



**SFB  
1287**

**Limits of Variability in Language**  
Cognitive, Grammatical, and Social Aspects

# DATA MANAGEMENT PLAN

**CRC 1287 | 2021 | Phase 1**  
**Project B01**

.....

## TABLE OF CONTENTS

<b>General Information .....</b>	<b>3</b>
Overview .....	3
Data Management Requirements .....	5
Financial Support .....	5
<b>Dataset Information .....</b>	<b>6</b>
Data Origin .....	6
Data Collection .....	6
Data Handling.....	7
Data Analysis .....	8
Data Reuse .....	9
<b>Legal and Ethics .....</b>	<b>10</b>
Personal Data.....	10
Property Rights.....	10
<b>Publication .....</b>	<b>11</b>
Publishing or Sharing Data .....	11
<b>Storage and Long-Term Preservation .....</b>	<b>12</b>
Archive .....	12

.....

## GENERAL INFORMATION

### Overview

#### **Project number**

B01

#### **Name of Experiment / Acronym / Number**

WP2 - Satz-Bild Studie 3

#### **Involved persons**

Huttenlauch, de Beer, Schneider

#### **PI or responsible person (head of the study)**

Isabell Wartenburger, Sandra Hanne

#### **Subject area**

- Production of locally ambiguous sentences
- Perception of locally ambiguous sentences

#### **Method / Type of data**

- Elicited productions (Experiment 3): wav files plus information about speakers, accompanying data: working memory (WM) measures of participants, hearing screening
- Perception data (2AFC, EyeTracking): response accuracy, reaction times, eye movements and pupil size, accompanying data: WM measures, hearing screening

## **Participants (of the study)**

Young adults (students, 18 - 35 years old)

## **Short description (of the study)**

Production of locally ambiguous sentences in two conditions (SVO "Das Mädchen grüßt nun den Opa" vs. OVS "Das Mädchen grüßt nun der Opa") to a virtual interlocutor in 2 contexts.

Contexts consist of two different female interlocutors: a young adult (C1 adult), an elderly adult (C2 elderly).

Perception of locally ambiguous sentences in SVO and OVS condition: new recordings in three prosody conditions (natural, exaggerated, neutral).

2AFC: perception of prosodic contours of SVO and OVS in the natural and exaggerated prosodic condition up to the disambiguating point of the sentence ("Das Mädchen grüßt nun...").

EyeTracking: perception of prosodic contours of SVO and OVS in all three prosodic conditions (natural, exaggerated, and neutral), sentence picture matching task.

## **Comments (optional)**

- Experiment 3: Produktion
- Perception: 2AFC (Master Thesis Kathleen Schneider)
- Perception: EyeTracking (Master Thesis Kathleen Schneider)
- WP2 (locally ambiguous sentences, young adults)
- Experiment 3: n=80 productions per participant
- 2AFC: n=336 items (21 verbs \* 2 word order conditions \* 2 prosodic conditions \* 4 recorded versions)
- EyeTracking: n=120 items (20 verbs \* 2 word order conditions \* 3 prosodic conditions) (+ 6 practice items)

## **Data Management Requirements**

**Are there requirements regarding the data management from your scholarly / scientific community?**

yes

**If yes, what are the requirements?**

- DFG Guidelines on the Handling of Research Data
- Handlungsempfehlungen zum Umgang mit Forschungsdaten University of Potsdam
- Technische und organisatorische Maßnahmen (TOM) gemäß Art. 32 Abs. 1 DSGVO

## **Financial Support**

**Who is funding the project?**

DFG - Deutsche Forschungsgemeinschaft e.V. (German Research Foundation)  
- <https://www.dfg.de/en/>

**In which funding line and / or which funding program is the project funded?**

Collaborative Research Centre 1287 - Project number 317633516

.....

## DATASET INFORMATION

### Data Origin

**Is the dataset being created or re-used?**

created

**If re-used, who created the dataset and under which address, PID or URL is the data set available?**

-

### Data Collection

**When does data collection start? (approximately / tentatively)**

- Experiment 3: 2019
- 2AFC / EyeTracking: 2021

**When does data collection end? (approximately / tentatively)**

- Experiment 3: 2019
- 2AFC / EyeTracking: 2021

.....

## **Data Handling**

### **Where is the dataset stored during the project?**

- CRC file server
- Box.UP university cloud
- researcher's computer
- computer in the laboratory

### **If data is stored on lab or personal computers, please describe the backup strategy.**

All data (raw and finally analyzed data) are stored in owncloud . Raw data is backed up on the CRC file Server. After finalizing data collection and analyses, the analyzed data will also be stored on the CRC file Server.

