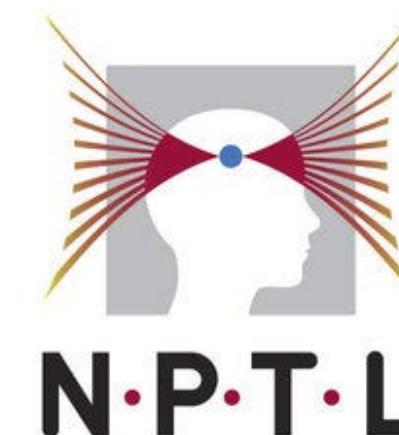


# Speech Brain-Computer Interfaces for Restoring Natural Communication

Past, now, and future

STANFORD UNIVERSITY



Neural Prosthetics Translational Laboratory

Chaofei Fan, CS224N

BRAINGATE  
TURNING THOUGHT INTO ACTION



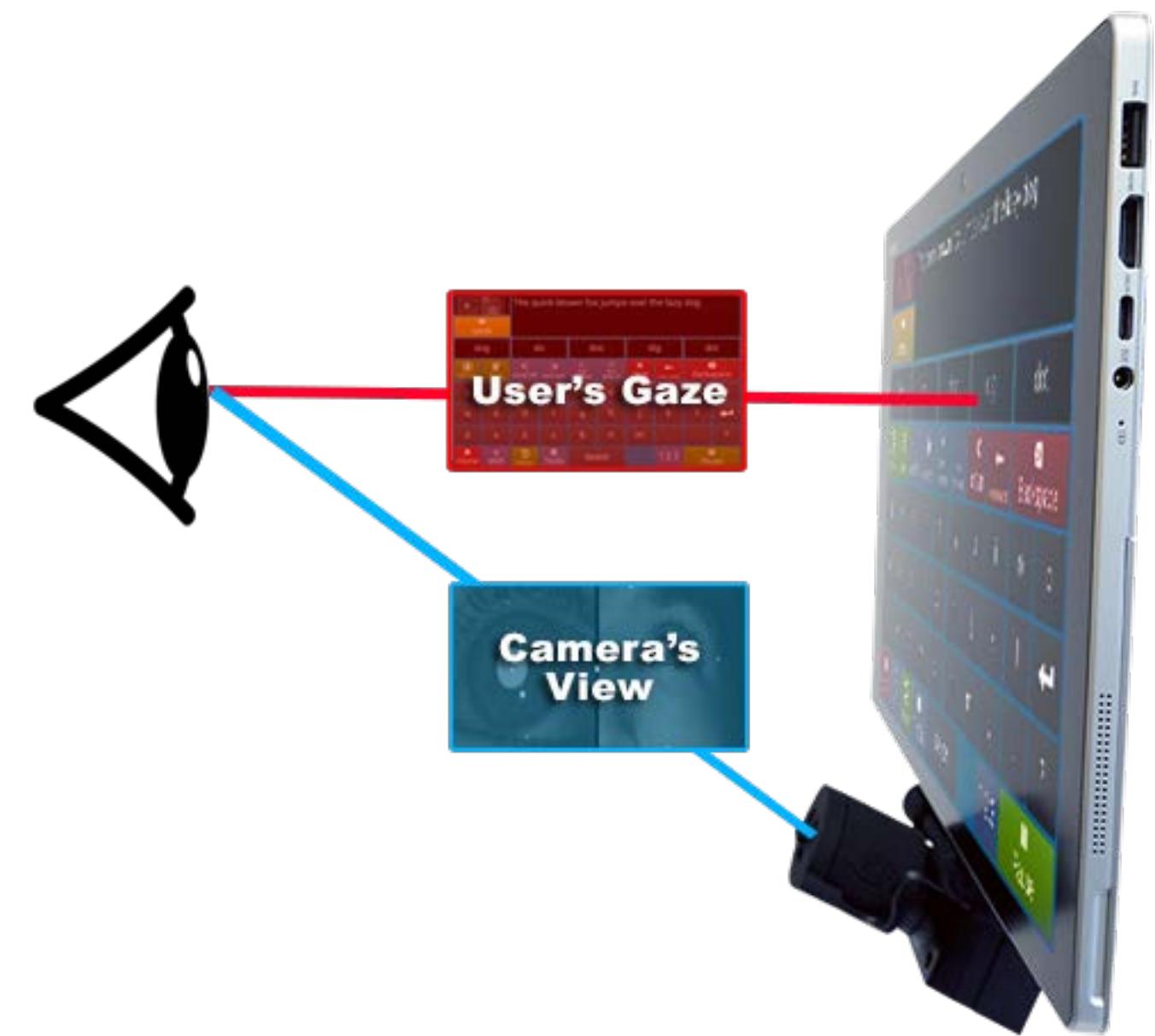


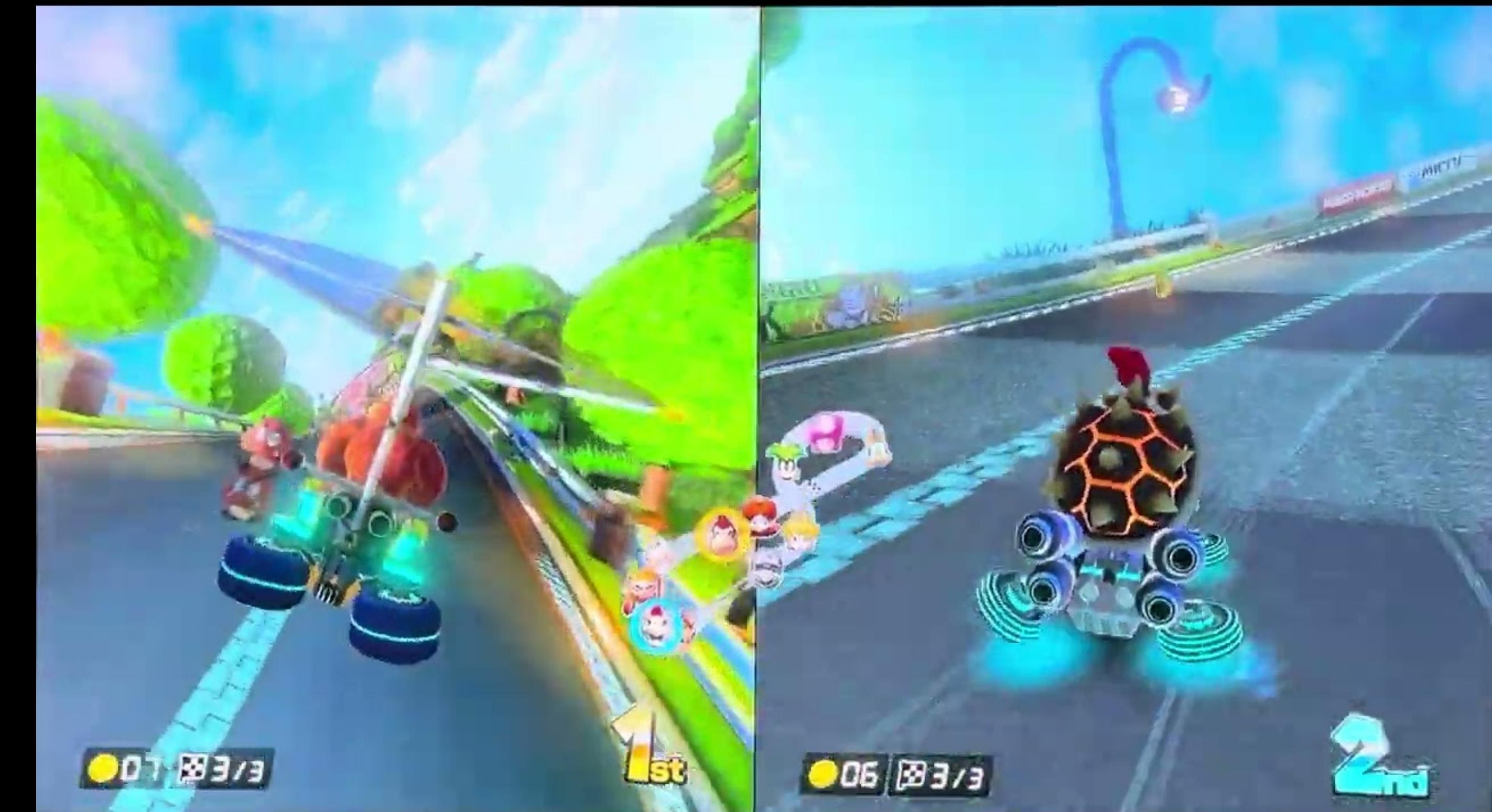
**DOING ACTIVITIES HAVING LESS ON YOUR MIND**

# What we take for granted is lost for some individuals

- Howard Wicks, 21, lost all his dreams after a sever stroke.
- Neurological disorders like brainstem stroke or Amyotrophic Lateral Sclerosis (ALS) can cause speech and motor impairment and even complete loss of speech.
- These individuals are facing extreme challenges in their lives.
- Communication with loved ones and caretakers is one of their most desperate needs.

# Assistive communication devices



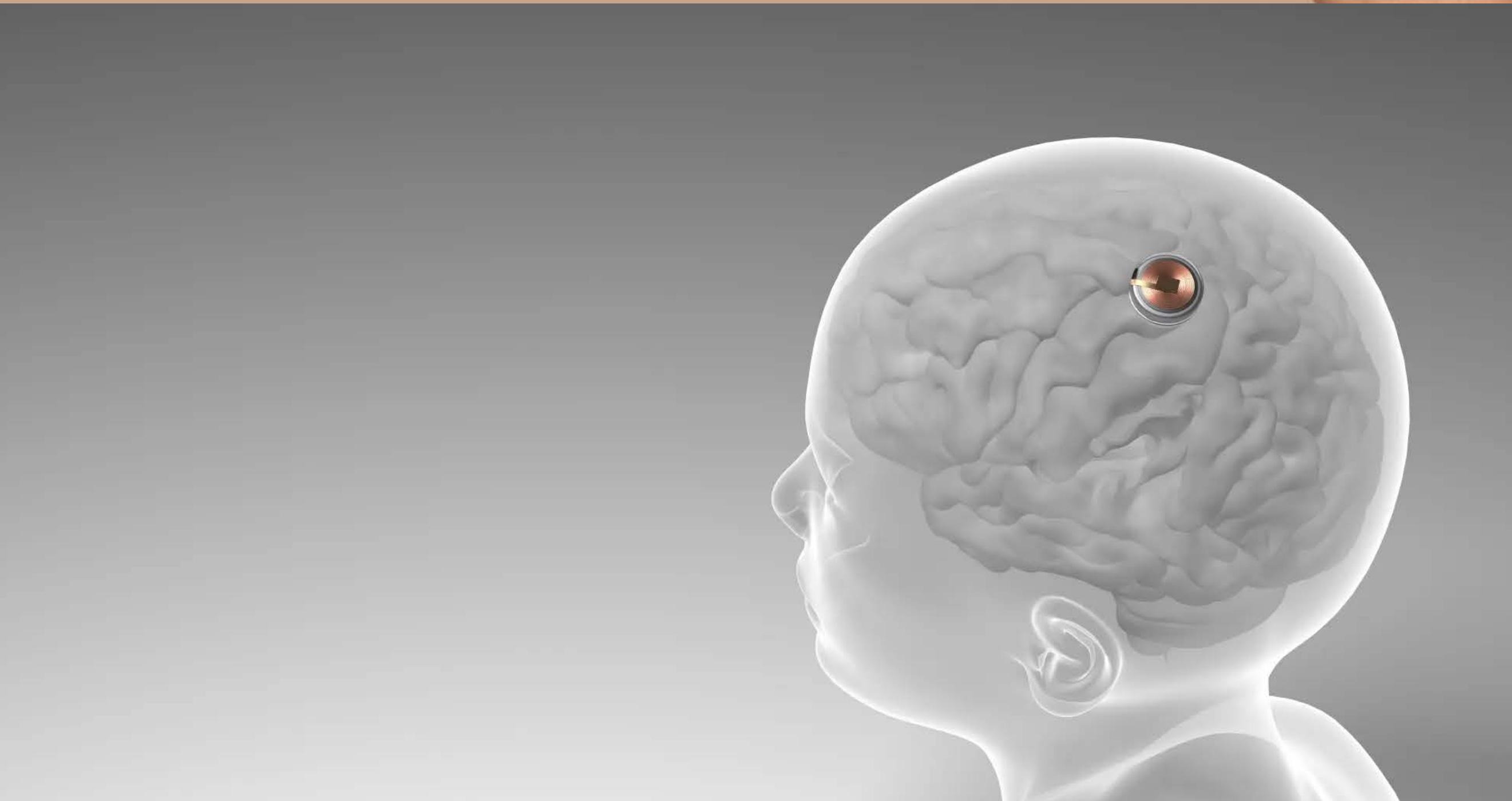


Neuralink Prime Study

Tap into the intact  
mind with brain-  
computer interfaces  
(BCIs)

“[The Link] has helped me reconnect with the world, my friends, and my family. It's given me the ability to do things on my own again without needing my family at all hours of the day and night.”

— Noland Arbaugh, PRIME Study participant



# A Brief History of BCI

# How does BCI work? Start from the beginning...

## Electricity in the brain

·278

THE BRITISH MEDICAL JOURNAL.

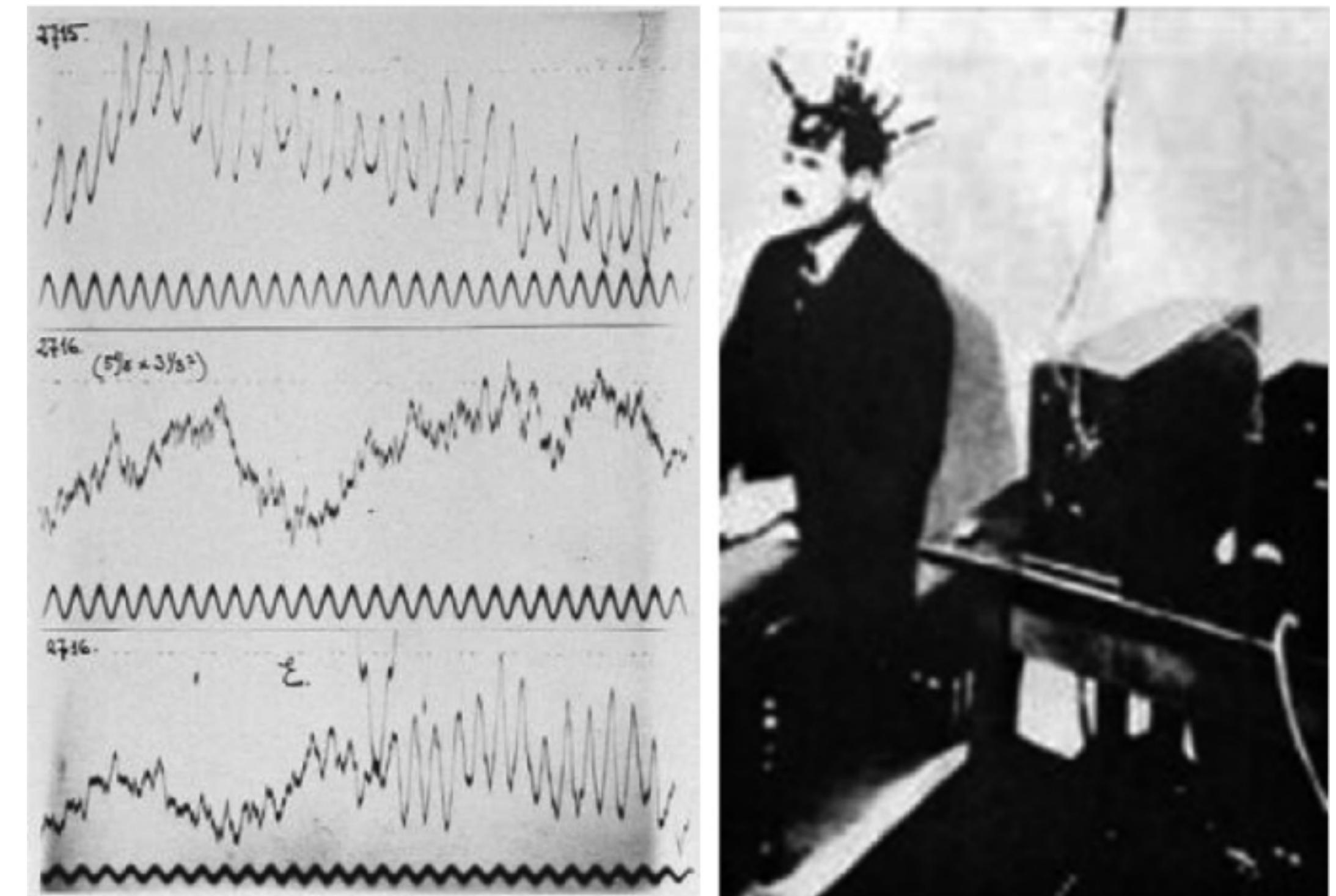
[August 28, 1875.]

