Eye Tracking Studies of Spoken Word Production in Children with Cochlear Implants

Richard G. Schwartz\textsuperscript{1,2}, Susan Steinman\textsuperscript{2}, Elizabeth Ying\textsuperscript{2}, Zara W. DeLuca\textsuperscript{1,2}, Georgia Drakopoulou\textsuperscript{1,2}, and Derek Houston\textsuperscript{3}

\textsuperscript{1}The Graduate Center of the City University of New York
\textsuperscript{2}New York Eye and Ear Infirmary of Mount Sinai
\textsuperscript{3}Indiana University Medical School

Supported by NIDCD 1RO1DC011041
Acknowledgements & Disclosure

Simon Henin
Ankit Rastogi
Jessie Schanker
The children and their parents

Disclosure: All the authors are paid employees of the NIH Grant. We have no other conflicts.
Examining a subject's eye gaze during response to a stimulus can provide further insight into how that stimulus is being processed.

Cohort effect. Upon hearing a word, all phonological and semantically related (competitor) words are activated. As identification proceeds, competitors are inhibited (or their activation fades) and the target activation remains.

Eye tracking enables us to quantify attention to a related stimulus or a foil, as well as create a timeline of when that attention occurs relative to stimulus onset and response.
Participants

All children:
Spoke primarily English at home
Passed a hearing test in sound field at 30dB
Completed the MLNT.
Scored within normal range on the TONI.

All children with CIs:
Had their implants for at least 3 years

All children with NH:
Within normal range on the CELF and PPVT

Exclusion Criteria:
Children whose primary language was not English
Children with Language disorders, ADD, Autism, or neurological disorders
Children who use ASL and spoken English
# ET Recognition

<table>
<thead>
<tr>
<th></th>
<th>CI (n = 21)</th>
<th>NH (n = 52)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Age</td>
<td>9.7 (1.4)</td>
<td>9.4 (1.3)</td>
</tr>
<tr>
<td>Implant Age</td>
<td>1.8 (1.3)</td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>94 (20.8)</td>
<td>118.8 (23.9)</td>
</tr>
<tr>
<td>CELF</td>
<td>88.5 (25.3)</td>
<td>109 (25.4)</td>
</tr>
<tr>
<td>MLNT</td>
<td>78 (30)</td>
<td>85.3 (32.5)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Bimodal</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Unilateral</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
# ET Production

<table>
<thead>
<tr>
<th></th>
<th>CI (n = 19)</th>
<th>NH (n = 63)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>9.7 (1.2)</td>
<td>9.6 (1.3)</td>
</tr>
<tr>
<td>Implant Age</td>
<td>2.0 (1.3)</td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>96.7 (22.5)</td>
<td>112 (28.2)</td>
</tr>
<tr>
<td>CELF</td>
<td>96.7 (33.1)</td>
<td>112.6 (22.9)</td>
</tr>
<tr>
<td>MLNT</td>
<td>74.4 (30)</td>
<td>89.1 (27.7)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Bimodal</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unilateral</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Design and Stimuli

Same pictures and auditory words for Recognition and Production.

Pictures were presented in quartets, with one target, two unrelated and one phonologically (onset or rhyme) related or one semantically related (associate or categorical) picture.

Within each category, half the words came from Low Density Neighborhoods (e.g., vacuum — very few phonologically similar words) and half from High Density Neighborhoods (e.g., bat — many similar words).

Eight blocks of stimuli: 34 trials per block and 272 total.
Experimental Procedure

Recognition
1. Four pictures appeared on the eye tracking monitor, separated by an inch of green space.
2. A word was presented over the speaker.
3. Children used the mouse to click on the picture that matched the word they heard.

Production
1. Four pictures appear on the eye tracking monitor, separated by an inch of green space.
2. 100ms after presentation, a pink border appeared around the target picture.
3. Children named the bordered picture as quickly as they could.
ET Recognition: Results

The NH children looked far more frequently at the target than the CI children.

The CI children looked more at the relative than the NH children.

The children with CIs looked more frequently at at least one of the unrelated pictures than the NH children.
ET Recognition
Mean Proportion of Fixation Time

[Graph showing mean proportion of fixation time for Target, Relative, Control1, and Control2 with CI and NH categories.]
Both groups looked longer at relatives when the target was from a sparse neighborhood than when it was from a dense neighborhood. CI subjects looked more at the relatives than did NH subjects.
Lexical neighborhood density affects lexical access for both groups. Deactivation of non-target relatives from sparse neighborhoods is slower than those from dense ones.

Overall, CI subjects attended to non-target pictures longer than their NH peers, suggesting a slower timeline of processing for lexical access, and potentially a different lexical organization.
The groups did not differ in their fixations towards the target.
Thank you!

Language and Hearing Research Laboratory
New York Eye and Ear Infirmary of Mount Sinai