Decoding Music Representation in the Human Brain using EEG Data

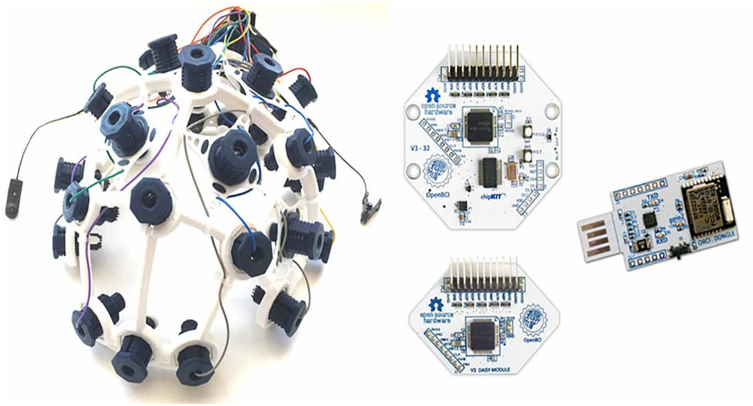
Chris Foster, Dhanush Dharmaretnam, Haoyan Xu,
Alona Fyshe, and George Tzanetakis

University of Victoria

University of Alberta

Why investigate EEG data?

- Prior research in Computational Linguistics
- Study language representations
- Augment features
- Young growing field in MIR
- Datasets: NMED-H, OpenMIIR etc.



www.opensourceimaging.org

Goals

- Study the correlation between music features and EEG data
- Evaluate first steps to creating semantic audio feature vectors
- Exploratory analysis on a popular EEG music dataset

OpenMIIR Dataset

- 10 subjects listening to and imagining 12 short music fragments

  - each 7s-16s long

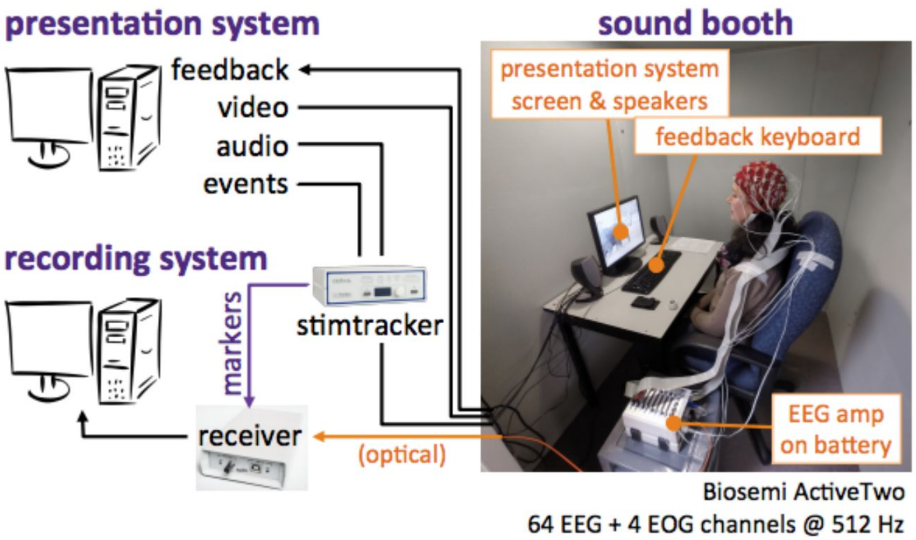
- Songs are from different genres, have varying tempos, some use vocals, etc
- There were 5 trials per participants with songs in random order
- 60 songs per participant (12 songs * 5 Trials)

OpenMIIR
public domain dataset of EEG recordings
for music imagery information retrieval

S Stober, 2015

Collection Methodology

1 Chim Chim Cheree (lyrics)
2 Take Me Out to the Ballgame (lyrics)
3 Jingle Bells (lyrics)
4 Mary Had a Little Lamb (lyrics)
5 Chim Chim Cheree
6 Take Me Out to the Ballgame
7 Jingle Bells
8 Mary Had a Little Lamb
9 Emperor Waltz
10 Hedwig's Theme (Harry Potter)
11 Imperial March (Star Wars Theme)
12 Eine Kleine Nachtmusik