*The Electric Currents of the Brain.* By RICHARD CATON, M.D., Liverpool.—After a brief *résumé* of previous investigations, the author gave an account of his own experiments on the brains of the rabbit and the monkey. The following is a brief summary of the principal results. In every brain hitherto examined, the galvanometer has indicated the existence of electric currents. The external surface of the grey matter is usually positive in relation to the surface of a section through it. Feeble currents of varying direction pass through the multiplier when the electrodes are placed on two points of the external surface, or one electrode on the grey matter, and one on the surface of the skull. The electric currents of the grey matter appear to have a relation to its function. When any part of the grey matter is in a state of functional activity, its electric current usually exhibits negative variation. For example, on the areas shown by Dr. Ferrier to be related to rotation of the head and to mastication, negative variation of the current was observed to occur whenever those two acts respectively were performed. Impressions through the senses were found to influence the currents of certain areas; *e. g.*, the currents of that part of the rabbit's brain which Dr. Ferrier has shown to be related to movements of the eyelids, were found to be markedly influenced by stimulation of the opposite retina by light.

# Listen to the brain from the outside

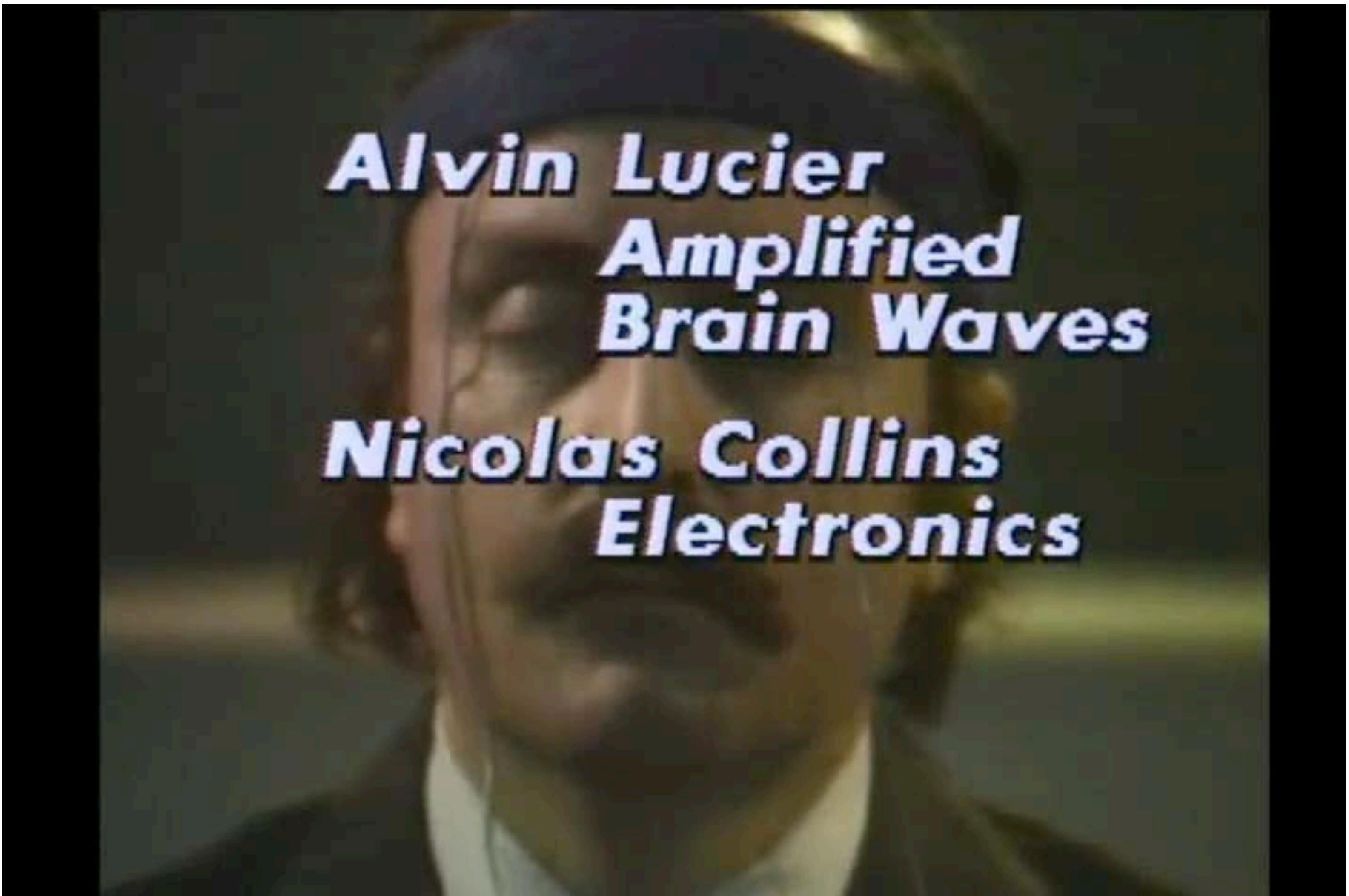
## Electroencephalogram (EEG)

Hans Berger, a German psychiatrist, invented EEG and succeeded in recording the first human EEG in 1924.

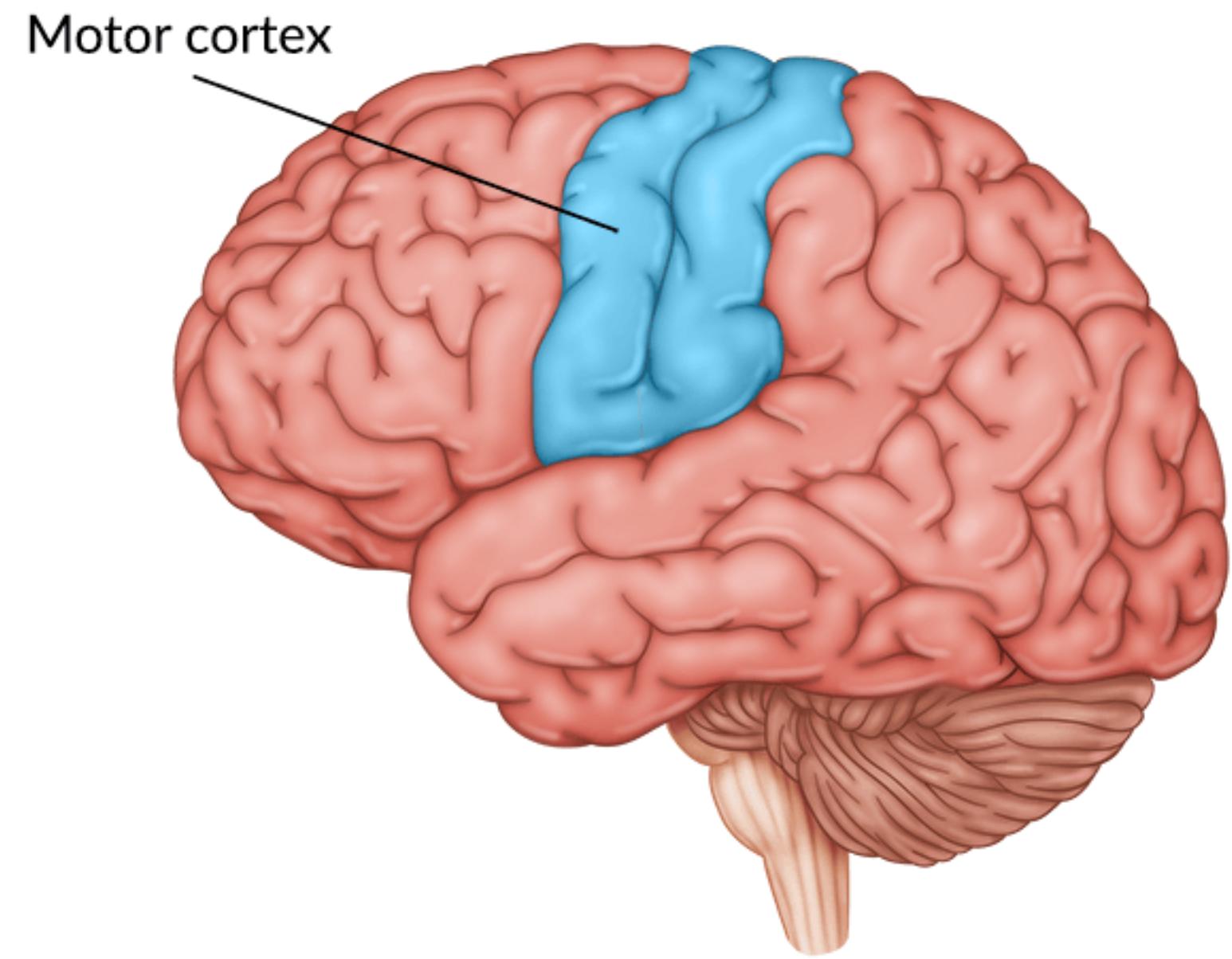
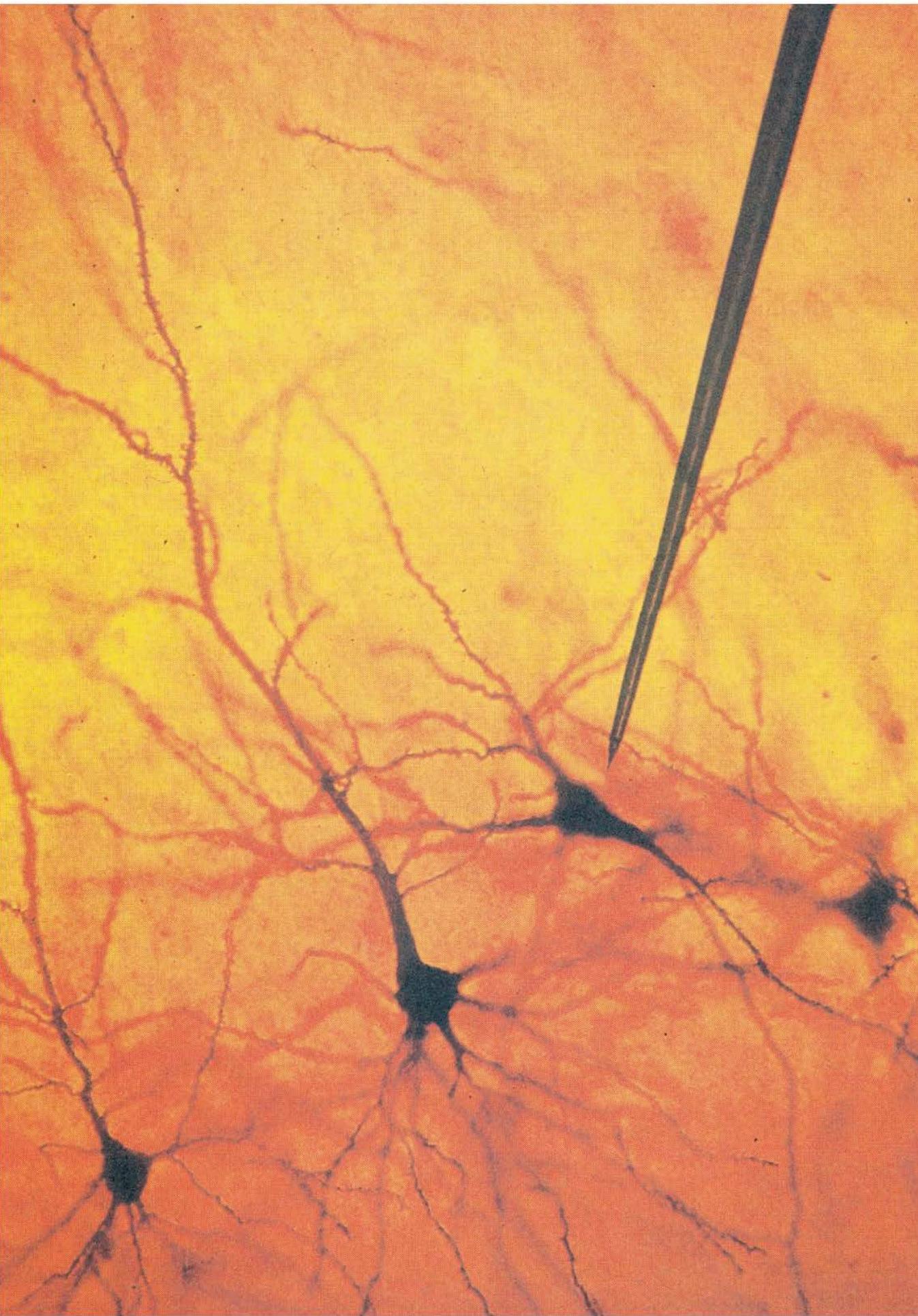


# Perform with brain waves

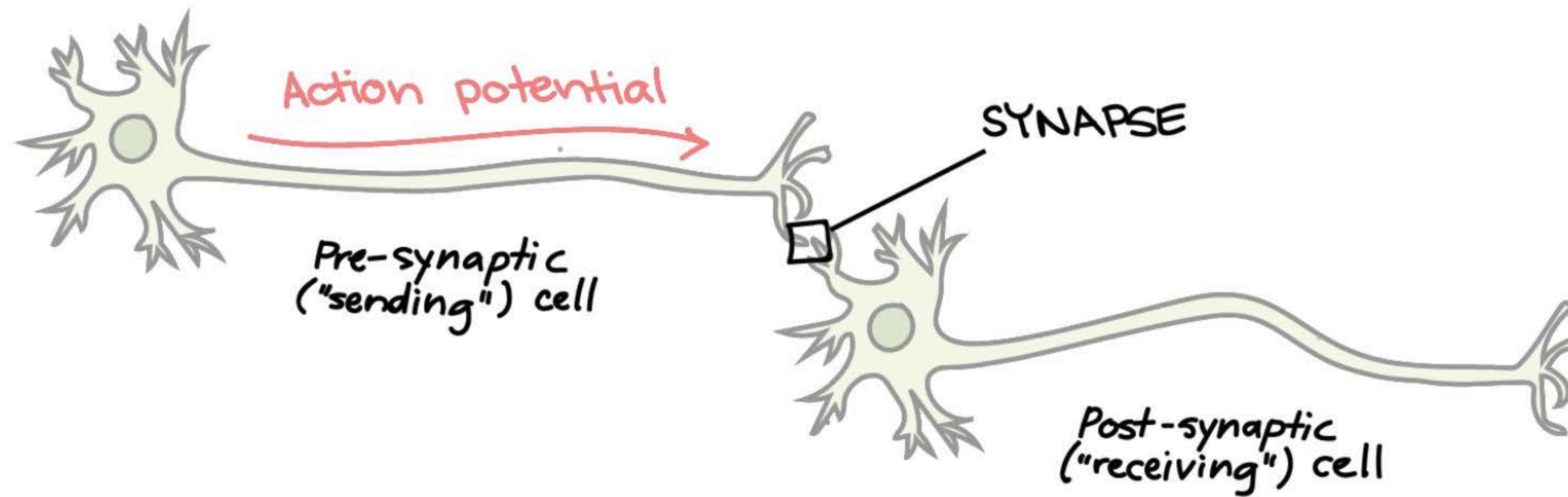
Electroencephalogram (EEG)



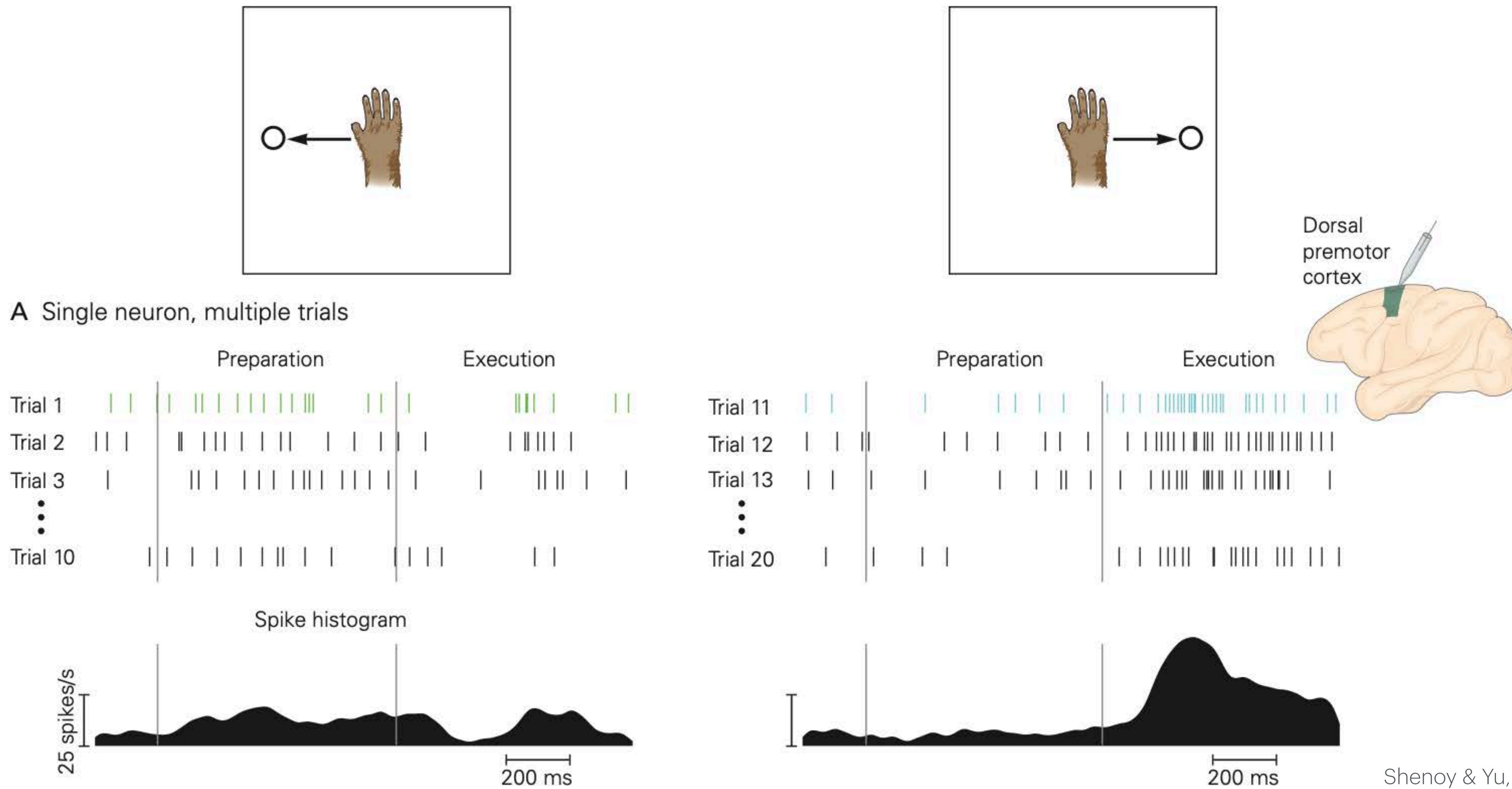
# Listen to single neurons in motor cortex



