Computational
Sensorimotor
Systems Lab

Integrating Temporal Response Function Estimation with Sparse Source Localization

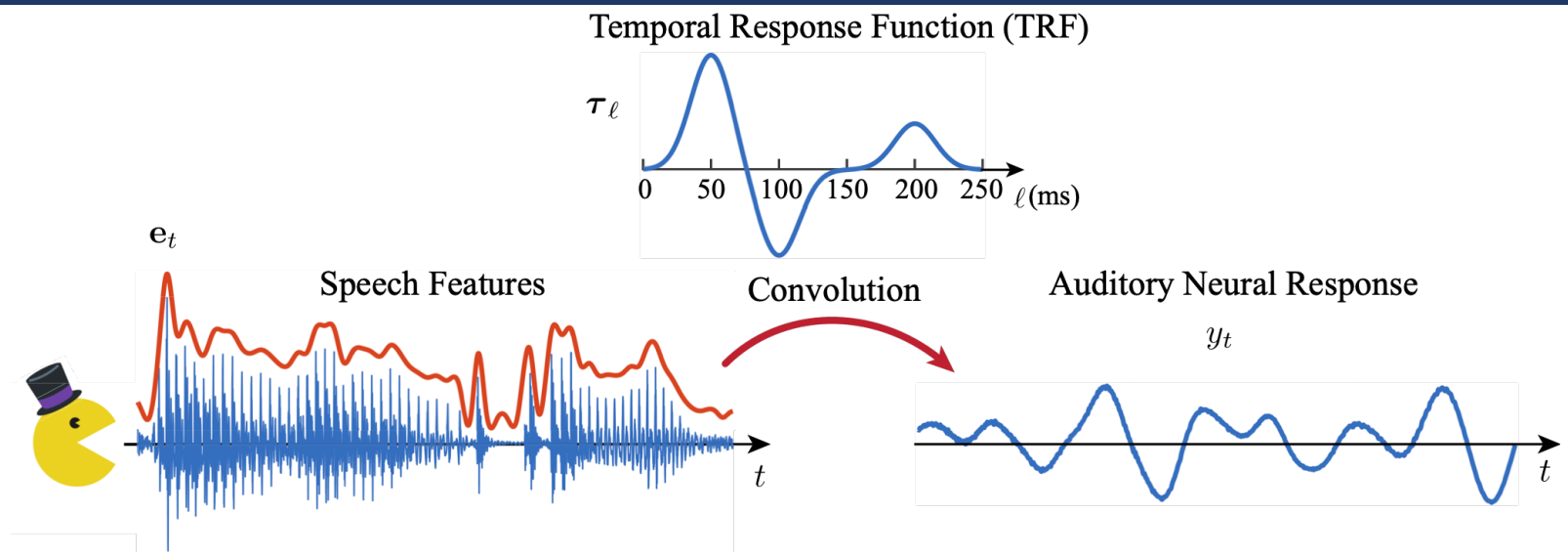
Proloy Das[1], Christian Brodbeck[2], Jonathan Z Simon[3], Behtash Babadi[3]
[1]Massachusetts General Hospital
[2]University of Connecticut
[3]University of Maryland

these slides
available at

Funding:

- National Science Foundation (Award No: 1552946 and 1734892)
- Defense Advanced Research Projects Agency (Award: N6600118240224)
- National Institutes of Health (Award No: R01-DC014085)

Encoding Model of Speech Processing



Linear filter model:

- predict the M/EEG response from continuous stimulus

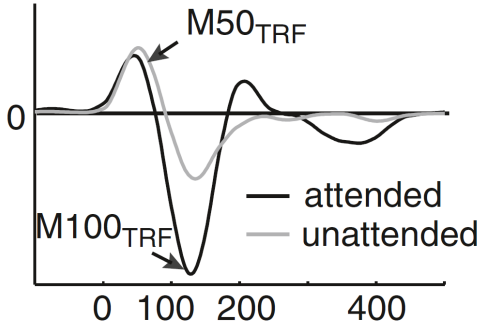
Functional Role:

- Different peaks are associated with different modes of processing
- M50 (c.f. EEG P1): encoding of acoustic-level features
- M100 (c.f. EEG N1): encoding modulated by attention

Beyond acoustic processing:

- phoneme level processing — Di Liberto et. al., 2015
- lexical processing — Brodbeck et. al., 2018
- semantic processing — Broderick et. al., 2018

Lalor et. al., 2006, 2009, 2010



Ding and Simon, 2012

Cortical Origins?

Cortical Origins of TRFs?

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Mesgarani et. al., 2008, 2014

Existing M/EEG approaches works in two-stage:

- Estimate the TRFs for each sensor → localize them within cortex
- Decompose MEG signals to source time courses → estimate TRF for each source

Lalor et. al., 2009

Brodbeck et. al., 2018

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model mismatch

Cortical Origins of TRFs?

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Mesgarani et. al., 2008, 2014

Existing M/EEG approaches works in two-stage:

- Estimate the TRFs for each sensor → localize them on cortical mantle
- Decompose MEG signals to source time courses → estimate TRF for each source

Lalor et. al., 2009

Brodbeck et. al., 2018

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model

MEG observations

Lead-field matrix

Source time courses

$$\mathbf{Y}_{N\times T} = \mathbf{L}_{N\times 3M}\boldsymbol{J}_{3M\times T} + \mathbf{W}_{N\times T}$$

Measurement noise

$\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_w)$

Quasi-static solution to Maxwell's equations

$$\boldsymbol{J}_{3M\times T} = \boldsymbol{\Phi}_{3M\times L}\mathbf{S}_{L\times T} + \mathbf{V}_{3M\times T}$$

Background nuisance

$\mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$

Neuro-current response function "NCRF"

Auditory covariates

N : sensors T : time points M : sources L : lags

Cortical Origins of TRFs?

Insight from intracranial recordings

- Electrophysiology, electrocorticography
- Limited spatial range

Mesgarani et. al., 2008, 2014

Existing M/EEG approaches works in two-stage:

- Estimate the TRFs for each sensor → localize them on cortical mantle
- Decompose MEG signals to source time courses → estimate TRF for each source

Lalor et. al., 2009

Brodbeck et. al., 2018

Challenges:

Bias propagation, spatio-temporal leakage, sensitivity to forward model

MEG observations

Lead-field matrix

Source time courses

$$\mathbf{Y}_{N\times T} = \mathbf{L}_{N\times 3M}\boldsymbol{J}_{3M\times T} + \mathbf{W}_{N\times T}$$

Measurement noise

$\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_w)$

Quasi-static solution to Maxwell's equations

$$\boldsymbol{J}_{3M\times T} = \boldsymbol{\Phi}_{3M\times L}\mathbf{S}_{L\times T} + \mathbf{V}_{3M\times T}$$

Background nuisance

$\mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma})$

Neuro-current response function "NCRF"

Auditory covariates

N : sensors T : time points M : sources L : lags

NCRF Learning: Bayesian Estimation

Measurement noise: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_w)$

$$p\left(\mathbf{Y}|\mathbf{J}\right) = \left|(2\pi)\mathbf{\Sigma}_w\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|^2_{\mathbf{\Sigma}_w^{-1}}\right)$$

