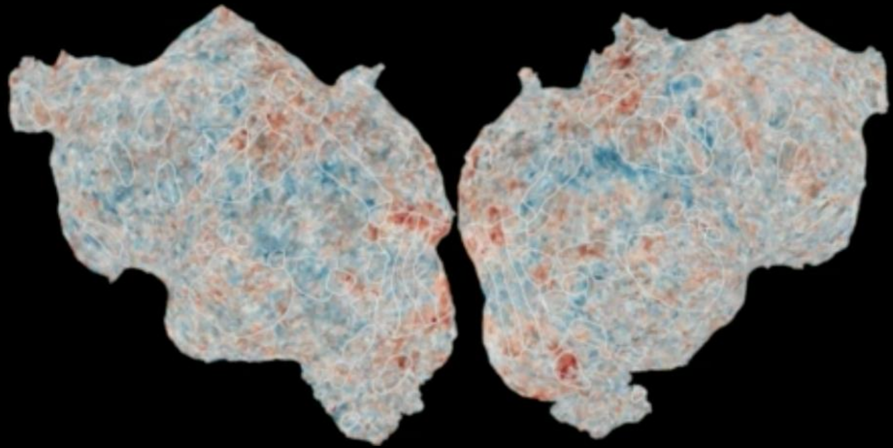
Deep multi-view representation learning of brain responses to natural stimuli



Leila Wehbe*, Anwar Nunez-Elizalde*,
Alex Huth, Fatma Imamoglu, Natalia Bilenko,
Jack Gallant

UC Berkeley

How does the human brain represent information?

- Machine Learning

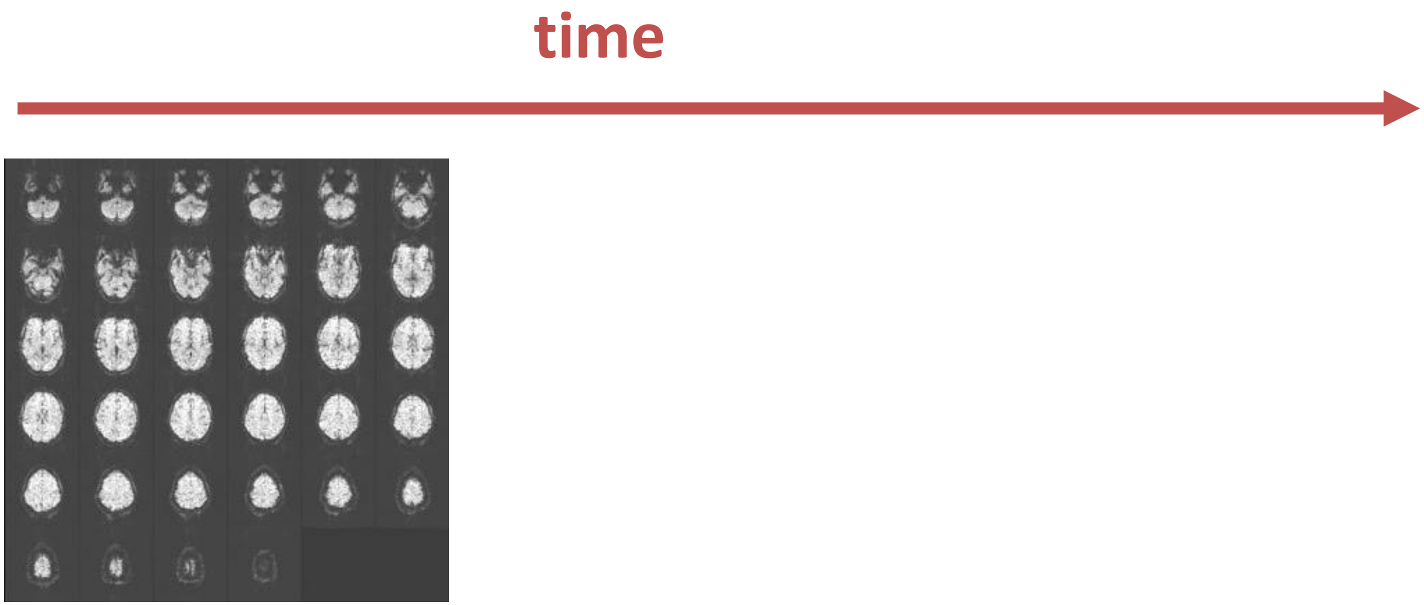
- Non-invasive imaging, e.g. fMRI

An fMRI experiment

- generate time series of functional volumes

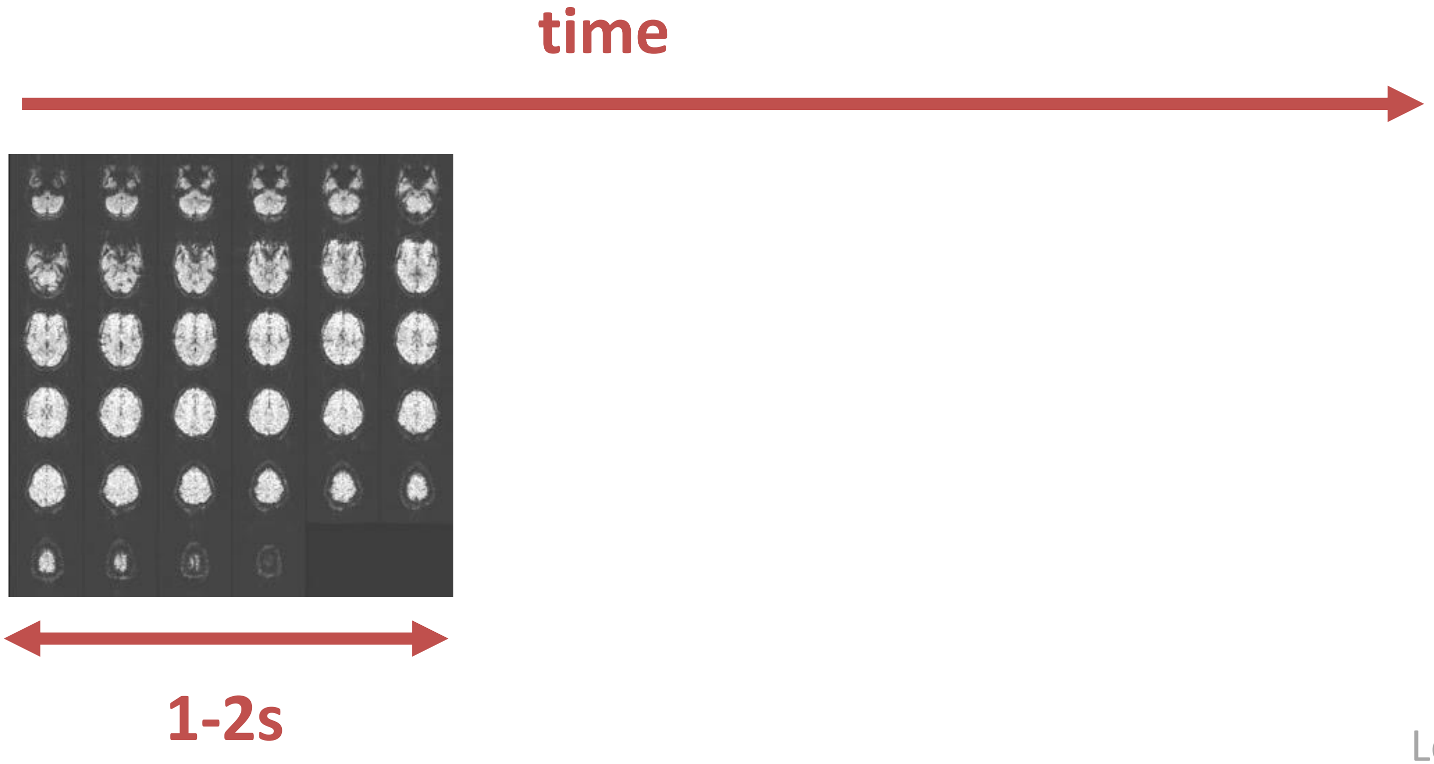
An fMRI experiment

- generate time series of functional volumes



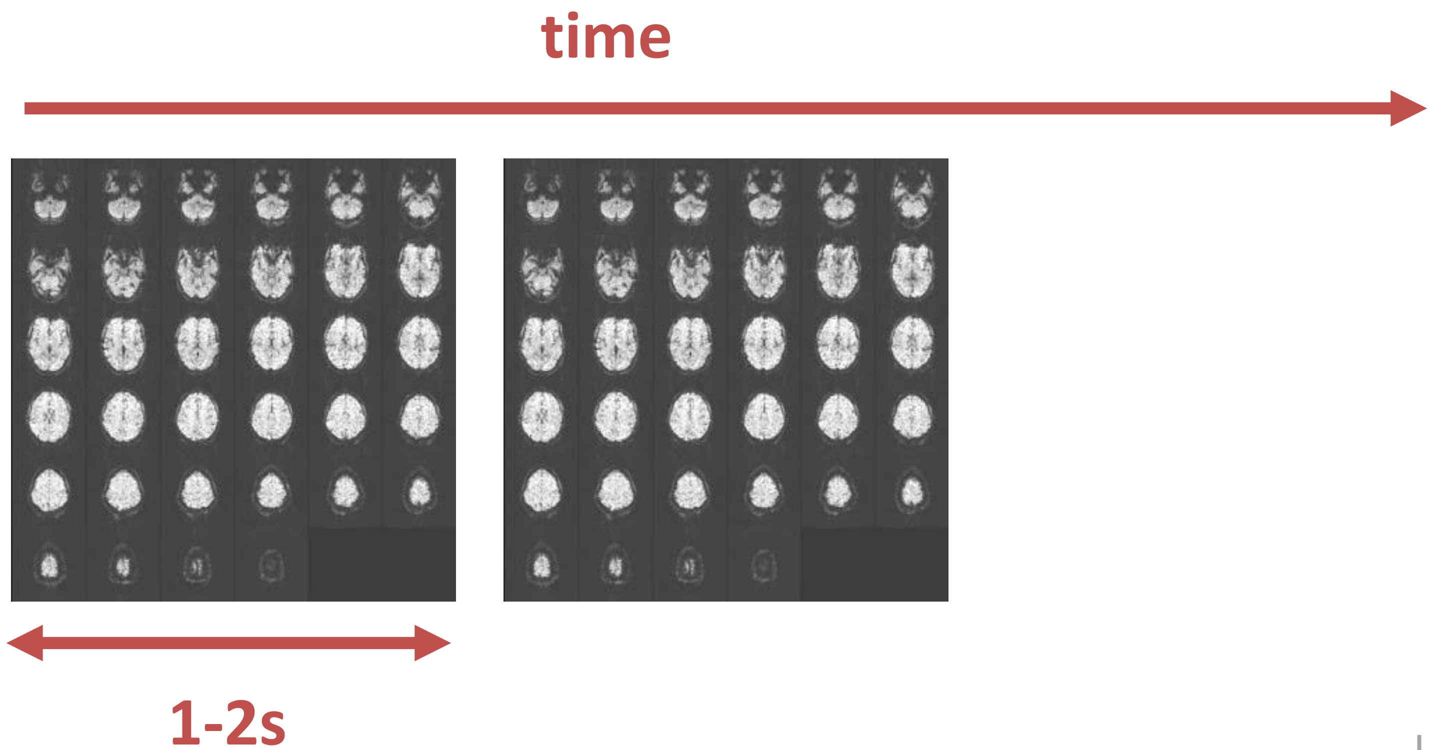
An fMRI experiment

- generate time series of functional volumes



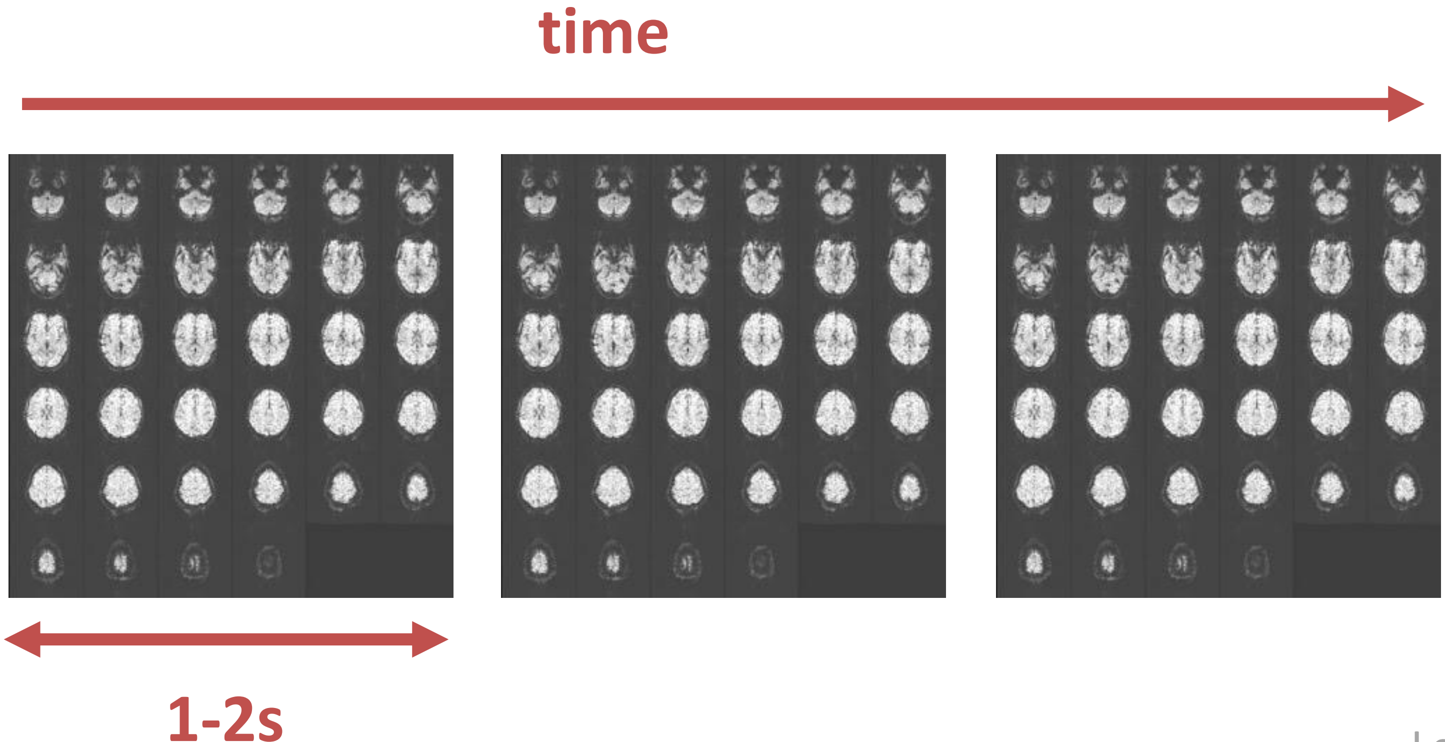
An fMRI experiment

- generate time series of functional volumes



An fMRI experiment

- generate time series of functional volumes



An fMRI experiment

An fMRI experiment

- Typically:
  - Isolate a specific
  - Two or a few conditions
  - Find regions that differ in activity

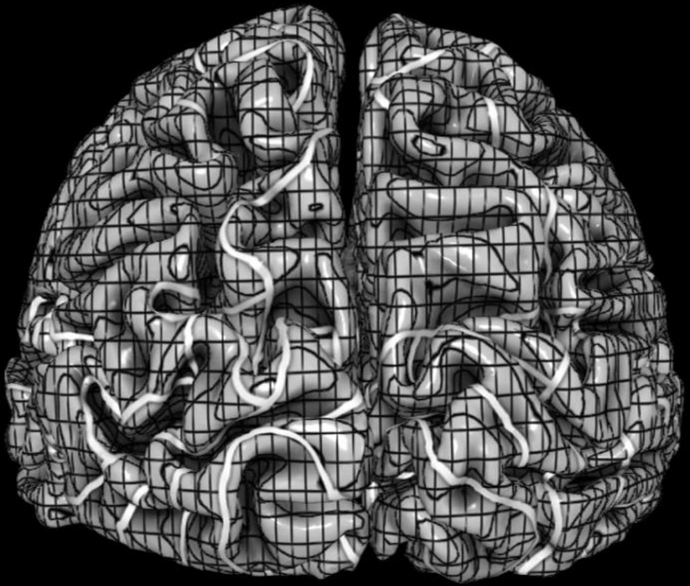
An fMRI experiment

- Typically:
  - Isolate a specific
  - Two or a few conditions
  - Find regions that differ in activity

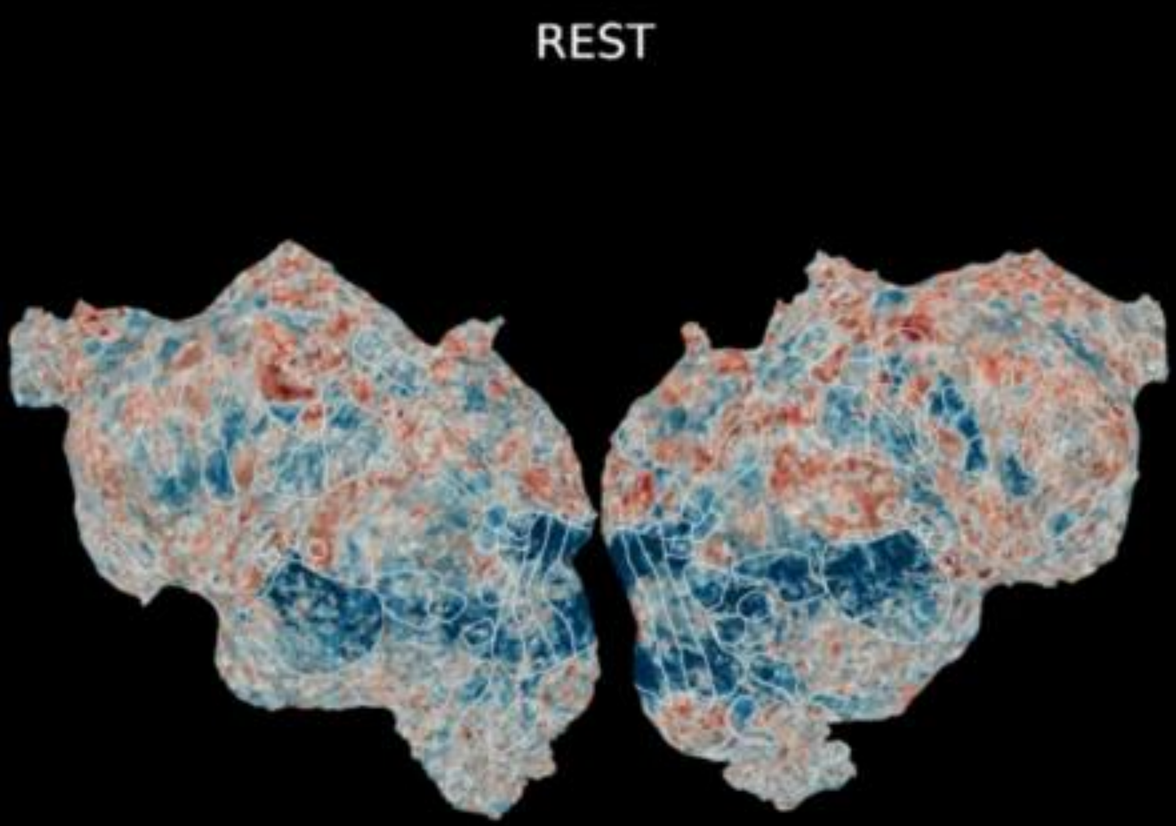
- Problems:
  - Hard to generalize
  - Infinite number of binary comparisons

Naturalistic experiments

- Make subjects do a real life complex task:
  - Watch movies
  - Listen / read stories



Video by James Gao and Anwar Nunez-Elizalde



Video by James Gao and Anwar Nunez-Elizalde

Naturalistic Experiments

Naturalistic Experiments

- No clear classes

Naturalistic Experiments

- No clear classes
  - Classification techniques are not useful/interesting here

Naturalistic Experiments

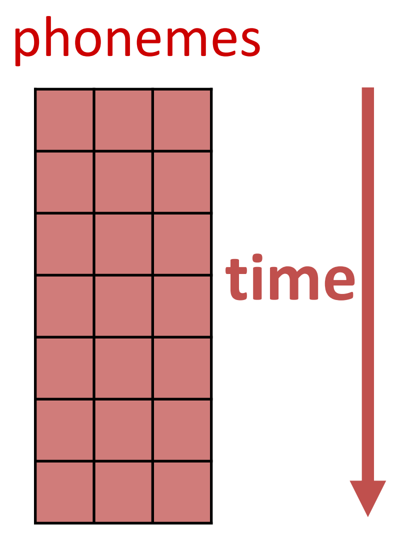
- No clear classes
  - Classification techniques are not useful/interesting here
- Highly complex input varying along multiple levels

Naturalistic Experiments

- No clear classes
  - Classification techniques are not useful/interesting here
- Highly complex input varying along multiple levels
  - Model it!