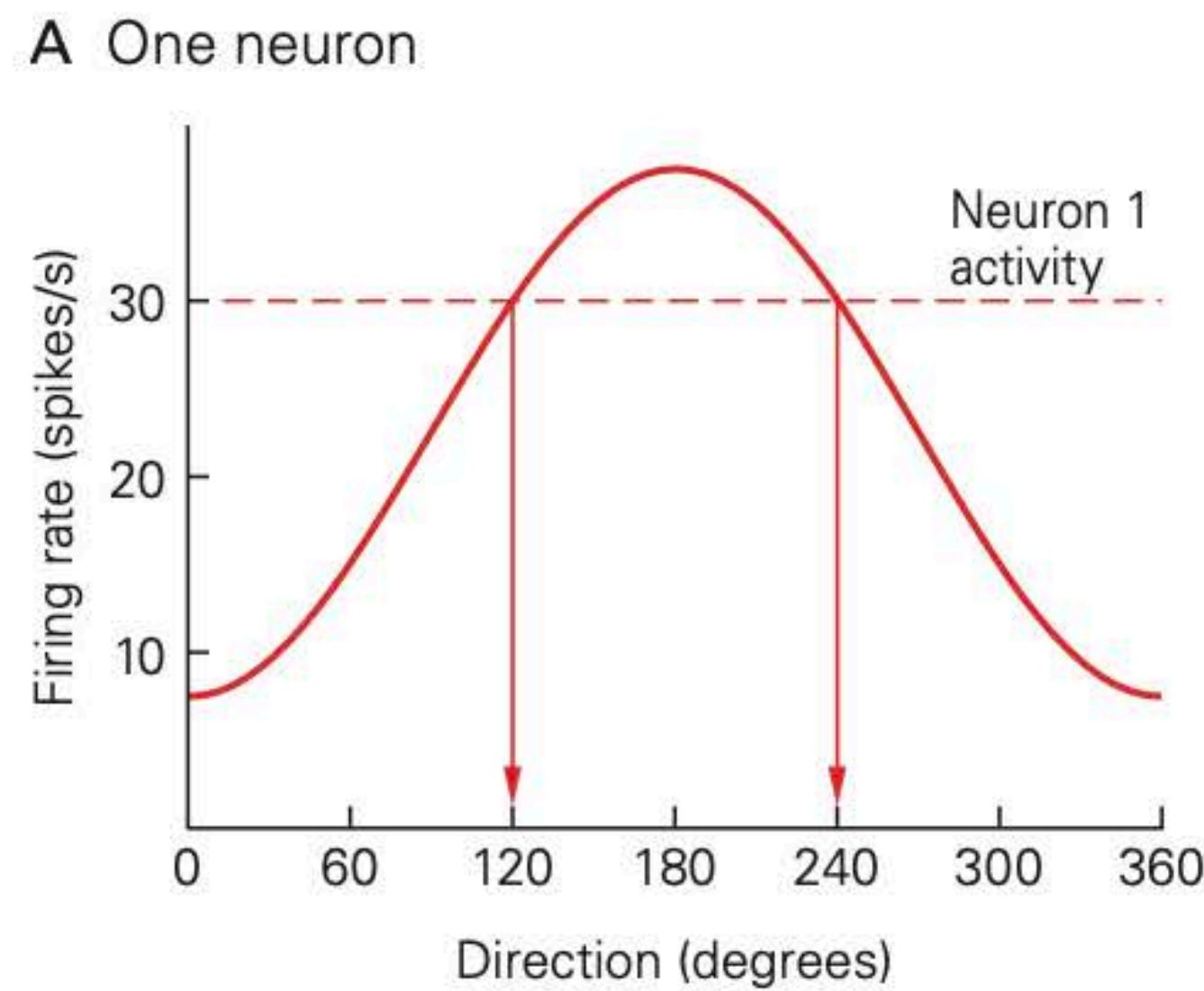
# Neurons communicate with spikes



# Single neurons encode movement directions

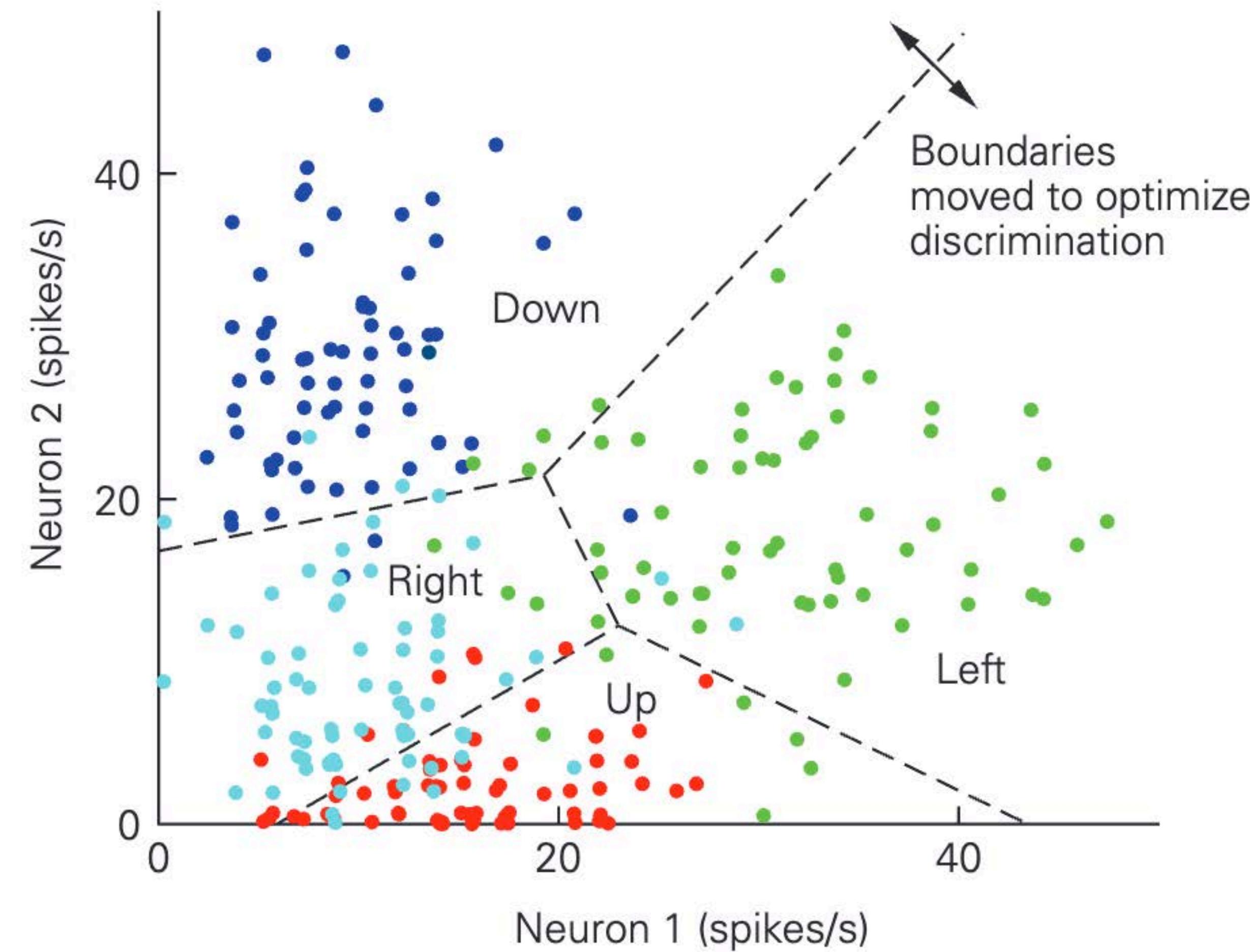


# Multiple neurons for accurate decoding

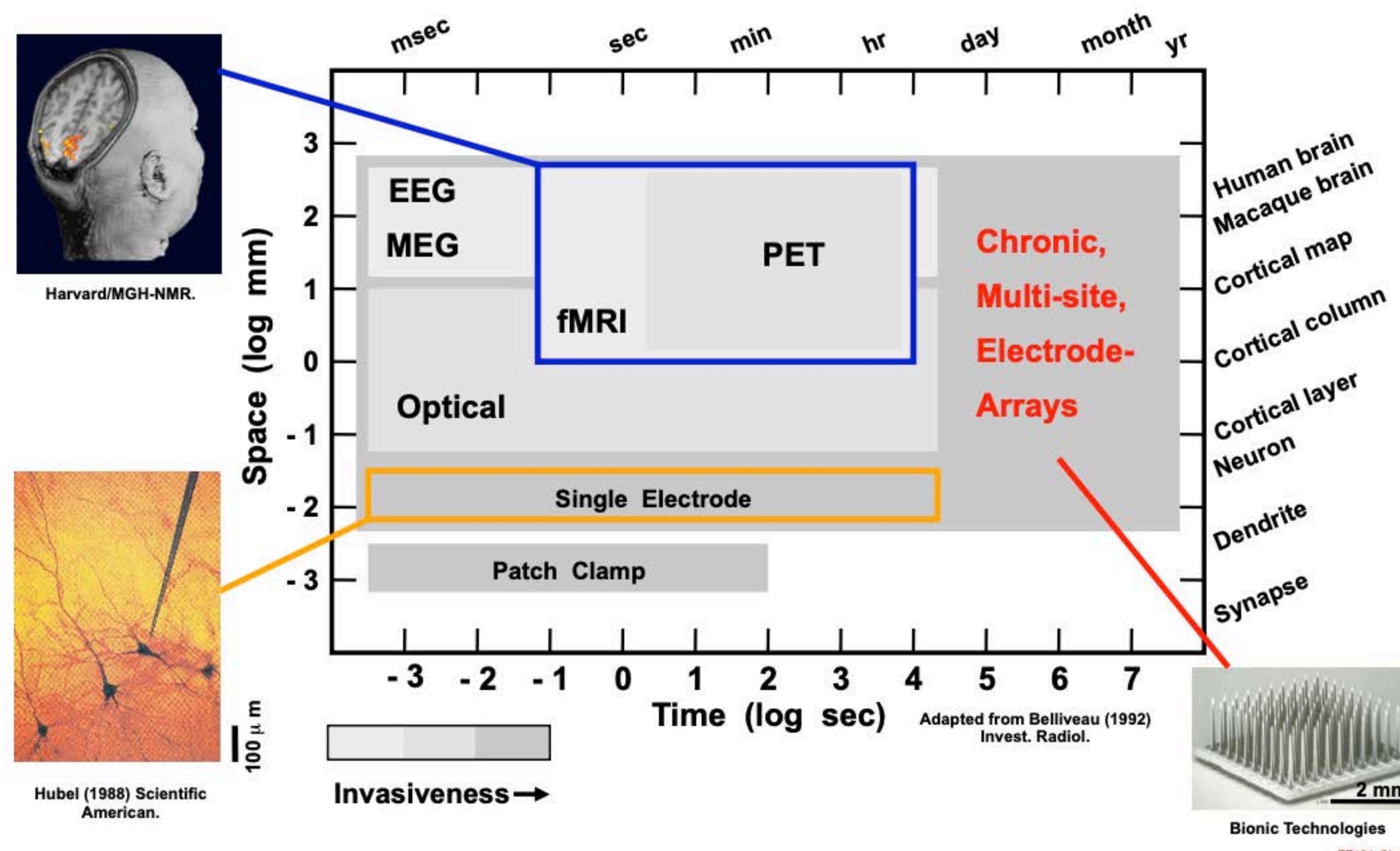


# Multiple neurons for accurate decoding

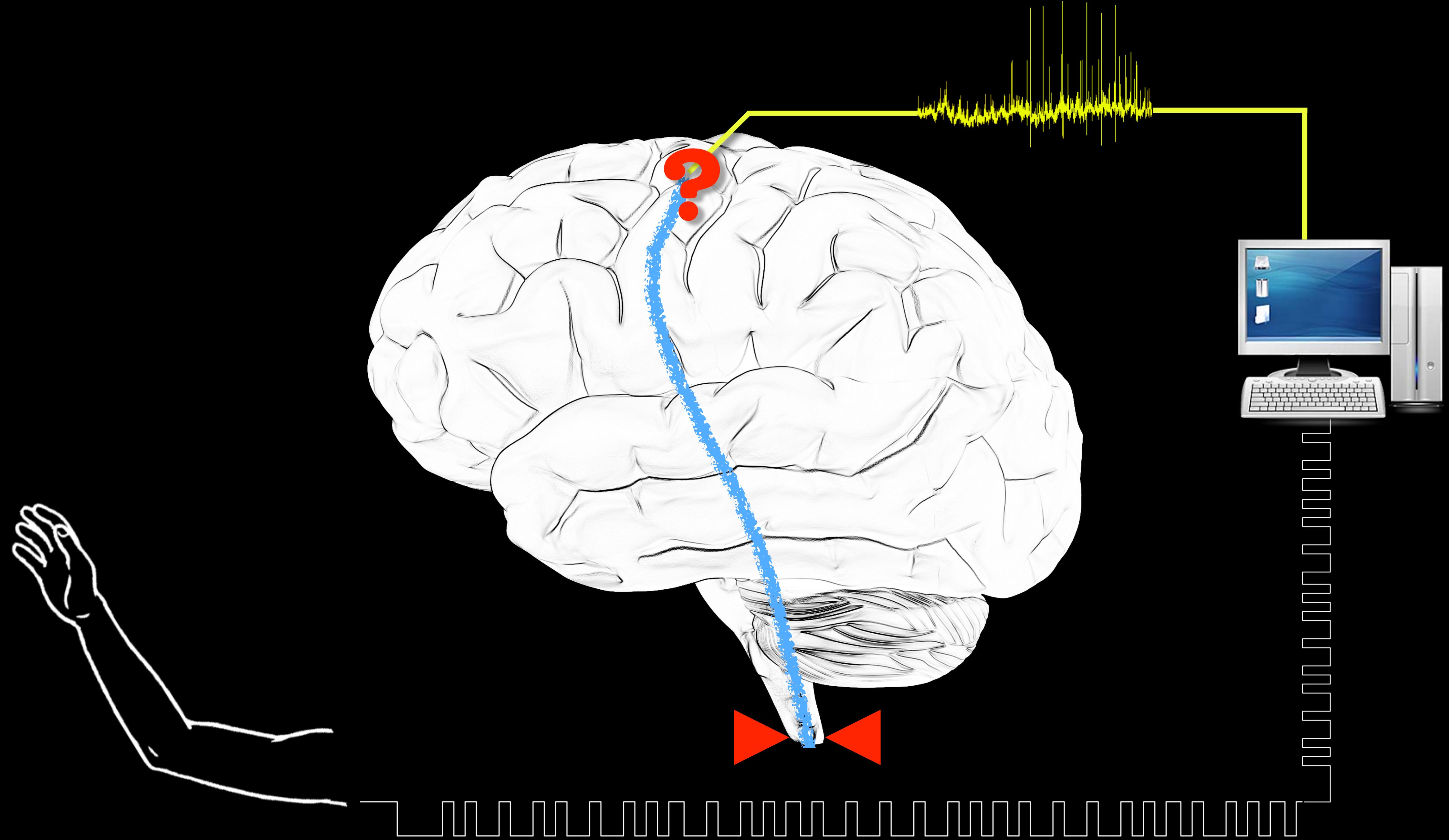
