### Cortical Encoding of Auditory Objects at the Cocktail Party

Jonathan Z. Simon University of Maryland

Computational Audition Workshop, June 2013

### Introduction

- Auditory Objects
- Magnetoencephalography (MEG)
- Neural Representations of Auditory Objects in Cortex: Decoding
- Neural Representations of Auditory Objects in Cortex: Encoding

# Auditory Objects

- What is an auditory object?
  - perceptual construct (not neural, not acoustic)
  - commonalities with visual objects
  - several potential formal definitions

### Auditory Object Definition

- Griffiths & Warren definition:
  - corresponds with something in the sensory world
  - object information separate from information of rest of sensory world
  - abstracted: object information generalized over particular sensory experiences

















Ding & Simon, PNAS (2012)

### Experiments











# Magnetoencephalography (MEG)



- •Direct electrophysiological measurement
  - not hemodynamic
  - •real-time
- •No unique solution for distributed source
- Measures spatially synchronized cortical activity
- •Fine temporal resolution (~ 1 ms)
- •Moderate spatial resolution (~ 1 cm)

Magnetic MEG orientation Dipolar recording of magnetic Field scalp surface field EEG Projection skull CSF tissue current \_ flow

Photo by Fritz Goro

# Magnetoencephalography (MEG)



Magnetic MEG orientation Dipolar recording of magnetic Field scalp surface field EEG Projection skull CSF tissue current \_ flow

Photo by Fritz Goro

- Direct electrophysiological measurement
  - not hemodynamic
  - •real-time
- •No unique solution for distributed source
- Measures spatially synchronized cortical activity
- Fine temporal resolution (~ 1 ms)
- Moderate spatial recolution (~ 1 cm)

### Phase-Locking in MEG to Slow Temporal Modulations



Ding & Simon, J Neurophysiol (2009) Wang et al., J Neurophysiol (2012)

### Phase-Locking in MEG to Slow Temporal Modulations



MEG activity is precisely phase-locked to temporal modulations of sound

Ding & Simon, J Neurophysiol (2009) Wang et al., J Neurophysiol (2012)



# MEG Responses to Speech Modulations



# MEG Responses Predicted by STRF Model



### Neural Reconstruction of Speech Envelope



## Neural Reconstruction of Speech Envelope



2 s

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



# Speech Stream as an Auditory Object

- corresponds with something in the sensory world
- information separate from information of rest of sensory world
   e.g. other speech streams or noise
- abstracted: object information generalized over particular sensory experiences
   e.g. different sound mixtures

# Neural Representation of an Auditory Object

- neural representation is of something in sensory world
- when other sounds mixed in, neural representation is of auditory object, not entire acoustic scene
- neural representation invariant under broad changes in specific acoustics









# Unselective vs. Selective Neural Encoding





# Unselective vs. Selective Neural Encoding













**Identical Stimuli!** 



**Identical Stimuli!** 

### Single Trial Speech Reconstruction



### Single Trial Speech Reconstruction



### Overall Speech Reconstruction



Distinct neural representations for different speech streams









-









**Gain-Control Models** 



#### **Gain-Control Models**





#### **Gain-Control Models**







#### **Gain-Control Models Object-Based** correlation 2 attended backarounc -8 -5 8 5 ()Speaker Relative Intensity (dB) Stimulus- Based correlation .2 attended background -8 -5 5 8 0

Speaker Relative Intensity (dB)





Stream-based not stimulus-based
Neural representation is invariant to acoustic changes.

# Neural Representation of an Auditory Object

- ✓ neural representation is of something in sensory world
- ✓ when other sounds mixed in, neural representation is of auditory object, not entire acoustic scene
- ✓ neural representation invariant under broad changes in specific acoustics

### Forward STRF Model



Spectro-Temporal Response Function (STRF)

### Forward STRF Model



### 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: PT
- •M50<sub>STRF</sub> source is anterior and medial to M100 (same as M50?): HG



# Cortical Object-Processing Hierarchy



- •M100<sub>STRF</sub> strongly modulated by attention, but not M50<sub>STRF</sub>.
- •M100<sub>STRF</sub> invariant against acoustic changes.
- •Objects well-neurally represented at 100 ms, but not 50 ms.

# Not Just Speech



Elhilali et al., PLoS Biology (2009)







Neural Enhancement for Foreground/Background



### Competing Tone Streams



### Competing Tone Streams



### Competing Tone Streams



# Summary

- Cortical representations of speech found here:
  - consistent with being *neural* representations of auditory (*perceptual*) objects
  - ✓ meet 3 formal criteria for auditory objects
- Object representation fully formed by 100 ms latency (PT), but not by 50 ms (HG)
- Not special to speech

### Acknowledgements

#### **Grad Students**

Francisco Cervantes Marisel Villafane Delgado Kim Drnec Krishna Puvvada

#### **Past Grad Students**

Nayef Ahmar Claudia Bonin Maria Chait Nai Ding Victor Grau-Serrat Ling Ma Raul Rodriguez Juanjuan Xiang Kai Sum Li Jiachen Zhuo

#### Funding NIH R01 DC 008342

#### **Undergraduate Students**

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

#### **Collaborators' Students**

Murat Aytekin Julian Jenkins David Klein Huan Luo

#### Collaborators

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

#### **Past Postdocs**

Dan Hertz Yadong Wang

### Thank You



### Reconstruction of Same-Sex Speech



### Speech in Noise: Stimuli



### Speech in Noise: Results



### Speech in Noise: Results



### Speech in Noise: Results