S Stober, 2015

EEG Data Preprocessing

- Remove bad channels
- Remove the ICA features
- Normalize the dataset

Prior work on the music EEG domain

- Correlate music features such as tempo, MFCC, zero crossings, CQT with brain data (NMED-H dataset) [1]
- S Stober trained an SVM to distinguish between 12 songs from the EEG data (18-27% accuracy with feature extraction) [2]
- Predicted tempo and rhythm perception from the EEG data [3]

[1] Cong, 2013
[2] S Stober, 2016
[3] S Stober, 2016

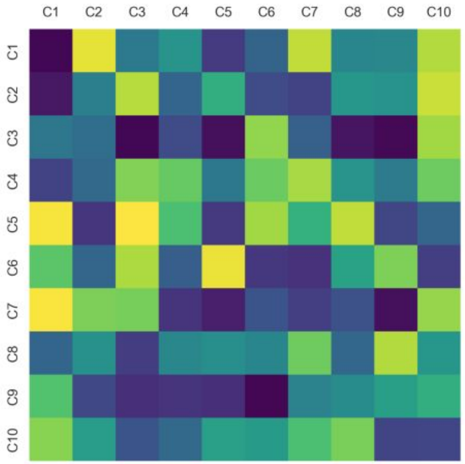
Creating audio vectors

- Inspired by “semantic word vectors” for computational linguistics
- Built audio vectors utilizing MFCCs, RMSE, spectral centroid, chroma STFT, spectral roll-off, tempogram, harmonic, and beat features
- Built vectors from tag-classification neural networks

Representational Similarity Analysis

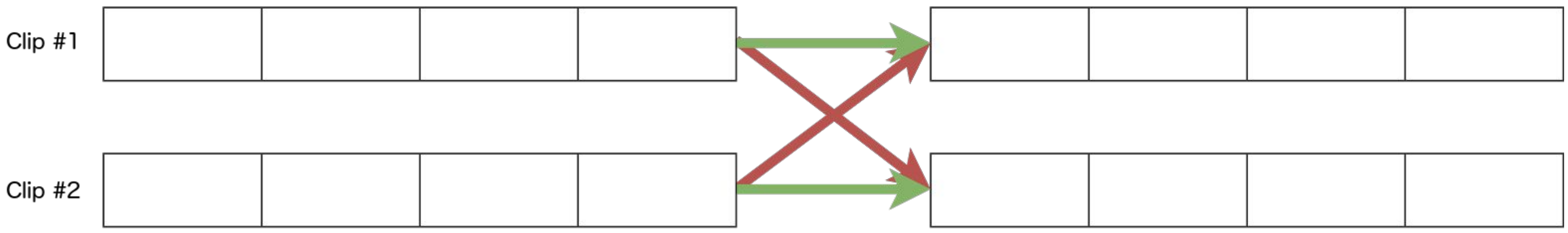
- Calculate a pairwise correlation matrix of EEG for the set of songs
- Calculate a pairwise correlation matrix of audio vectors for the set of songs

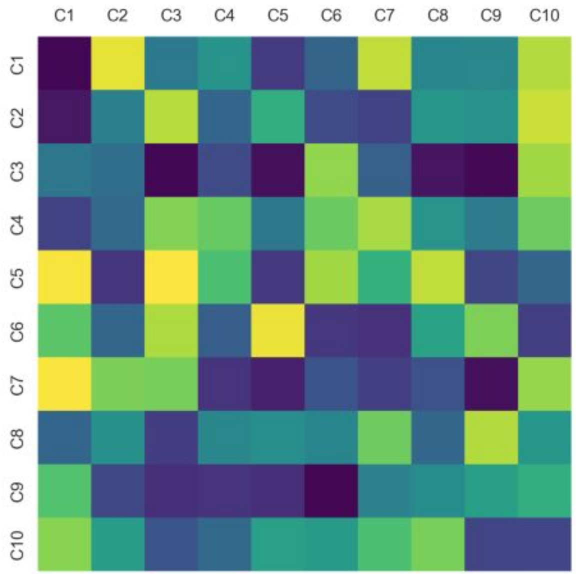
	Sing 1	Song 2	Song 3
Sing 1	1	0.2	0.8
Sing 2	0.2	1	0.3
Song 3	0.8	0.3	1

