Revealing phonetic uncertainty in cochlear implant listeners

UNIVERSITY of WASHINGTON Listen Lab

Steven Gianakas Matthew B. Winn

Conference on Implantable Auditory Prostheses 2017

INTRODUCTION

When the speech signal is degraded, listeners need to guess certain phonemes and words they missed in the conversation.

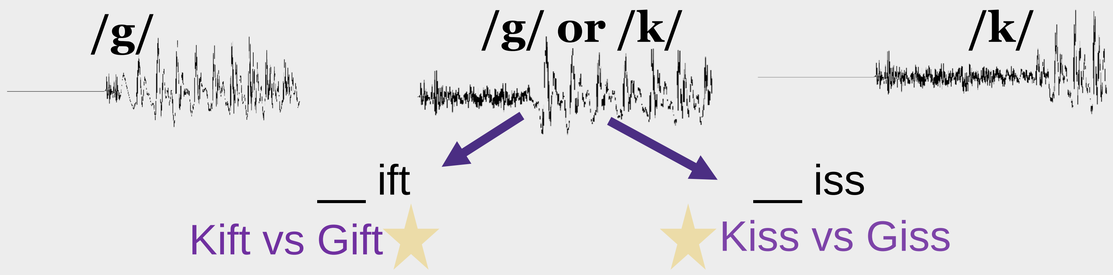
The speech signal becomes compromised when a person uses a cochlear implant (CI). CI listeners might need to guess more often.

In this study we are studying the tendency to resolve phonetic ambiguities using experience (lexical bias) rather than your ears (the signal)

The Ganong Effect

The Ganong effect is the tendency to perceive an ambiguous speech sound as a phoneme that would complete a real word, rather than completing a nonsense/fake word. [1]

Example:



A sound that could be heard as either /g/ or /k/ tends to be perceived as /g/ when followed by "ift" and as /k/ when followed by "iss", presumably because those sound sequences produce real words

This shows

Even though the acoustic signals (g/k) are the same across contexts, perception changes; the effect reflects top-down influence rather than pure reliance on the signal.

This is influenced by...

- **Ambiguity of Speech [1, 2]**
- Frequency of word in spoken language
- Semantic Context [4]
- Phonotactic probability [5]
- Lexical context [6]
- Stimulus blocking [7]

Why is this important?

- When speech sounds are ambiguous, we rely more heavily on our lexical knowledge
- This situation is probably very common for people with hearing loss
- **Hypothesis:** Degrading speech stimuli as if listening with a hearing loss or cochlear implant, should render the speech more ambiguous, thus yielding an increased reliance on top-down processing, seen as an increased "Ganong effect".

METHODS

PARTICIPANTS: 43 young listeners with normal hearing (ages 18-35y)
10 listeners with cochlear implants (ages 44-87 y)

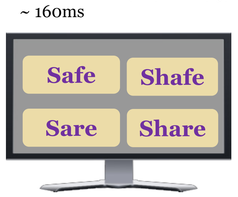
STIMULI: Three 7-step speech continua of spectral phonetic contrasts:

PROCEDURE: Listeners selected the word they heard out of four choices

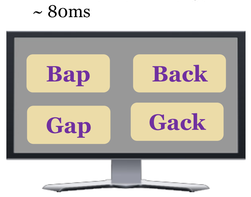
Slow: /æ/ vs /ɑ/ ~ 240ms

If the word ends in /t/, you'll think it's "dot"; if it ends in /ʃ/, you'll think its dash, even if it's the SAME vowel being heard

Medium: /s/ vs /ʃ/ ~ 160ms



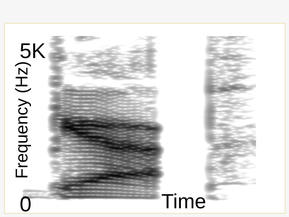
Fast: /b/ vs /g/ ~ 80ms



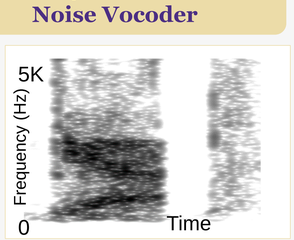
Stimulus Creation

- All stimuli were made using modified natural speech sounds
- For all stimuli, phonetic environments were kept consistent across stimulus sets using a cross-fading/ splicing
- Formants for /æ/-/ɑ/ and /b/-/g/ modified using LPC decomposition in Praat (see Winn & Litovsky 2015)
- /s/ - /ʃ/ continuum made from natural tokens of these fricatives combined with gradual attenuation
- Words were chosen from the HML database to control for familiarity and frequency in the English language

