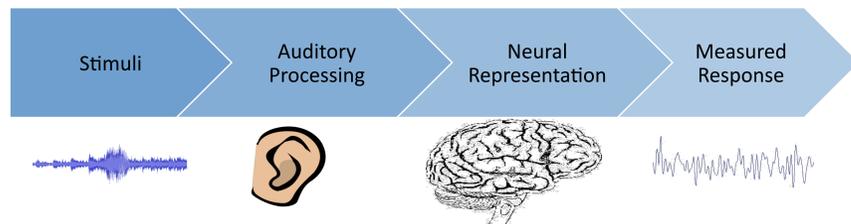


## Introduction

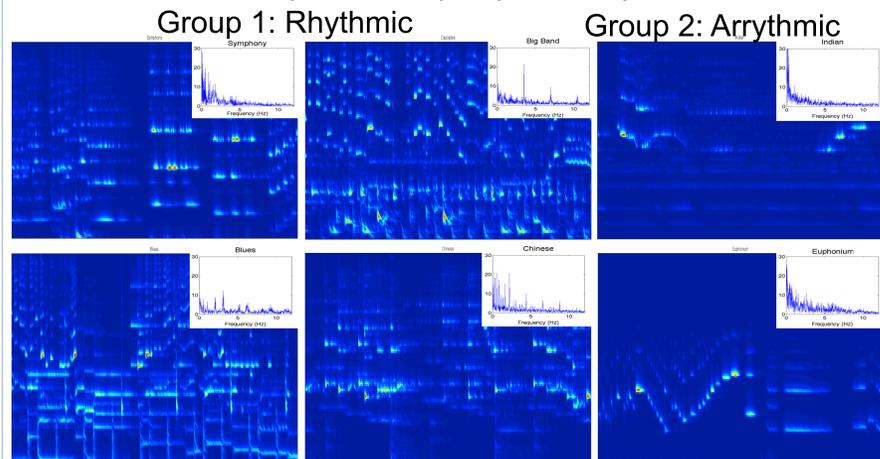
- Cortical neural representations of music in humans are investigated using analysis techniques already proven for cortical representations of speech (Ding and Simon, 2012).
- Neural responses to diverse samples of music are acquired with magnetoencephalography (MEG) and characterized by a temporal response function (TRF).
- Correlations between actual and predicted responses are comparable to those of speech stimuli.



## Methods

### Stimuli

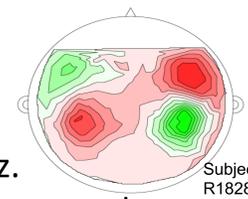
- Six 60 s samples of spectrally and rhythmically diverse instrumental music from different genres were played binaurally to 12 subjects.
- The samples were pseudorandomly presented in three trials each, within 6 stimulus blocks.
- Subjects were probed on whether a later 5-second clip was a sample excerpt or not (performance=81%).
- Stimuli can be separated by rhythmicity.



## Methods

### MEG Recordings

- 157 MEG channels sampled at 1 kHz over the whole head, resampled to 40 Hz.
- All stimuli generate a robust auditory response, shown above for a representative subject.



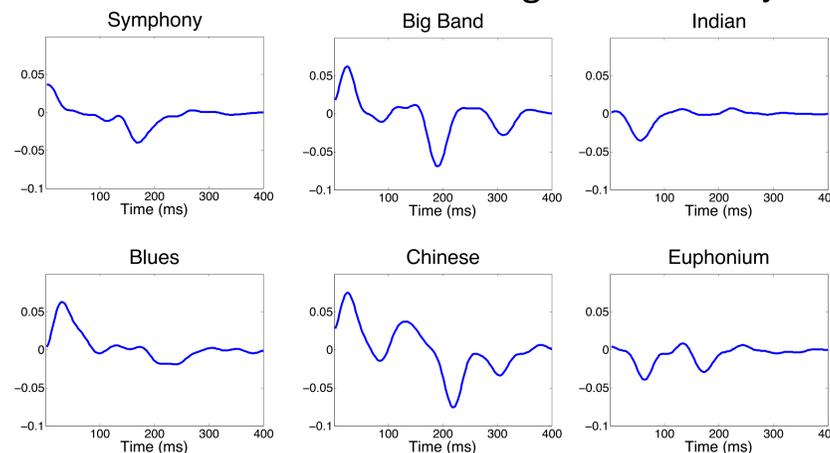
### Analysis: Temporal Response Function

- A TRF is convolved with the stimulus envelope to generate the predicted response.
- TRFs are estimated for each stimulus using boosting with 10-fold cross validation.

$$r(t) = \int_{\tau} TRF(\tau) S(t - \tau) d\tau + \epsilon(t)$$

## Results

### TRFs for each stimulus averaged across subjects



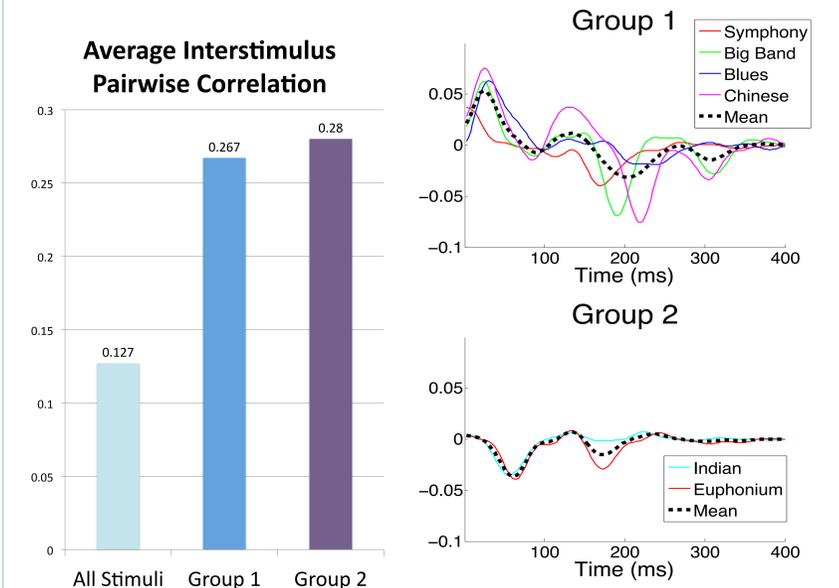
- TRFs had significant (>0.15) predictive power and intersubject pairwise correlations
- Predictive power averaged across subjects and stimuli (0.22) comparable to speech case
- Intersubject pairwise correlation averaged across stimuli (0.31) comparable to speech case
- TRFs more correlated across subjects than across stimuli

### References and Acknowledgments

- NSF OCI #1063035
- NIH R01-DC-008342
- Ding, N. and Simon, J.Z. (2012), "The Emergence of Neural Encoding of Auditory Objects While Listening to Competing Speakers", PNAS 109(29).

## Observations

- Separating TRFs by the polarity of their initial peak results in a large increase in inter-stimulus pairwise correlation (mean across subjects).
- Neural responses were predicted by qualitatively different models depending on the presence or absence of rhythm in the stimulus (see spectrogram).



## Summary

### Conclusions

- Linear impulse response models of significant predictive power were obtained per musical scene, and their profiles were highly consistent across subjects.
- Classification of TRF profiles by polarity of the first component is consistent with the presence/absence of strong rhythm in the music.

### Further questions

- Streaming of musical scenes/instruments: what does attending to a single scene/instrument imply for the competing background?