A Calibration phase

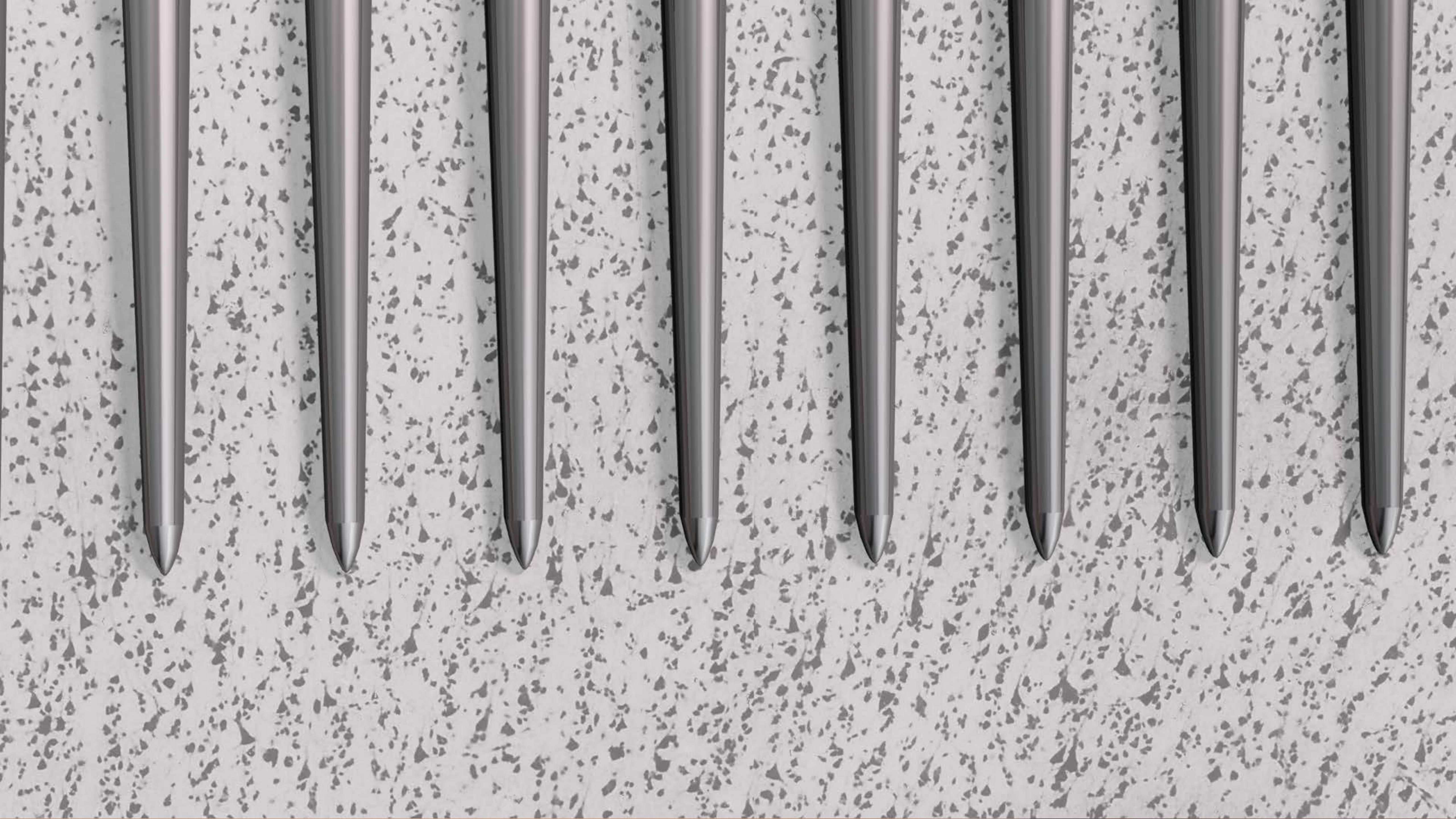


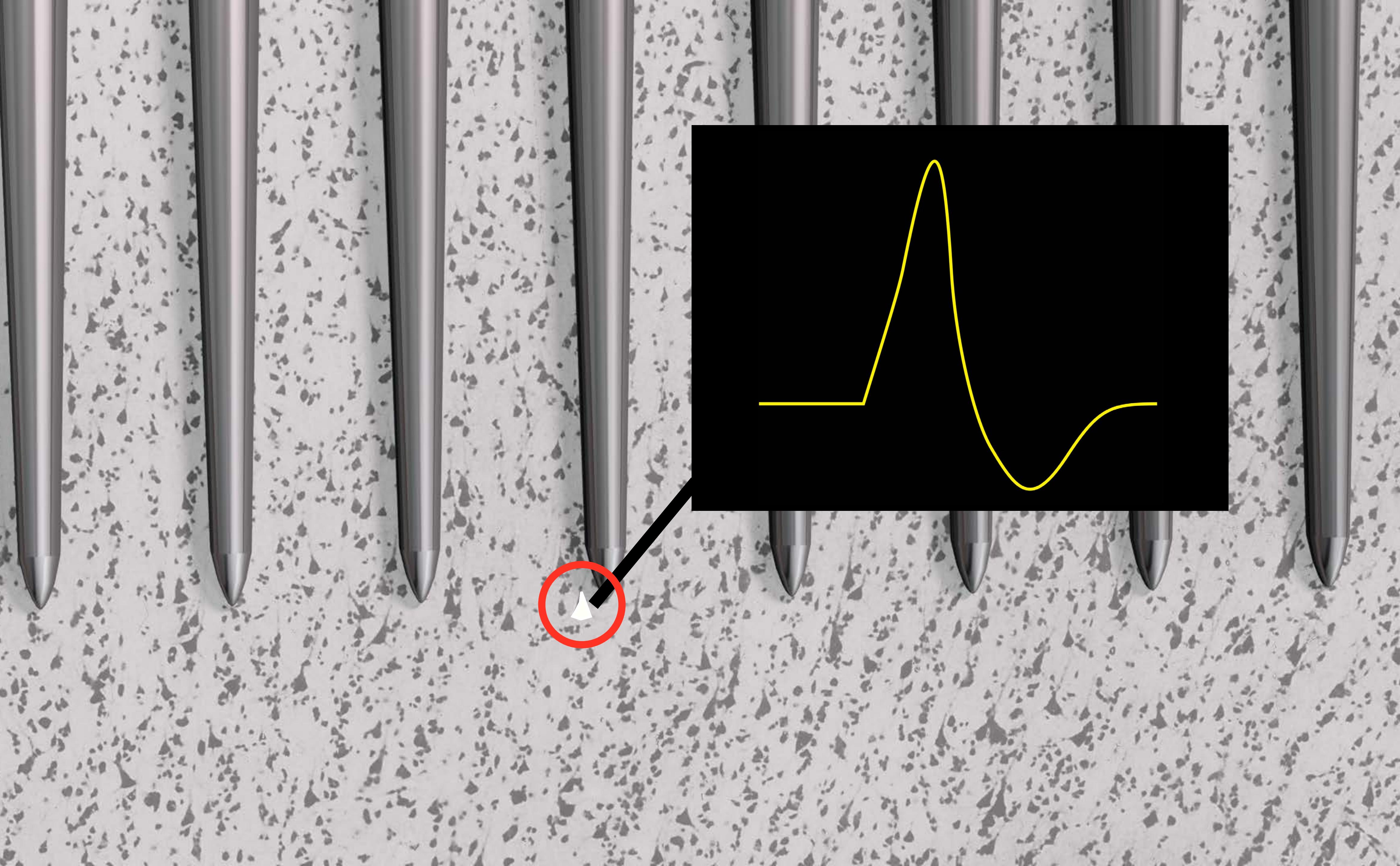
# Techniques for measuring neural activity

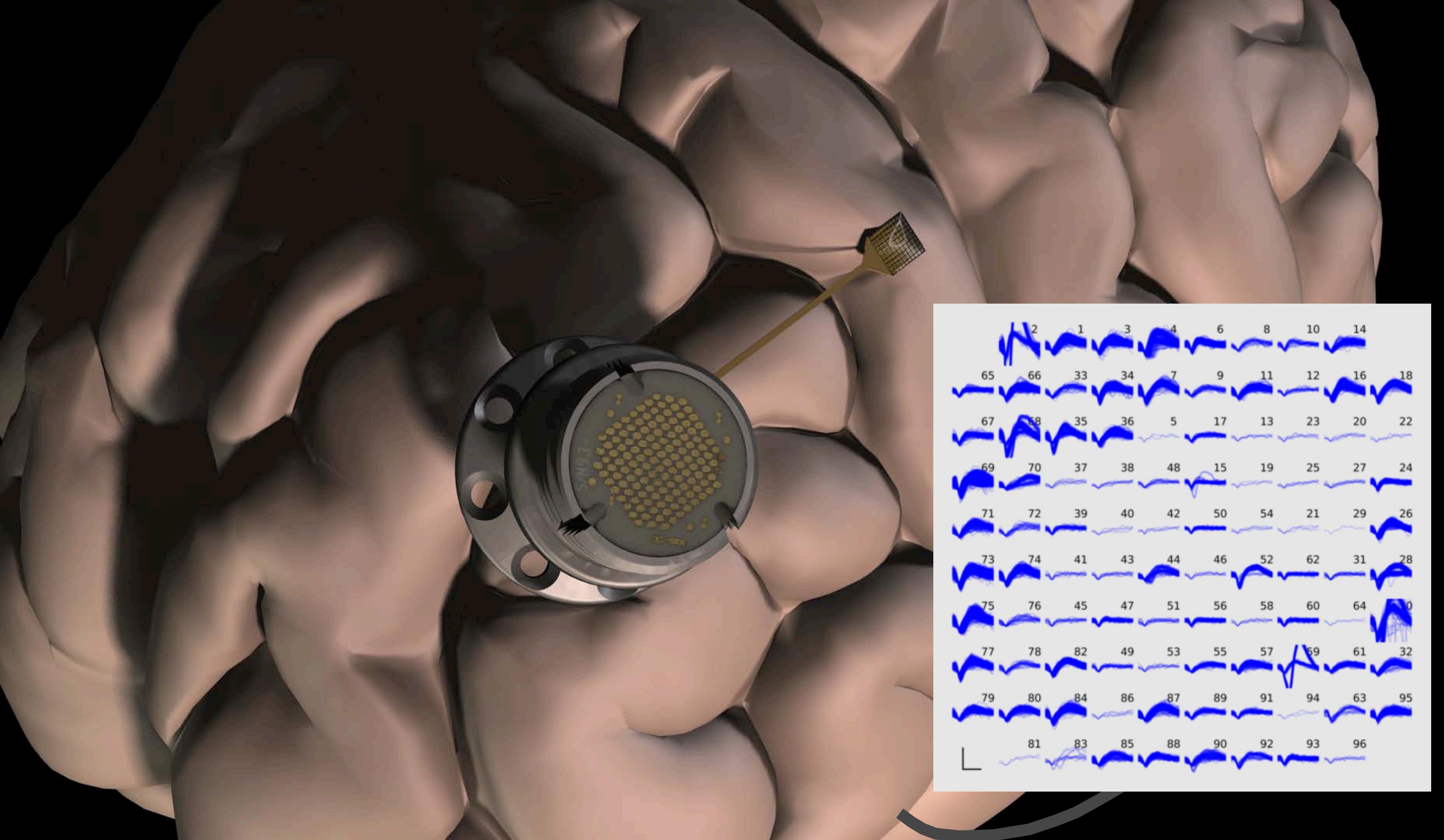


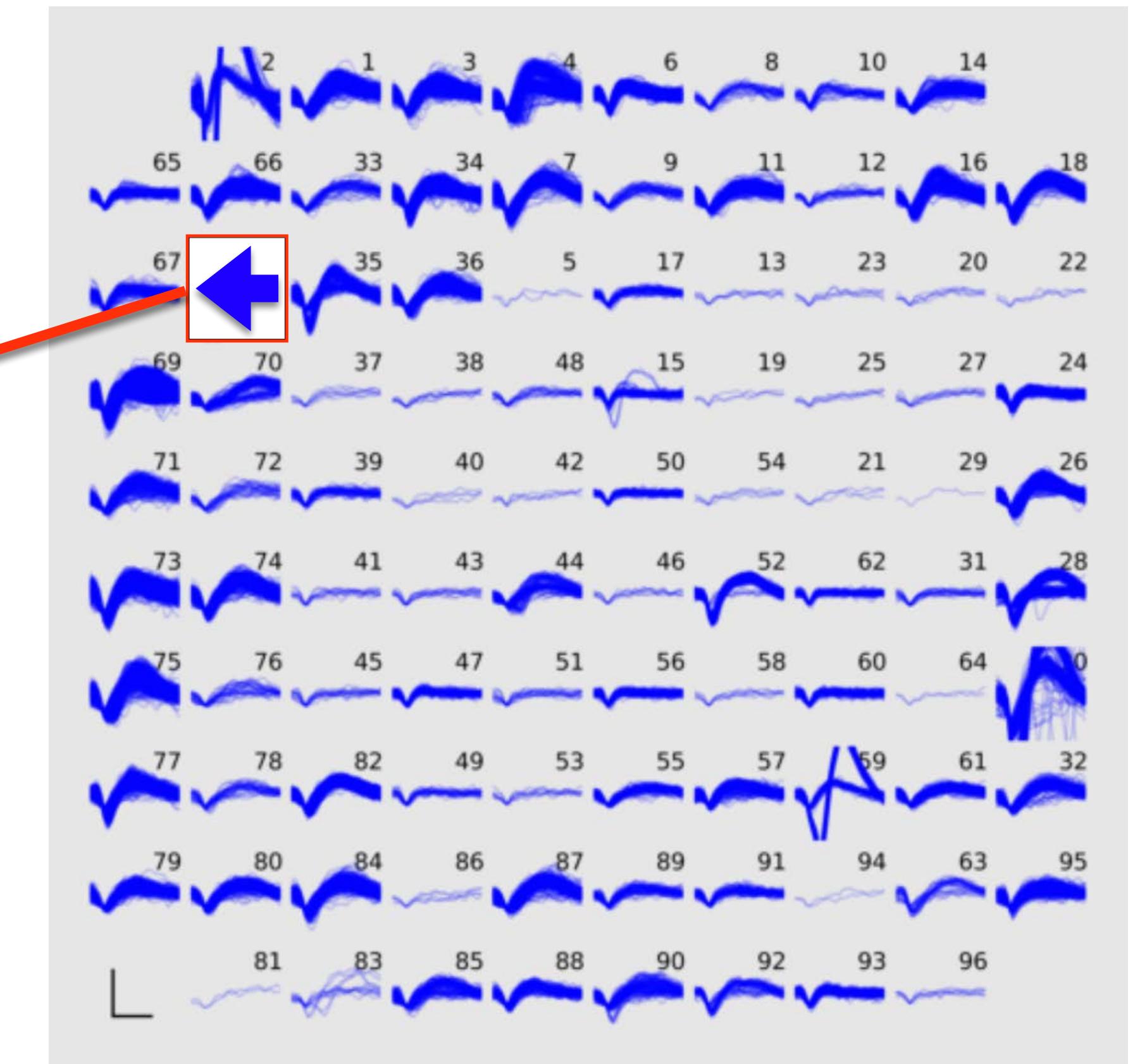
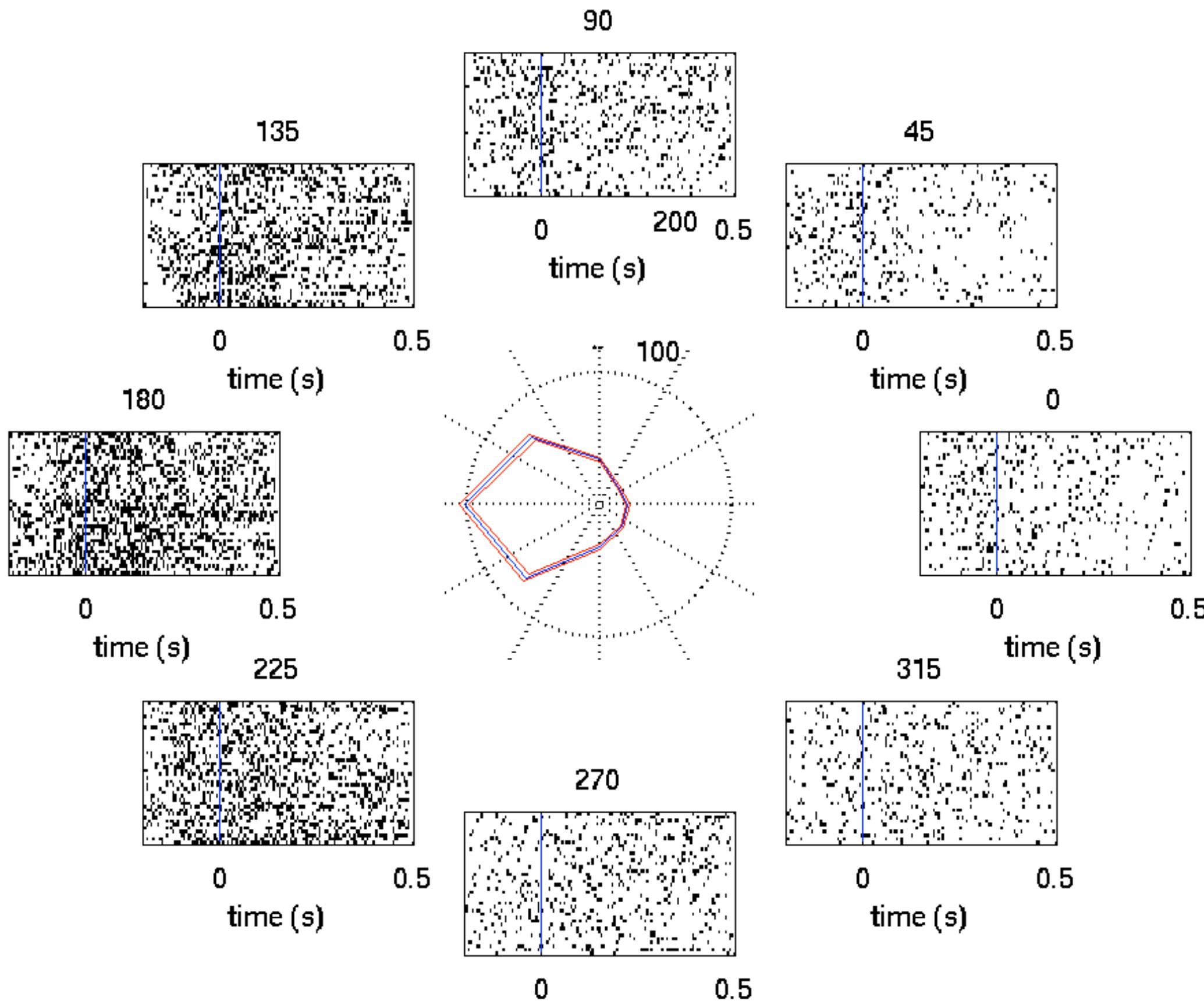