Build feature spaces!

- **Stories** have acoustic and semantic properties:
  - phonemes: Count of the occurrence of 39 phonemes



Leila Wehbe

Build feature spaces!

- **Stories** have acoustic and semantic properties:
  - word2vec: Bag of words model of the words occurring at each 2s

Build feature spaces!

- **Movies** have visual and semantic properties:
  - word2vec: Bag of words model of the objects occurring at each 2s

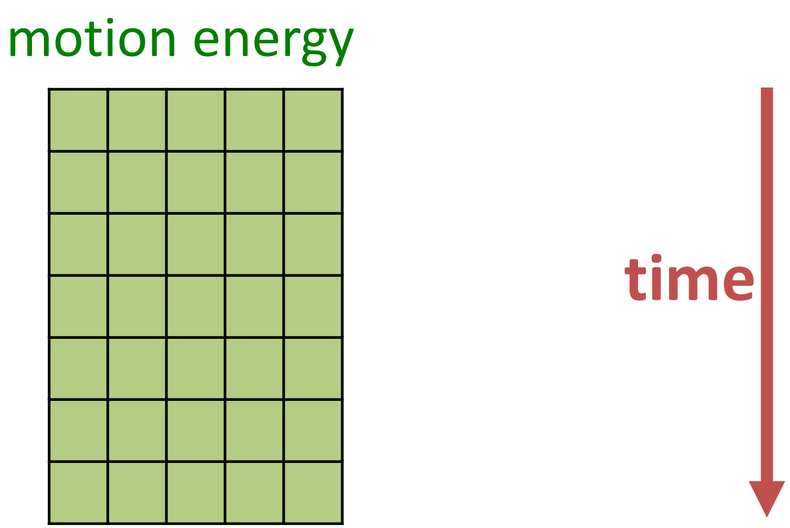
Build feature spaces!

- Movies have visual and semantic properties:
  - motion energy filters: spatio-temporal Gabor pyramids

Video by Mark Lescroart

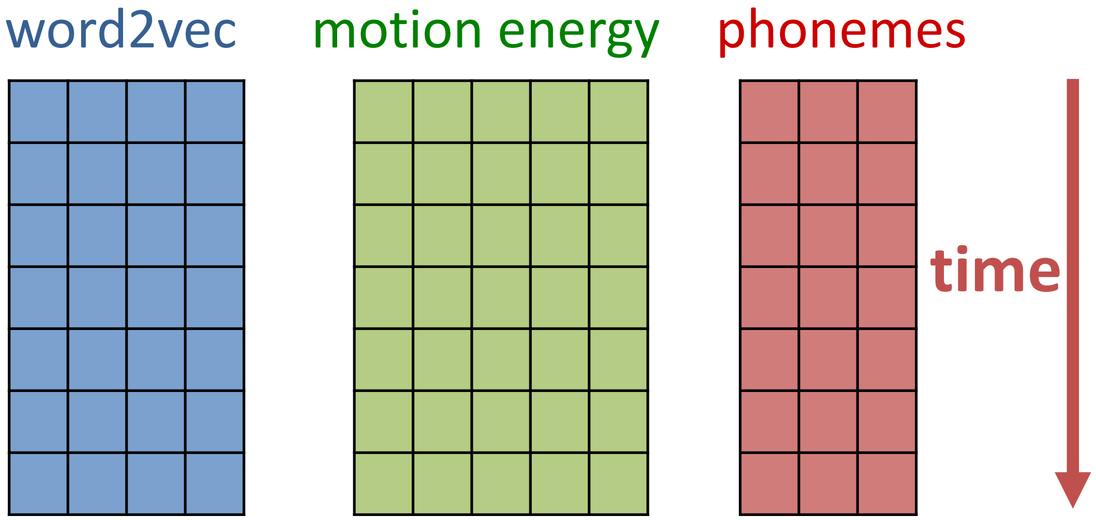
Build feature spaces!

- Movies have visual and semantic properties:
  - motion energy: spatio-temporal Gabor pyramids



Leila Wehbe

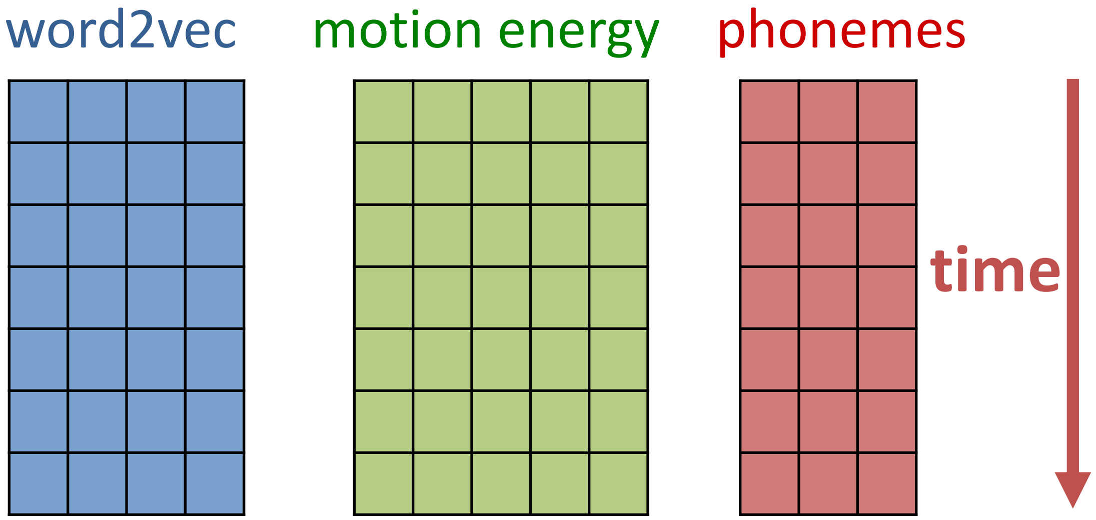
Build feature spaces!



Leila Wehbe

Build feature spaces!

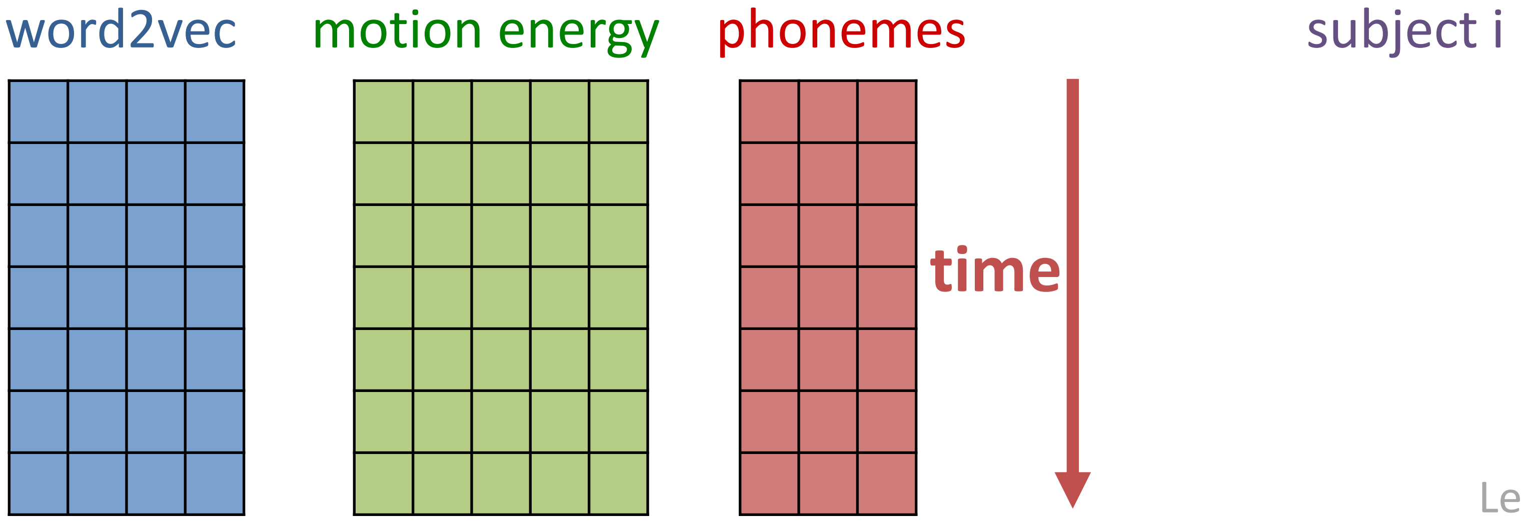
- Data from 4 subjects for both experiments
  - Story Listening (Huth et al. 2016)
  - Natural Movies (Nishimoto et al. 2011)
- Each subject has 30-50 thousand voxels
  - And ~ 3600 time points per experiment



Leila Wehbe

Build feature spaces!

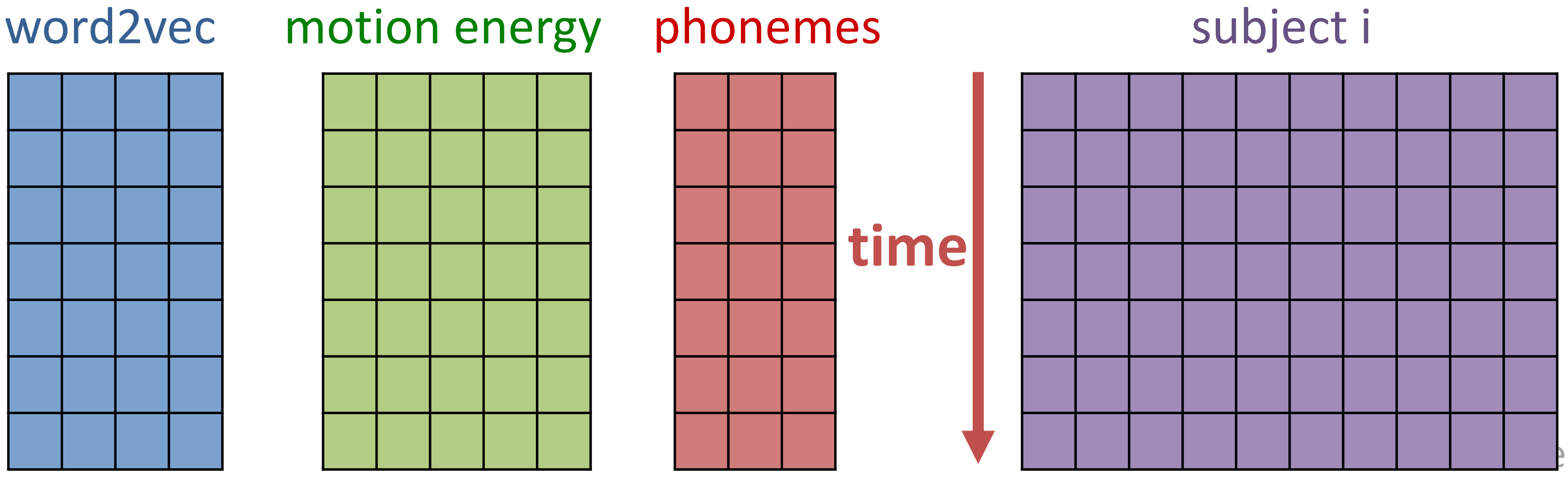
- Data from 4 subjects for both experiments
  - Story Listening (Huth et al. 2016)
  - Natural Movies (Nishimoto et al. 2011)
- Each subject has 30-50 thousand voxels
  - And ~ 3600 time points per experiment



Leila Wehbe

Build feature spaces!

- Data from 4 subjects for both experiments
  - Story Listening (Huth et al. 2016)
  - Natural Movies (Nishimoto et al. 2011)
- Each subject has 30-50 thousand voxels
  - And ~ 3600 time points per experiment



Small data / complex brains

Small data / complex brains

- Small data scenario!
  - … and low SNR

Small data / complex brains

- Small data scenario!
  - … and low SNR
- We want to generalize:
  - Across subjects
  - Across feature spaces
  - And across experiments and cognitive domains

Useful Neuroimaging Tasks

Useful Neuroimaging Tasks

- Predict data from features

Useful Neuroimaging Tasks

- Predict data from features
- Predict features from data (decoding)

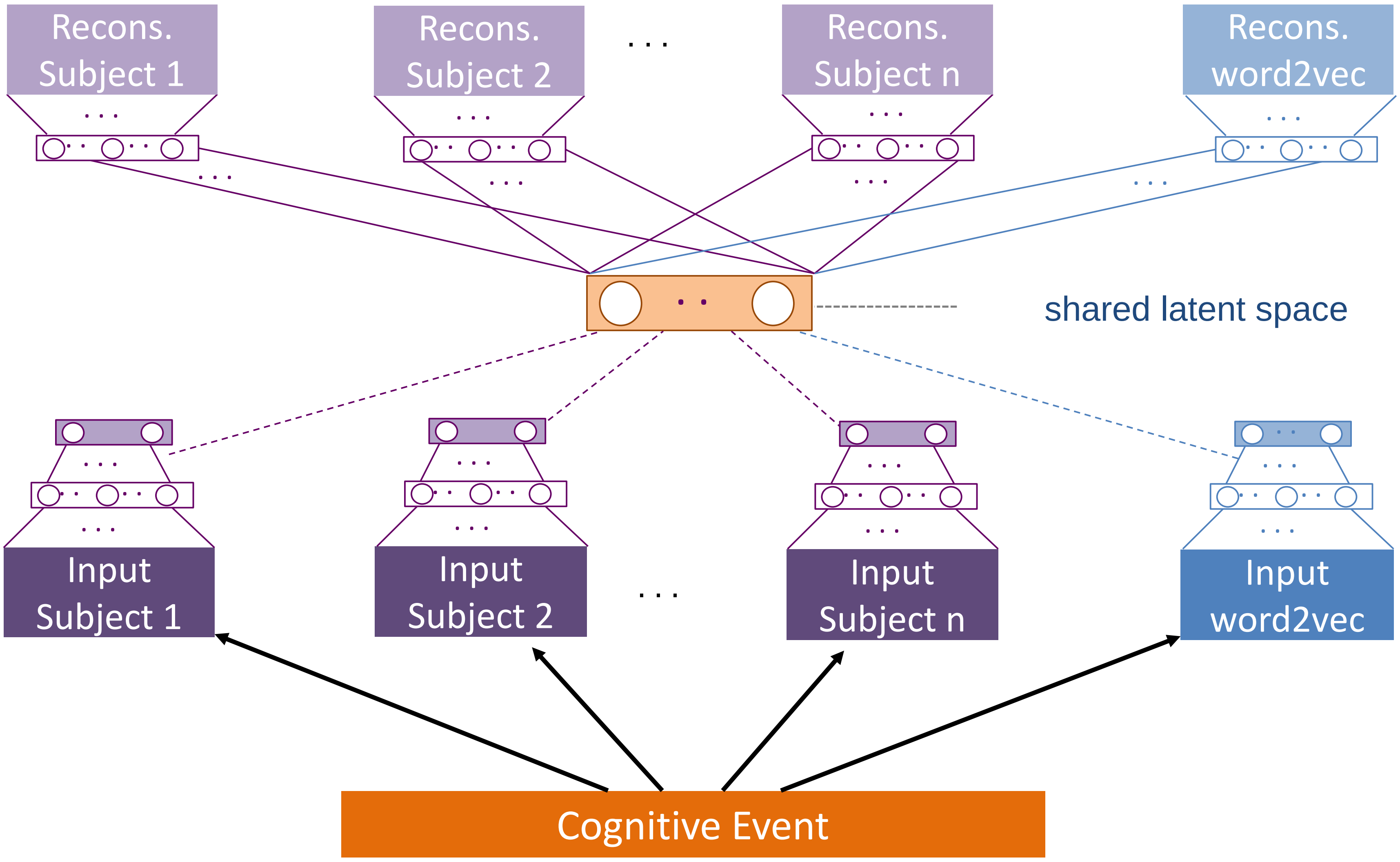
Useful Neuroimaging Tasks

- Predict data from features
- Predict features from data (decoding)
- Combine data across subjects

Useful Neuroimaging Tasks

- Predict data from features
- Predict features from data (decoding)
- Combine data across subjects
- Predict one subject from another

