Response of the auditory brainstem to running speech and selective attention

Tobias Reichenbach

Department of Bioengineering,
Imperial College London

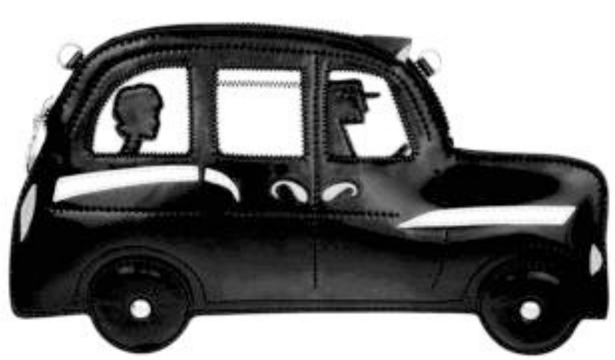
Physics of Hearing:
From Neurobiology to Information Theory and Back

KITP, June 14 2017

Auditory-scene analysis and attention

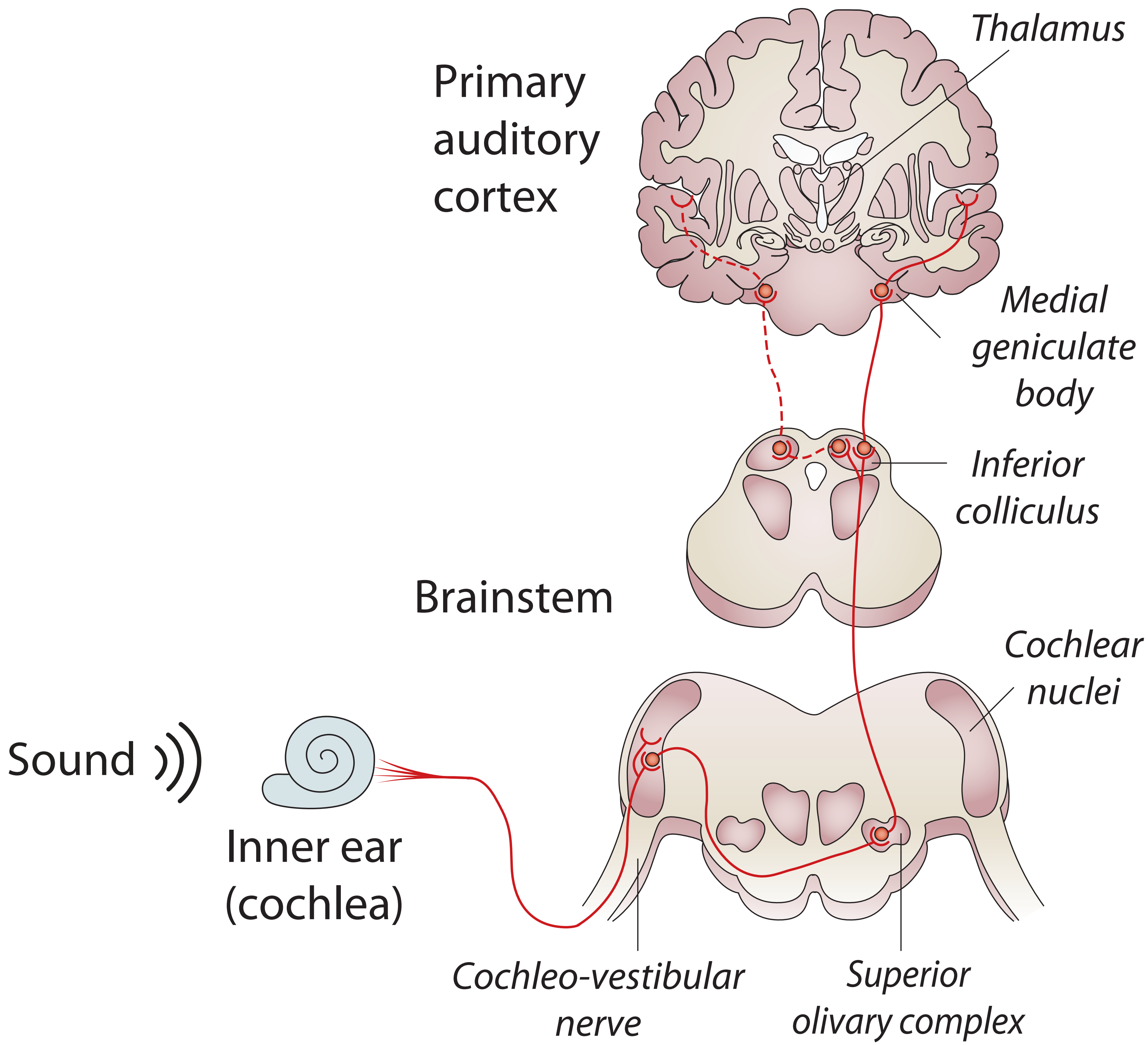
