

Title: Connected speech as an expression of brain health

Data management plan C14/21/109. Promotor Rik Vandenberghe, Copromotor Patrick Dupont

Project summary: Spontaneous speech is an expression of who we are. Here, we examine in older adults how connected speech can also express brain health, the integrity of cognitive brain circuitry, indicate subclinical neurodegeneration and predict future cognitive decline. Recently Natural Language Processing (NLP) has made phenomenal strides and is currently used in a wide variety of applications. We will apply NLP to characterize speech fragments collected in deeply phenotyped cognitively intact older adults. Speech recordings will be collected longitudinally through a secure remote connection in response to questions probing autobiographical memory, episodic memory for personally experienced events and for public events, and semantic knowledge. In the same individuals we will collect volumetric MRI, resting-state functional MRI, as well as amyloid and tau PET imaging. The primary outcome analysis will be based on compositional distributional semantics models. Using fMRI in healthy young adults we will also examine the underlying functional neuroanatomy of key component processes and classes of NLP.

1. General information:

Name applicants: Rik Vandenberghe PI, Patrick Dupont, coPI

Project number and title: Connected speech as an expression of brain health C14/21/109

Affiliation: KU Leuven

2. Datatypes the research will collect and generate

New data will be generated

PET scans, MRI scans (dicom format)

Processed imaging data (nifti format)

Written transcripts of interviews

Processed data from the transcripts

Pseudonymized data with age, sex, educational level, and other individual features

Blood samples

3. Specify in which way the following provisions are in place in order to preserve the data after the end of the research.

All data will be stored on the system provided by the IT department of the group Biomedical Sciences on mirrored hardware on redundant NAS servers. This system is continuously monitored and is secured according to the latest technology. Access to data is only allowed for researchers involved in this project. LCN has access to a total of 2 TByte disks. Archiving is paid from research funds.

Fluid biosamples will be stored in the UZ Leuven biobank after approval of the project by the UZ Leuven Biobank board.

The data will be stored for maximally 20 years after the end of the project in a safe, secure and sustainable ways for purposes of reproducibility, verification and potential reuse.

Metadata of the final datasets will be registered in the academic bibliographic database of the KU Leuven.

Original (raw) data will be stored in a read only format. When data are published, the data on which the reported study results rely will be archived for the different processing steps, together with the exact scripts and version numbers used to generate the study output, and the details of the experiment and study procedures (study report).

4. The project will be conducted in accordance with ICH-GCP guidelines and will be initiated only after approval by the Ethics Committee UZ/KU Leuven Klinische Studies. If a participant withdraws consent, the already acquired data from that participant will be preserved to ensure the validity of the study but no additional data will be acquired.

5. The project will be conducted in accordance with the GDPR regulations and the Belgian legislation regarding the protection of natural persons with regards to processing of privacy data. The data controller will be UZ Leuven. All data will be pseudonymized. The pseudonymization key will be generated based on preset rules (including study identification and subject number). The code will not contain any elements that could lead to the identifiability of the individual.

Only the PI and the team members who need to have direct contact with the study participants for the conduct of the study will have access to the key.

6. Data sharing will be allowed in so far that the ICF permits and only after appropriate Material Transfer Agreements have been put into place.

7. Responsibilities

Patrick Dupont, coPI, is responsible for archiving and storage of data.