Cognitive Space Multi-view Autoencoder



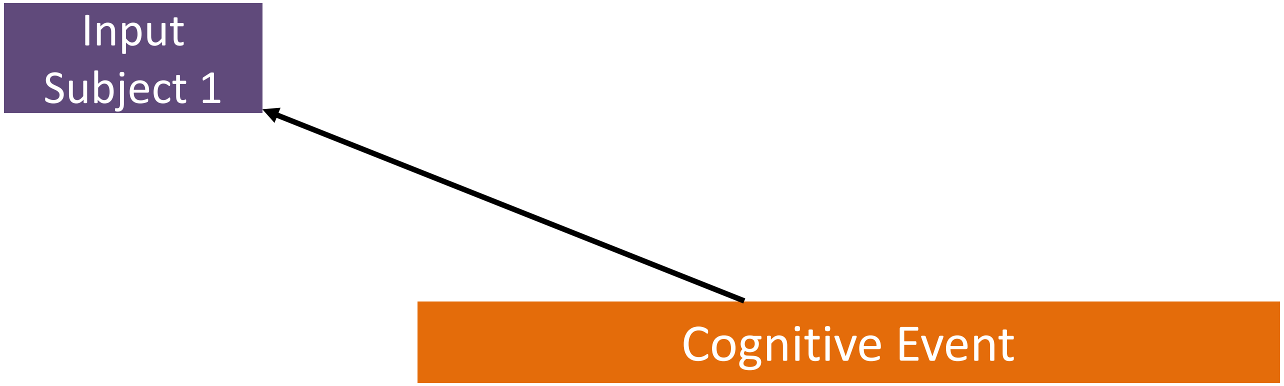
Cognitive Space Multi-view Autoencoder

Cognitive Event

E.g., perception of:

"… was running away as fast as he could …"

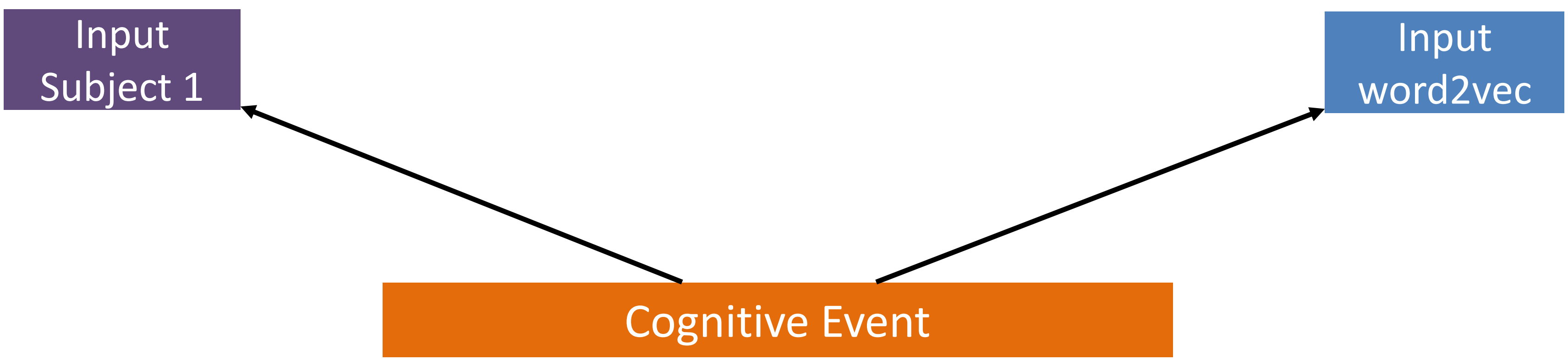
Cognitive Space Multi-view Autoencoder



E.g., perception of:

"… was running away as fast as he could …"

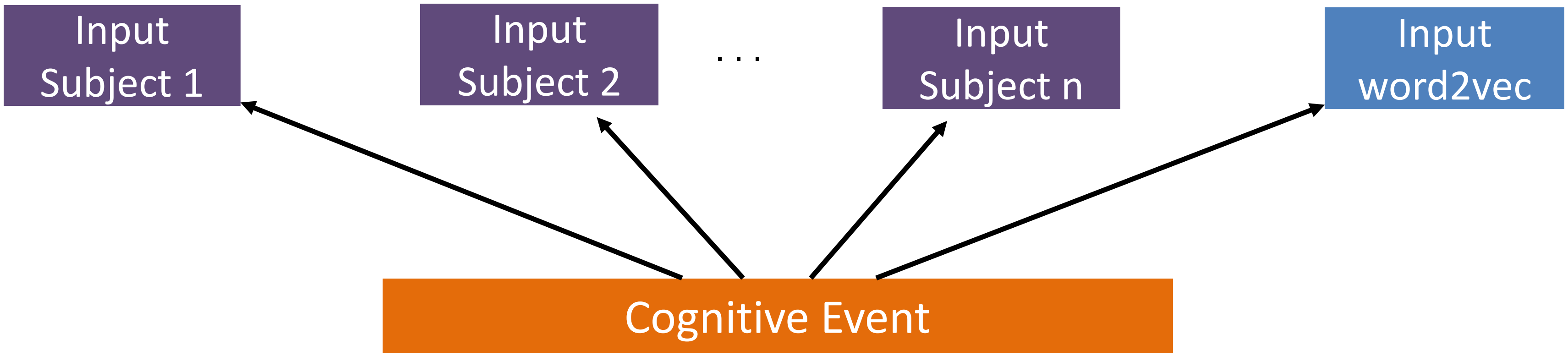
Cognitive Space Multi-view Autoencoder



E.g., perception of:

"… was running away as fast as he could …"

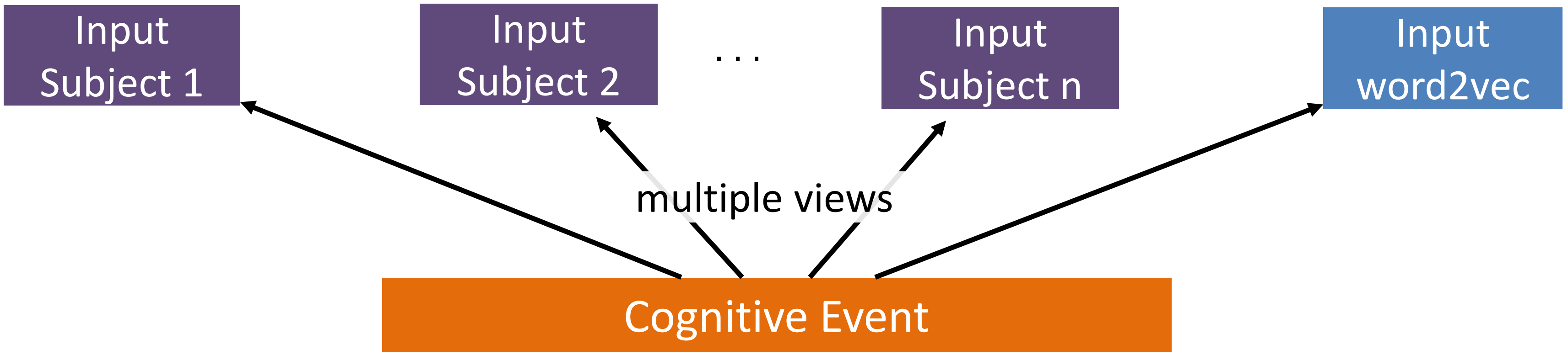
Cognitive Space Multi-view Autoencoder



E.g., perception of:

"… was running away as fast as he could …"

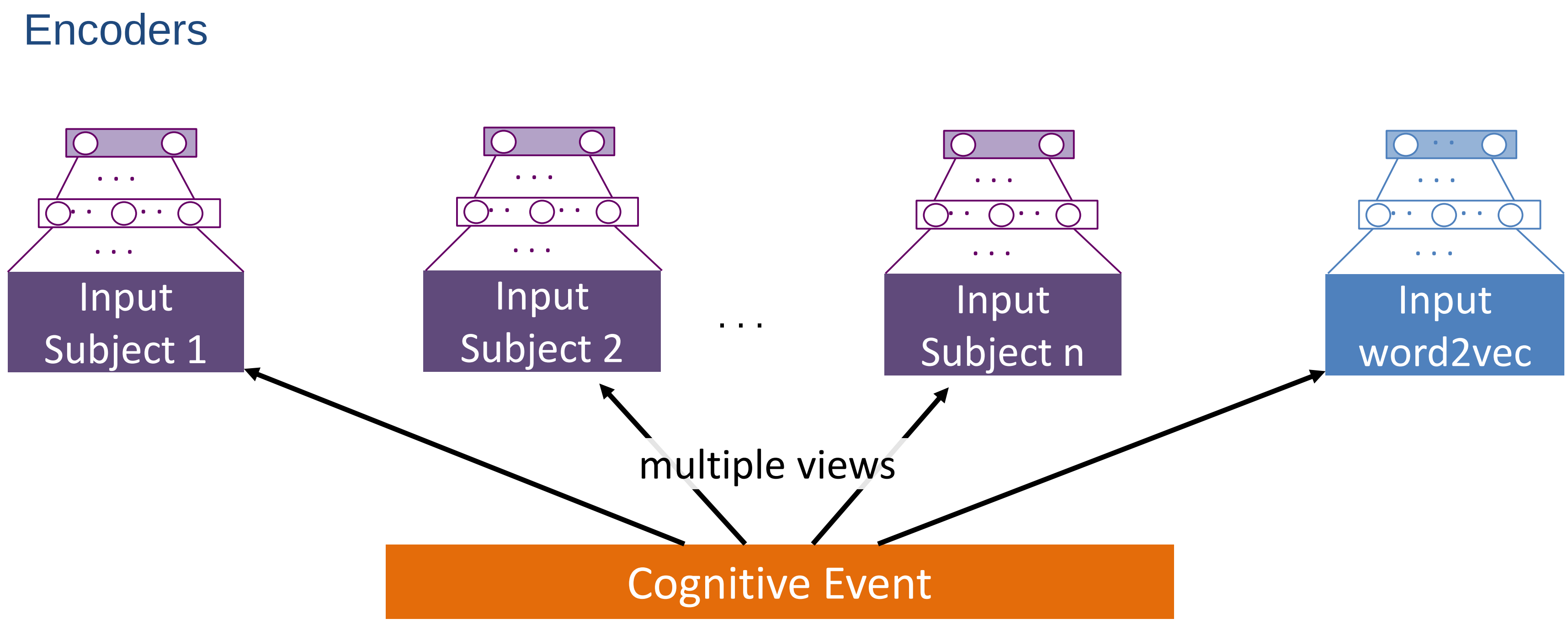
Cognitive Space Multi-view Autoencoder



E.g., perception of:

"… was running away as fast as he could …"

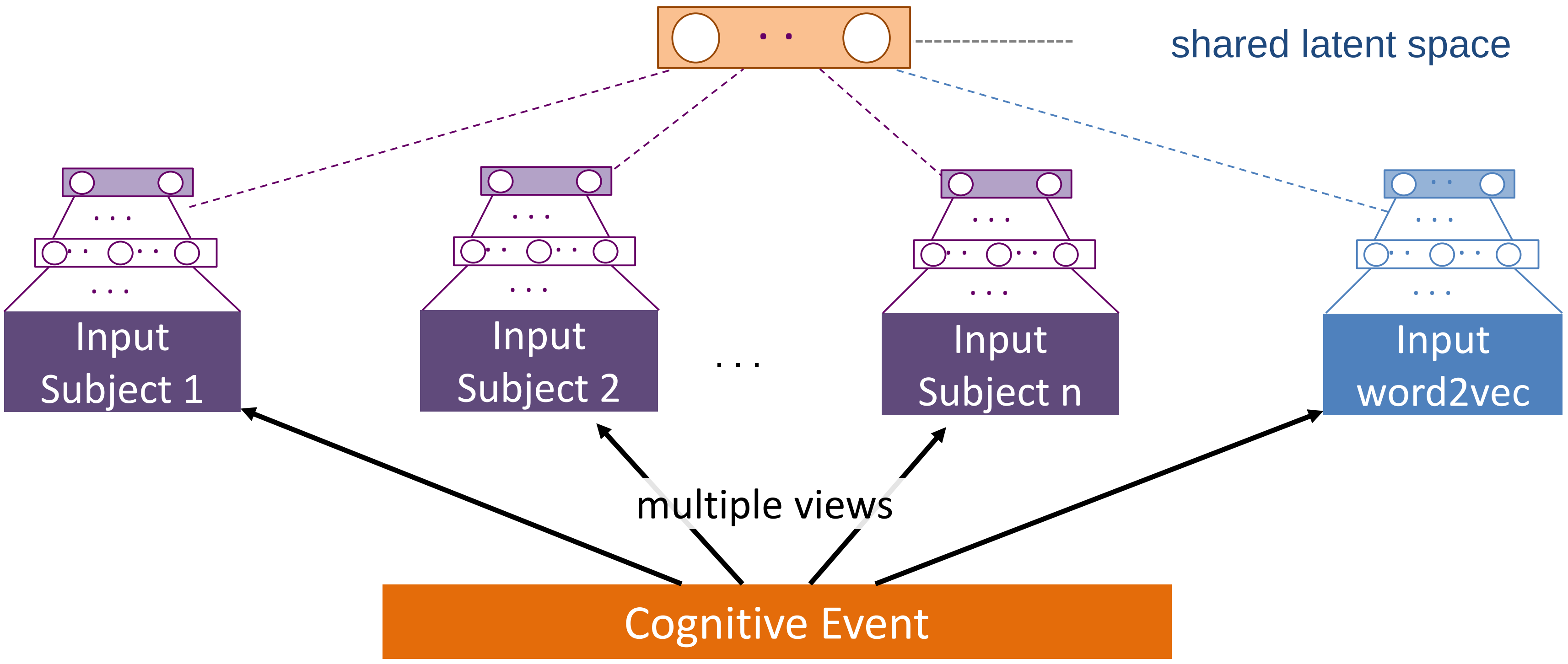
Cognitive Space Multi-view Autoencoder



E.g., perception of:

"… was running away as fast as he could …"

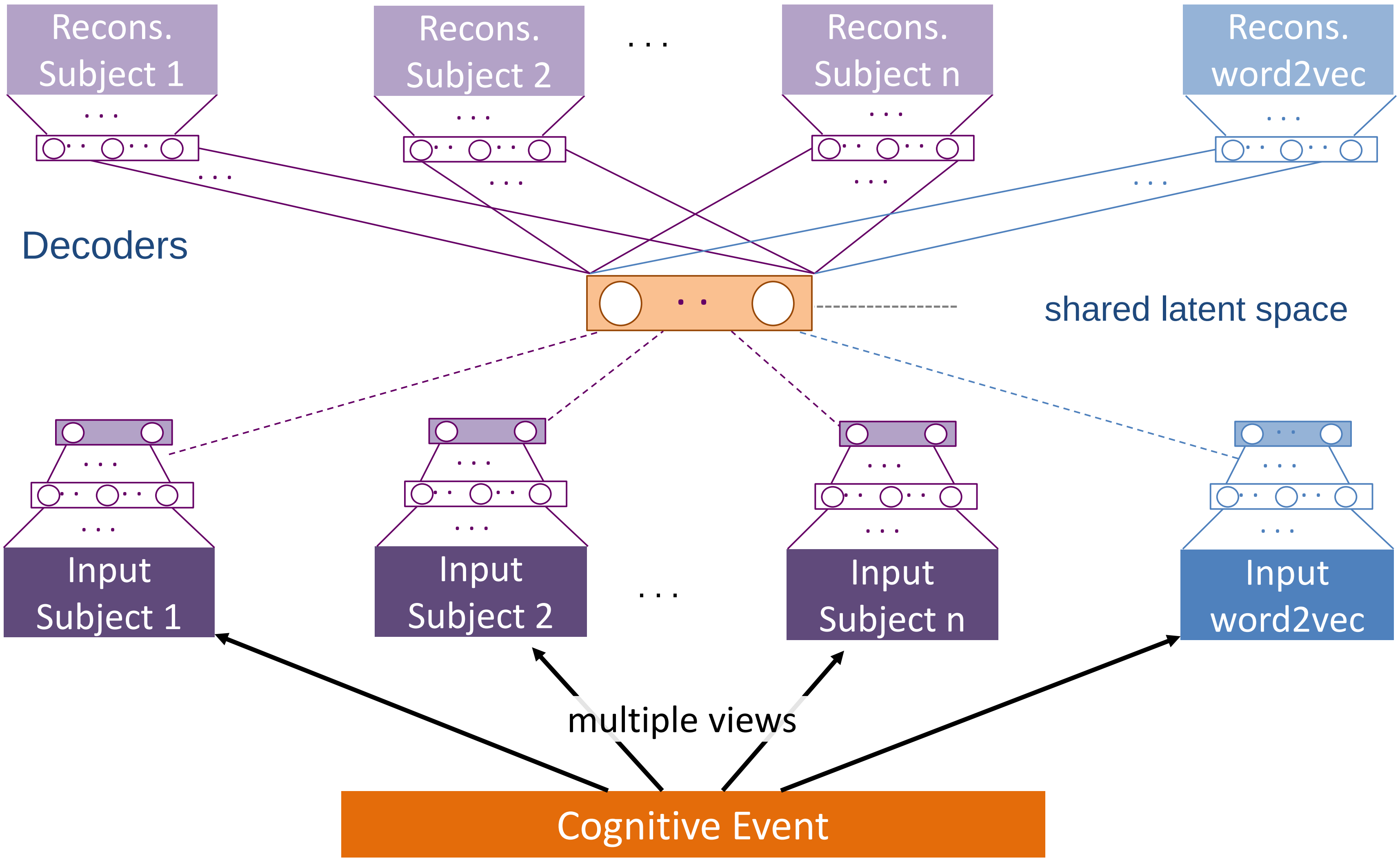
Cognitive Space Multi-view Autoencoder



E.g., perception of:

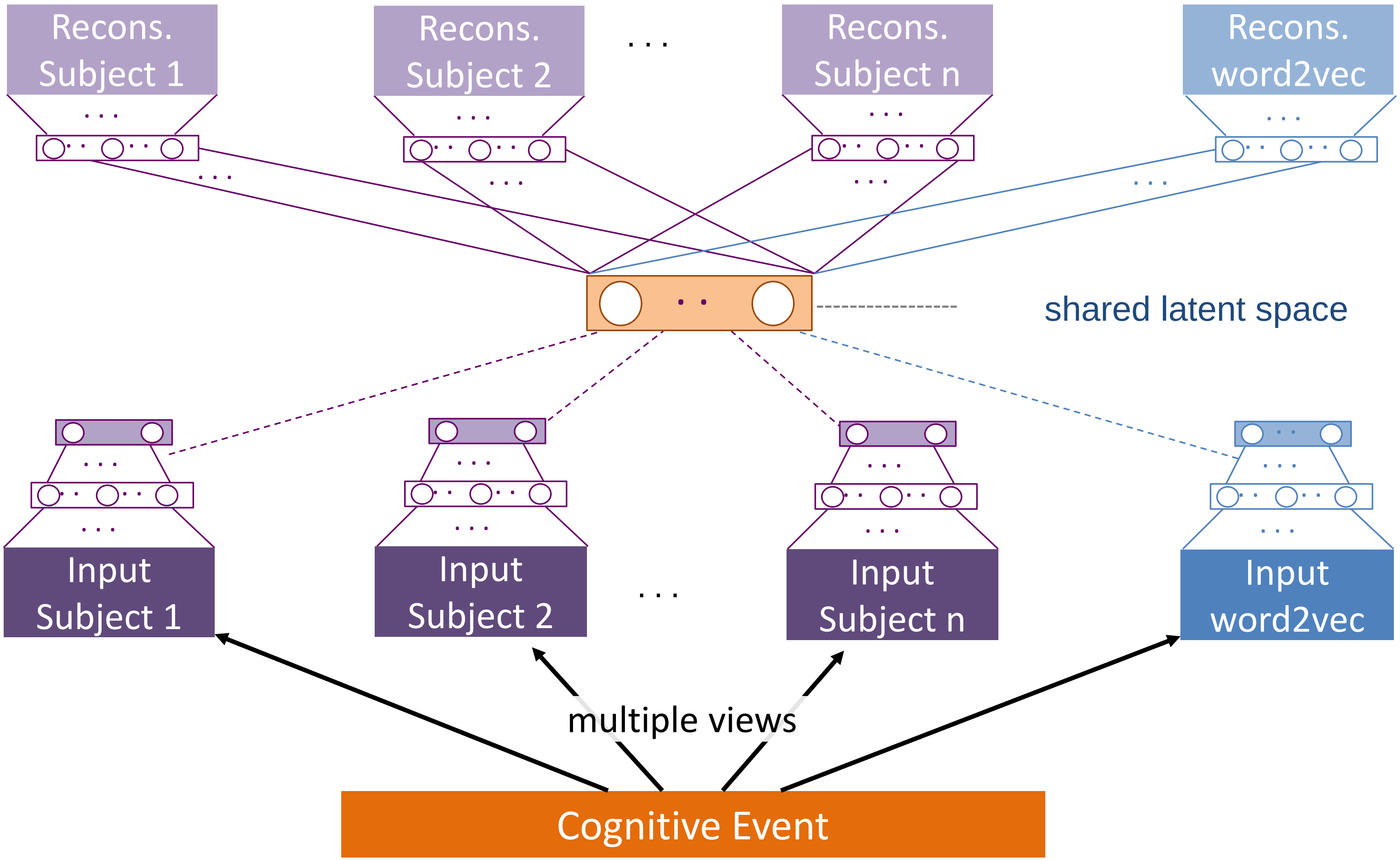
"… was running away as fast as he could …"

Cognitive Space Multi-view Autoencoder



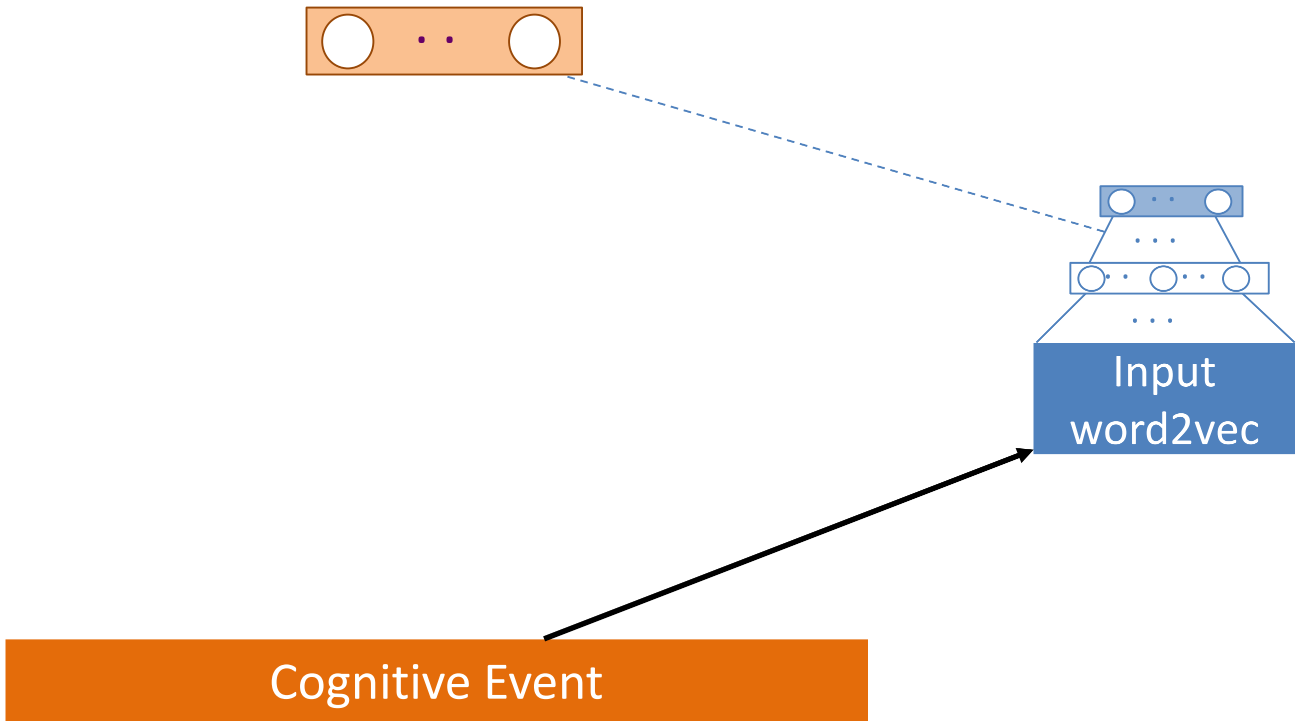
E.g., perception of:
"… was running away as fast as he could …"

Cognitive Space Multi-view Autoencoder

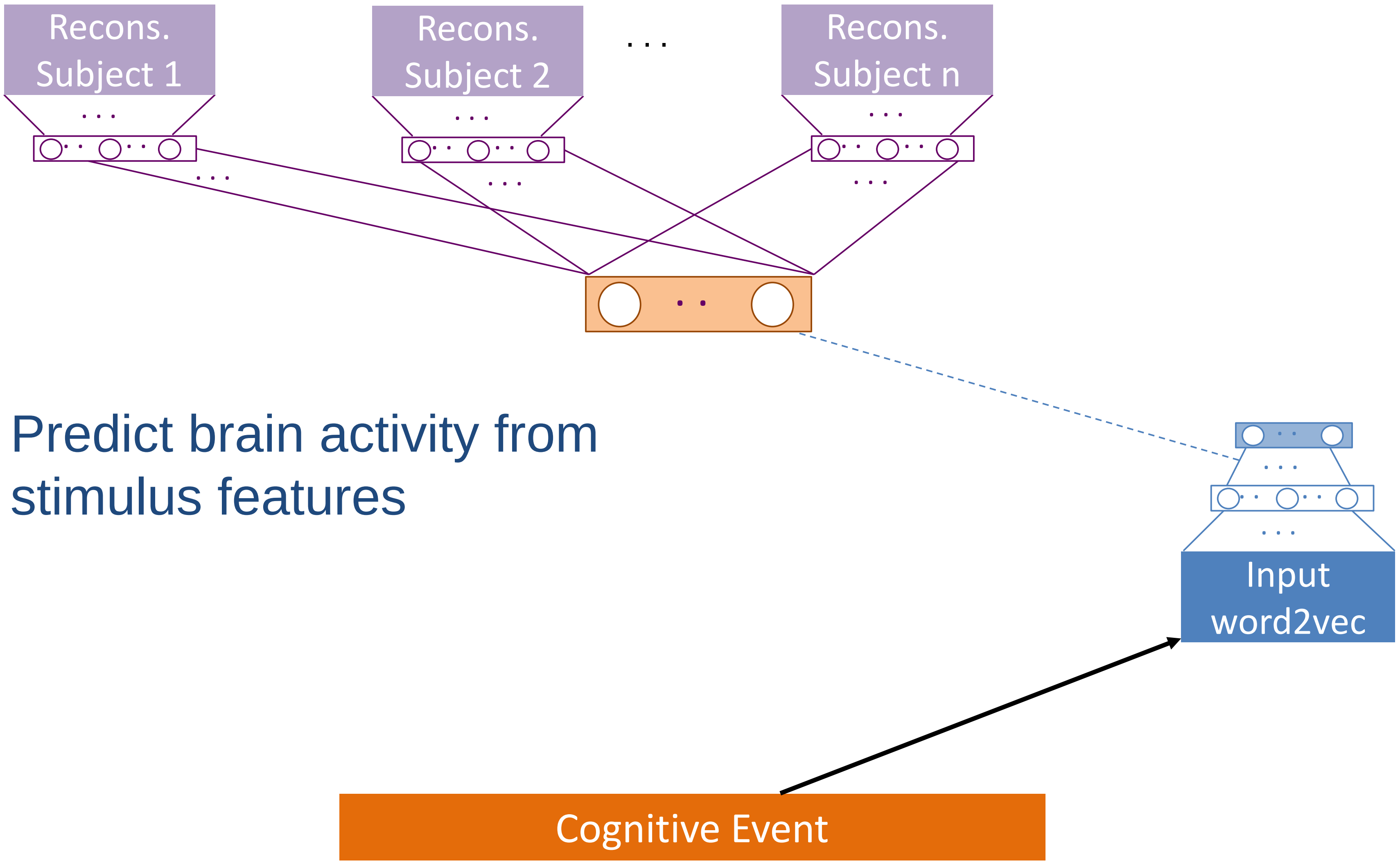


E.g., perception of:
"… was running away as fast as he could …"