Bird songs

Conversations

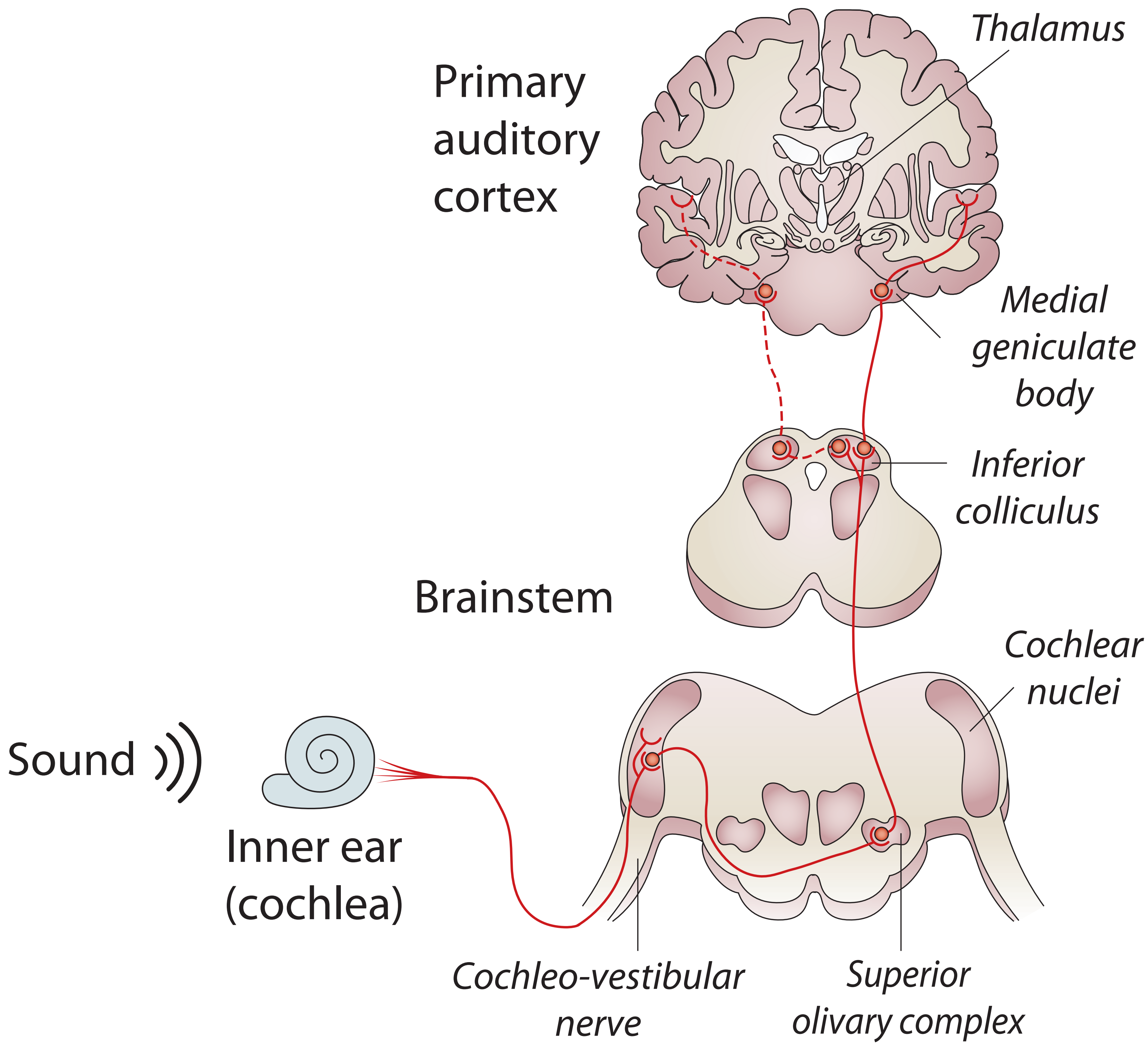


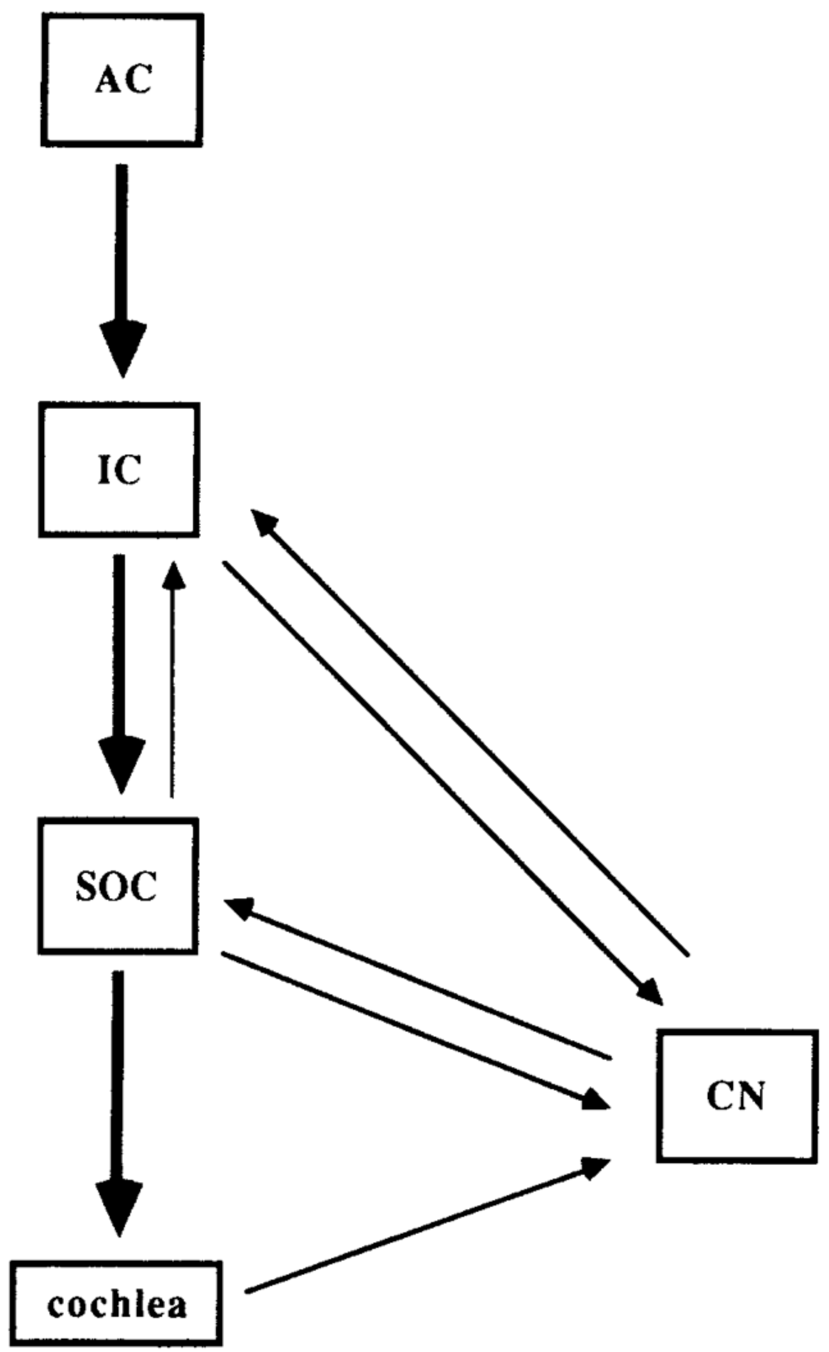
Traffic

The auditory system: neural feedback loops



The auditory system: neural feedback loops

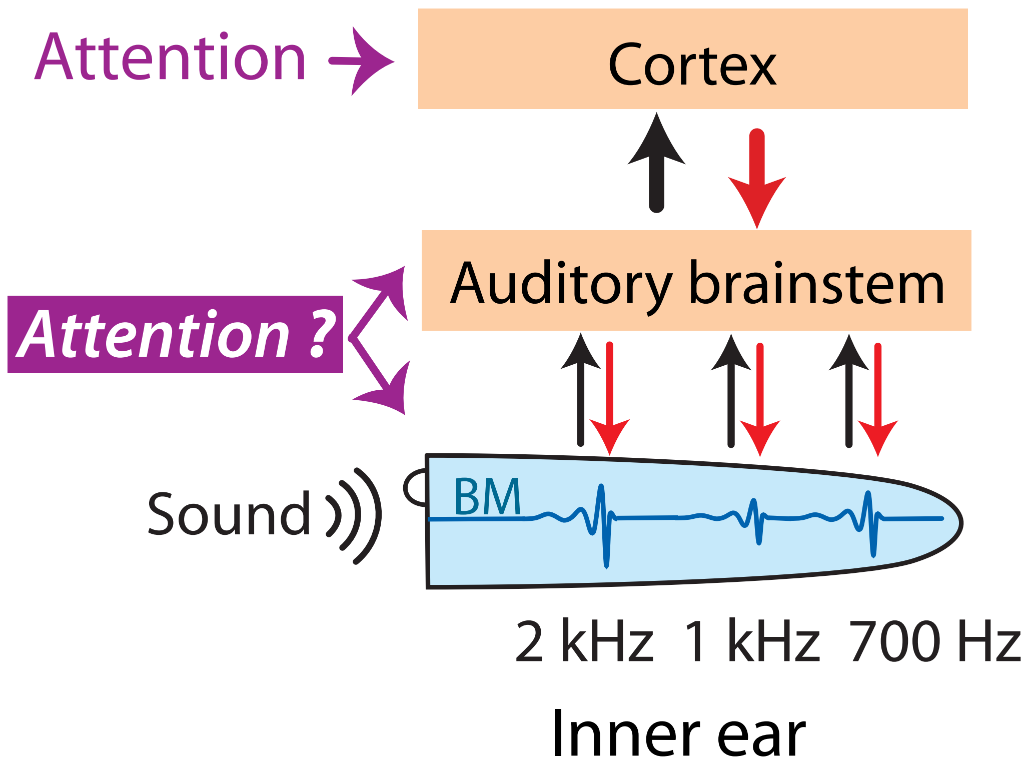




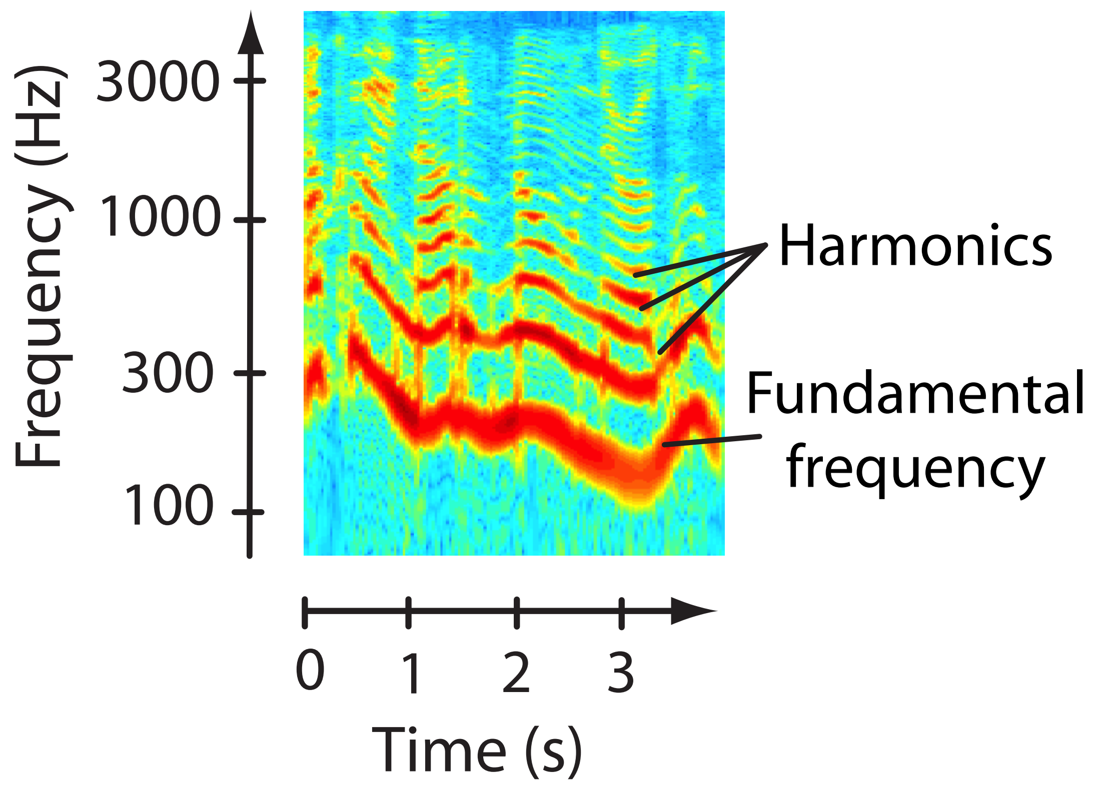
Huffman & Henson, *Brain Res. Rev.* (1990)

