Using EEG to Decode Semantics During an Artificial Language Learning Task

Thesis Oral Defense
Chris Foster

University of Victoria

Can machine learning algorithms be used to track the acquisition of a new language in adults, based solely on electrical signals from the brain?

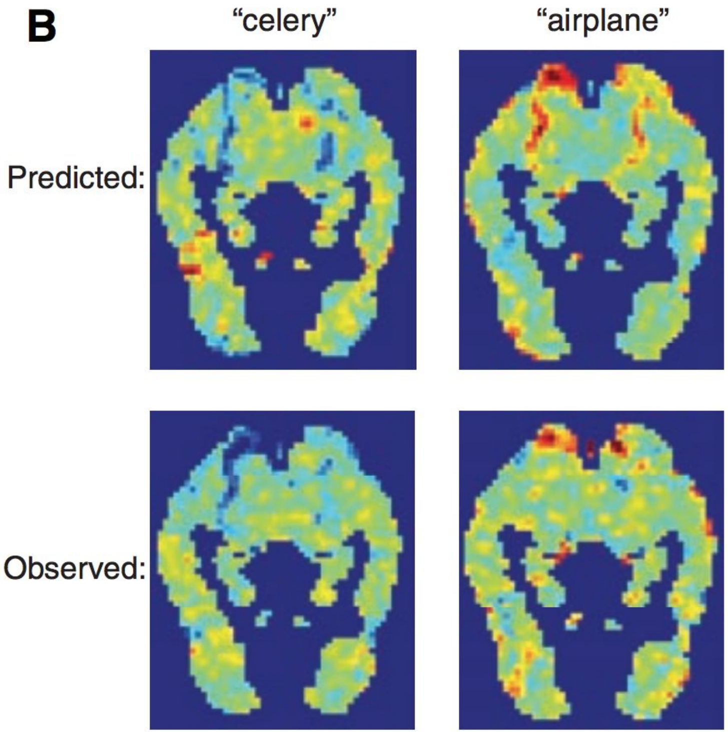
Current applications

Neuroscience

Accessibility

Linguistics

Predicting fMRI activity

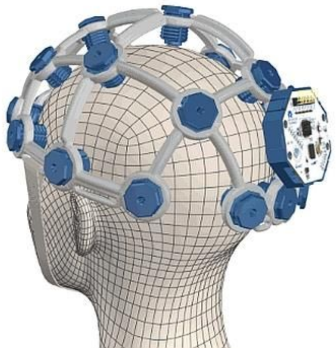


Mitchell, et al (2008)

Can we do this with EEG?

EEG is great due to its:

- Low cost of collection
- Mobility for collecting in various environments
- High time resolution



But, EEG has:

- Generally poorer spatial resolution
- Lower signal-to-noise ratio

Image per OpenBCI

Experiment paradigm

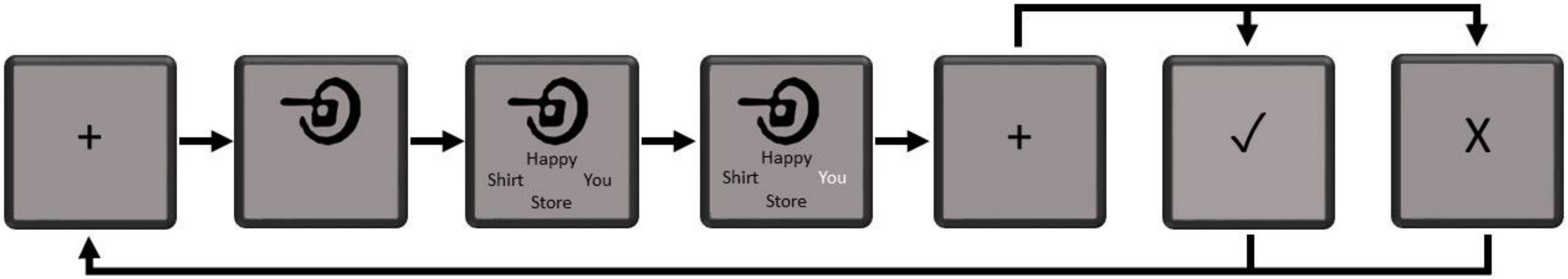


Figure per Chad Williams

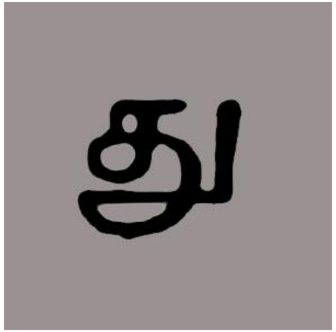

MARKET


CINEMA

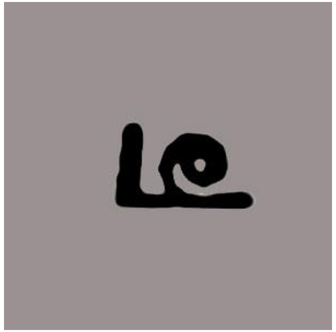

WE


TIRED

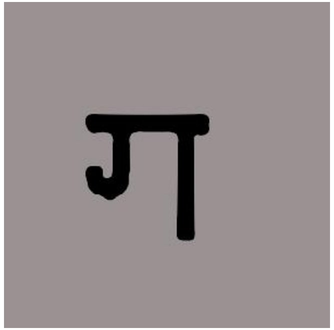

HUNGRY


PENCIL


