



**SFB
1287**

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

DATA MANAGEMENT PLAN
SFB 1287 / 2025 / PHASE 2

PROJECT C08

Table of contents

General Information 1

 Overview 1

 Data Management Requirements 3

 Financial Support 3

Dataset Information 4

 Data Origin 4

 Data Collection 4

 Data Handling 4

 Data Analysis 6

 Data Reuse 6

Legal and Ethics 7

 Personal Data 7

 Property Rights 7

Publication 8

 Publishing or Sharing Data 8

Storage and Long-Term Preservation 9

 Archive 9

General Information

Overview

Project number

C08

Name of Experiment / Acronym / Number

Consequences of Head Argument Order for Syntax / CHAOS

Involved persons

Martin Salzmann; Andreas Hölzl

PI or responsible person (head of the study)

Martin Salzmann

Subject area

Theoretical Linguistics

Method / Type of data

- *Method: linguistic field work*
- *Data: acceptability/grammaticality judgements, sentence elicitations*

Participants (of the study)

- *Primary: linguistically trained experts*
- *Secondary: linguistically naive native speakers of various languages*

Short description (of the study)

The project CHAOS investigates whether there are systematic overall differences in the grammar of VO- and OV-languages that go beyond head-dependent ordering and that are intrinsically related to or caused by the structures linked to the different linearizations of the verb and the object. To achieve this goal, we collect highly comparable data for languages from different areas and language families. On the basis of this data collection, we will find out what influence areal and diachronic factors have on the correlates of the verb-object-position, and, above all, we will identify other factors that interact with the verb-object-position in determining the grammatical profile of a language, and determine the relative strength these factors have in explaining possible correlations.

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?

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)

01.02.2022

When does data collection end? (approximately / tentatively)

01.03.2025

Data Handling

Where is the dataset stored during the project?

- *UP Server & SFB server's*
- *GitUp server:* <https://gitup.uni-potsdam.de/chaos/ProjectC08>
- *Linguistic Database*

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

-

Which file formats are used?

- *PDF*
- *plain text - UTF-8 encoded (required for international scripts)*
- *DOCX and ODT*
- *CSV, JSON*
- *MySQL database*

Which measures of quality assurance are taken for this dataset?

4+-eyes principle; extensive commenting; use common file formats

Data Analysis

When does data analysis start? (approximately / tentatively)

01.07.2022

When does data analysis end? (approximately / tentatively)

31.12.2025

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?

The global morphosyntactic research community would be interested in reusing this dataset. It is a central aim to provide the dataset to the public for reuse. As such, reuse is not only permitted but a design feature. The researchers collaborating with us in the collection of the data provide their consent in written form to the public use of their data.

Legal and Ethics

Personal Data

Does this dataset contain personal data?

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?

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?

Creative Commons Namensnennung-ShareAlike (CC-BY-SA)

If yes, when will the data be published?

Recommended procedure: Upload data and obtain digital identifier (e.g., DOI, OSF link) when submitting the first paper; thus, you can cite the data in the paper. If necessary, restrict public access (embargo) until last paper published (max. 2 years).

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)*
- *UP server's GitUP server: <https://gitup.uni-potsdam.de/chaos/ProjectC08>*

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>