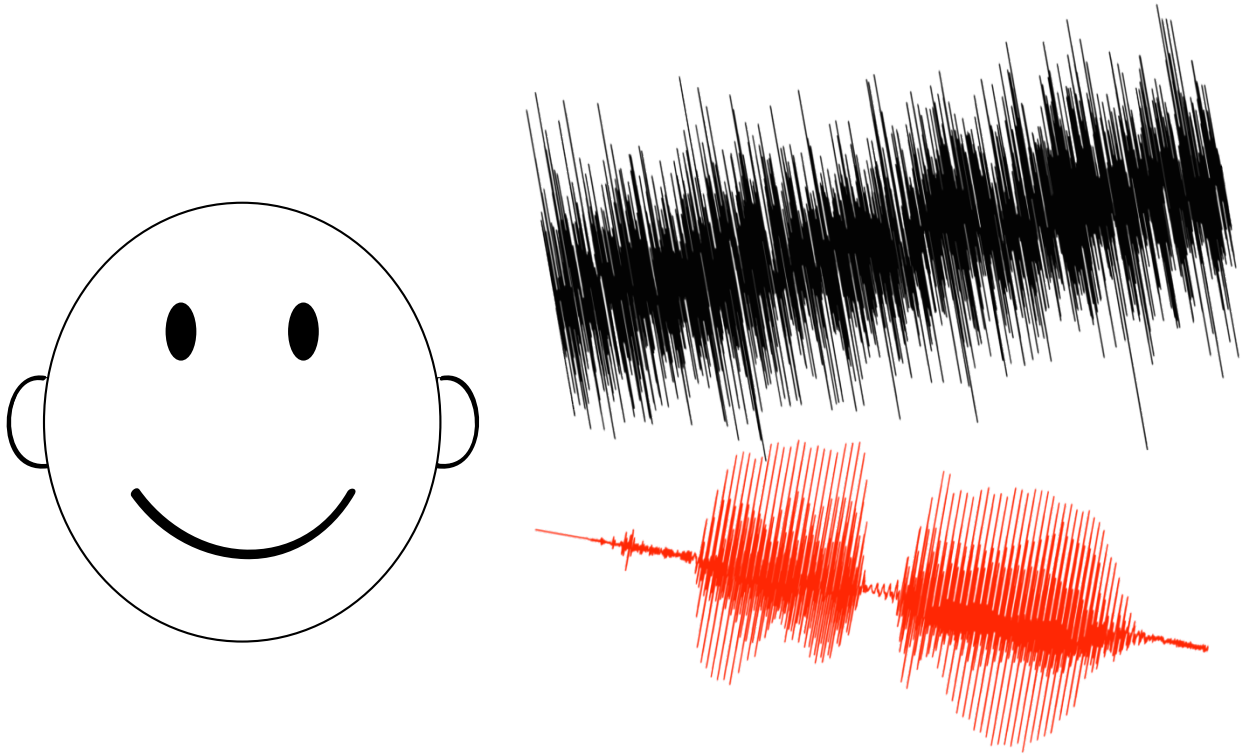
Spectro-temporal Fine Structure Critical for Robust Neural Synchronization to Speech

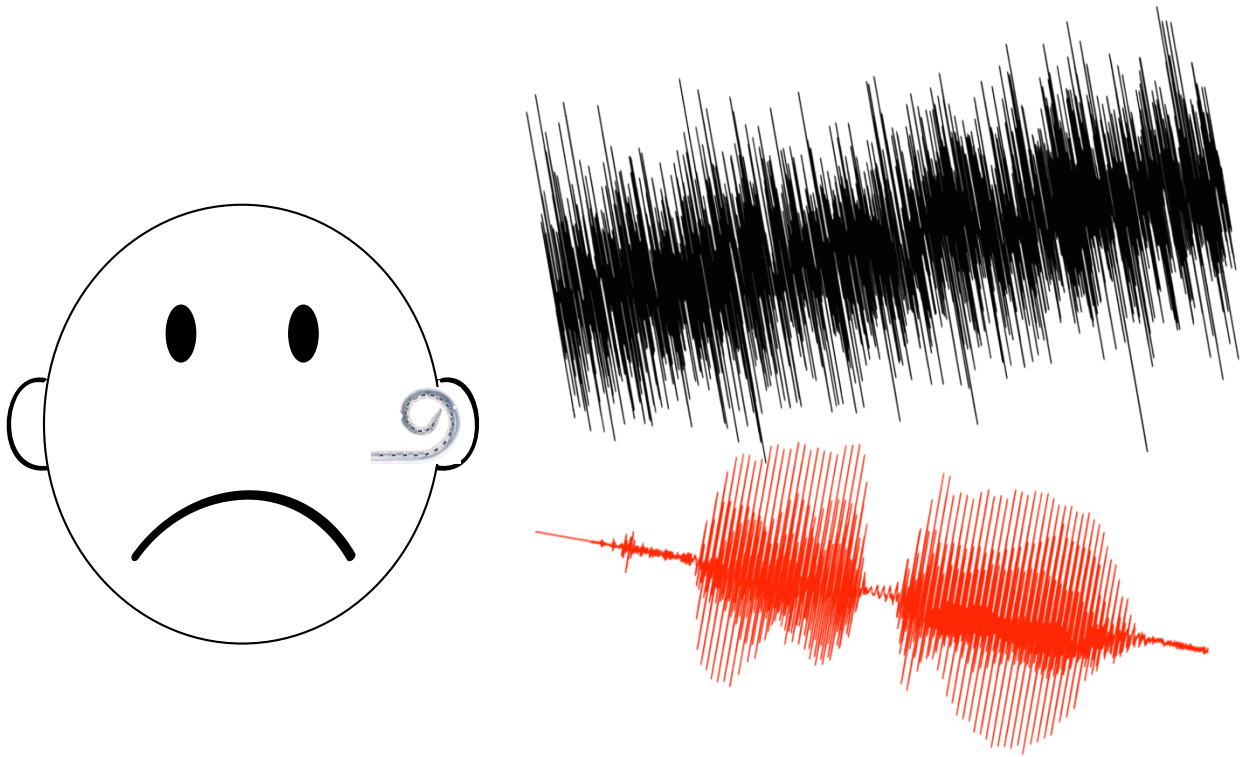
Nai Ding, Monita Chatterjee, Jonathan Z. Simon

University of Maryland, College Park
Boys Town Hospital, Omaha

Normal hearing listeners can recognize speech reliably in noisy environments.



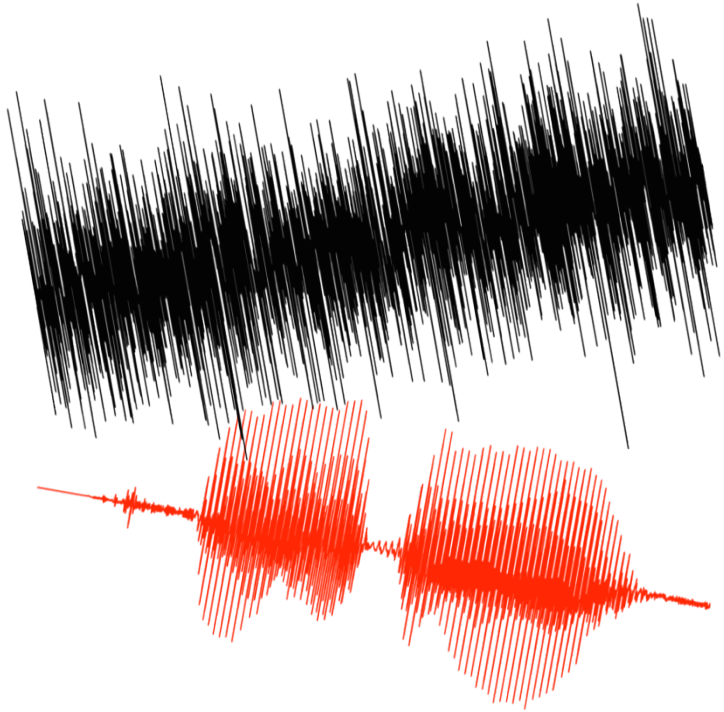
Cochlear implant users, however, have difficulty recognizing speech in noise.



