



**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	4
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 - Definite Pseudocleft/Cleft

Involved persons

Joseph P. De Veugh Geiss, Swantje Tönnis, Emilie Destruel, Agata Renans

PI or responsible person (head of the study)

Malte Zimmermann, Edgar Onea

Subject area

Linguistics > Semantics > definiteness > cross linguistic comparisons, in particular with understudied language Akan; homogeneity

Method / Type of data

- Behavioral response and reaction time data
- CSV/ODS files for stimuli
- python scripts for experimental software
- audio files for experiment

Participants (of the study)

University students and staff at the following universities:

- Universität Potsdam/Göttingen, Germany;
- University of Ghana at Legon, Ghana; University of Albi or University of Pau, France;
- Prolific (crowd sourced)

Short description (of the study)

Nouns in Akan exhibit a three way definite system: definites, specific indefinites, and bare nouns with either unique or indefinite interpretation; in contrast, German, English, and French have a two way system divided between definiteness and indefinites. Moreover, for all languages parallelism between definite pseudoclefts and clefts has been proposed.

Comments (optional)

Studies conducted in part with funding from "Exhaustivity in Clefts" project in the XPrag.de priority program (SPP 1727)

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 317633493

.....

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)

4-6 weeks per experiment

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?

- Responses / Reaction time data (raw and processed): *.csv --/*.tsv files
- Presentation: *.csv --/*.ods files for stimuli; *.py scripts and *.html files for experiment (Python/Onexp); *.wav files for auditory stimuli; *.jpg files for images
- Analysis s

Which measures of quality assurance are taken for this dataset?

-

Data Analysis

When does data analysis start? (approximately / tentatively)

3-6 weeks

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?

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?

- Images used in Akan experiment reported in De Veauh Geiss (2020) according to the terms of the Creative Commons BY SA license:
<https://creativecommons.org/licenses/by-sa/2.0/>.

- The only modifications made to them were (i) cropping the images and (ii) converting them to black and white. The four photos were taken by Mark Fischer and are available for download at the following URLs:

- <https://www.flickr.com/photos/fischerfotos/16360022790>

- <https://www.flickr.com/photos/fischerfotos/22986270194>

- <https://www.flickr.com/photos/fischerfotos/23519903990>

- <https://www.flickr.com/photos/fischerfotos/23484547082>

.....

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>