

**Cross-linguistic comparison of speech features in older adults with
Alzheimer's Disease and related disorders
(ELEMENT)**

Promoteur de l'étude : Association IA - Institut Claude Pompidou - 10 rue Moliere 06100
Nice cedex.

Tel: +33 4 92034770 - *Fax*: +33 4 92034562

Investigateur principal: **Pr Philippe Robert** –CoBTeK lab_ - Institut Claude Pompidou - 10 rue
Moliere 06100 Nice cedex.

Tel: +33 4 92034773 - *Fax*: +33 4 92034562

Summary

Due to the rapidly ageing population, the number of people with dementia in the EU will triple by 2050. The proposed speech-based screening app supports early detection and intervention, which in turn significantly reduces cost of care and preserves quality of life. The people best placed to spot early cognitive decline are carers, social workers, and family. But there is a clear lack of affordable, usable screening apps that people without medical training can use to validate these concerns and to provide actionable data for medical professionals. The approach can also be used to track mental health and other neurological conditions.

The proposed solution supplements neuropsychological assessment with sophisticated and unobtrusive natural biomarkers extracted from speech data that is collected outside of medical consultations. It provides rich information about cognitive and emotional characteristics and can be used to inform clinical judgment during consultations, saving time and money.

The project aims to validate a new product for fully-automated, reliable, unobtrusive, self-managed screening for cognitive decline, in particular dementia, and other cognitive disorders. It will allow earlier detection and, through that, more effective interventions resulting in the reduction of overall costs associated with treatment and rehabilitation. For users it will offer the comfort of flexible usage without visiting professional physicians.

The main objective is to identify speech and language features in French, which are correlated with cognitive performance and neuropsychiatric symptoms of dementia (e.g., depression, apathy, aggression), as measured by standard clinical cognitive and behavioral assessments as well as video sensors. To achieve this objective, a French speech corpus consisting of 250 speech samples of patients suffering from cognitive and behavioral disorders with different degrees of severity, and 50 samples of control subjects will be collected. The creation of such corpus work will ultimately allow automatic detection of these disorders on the basis of speech analysis.

The corpus will contain ca. 250 samples of speech of patients divided into diagnosis

subgroups with various severity levels of the syndrome, as well as other cognitive and behavioral disorders and ca. 50 samples of healthy people as a control group. All participating individuals will be asked to perform a set of vocal tasks such as describing the content of a series of images. Then, all speech samples will be transcribed (transformed into text) and annotated by professional clinicians to make the corpus suitable for machine learning. The processes, tools and experiences will be then presented in the blueprint for transferring the system into a new language.

Research Objectives and Hypotheses

Our main objectives for a collaborative study between the Computational Linguistics group at the University of Toronto in Toronto, Canada, and the Nice Memory Centre for Care and Research (CMRR) in Nice, France, are to:

1. Identify speech and language features in French, which are correlated with cognitive performance and neuropsychiatric symptoms of dementia (e.g., depression, apathy, aggression), as measured by standard clinical cognitive and behavioral assessments.
2. Collect a large French-language dataset of spontaneous speech from healthy older adults, people with Mild Cognitive Impairment (MCI), people with dementia due to Alzheimer's disease (AD), and other cognitive and behavioral disorders elicited through a standard clinical task consistent with similar protocols conducted in English, in order to facilitate comparative cross-linguistic research herein and in the future.
3. Identify speech and language features, which are predictive of dementia symptomatology but are robust across English and French.
4. Support the proposed speech analysis by a complementary computer-vision based analysis. Such analysis exploits advanced methods related to automated face analysis, tracking, detection and recognition, as well as human behavior analysis.

Participants / Study type:

- We are aiming to enroll 50 healthy subjects, 250 subjects with either mild cognitive impairment (MCI) (Minor cognitive disturbances DSM5), or with Alzheimer's disease (AD) (major cognitive disturbance DSM5) including mixed dementia or other minor / major cognitive or behavioral disorders, depressive / anxiety disorders according to DSM5). The participants in each of the three groups will be age-, education-, and gender-matched.

- Non-interventional observational study. Start of inclusion: February 2017. Study length: 9 months

Methodology

To address the objectives above, we will:

- Extract speech and language features from French-language recordings from participants at the CMRR memory clinic, and construct statistical supervised machine learning models for classification and regression (predicting clinical scores of cognition). We will assess the performance of the features with accuracy, precision, recall, and F-measure in classification experiments, and mean absolute error (MAE) and correlation in regression experiments.
- Extract speech and language features from (i) French-language data collected at the CMRR memory clinic as part of this study, and (ii) English-language data collected at the University of Pittsburgh (DementiaBank), and the University of Toronto (Talk2Me). Compare the performance of the features on each dataset by computing correlation and using statistical analyses to determine which features are predictive of cognitive status (healthy vs AD) but robust to language.

Innovative aspect / Expected benefits

The target customer group can be characterized as individuals 60+ living either at home or in residential care facilities, as well as their families, caregivers, charities, social services, other stakeholders involved in supporting older persons. The project will enable to validate a first-of-that-kind product allowing implementing sophisticated and unobtrusive neuropsychological assessment within minutes right at home or at easy reachable locations without the support of professional clinicians. The initial target markets are France and UK, with the goal to start focused marketing there at the of 2017 – beginning 2018. The focus of initial marketing will be threefold: (1) residential care facilities as access points for groups of users, (2) social services providing homecare, and (3) pharmacies with the modern trend of turning them from sales points to service providers.

Individual benefits:

The proposed study aims to validate a tool allowing automated early detection of voice-based dementia. This tool will make it possible to benefit from an accurate evaluation of the cognitive abilities, which can lead to a better management and advice.

Collective benefits:

Demonstration of the effectiveness of this technology may later make it possible to diffuse its use in clinics or consultation centers, and thus to improve the early diagnosis of neurodegenerative pathologies, while providing data crucial for basic research.