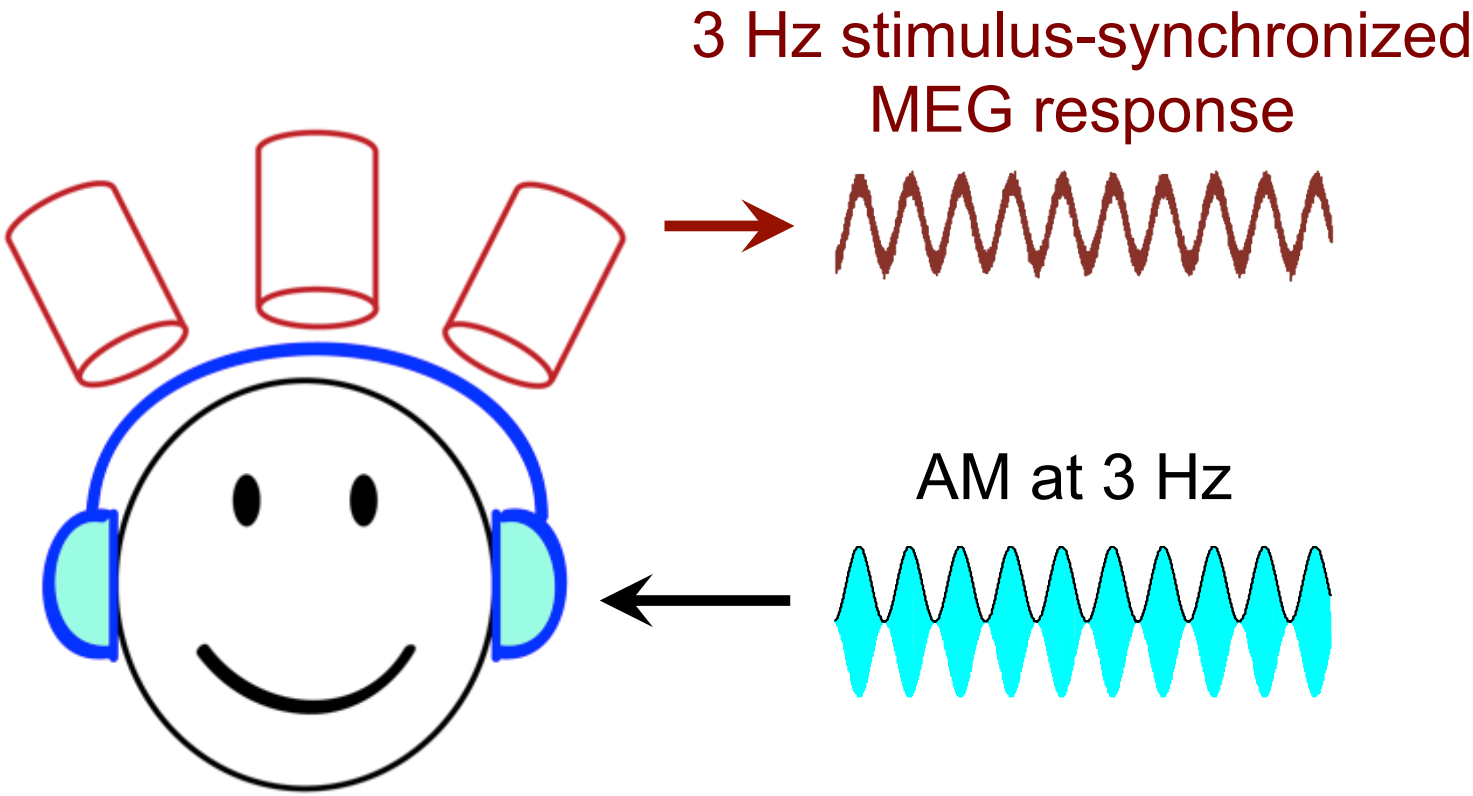
Cortical auditory processing of speech in noisy environments is studied using MEG.

magnetoencephalography

(MEG)



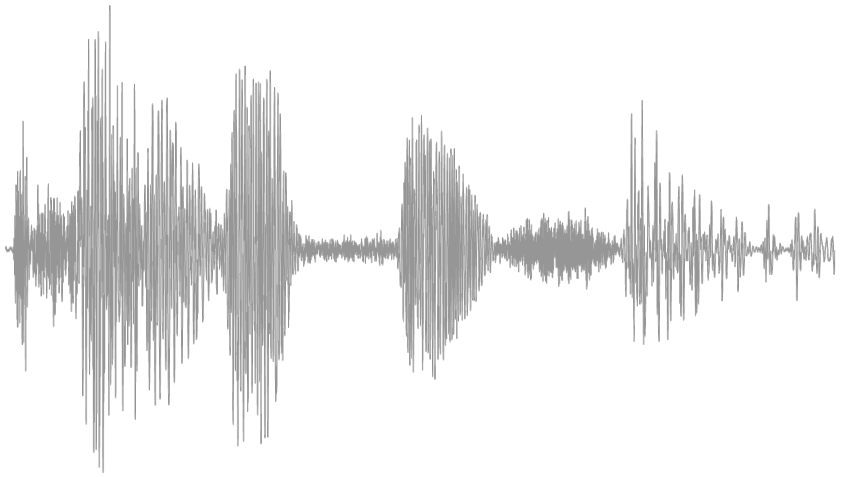
MEG can record neural activity phase-locked to slow temporal modulations



Ding & Simon, J Neurophys (2009)