Estimation

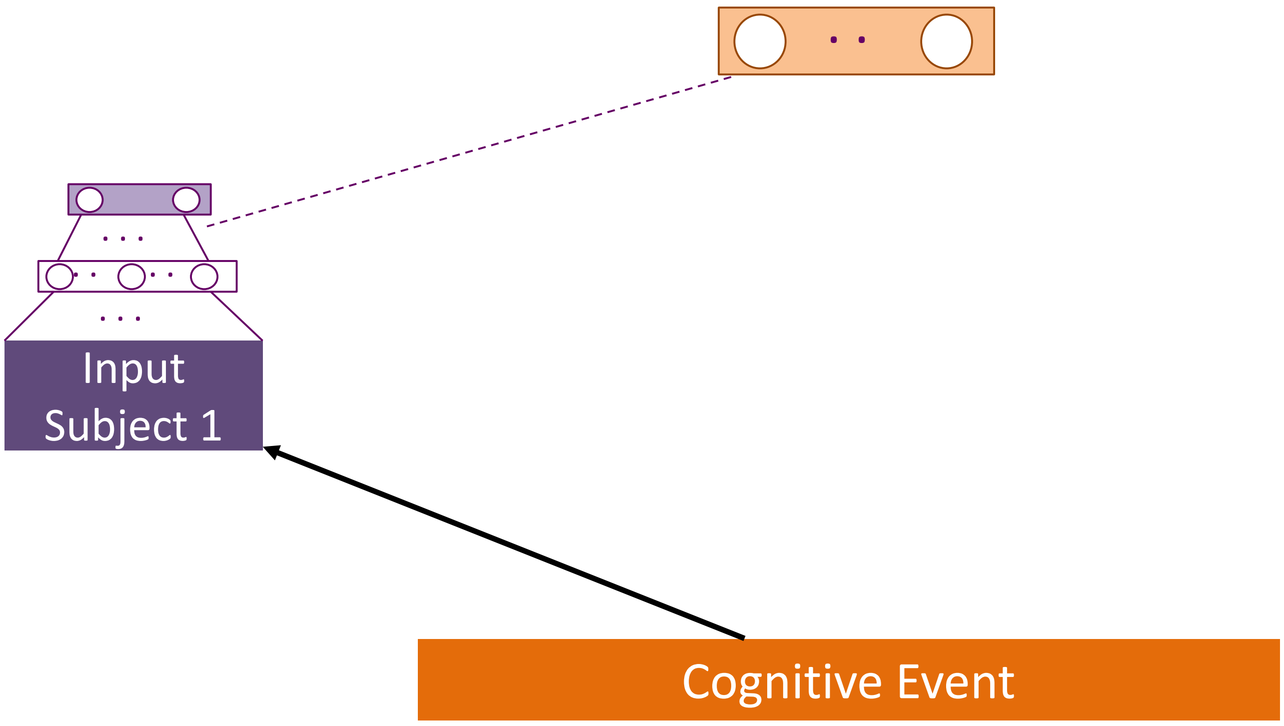


Estimation

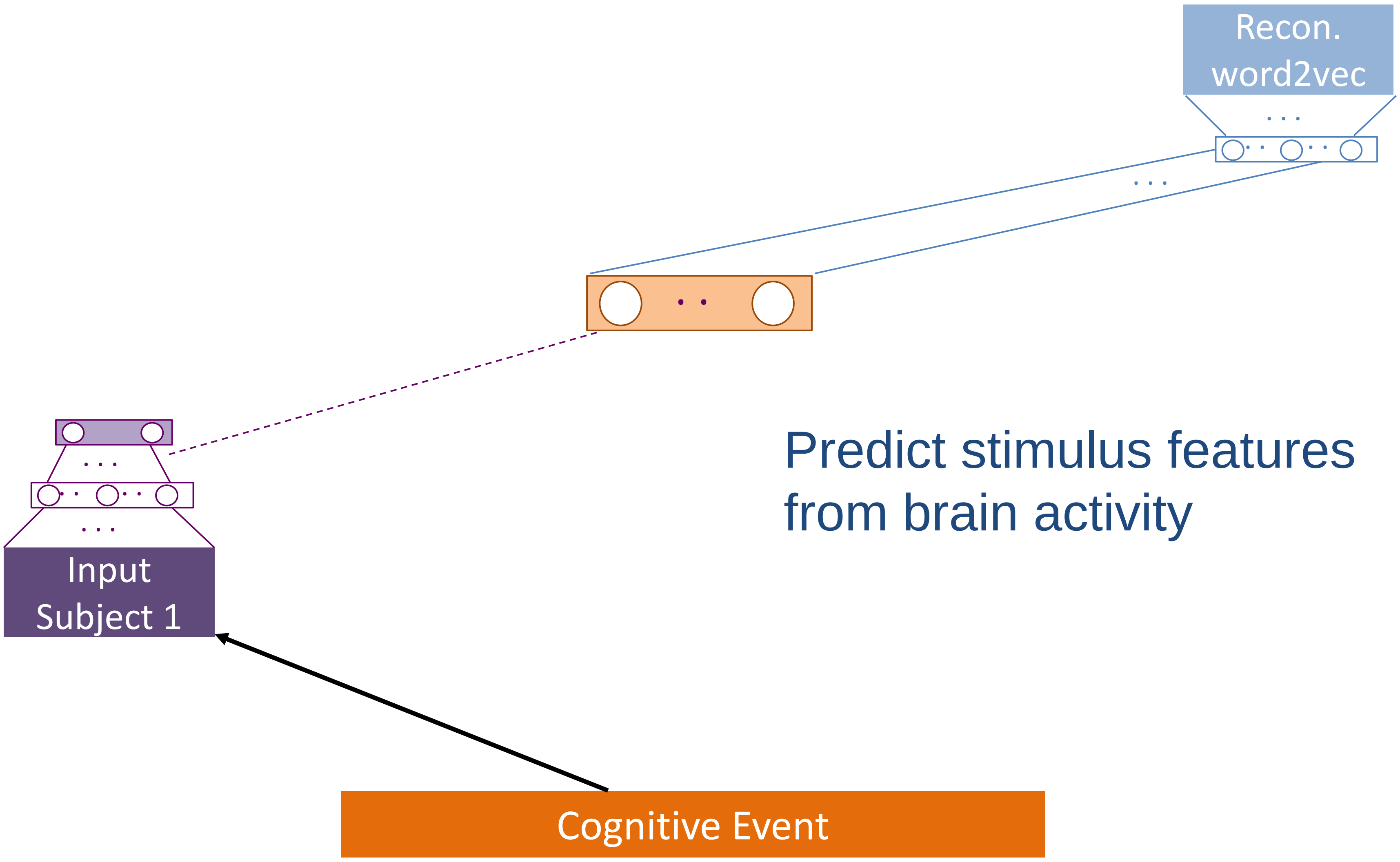


Predict brain activity from stimulus features

Estimation

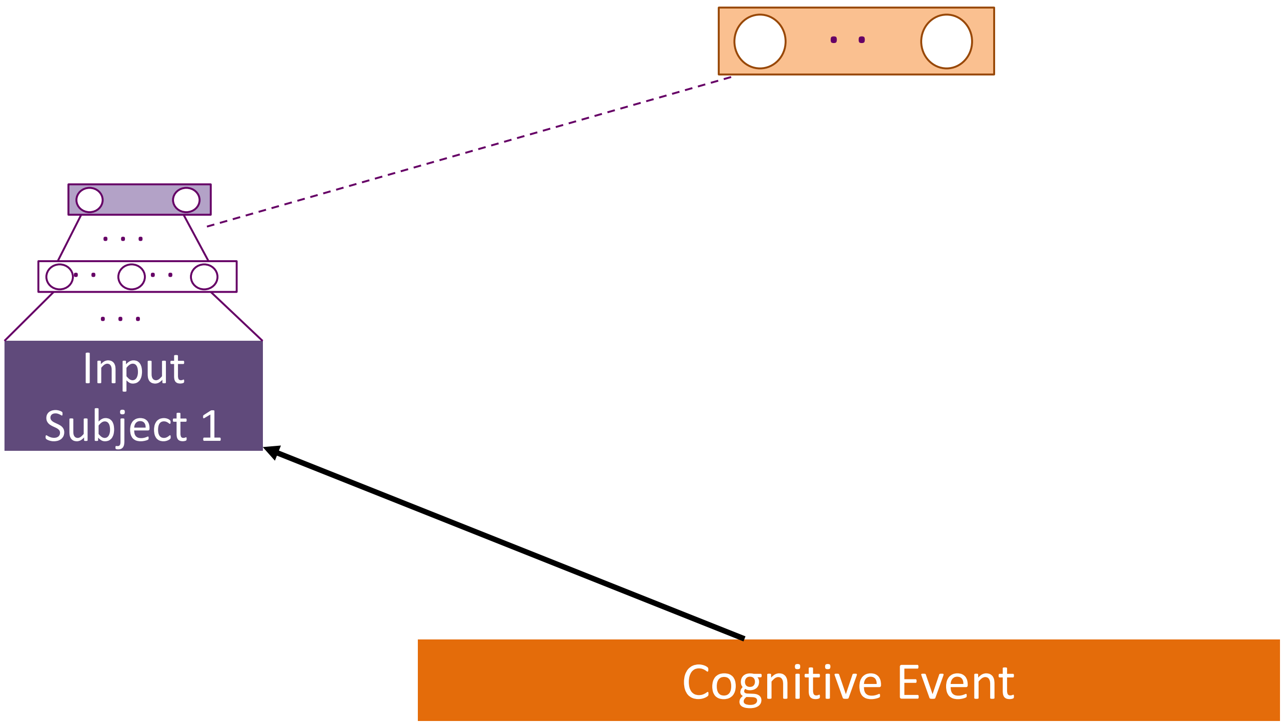


Estimation

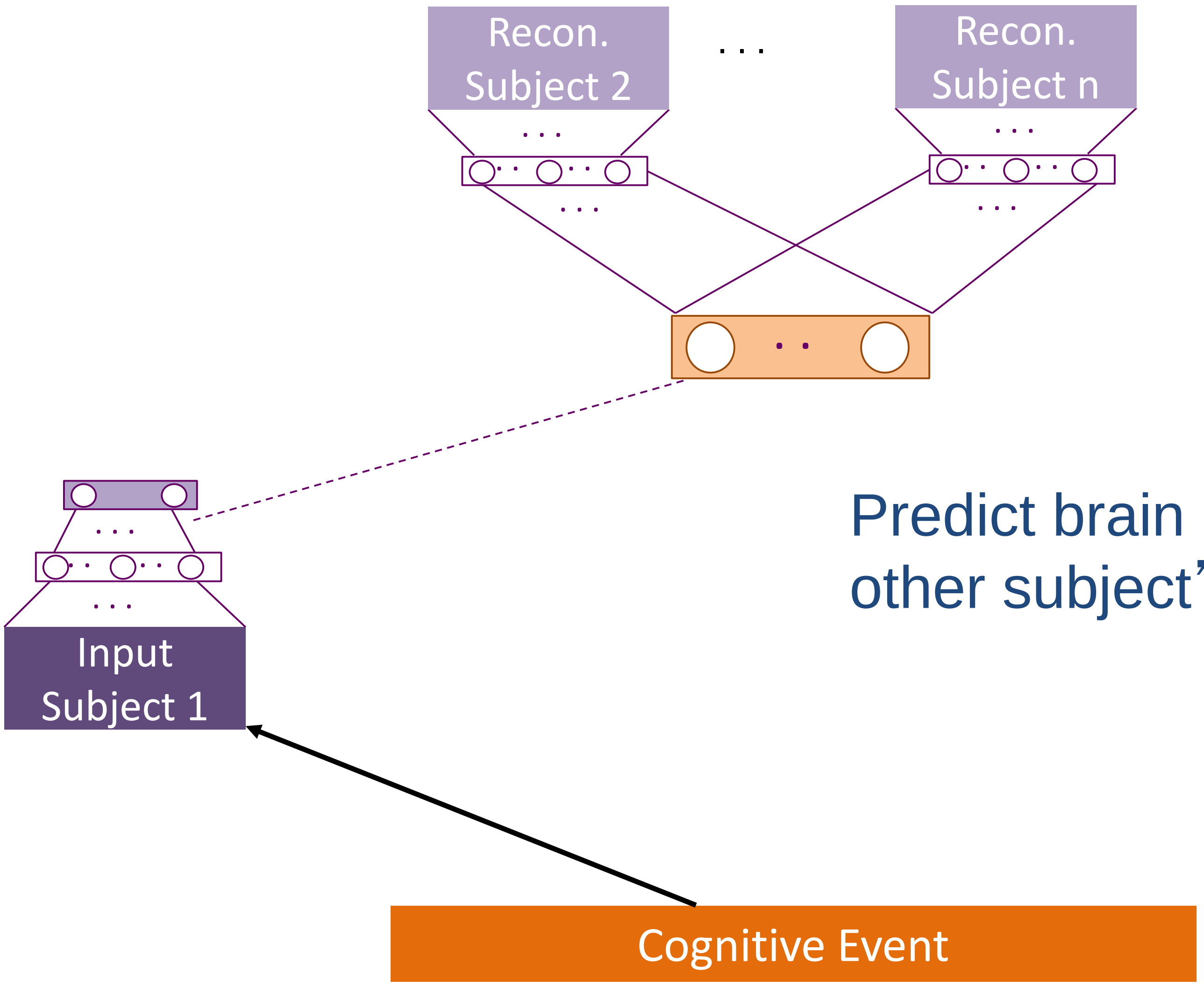


Predict stimulus features from brain activity

Estimation

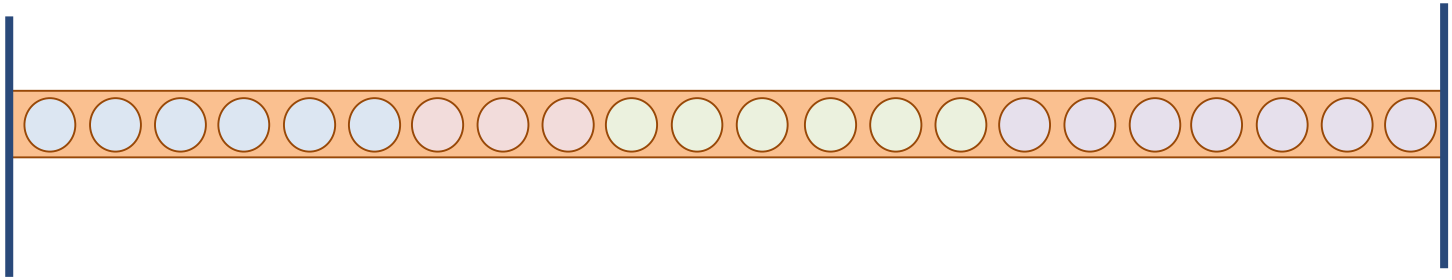


Estimation

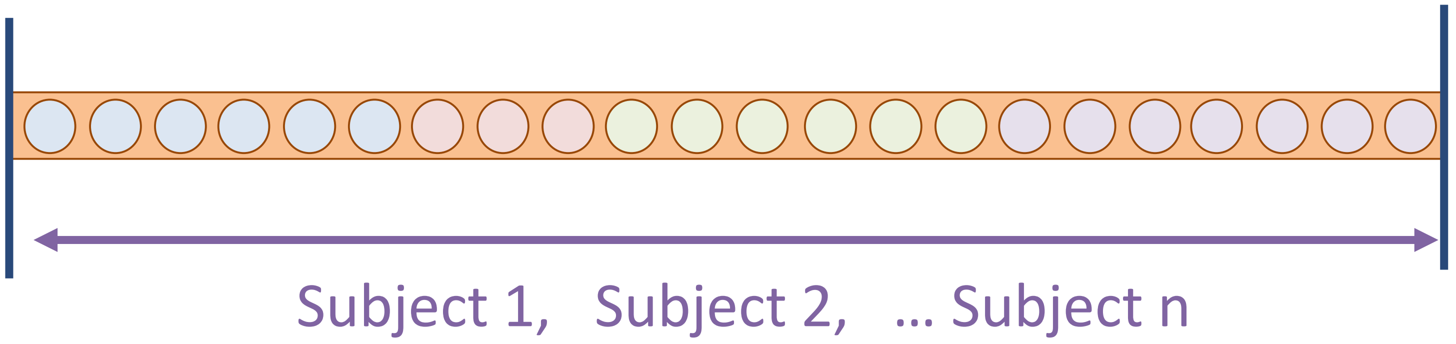


Predict brain activity from other subject's activity

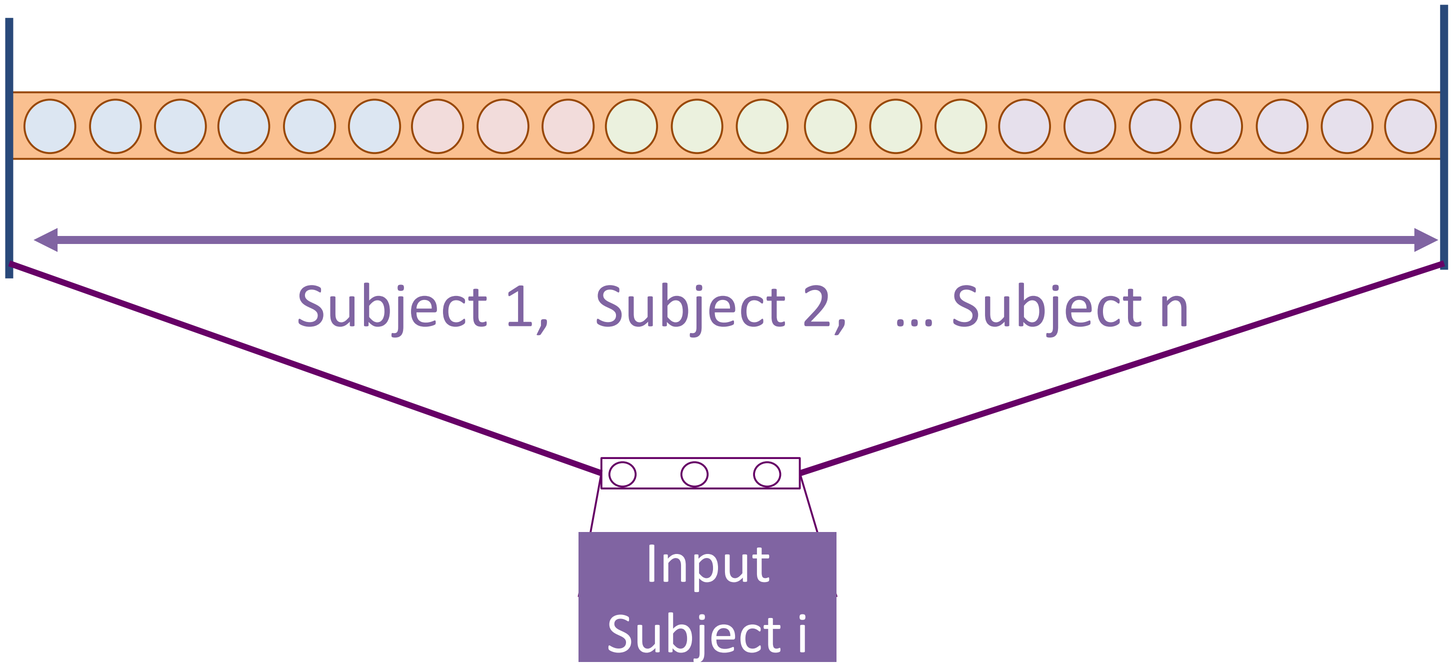
Bottleneck Layer



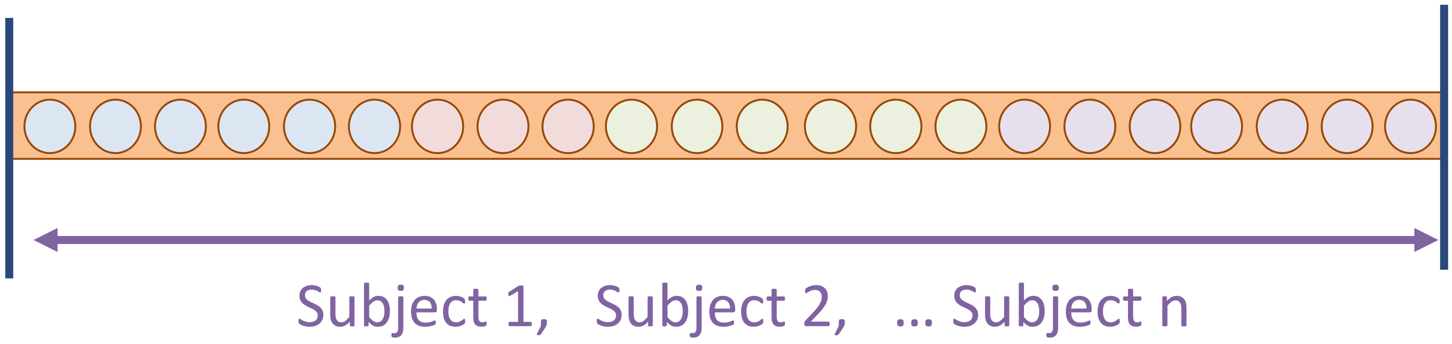
Bottleneck Layer



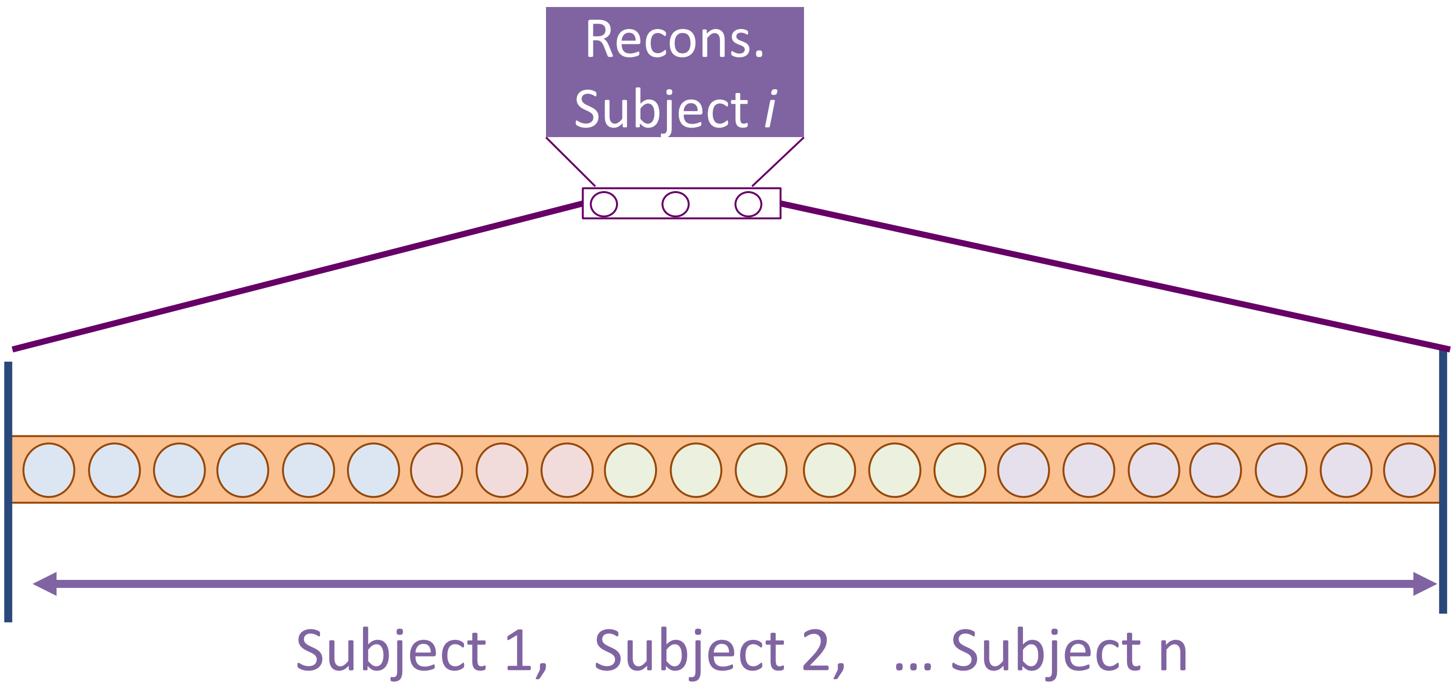
Bottleneck Layer



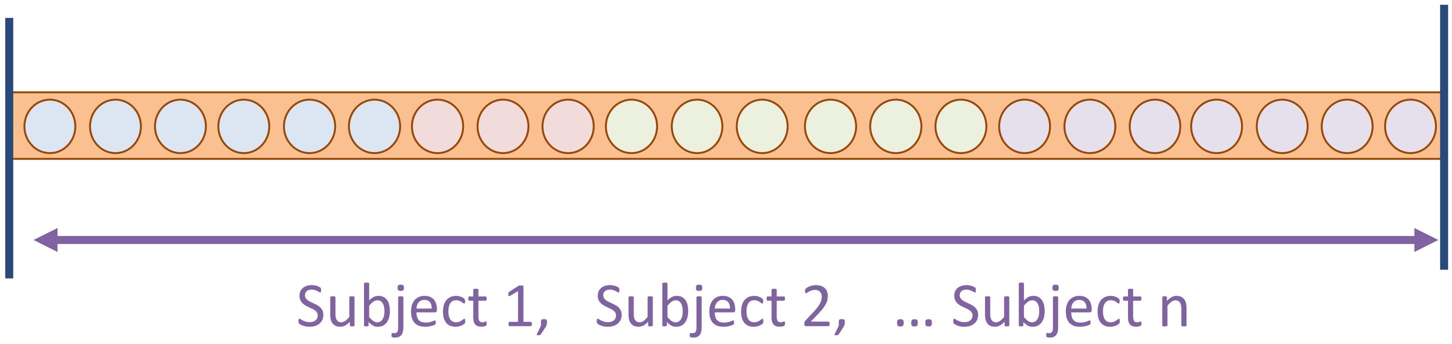
Bottleneck Layer



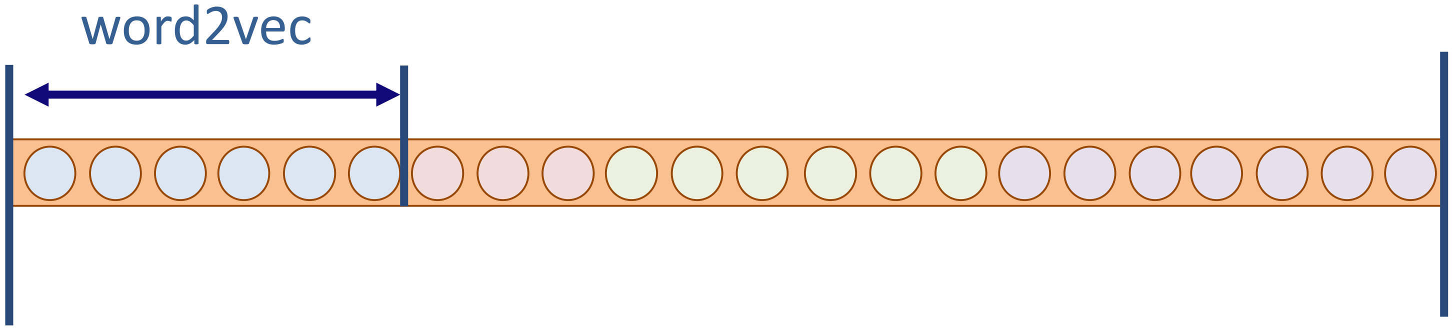
Bottleneck Layer



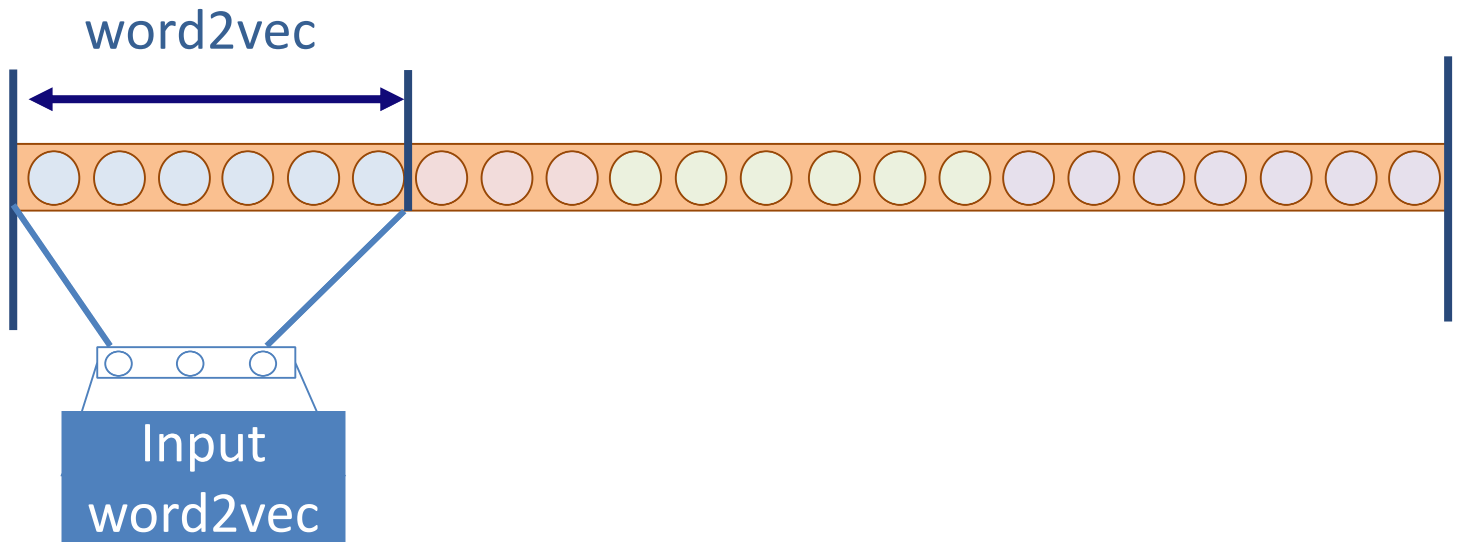
Bottleneck Layer



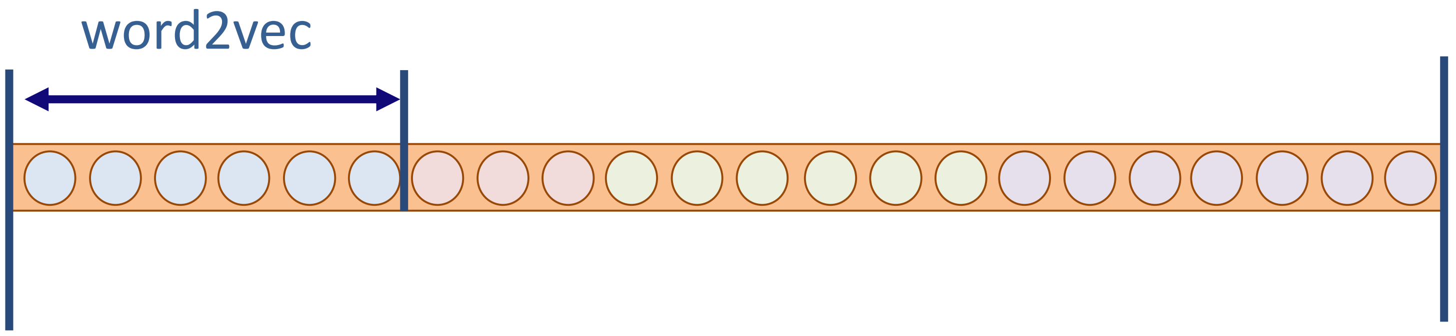
Bottleneck Layer



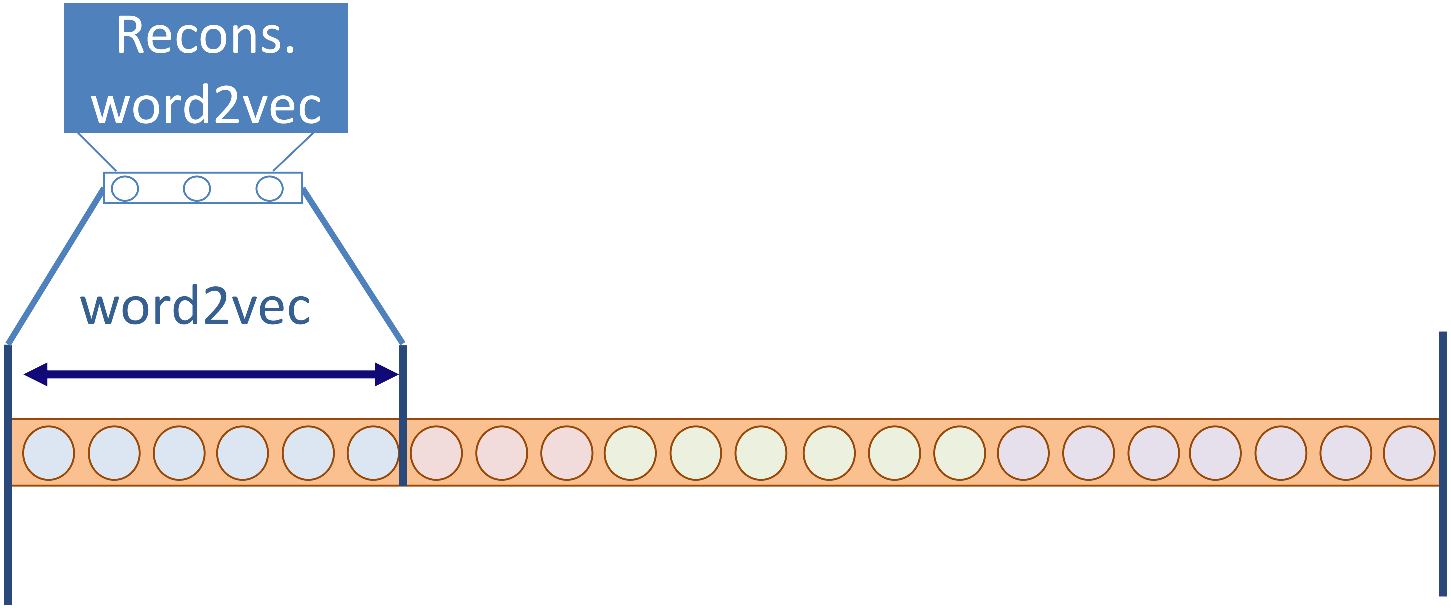
Bottleneck Layer



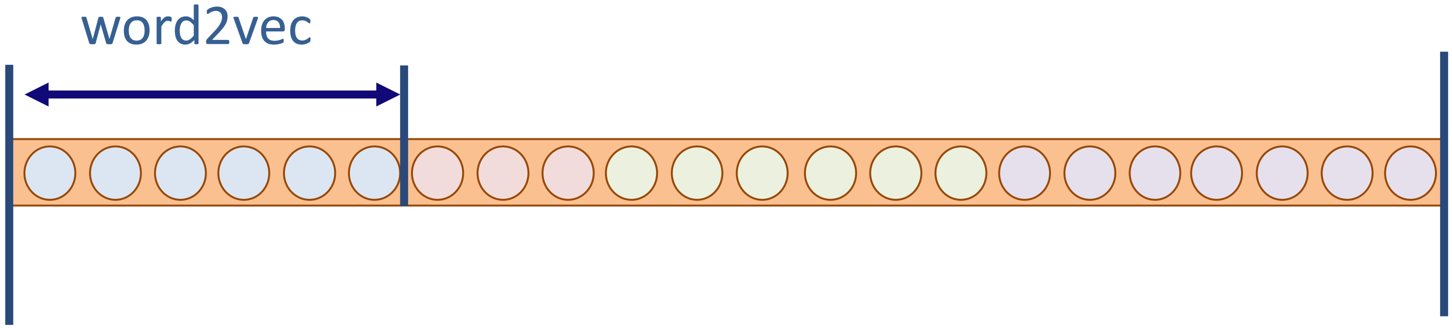
Bottleneck Layer