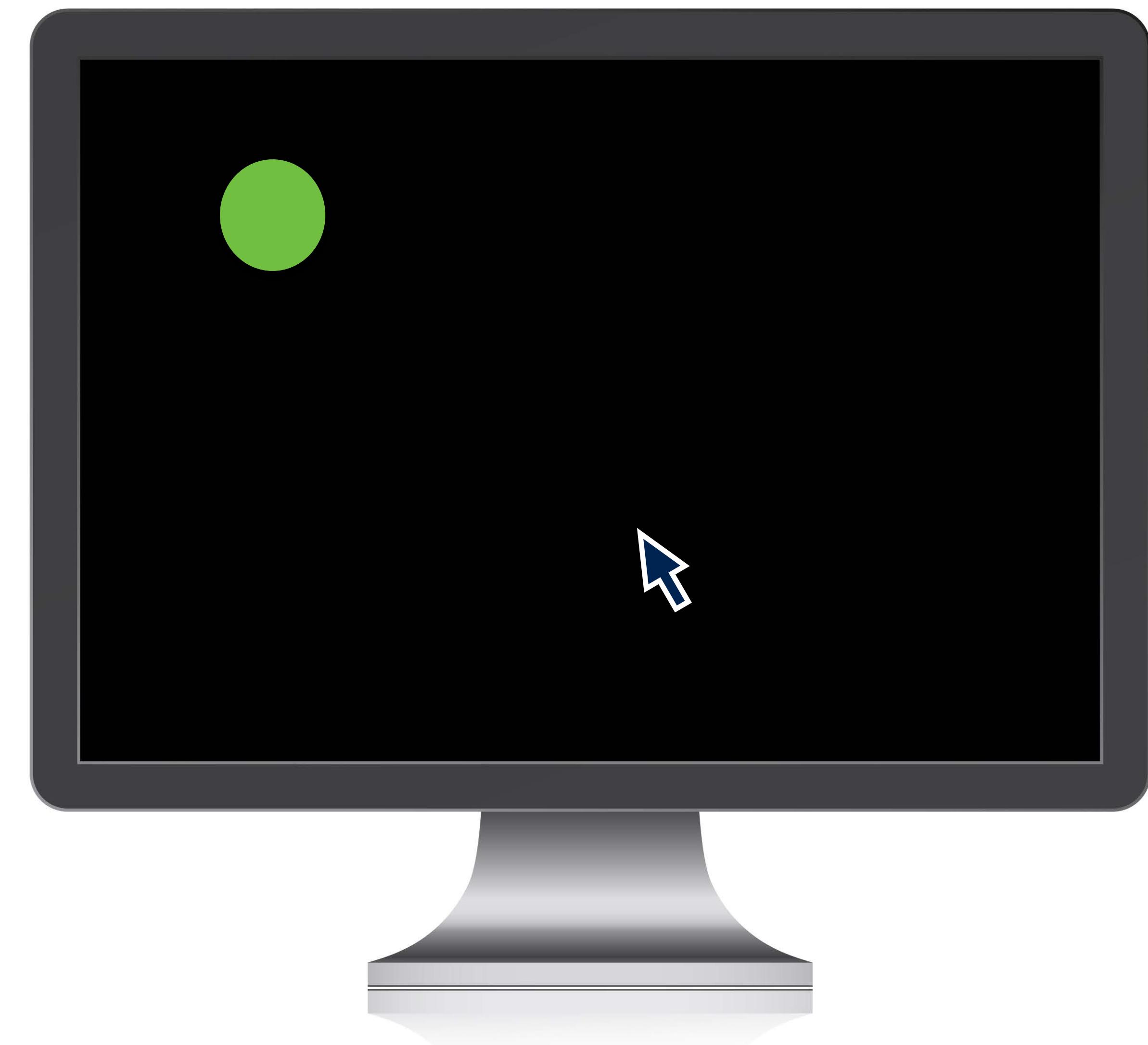
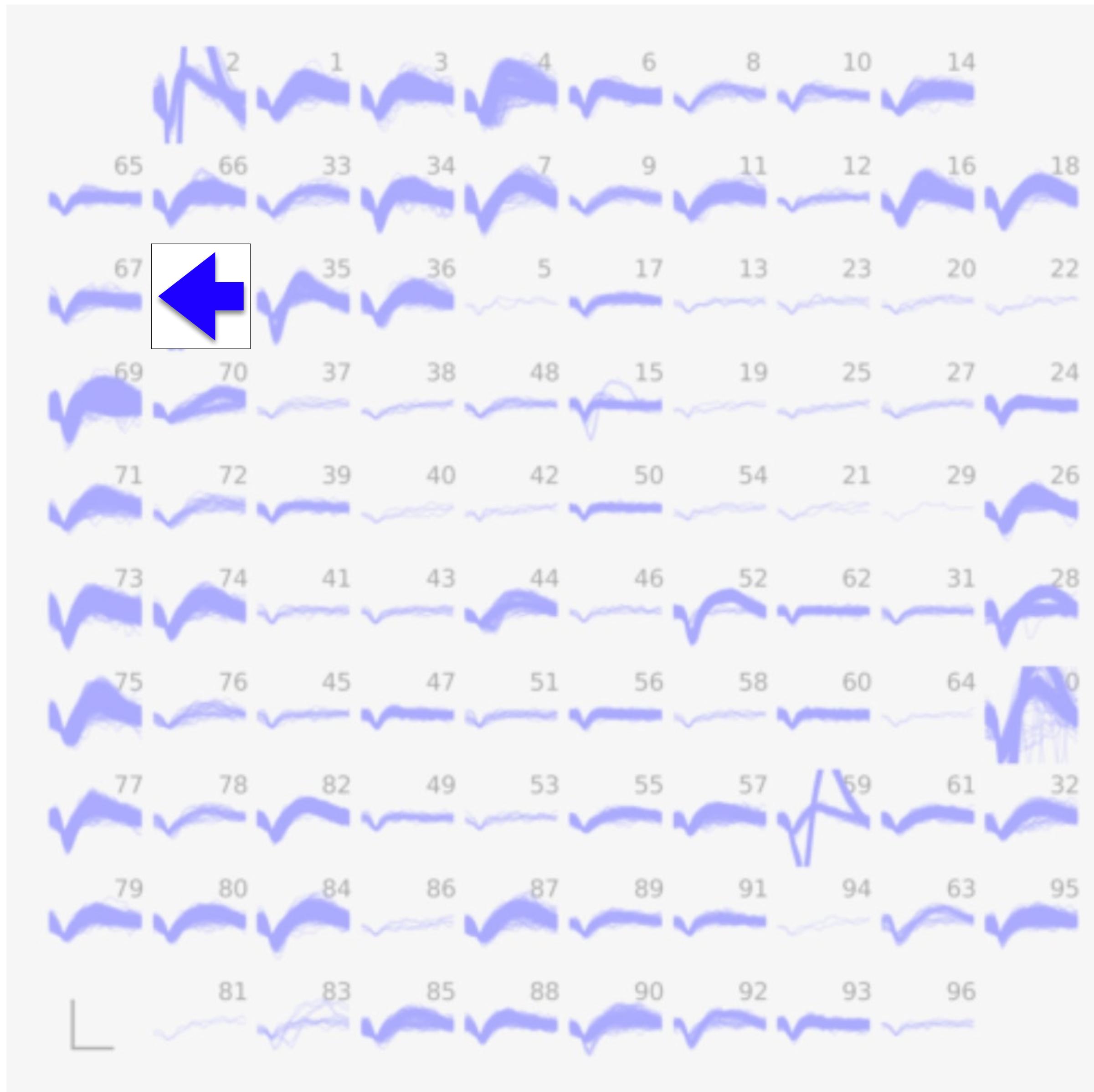