Attention and the fundamental frequency of speech

Inner ear and neural pathways



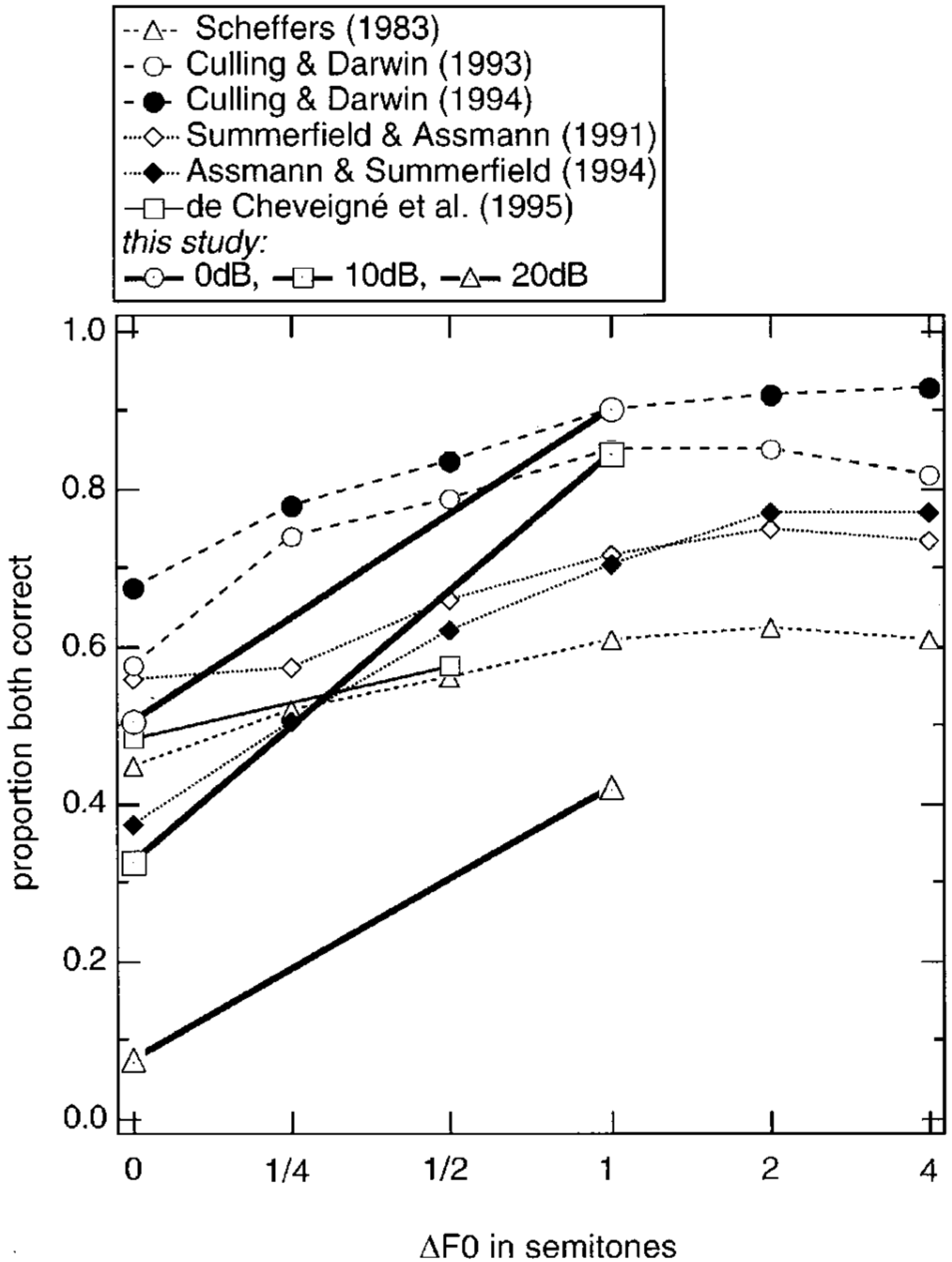
Spectrogram of natural speech



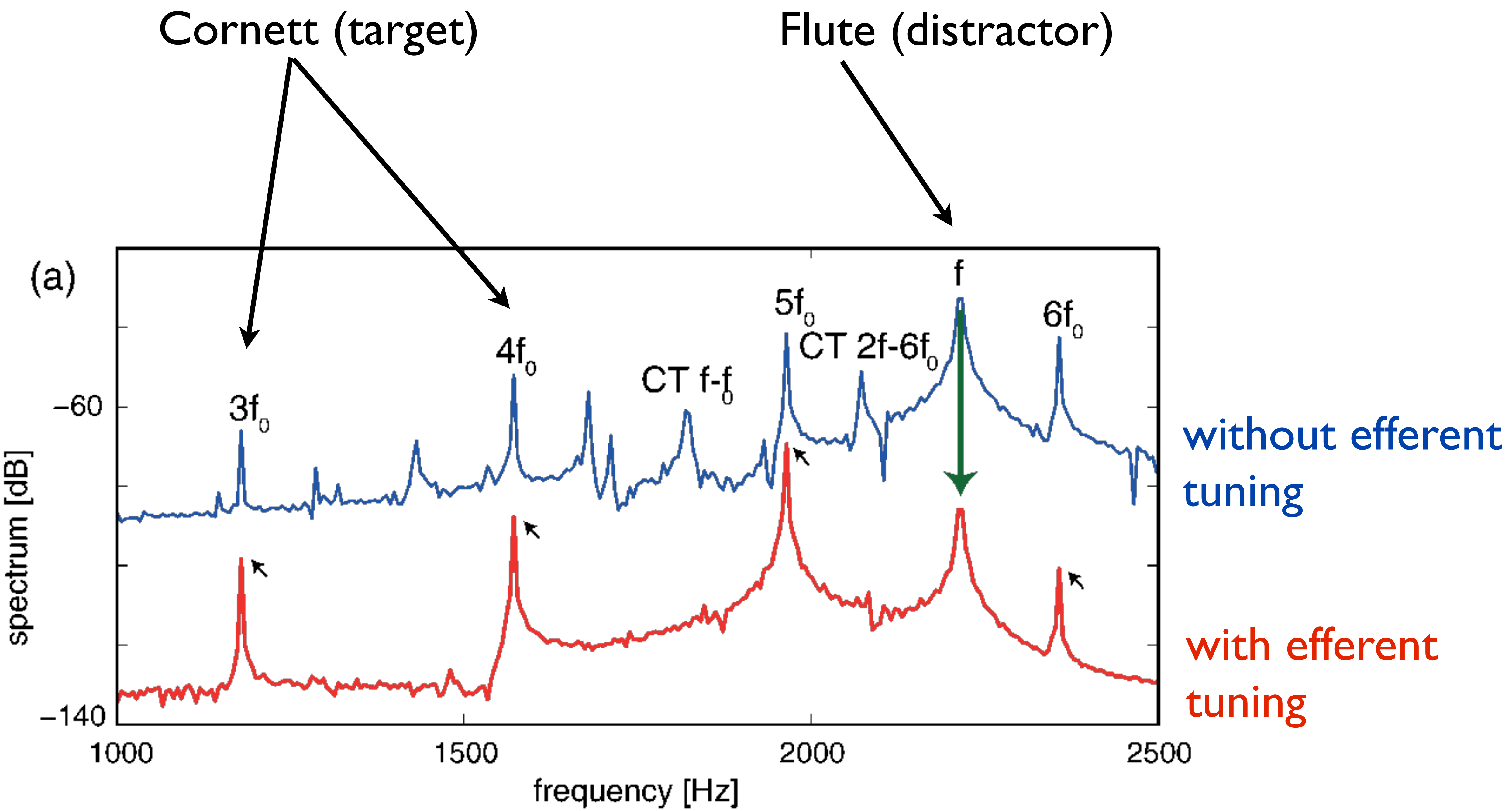
Psychoacoustics: selective attention to vowels

Separation of concurrent vowels is easier if their fundamental frequencies differ

Cheveigne et al., *JASA* (1997)



Pitch segregation in a cochlear model

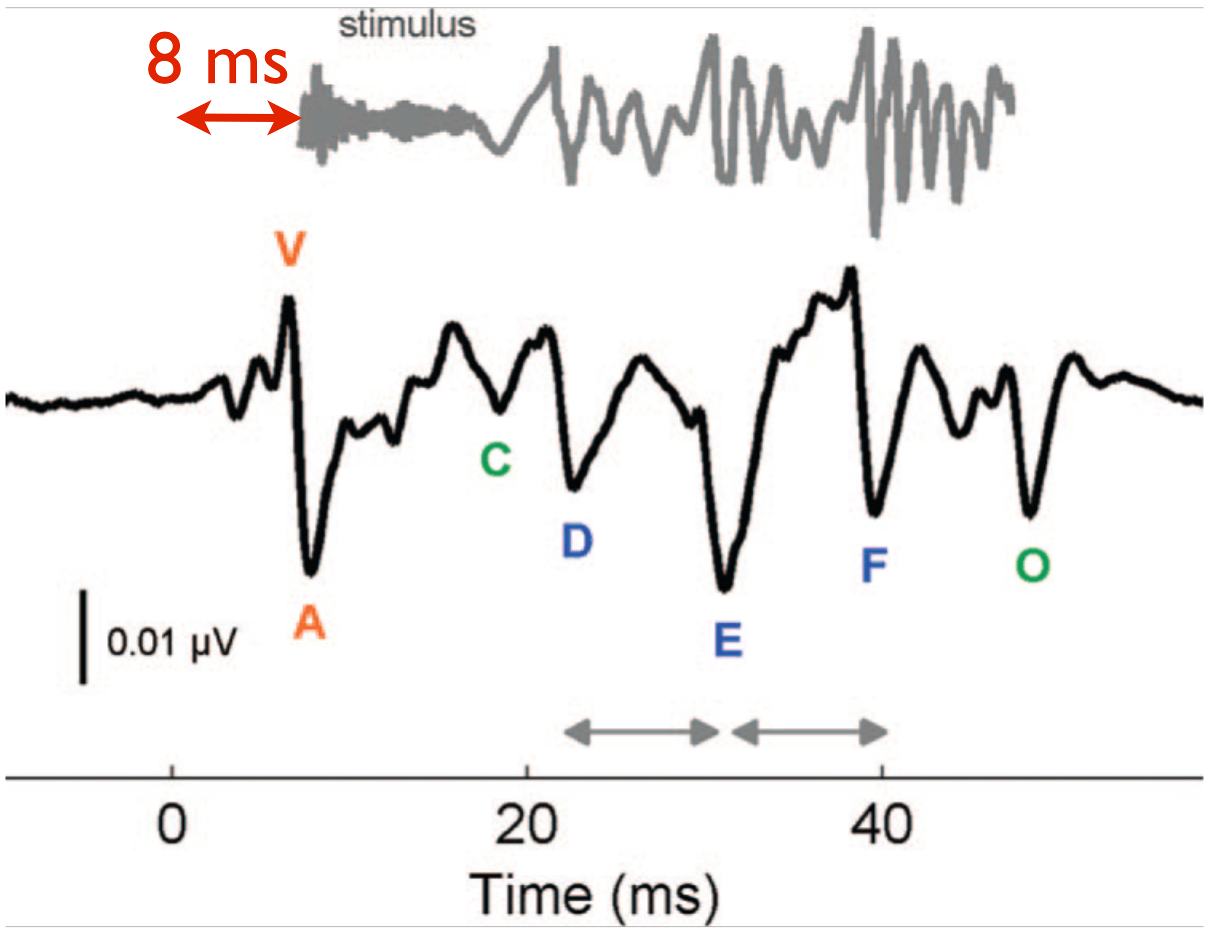


Gomez, Saase, Buchheim, Stoop, *Appl. Phys. Lett.* (2014)

