#### Neural Representations of the Cocktail Party in Human Auditory Cortex

#### Jonathan Z. Simon

Department of Biology Department of Electrical & Computer Engineering Institute for Systems Research University of Maryland

http://www.isr.umd.edu/Labs/CSSL/simonlab

CHSCOM2015, 15 June 2015

## Acknowledgements

#### Current (Simon Lab & Affiliates)

Francisco Cervantes Natalia Lapinskaya Mahshid Najafi Alex Presacco Krishna Puvvada Lisa Uible

Peng Zan

#### Past (Simon Lab & Affiliate Labs)

Nayef Ahmar Sahar Akram Murat Aytekin Claudia Bonin Maria Chait Marisel Villafane Delgado Kim Drnec Nai Ding

Victor Grau-Serrat Julian Jenkins David Klein Ling Ma Kai Sum Li Huan Luo Raul Rodriguez Ben Walsh Juanjuan Xiang Jiachen Zhuo

#### Collaborators

Pamela Abshire Samira Anderson Behtash Babadi

Catherine Carr Monita Chatterjee Alain de Cheveigné Didier Depireux Mounya Elhilali Bernhard Englitz Jonathan Fritz Cindy Moss David Poeppel Shihab Shamma

#### Past Postdocs & Visitors

Aline Gesualdi Manhães Dan Hertz Yadong Wang

#### Undergraduate Students

Abdulaziz Al-Turki Nicholas Asendorf Sonja Bohr Elizabeth Camenga **Corinne Cameron** Julien Dagenais Katya Dombrowski Kevin Hogan Kevin Kahn Alexandria Miller Isidora Ranovadovic Andrea Shome Madeleine Varmer **Ben Walsh** 

#### Funding NIH (NIDCD, NIA, NIBIB); USDA

## The Cocktail Party



## The Cocktail Party



## The Cocktail Party



#### Outline

- Cortical Representations of Speech (Encoding vs. Decoding)
- Attended vs. Unattended Speech
- Studies in Progress:
  - Attentional Dynamics
  - Aging & Neural Representations of Speech

#### Magnetoencephalography

- Non-invasive, Passive, Silent Neural Recordings
- Simultaneous Whole-Head Recording (~200 sensors)
- Sensitivity
  - high: ~100 fT (10<sup>-13</sup> Tesla)
  - low:  $\sim 10^4 \sim 10^6$  neurons
- Temporal Resolution: ~1 ms
- Spatial Resolution
  - coarse: ~ I cm
  - ambiguous



#### Time Course of MEG Responses

#### **Auditory Evoked Responses**

- MEG Response Patterns Time-Locked to Stimulus Events
- Robust
- Strongly Lateralized





## MEG Responses to Speech Modulations



# MEG Responses Predicted by STRF Model



## Neural Reconstruction of Speech Envelope



## Neural Reconstruction of Speech Envelope



Ding & Simon, J Neurophysiol (2012) Zion-Golumbic et al., Neuron (2013) Reconstruction accuracy comparable to single unit & ECoG recordings



## Neural Representation of Speech: Temporal



# Speech in Noise



Ding & Simon, J Neuroscience (2013)

# Speech in Noise



Ding & Simon, J Neuroscience (2013)

Neural Reconstruction of Underlying Speech Envelope



Neural Reconstruction of Underlying Speech Envelope



Neural Reconstruction of Underlying Speech Envelope



**Contrast Index** 



Ding & Simon, J Neuroscience (2013)

Neural Reconstruction of Underlying Speech Envelope





**Reconstruction Accuracy** 



Neural Reconstruction of Underlying Speech Envelope





correlation

**Reconstruction Accuracy** 



Ding & Simon, J Neuroscience (2013)

Neural Reconstruction of Underlying Speech Envelope





Ding & Simon, J Neuroscience (2013)

Correlation with Intelligiblity

# Cortical Speech Representations

- Neural Representations: Encoding & Decoding
- Linear models: Useful & Robust
- Speech Envelope only (as seen by MEG)
- Envelope Rates: ~ I I0 Hz

## Auditory Objects at the Cocktail Party



## Auditory Objects at the Cocktail Party



### Experiments





## Selective Neural Encoding



# Selective Neural Encoding





# Unselective vs. Selective Neural Encoding





# Selective Neural Encoding









## Stream-Specific Representation



Identical Stimuli!

Ding & Simon, PNAS (2012)

## Single Trial Speech Reconstruction



Ding & Simon, PNAS (2012)

## Single Trial Speech Reconstruction



# Invariance Under Relative Loudness Change?



# Invariance Under Relative Loudness Change?





# Invariance under Relative Loudness Change



- Neural representation invariant to relative loudness change
- Stream-based Gain Control, not stimulus-based

### Forward STRF Model



Spectro-Temporal Response Function (STRF)

### Forward STRF Model

Spectro-Temporal Response Function (STRF)

#### STRF Results



STRF separable (time, frequency)
300 Hz - 2 kHz dominant carriers
M50<sub>STRF</sub> positive peak
M100<sub>STRF</sub> negative peak

### STRF Results



time (ms)

### STRF Results



- STRF separable (time, frequency)
  300 Hz 2 kHz dominant carriers
  M50<sub>STRF</sub> positive peak
  M100<sub>STRF</sub> negative peak
- •M100<sub>STRF</sub> strongly modulated by attention, *but not M50<sub>STRF</sub>*



## Neural Sources

- •M100<sub>STRF</sub> source near (same as?) M100 source: Planum Temporale
- •M50<sub>STRF</sub> source is anterior and medial to M100 (same as M50?): Heschl's Gyrus



•PT strongly modulated by attention, *but not HG* 

## Studies In Progress

- Attentional Dynamics
- Aging & Neural Representations of Speech

#### Attentional Dynamics

Attend to Speaker 1



#### Attentional Dynamics

Attend to Speaker 1



## Younger vs. Older Listeners

**Older Adults** Younger Adults Speech Reconstruction \*\* \*\* In Quiet 0.25 0.25 \* In Quiet 0.2 0.2 with with Competing Competing Speaker 0.15 0.15 Speaker ٦ 0.1 0.1 500 400 300 200 100 500 400 300 200 100

Integration window (ms)

## Younger vs. Older Listeners



Integration window (ms)

## Summary

- Cortical representations of speech
  - representation of envelope (up to ~10 Hz)
- Cortical Processing Hierarchy: Consistent with being neural representation of auditory perceptual object
- Object representation at 100 ms latency (PT), but not by 50 ms (HG)

#### Thank You