Background activity: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$

$$p(\mathbf{V}|\mathbf{\Gamma}) = \left(\prod_{m=1}^{M} \left|(2\pi)\mathbf{\Gamma}_m\right|^{-T/2}\right) \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \|\mathbf{V}_m\|^2_{\mathbf{\Gamma}_m^{-1}}\right)$$

$$p\left(\mathbf{J}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\mathbf{\Gamma}\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|^2_{\mathbf{\Gamma}^{-1}}\right)$$

not observed

Joint density of measurements ($\mathbf{Y}$) and current dipoles ($\mathbf{J}$):

$$p\left(\mathbf{Y}, \mathbf{J}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\mathbf{\Sigma}_w\right|^{-T/2}\left|(2\pi)\mathbf{\Gamma}\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|^2_{\mathbf{\Sigma}_w^{-1}} - \frac{1}{2}\|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|^2_{\mathbf{\Gamma}^{-1}}\right)$$

Marginal density of the measurements given NCRFs (Φ):

$$p\left(\mathbf{Y}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\left(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right)\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}\right)$$

Maximum likelihood estimate of NCRFs (Φ):

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation

Measurement noise: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_w)$

$$p\left(\mathbf{Y}|\mathbf{J}\right) = \left|(2\pi)\mathbf{\Sigma}_w\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|^2_{\mathbf{\Sigma}_w^{-1}}\right)$$

Background activity: Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$

$$p(\mathbf{V}|\mathbf{\Gamma}) = \left(\prod_{m=1}^{M} \left|(2\pi)\mathbf{\Gamma}_m\right|^{-T/2}\right) \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \|\mathbf{V}_m\|^2_{\mathbf{\Gamma}_m^{-1}}\right)$$

$$p\left(\mathbf{J}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\mathbf{\Gamma}\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|^2_{\mathbf{\Gamma}^{-1}}\right)$$

not observed

Joint density of measurements ($\mathbf{Y}$) and current dipoles ($\mathbf{J}$):

$$p\left(\mathbf{Y}, \mathbf{J}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\mathbf{\Sigma}_w\right|^{-T/2}\left|(2\pi)\mathbf{\Gamma}\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{J}\|^2_{\mathbf{\Sigma}_w^{-1}} - \frac{1}{2}\|\mathbf{J} - \mathbf{\Phi}\mathbf{S}\|^2_{\mathbf{\Gamma}^{-1}}\right)$$

Marginal density of the measurements given NCRFs (Φ):

$$p\left(\mathbf{Y}|\mathbf{\Phi}, \mathbf{\Gamma}\right) = \left|(2\pi)\left(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right)\right|^{-T/2} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}\right)$$

Maximum likelihood estimate of NCRFs (Φ):

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

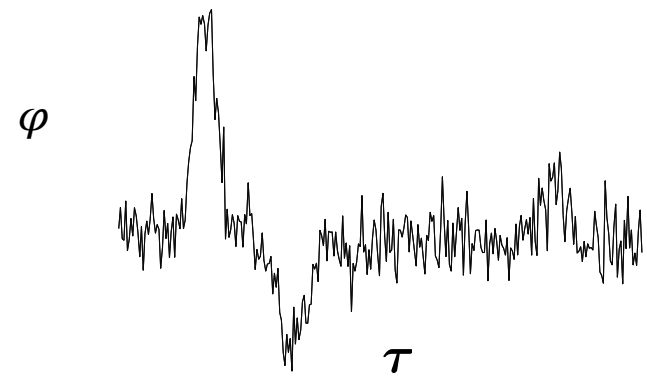
NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}$$

But

- NCRF estimates contain high frequency noise



P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}$$

But

- NCRF estimates contain high frequency noise

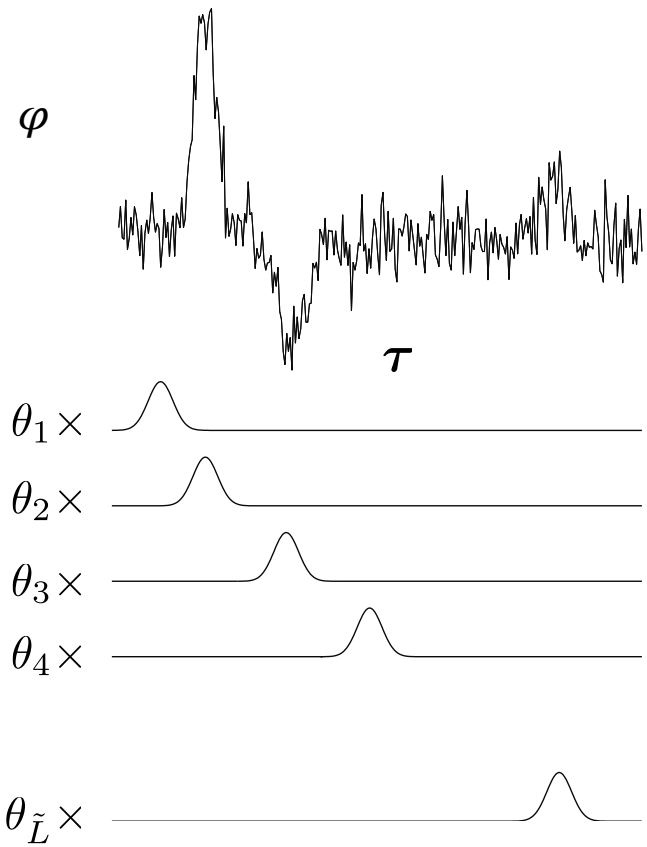
Solution:

Represent $\mathbf{\Phi}$ using a Gabor dictionary, $\mathbf{G}$

NCRF dictionary coefficients

$$\mathbf{\Phi} = \mathbf{\Theta}\mathbf{G}^\top$$
$$\widetilde{\mathbf{S}} := \mathbf{G}^\top\mathbf{S}$$

$$\min_{\Theta} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|_{\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^{\top -1}}$$



P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Phi}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Phi}\mathbf{S}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}}$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation

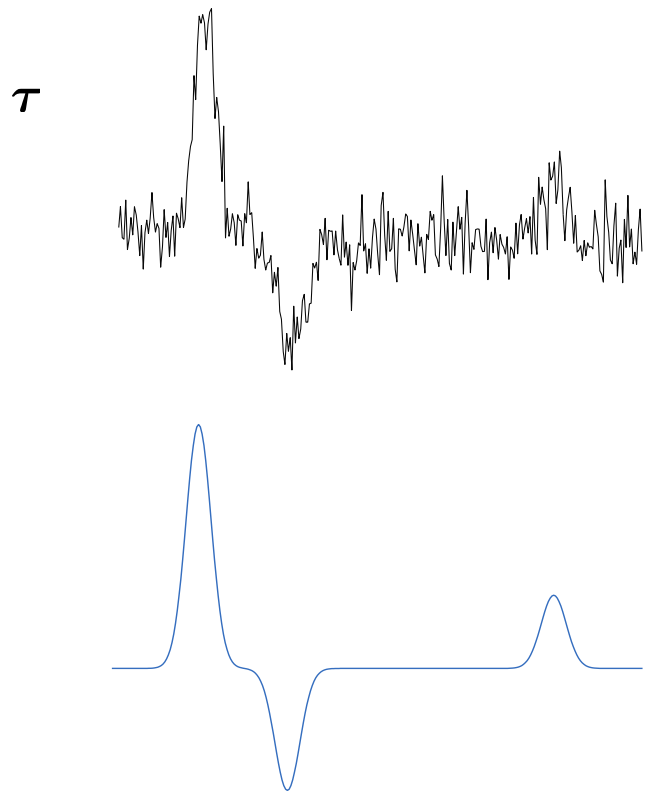
Solution:

Represent $\mathbf{\Phi}$ using a Gabor dictionary, $\mathbf{G}$

NCRF dictionary coefficients

$$\mathbf{\Phi} = \mathbf{\Theta}\mathbf{G}^\top$$
$$\widetilde{\mathbf{S}} := \mathbf{G}^\top\mathbf{S}$$

$$\min_{\Theta} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|_{\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^{\top -1}}$$



P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\Theta} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top}{}^{-1}$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem

  $N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\Theta} \qquad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top}{}^{-1}$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem

  $N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\Theta} \qquad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top}{}^{-1}$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem

  $N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge:

- Most of brain does not process this stimulus

  Enforce spatial sparsity

- TRFs dominated by peaks and troughs

  Enforce lag-domain sparsity

- Dipole currents are vectors

  Enforce coordinate rotational invariance

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\Theta} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|_{\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top}{}^{-1}$$

But

- NCRF estimates contain high frequency noise → Use Gabor dictionary representation
- It is an ill-posed problem → Enforce spatial and lag-domain sparsity

$N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge:

- Most of brain does not process this stimulus

  Enforce spatial sparsity

- TRFs dominated by peaks and troughs

  Enforce lag-domain sparsity

- Dipole currents are vectors

  Enforce coordinate rotational invariance

$$\mathcal{P}_{2,1,1}(\boldsymbol{\Theta}) = \sum_m \sum_l \|\boldsymbol{\theta}_{m,l}\|_2$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

Regularization parameter

$$\min_{\boldsymbol{\Theta}} \ \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|^2_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRF estimates contain high frequency noise → Use Gabor dictionary representation
- It is an ill-posed problem → Enforce spatial and lag-domain sparsity

$N \sim 10^2 \ll M \sim 10^4$

Incorporate prior knowledge:

- Most of brain does not process this stimulus

  Enforce spatial sparsity

- TRFs dominated by peaks and troughs

  Enforce lag-domain sparsity

- Dipole currents are vectors

  Enforce coordinate rotational invariance

$$\mathcal{P}_{2,1,1}(\boldsymbol{\Theta}) = \sum_m \sum_l \|\boldsymbol{\theta}_{m,l}\|_2$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Theta}} \ \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\mathbf{\Gamma}$) is unknown $\longrightarrow$ Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$

Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$:

$$\min_{\mathbf{\Theta},\mathbf{\Gamma}} \quad \frac{T}{2}\log\left|\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Theta}} \ \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\mathbf{\Gamma}$) is unknown $\longrightarrow$ Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$

Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$:

$$\min_{\mathbf{\Theta},\mathbf{\Gamma}} \quad \frac{T}{2}\log\left|\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Theta}} \ \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\mathbf{\Gamma}$) is unknown $\longrightarrow$ Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$

Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$:

$$\min_{\mathbf{\Theta},\mathbf{\Gamma}} \quad \frac{T}{2}\log\left|\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\mathbf{\Theta},\mathbf{\Gamma}} \quad \frac{T}{2}\log\left|\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\mathbf{\Theta}\widetilde{\mathbf{S}}\|^2_{(\mathbf{\Sigma}_w + \mathbf{L}\mathbf{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\mathbf{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\mathbf{\Gamma}$) is unknown $\longrightarrow$ Estimate $\mathbf{\Gamma}$ jointly with $\mathbf{\Theta}$
- Hard to solve the non-convex problem

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \quad \frac{T}{2}\log\left|\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|^2_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\boldsymbol{\Gamma}$) is unknown $\longrightarrow$ Estimate $\boldsymbol{\Gamma}$ jointly with $\boldsymbol{\Theta}$
- Hard to solve the non-convex problem

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

NCRF Learning: Bayesian Estimation (more modeling)

Maximum likelihood estimate:

$$\min_{\boldsymbol{\Theta},\boldsymbol{\Gamma}} \quad \frac{T}{2}\log\left|\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top\right| + \frac{1}{2}\|\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}}\|^2_{(\boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top)^{-1}} + \eta\mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

But

- NCRF estimates contain high frequency noise $\longrightarrow$ Use Gabor dictionary representation
- It is an ill-posed problem $\longrightarrow$ Enforce spatial and lag-domain sparsity
- Nuisance covariance ($\boldsymbol{\Gamma}$) is unknown $\longrightarrow$ Estimate $\boldsymbol{\Gamma}$ jointly with $\boldsymbol{\Theta}$
- Hard to solve the non-convex problem $\longrightarrow$ Champ Lasso algorithm

Idea: coordinate descent approach

Initialize $\boldsymbol{\Theta}^{(0)} = \mathbf{0}$

Champ-Lasso Algorithm

$$\mathbf{C}_v{}^{(r)} = \frac{1}{T}(\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}^{(r)}\widetilde{\mathbf{S}})(\mathbf{Y} - \mathbf{L}\boldsymbol{\Theta}^{(r)}\widetilde{\mathbf{S}})^\top$$

- Update $\boldsymbol{\Gamma}$

$$\boldsymbol{\Gamma}^{(r+1)} = \underset{\boldsymbol{\Gamma}}{\operatorname{argmin}} \quad \operatorname{tr}\left(\boldsymbol{\Sigma}_v^{-1}\mathbf{C}_v^{(r)}\right) + \log|\boldsymbol{\Sigma}_v|$$
$$\text{s.t. } \boldsymbol{\Sigma}_v = \boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top$$

Champagne
Wipf et. al, 2010

- Update $\boldsymbol{\Theta}$

$$\boldsymbol{\Theta}^{(r+1)} = \underset{\boldsymbol{\Theta}}{\operatorname{argmin}} \quad \frac{1}{2}\|\mathbf{L}\boldsymbol{\Theta}\widetilde{\mathbf{S}} - \mathbf{Y}\|^2_{\boldsymbol{\Sigma}_v{}^{(r+1)^{-1}}} + \eta\mathcal{P}_{2,1,1}(\boldsymbol{\Theta})$$

Forward-Backward splitting
Nesterov, 2005
Beck & Teboulle, 2009

$$\boldsymbol{\Sigma}_v{}^{(r+1)} = \boldsymbol{\Sigma}_w + \mathbf{L}\boldsymbol{\Gamma}^{(r+1)}\mathbf{L}^\top$$