YOU


WORK

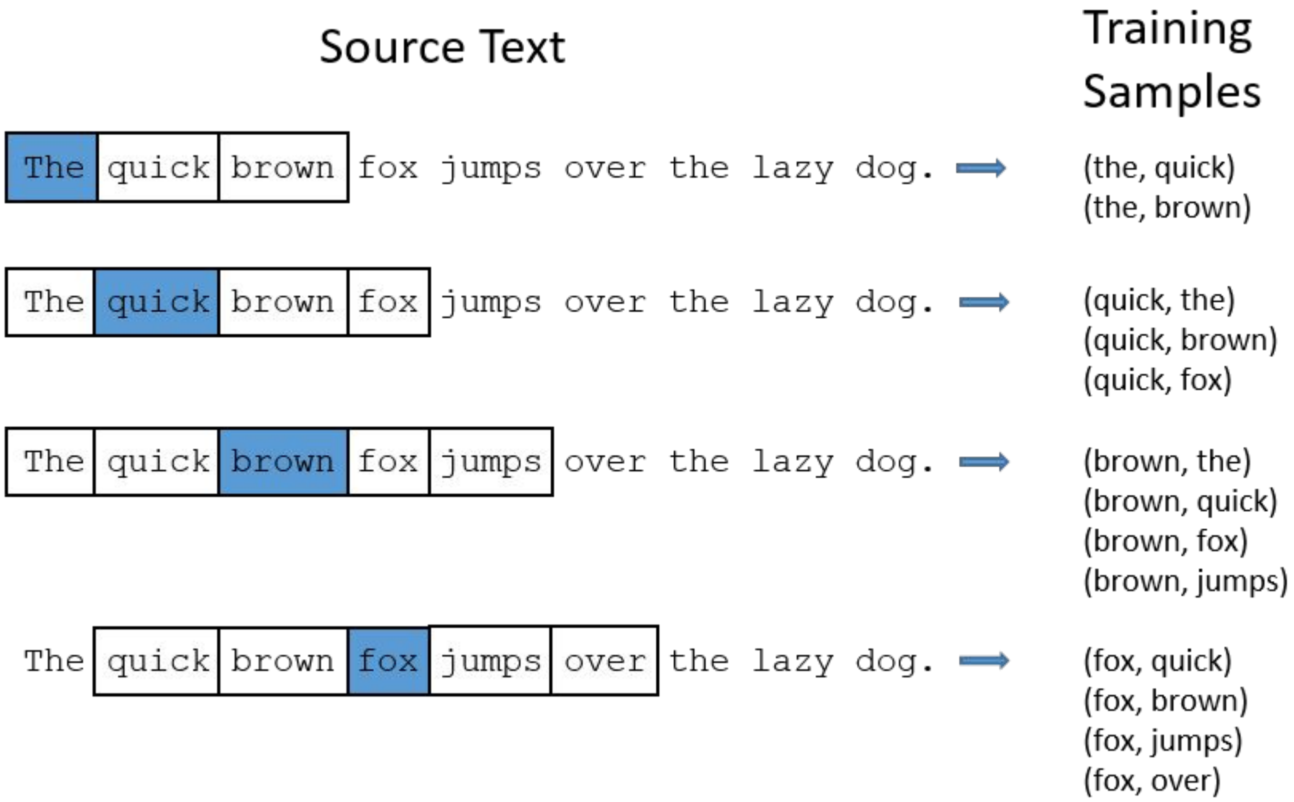

SALESMAN


CAMPING

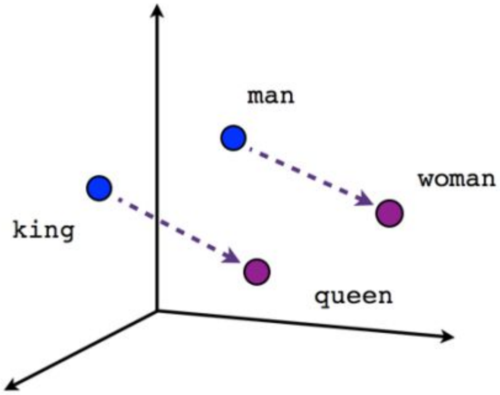
Machine learning analysis of brain data

Three components to discuss

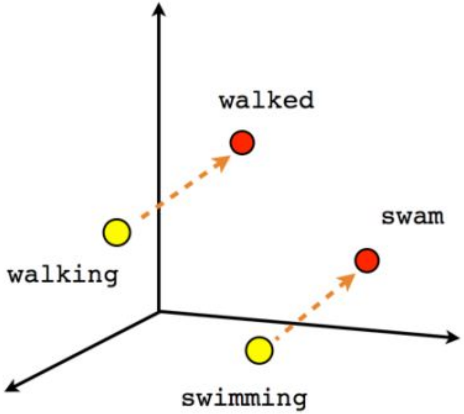
1. Input features
2. Regression model
3. Prediction target



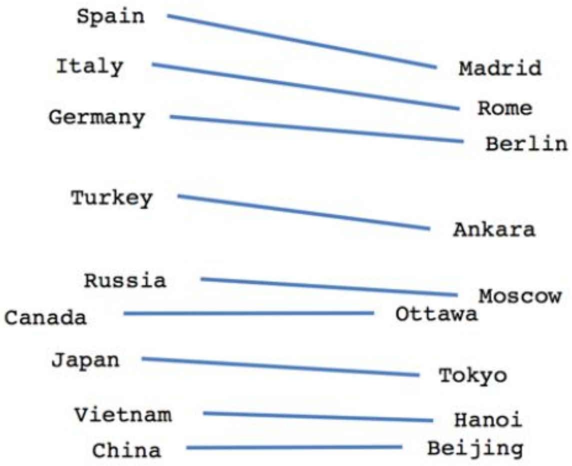
McCormick, C. (2016)



