



**SFB  
1287**

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

**DATA MANAGEMENT PLAN**  
**SFB 1287 / 2025 / PHASE 2**

**PROJECT C02**

Table of contents

General Information ..... 1

    Overview ..... 1

    Data Management Requirements ..... 1

    Financial Support ..... 2

Dataset Information ..... 3

    Data Origin ..... 3

    Data Collection ..... 3

    Data Handling ..... 3

    Data Analysis ..... 4

    Data Reuse ..... 4

Legal and Ethics ..... 5

    Personal Data ..... 5

    Property Rights ..... 5

Publication ..... 6

    Publishing or Sharing Data ..... 6

Storage and Long-Term Preservation ..... 7

    Archive ..... 7

## General Information

---

### Overview

Project number

*C02*

Name of Experiment / Acronym / Number

*WPB 1 – Bare Nouns*

Involved persons

*Giuliano Armenante, Malte Zimmermann*

PI or responsible person (head of the study)

*Prof. Dr. Malte Zimmermann*

Subject area

*Definiteness*

Method / Type of data

*Fieldwork data, experimental data*

Participants (of the study)

*Language consultants, experiment participants from students of University of Potsdam*

Short description (of the study)

*The studies in this project investigate the conditions licensing article drop for single count nouns, with focus on goal-achieving contexts.*

Comments (optional)

-

## 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*
- *Data Management in Psychological Science*
- *„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 317633480*

## Dataset Information

---

### Data Origin

Is the dataset being created or re-used?

*Fieldwork data and experiment data involve new data, collected through fieldwork/online experiment. The BN\_Ger file involves re-used data from corpus (as specifically indicated)*

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

*Data extracted from the German corpus DeTenTen10 through <https://www.sketchengine.eu/>*

### Data Collection

When does data collection start? (approximately / tentatively)

*September 2023*

When does data collection end? (approximately / tentatively)

*May 2025 (fort the Boris Becker online study)*

### Data Handling

Where is the dataset stored during the project?

*SFB 1287 File-Server, Box.UP, Lab computer*

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

*University server, external hard drives*

Which file formats are used?

*xlsx, pdf, text, md, R, png, js, css, html, wav,*

Which measures of quality assurance are taken for this dataset?

*Sentences for fieldwork sentences were constructed with the help of native speakers. Inconclusive judgments were double checked with an additional speaker. Empirical findings were presented in several international conferences*

*The design of the study "Boris Becker" was refined through multiple presentations in lab meetings and conferences. The items were proofread by a native speaker. Empirical findings were reviewed and discussed in lab meetings.*

## Data Analysis

When does data analysis start? (approximately / tentatively)

*November 2024*

When does data analysis end? (approximately / tentatively)

*May 2025*

## 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?

*Linguists, language teachers*

## Legal and Ethics

---

### Personal Data

Does this dataset contain personal data?

*The empirical study Boris Becker contains metadata of the participants as well as unique codes related to their SONA IDs*

Are these data anonymised?

*SONA IDs are pseudoanonymised and erased after data collection.*

### Property Rights

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

*no*

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

-

## 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?

*See licenses. Data additionally shared via OSF and Zenodo*

If yes, when will the data be published?

*Pending acceptance of paper submission. Data sets available on Zenodo and OSF from June/July 2025..*

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?

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

**coordinated by:**



<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>