LINGUISTIC AND EXTRA-LINGUISTIC PARAMETERS FOR EARLY DETECTION OF COGNITIVE IMPAIRMENT

KRISTINA LUNDHOLM FORS

Talk outline

- Project and research questions
- Cognitive decline
- Participants and data collection
- Data processing
- Analysis and results
  - first round of recordings
  - second round of recordings
- Ongoing and future analyses

Talk outline

- Project and research questions
- Cognitive decline
- Participants and data collection
- Data processing
- Analysis and results
  - first round of recordings
  - second round of recordings
- Ongoing and future analyses

Dialogue!

- 4 year project financed by The Swedish Foundation for Humanities and Social Sciences
- Project leader: Dimitrios Kokkinakis
- Interdisciplinary: language technology, linguistics, phonetics, speech language pathology, psychology, computer science…

Research question

Is it possible to identify early, subtle linguistic signs that precede dementia in a person's spoken language and reading patterns?

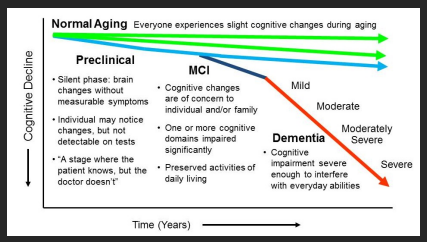
Dementia

- Alzheimer's disease is the most common type of dementia
- Characterized by progressive deficits, primarily in executive functioning and memory
- Linguistic impairments, specifically with regards to semantic processing
- Symptoms may be present for years, even decades, before clinical diagnosis

Cognitive impairment

- Mild cognitive impairment (MCI)
  - Minor problems with cognition
  - Not severe enough to warrant dementia diagnosis
  - Problems do not interfere significantly with daily life
  - High degree of conversion to dementia
- Subjective cognitive impairment (SCI)
  - Perceived cognitive problems
  - Not measurable on standard tests
  - Risk factor for developing MCI and later dementia

Cognitive decline

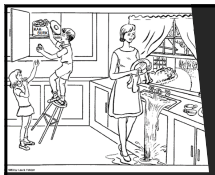


Participants

- 3 groups:
  - persons with subjective cognitive impairment (SCI)
  - persons with mild cognitive impairment (MCI)
  - healthy controls
- Cognitive assessments, neuropsychological tests, MRI etc at the university hospital memory clinic

Participants

- 3 groups:
  - persons with subjective cognitive impairment (SCI)
  - **persons with mild cognitive impairment (MCI)**
  - **healthy controls**
- Cognitive assessments, neuropsychological tests, MRI etc at the university hospital memory clinic



First recording (2016)

approximately 90 participants

- description of cookie theft picture
- reading two short texts with eye-tracking, followed by comprehension questions
  - reading silently
  - reading outloud

Second recording (2018)

approximately 75 returning participants

3 added tasks

- semantic verbal fluency
- "Trip to Stockholm" – planning task
- map task

Follow-up at memory clinic (2018-2019)

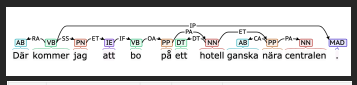
- Updated cognitive assessment (GDS; global deterioration scale)
- Neuropsychological tests

Data processing

- manual, orthographic transcription of recorded data
  - utterance segmentation
  - annotation of disfluencies such as fillers and false starts
- automatic alignment of text and audio files (manual corrections)

Data processing, cont.

- part of speech-tagging
- lemmatizing
- dependency parsing
- constituent parsing



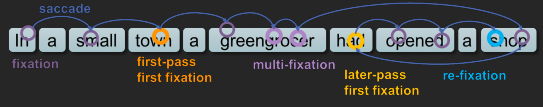
token	msd	lemma	lex	sense
Där	AB	där	där..ab.1	där..1
kommer	VB. PRS. AKT	komma	komma..vb.1	komma..1 (0.596), komma..3 (0.404)
jag	PN. UTR. SIN. DEF. SUB	jag	jag..pn.1	jag..1
att	IE	att	att..sn.1	att..1
bo	VB. INF. AKT	bo	bo..vb.1	bo..1
på	PP	på	på..pp.1	på..1
ett	DT. NEU. SIN. IND	en	en..at.1	den..1, en..2
hotell	NN. NEU. SIN. IND. NOM	hotell	hotell..nn.1	hotell..1