Male-Female



Verb tense



Country-Capital

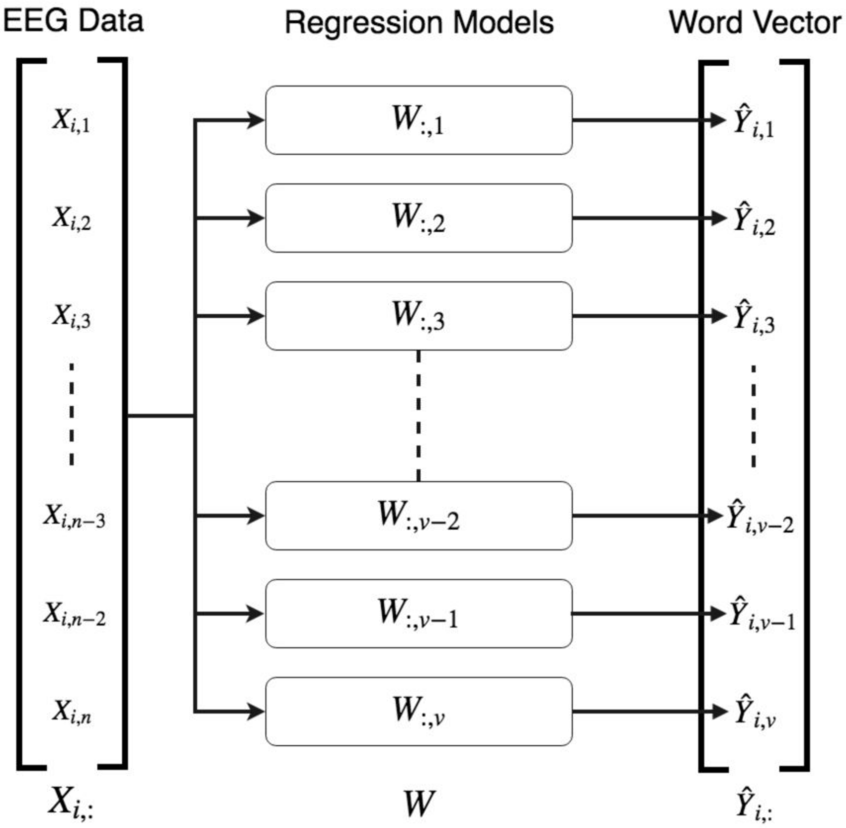
Ridge regression

Linear regression model with regularization via weight decay

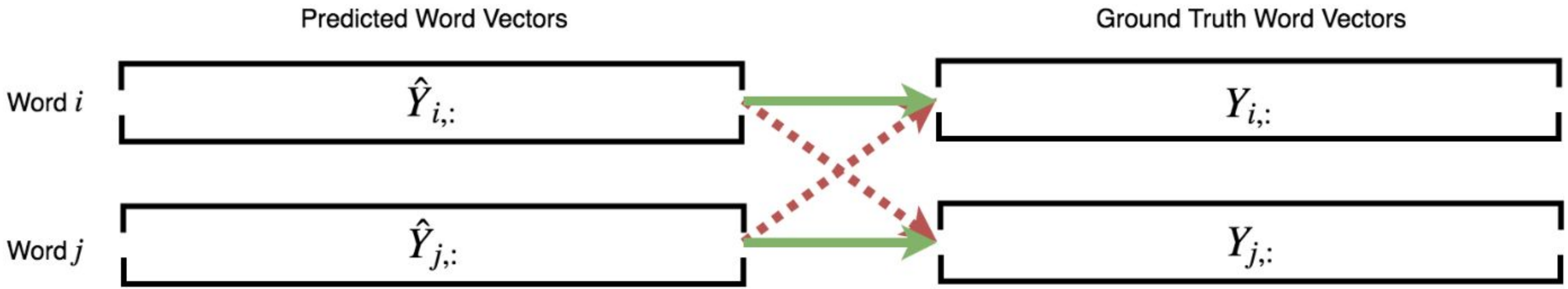
Basic linear regression: $$y = w_0 x_0 + w_1 x_1 + \ldots + w_m x_m = \sum_{j=0}^{m} = \mathbf{w}^T \mathbf{x}$$

Gradient descent using a loss function: $$J(\mathbf{w}) = \frac{1}{2} \sum_i (\text{target}^{(i)} - \text{output}^{(i)})^2$$