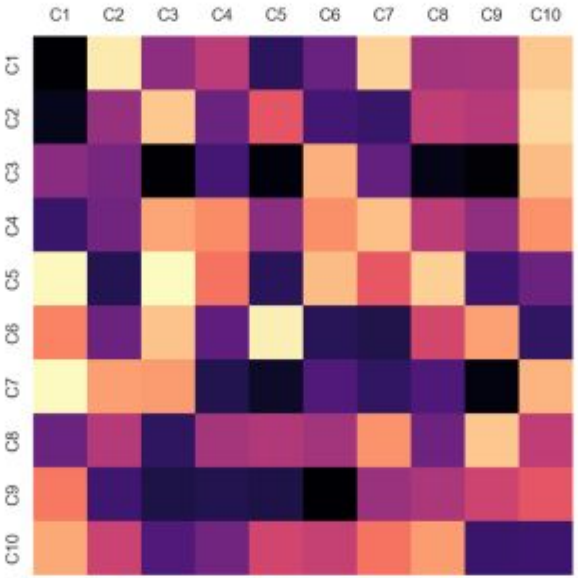


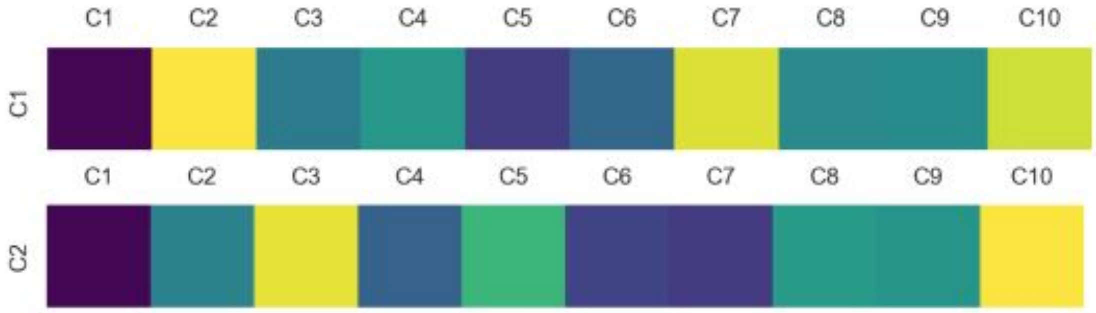
“2 vs 2” Correlation Test

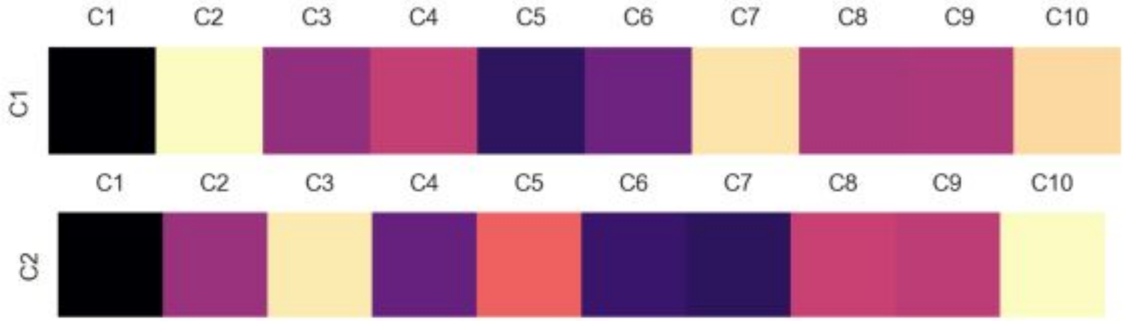
- Simplify vector comparisons into binary classification task
- Used often in computational neurolinguistics
- “Leave two out” benchmark analysis

