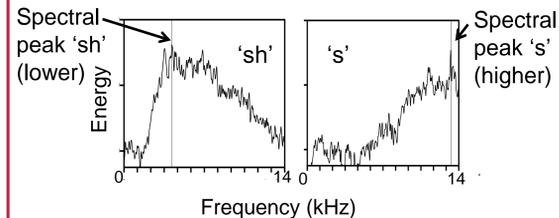


## INTRODUCTION

A **cochlear implant (CI)** can provide an auditory signal to deaf individuals, but with less frequency information than what listeners with **normal hearing (NH)** can access. Perhaps due to device limitations, the speech of children with **CIs** is somewhat less intelligible than that of peers with **NH**.



Frequency information can help listeners tell speech sounds apart. **Spectral peak** refers to the frequency in a sound with the most energy. The sound 's' has a higher-frequency spectral peak than the sound 'sh'.



Children with **CIs** produce 's' with a lower spectral peak than their peers with **NH**, but both groups of children produce 'sh' similarly [1]. This effect is present even for productions that are judged as correct by a trained transcriber.

**Judged as Correct ≠ Acoustically Identical**

**Question:** Is the subtle acoustic difference in 's' between **NH** and **CI** groups perceptible to adults?

**Motivation:** Begin to understand if the acoustic difference impacts listeners' ease of understanding.

## HYPOTHESES

- Adults will rate the 's' and 'sh' productions by children with **CIs** more closely together than those by children with **NH** on a scale from 's' to 'sh'. Previous acoustic analyses have found that both groups produced 'sh' similarly, but children with **CIs** produced 's' with less contrast from 'sh' than children with **NH** [1].
- Adults will respond most slowly to 's' productions by children with **CIs**. Children with **CIs** produced less acoustic contrast between 's' and 'sh' than peers with **NH** [1]. Listeners respond more quickly when sound categories are more distinct from each other [2].

## PARTICIPANTS

**Talkers:** 21 children with **NH**, 21 children with **CIs** 4-7 years old

- ✓ Age-matched within 4 months
- ✓ Children with **CIs** implanted before 2.5 years old
- ✓ Typical development (except hearing loss for **CI** group)

✓ Native English speakers

**Listeners:** 36 adults (19 female, 17 male) 19-27 years old

- ✓ Native English speakers
- ✓ No training in phonetic transcription

## METHODS

### Stimuli

- 220 speech samples (tokens)
- Subset from acoustic analysis [1]
- Word initial consonant-vowel sequences
- Judged as correct by trained transcriber
- Balanced by vowel context
- Amplitude normalized

55 NH-s, 55 CI-s, 55 NH-sh, 55 CI-sh

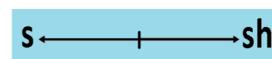
# of tokens by Hearing Type (NH vs CI) and Sound ('s' vs 'sh')

### Design

- 3 perception tasks presented in varied order + Acoustic Analysis [1]
- All 220 tokens presented in random order in each task

### SECTION

**Visual Analog Scale**  
Click along the line



### MEASURE

Scale Rating

**Respond to Is it 's'?**  
Yes or No



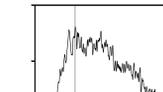
Token Judgment  
Reaction Time

**Respond to Is it 'sh'?**  
Yes or No



Token Judgment  
Reaction Time

**Acoustic Analysis**  
As in [1]



Spectral Peak

### Testing Setup

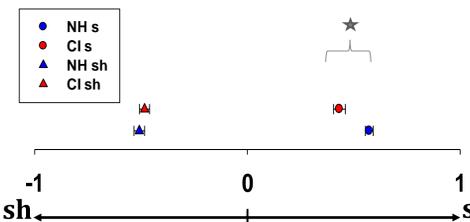
- Tested in a quiet room
- Equipment: laptop and headphones
- Responded using number keys and mouse

### Analysis

- Removed reaction time and scale rating outliers
- Took the log of reaction time
- Normalized each listener's scale ratings

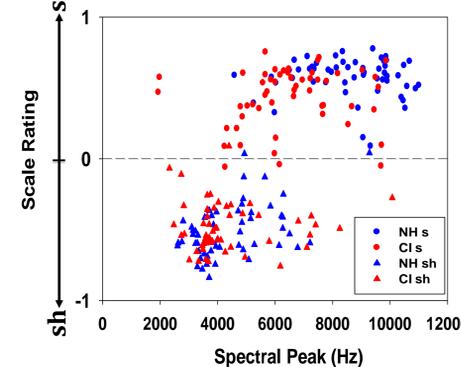
## RESULTS

### A. Visual Analog Scale Ratings



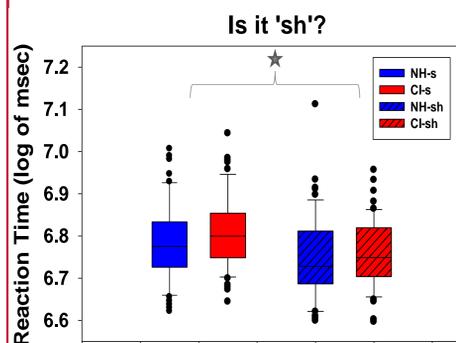
Listeners rated **CI-'s'** tokens closer to the midline than **NH-'s'** tokens ( $p < .01$ ). There was no difference in rating for 'sh' tokens. This mirrors the acoustic findings.

