SFB 1287 Limits of Variability in Language
Cognitive, Computational, and Grammatical Aspects

DATA MANAGEMENT PLAN
SFB 1287 / 2025 / PHASE 2

PROJECT C09

Table of contents

General Information

Overview

Project number

C09

Name of Experiment / Acronym / Number

Fieldwork in Felipe Carrillo Puerto, Mexico (2017)

Involved persons

Melanie Uth, Jan Fliessbach, Patrick Auhagen, Iñaki Cano, Laura Piracún, Álvaro Ugarte Isla, Nuria Martínez García

PI or responsible person (head of the study)

Prof. Dr. Melanie Uth

Subject area

Sociolinguistics, phonetics

Method / Type of data

This dataset comprises two types of data:

- *Elicitation experiments of focus constructions.*
- *Sociolinguistic interviews.*

Participants (of the study)

41 adult speakers: 21 Spanish monolinguals or Spanish-dominant speakers, and 20 Spanish Maya bilinguals.

Short description (of the study)

The elicitation task was picture-based and designed to elicit broad focus and contrastive focus in subject and direct object position. It consisted of a total of 40 stimuli per participant which were presented in pseudo-randomized order. Each stimulus consists of an image in which the depicted action and involved figures are accompanied by written descriptions. The

sociolinguistic interviews involved questions on cultural topics such as preferences of taste (music and food) or recommendations for visitors. The interviews were conducted as a dialogue between the speaker and one of the researchers, based on pre-prepared guiding questions. Interview time ranges from 15 to 30 minutes.

Comments (optional)

This dataset was collected in the context of the SFB 1252 Prominence in Language (Project A05) and is being reused for the current research project.

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?

reused

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

This dataset was collected in the context of the SFB 1252 Prominence in Language (Project A05) and is being reused for the current research project

Data Collection

When does data collection start? (approximately / tentatively)

-

When does data collection end? (approximately / tentatively)

-

Data Handling

Where is the dataset stored during the project?

SFB File-Server, External hard drive

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

Monthly backup to external hard drive

Which file formats are used?

.wav, .eaf, .docx,.txt, .xlsx, .R

Which measures of quality assurance are taken for this dataset?

Four-eyes principle,Using common file formats, Good documentation

Data Analysis

When does data analysis start? (approximately / tentatively)

-

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?

-

Legal and Ethics

Personal Data

Does this dataset contain personal data?

yes

Are these data anonymised?

-

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?

-

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

-

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

-

If yes, when will the data be 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?

- *SFB 1287 File-Server*

- *OSF*

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