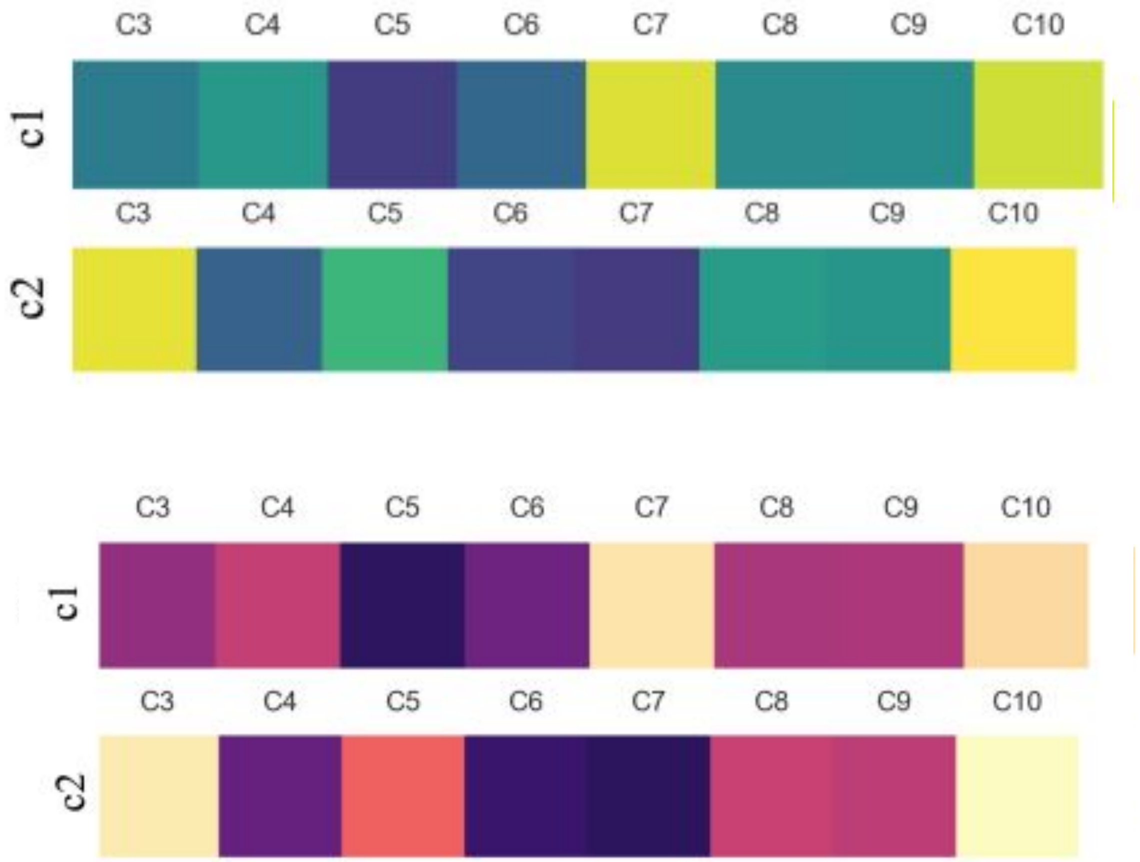


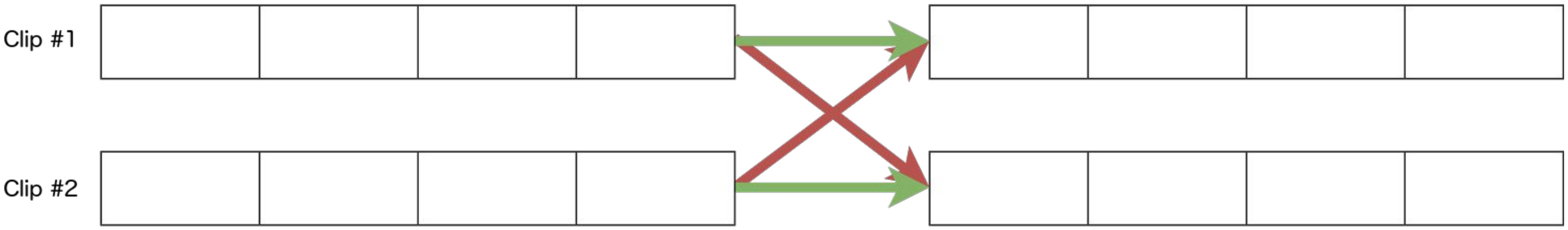

C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10

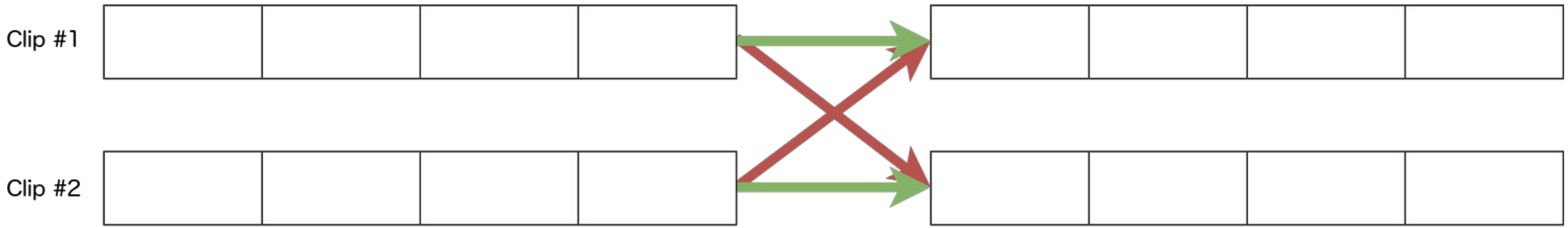

C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10

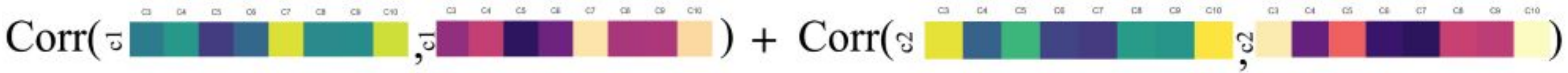