Brainstem response to speech

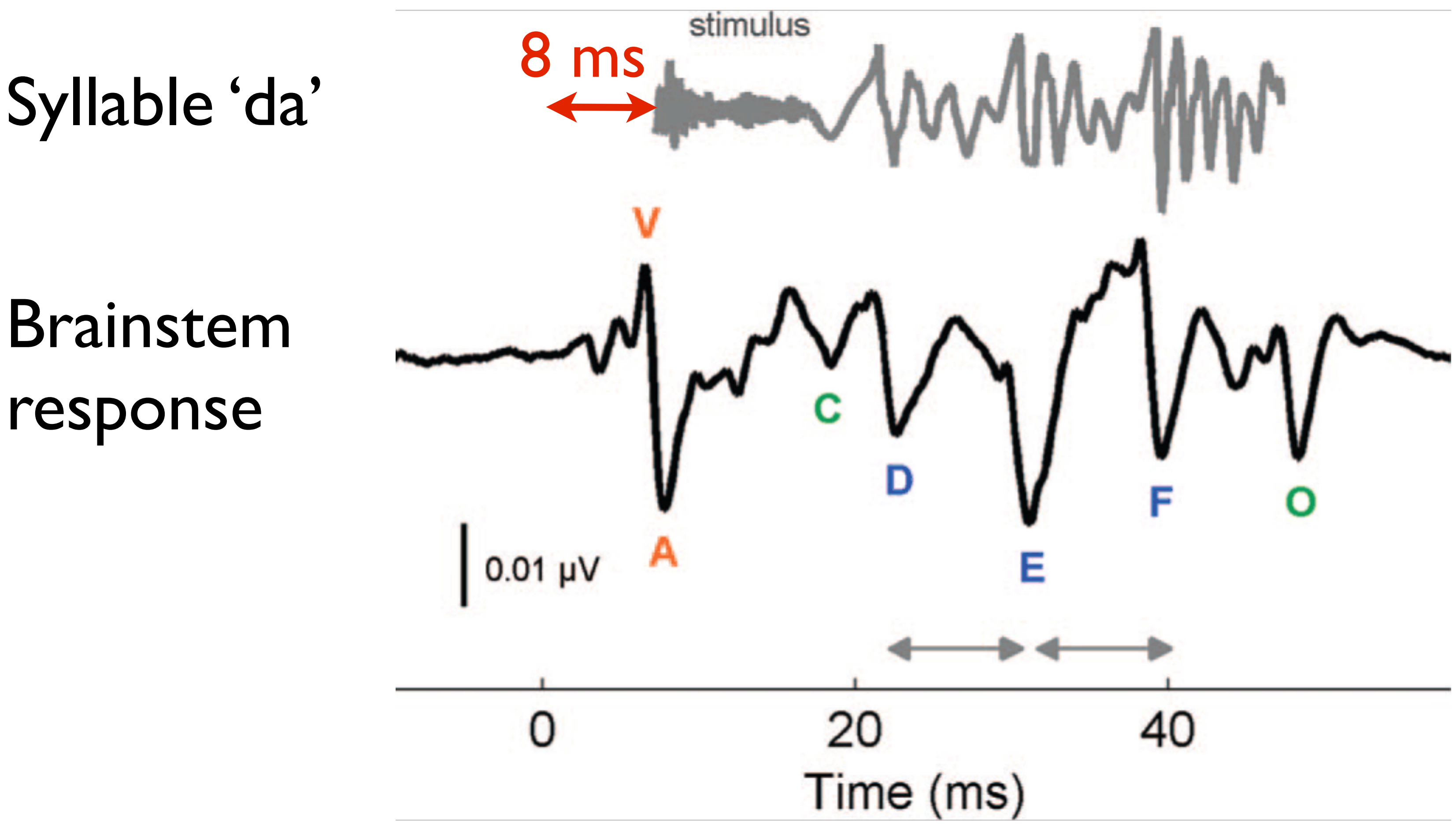
Syllable 'da'

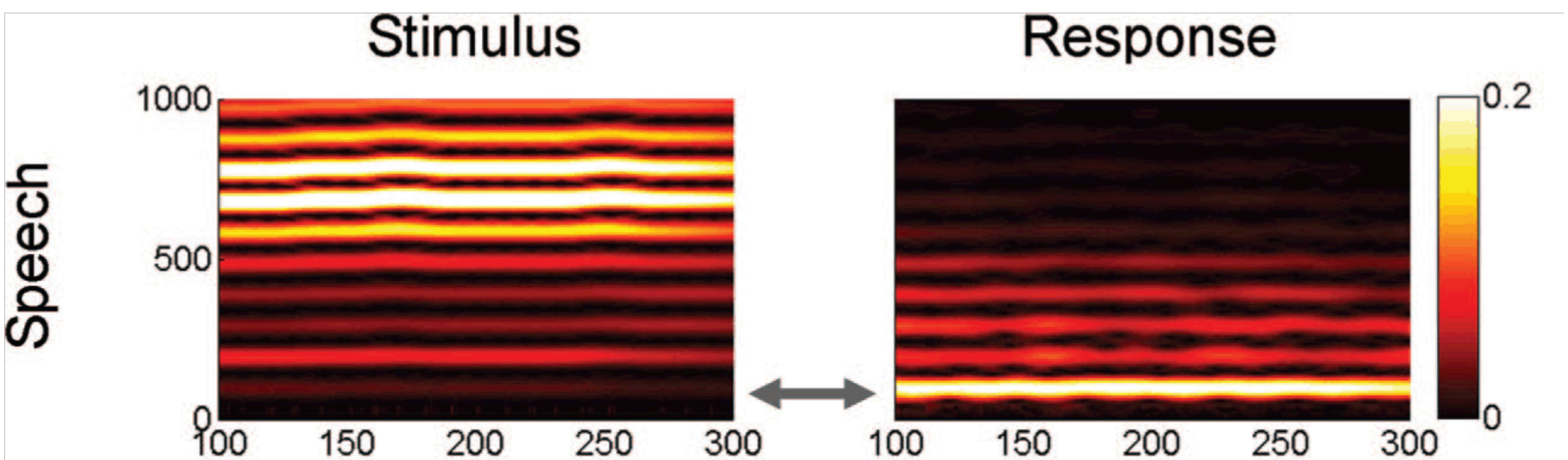
Brainstem response



Skoe & Kraus, *Ear Hear.* (2010)

Brainstem response to speech





Skoe & Kraus, *Ear Hear.* (2010)

Previous studies on attention in the brainstem

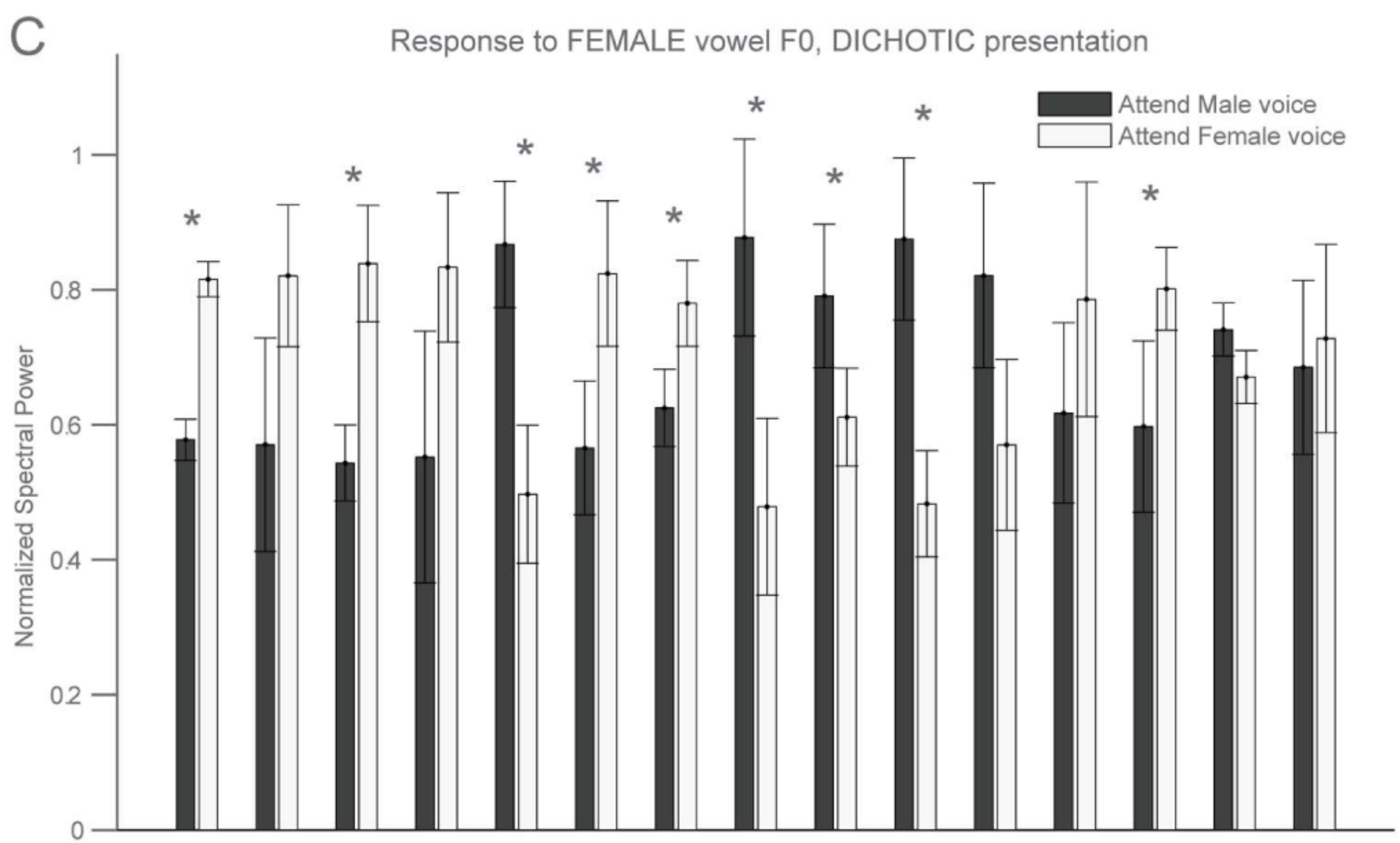
<u>Inconsistent results:</u>

effect of attention:
yes/no/subject-specific

- short speech segments (vowels, syllables, single digits)