Analyses based on the first round of recordings

- Using eye-movements during reading to detect MCI
- Adding data from another language
- Learning from multiple tasks and modalities

Using eye-movements during reading to detect MCI

- Extract eye-movement features from silent and aloud reading
- Augment them with linguistic features of the text
- Train machine learning classifier to distinguish MCI vs HC

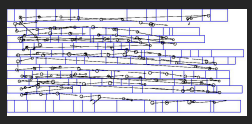


Fraser, K. C., Lundholm Fors, K. , Kokkinakis, D., & Nordlund, A. (2017). An analysis of eye-movements during reading for the detection of mild cognitive impairment. In *Proceedings of the 2017 Conference on EMNLP* (pp. 1016-1026).

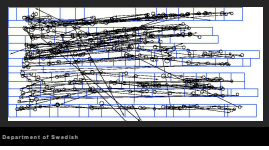
Results

- **Best configuration:** 86% accuracy
- **Silent versus aloud:** For each classifier and feature set, the accuracy was higher in the silent trial.
- **Augmenting features with linguistic info:** On average, the baseline features led to better performance in all cases.
- **Individual versus combined trials:** On average, using silent reading data alone led to better results than concatenating with aloud reading data, but the best result was achieved by merging data from the two trials.
- **Relevant features**: The only features which differed significantly between the two groups were a *decreased number of first-pass first fixations* and an *increase in later-pass first fixations* in the MCI group

Examples



Control participant
(Trial 1)



MCI participant
(Trial 1)

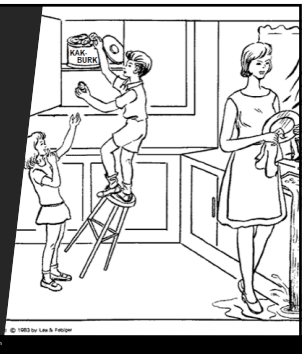
Adding data from another language

- Research groups all over the world are collecting these small clinical data sets in different languages
- How can we combine these data to improve statistical power and machine learning accuracy?

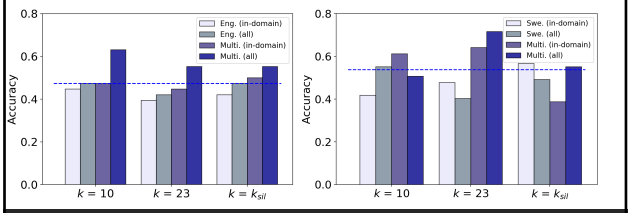
Fraser, K. C., Lundholm Fors, K., & Kokkinakis, D. (2018). Multilingual word embeddings for the assessment of narrative speech in mild cognitive impairment. *Computer Speech & Language*, *53*, 121–139.

Approach

- Use English and Swedish word embeddings
- Rotate them into into the same space
  Train **multilingual cluster models**
- Extract features in English and Swedish
- Classify MCI vs HC

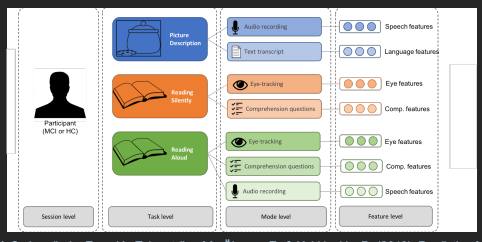


Results



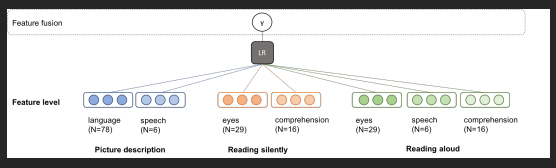
Multilingual topic models lead to significantly better ($p = 0.01$) accuracies than unilingual topic models in both English and Swedish.

