

Manual and Automatic Transcriptions in Dementia Detection from Speech



Jochen Weiner, Mathis Engelbart, Tanja Schultz

jochen.weiner@uni-bremen.de

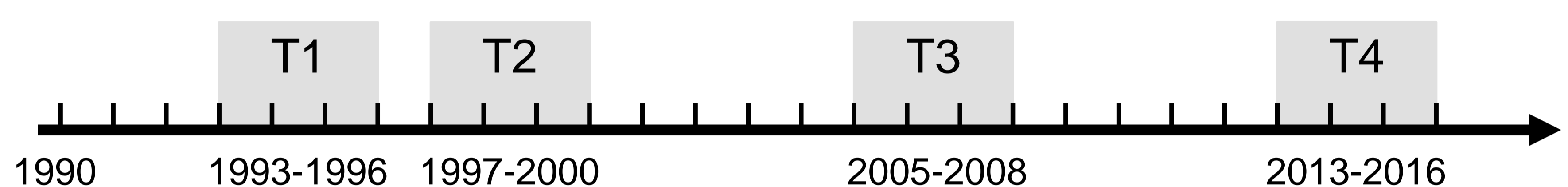
Motivation

- The population in Germany is ageing rapidly
- With age comes a higher risk for developing dementia
- Society needs fast and affordable means of detecting dementia
- Automatically detect dementia from speech
- We can detect dementia from speech using
 - Acoustic features
 - Linguistic features → Needs transcription
- Automatic detection relies on automatic transcriptions
- Investigate how our dementia detection works with automatic transcriptions

Database: ILSE

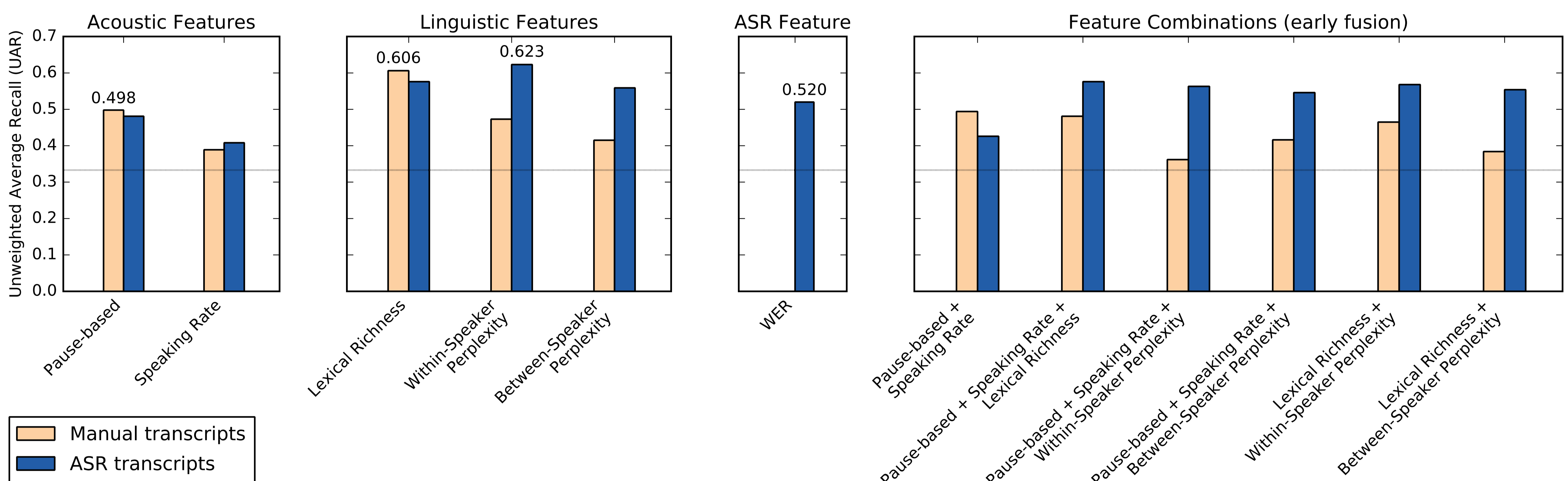
- Biographic interviews
 - Several hours duration
- Cognitive diagnoses made by psychiatrists
- Three measurements
 - Over time some participants developed dementia
- Real world data set
 - Natural prevalence of dementia
- Selected data:
 - Cognitive diagnoses available
 - Transcription available

		Diagnoses		
		control	AACD	AD
Measurement	T1	51	4	-
	T2	19	8	-
	T3	10	1	5
Total		80	13	5



Dementia Detection

- Leave-one-person-out cross-validation
- Mutual information feature selection
- Gaussian classifier
- Compare features from manual and ASR transcriptions
- Metric: Unweighted Average Recall (UAR)



- We can detect dementia using ASR transcripts
 - No need for manual transcription
 - Features are robust against ASR errors
- Automatic versions of the features outperform their manual counterparts
- Word error rate itself is a strong feature
- Feature selection + feature-level fusion finds no better features
- Future work: Dementia detection using speaker diarization and automatic transcription

Automatic Speech Recognition System

- Trained on 256 hours of ILSE interviews
 - Acoustic Model: DNN, 6 hidden layers
 - Language Model: 3-gram KN (ppl 199.38)
 - Dictionary: Train vocab (oov 1.49 %)

Diagnosis	Overall WER	mean WER	std WER
control	56.0 %	58.2 %	14.9 %
AACD	60.4 %	60.8 %	12.1 %
AD	70.2 %	70.1 %	7.0 %
All	58.5 %	59.2 %	14.4 %