Linguistic Features During Speech Utterances in the Context of Social Skills Training

Hiroki Tanaka¹, Hidemi Iwasaka², Hideki Negoro³, Satoshi Nakamura¹

¹Nara Institute of Science and Technology (NAIST), Japan
²Nara Medical University, Japan
³Nara University of Education, Japan

EMBC 2020
(Research posters paper)
Social Skills Training (SST)

- Psychosocial treatment
- For people with autism, schizophrenia [Bellack, 2013]

To obtain appropriate social skills through
- Instruction
- Modeling
- Role-play
- Feedback
- Homework

Participant

Trainer
Automated social skills trainer

- People with autism have good systemizing skills
- Embodied conversational agents
- Speaking skills [Tanaka, et al., 2015] and listening skills [Tanaka, et al., 2020]
Role-play and feedback

Speech contents

- Words per minute, 6 letters, fillers
- F0 variation, amplitude, voice quality, pauses
- Ratio of smiling, yaw, pitch

Sensing & analysis from video

Behavior generation

Feedback

Online feedback

Summary feedback
Speech contents analysis

- 36 data of adults and 18 data of children/adolescents with autism spectrum disorders
- Speak about a recent fun story for one minute to a person
- Human social skills trainer rated the score (1 to 7)
- Extracted linguistic features:
  - # of Tokens
  - Type-token ratio (TTR)
  - # of conjunction
  - # of fillers
  - Similarity of word embedding
    *sequentially shifting ten words
  - Parse tree depth (depth)
Correlations ($r$) to overall speaking skills

*represents $p < 0.05$, and †represents $p < 0.1$ compared to no correlation

### Adults

<table>
<thead>
<tr>
<th></th>
<th>Tokens</th>
<th>TTR</th>
<th>Conj.</th>
<th>Fillers</th>
<th>Similarity</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>0.50*</td>
<td>-0.09</td>
<td>0.03</td>
<td>0</td>
<td>0.10</td>
<td>0.38*</td>
</tr>
</tbody>
</table>

### Children/adolescents

<table>
<thead>
<tr>
<th></th>
<th>Tokens</th>
<th>TTR</th>
<th>Conj.</th>
<th>Fillers</th>
<th>Similarity</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>0.57*</td>
<td>0</td>
<td>0.53*</td>
<td>0.29</td>
<td>-0.07</td>
<td>0.43†</td>
</tr>
</tbody>
</table>
Conclusions

- We have proposed automated social skills trainer
- This study analyzed speech contents
- Adults: # of tokens, the tree depth are significantly correlated to overall speaking skills
- Children/adolescents: # of tokens, the tree depth, # of conjunctions are correlated to overall speaking skills
- These features can be used for feedback in future automated SSTs