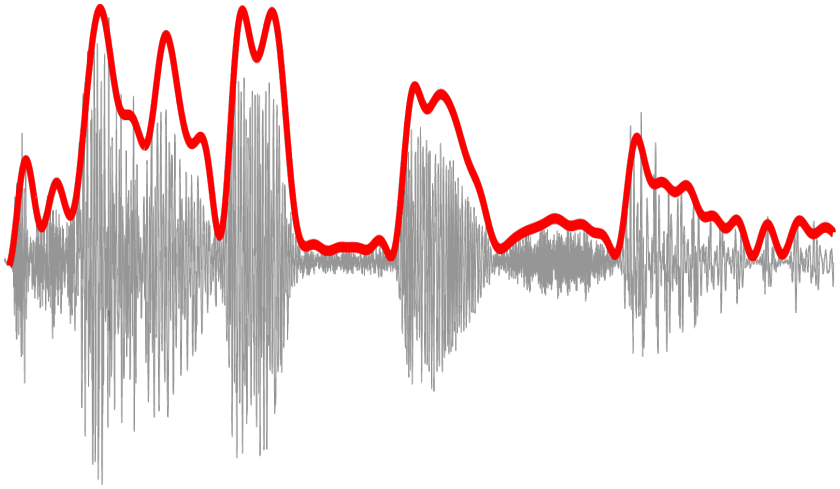
Envelope & Fine Structure of Speech

Speech Waveform



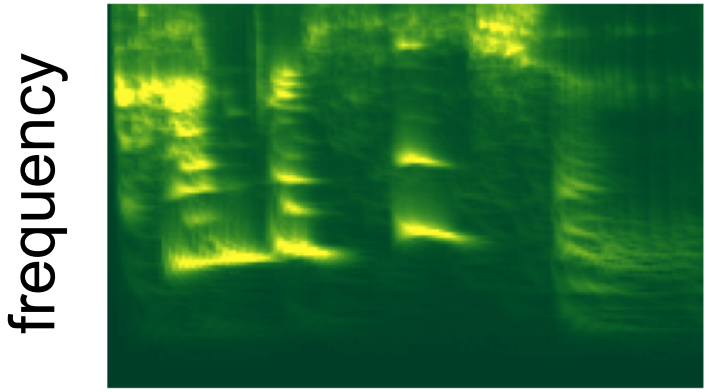
Envelope & Fine Structure of Speech

Speech Waveform



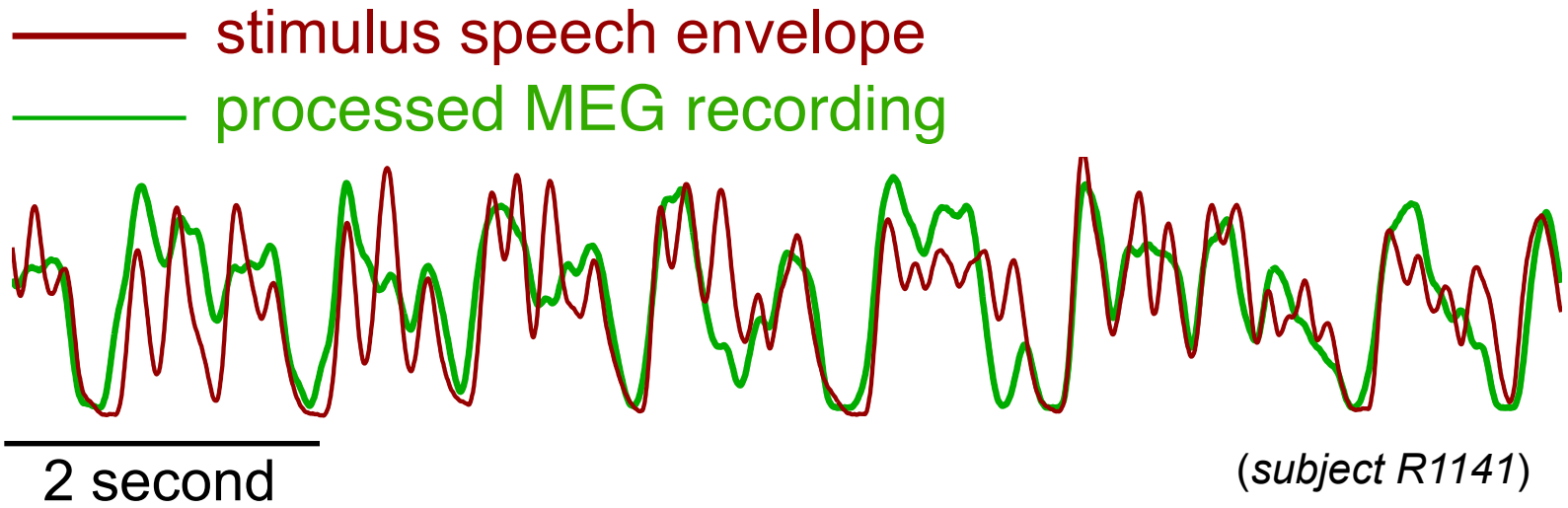
Speech Envelope

Speech Spectrogram



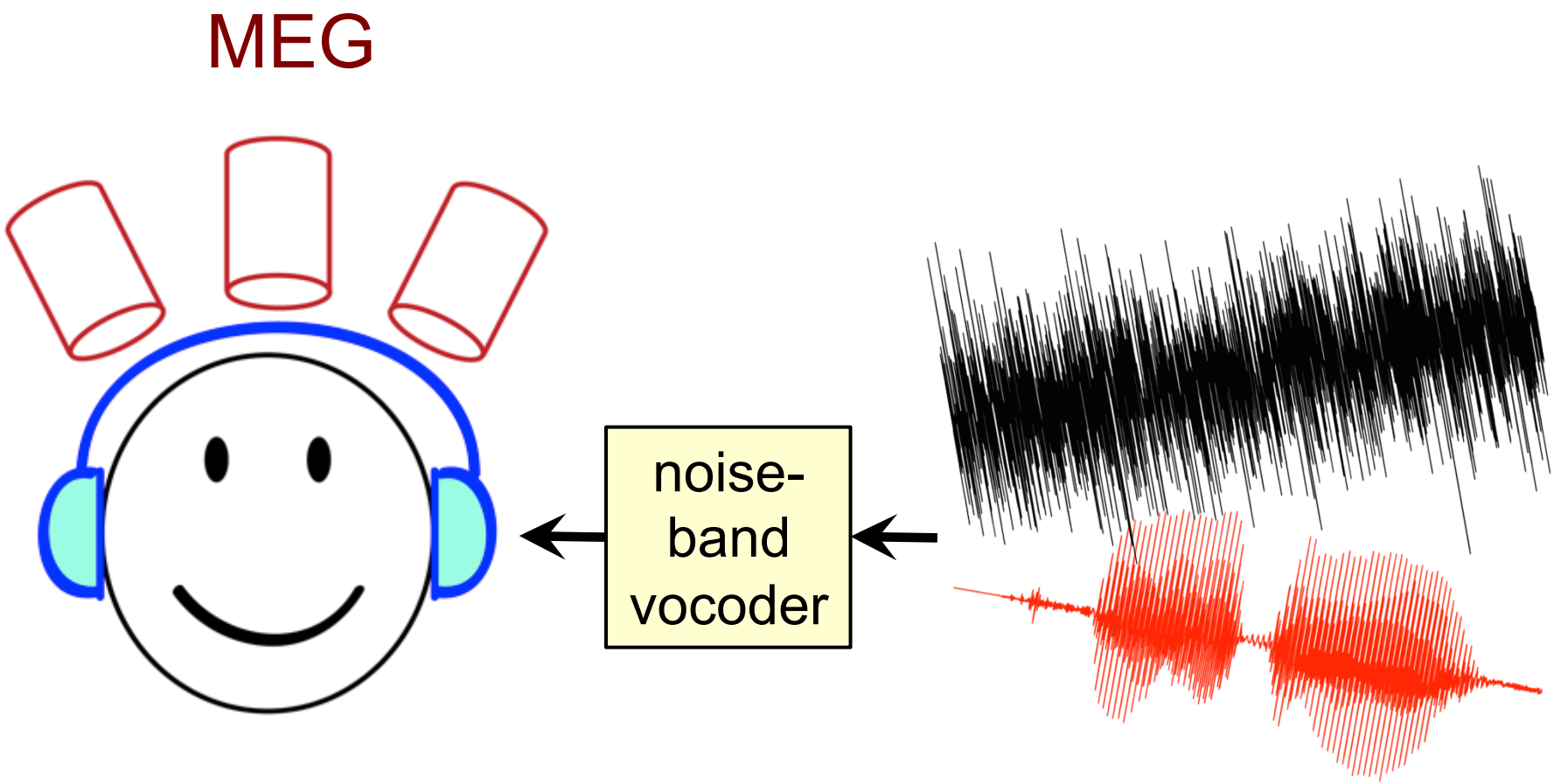
MEG Response to Speech

The MEG response is precisely synchronized to the temporal envelope of speech.

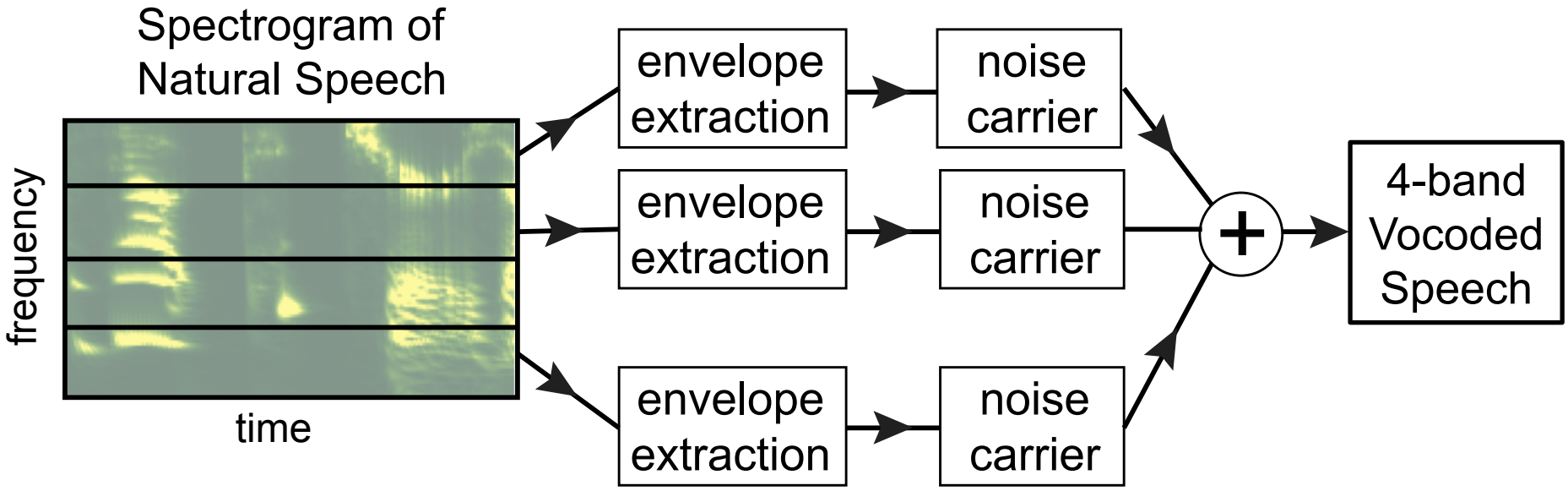


Ding & Simon, J Neurophys (2012)
Ding & Simon, PNAS (2012)

Cochlear implant stimuli are *simulated* acoustically using noise band vocoding.

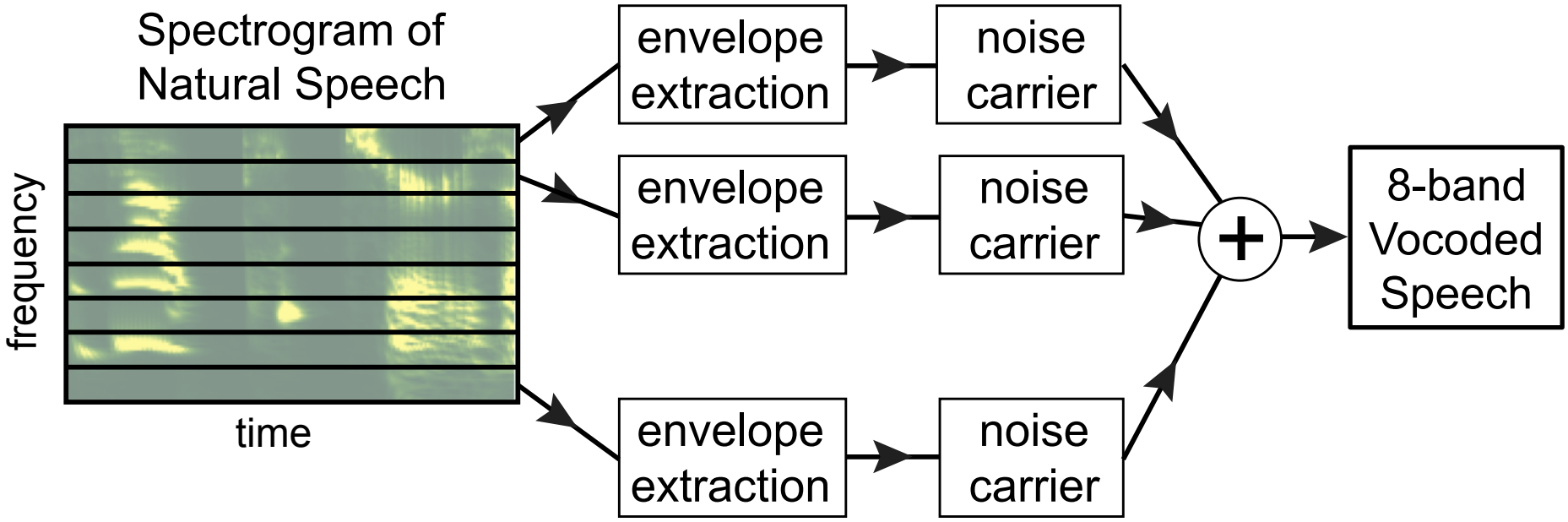


Noise-vocoded Speech



Noise-vocoded speech preserves the envelope of speech, but removes the spectro-temporal fine structure.

