A multi-modal training approach to improve cochlear implant users' ability to handle simultaneous talk

Amy Beeston and Emina Kurtić
Multimodality in Language Research
Leeds · 26 June 2014
Project team

University of Sheffield
• Prof Bill Wells, Human Communication Sciences
• Prof Guy Brown, Computer Science
• Dr Emina Kurtić and Amy Beeston, Research Associates, Computer Science
• Ella Page, project student, Human Communication Sciences

Sheffield Teaching Hospitals NHS Trust
• Dr Harriet Crook, Clinical Scientist, Neurotology
• Erica Bradley, Speech & Language Therapist, Neurotology

Cochlear implant (CI) user panel
• 5 cochlear-implanted participants recruited from Sheffield Cochlear Implant Service
Introduction

• Losing your hearing is not simply the absence of sound
  – for individual: loss of capacity to take part in social life
  – additional societal stigma: being treated differently*
• Cochlear implantation is not a cure for deafness
  – it is a prosthetic substitute.
• Approx 324,200 people worldwide implanted by end of 2012†
  – 10,000 cases in UK†
  – c. 113 per month in UK‡
• Insufficient evidence/training materials to guide clinicians to
  help cochlear implant (CI) users deal with simultaneous talk.

Electric hearing

- Acoustical vibrations are delivered to brain as electrical signals
Challenges for CI users

• Cochlear implant (CI) users need optimum conversational settings (quiet, one partner, awareness to avoid overlap).
• Q: “When no background noise is present, can you effortlessly participate in conversations with friends or family members (e.g., after dinner)?”

Mean response for 26 adults = 5.7

MED-EL Hearing Implant Sound Quality Index (HISQUI) v23015_2.0, English.
Challenges for CI users

• Q: “When multiple people are talking simultaneously, can you effortlessly follow discussions of friends and family members?”

   never 1 2 3 4 5 6 7 always
   “rarely” occasionally “always”

• Mean response for 26 adults = 2.8
• Unfamiliar voices are still more difficult to follow
• Limits of the device? Conversational/social experience?

MED-EL Hearing Implant Sound Quality Index (HISQUI) v23015_2.0, English.

Multimodality in language research
26 June 2014 · Leeds · Amy Beeston and Emina Kurtić
Social signals and behaviours

• Multiple behavioural cues combine to produce a social signal
  – Physical appearance
  – Gesture and posture
  – Face/eyes behaviour
  – Space and environment
  – Linguistic content
  – Vocal behaviour *

* prosodic features, e.g., pitch, loudness, duration, silence, ...

• Stronger reliance for CI users on non-auditory cues?
• Is vocal non-verbal behaviour informative or distracting?

CI sound encoding

- 22 electrodes vs. 30,000 hair cells.
- Low spectral resolution: hard to distinguish one talker from another (especially concurrent talkers).
- Poor representation of voice F0: hard to identify social actions realised by prosodic cues (e.g., turn competition).
Scene analysis problem

- Parse single-channel stream without cues from pitch or location

Conversational dataset

- 3 hours of informal talk, native English (UK)

- 16% of total talk time is overlapping talk
- 41% of speaker turns are overlapped by another speaker

- Lip-reading impossible; Gaze, gesture, posture, ...
- Segmentation by turn construction unit (TCU), transcription, overlap competitiveness, xml data in ELAN annotation format

Cl-user feedback

• Web-based questionnaire
  – hearing conversations
  – taking part in conversations
  – age, gender, and hearing status (implant type) per ear

• Cl-user panel
  – 5 expert Cl-users recruited for duration of project
  – conversational experiences, awareness of overlap
  – grading speech material
  – grading user tasks
  – evaluate training software
    – listening skills
    – (later, speaking skills)
Listening skill development

- Objective – Understand multi-party conversation
- Visualisation of acoustic information
- CI-user difficulty ratings for speech material and tasks

Reported difficulties
- identify talker
- recognise speech
- understand action
i. Speech material

- Graded exposure to multi-party conversation
  - Incremental learning through repetition
- Classify by number/gender of talkers, turn transitions
ii. User tasks

- Widen conversational experience
  - Rate difficulty for CI user
- Focus on non-linguistic aspects
  - identify number of talkers
  - identify same/different gender
  - identify clear/overlap

Number of talkers, N
(single = 0, multiple = 1)

Gender, G
(same = 0, different = 1)

Turn exchange, T
(clear = 0, in overlap = 1)

<table>
<thead>
<tr>
<th>N</th>
<th>G</th>
<th>T</th>
<th>speech collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 talker</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2 talkers, same gender, clear</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2 talkers, same gender, overlap</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2 talkers, diff. gender, clear</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2 talkers, diff. gender, overlap</td>
</tr>
</tbody>
</table>
Speaking skill development

- Objective: participate in multi-party conversation
  - 1. gather voice norms (pitch, intensity, tempo)
  - 2. use overlaps appropriately

- Identify entry times (especially in overlap)
- Control pitch/loudness (show competitiveness)
Conclusions and further work

• Some CI-users keen to develop multi-party conversational skills
  – varied population
  – different experiences/strategies
• Informal talk is extremely challenging
  – missing some turn-taking cues
  – multimodal presentation of acoustic cues
• Learning interactional strategies for handling overlapping talk
  – complement technological advances
  – increase social participation
• Software may be more widely relevant
  – hearing or communicative difficulties
Thank you for your attention

• Any questions?