### B. Linking Acoustics and Scale Rating



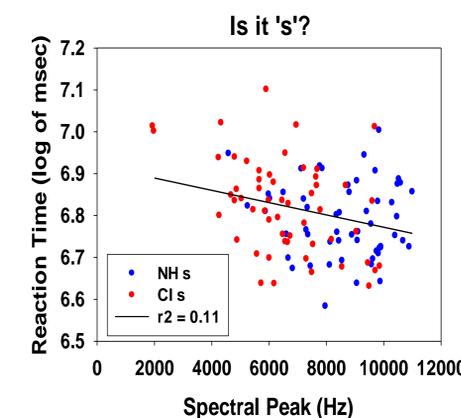
Spectral peak predicted scale rating for 'sh' only ( $p = .017$ ), not 's' for either hearing type. Listeners rated **CI-'s'** with greater variability than they rated **NH-'s'** tokens.

### C. Reaction Times



For the question *Is it 'sh'?* listeners responded more slowly to 's' tokens by both hearing type groups ( $p < .01$ ). This result was not significant for *Is it 's'?* Listeners did not respond slowest to **CI-'s'**.

### D. Linking Acoustics and Reaction Time

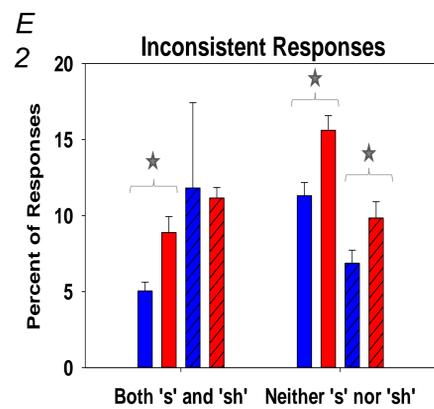
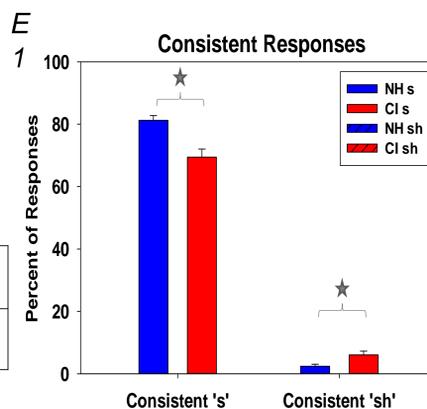


Listeners responded faster to 's' tokens with higher spectral peaks (more characteristic of an 's';  $p < .01$ ). Spectral peak did not predict reaction time for 'sh' tokens. Listeners displayed this pattern for both questions *Is it 's'?* and *Is it 'sh'?*

### E. Listener Response Consistency Between Tasks

Because listeners judged each sound twice, we could assess listener consistency.

| Is it 'sh'?       | Is it 's'?        |                      |
|-------------------|-------------------|----------------------|
|                   | YES               | NO                   |
| Consistent 's'    | Consistent 's'    | Neither 's' nor 'sh' |
| Both 's' and 'sh' | Both 's' and 'sh' | Consistent 'sh'      |



Listeners responded *less consistently* to **CI** tokens than to **NH** tokens ( $p < .01$ ). Figures E1 and E2 suggest that **CI** tokens were less perceptually distinct than **NH** tokens.

## SUMMARY

### Addressing Hypotheses

- Adults rated the **CI-'s'** productions closer to the middle of the scale than **NH-'s'**, mirroring the acoustic findings as predicted.
- Adults did NOT respond slowest to **CI-'s'**. Listeners did respond slower to 's' than 'sh' in one task, and slower to 's' tokens with lower spectral peaks in both timed tasks. This discrepancy may suggest a lack of adequate power.

### Additional Findings

- Listeners identified **CI** tokens less accurately and consistently than they identified **NH** tokens, suggesting that **CI** tokens may be less perceptually distinct.
- Reaction time and scale rating both had some relationship to spectral peak: spectral peak predicted scale rating for 'sh' and reaction time for 's'.

## DISCUSSION

Scale ratings showed that the subtle acoustic differences between correct 's' productions by children with **NH** and children with **CIs** were perceptible to adults.

For children with **CIs**, it is possible that clinicians should not rely only on transcription accuracy. A production may be correct, but the child may not have a robust contrast.

Using additional measures such as a visual analog scale in the clinic may be important. Future research is needed to determine whether this finding can explain the reduced intelligibility of speech of children with **CIs** relative to that of children with **NH**.

### REFERENCES

- Todd, A. E., Edwards, J. R., Litovsky, R. Y., (2011). Production of contrast between sibilant fricatives by children with cochlear implants. *J. Acoust. Soc. Am.*, 130, 3969-3979.
- Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *J. Acoust. Soc. Am.*, 109, 1181-1196.

### ACKNOWLEDGEMENTS

We would like to thank the listeners who participated in this study, as well as the children who produced the stimuli and their families. We would also like to thank the members of the Learning To Talk lab for their support and input. This work was supported by a Hilldale Undergraduate Research Fellowship to Sara Bernstein and by NIDCD grants 02932 to Jan Edwards and 008365 to Ruth Litovsky.