Noise-vocoded Speech

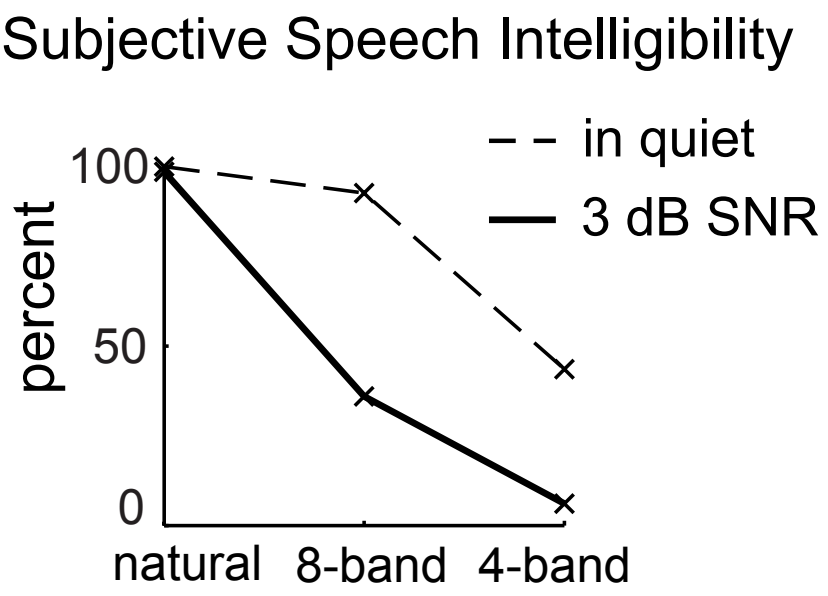


Vocoded speech is more intelligible when more frequency bands are used.

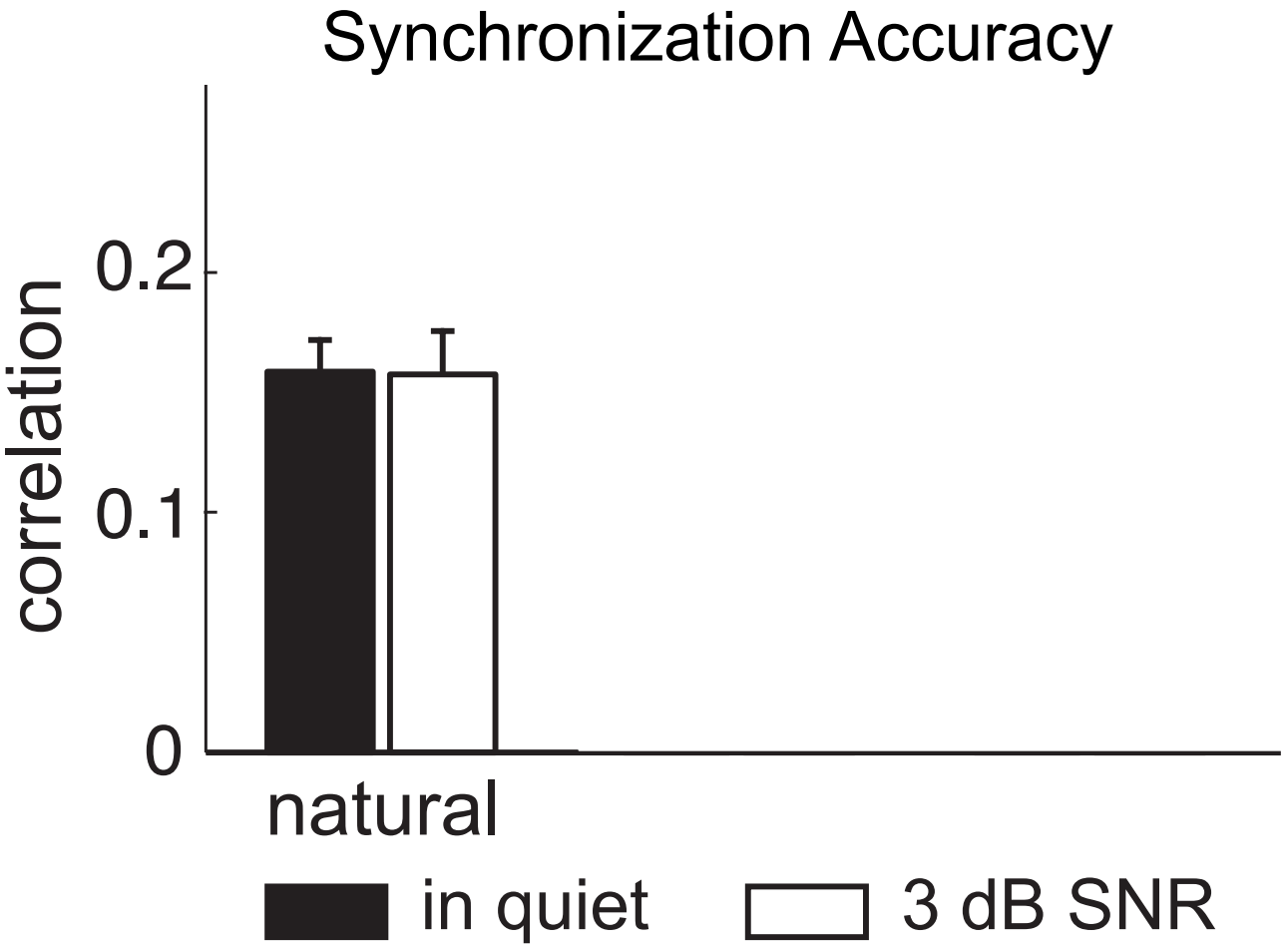
Experiment Design

- **Stimulus Conditions**
  - Three types of speech:
    - Natural Speech
    - 8-band Vocoded Speech
    - 4-band Vocoded Speech
  - Each stimulus is either presented in quiet or in stationary noise at -3 dB SNR.
- **Procedure**