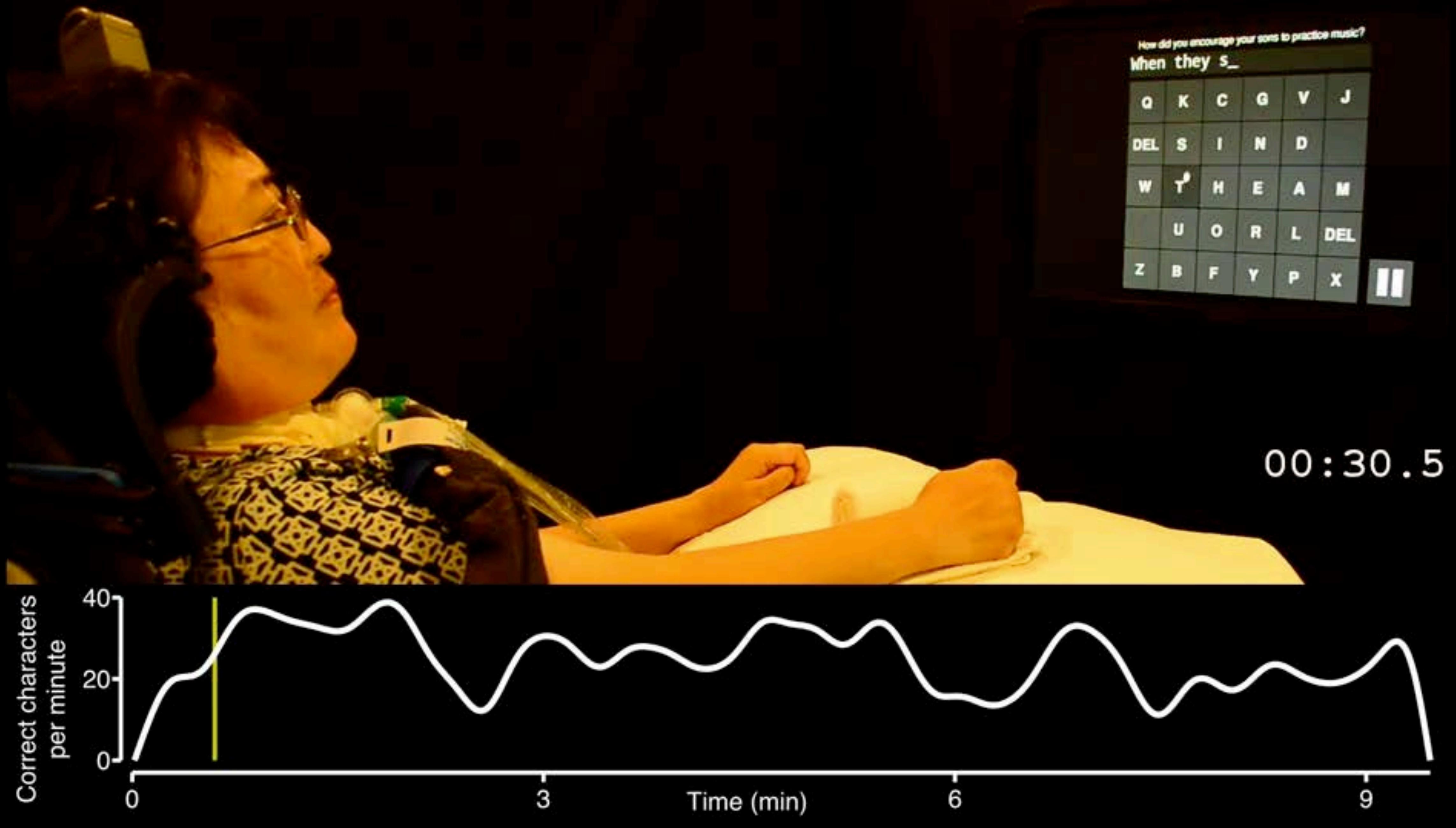












Pandarinath\*, Nuyujukian\*, et al. (2017) eLife

# Beyond 2D control

Control robotic arms



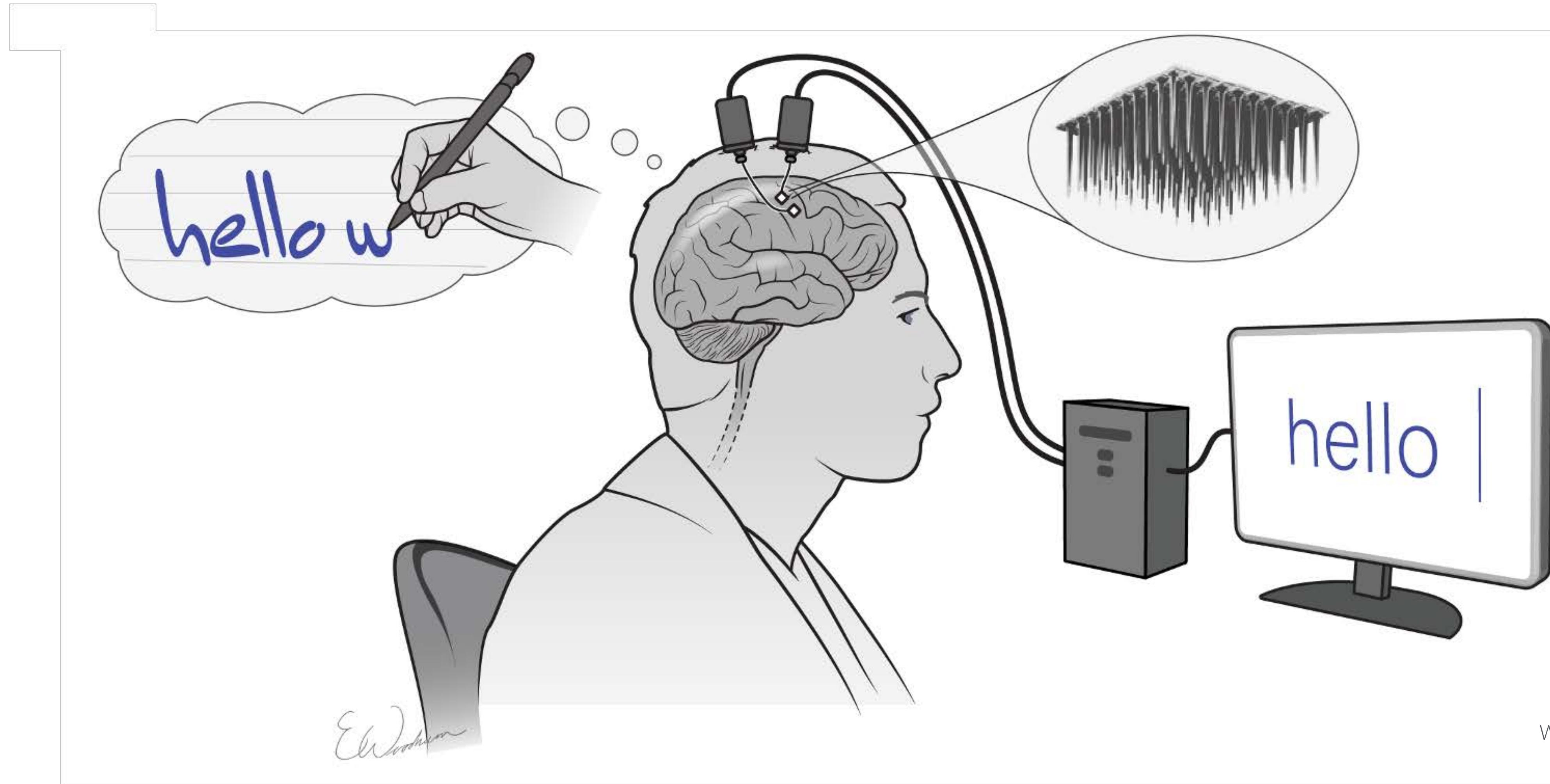
# Beyond 2D control

Restore walking



# Beyond 2D control

Communication through handwriting

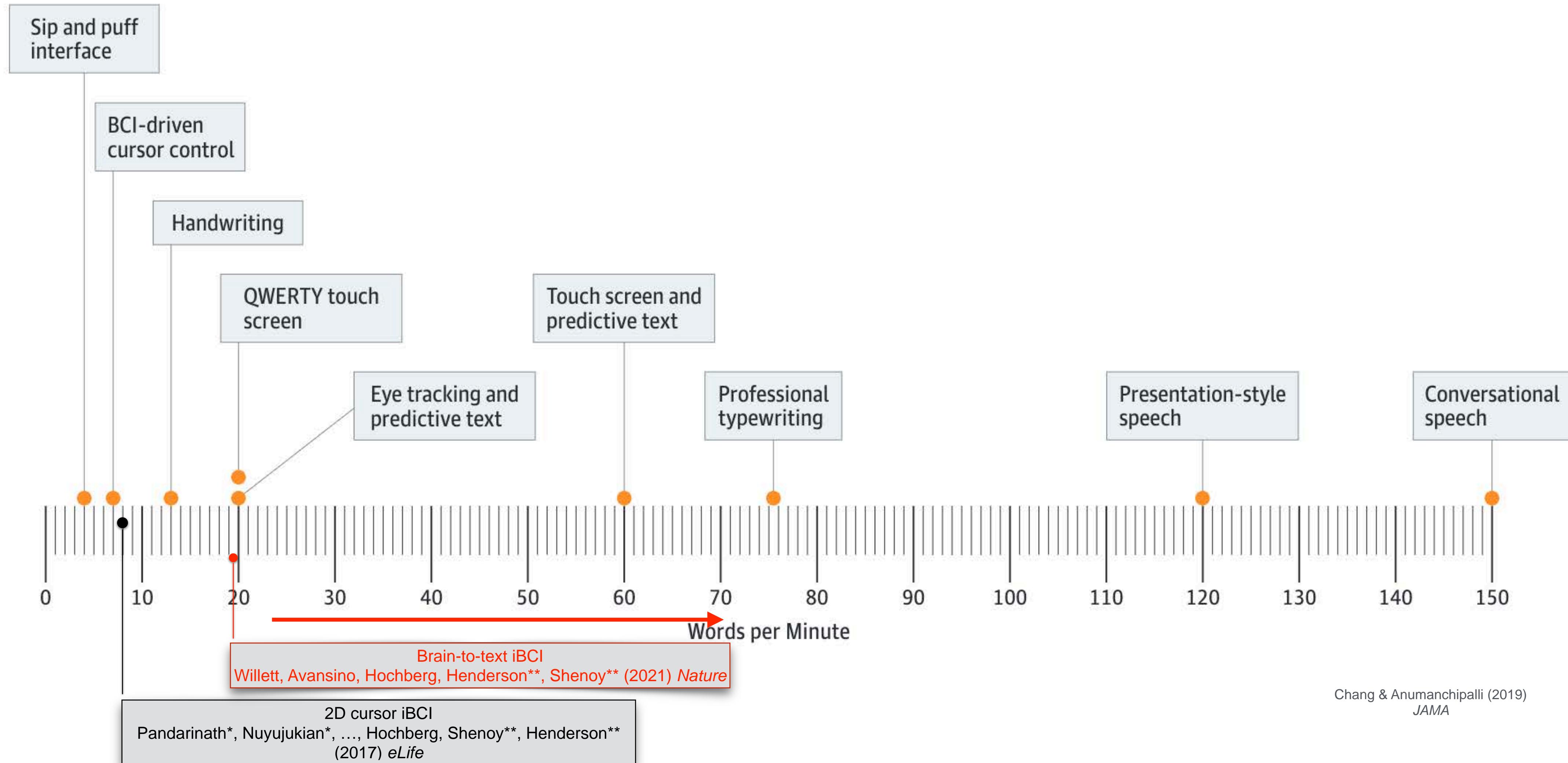


# 90 char/min with 95% accuracy

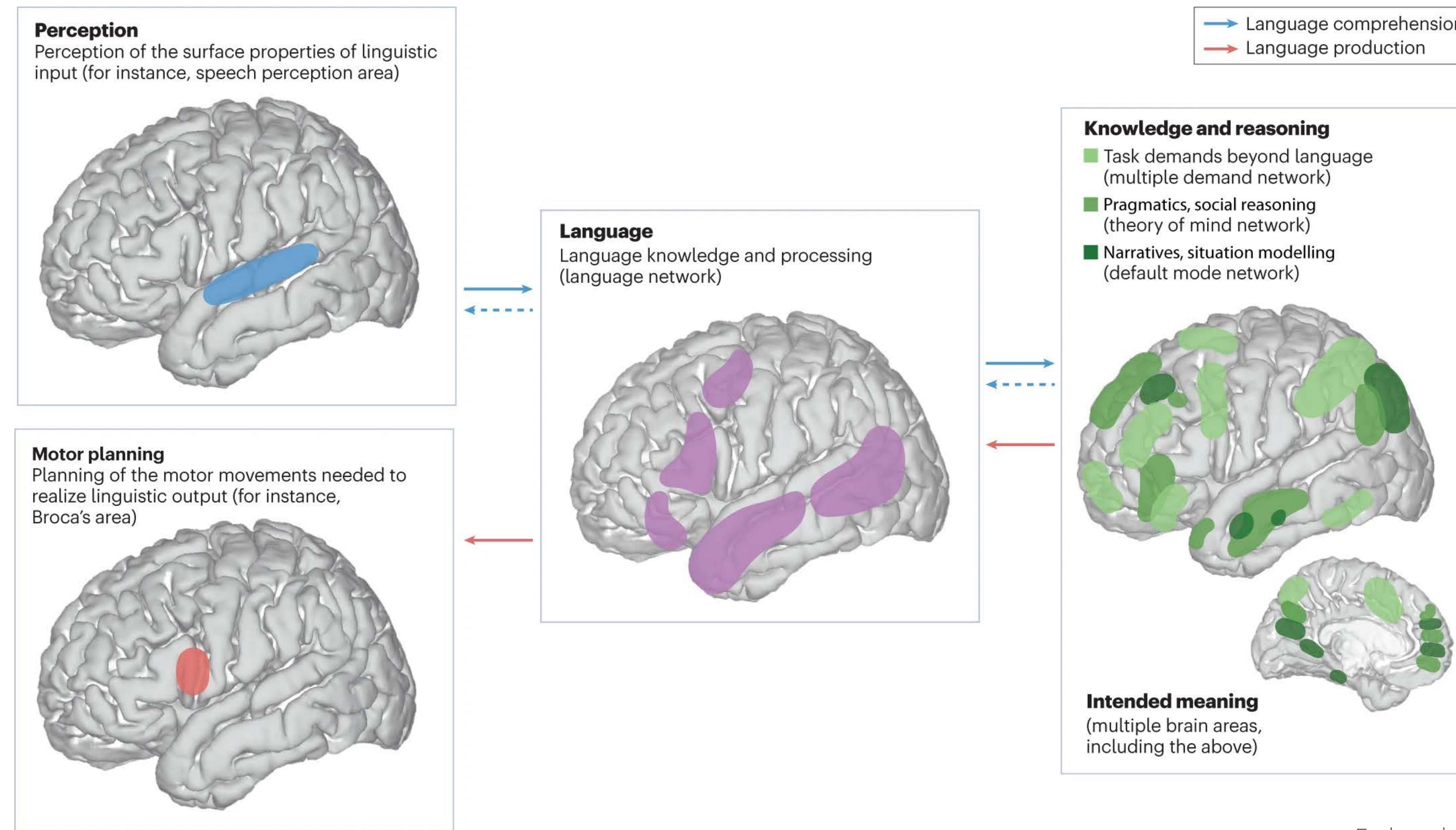
Communication through handwriting



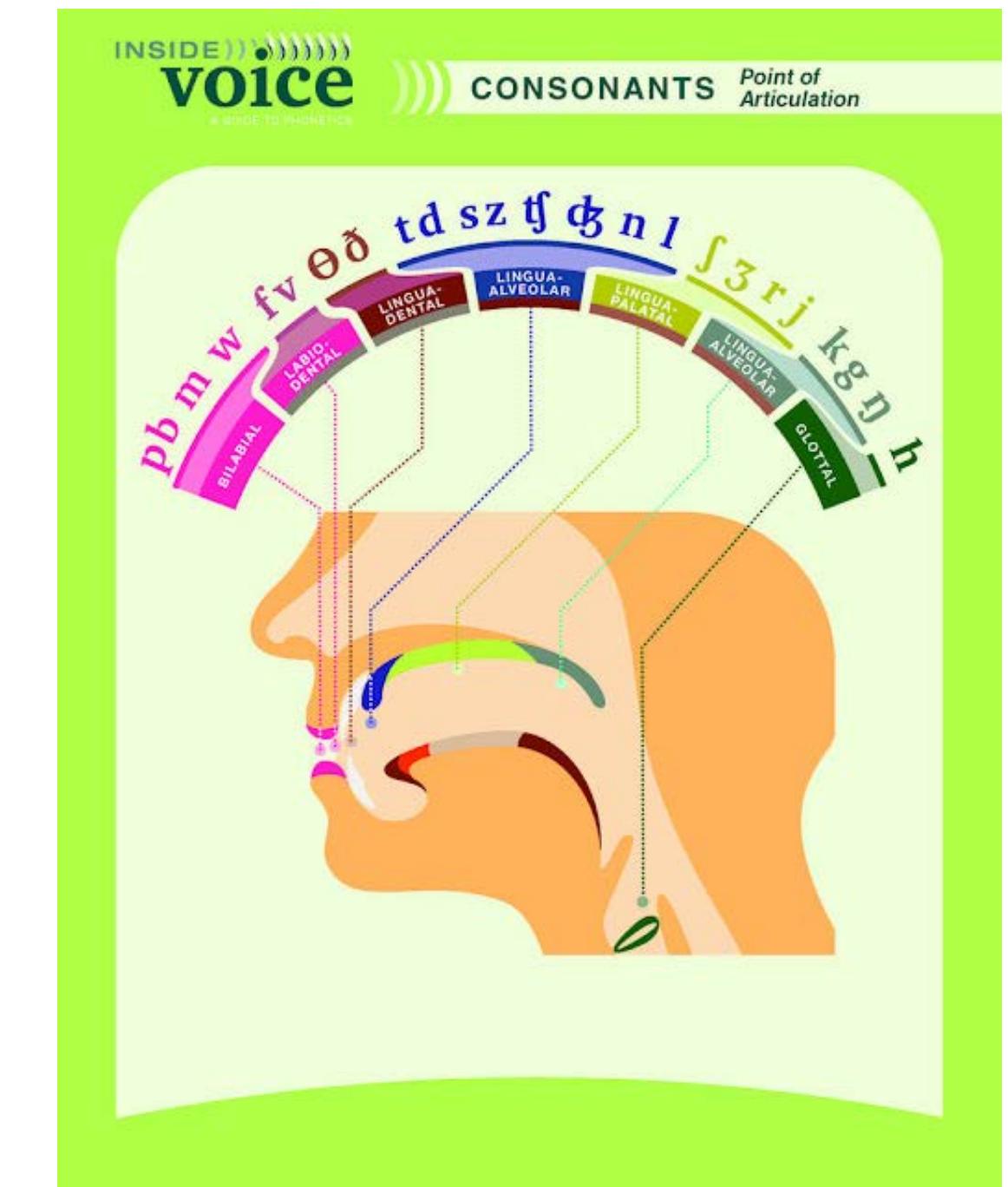
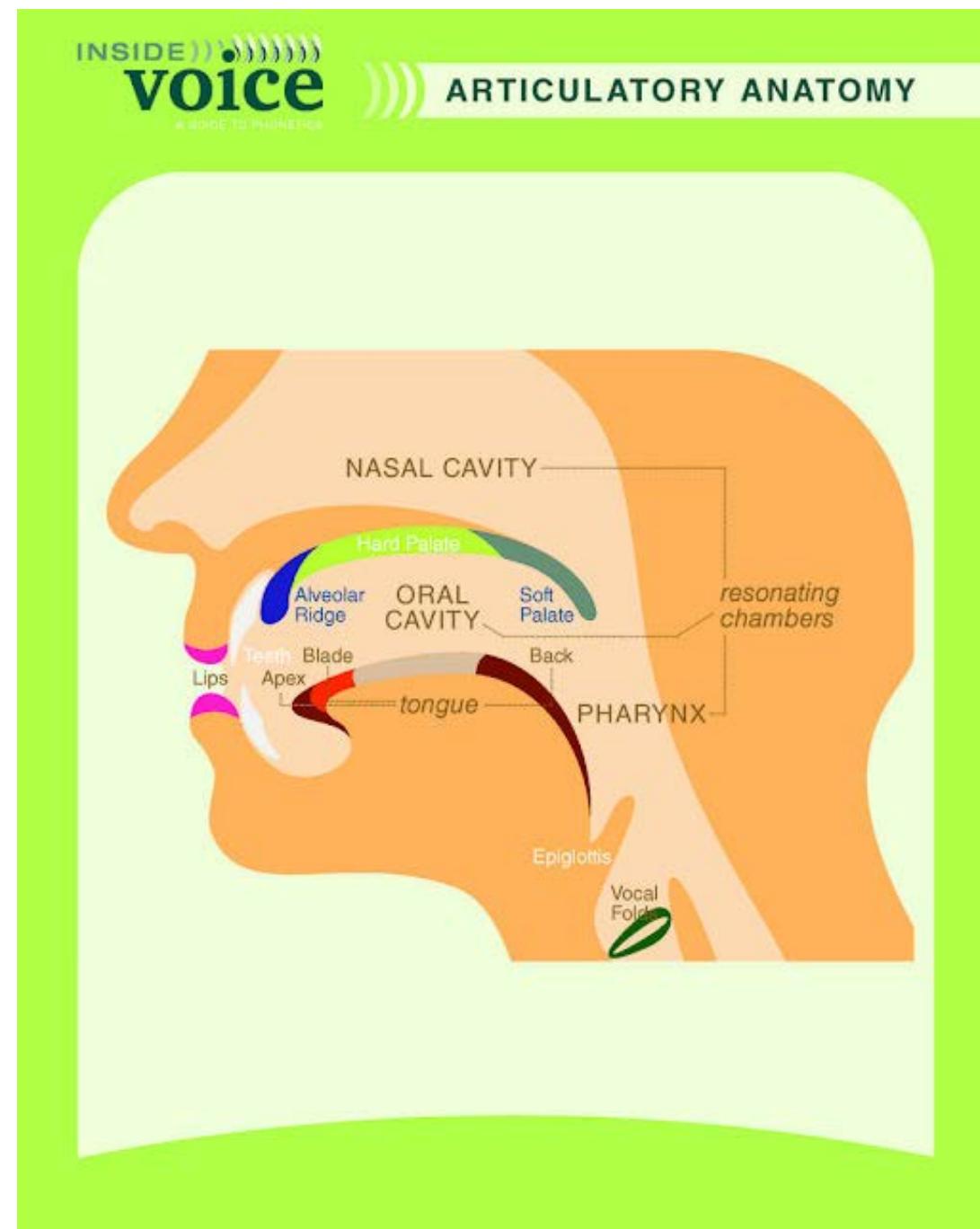
# Restore natural speech?



