#### Auditory Cortex Tracks Acoustic Onsets of Ignored Speech: A Potential Mechanism in Stream Segregation Christian Brodbeck, L. Elliot Hong & Jonathan Z. Simon University oMaryland

christianbrodbeck@me.com





# Outline

### Cocktail-party problem

Selectively listening to one of two talkers

## Processing of ignored speech?

### MEG response to competing speakers

- Participants listen to two competing audiobook segments.
- Continuous neural response model
- MEG responses modeled to determine whether features of the ignored speech are represented





## Cocktail party problem



#### Acoustic scene

- Acoustic mixture (acoustic scene, representation in auditory nerve)
- Acoustic sources (speakers)

#### **Cortical representations**

- Early (~50 ms) acoustic mixture (Puvvada & Simon, 2017)
- Later (~100 ms) preferential processing of the attended speech source (Ding & Simon, 2012)

#### Is ignored speech separated from the mixture in auditory cortex?







## Representation of ignored speech



#### Indirect evidence

- Your name may attract attention (Cherry, 1953)
- Background speech is more distracting than other noises (e.g. Brungart, 2001)
- But less so when you don't know the language in the background (Van Engen & Bradlow, 2007)
- Identity priming from unattended words (Rivenez et al., 2006)

Retrospective access to no more than one speaker (Kidd et al., 2005) Hard to distinguish consistent lexical processing from attention switches No time-locked lexical processing based on MEG (Brodbeck et al., 2018)

#### Paradigm

- Two speakers, equal loudness (female & male)
- Instructions: Attend to one, ignore the other
- Task: After each segment, answer a question about the content of the attended stimulus

# Temporal response function (TRF)



### Mathematically

• We model the response (r) as convolution (\*) of the stimulus (s) with a response function (h):

$$r = s * h$$

I.e., each point in the response is a weighted sum of the stimulus preceding it:

$$h_t = \sum_{\tau} h_{\tau} \cdot s_{t-\tau}$$

 Stimulus and response are known, kernel is to be estimated





# Spectro-temporal response function (STRF)



Amplitude in frequency bins Spectrogram

#### Multiple predictor variables

The measured response is the sum of the individual responses

#### **Physiological motivation**

- Neural sub-populations respond to different stimulus features
  - E.g. frequency tuning
- Electrical activity is locally additive

#### **Spectro-temporal response** function (STRF)

- Brain response to acoustic stimulus
- TRF can differ depending on the acoustic frequency











## Single speaker



#### Acoustic onsets

- Acoustic edge detector model (Fishbach, Nelken, & Yeshurun, 2001)
- Relevant for auditory object perception



## Single speaker

Significant prediction







#### **ROI** for TRF analysis



#### **Acoustic onsets**

- Acoustic edge detector model (Fishbach, Nelken, & Yeshurun, 2001)
- Source localization consistent with superior temporal Gyrus



Onsets

Envelope



## Single speaker



#### **Acoustic onsets**

- Acoustic edge detector model (Fishbach, Nelken, & Yeshurun, 2001)
- Source localization consistent with superior temporal Gyrus
- Typical response pattern:
  - + peak
  - peak





## Two speakers

#### Acoustic mixture



#### Potential representations

- Acoustic input (mixture)
- Recovered source signals
  - Attended source
  - Ignored source?

#### Significant responses

- Significant response to onsets in the ignored source
- After accounting for mixture and attended source



Э

## Masked onsets

#### **Masked onset**



**Overt onset** 

#### Intuition

- Sources are represented in addition to mixture
- The auditory cortex has to recover features in the source that are masked in the mixture

#### New predictors

- **Overt onsets:** Onsets in a source that are visible in the mixture
- Masked onsets: Onsets in a source that are masked in the mixture

#### $\rightarrow$ New model

 Overtness (overt, masked) × Source (attended, ignored)



## Masked onsets



#### + peak: Overt > masked

– peak: Attended > ignored

#### **Delayed response to masked** onsets

- Delay not uniform as previously assumed (cf. Ding & Simon, 2013)
- Relation to conscious experience?





## Summary



#### **Increasing abstraction**

- ▶ 74 ms: Bottom-up, stimulus-driven
- ▶ 93 ms: reconstructed onsets
- >120 ms: reconstructed onsets same amplitude as overt onsets
- Increasing selectivity for attended source

#### **Representations of ignored speech could** explain

- Why speech is more distracting than stationary noise
- Intrusions from ignored speech (cf. Brungart, 2001)
- Detection of over-learned words such as one's name (cf. Woods & McDermott, 2018)



# Thank You!

# Acknowledgements

### Advisor

Jonathan Z. Simon

## Experiment design

Krishna Puvvada

### **MEG** data collection

Natalia Lapinskaya

### **Undergraduate students**

- Alex Jiao
- Ross Baehr

## Collaborator

L. Elliot Hong

## Funding

- National Institutes of Health (R01-DC-014085 to J.Z.S.)
- University of Maryland Seed Grant (to L.E.H. and J.Z.S.)

