Real world environments of toddlers: How accurate is cochlear implant scene classification?

Hillary Ganek
Deja Forde-Dixon
Sharon Cushing
Blake Papsin
Karen Gordon
Cochlear Implant Team

DIRECTORS
- Karen Gordon
- Blake Papsin
- Sharon Cushing

RESEARCH COORDINATORS
- Jaina Negandhi
- Alan Blakeman

GRADUATE STUDENTS
- Melissa Polonenko
- Daniel Smieja
- Ashley Deonarain
- Claire McSweeny
- Robel Alemu

POST-DOCTORAL FELLOWS
- Homira Osman
- Hillary Ganek
- Hyo-Jeong Lee

COLLABORATORS

Local – SickKids
- Robert Harrison
- Susan Blaser
- Adrian James
- Paolo Campisi
- Vicky Papaioannou
- Ben Dunkely
- Elizabeth Pang

Local – External
- Sandra Trehub
- Frank Russo
- Sam Doesburg
- Andrew Dimeitrijvic
- Vincent Lin
- Claude Alain

International
- Robert Cowen

CLINICAL TEAM
- Gina Goulding
- Laurie MacDonald
- Mary-Lynn Feness
- Pat Di Santo
- Susan Druker
- Valerie Simard
- Naureen Sohail
- Rebecca Malcolmson
- Jill Liberman

SUMMER STUDENTS
- Jonah Gorodensky
- Robel Alemu
- Meera Mehta
- Jacob McLennan
- Kavina Shanmugalingam

FUNDING
- CIHR
- SickKids Foundation
Disclosures

- Drs. Cushing, Papsin, and Gordon are on Cochlear’s and Advanced Bionics’s Speaker Bureau
Objective
Language acquisition requires

- **Quality stimulation** (Hoff & Naigles, 2002)
- **Quantity stimulation** (Gilkerson et al., 2017; Hart & Risley, 1995)
- **Quiet environment** (Caldwell & Nittrouer, 2013; Gremp & Easterbrooks, 2018)
- **Quick-time (and other forms of electronic media)** (Christakis et al., 2009; Sosa, 2016; van den Heuvel et al., 2019; Zimmerman et al., 2009)
Cochlear Scene Classification

- 597 datalogs
- Unilateral, bimodal, & bilateral CI users
- Datalog lengths: 9.9±3.4 hours

Easwar et al., 2016
Hypotheses

- Cochlear implants may miss nuances of sound that are classified as quiet
- Children will interact frequently with electronic sounds and those instances will be missed by the cochlear implant
Participants

- 11 datalogs
- Datalog length: $10.1 \pm 1.8$ hours
- Recording length: $14.7 \pm 1.6$ hours
Sound Categories identified by human coders

- **Quiet**
  - Silence
  - Quiet Speech
    - Electronic
  - Quiet Noise
    - Electronic
  - Quiet Music
    - Electronic

- **Speech**
  - Speech-in-Noise
  - Speech
    - Electronic

- **Noise**
  - Noise
    - Electronic

- **Music**
  - Music
    - Electronic

High interrater reliability
(ICC=.95, p<.001)
Scene classification detects primarily ‘Quiet’ and ‘Speech’
Humans detected primarily ‘Speech’
'Quiet' Sounds

![Bar chart showing average proportion of different sound categories: Quiet, Speech, Speech in Noise, Music, Noise, Wind.](chart)

<table>
<thead>
<tr>
<th></th>
<th>Average Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quiet</td>
<td>30</td>
</tr>
<tr>
<td>Speech</td>
<td>40</td>
</tr>
<tr>
<td>Speech in Noise</td>
<td>30</td>
</tr>
<tr>
<td>Music</td>
<td>20</td>
</tr>
<tr>
<td>Noise</td>
<td>10</td>
</tr>
<tr>
<td>Wind</td>
<td>5</td>
</tr>
</tbody>
</table>

Human Coder Environments:
- Silence
- Quiet Speech
- Quiet Electronic Speech
- Quiet Noise
- Quiet Electronic Noise
- Quiet Music
- Quiet Electronic Music
- Speech
- Electronic Speech
- Speech-in-Noise
- Music
- Electronic Music
- Noise
- Electronic Noise

CI2019 Hollywood, Florida, July 10-13, 2019
'Speech-in-Noise' Sounds

[Bar chart showing average proportion across different conditions: Quiet, Speech, Speech-in-Noise, Music, Noise, Wind.]

Human Coder Environments

- Silence
- Quiet Speech
- Quiet Electronic Speech
- Quiet Noise
- Quiet Electronic Noise
- Quiet Music
- Quiet Electronic Music
- Speech
- Electronic Speech
- Speech-in-Noise
- Music
- Electronic Music
- Noise
- Electronic Noise
'Noise' Sounds

<table>
<thead>
<tr>
<th></th>
<th>Quiet</th>
<th>Speech</th>
<th>Speech in Noise</th>
<th>Music</th>
<th>Noise</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

Human Coder Environments

- Silence
- Quiet Speech
- Quiet Electronic Speech
- Quiet Noise
- Quiet Music
- Quiet Electronic Music
- Speech
- Electronic Speech
- Speech-in-Noise
- Music
- Electronic Music
- Noise
- Electronic Noise

CI2019 Hollywood, Florida, July 10-13, 2019
Scene classification is relatively accurate

*(ICC=.82, p=.05)*
Scene classification over-identifies 'Quiet' sounds
Scene classification misses some ‘Speech’ sounds
Electronic sounds were found in most scene classes
Electronic sounds identified by human coders

7.8% of the day was electronic sound (47 minutes)
Summary

- Cochlear implant scene classification accurately reflects children’s auditory environments
- ‘Quiet’ may be over-identified and ‘Speech’ maybe under-identified by the algorithm
- The algorithm does not adequately account for electronic sounds
Thank you to all of our participants