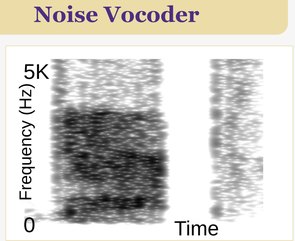
Unprocessed (normal) speech



24 Channel Noise Vocoder



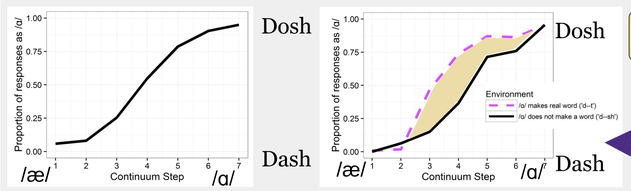
8 Channel Noise Vocoder



When the spectrum is degraded, Formants smear/merge together

ANALYSIS

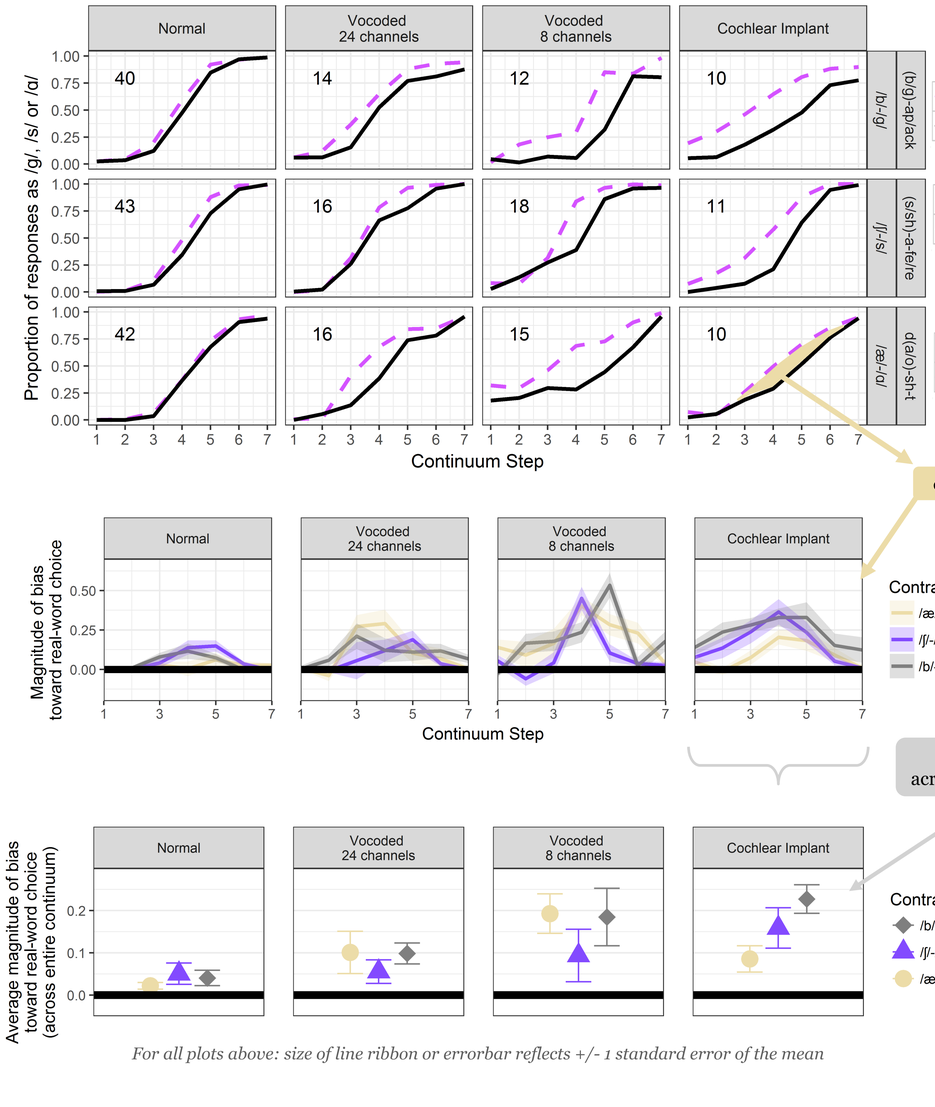
Understanding the perceptual shift



The space between dashed and solid lines reflects the influence of lexical bias.

If dashed line is higher than solid line, listener shows bias to hear phoneme as sound that makes a real word

RESULTS



People with cochlear implants show increased lexical bias with faster phonetic contrasts

Lexical bias effect for /s/ vs /ʃ/ and /b/ vs /g/ were both greater than /æ/ vs /ɑ/ ($p < 0.001$)
Effect for /b/ vs /g/ appears to be largest, but was not statistically different than that for /s/ vs /ʃ/ ($p = 0.24$)

No statistically significant difference in effect size for NH listeners

For all plots above: size of line ribbon or errorbar reflects +/- 1 standard error of the mean

CONCLUSIONS

➢ Phonetic perception is informed by lexicon knowledge (the same sound is perceived differently depending on lexical status)

➢ Listeners tend to rely more heavily on lexical knowledge when the auditory signal is spectrally degraded.

➢ CI listeners show greater influence of lexical knowledge across a wider range of stimuli, indicating that a wider range of sounds is acoustically (electrically) ambiguous

➢ CI listeners show greater dependence on lexical knowledge when acoustic spectral cues are shorter in duration.

[1] Ganong W (1980) Phonetic Categorization in Auditory Word perception. *J Exper Psych: Hum Perc Perf*
[2] Burton M. (1995) Lexical Effects on Phonetic Categorization: The Role of Stimulus Naturalness and Stimulus Quality. *J Exper Psych: Hum Perc Perf*
[3] Morton J. (1979). Word recognition IN Morton J. & Marshall J.C. (Eds.). *Psycholinguistics 2: Structure and processes*
[4] Borsky S., Tuller B., & Shapiro L.P. (1998) The effects of semantic and acoustic information on phoneme categorization. *JASA*
[5] McClelland J.L., Mirman D., Holt L.L. (2006) Are there interactive processes in speech perception? *Trends in Cognitive Sci*
[6] Magnuson J, McMurray B, Tanenhaus M, & Aslin R. (2003) Lexical effects on compensation for articulation: a tale of two systems? *Cognitive Science*
[7] Pitt M.A. & Samuel A.G. (1993) An empirical and meta-analytic evaluation of the phoneme identification task. *J Exper Psych: Hum Perc Perf*

www.UWListenLab.com

Ashley Moore, David J Audet Jr., and Tiffany Mitchel assisted with data collection.
Financial Support provided NIH-NIDCD 1R03 DC014309-01A1 (Winn), NIH Loan Repayment Program (Winn)