Code available on Github

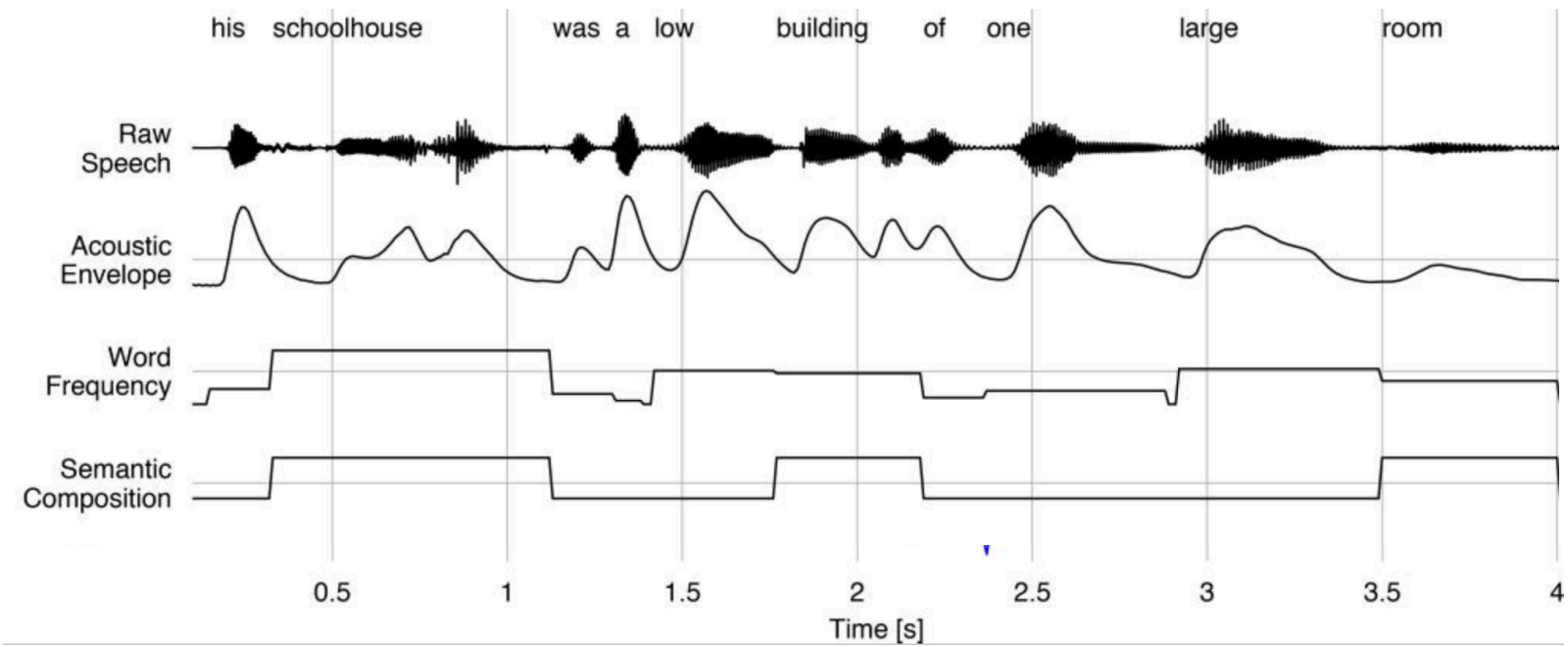
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

MEG Data:

Brodbeck & Simon, 2018

- 17 young-adult participants.
- Two 60-second segments from 'The Legend of Sleepy Hollow' by W. Irving.
- 3 repetitions for each segment.

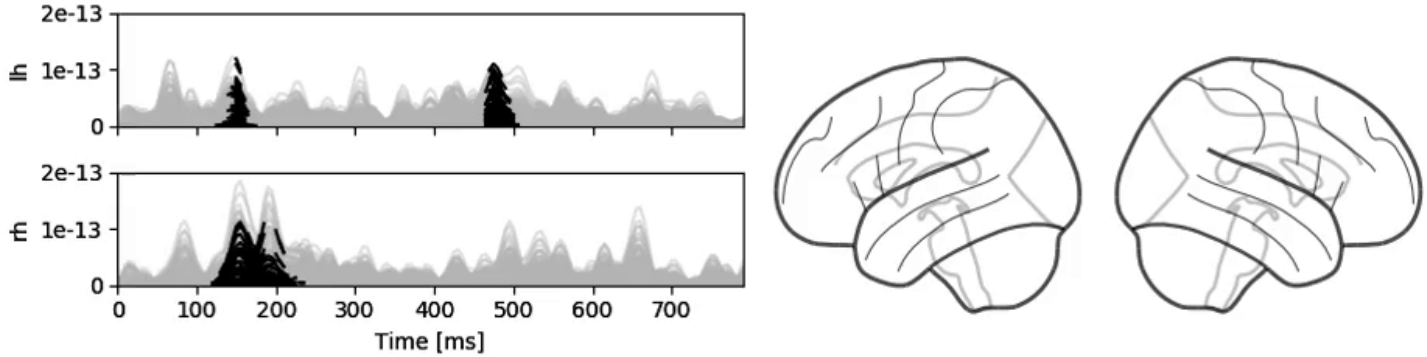


- No MRI was available:
  'fsaverage' morphed in individual head shape

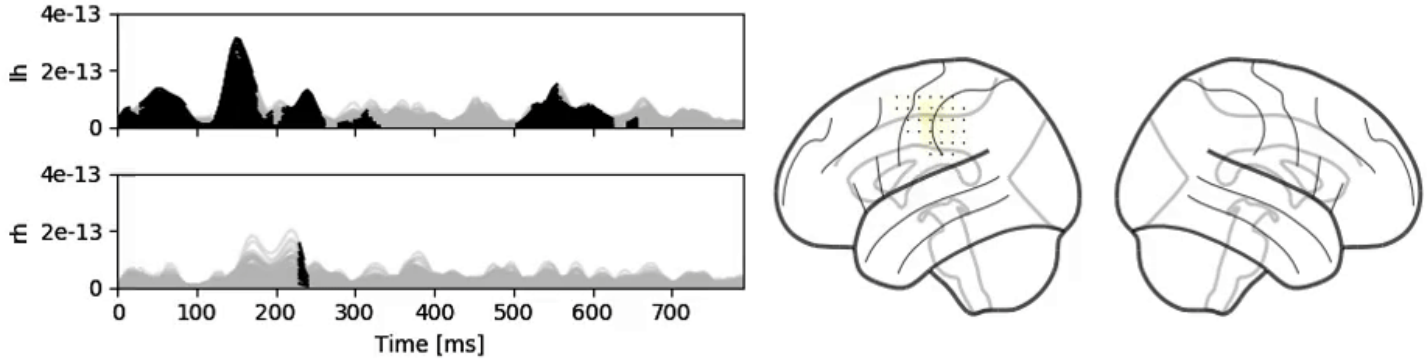
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

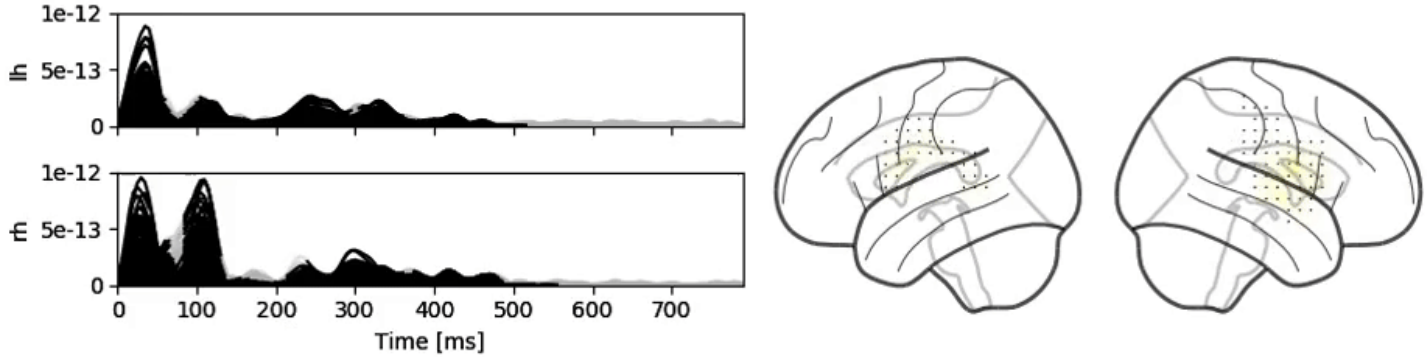
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