Multivariate regression model



Model evaluation



Research questions

1. Can we identify the semantics of the English word?

2. How much participant practice is needed to identify semantics?

3. How long does the brain take to process the semantics?

4. Are there areas of the brain that contribute more to identifying semantics?

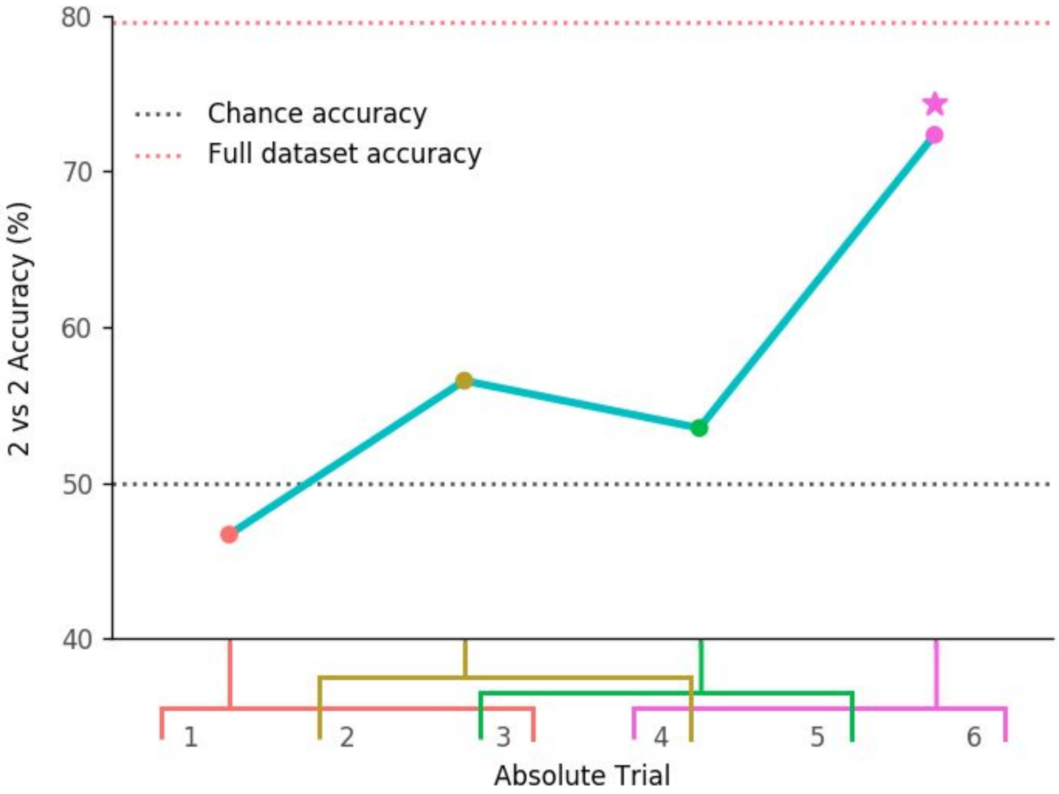
Can we identify the semantics of the English word?

Our machine learning model shows an accuracies of:

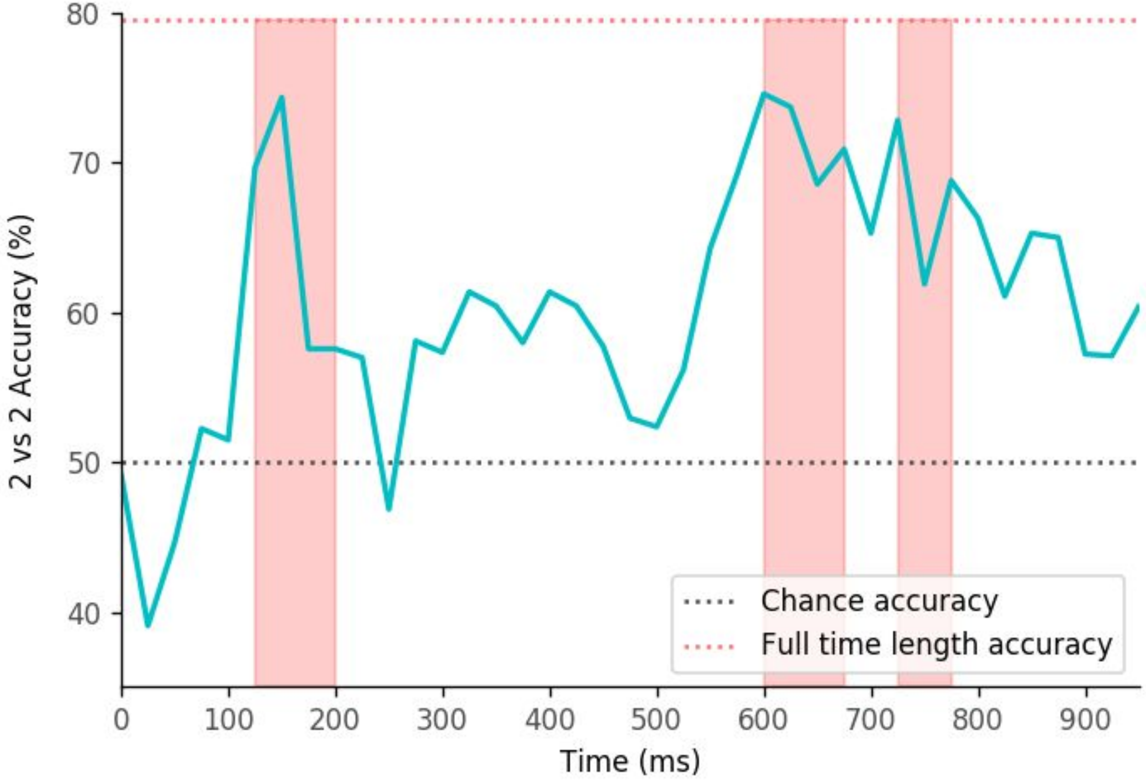
- **79.54%** over 0 - 700ms
- **69.15%** over 0 - 500ms

