additional tasks need to be created as well
- Filler items are important!

Materials

The two conditions and their levels are

- Contrast: Contrast/No contrast
- Subject Type: Animate Subject/Inanimate Subject

These conditions are compared in a set of four sentences with the same structure:

- Matrix clause, Subject Object Adverb Noun
- The matrix clause gives context for the rest of the clause
- All nouns in the subordinate clause are masculine. This is to avoid ambuiguity in case marking
- Each set is based on a verb, each verb is used twice, but with different nouns

Example Material

(1)
a. Anna sagt, dass der Kritiker einen jungen Schauspieler besonders empfiehlt. → **animate subject/no contrast**
b. Anna sagt, dass der Kritiker einen neuen Gedichtband besonders empfiehlt. → **animate subject/contrast**
c. Anna sagt, dass der Artikel einen jungen Schauspieler besonders empfiehlt. → **inanimate subject/contrast**
d. Anna sagt, dass der Artikel einen neuen Gedichtband besonders empfiehlt. → **inanimate subject/no contrast**

Materials

- 160 sentences, each participant reads 40 - Latin Square Design
- 80 filler sentences, with a similar structure to the critical items, meant to sound natural
- A yes/no question for each item (both filler and critical) - 50:50 y/n answers

Collecting the data

Issues to keep in mind:

- Finding participants for an experiment can be difficult and time-consuming, especially when the criteria are quite specific
- Some experiments can be web-based which makes it easier to reach participants. This applies to a lesser extent to mobile experiments.
- (Paying participants helps)
- Presenting the items also needs to be considered. Free programmes such as OpenSesame or Linger are available.

Participants & Procedure

- 41 participants, 8 male; age: mean: 24,4 (sd=5,1)
- Before the experiment, participants signed a consent form giving them information about the use of their data
- The experimental method was a non-cumulative word-by-word self-paced reading experiment
- The experiment took place on a desktop computer in the VLAB at Potsdam University
- Instructions were given on the screen and 5 trial sentences preceded the experiment, which took about 20 min with breaks

Analysing the data

- Determine the analysis that is planned, including exclusions
- The kind of analysis depends on the type of data
- Possible tools for analysis are SPSS and the free and powerful software R (though that includes an ongoing learning process)
- Descriptive statistics before inferential statistics!

Analysis

- Reading times were measured for every word
- The regions of interest are the object noun (critical region), adverb (postcritical region) and verb (spillover region)
- Data points with a reading time below 150ms and above 3000ms were excluded
- Accuracy measured for questions, exclusion of participants with an accuracy of below 80%
- Linear mixed models performed on each region of interest (then compared to a null model using an ANOVA)
- Fixed effects: Contrast, Subject Type
- Random effects: Participant, Item
- Reading times are log-transformed to ensure a normal distribution of the data