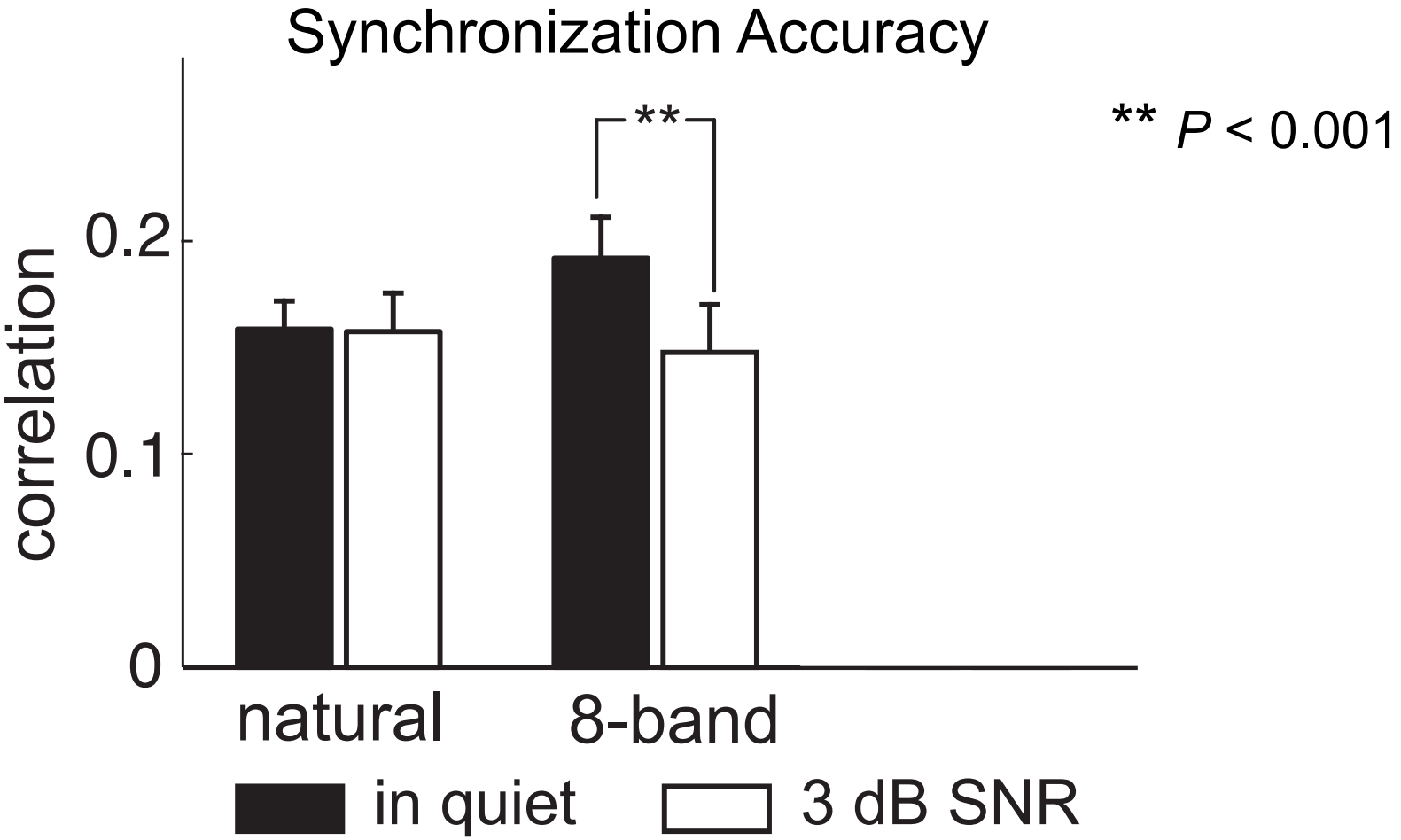
  Subjects listened to a story, answered questions and rated speech intelligibility.



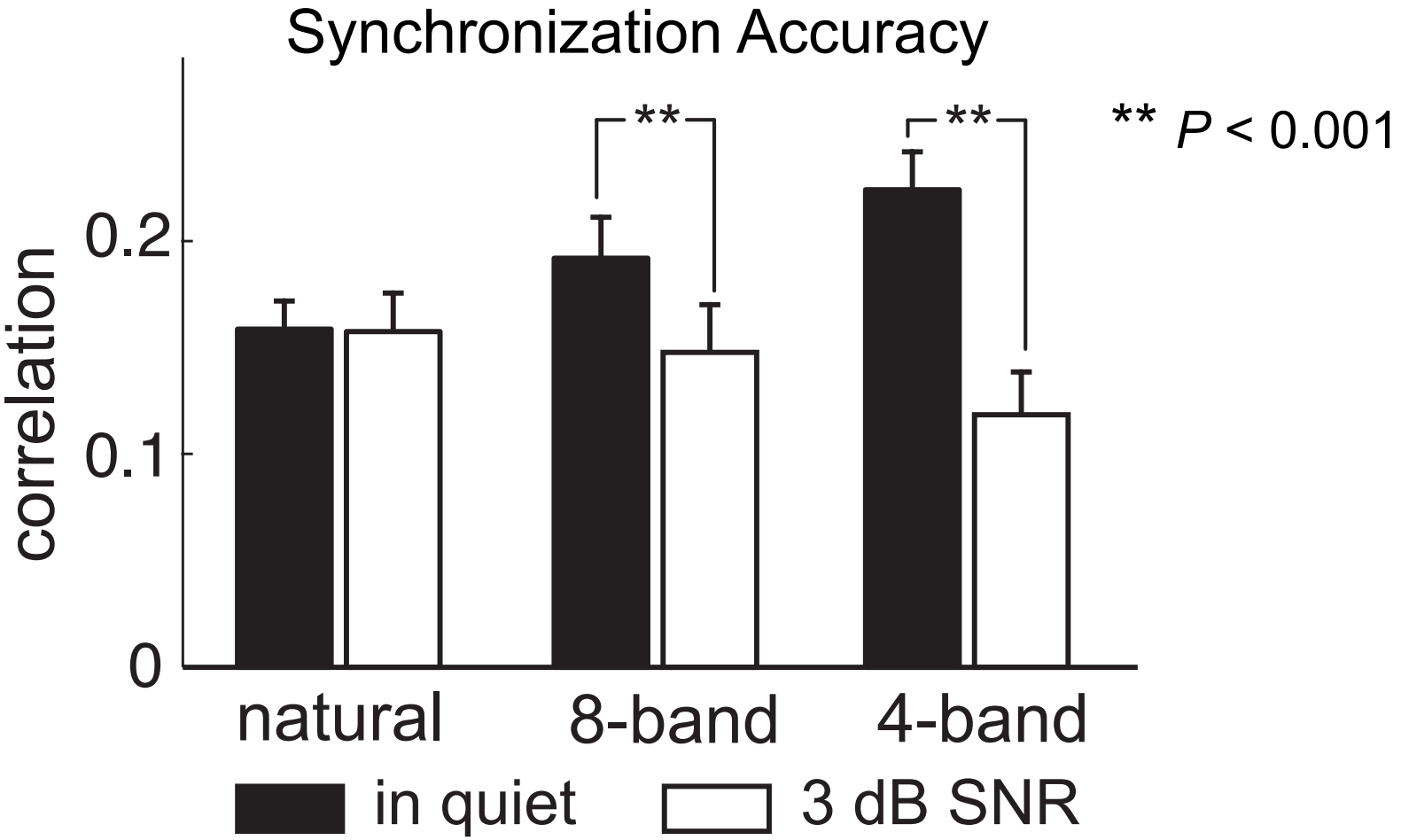
Neural synchronization to natural speech is resilient to noise



Neural synchronization to vocoded speech is susceptible to noise



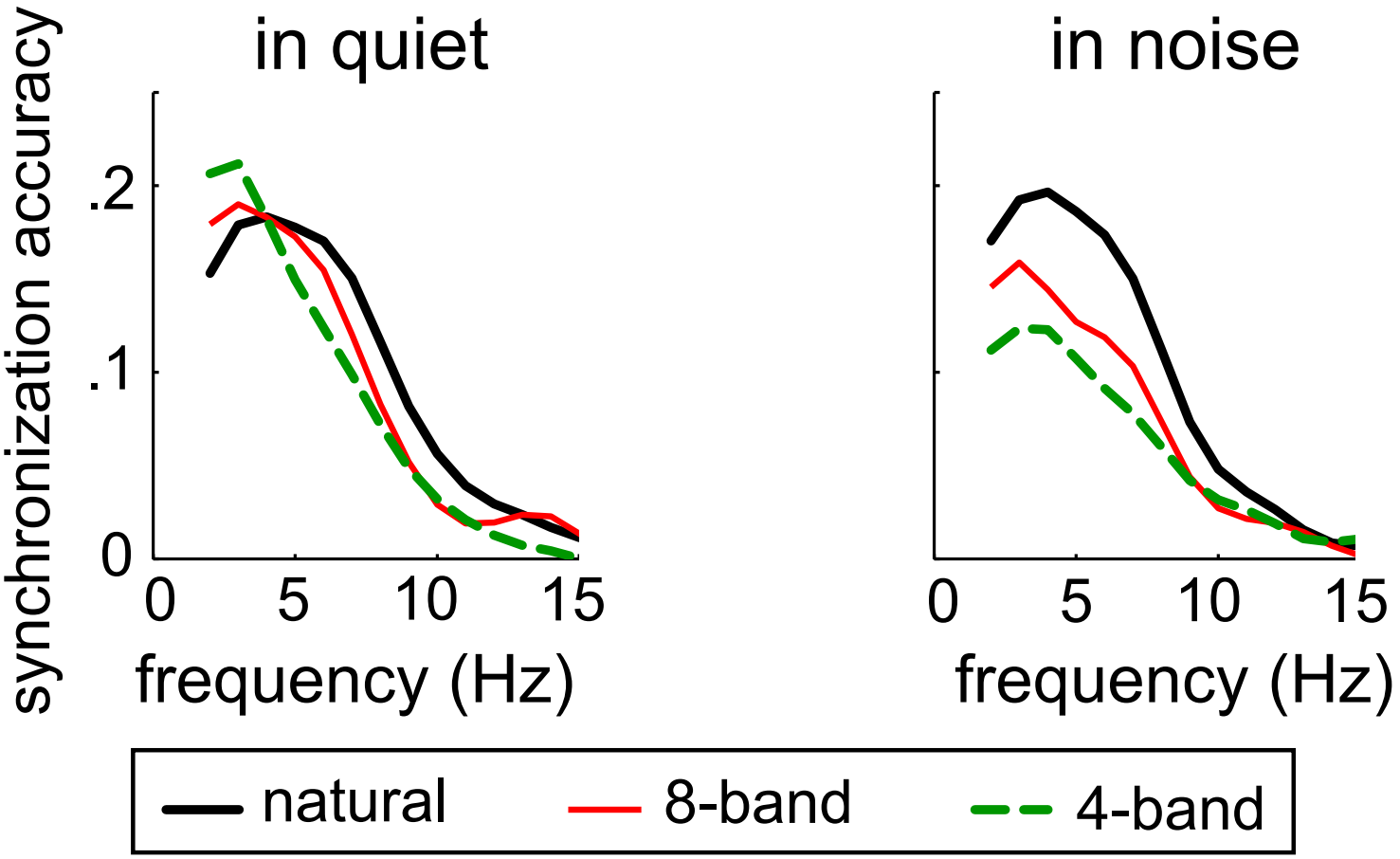
Neural synchronization is robust to noise, for natural but not vocoded speech.