C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
C1
C2

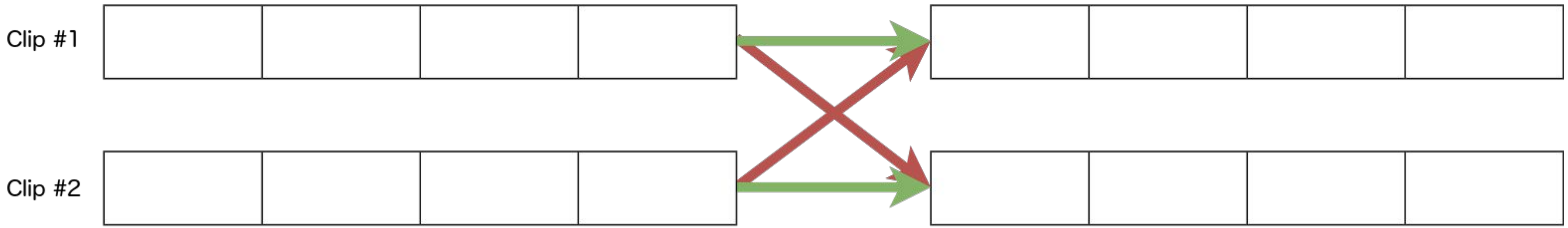

C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
C1
C2

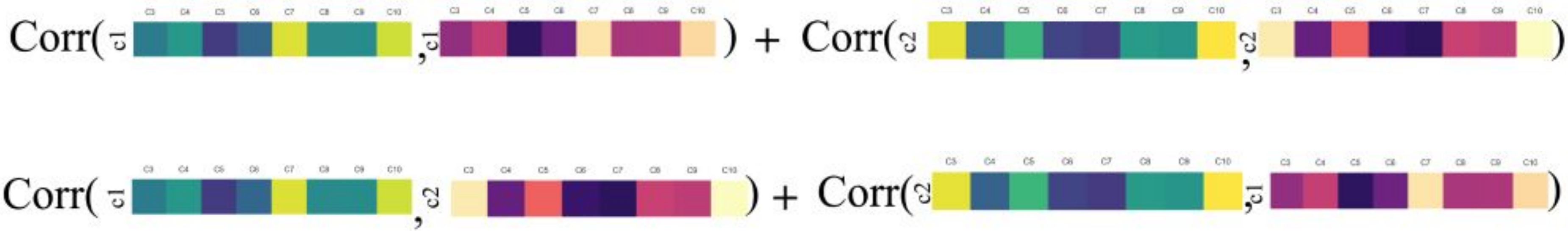

C3
C4
C5
C6
C7
C8
C9
C10
c1
C3
C4
C5
C6
