Neural Representations of Speech, and Speech in Noise, 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

ASA 2016, 24 May 2016

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

Outline

- Cortical Representations of Speech (via MEG)
 - Encoding vs. Decoding
- Cortical Representations of Speech in Noise
- Recent Studies:
 - Attentional Dynamics
 - Aging & Cortical Representations of Speech
 - Higher Level Interference & Noise

Magnetoencephalography (MEG)

- Non-invasive, Passive, Silent Neural Recordings
- MEG Response Patterns Time-Locked to Stimulus Events
- Robust
- Strongly Lateralized
- Cortical Origin Only





MEG Responses to Speech Modulations



MEG Responses Predicted by STRF Model



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 Stationary Noise



Ding & Simon, J Neuroscience (2013)

Speech in Stationary Noise



Ding & Simon, J Neuroscience (2013)

Speech in Noise: Results

Neural Reconstruction of Underlying Speech Envelope



Speech in Noise: Results

Neural Reconstruction of Underlying Speech Envelope





correlation

Reconstruction Accuracy



Ding & Simon, J Neuroscience (2013)

Speech in Noise: Results

Neural Reconstruction of Underlying Speech Envelope





Ding & Simon, J Neuroscience (2013)

Correlation with Intelligiblity

Noise-Vocoded Speech



Intelligibility Reflected only in Delta Band (I–4 Hz)

Ding, Chatterjee & Simon, NeuroImage (2014)

Multiple Representations?

Di Liberto, et al. (2015) Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing

Kayser et al. (2015) Irregular Speech Rate Dissociates Auditory Cortical Entrainment, Evoked Responses, and Frontal Alpha

Ding et al. (2015) Cortical tracking of hierarchical linguistic structures in connected speech

Cortical Speech Representations

- Neural Representations: Encoding & Decoding
- Linear models: Useful & Robust
- Speech Envelope only (as seen in MEG)
- Envelope Rates: ~ I I0 Hz
- Intelligibility linked to lower range of frequencies (Delta)

Competing Speech Streams



Selective Neural Encoding



Selective Neural Encoding





Unselective vs. Selective Neural Encoding





Selective Neural Encoding









Selective Encoding: Results



Identical Stimuli!

Ding & Simon, PNAS (2012)

Single Trial Speech Reconstruction



Ding & Simon, PNAS (2012)

Single Trial Speech Reconstruction



Reconstruction of Same-Sex Speech



Ding & Simon, PNAS (2012)

STRF Results



attended

0

100

200

time (ms)

400

- •300 Hz 2 kHz dominant carriers
- •M50_{STRF} positive peak
- •M100_{STRF} negative peak
- •M100_{STRF} strongly modulated by attention, *but not M50_{STRF}*

Neural Sources

- •M100_{STRF} source near (same as?) M100 source: Planum Temporale
- •M50_{STRF} source is anterior and medial to M100 (same as M50?): Heschl's Gyrus



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

Recent Studies

- Attentional Dynamics
- Aging & Cortical Representations of Speech
- High Level Interference & Noise

Attentional Dynamics

Attend to Speaker 1



Akram et al., NeuroImage (2016)

Attentional Dynamics

Attend to Speaker 1



Akram et al., NeuroImage (2016)

Younger vs. Older Listeners



Integration window (ms)

High Level Interference



Summary

- Cortical representations of speech
 - representation of envelope (up to ~10 Hz)
 - robust against a variety of noise types
 - neural representation of perceptual object
- Object-based representation at 100 ms latency (PT), but not by 50 ms (HG)
- At least 2 different object-based representations, e.g., delta vs. theta; effect of language; phoneme acoustics vs. perception

Thank You