=> require many repetitions

=> may lead to neural adaptation



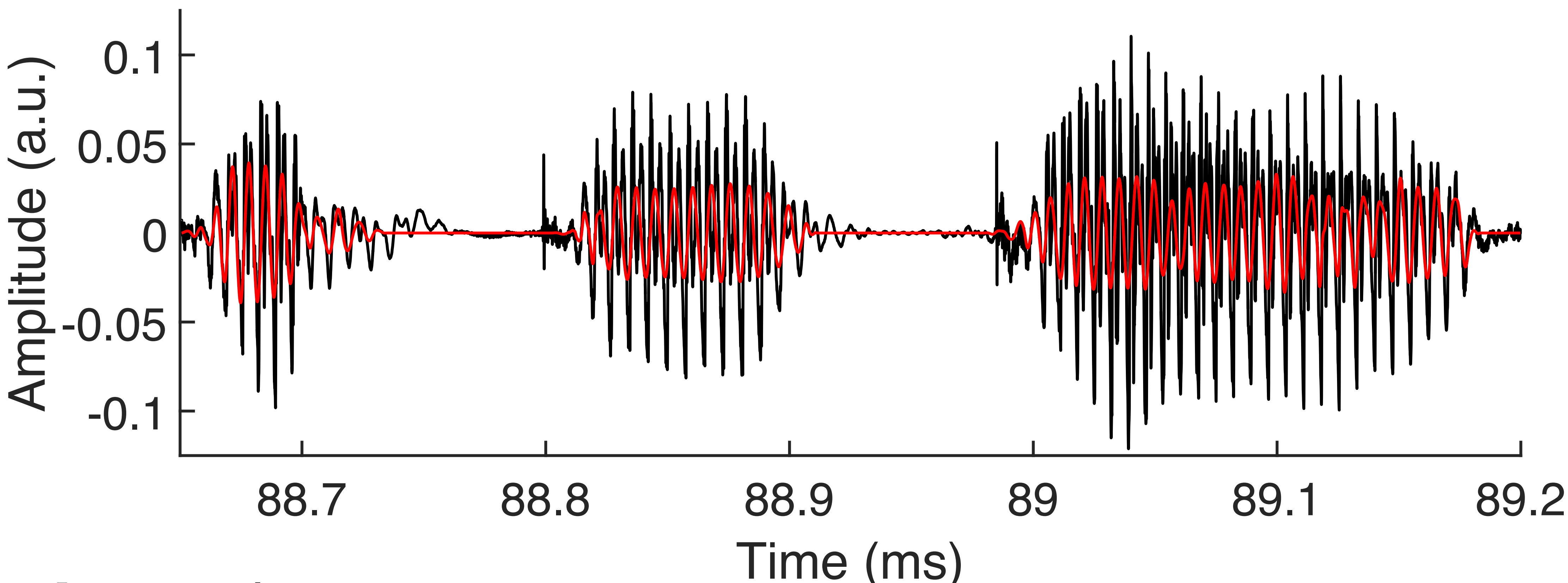
Lehmann & Schönwiesner, *PLoS One* (2014)

Auditory brainstem response to running speech

Goal: Measure response of the auditory brainstem to **running speech, no repetitions**

Problem: Fundamental frequency varies over time

Idea: Extract fundamental waveform that corresponds to oscillation at the fundamental frequency, correlate this to the neural recording



Black: speech

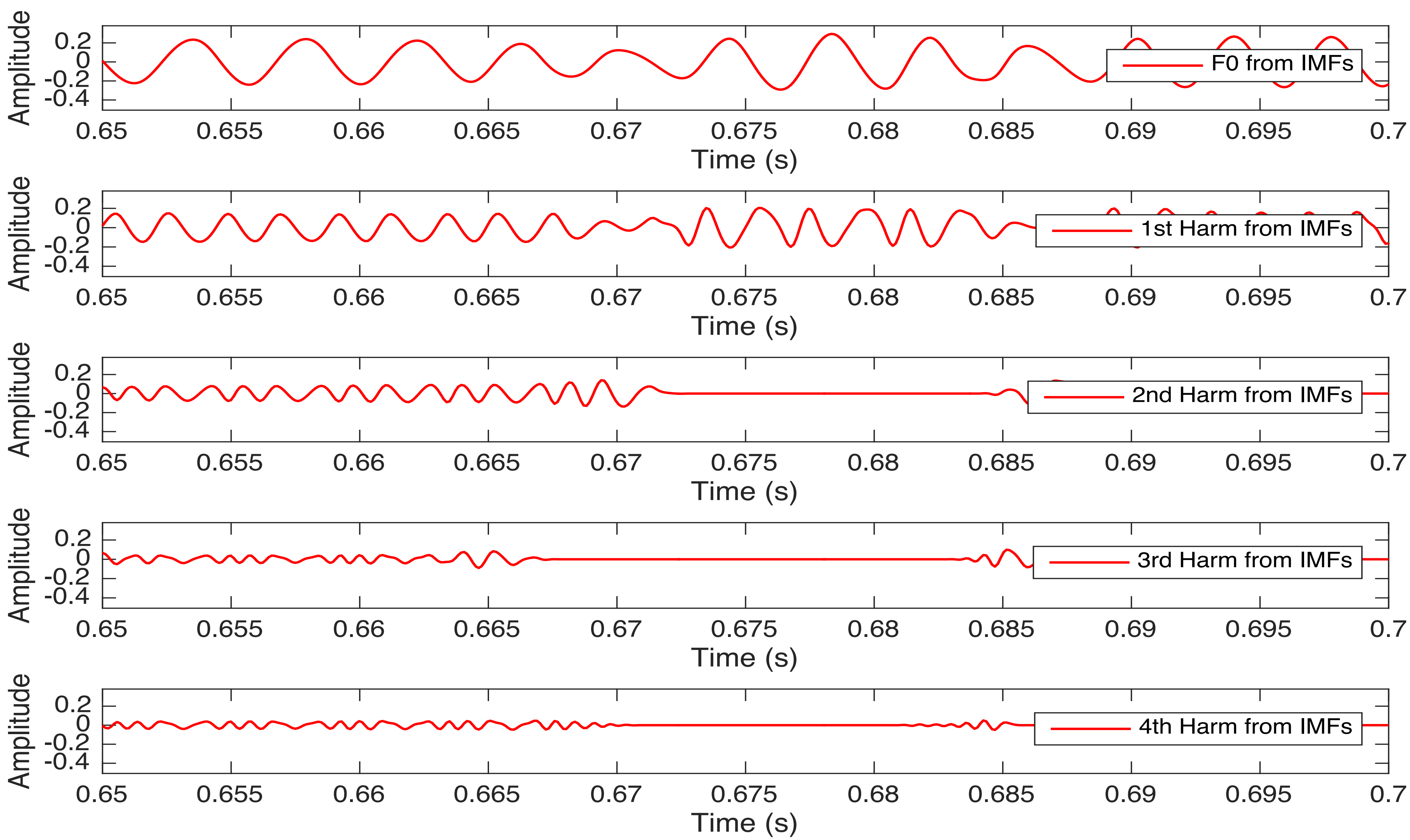
Red: fundamental waveform

Fundamental waveform: Empirical Mode Decomposition

Hilbert-Huang transform: decompose the speech signal into intrinsic mode function (IMFs)