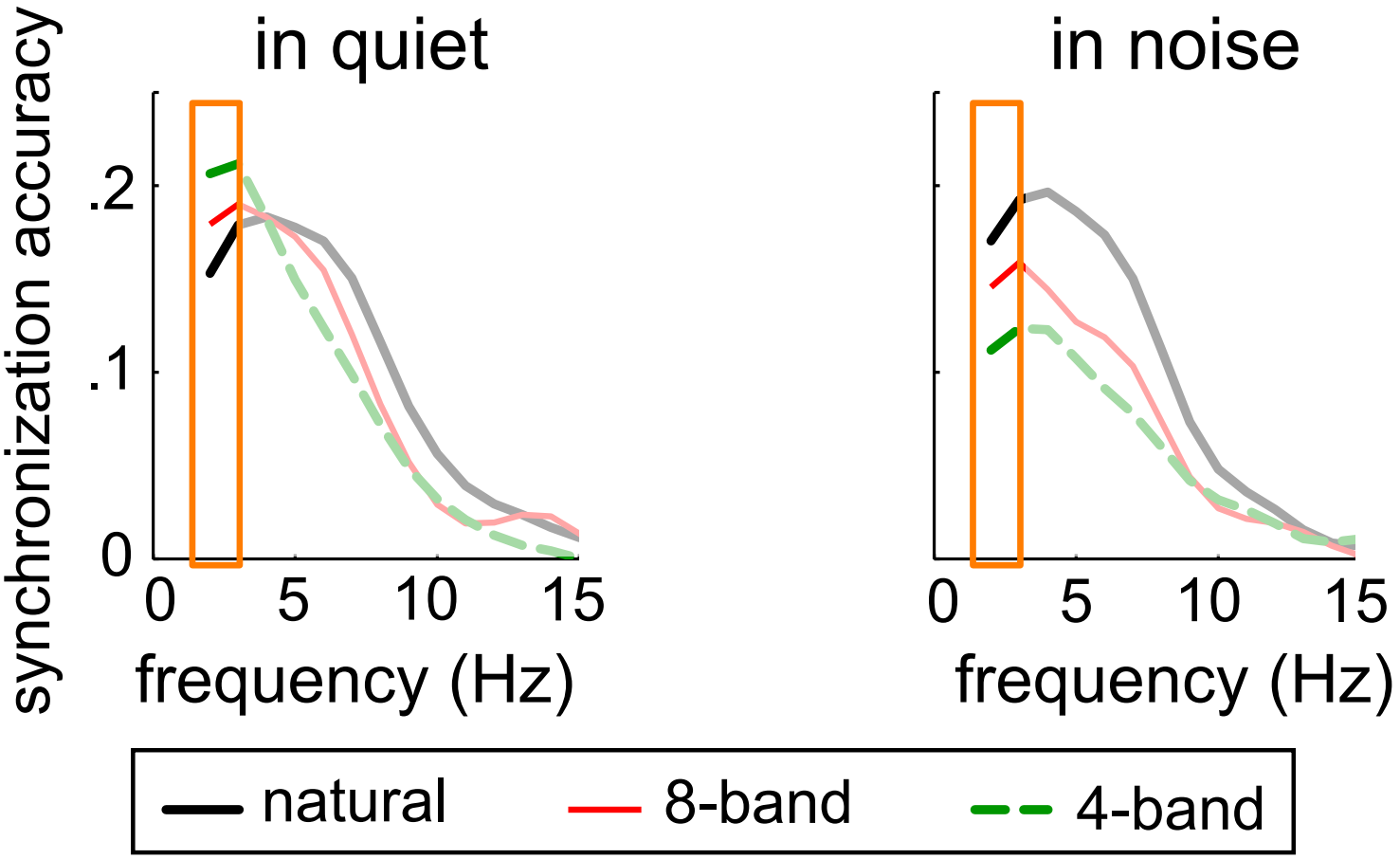
Summary I

The robustness of neural synchronization to speech relies on the spectro-temporal fine structure, and is lost when the fine structures are removed by noise vocoding.