C7
C8
C9
C10
c2
C3
C4
C5
C6
C7
C8
C9
C10
c1
C3
C4
C5
C6
C7
C8
C9
C10
c2


Clip #1
Clip #2


Clip #1
Clip #2

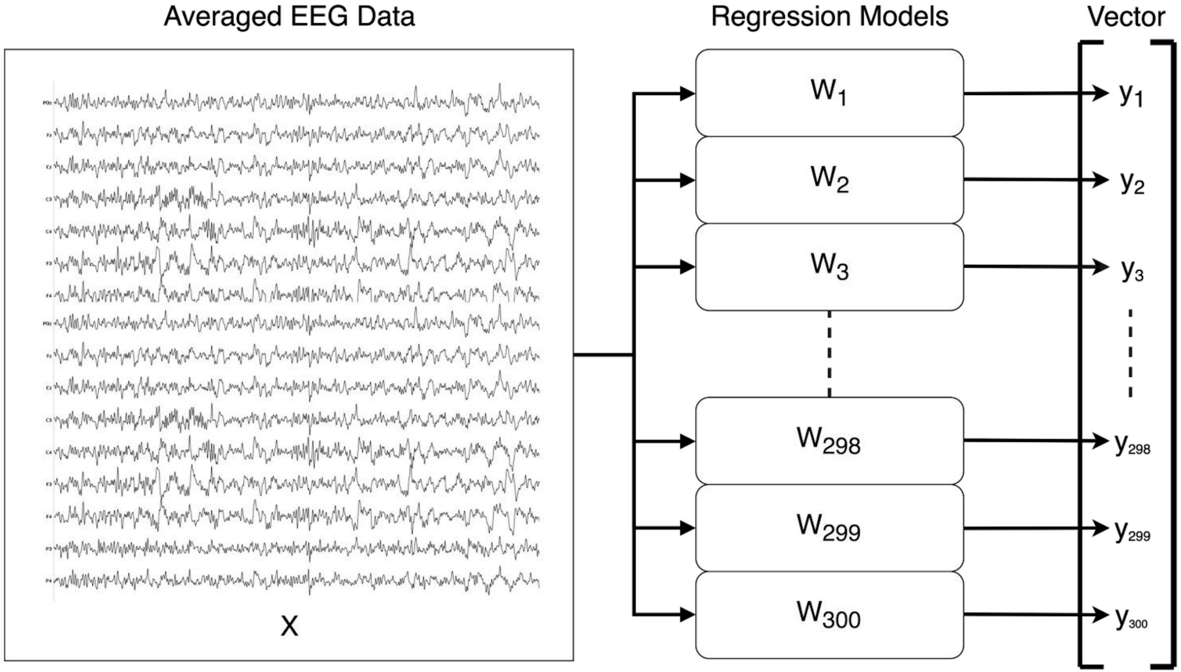

Corr(c1 , c1) + Corr(c2 , c2)


Clip #1
Clip #2


Corr(
c1
,c1
) + Corr(
c2
,c2
)
Corr(
c1
,c2
) + Corr(
c2
,c1
)

