A framework for multidisciplinary project work with cochlear implant users: Involving CI users in developing rehabilitation tools to improve performance in multi-speaker conversation

Harriet Crook1, Erica Bradley1, Bill Wells2, Guy Brown2, Amy Beeston2, Emina Kurtic2
1 Department of Neurology, Sheffield Teaching Hospitals NHS Trust. 2 University of Sheffield, UK.

Background

- Simultaneous or overlapping talk by two or more speakers (OT) is surprisingly frequent.
- In casual conversations among young British adults it occupies 16% of total talking time, while 41% of speaker turns are overlapped by another speaker (Kurtic et al. 2012).
- A recent survey undertaken for this project has shown OT is a common problem for CI users.

Sheffield corpus of informal conversation

- Previous studies explore single speaker in multi-speaker babble or background noise.
- This is not the same task as multi-speaker conversation.
- Here, 4 friends (3 female, 1 male), native (British) English, chatting in an informal setting.
- Corpus provides data for realistic audio-visual training materials.

CI performance on acoustic cues in OT

- Poorer performance by CI users on a range of acoustic tasks used in OT.
- Day to day difficulties in OT are signalled by how acoustic tasks combine in real listening.
- The complex task-specific nature of modelling speech processing supports use of real-time acoustic materials and training processes while further research generates sufficient laboratory evidence on the interaction of acoustic variables and processing algorithms in live speech settings.

Towards the training software

- Short-term – assess difficulty of multi-speaker conversational material for CI users.
- Mid-term – design user tasks of increasing difficulty level.
- Long-term – software to help develop CI users’ skills in (i) listening and (ii) speaking.

Assessing task difficulty

- Gradually increase CI users’ zone of comfort

Assessing conversational difficulty

- Gradually introduce more complex conversational material

Aims and methods

- Identify specific issues that OT raises for CI users: focus groups; web-based questionnaire survey; observation of conversations.
- Develop training materials that replicate task difficulty in realistic conversational settings: real conversational snippets selected from corpus of informal multi-speaker conversation.
- CI users included in software development process: piloting samples of training material to assess difficulty of tasks and software.
- Develop ways to improve CI users’ experience during conversation: devising training software and activities for CI users, their families and professionals.
- Training software promoted to CI users, their families and professionals.

Table 1: shows performance variation with some acoustic features dependent on the task.

<table>
<thead>
<tr>
<th>Acoustic Task/ Parameter</th>
<th>Role in OT</th>
<th>Evidence on CI performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0 and Intensity</td>
<td>Combined F0 and Intensity: most prominently used prosodic resource for turn competition (Kurtic et al., 2013); Stress and Pitch variation are used to identify turns and appropriate places to interject.</td>
<td>F0 best represented by temporal fine structure (TFS); TFS poorer in CI, current processing mainly codes slow varying envelope information rather than TFS; TFS may be useful in separating target from masker (Heng et al. 2011). Mulheim &amp; Cullington (2014): no statistical difference on talker CI between bimodal/binarial CI despite hypothesis that acoustic hearing may provide better F0 perception. Other studies do show benefit of F0 with addition of acoustic hearing (e.g., Zhang et al 2010). Peng et al (2008): CI users ability to assess whether an utterance is a question or assertion, with average score of 70% in CI users versus 97% in NH listeners.</td>
</tr>
<tr>
<td>Gender Identification</td>
<td>Tracking specific speaker and identifying OT</td>
<td>Chatterjee &amp; Peng (2008) showed similar performance between NH, NH with CI simulation, and CI.</td>
</tr>
</tbody>
</table>

Envisaged functionality

- Objective: understand multi-party conversation

Key skills:
- identify the main talker
- recognise relevant speech content
- understand the underlying social action

- Objective: join in multi-party conversation

Key skills:
- identify suitable start times
- control pitch, loudness, speech rate
- use overlaps appropriately

References

- Zhang, F., Dorman, M. F. and Spahr, A. (2010). ‘Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation’, Ear and hearing, 31(2), 83-91.