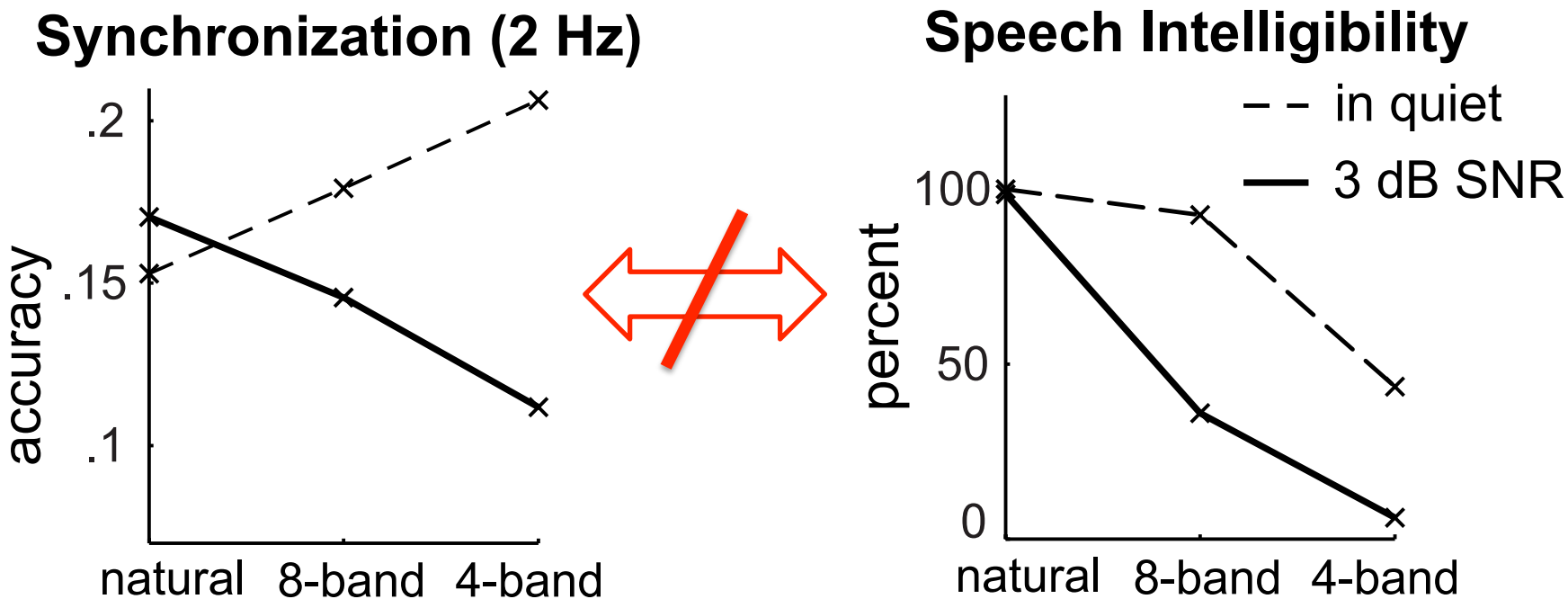
Synchronization Spectrum



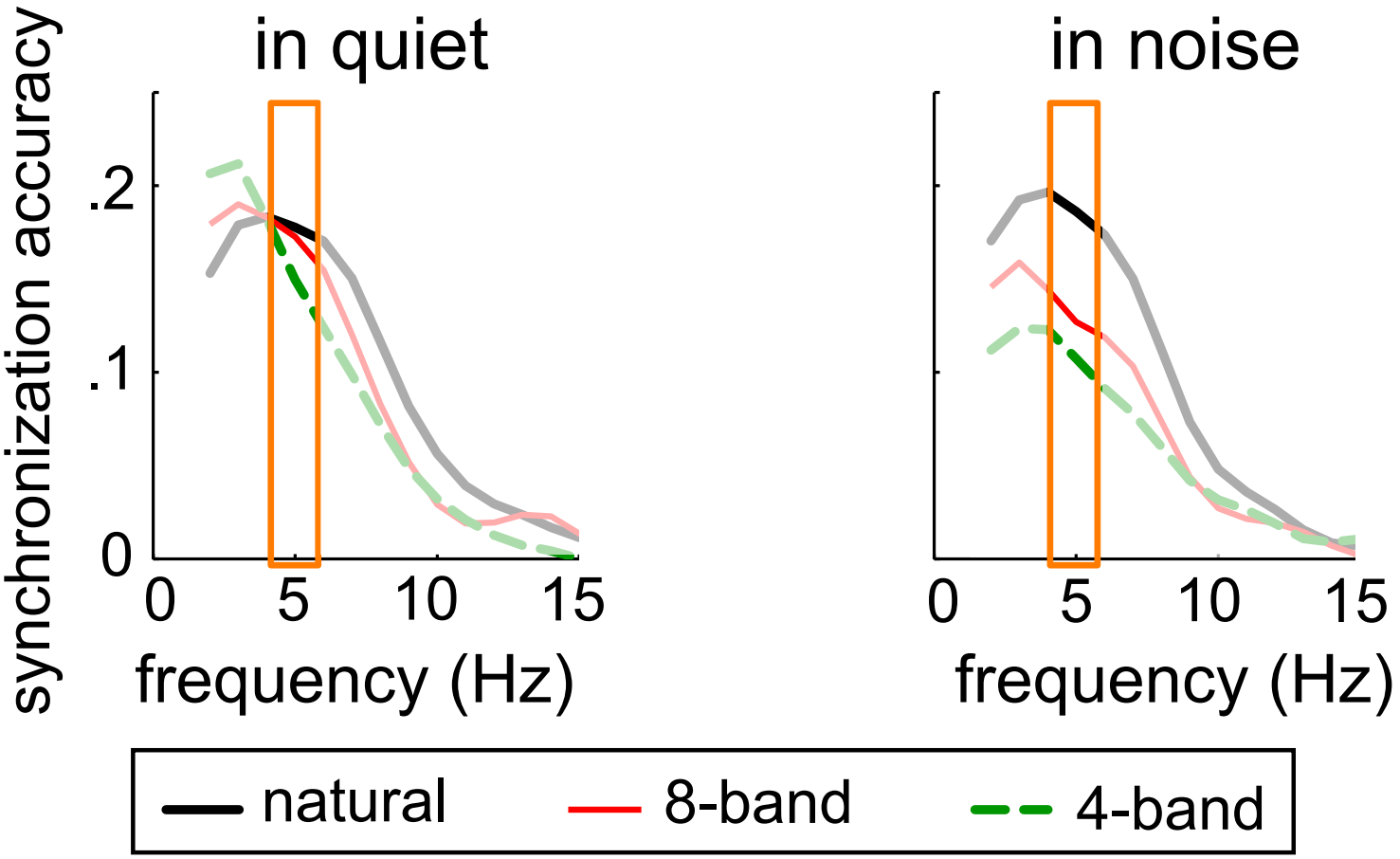
Synchronization Spectrum



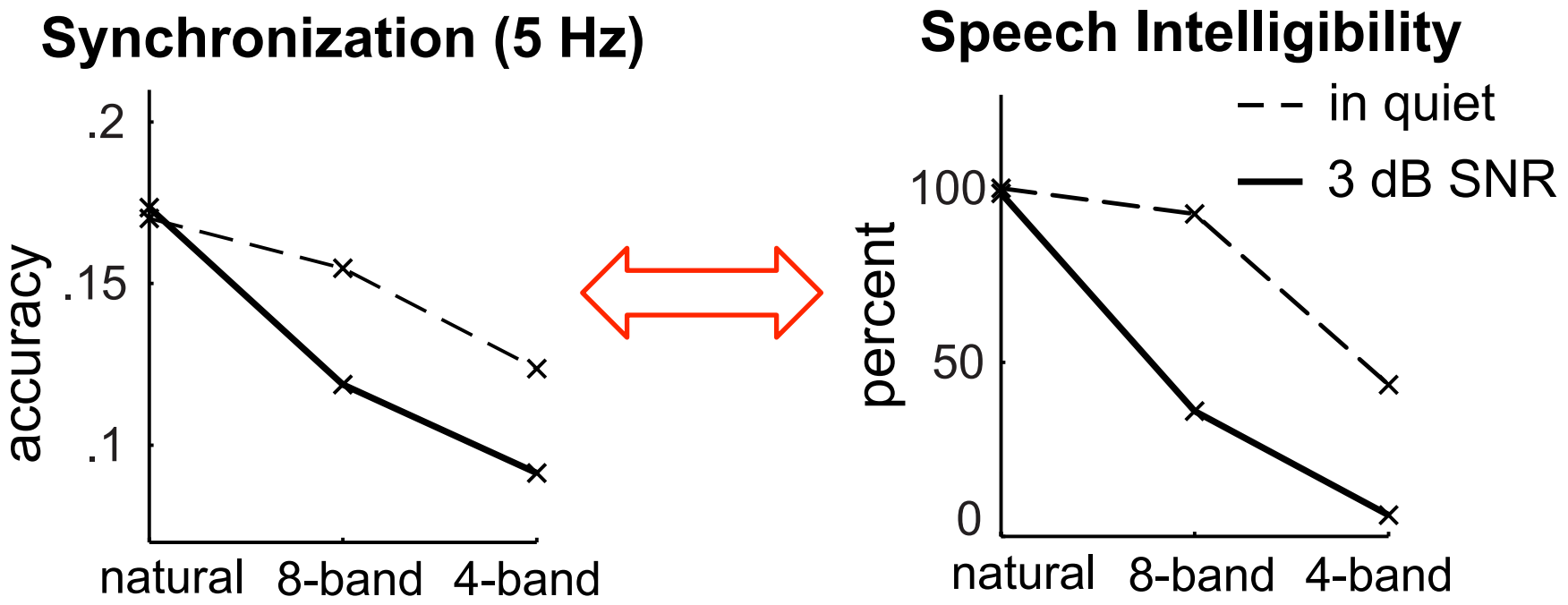
Synchronization at 2 Hz does not correlates well with speech intelligibility.



Synchronization Spectrum



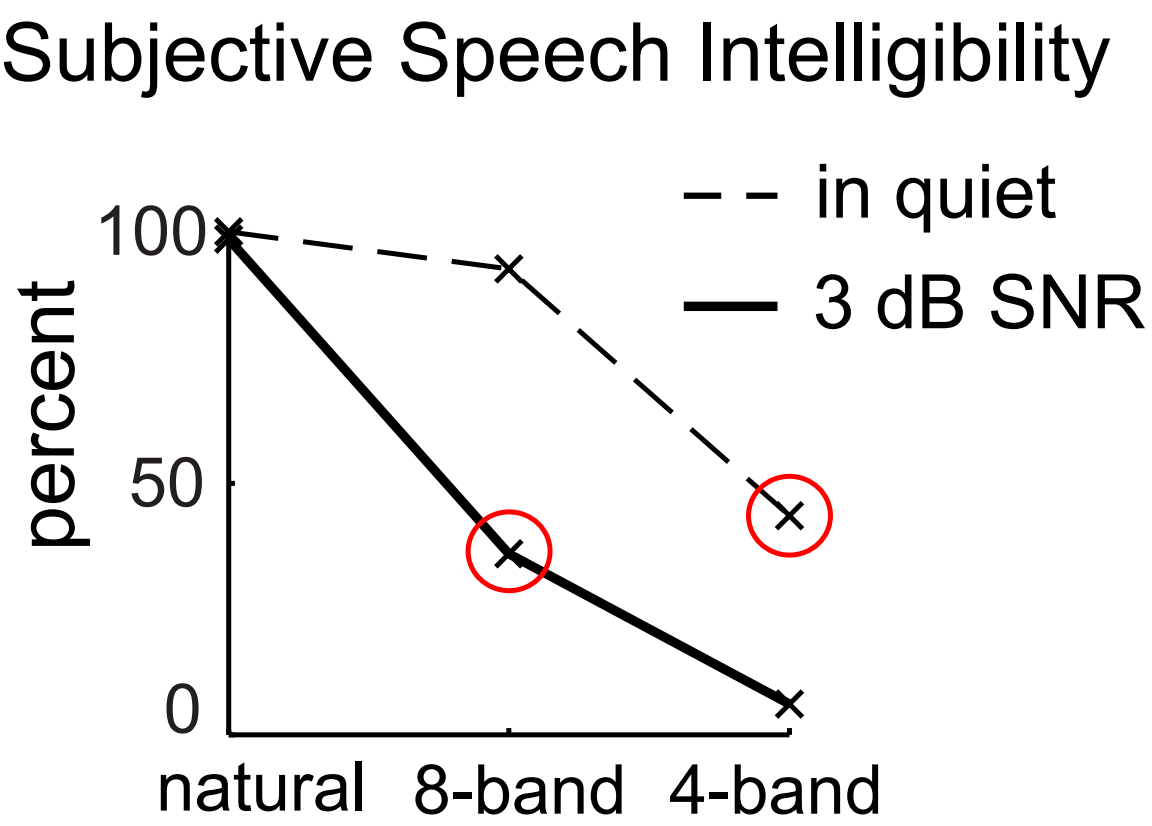
Synchronization at 5 Hz correlates well with speech intelligibility.