### **Which file formats are used?**

- Production data (raw and processed): \*.wav files, praat TextGrids, pitch objects
- Perception data: CSV and TXT files
- Presentation folder including scrip t, stimuli (videos of interlocutors, 2 videos per context), logfiles
- visual stimuli
- Instructions for participants
- Description of procedure
- Study protocol
- Participant information
- Covid 19 Agreement
- Empathy Quotient scores
- Scores from working memory tests

- Hearing screening
- Audio recordings: Recordings took place in a sound attenuated booth in the acoustics laboratory of the University of Potsdam via an Alesis io12 interface. Speakers wore a headset HSC 271 (AKG Acoustics by Harman, [www.akg.com](http://www.akg.com)) with over ear headphones and a condenser microphone and were seated in front of a wide screen monitor with 1920 x 1200 resolution. Picture and written stimuli were presented on Screen. The experiment was run from a Dell laptop standing outside of the recording booth using the software Presentation (Neurobehavioural Systems <https://www.neurobs.com/>; Version 20.1)

### **Which measures of quality assurance are taken for this dataset?**

Data will be analyzed for outliers and biases in the responses (by the person who analyzes the data).

- wav files are annotated.
- listening checks are implemented in the perception experiments.
- triggers as well as validity measures are implemented for

### **Data Analysis**

#### **When does data analysis start? (approximately / tentatively)**

- Experiment 3: 2019-2020
- 2AFC: 2021
- EyeTracking: 2021-2022

#### **When does data analysis end? (approximately / tentatively)**

- Experiment 3: 2019-2020
- 2AFC: 2021
- EyeTracking: 2021-2022



.....

## **Data Reuse**

**Which individuals, groups or institutions could be interested in re-using this dataset? What consequences does the reuse potential have for the provision of the data later?**

Non anonymizable data:

- Raw data (audio recordings from participants) are non anonymizable.
- Video stimuli and audio stimuli of experimental interlocutors (1 interlocutor per context): 2 videos + 1 audio file per context (2 contexts)
- Reuse of raw data and stimuli by other researchers/scientists is NOT possible because the identity of the participants/recorded person could be revealed (personal data). Raw data and stimuli can not be shared.
- Eye movements and pupil size are non anonymizable.

Anonymizable data:

- Extracted measures (duration, f0) could be shared.
- Response accuracy and reaction times could be shared.
- Test scores on working memory assessments (digit span, block span, digit symbol substitution, Mehrfachwahlwortschatztest), hearing test (single sound audiogram), Empathy Quotient, handedness (Edinburgh Handedness Inventory)

.....

.....

## LEGAL AND ETHICS

### Personal Data

**Does this dataset contain personal data?**

yes

**If yes, are these data anonymised?**

yes

### Property Rights

**Does the project use and/or produce data that is protected by intellectual or industrial property rights?**

yes

**If yes, please explain which data protected by intellectual or industrial property rights?**

Software Presentation (Neurobehavioural Systems <https://www.neurobs.com/>;  
Version 20.1)

.....

.....

## PUBLICATION

### **Publishing or Sharing Data**

**Will this dataset be published or shared?**

yes

**If yes, the principal investigator of the study ensured that the consent form / subject information sheets support publishing of the data?**

yes

**If yes, under which terms of use or license will the dataset be published or shared?**

CC0 (Public Domain)

**If yes, when will the data be published?**

when the paper is published

**If no, please explain why not. Please differentiate between legal and contractual reasons and voluntary restrictions.**

-

.....

## STORAGE AND LONG-TERM PRESERVATION

### Archive

**Does this dataset have to be preserved for long-term?**

yes

**How long does the data need to be stored?**

The DFG expects primary data that is the basis of a publication to be stored in the researcher's own institution or an appropriate nationwide infrastructure long-term (for at least 10 years).

**What are the reasons this dataset must be preserved for the long-term?**

- Use in a publication / Evidence of good scientific practice
- Reuse (if anonymizable data) in subsequent projects or by others
- Legal obligations
- Documentation because it is socially relevant
- Self-commitment
- Evidence of good scientific practice
- DFG requirements

.....

**Where will the data (including metadata, documentation, and relevant code) be stored or archived after the end of the project?**

- CRC 1287 File-Server
- OSF
- Research Data Server from Project IN-FDM-BB (a.t.m. not available)

.....

coordinated by:



Universität Potsdam

<https://www.uni-potsdam.de>

funded by:

**DFG** Deutsche  
Forschungsgemeinschaft

<https://www.dfg.de/>

in cooperation with:

**RUHR  
UNIVERSITÄT  
BOCHUM**

**RUB**

<https://www.ruhr-uni-bochum.de>