This provides evidence that EEG activity is correlated with the representation of word semantics in the brain

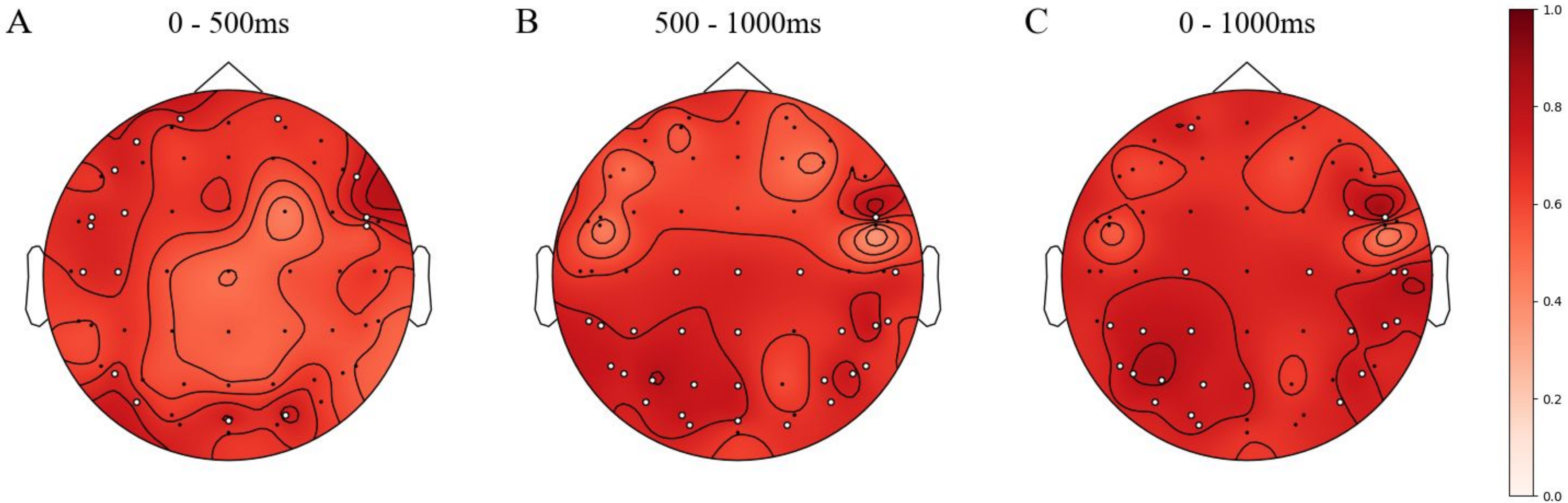
How much participant practice is needed?



How long does the brain take to process semantics?



Are there areas of the brain that contribute more?