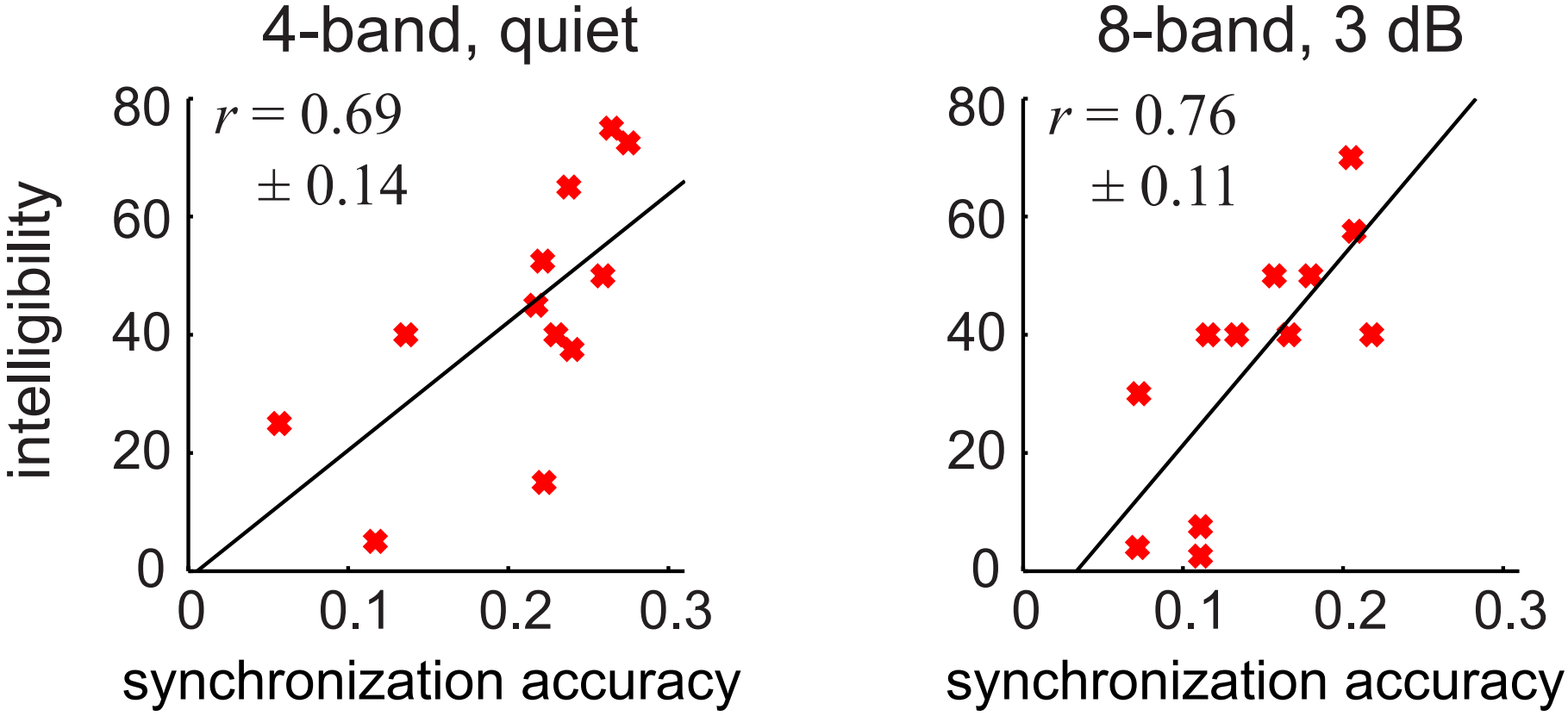
Summary II

In the grand average, neural synchronization at ~5 Hz but not ~2 Hz predicts how speech intelligibility varies across conditions.

Predictors for Individual Difference



Only synchronization near 2 Hz predicts individual's speech intelligibility.



Subjects showing better neural synchronization to speech tend to understand speech better.

Summary III

Neural synchronization at very the low frequencies near 2 Hz predicts individual differences in speech perception ability.

Conclusions

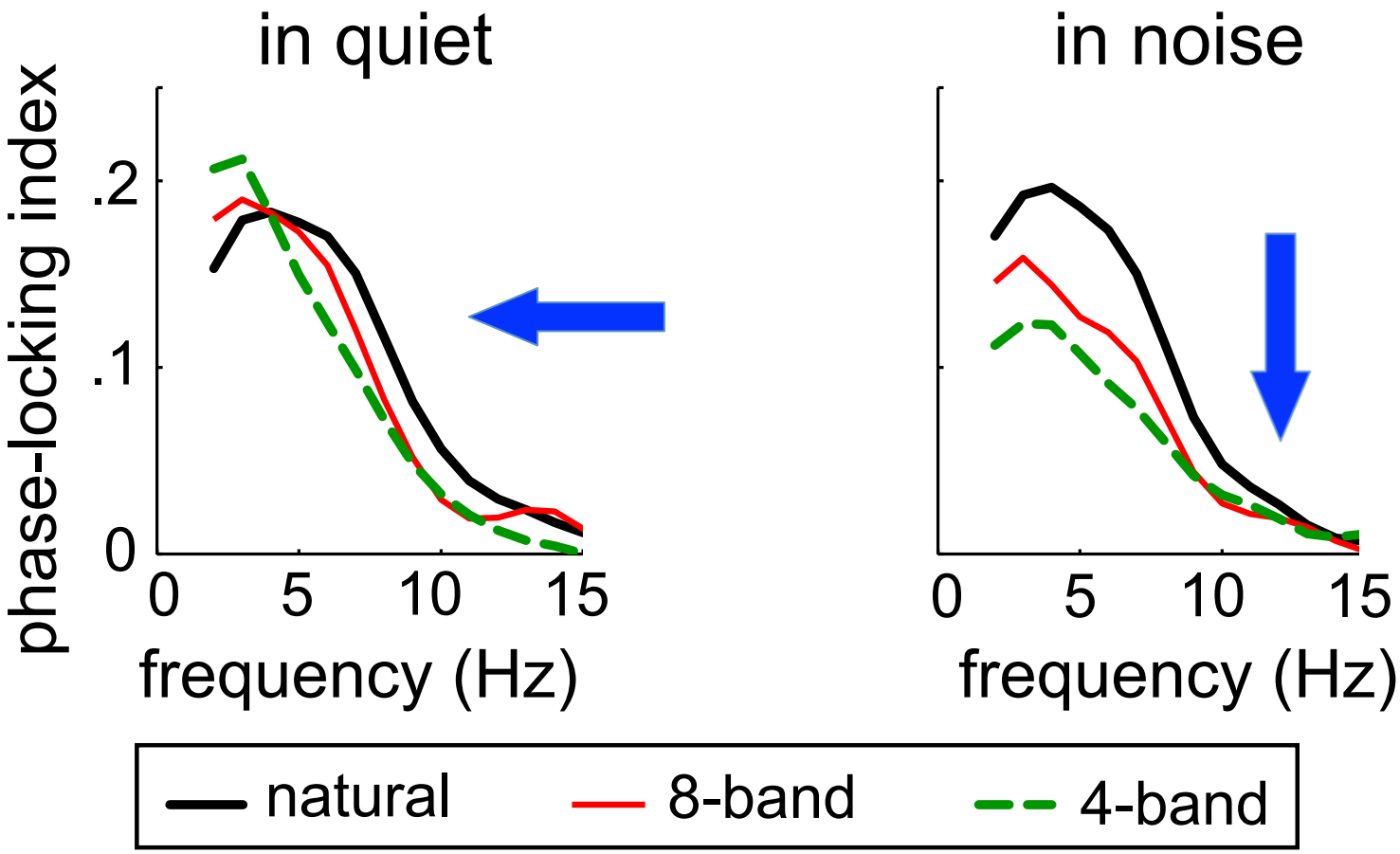
- ✓ Cortical synchronization to speech envelope is robust to noise, but this robustness relies on the spectro-temporal fine structure.

- ✓ Cortical synchronization near 5 Hz (theta band, syllable rate) reflects how speech intelligibility varies across different stimuli.

- ✓ Cortical synchronization near 2 Hz (delta band, stressed syllable) reflects how the intelligibility of the same stimulus varies across subjects.

Thank you!

Synchronization Spectrum



vocoding induced low-frequency shift

vocoding induced reduction in synchronization accuracy