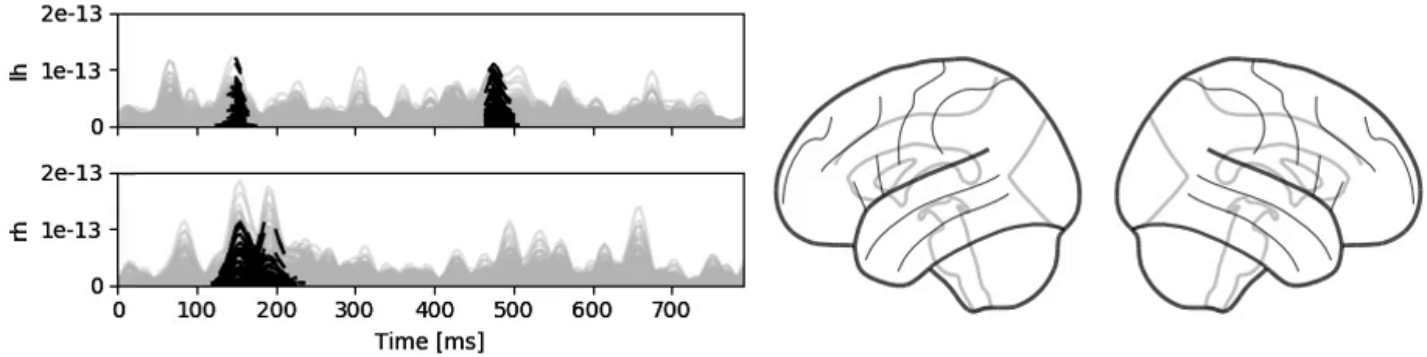
Acoustic envelope
Yang et. al., 1992



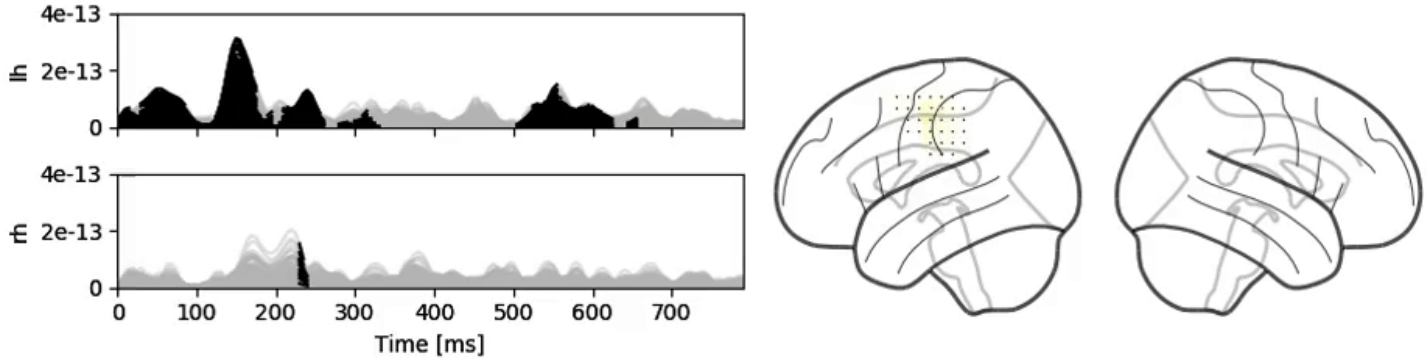
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

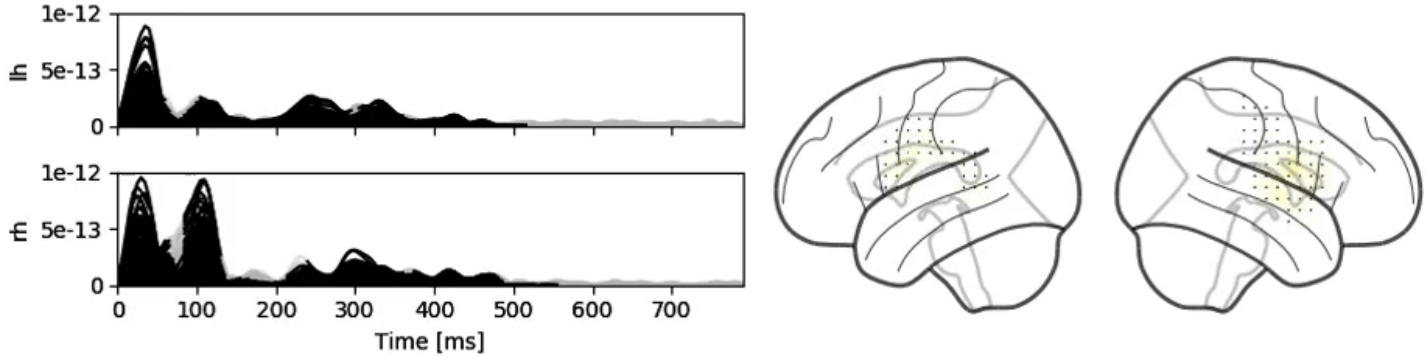
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