Conclusion

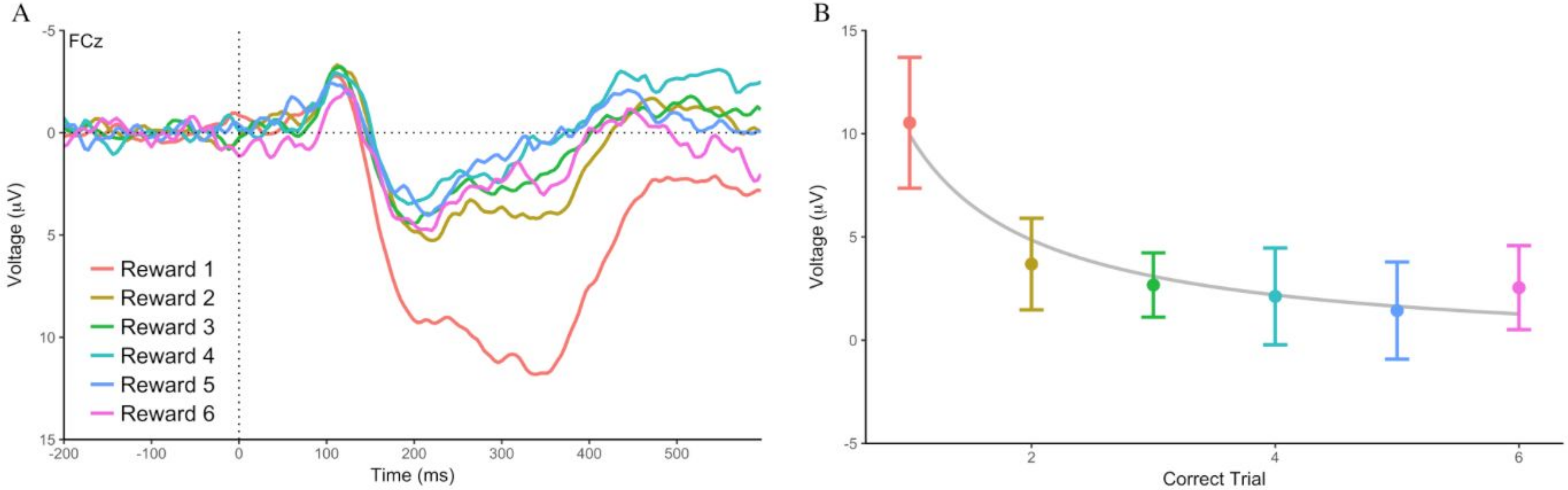
We show that semantics can be detected via EEG, and further that we can detect learning of semantic concepts as they develop a language mapping in the brain

This opens new avenues for studying language semantics and learning via EEG

Thanks!

CIFAR University of Victoria NSERC CRSNG

How does this compare to traditional methods?



Is there a relationship to the participant’s choices?

Our machine learning model shows an accuracies of:

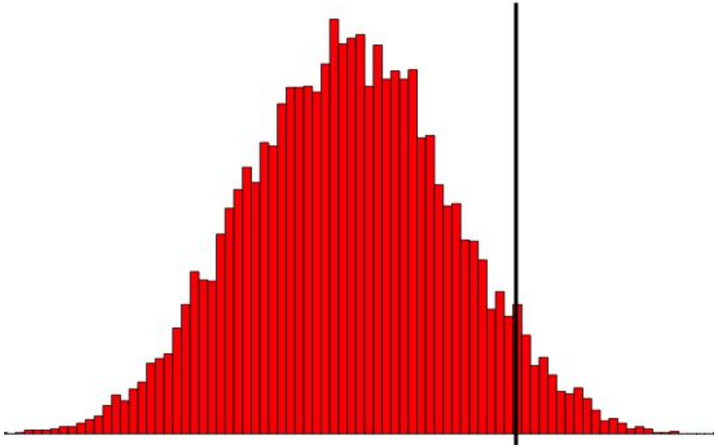
- **65.13% vs 59.71%** over 0 - 700ms
- **57.55% vs 57.47%** over 0 - 500ms

This effect might not be very powerful due to the tasks repetition of blocks with lower participant accuracy

