



**SFB
1287**

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

DATA MANAGEMENT PLAN

CRC 1287 | 2021 | Phase 1
Project C02

TABLE OF CONTENTS

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

GENERAL INFORMATION

Overview

Project number

C02

Name of Experiment / Acronym / Number

WP2 - Bare Nouns/Definiteness

Involved persons

Joseph P. De Vaugh Geiss, Hieronymus Lanfer, Lia Bloom, Alexander Schreiber, Mareike Phillipp

PI or responsible person (head of the study)

Malte Zimmermann, Alexander Koller

Subject area

Linguistics > Semantics > bare nouns, definiteness, determiner drop, Kiezdeutsch, noun incorporation

Method / Type of data

- Corpus data
- Behavioral response data
- Spreadsheets for stimuli
- Materials for experimental software
- R scripts

Participants (of the study)

- Corpus study done on KiezDeutsch Korpus (KiDKo)
- Participants for pretests and Study v1 via Prolific (crowd sourced)
- Participants for Study v2 (in preparation) to be determined

Short description (of the study)

Bare SG count nouns may be licit because of (i.) the existence of SG count NPs in lexical incorporation structures and nominal compounding; (ii.) the emergence of bare SG count NPs in German under language contact and a general decline of article use in spoken in formal German; and (iii.) the availability of compositional semantic mechanisms for interpreting such structures in natural language. Pilot study v1 explored contexts licensing definite vs. indefinite vs. specific indefinite bare SG count NPs. Study v2 is in preparation.

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
- 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 317633489

.....

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)

The Prolific platform typically yields results within 24 hours or less.

When does data collection end? (approximately / tentatively)

-

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.

- synchronization with BoxUP / Regular full system
- backups using backup program (deja dup).

Which file formats are used?

Test data Experimental:

- Corpus data: *.csv files
- Presentation: *.csv files for stimuli
- Responses data (raw and processed): *.csv/*.tsv files
- L Rex experimental materials: *.zip

Data processing (preprocessing steps):

- Analysis scripts: *.Rmd code

Which measures of quality assurance are taken for this dataset?

-

Data Analysis

When does data analysis start? (approximately / tentatively)

2-4 weeks initial analysis

When does data analysis end? (approximately / tentatively)

-

.....

.....

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?

- Raw data are not anonymous; reuse of raw data by other researchers/scientists is NOT possible because the identity of the participants could be revealed (personal data).
- However, after data preparation the (pesudo anonymized data can be released to others.

.....

.....

LEGAL AND ETHICS

Personal Data

Does this dataset contain personal data?

yes

If yes, are these data anonymised?

psedonyzed only

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?

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>