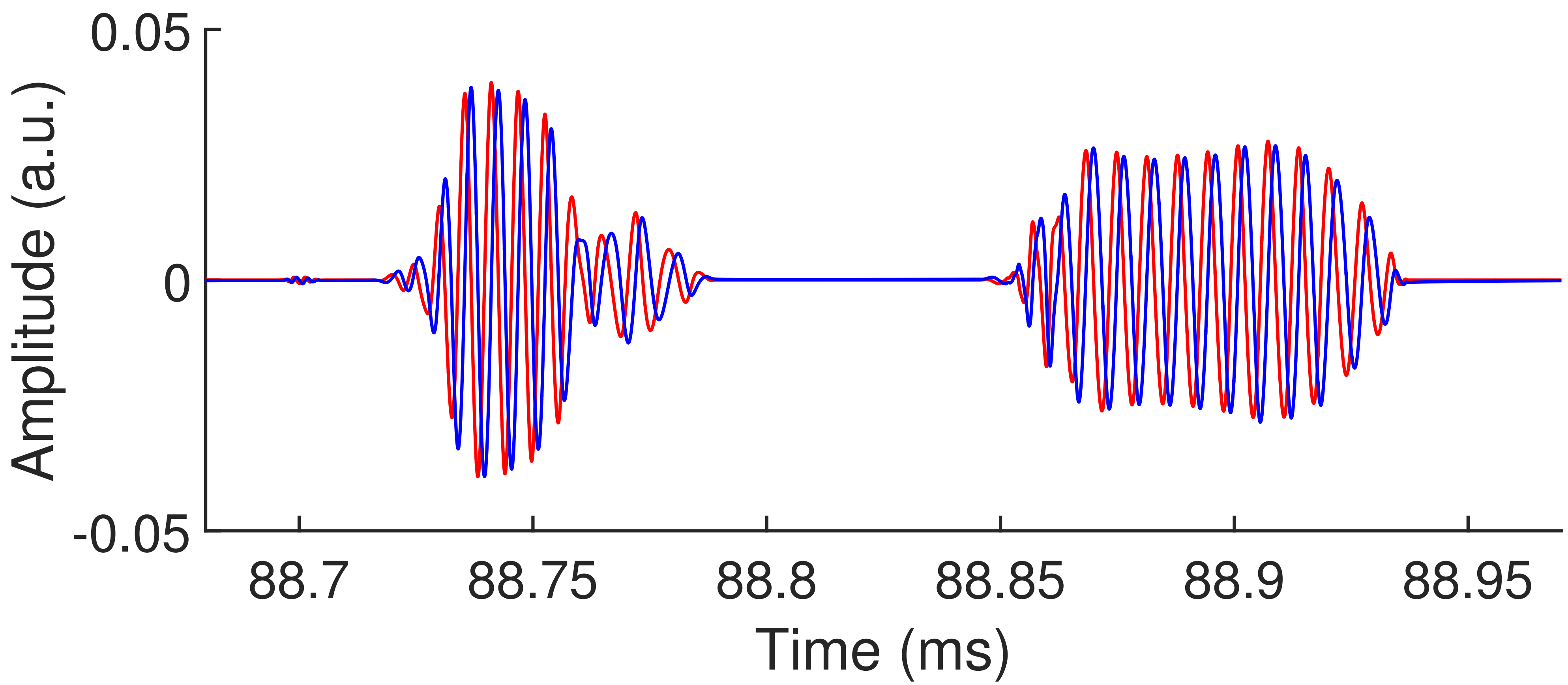
IMFs must satisfy two properties:

- Number of extrema = number of zero crossing (or differ by 1 at most)
- Mean of maximum and minimum envelope is zero



Huang & Pan, *Signal Proc.* (2006)

Fundamental waveform and its Hilbert transform

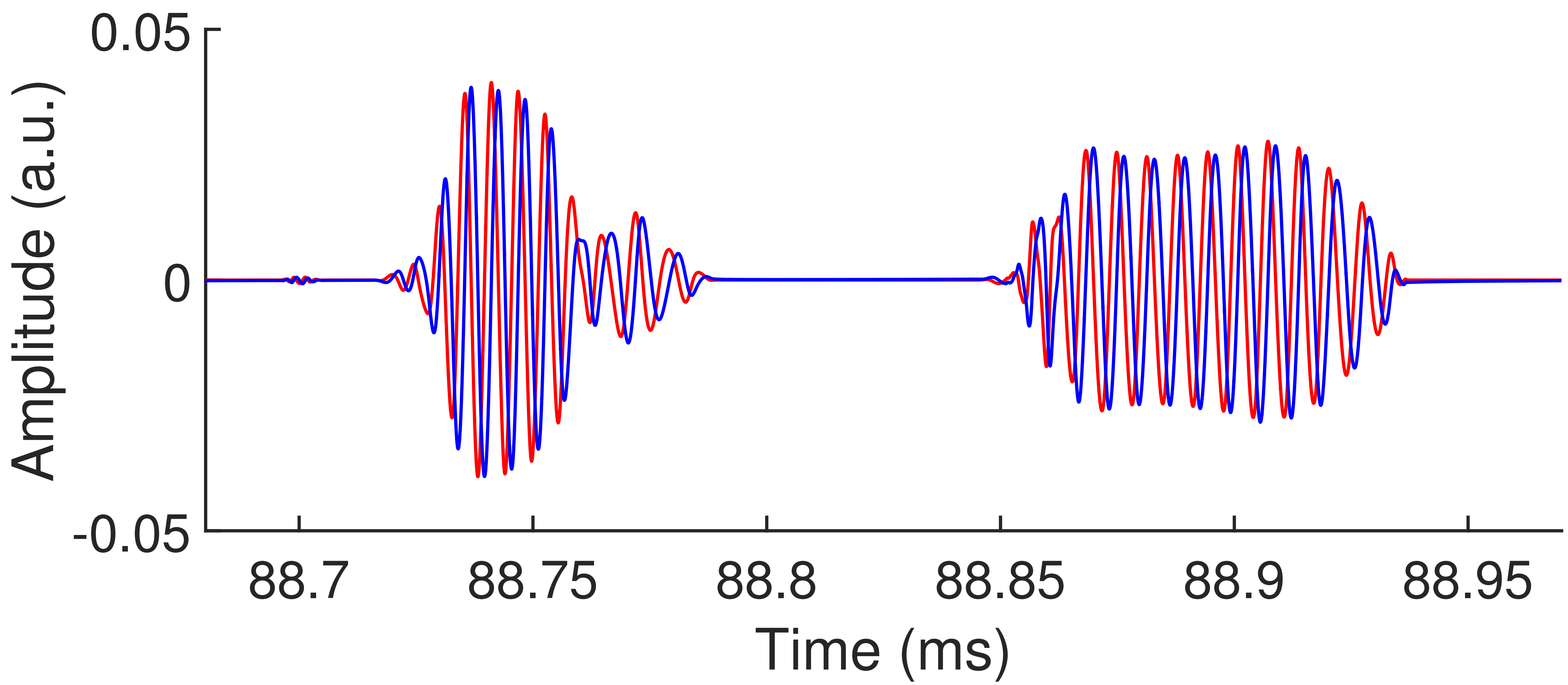


w(*t*) ... fundamental waveform

=> complex signal $z(t) = w(t) + i*H[w(t)]$

=> use correlation Corr(*x*,*z*) of neural response *x*(*t*) with *z*(*t*) to extract component that oscillates at the fundamental frequency at a certain phase

Fundamental waveform and its Hilbert transform



$w(t)$ ... fundamental waveform

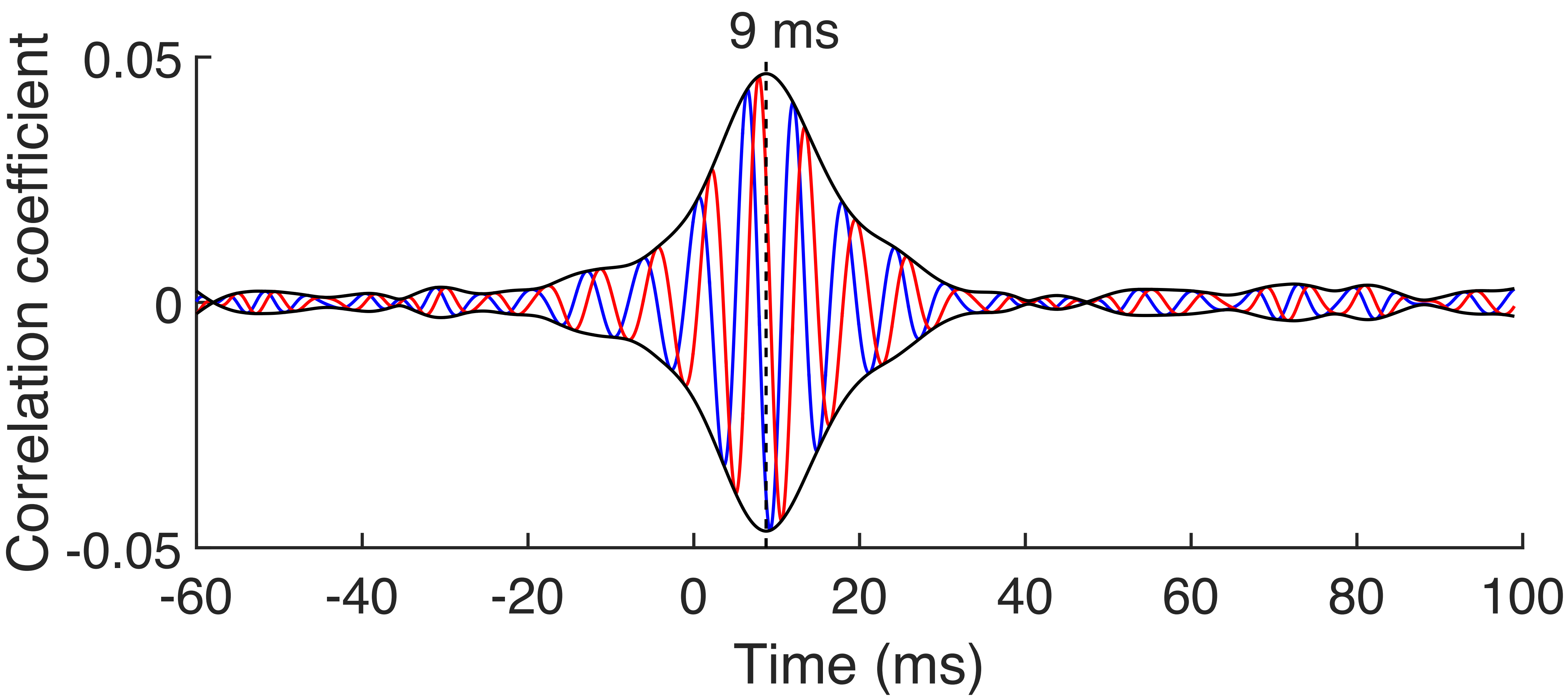
=> complex signal $z(t) = w(t) + i*H[w(t)]$

=> use correlation Corr(x,z) of neural response $x(t)$ with $z(t)$ to extract component that oscillates at the fundamental frequency at a certain phase

Reminder: Fourier analysis of signal $x(t)$:

Use correlation of $x(t)$ with complex function $cos(\omega t)+i*sin(\omega t)$ to extract component that oscillates at angular frequency ω