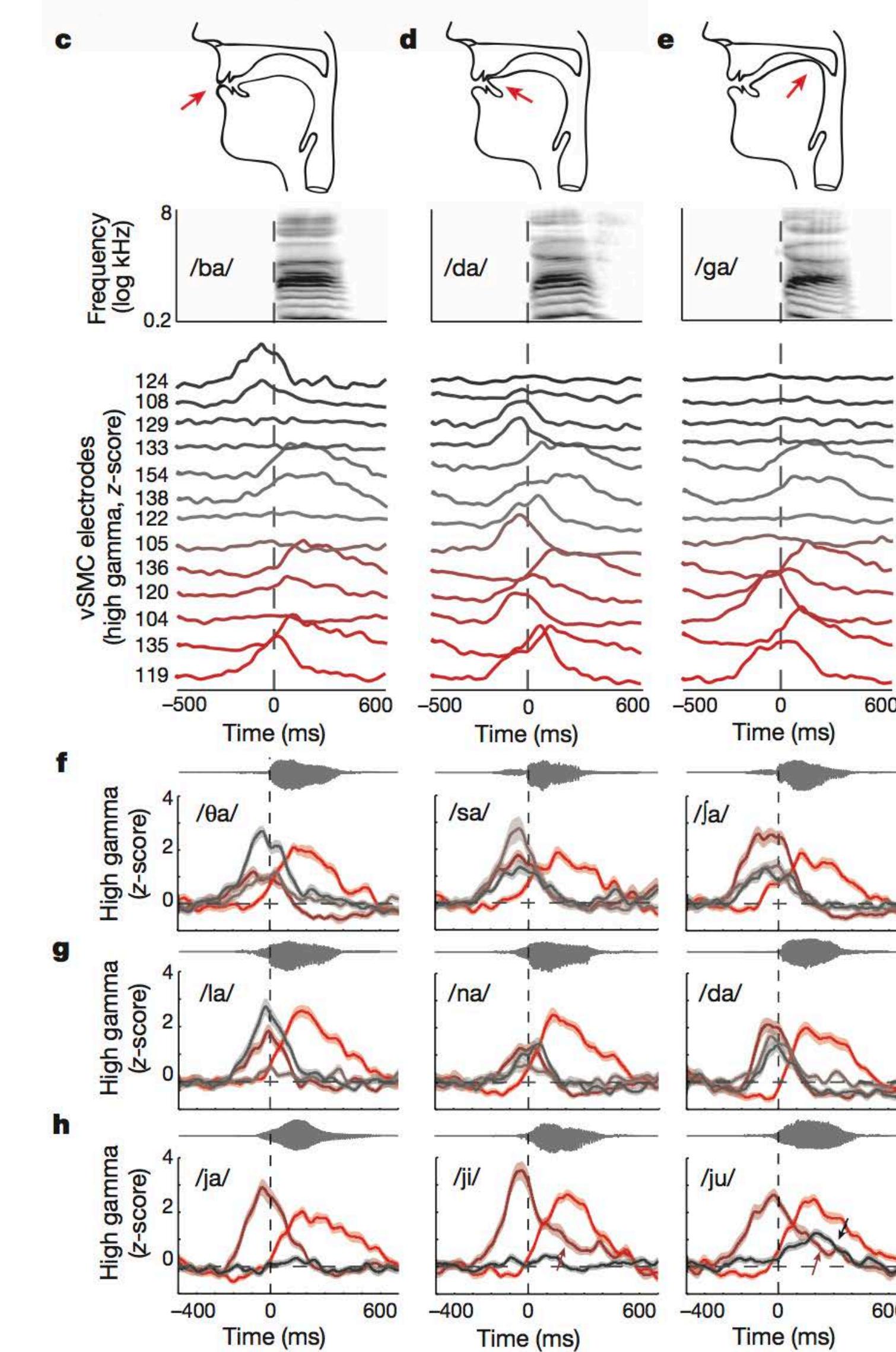
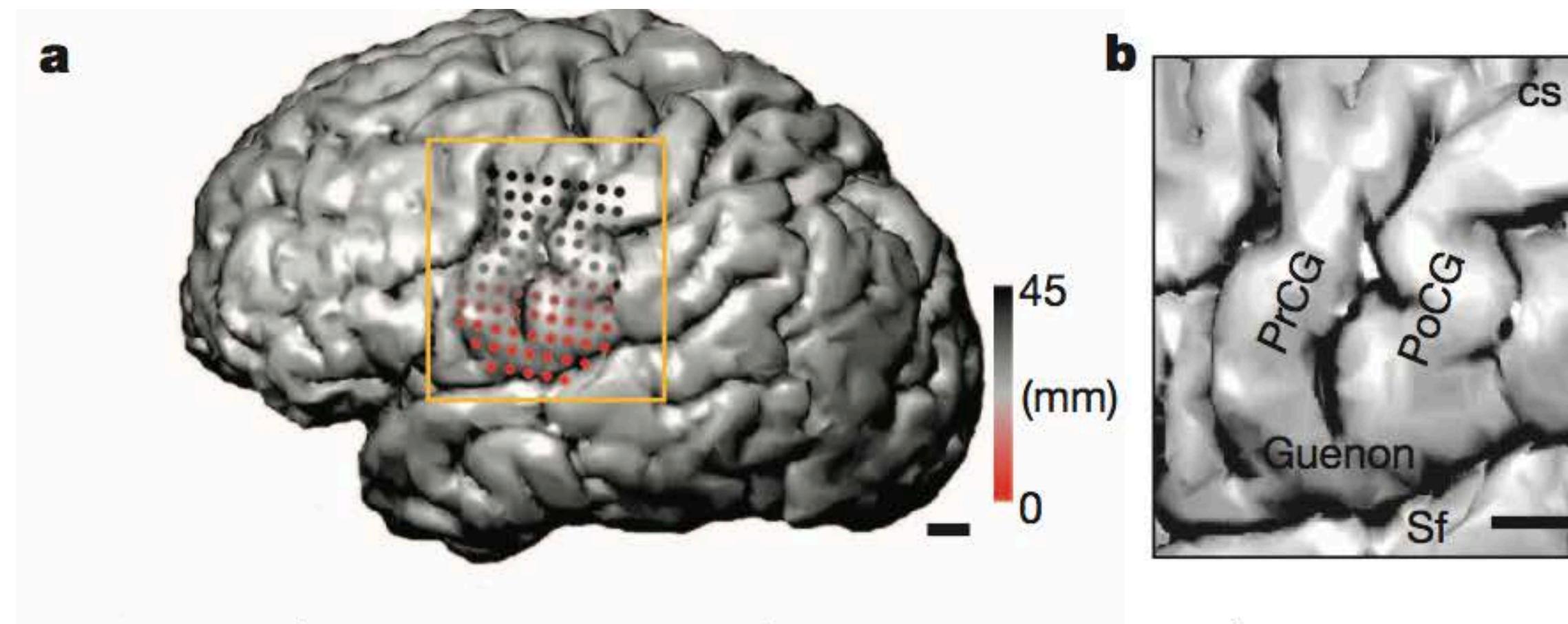
# Language processing in the brain



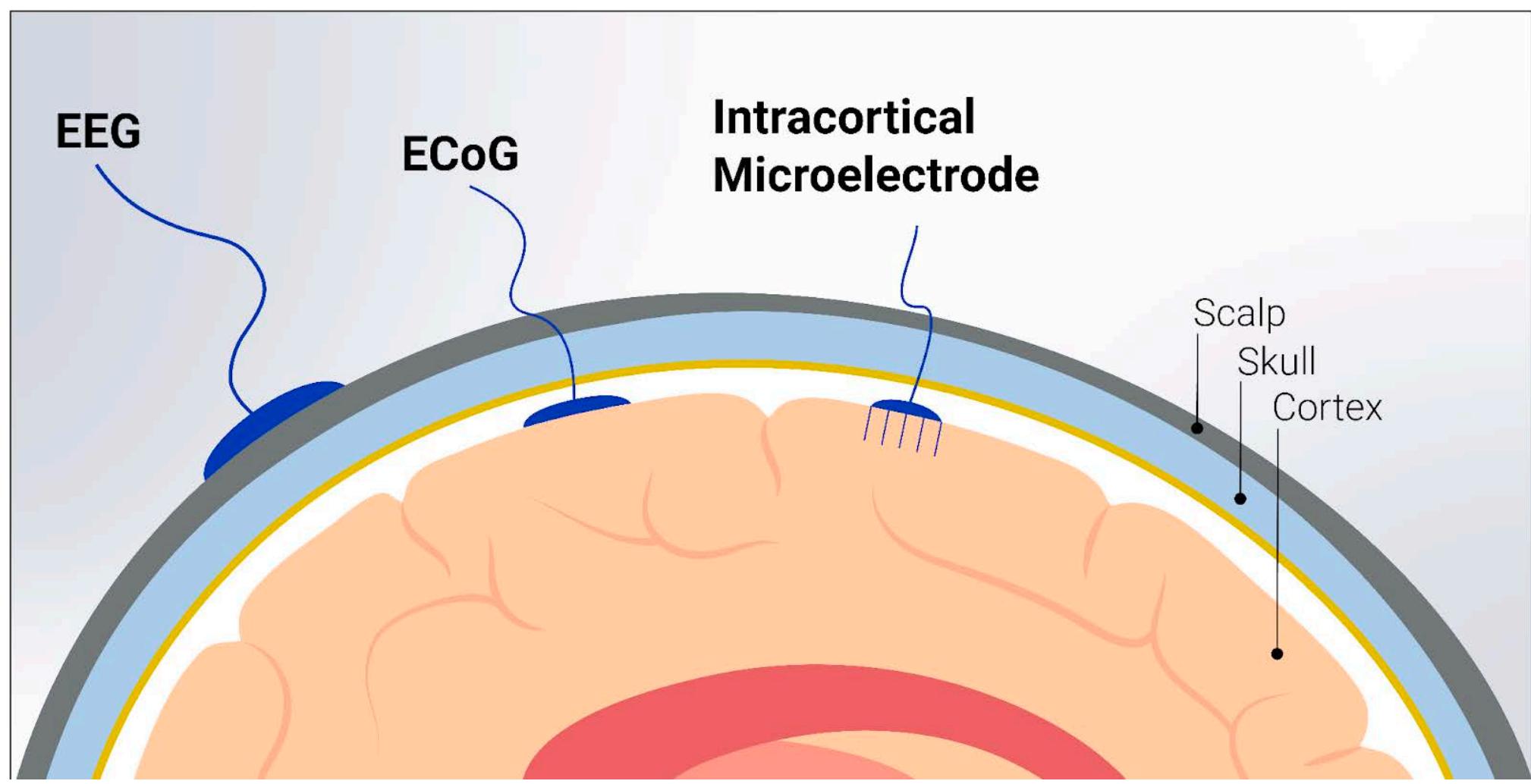
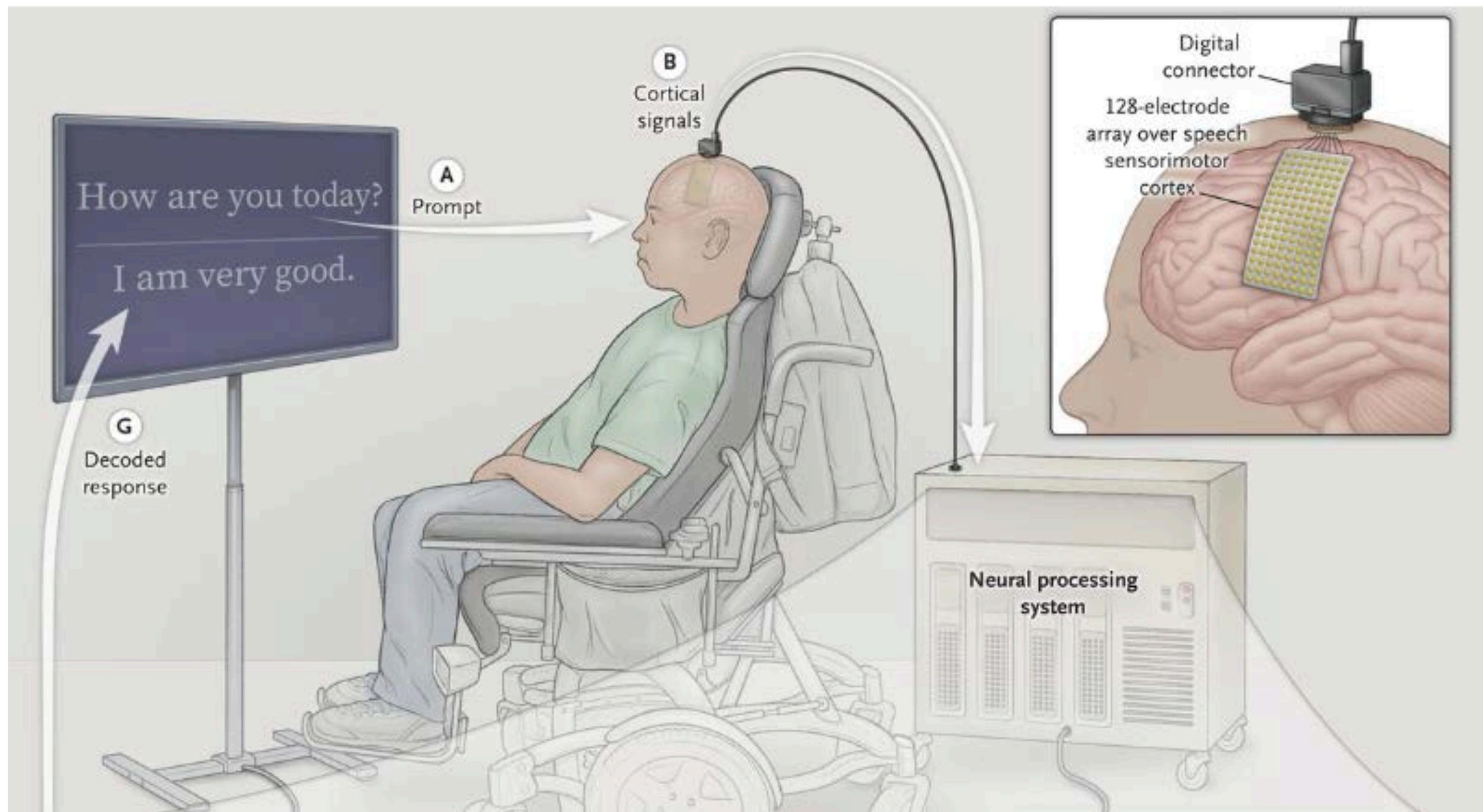
# How do we produce speech sound



# Motor cortex encodes articulatory and phonemic information



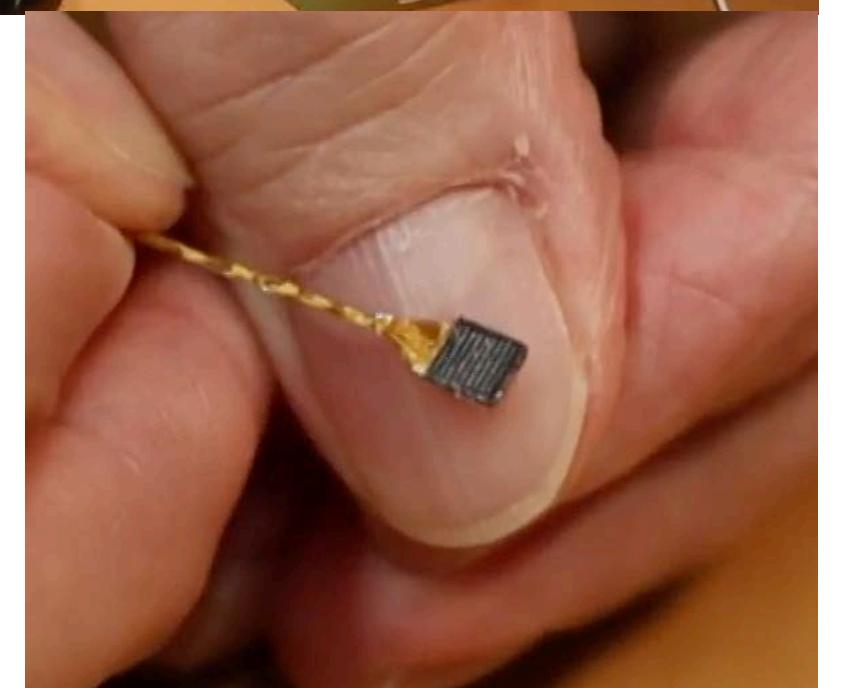
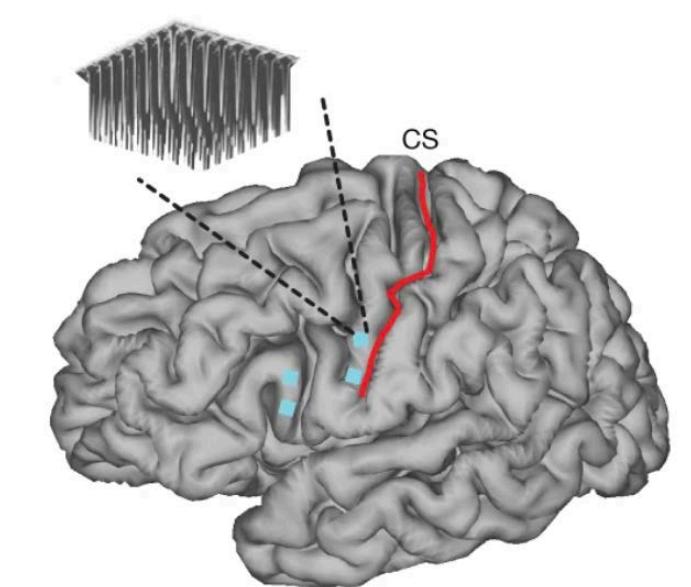
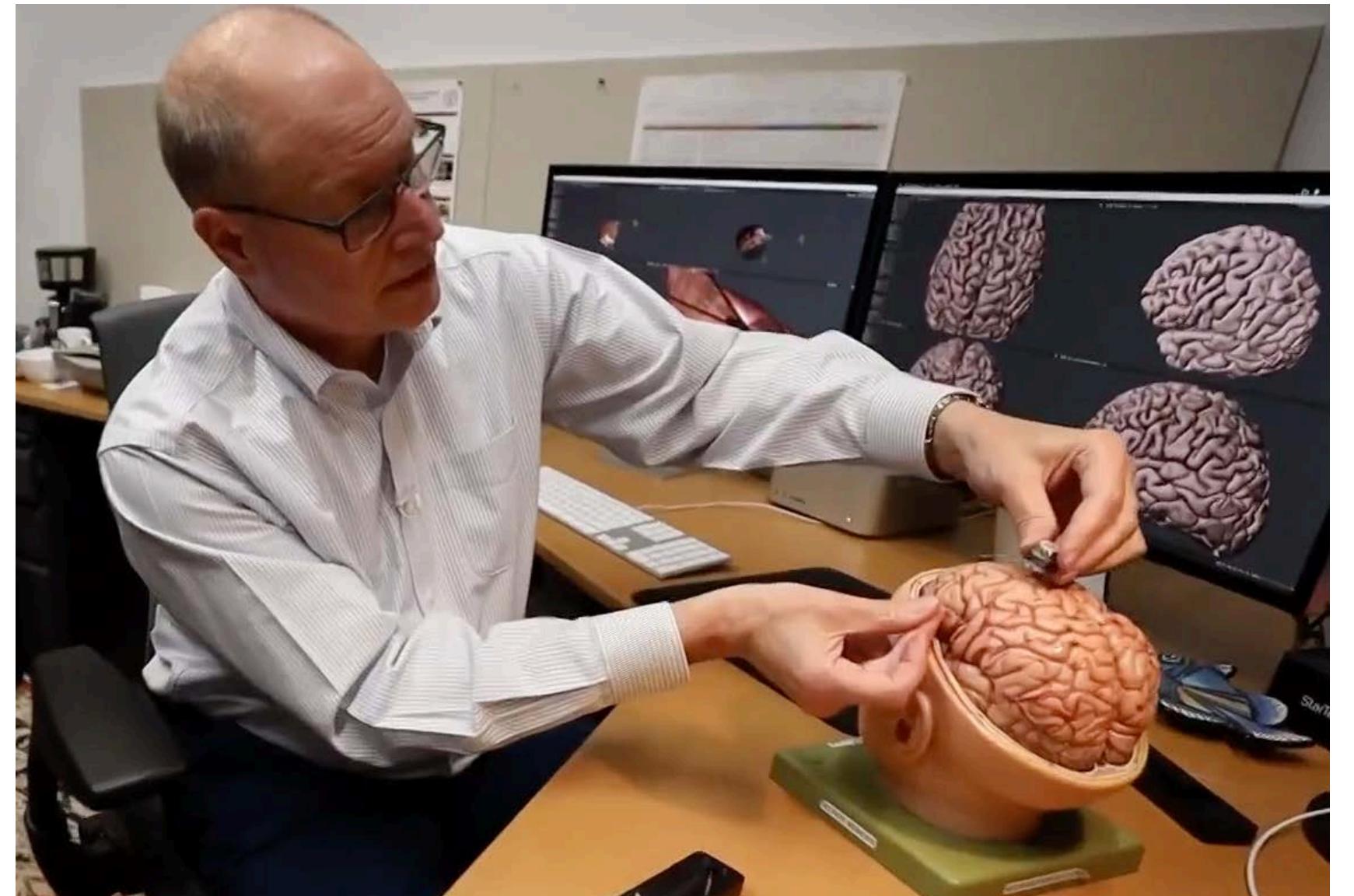
# Small vocabulary speech BCI with ECoG



