Manual and Automatic Transcriptions in Dementia Detection from Speech
Jochen Weiner, Mathis Engelbart, Tanja Schultz

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

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)

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 %)

Diagnoses

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Overall WER</th>
<th>mean WER</th>
<th>std WER</th>
</tr>
</thead>
<tbody>
<tr>
<td>control</td>
<td>56.0 %</td>
<td>58.2 %</td>
<td>14.9 %</td>
</tr>
<tr>
<td>AACD</td>
<td>60.4 %</td>
<td>60.8 %</td>
<td>12.1 %</td>
</tr>
<tr>
<td>AD</td>
<td>70.2 %</td>
<td>70.1 %</td>
<td>7.0 %</td>
</tr>
<tr>
<td>All</td>
<td>58.5 %</td>
<td>59.2 %</td>
<td>14.4 %</td>
</tr>
</tbody>
</table>