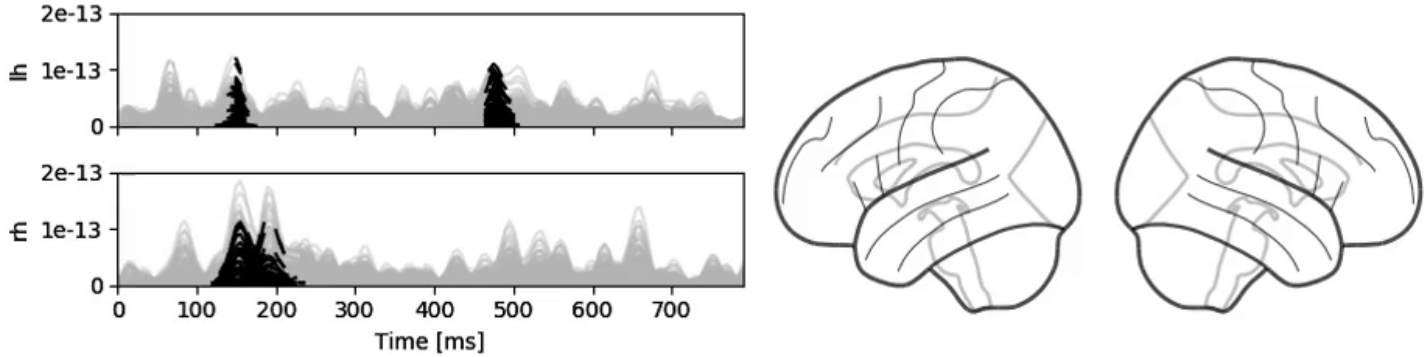
Acoustic envelope
Yang et. al., 1992



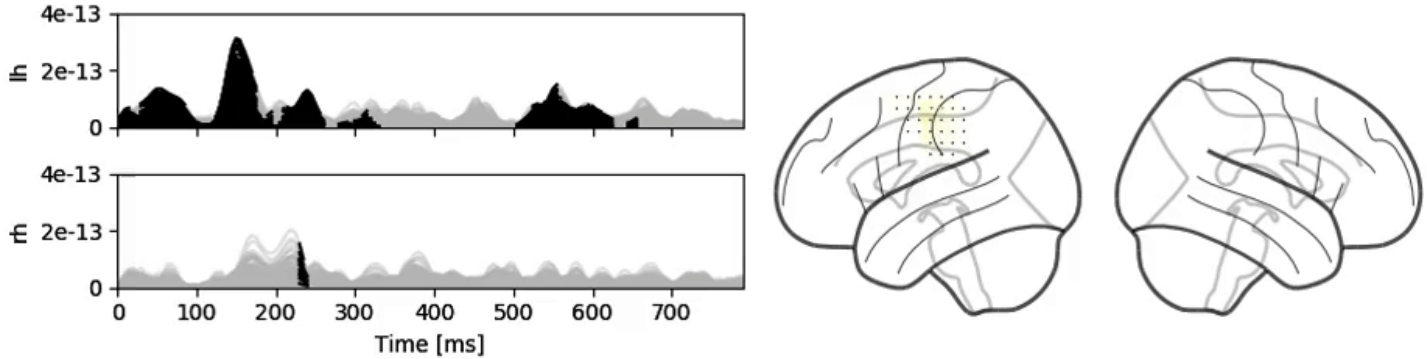
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

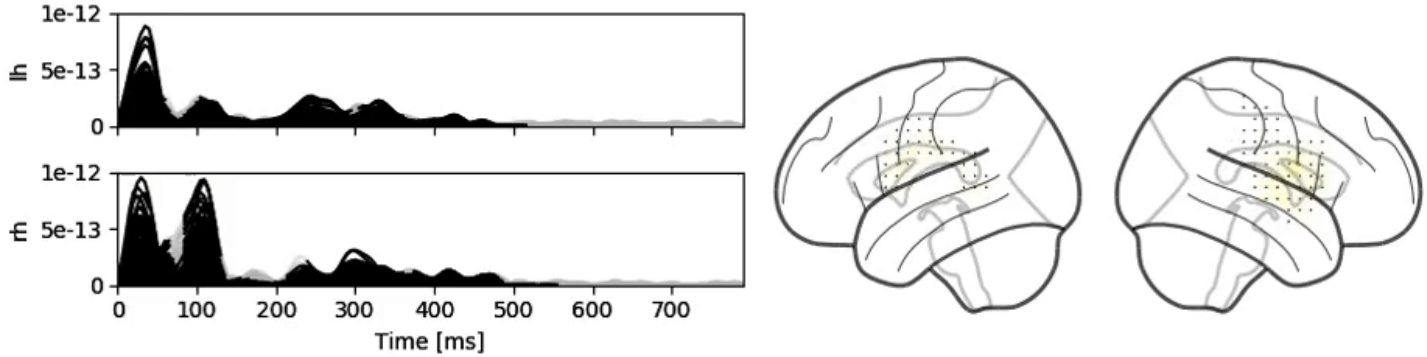
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



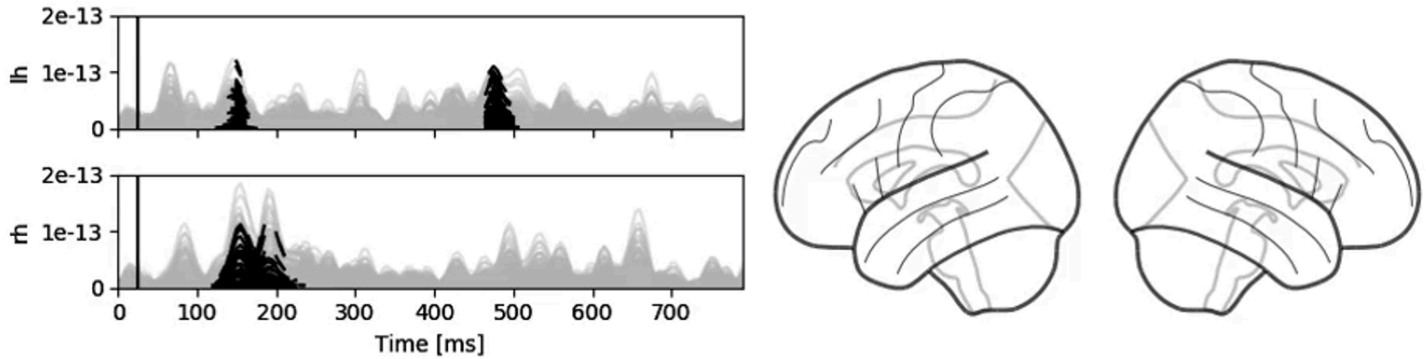
Acoustic envelope
Yang et. al., 1992



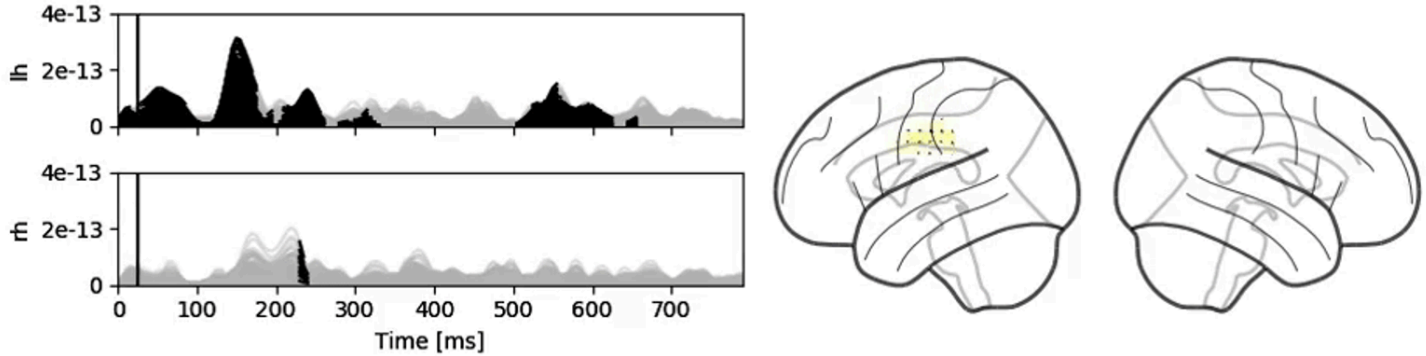
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

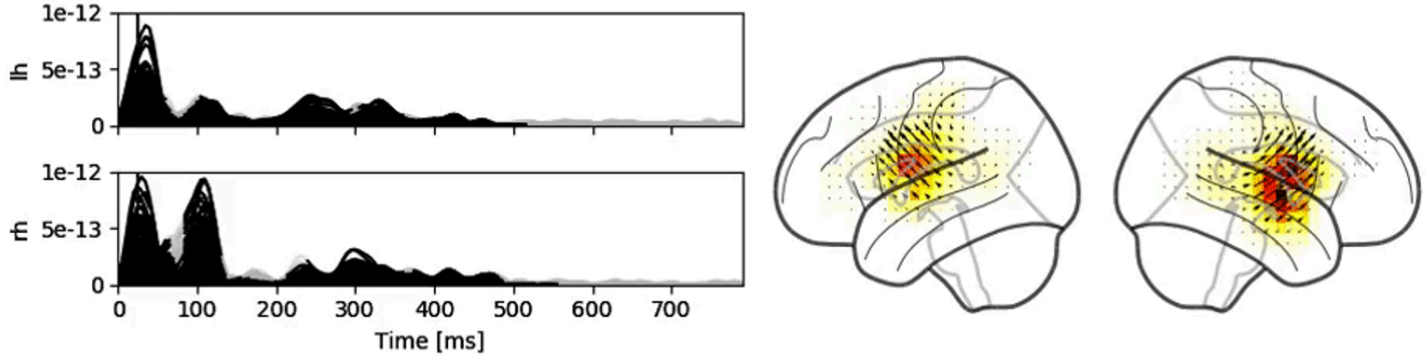
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