Correlation reveals brainstem response



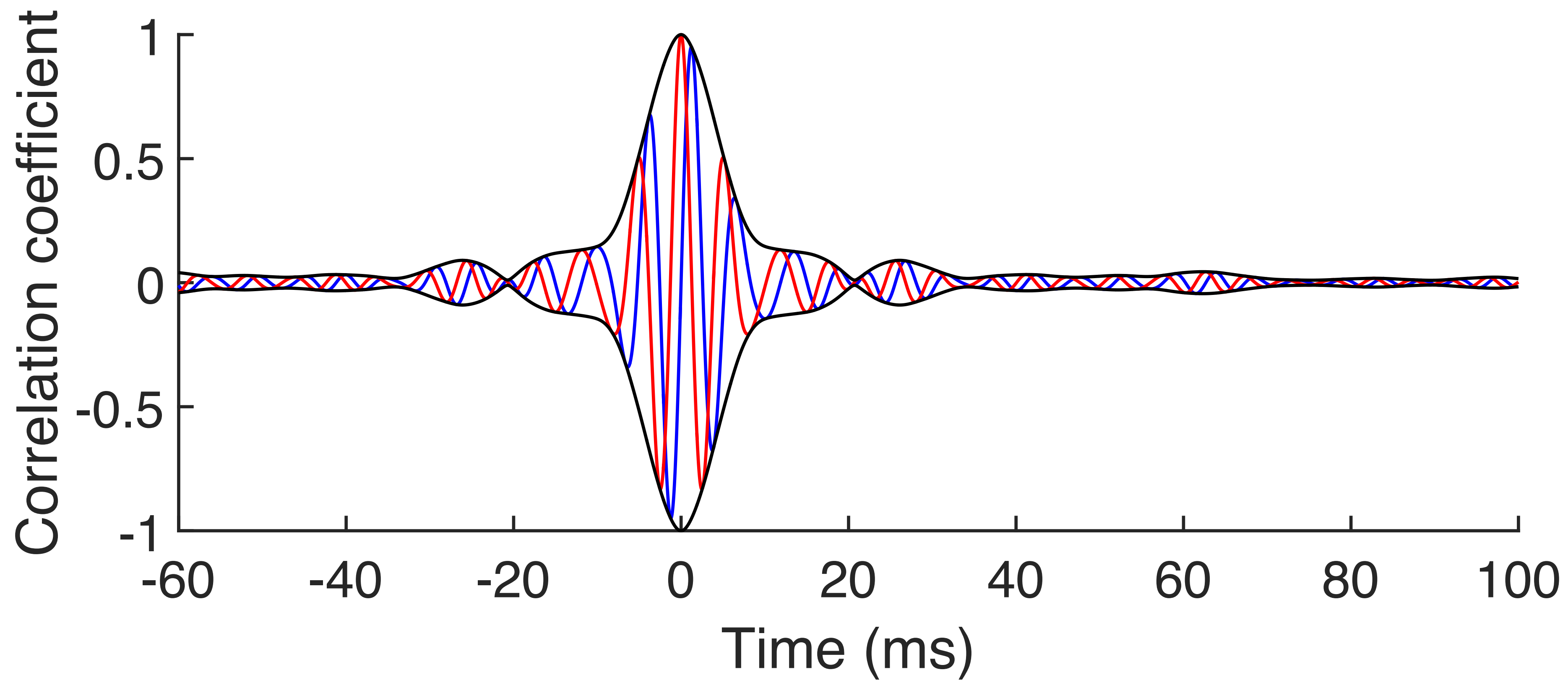
Red: correlation with the fundamental waveform

Blue: correlation with the Hilbert transform of the fundamental waveform

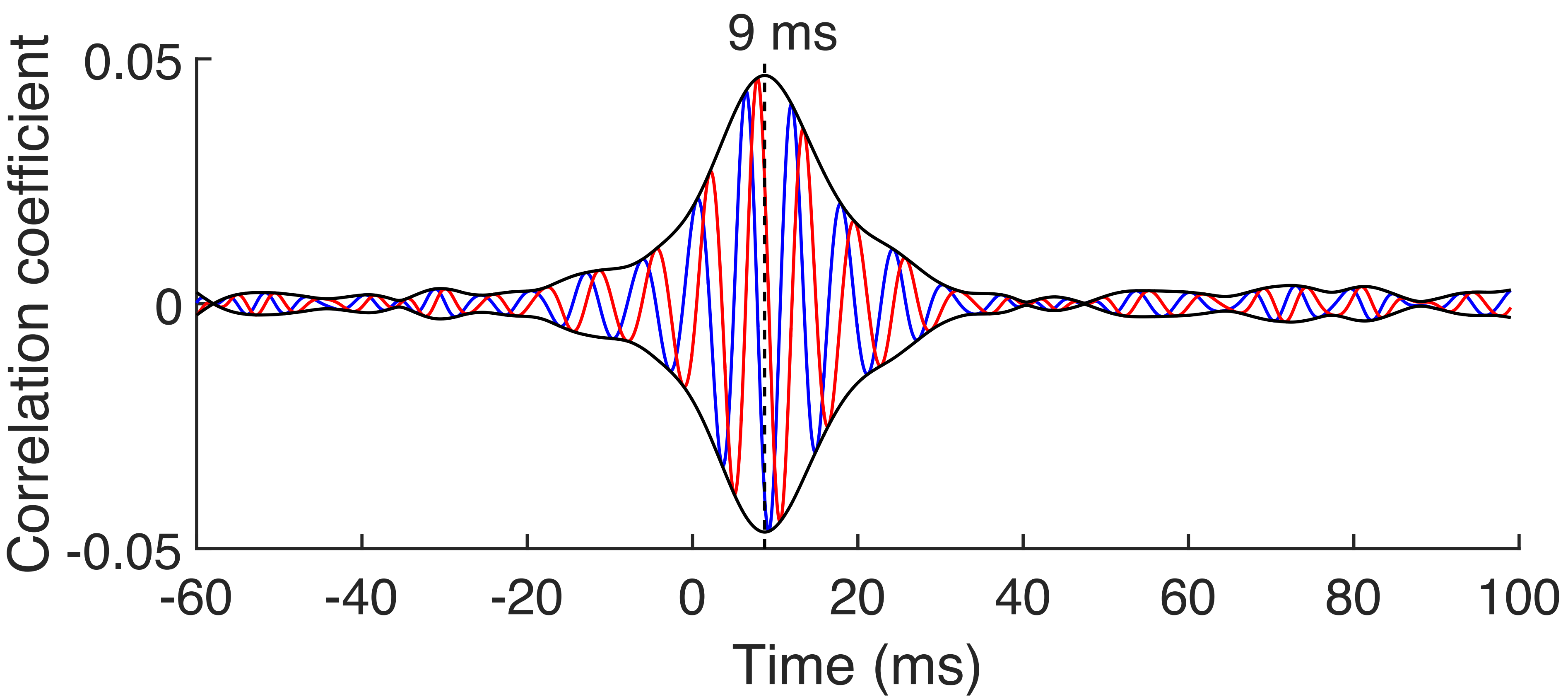
Black: absolute value of the complex correlation

Comparison to autocorrelation

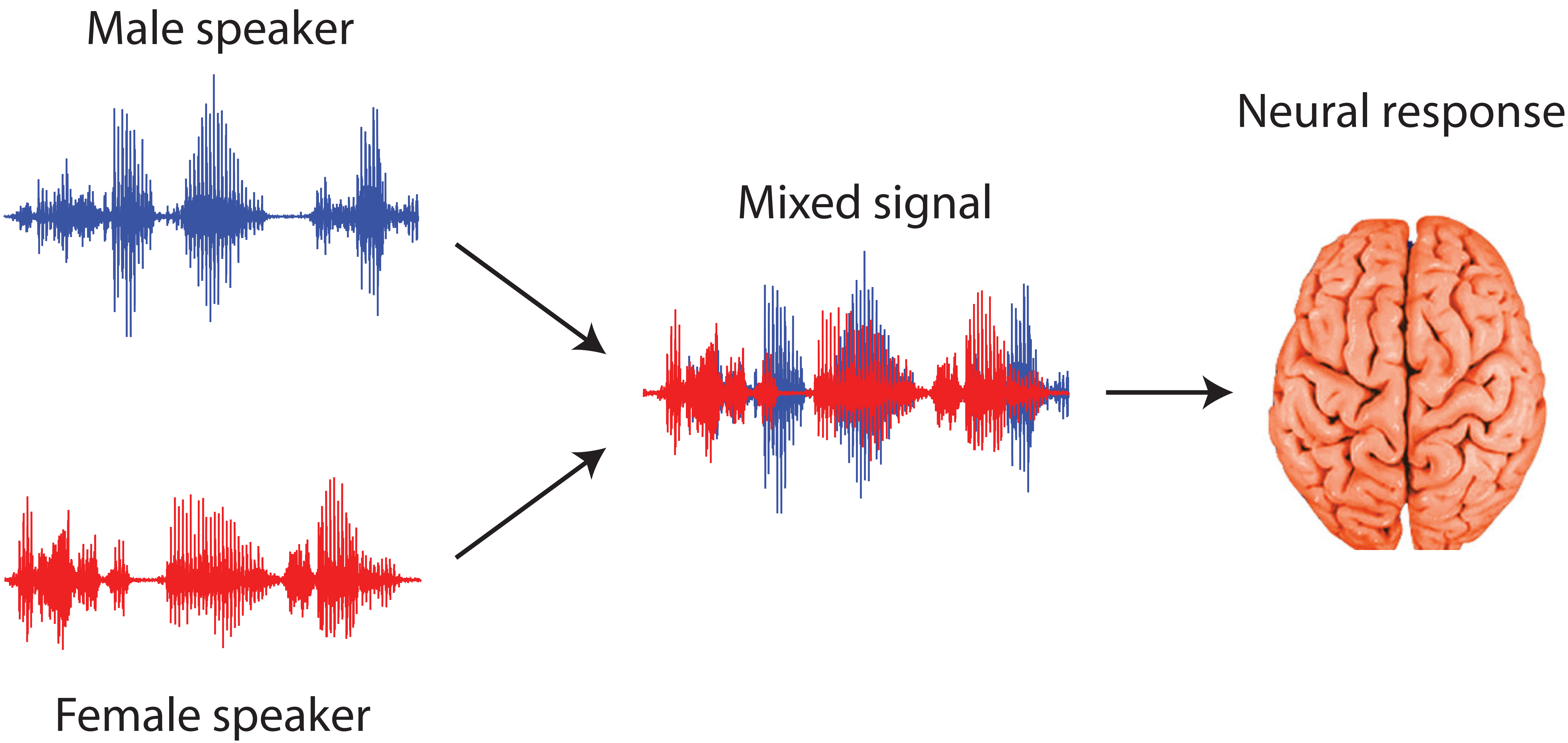
Autocorrelation of the fundamental waveform