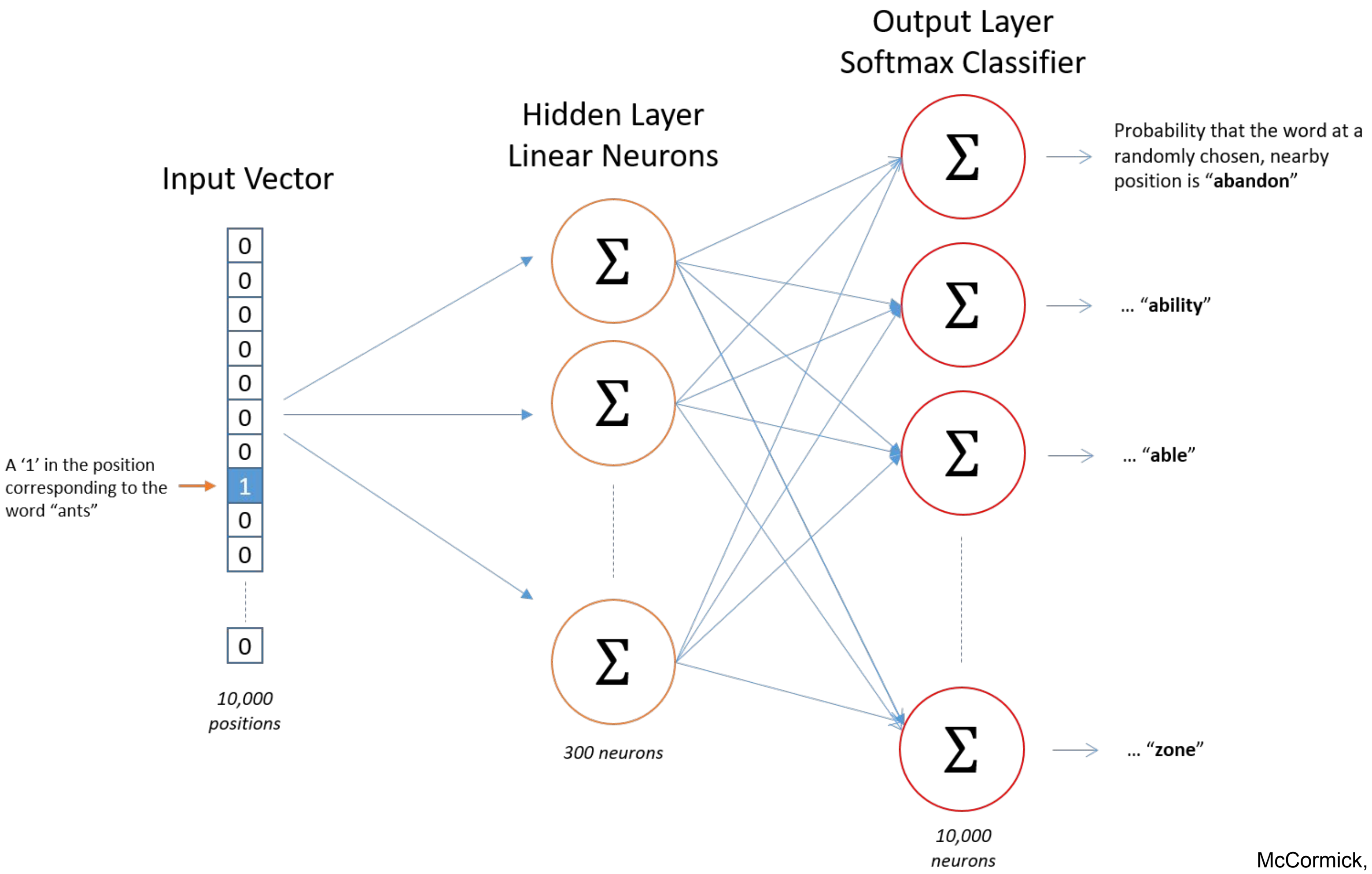
Measuring significance

Significant is measured with a permutation test

1. Randomly permute the word vectors so that they no longer pair to the correct EEG
2. Run the entire model on the permuted data
3. Repeat 1,000 times



This creates a null distribution we can use


Output Layer
Softmax Classifier
Hidden Layer
Linear Neurons
Input Vector
Probability that the word at a randomly chosen, nearby position is "abandon"
... "ability"
... "able"
... "zone"
A '1' in the position corresponding to the word "ants"
0
0
0
0
0
0
0
1
0
0
0
10,000 positions
300 neurons
10,000 neurons

Hidden Layer Weight Matrix → *Word Vector Lookup Table!*