Learning from multiple tasks and modalities



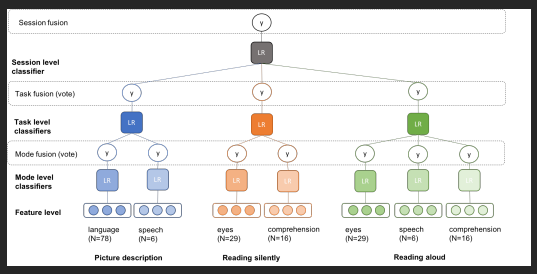
Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F., & Kokkinakis, D. (2019). Predicting MCI Status From Multimodal Language Data Using Cascaded Classifiers. *Frontiers in Aging Neuroscience*, *11*, 205.

Feature fusion



accuracy: 0.70

Late fusion (at level of mode, task, session)



0.80

0.83

0.76

Analyses that include data from second round of recordings

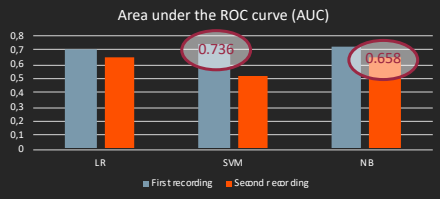
- longitudinal study of eye movements when reading
- comparison of picture description task and planning task

Longitudinal study of eye movements when reading

- extraction of eye-tracking features for participants that took part in both recordings
- compare classification results

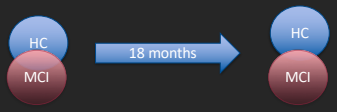
Lundholm Fors, K., Antonsson, M., Kokkinakis, D., & Fraser, K. C. (2019). Reading and mild cognitive impairment. In *Proceedings of the 10th International Conference of Experimental Linguistics.*

Classification results



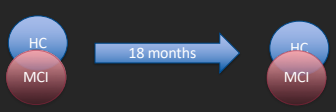
Worse classification results – why?

- Expectation: cognitive differences should increase over time



Worse classification results – why?

- Expectation: cognitive differences should increase over time



Worse classification results – why?

- Expectation: cognitive differences should increase over time
- But: up to 44% of persons with MCI return to normal within a year (Gauthier et al., 2006)
- Prevalence of MCI increases with age

This underlines the need for updated cognitive assessments.

Comparison of picture description and planning task

- planning task was added to provide a more cognitively and linguistically demanding task
- does the performance on the planning task differ more between the groups than the performance on the picture description task?

Antonsson, M., Lundholm Fors, K & Kokkinakis, K. (2019). Language tasks and mild cognitive impairment. In *Proceedings of the 10th International Conference of Experimental Linguistics.*

Results

MCI vs HC
no significant difference

MCI vs HC
F(5, 50)=2.65, p = 0.03

	Cookie theft			Planning task		
	MCI	HC	sig.	MCI	HC	sig.
Words (n)	142.0	172.7	n.s	325.4	426.4	t = 5.8, p = 0.019
Sentences (n)	11.5	12.7	n.s	22.8	30.7	t = 8.18, p = 0.006
Pronoun index	0.36	0.36	n.s	0.55	0.58	n.s
False start ratio	0.005	0.005	n.s	0.006	0.01	n.s.
Fillers ratio	0.02	0.02	n.s	0.03	0.03	n.s.

	Cookie theft			Planning task		
	MCI	HC	sig.	MCI	HC	sig.
Words (n)	142.0	172.7	n.s	325.4	426.4	t = 5.8, p = 0.019
Sentences (n)	11.5	12.7	n.s	22.8	30.7	t = 8.18, p = 0.006
Pronoun index	0.36	0.36	n.s	0.55	0.58	n.s
False start ratio	0.005	0.005	n.s	0.006	0.01	n.s.
Fillers ratio	0.02	0.02	n.s	0.03	0.03	n.s.

Ongoing and future analyses

- Incorporate results from renewed cognitive assessments
- Thematic analysis of planning task
- "Modalizations"
- Swedish Framnet++ for analysis of semantic content in planning task
- Analysis of map task: dialogue features, coherence and cohesion etc
- Investigate predictive power of linguistic features


GÖTEBORGS UNIVERSITET
SPRÅKBANKENTEXT
RIKSBANKENS JUBILEUMSFOND
Thank you for listening!
Questions?
GÖTEBORGS UNIVERSITET
The Swedish Language Bank, Department of Swedish