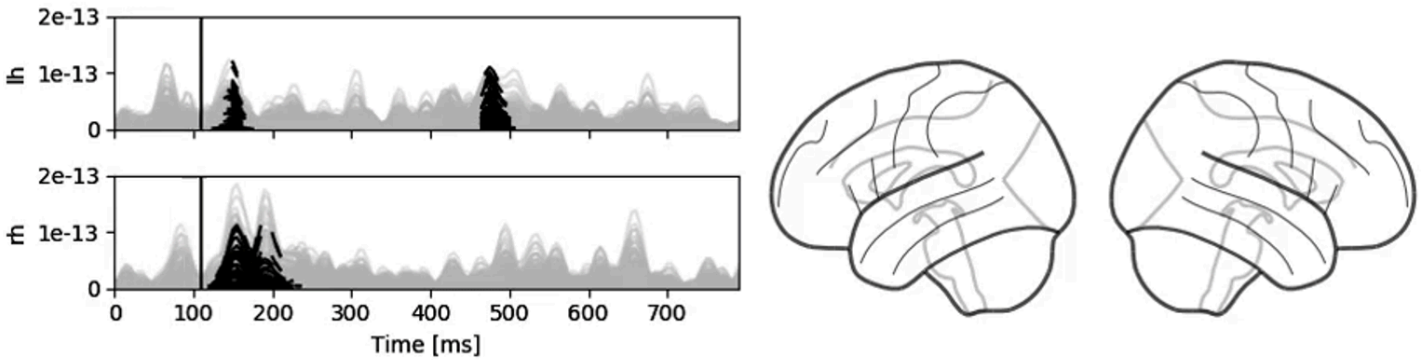
Acoustic envelope
Yang et. al., 1992



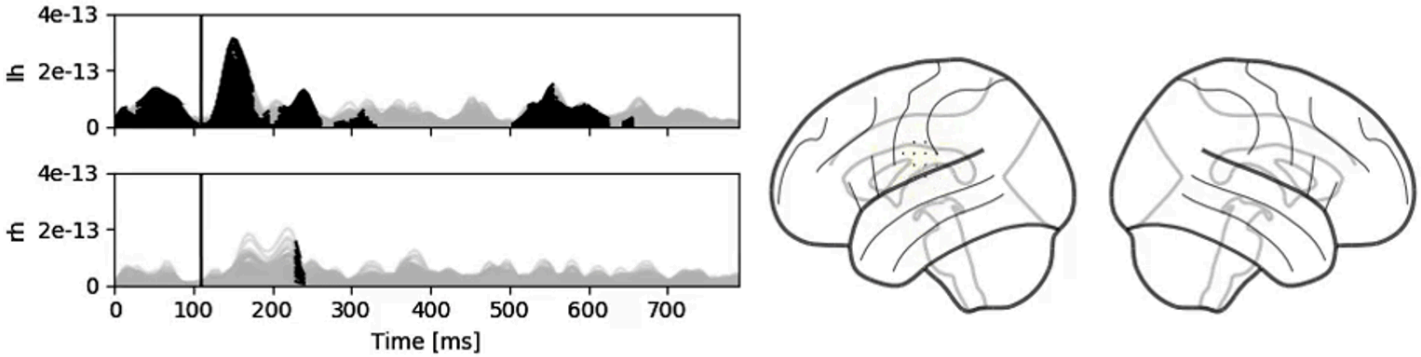
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

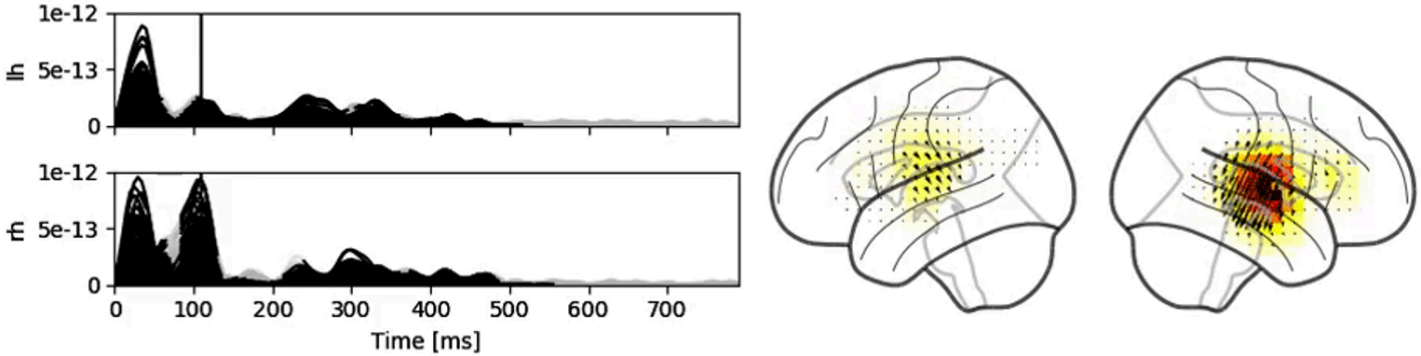
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