Descriptive statistics

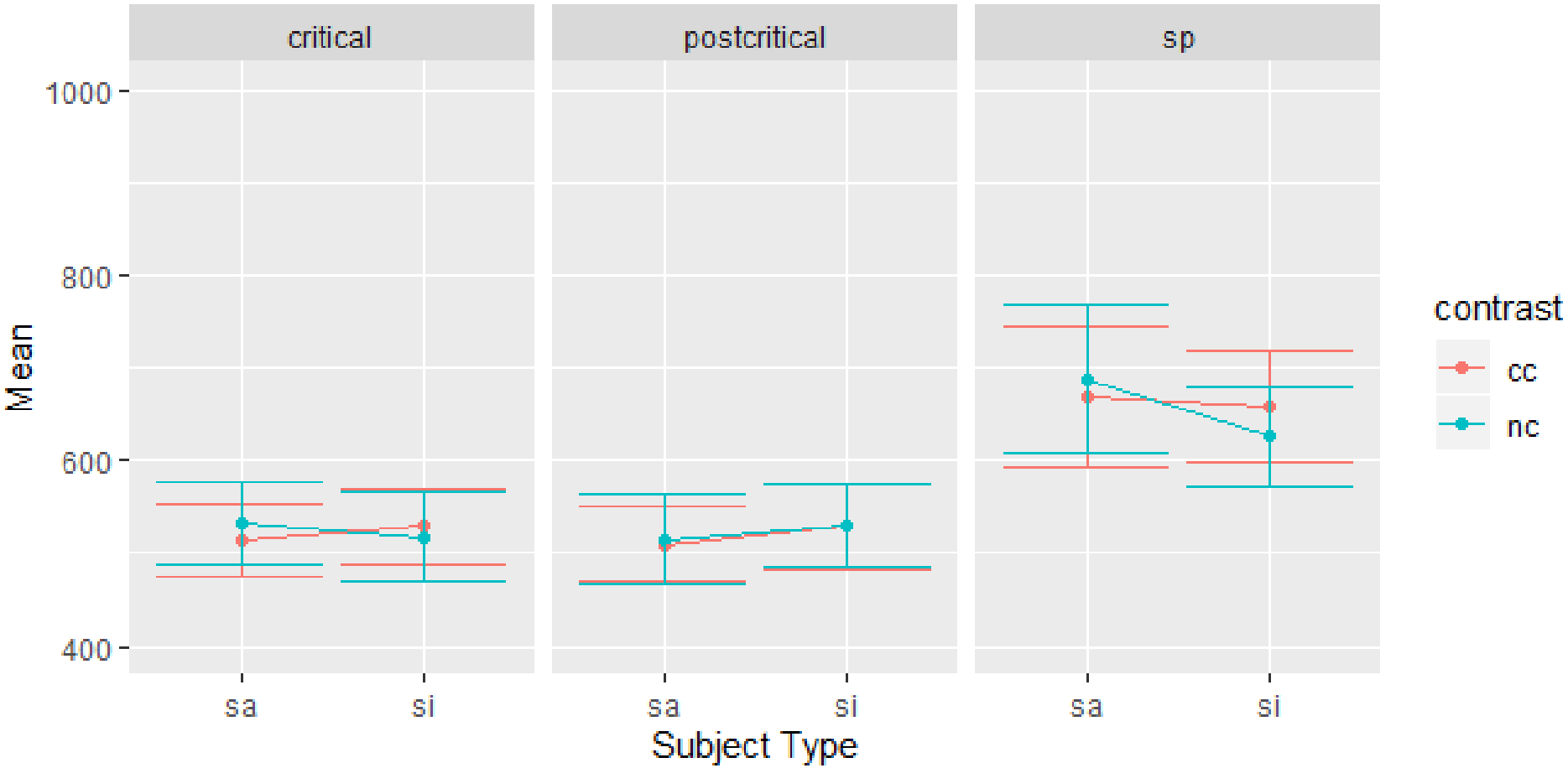

Figure 1: Means of RTs for Subject Type and Contrast

Inferential statistics

The following are the main results from the ANOVAs comparing the linear mixed models to a null model without the effect in question:

- No main effect for contrast in any region
- Significant main effect of subject type in postcritical region ($\chi^2 = 6.44$, p= 0.01)
- Significant interaction between contrast and subject type in critical region ($\chi^2 = 6.34$, p= 0.01)

Giving the results meaning

- The results are just evidence for or against a claim
- Explain what they mean and how they fit into the wider context
- If there are no significant results, explain what went wrong and how the experiment could be improved

Preliminary discussion

- The lack of a significant main effect for contrast implies that contrast is not the most important factor in processing
- The significant main effect for subject type in the postcritical region is evidence for its importance
- The interaction between contrast and subject type in the critical region is evidence for the claim that the contrast does affect processing, but is not independent of subject type

... this is a work in progress.

Take-home message

- Experiments take a long time
- Things go wrong
- But a lot of problems can be avoided by planning the steps beforehand

Thank you for your attention!

References

Czypionka, Anna, Katharina Spalek, Isabell Wartenburger & Manfred Grifka. 2017. On the interplay of object animacy and verb type during sentence comprehension in German: ERP evidence from the processing of transitive dative and accusative constructions. *Linguistics* 55(6). 1383–1433.

Frisch, Stefan & Matthias Schlesewsky. 2001. The N400 reflects problems of thematic hierarchizing. *Neuroreport* 12(15). 3391–3394.