Correlation between the fundamental waveform and the neural recording



Experiment on selective attention



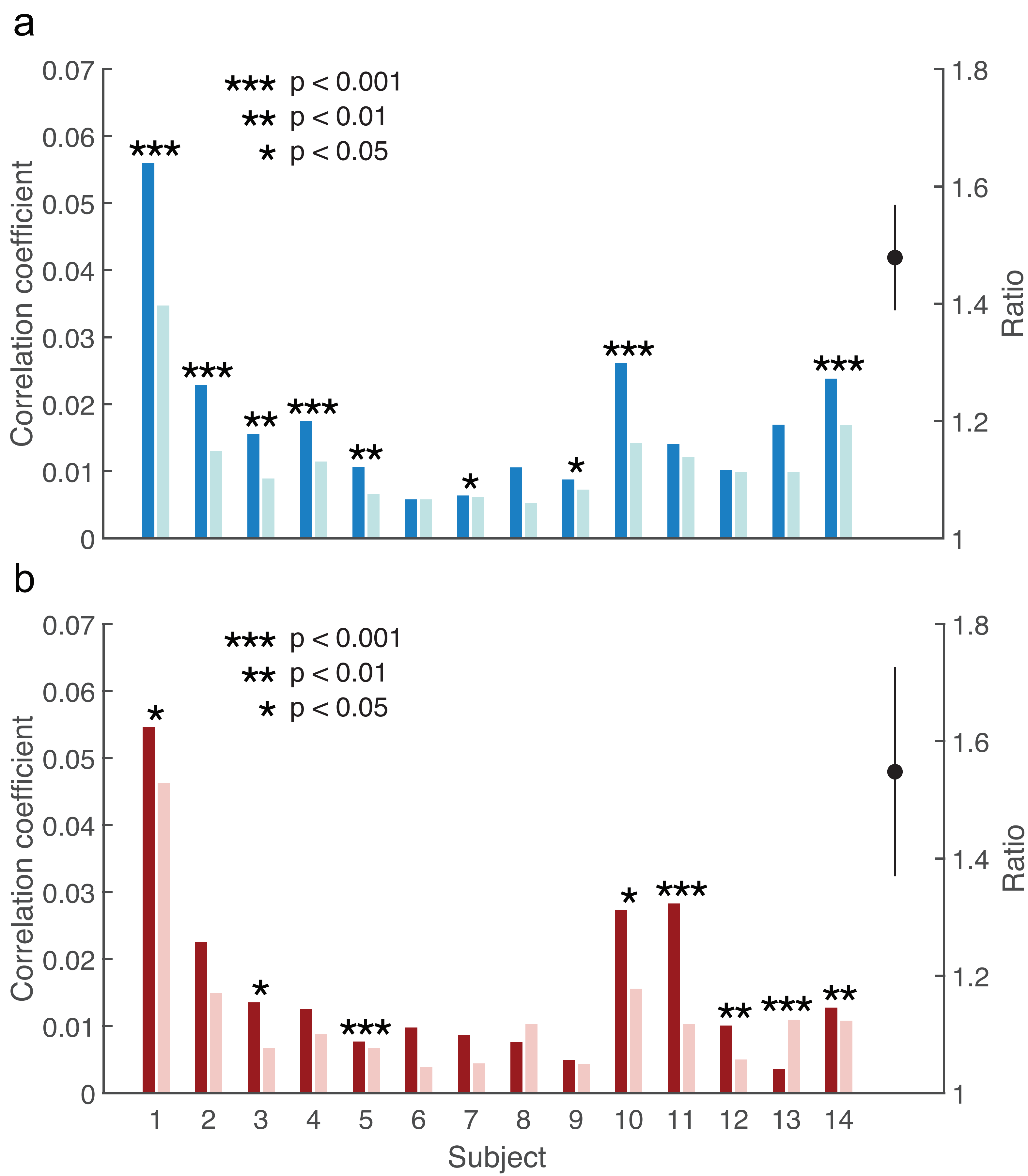
Modulation of the brainstem response by attention

Dark: attend
Light: ignore

Blue: responses to the male speaker

Red: response to the female speaker

Larger response to the attended then to ignored voice



Measuring the brainstem response from EEG

Idea: Measure the brainstem response to speech from high-density EEG

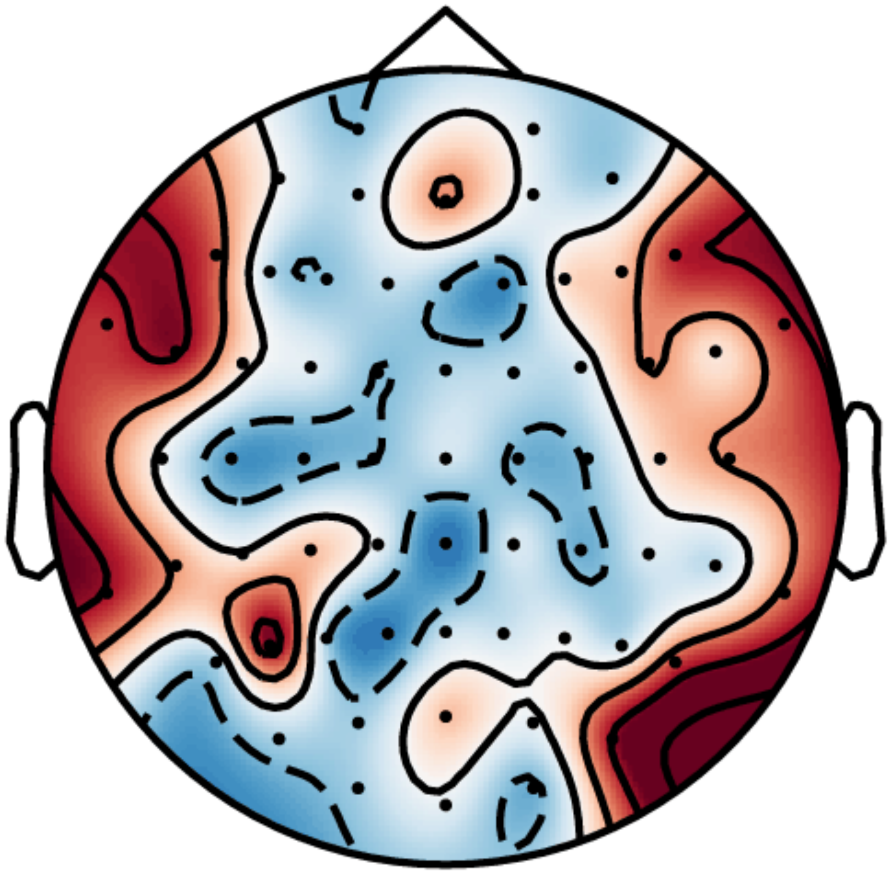
Challenge: How to select/combine channels to obtain the best signal?

Method: Statistical model to express the fundamental waveform through the EEG recordings

$$\hat{s}(t) = \sum_{c=1}^{N_c} \sum_{\ell=0}^{N_\ell - 1} \beta_c^\ell \cdot x_c(t+\ell) = \boldsymbol{X} \cdot \boldsymbol{\beta}$$

$$\text{Ridge solution:} \begin{cases} \hat{\boldsymbol{\beta}} = \operatorname{argmin}_{\boldsymbol{\beta}} \|\hat{\boldsymbol{s}} - \boldsymbol{s}\|^2 \ \text{ subject to } \|\boldsymbol{\beta}\|^2 \leq \tilde{\lambda} \\ \text{or, equivalently:} \\ \hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T \cdot \boldsymbol{X} + \lambda \boldsymbol{I})^{-1} \boldsymbol{X}^T \cdot \boldsymbol{s} \end{cases}$$

Ridge regression (linear regression with L_2 regularization)



Model coefficients

Conclusions

- Brainstem response to non-repetitive natural speech can be measured reliably through extracting the fundamental waveform of speech and correlating to the neural response
- The response is affected by selective attention to one of two competing speakers
- The brainstem response to speech can also be measured effectively from high-density EEG using advanced statistical modeling
- Attentional modulation is frequency-specific
- Attention may begin in the cochlea
- Relevant for biologically-inspired speech processing?

Acknowledgement

Antonio Elia Forte
Octave Etard
Mikolaj Kegler
Hugo Weissbart

and the rest of the lab

<u>Collaborators</u>

Steve Bell David Simpson	*University of Southampton*
Karolina Kluk	*University of Manchester*
James Harte	*Interacoustics*
Paul Iverson	*University College London*

EPSRC
Pioneering research and skills

Center for Neurotechnology

THE ROYAL BRITISH LEGION

Center for Blast Injury Studies

THE ROYAL SOCIETY

wellcome trust