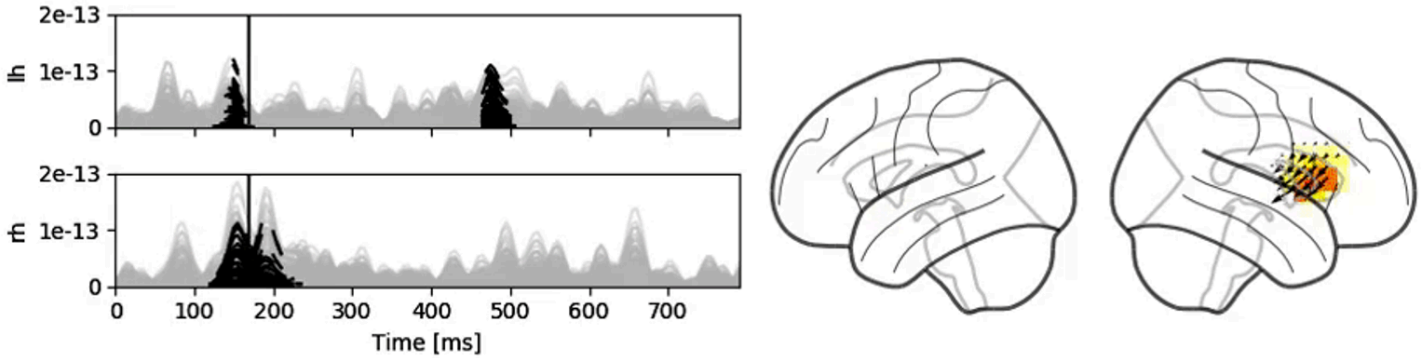
Acoustic envelope
Yang et. al., 1992



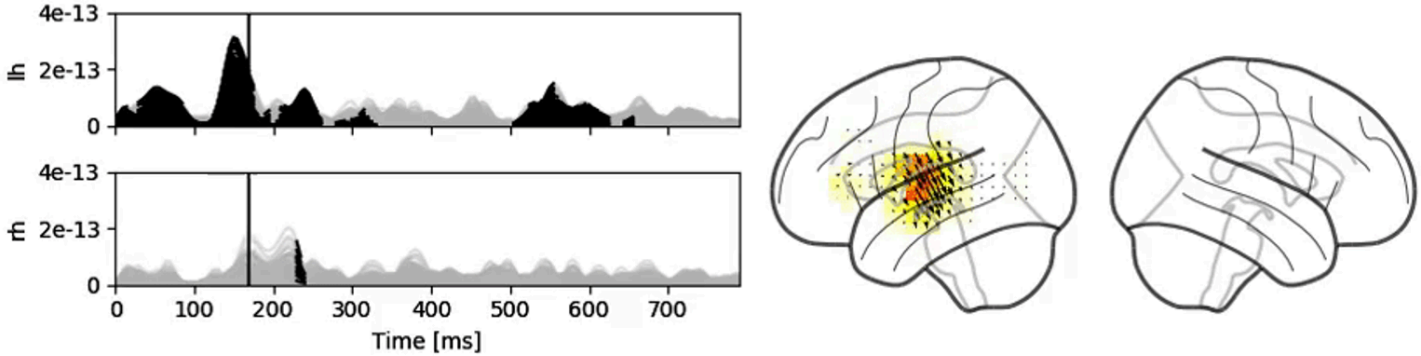
P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

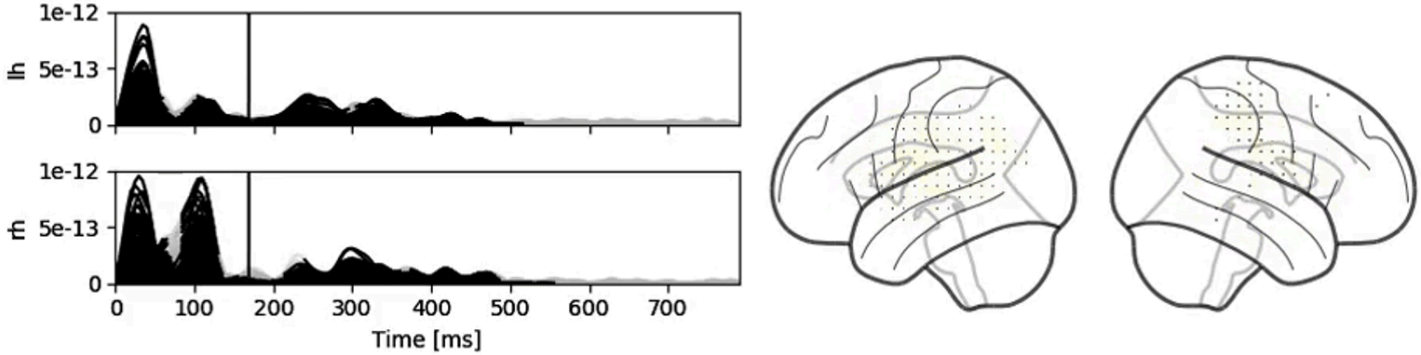
Semantic composition
Westerlund et. al., 2015



Word frequency
Brysbaert et. al., 2009



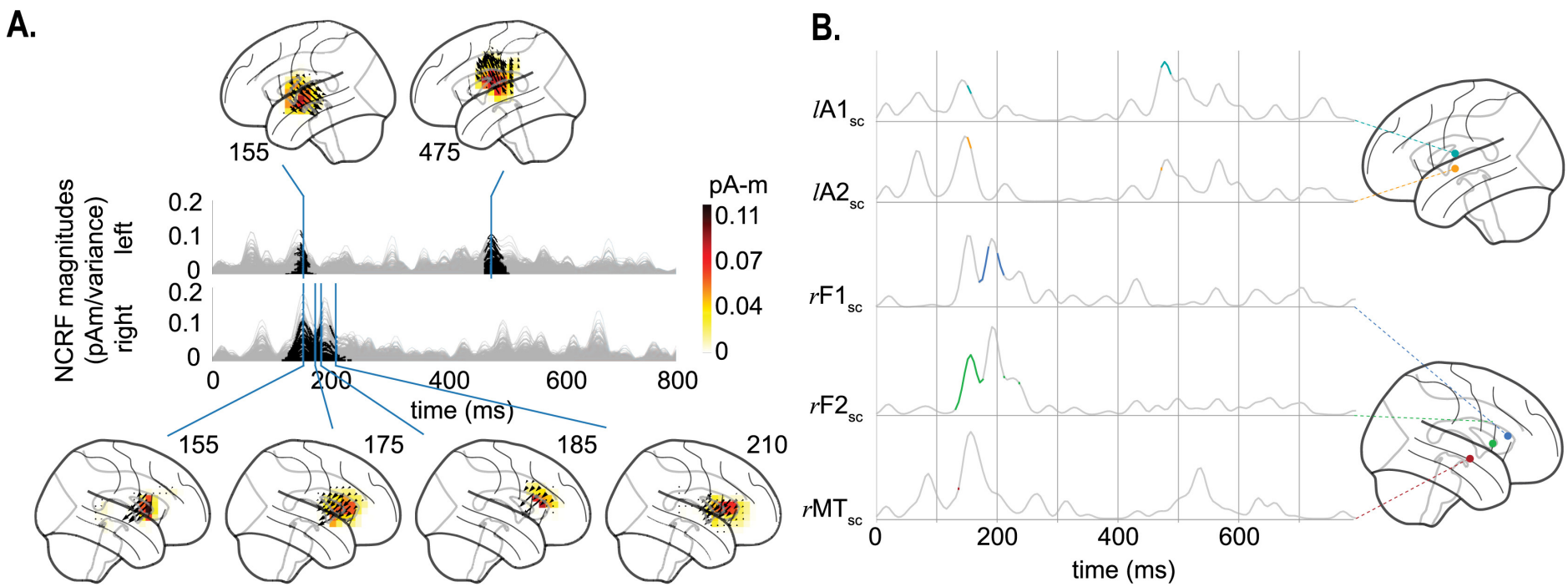
Acoustic envelope
Yang et. al., 1992



P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

MEG Application to Cortical Processing of Speech

Semantic composition:



- Bilateral auditory component at ~ 155 ms, late auditory component at ~ 475 ms.
- Auditory-frontal dynamics at ~ 175-210 ms ($r\mathrm{MT}_{\mathrm{sc}} \rightarrow r\mathrm{F2}_{\mathrm{sc}} \rightarrow r\mathrm{F1}_{\mathrm{sc}} \rightarrow r\mathrm{F2}_{\mathrm{sc}}$).

P. Das, C. Brodbeck, J. Z. Simon, and B. Babadi, NeuroImage, 2020

Summary

- NCRFs: A tool for directly extracting the cortical dynamics that underlie **continuous stimulus processing** from MEG.
- Novel spatiotemporal prior that not only combats overfitting and spatio-temporal dispersion but is also robust in absence of MR scan.
- The NCRFs are readily interpretable in a without post-hoc processing.

NCRFs as a powerful source localization tool for continuous stimulus experiments.

Thank you!

slides available at

github repo at