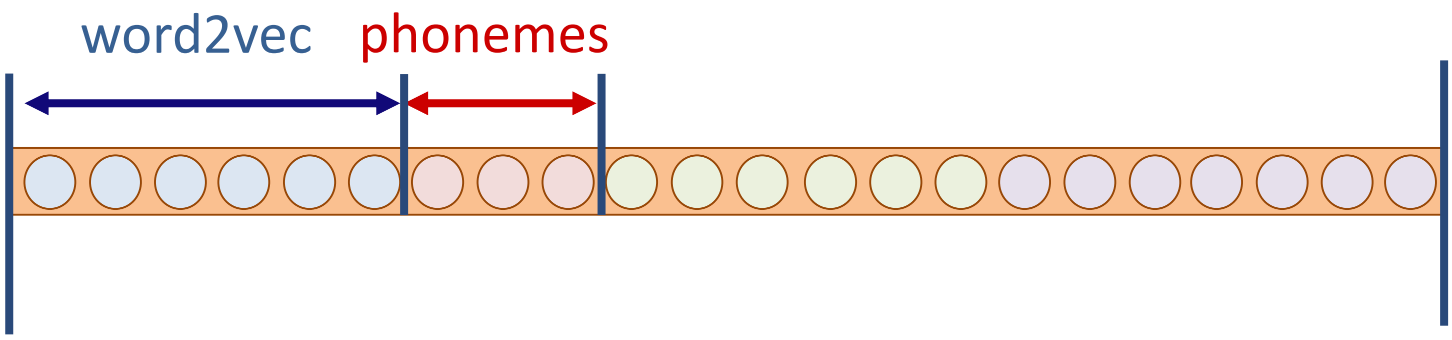
Bottleneck Layer



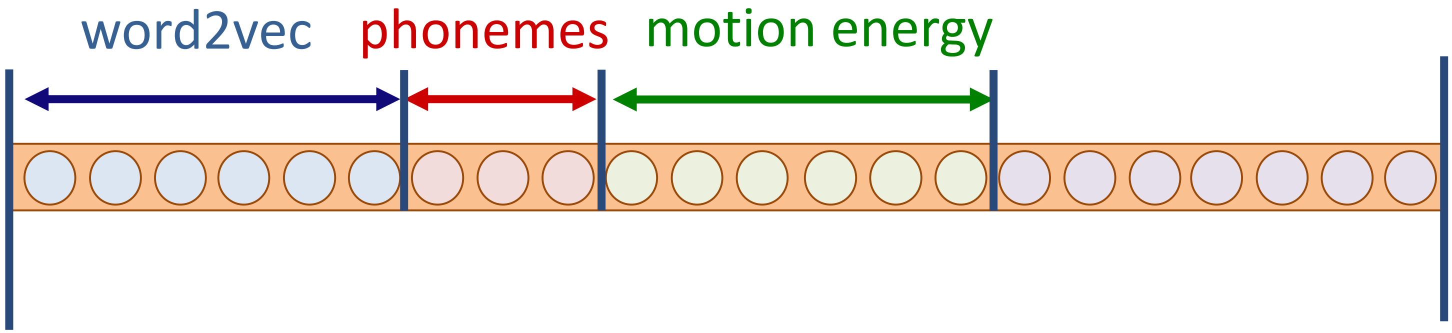
Bottleneck Layer



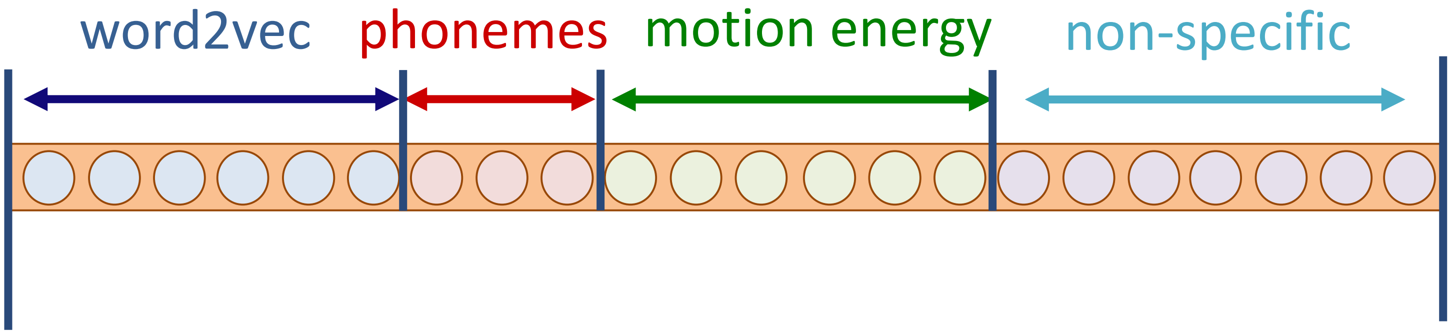
Bottleneck Layer



Bottleneck Layer



Bottleneck Layer



Multimodal representations

- We saw this in the school already
- Bimodal / split auto encoder (Ngiam et al. 2011)
- DCCA, DCCAE (Wang et al. 2015)

Multimodal representations

- We saw this in the school already
- Bimodal / split auto encoder (Ngiam et al. 2011)
- DCCA, DCCAE (Wang et al. 2015)

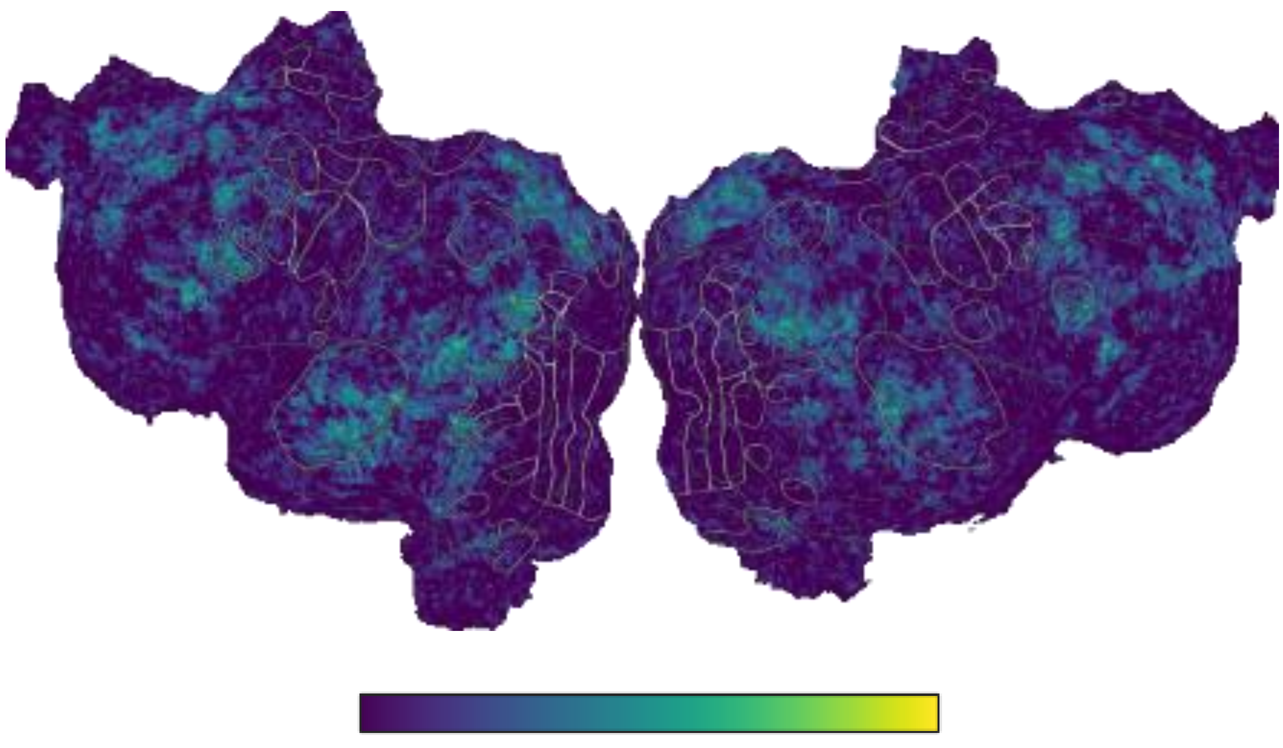
Multimodal fusion of brain structural and functional imaging with a deep neural machine translation approach

Md Faijul Amin[*], Sergey M. Plis[*], Eswar Damaraju[*], Devon Hjelm[*], KyungHyun Cho[†], Vince D. Calhoun[*‡]
[*]The Mind Research Network, 1101 Yale Blvd, Albuquerque, NM 87106, USA
[†]Courant Institute & Center for Data Science, New York University, New York, NY 10012, USA
[‡]Department of ECE, University of New Mexico, Albuquerque, NM 87106, USA