Linear Models Approach

- Machine learning model - set of linear ridge regressors

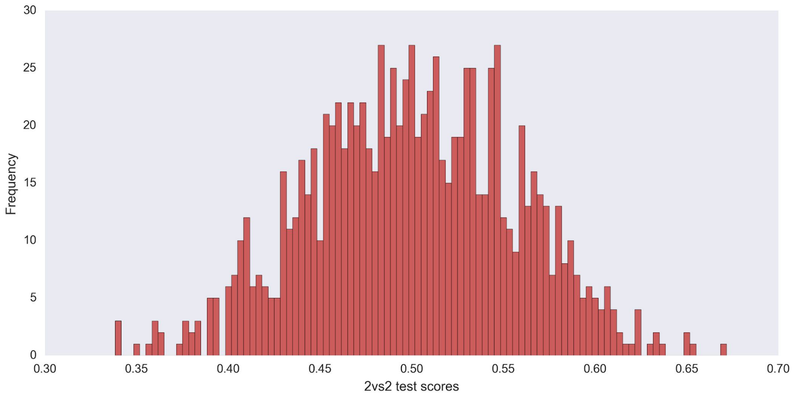


Correlation Results

- We've explored various features for correlating with the brain data
  - MFCC coefficients, tempo, beat features, chroma, mel spectrogram features....
- We've detected evidence of correlation!

MFCCs: **0.62** Tempogram: **0.63**

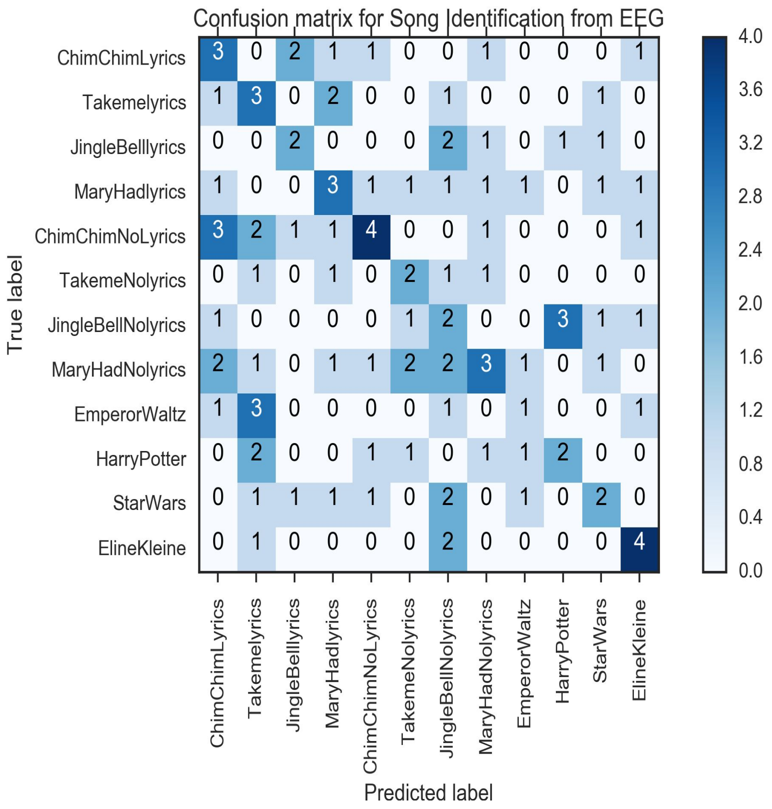
- However, more thorough testing and permutation tests are needed to confirm these results



Performing Song Identification

- Searching for correlation is a good opportunity to explore the dataset!
- We improved on the original song identification via logistic regression

Model	Accuracy
SVM (Raw)	18%
SVM (SCE)	27%
LogReg (Raw)	**29%**



Why might correlation be hard to find?

- This is a difficult task with very high noise
- There is a lot of data, and there is little data!
- Trials captured over multiple recordings
- Much of the data is mixed classes
- The “semantic song space” is small
- The “semantic song space” is narrow

The Perfect Experiment™

- Standardized track length
- Full 64 sensor recording
- Larger spread of audio genres
- Single-session recording
- Larger number of audio tracks
- Larger number of participants

Future Work

- Exploring regularization improvements
- Exploring combination of subject data
- Continuing iteration on audio feature vectors
- Further improvements / classifications on dataset
- Perform the Perfect Experiment™?

Summary

- Designed and built an experiment methodology for applying techniques from computational neurolinguistics to a new and exciting audio-EEG dataset
- Found evidence of correlation between the EEG activity of participants listening to songs and the features which we extracted from those songs
- However, this dataset has a number of limitations for this task!
- Improved over previous work on this dataset, using a greatly simplified model

Questions?