Results: predicting brain activity

Input = word2vec

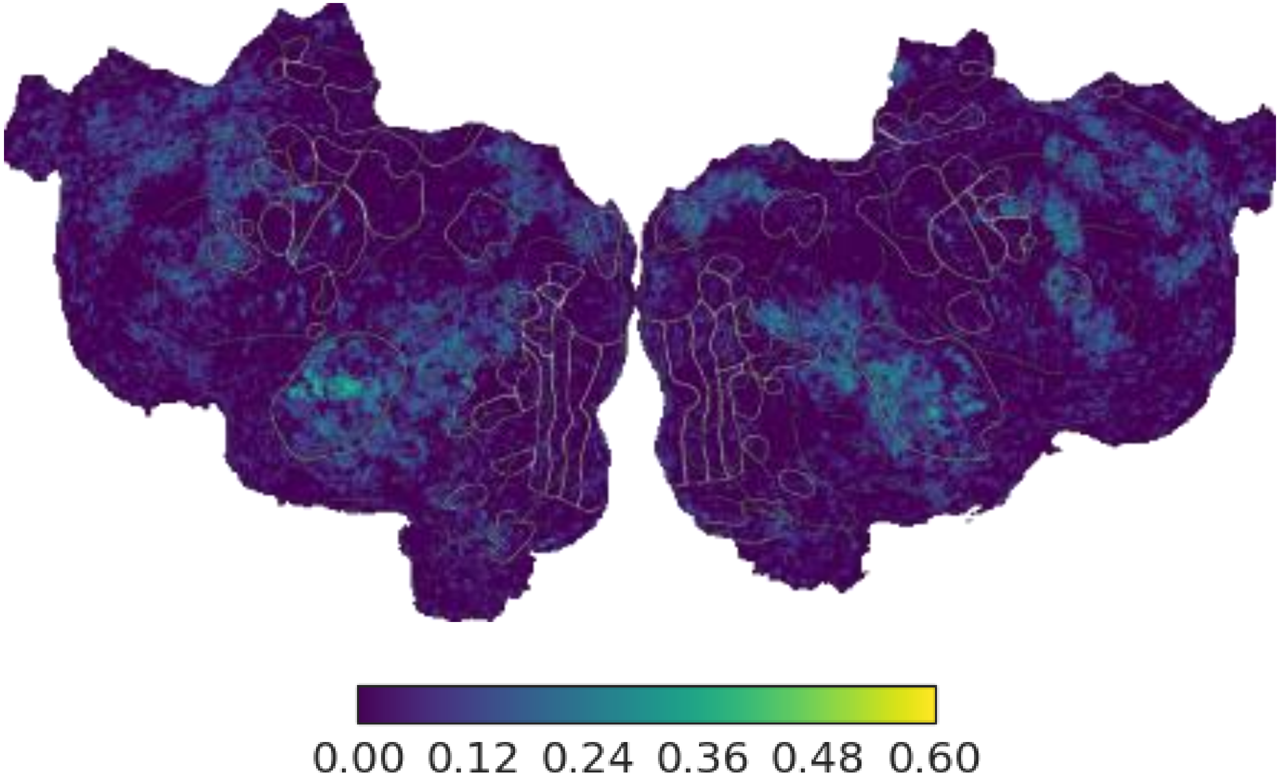
Output = story data



Leila Wehbe

Results: predicting brain activity

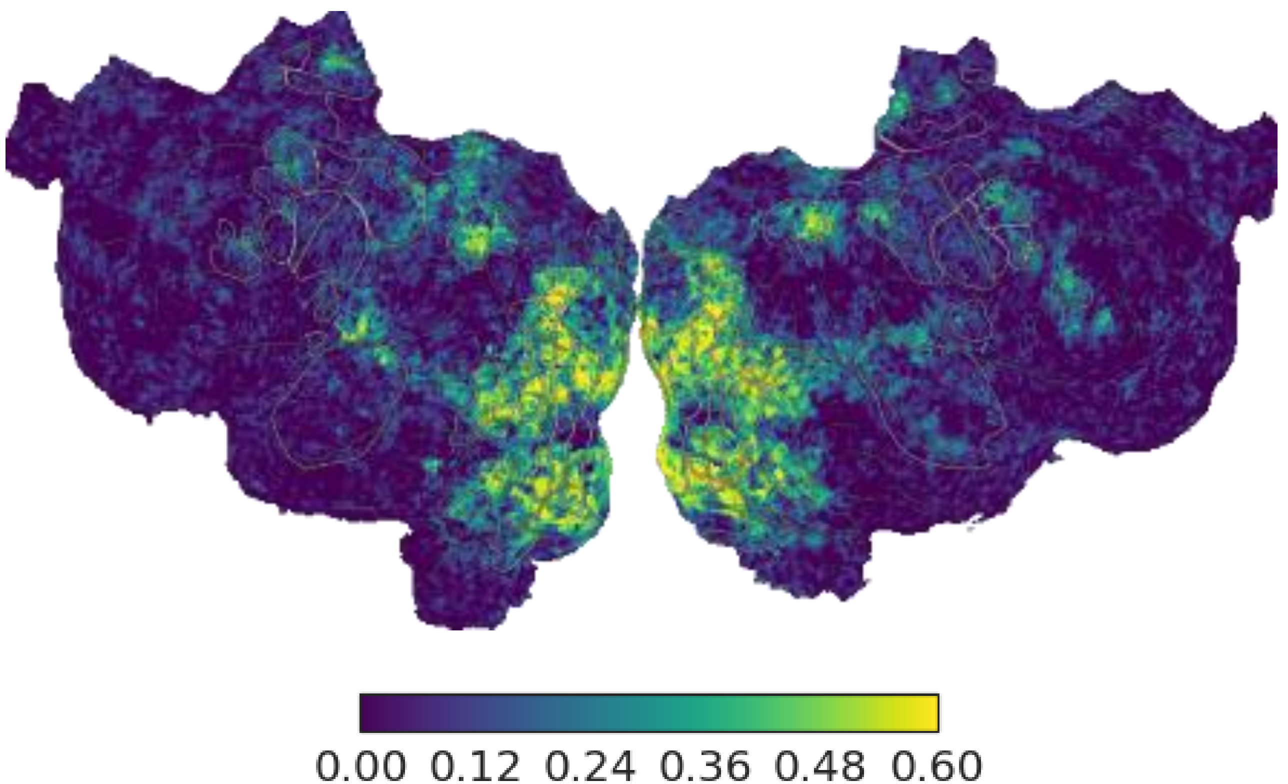
Input = phonemes
Output = story data



Leila Wehbe

Results: predicting brain activity

Input = motion energy
Output = movie data

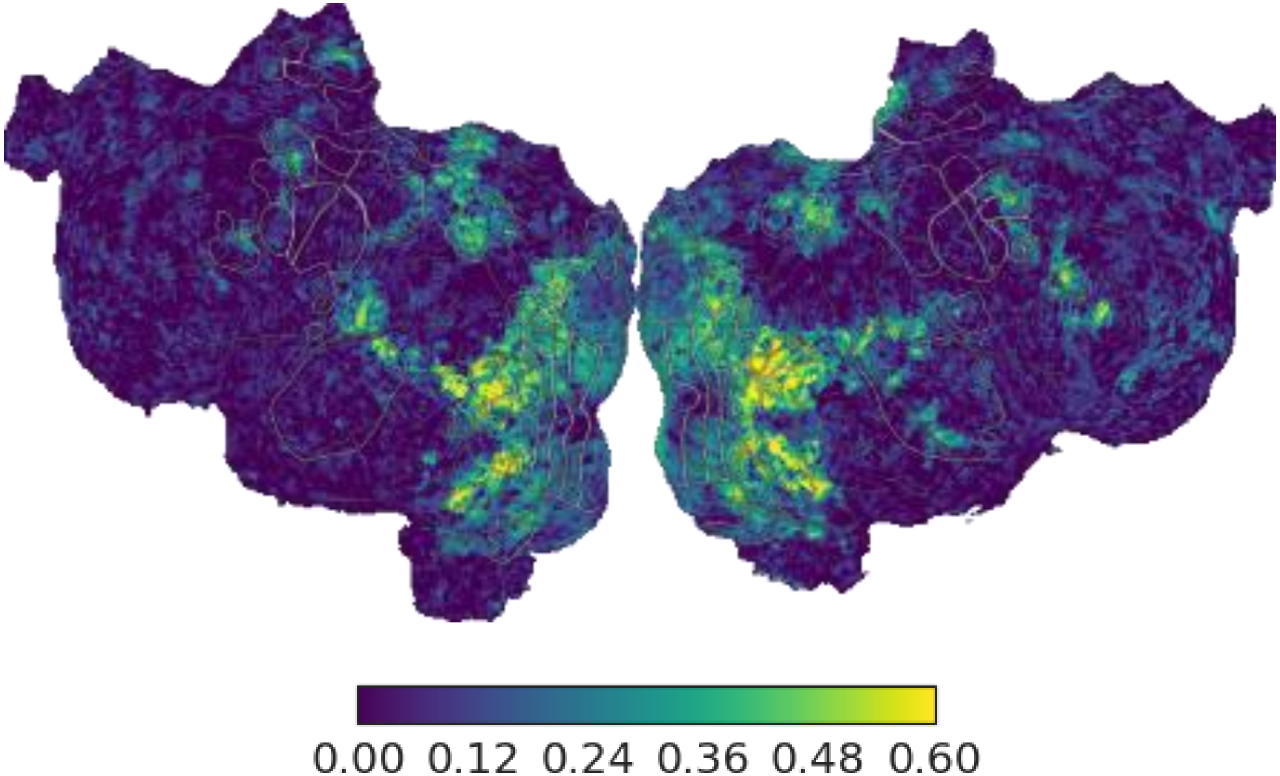


Leila Wehbe

Results: predicting brain activity

Input = word2vec

Output = movie data

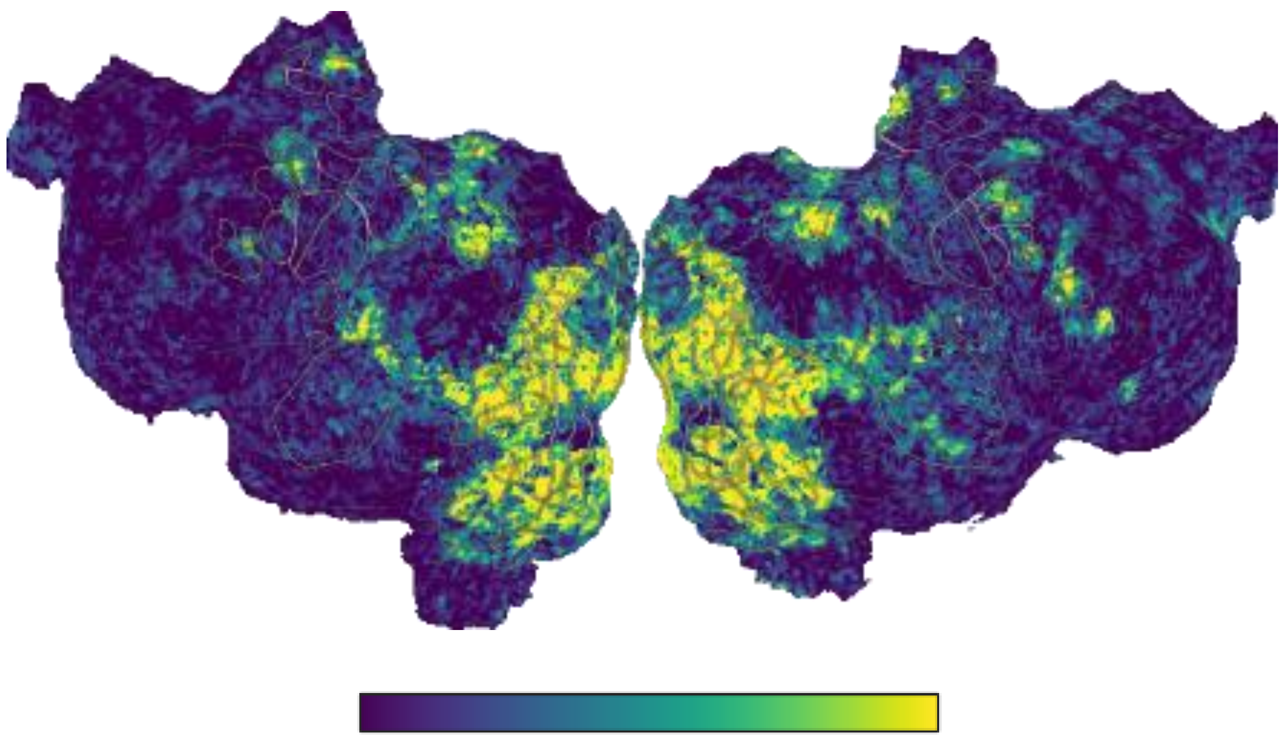


Leila Wehbe

Results: predicting brain activity

Input = S2

Output = movie data (S1)

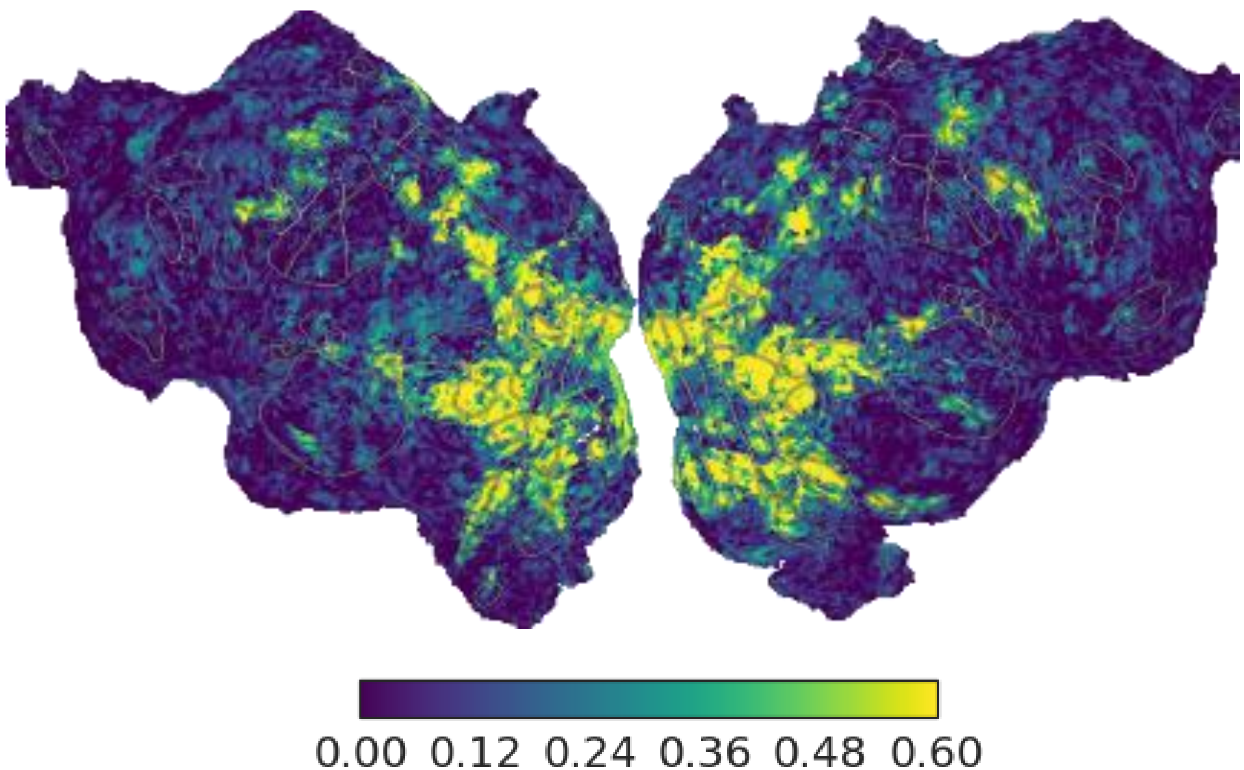


Leila Wehbe

Results: predicting brain activity

Input = S1

Output = movie data (S2)



Leila Wehbe

A unified model

A unified model

- For experiments, subjects and feature spaces

A unified model

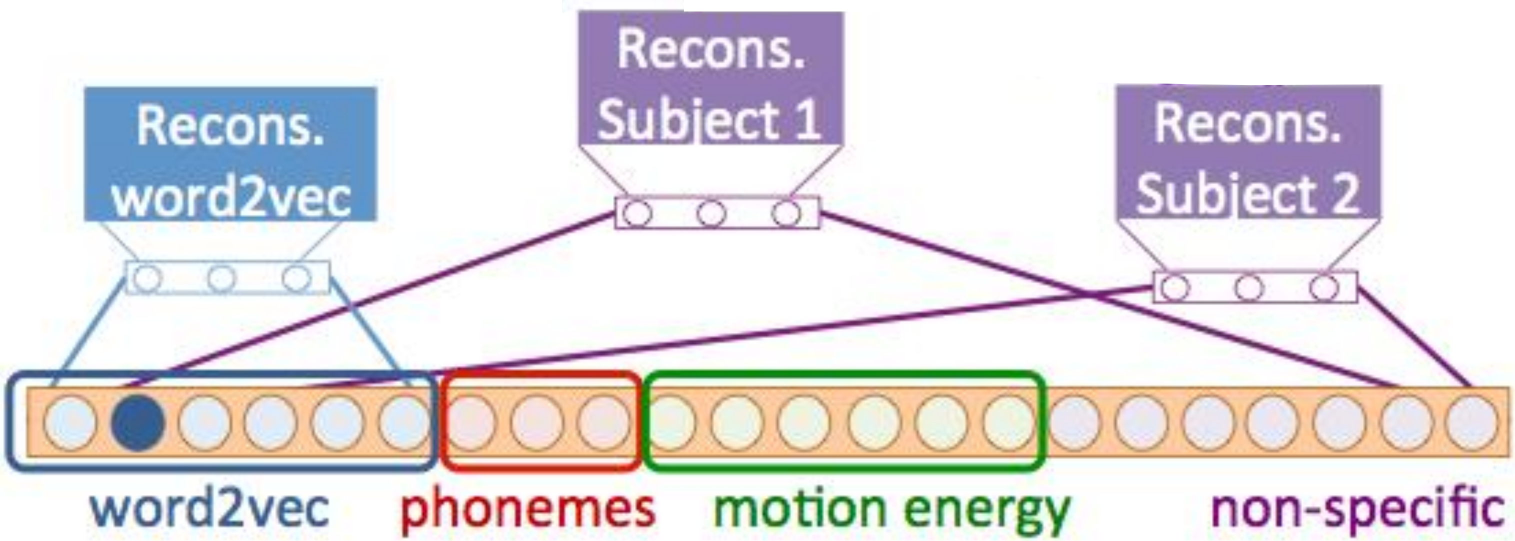
- For experiments, subjects and feature spaces
- Enables standard neuroimaging tasks
  - Good performance

A unified model

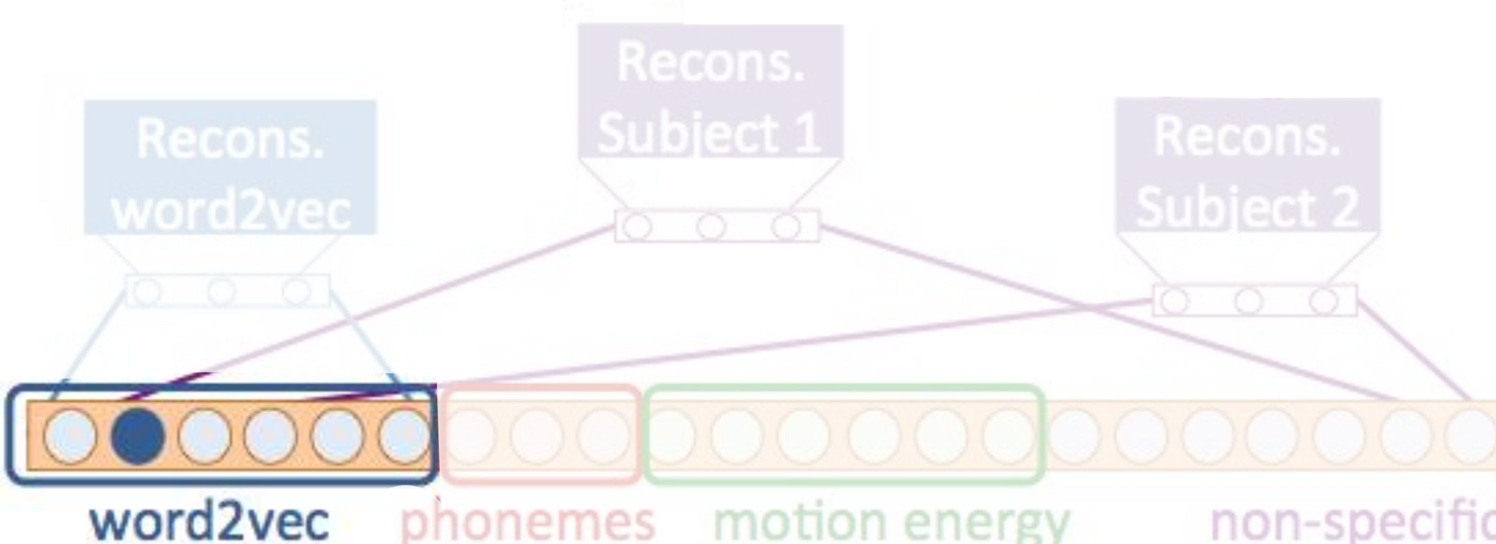
- For experiments, subjects and feature spaces
- Enables standard neuroimaging tasks
  - Good performance

- Also learns an embedding space for brain responses / stimulus features

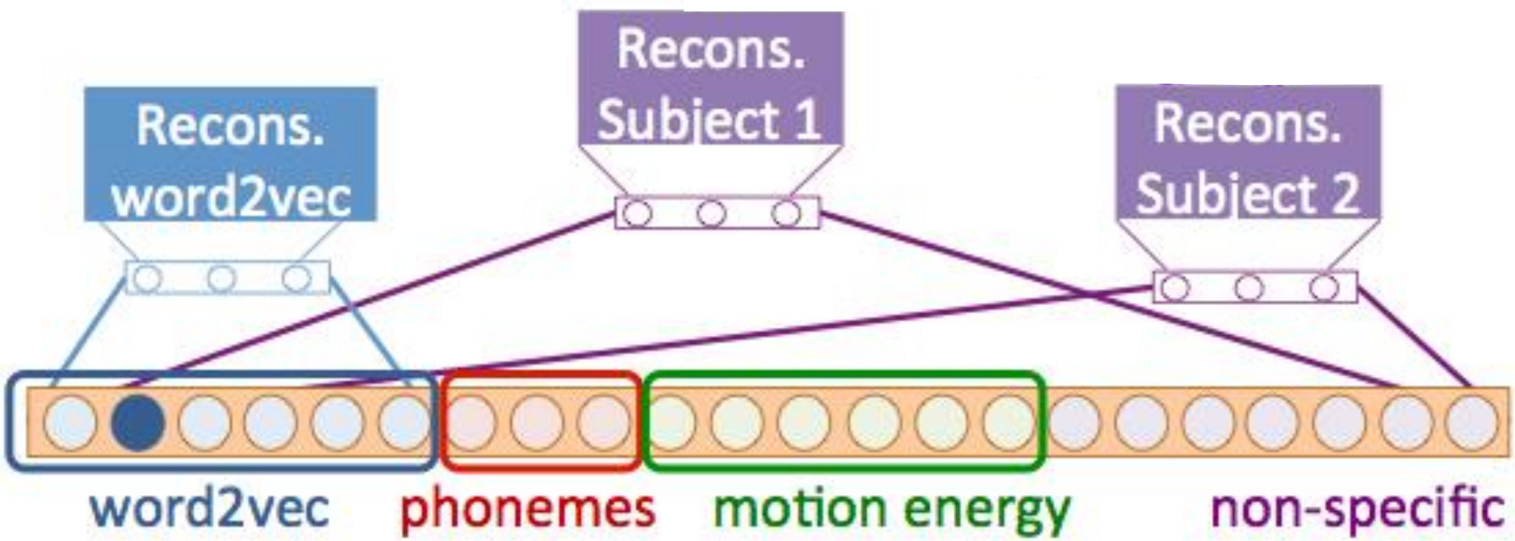
Exploring the learned space



Exploring the learned space



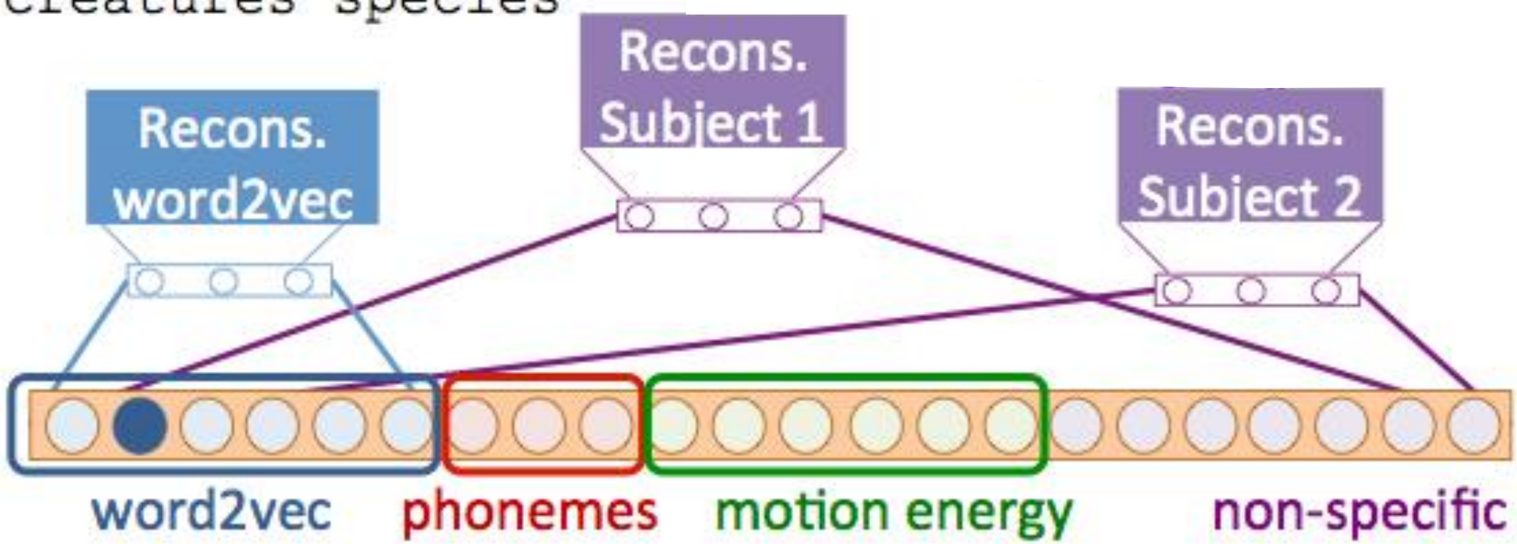
Exploring the learned space



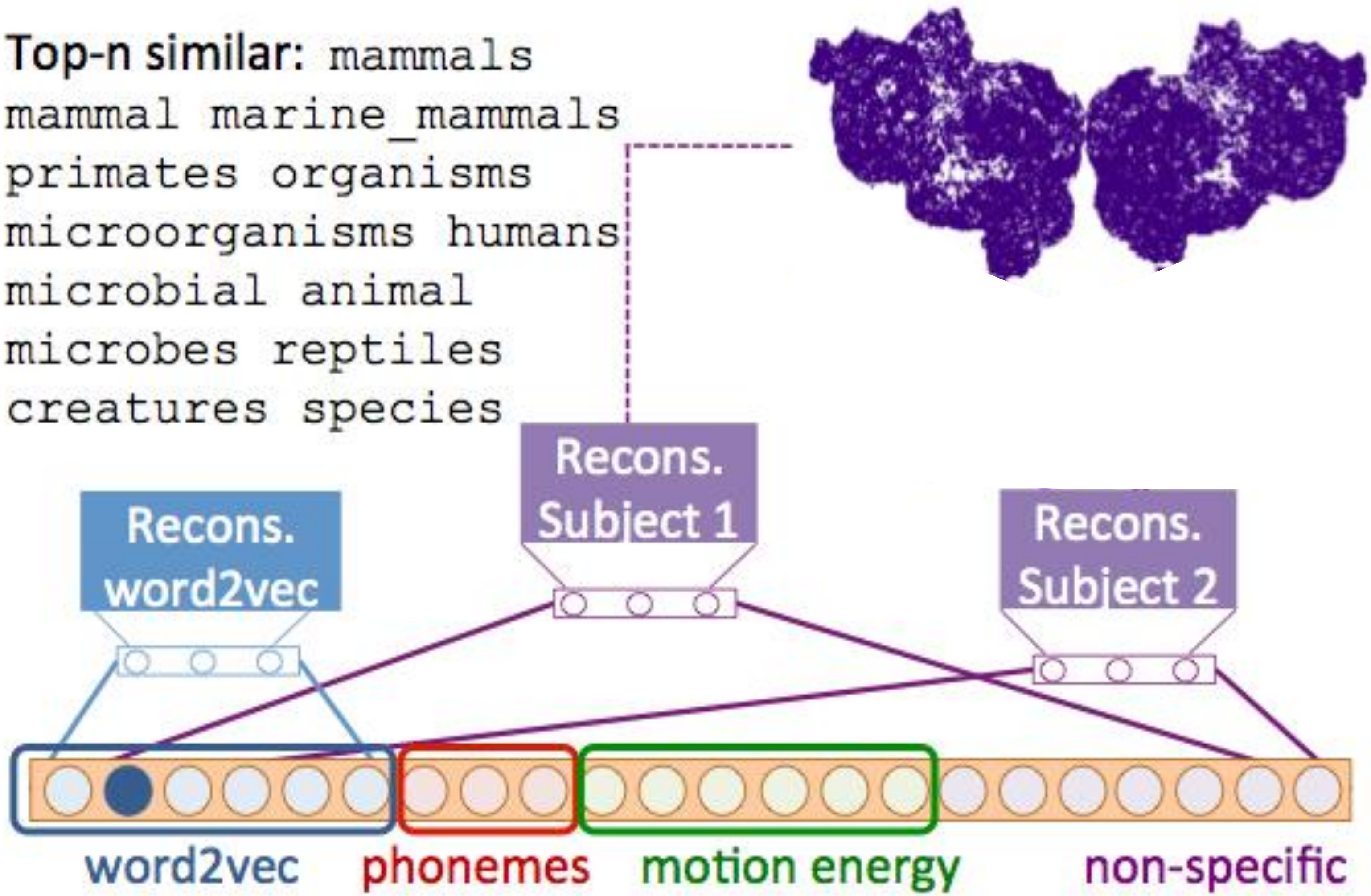
Exploring the learned space

Top-n similar: mammals

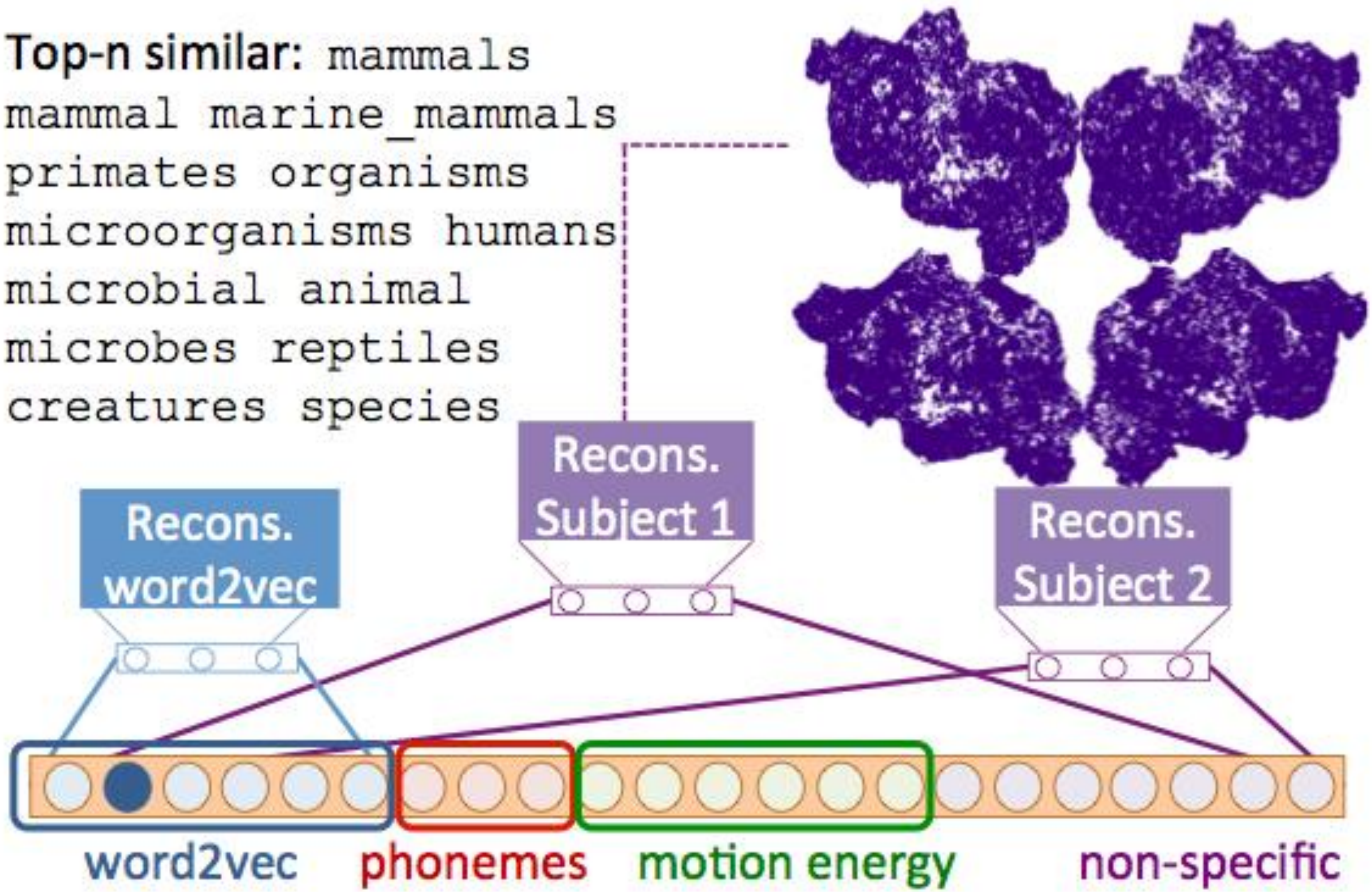
mammal marine_mammals
primates organisms
microorganisms humans
microbial animal
microbes reptiles
creatures species



Exploring the learned space



Exploring the learned space



Future work

Future work

- Use spatial information

Future work

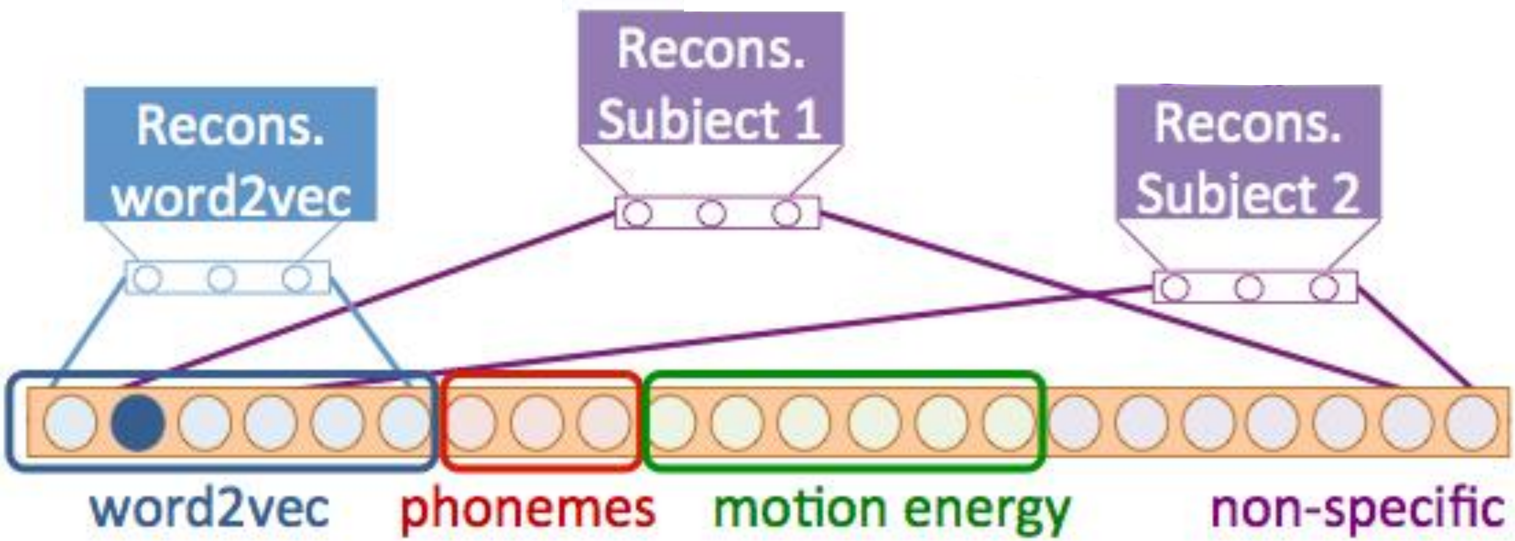
- Use spatial information

- Use temporal information

Future work

- Use spatial information
- Use temporal information
- Learn allocation of feature spaces automatically, not as hyper parameter

Thank you


Recons.
word2vec
Recons.
Subject 1
Recons.
Subject 2
word2vec
phonemes
motion energy
non-specific