# A High-Performance Speech Neuroprosthesis

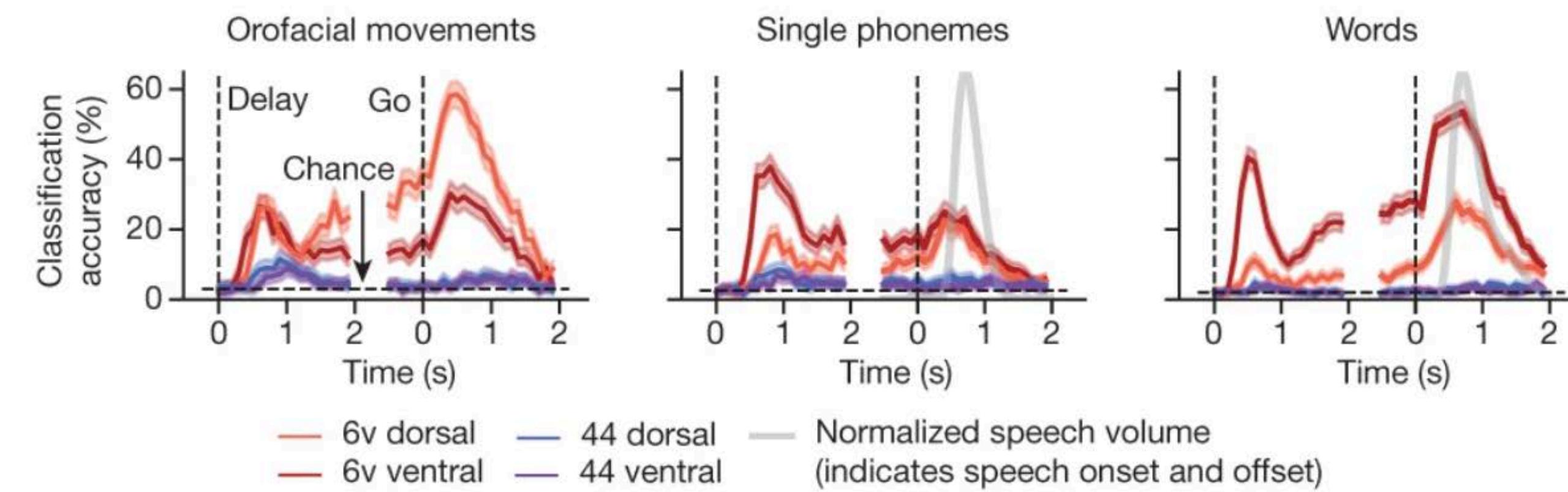
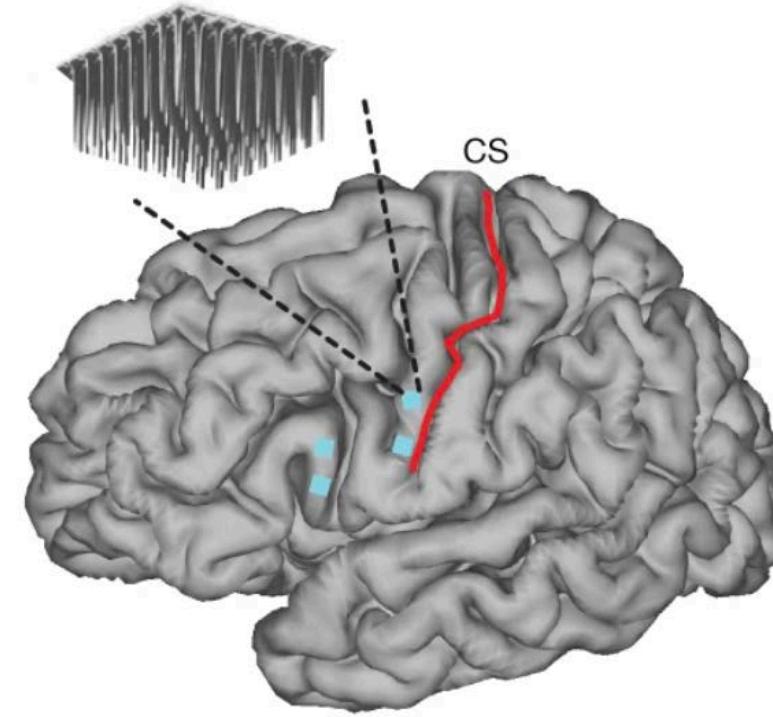
# Implanting microelectrode arrays into BrainGate2 clinical trial participant T12

- T12 has bulbar-onset Amyotrophic Lateral Sclerosis (ALS)
- She retains some limited orofacial movement and an ability to vocalize, but is unable to produce intelligible speech.
- Four 64-channel Utah arrays
  - Two in area 6v (ventral motor cortex)
  - Two in area 44 (part of Broca's area)

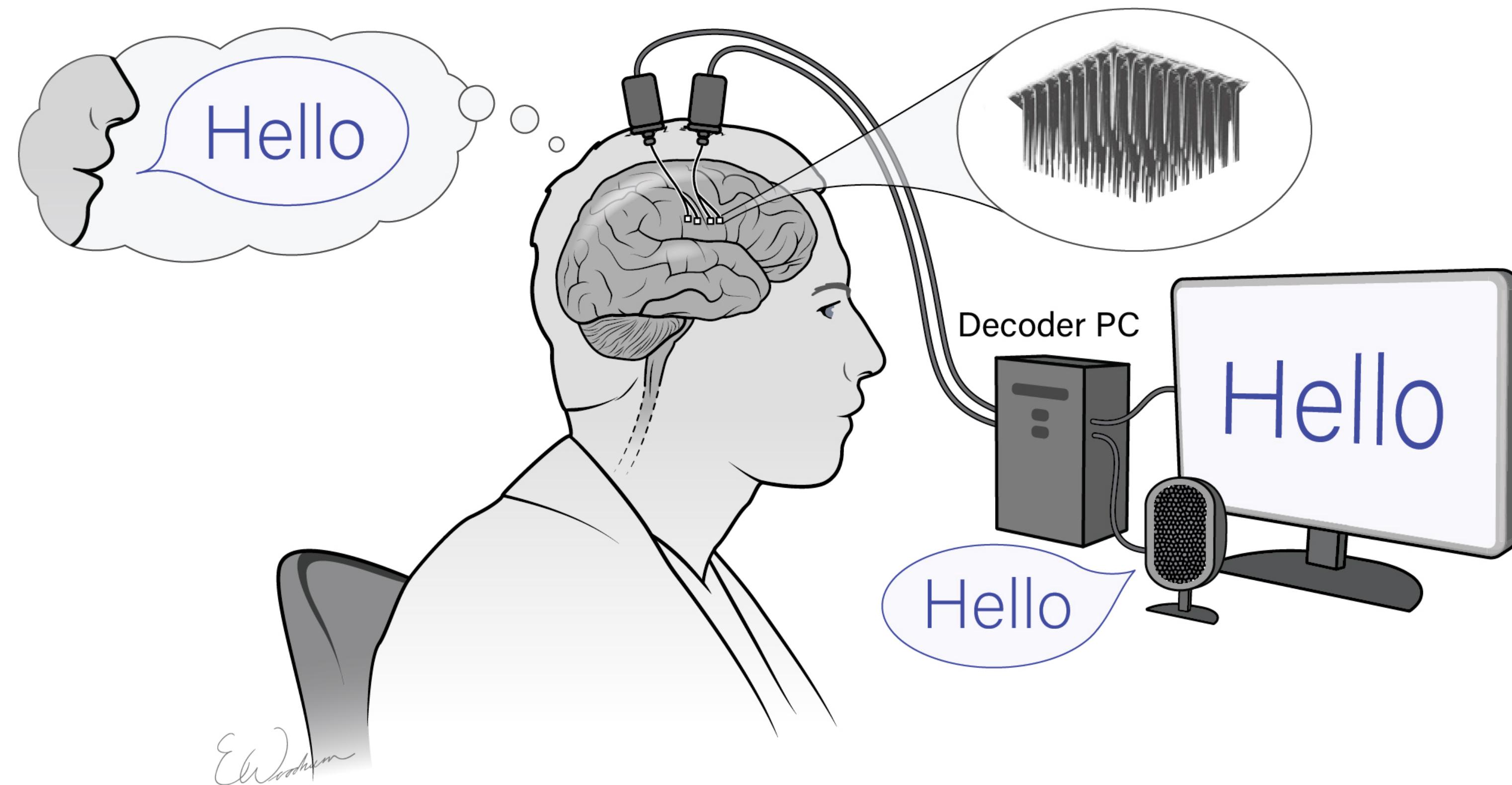


Willett, Kunz, Fan, et al. 2023

# Neural representation of orofacial movements and speech



# Real-time brain-to-text BCI



I don't want to call her a baby sitter.

Block: 17  
Trial: 32

What are you proud of?



Block: 14  
Trial: 15

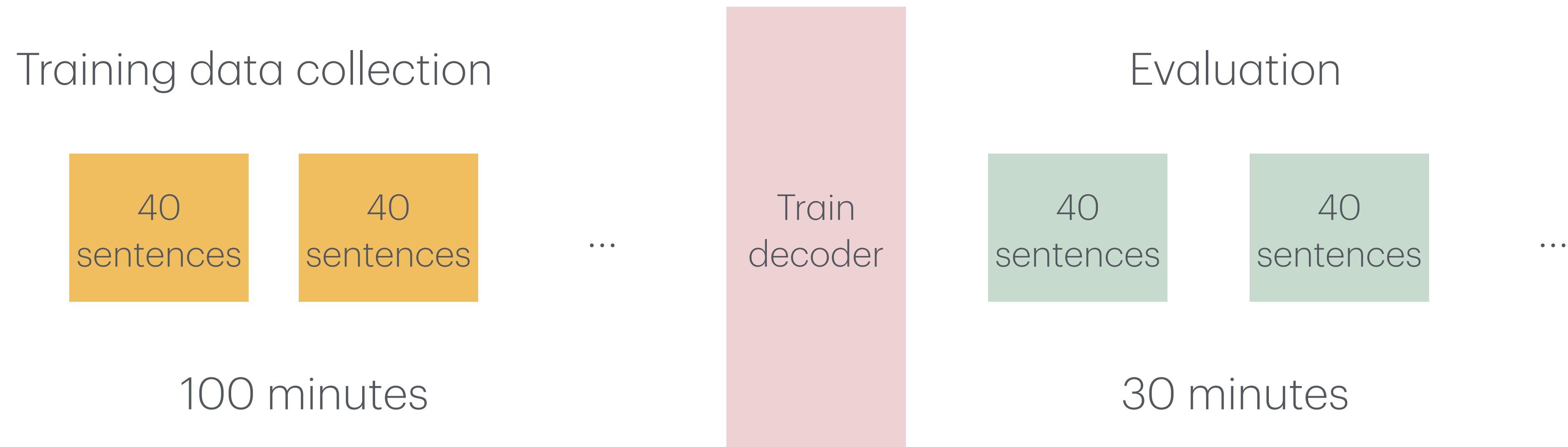
I do not have much to compare it to.

Block: 17  
Trial: 29

# Data collection



# Data collection



Training and evaluation sentences are randomly selected from the Switchboard corpus of telephone conversation (~1,0000 sentences)

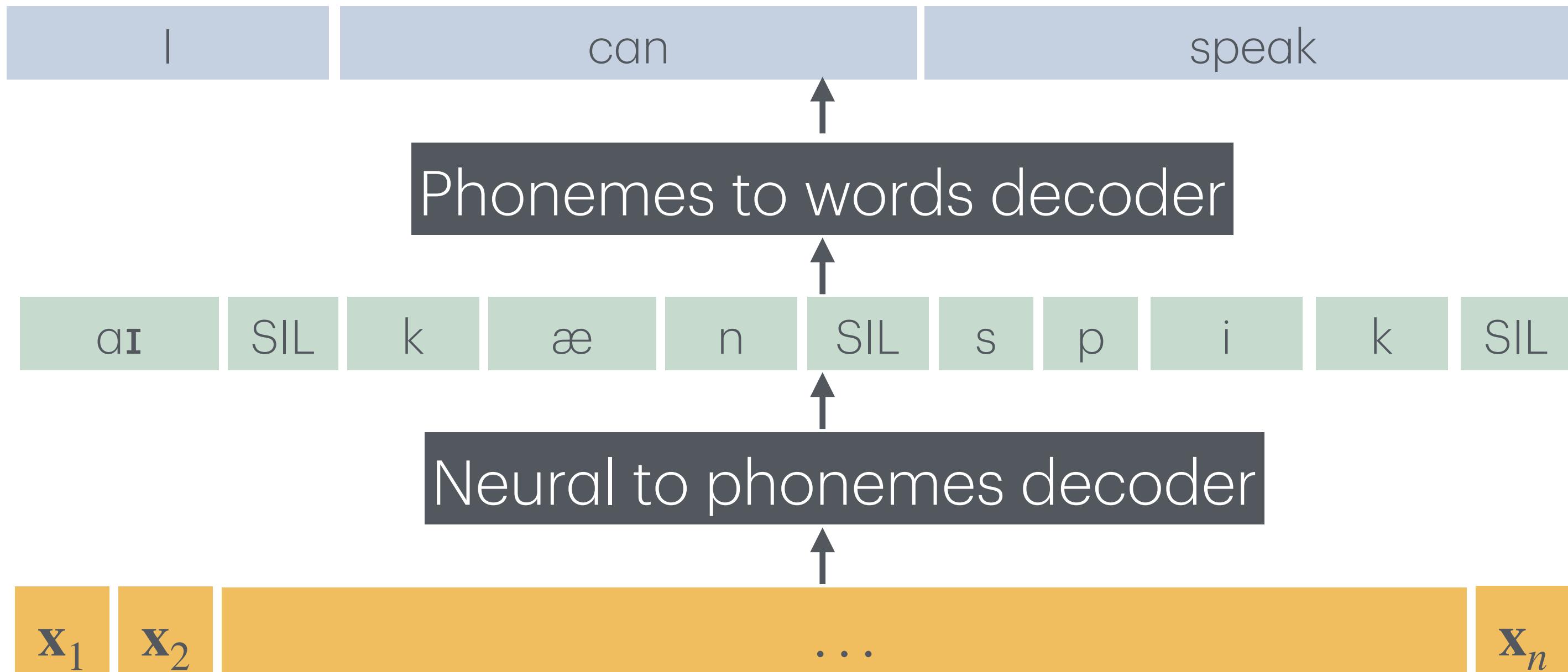
# Problem definition

- Neural feature inputs:  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \quad \mathbf{x}_i \in \mathbb{R}^{d \times 1}$
- Words outputs:  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\} \quad \mathbf{y}_i \in \mathbb{R}^{V \times 1}$



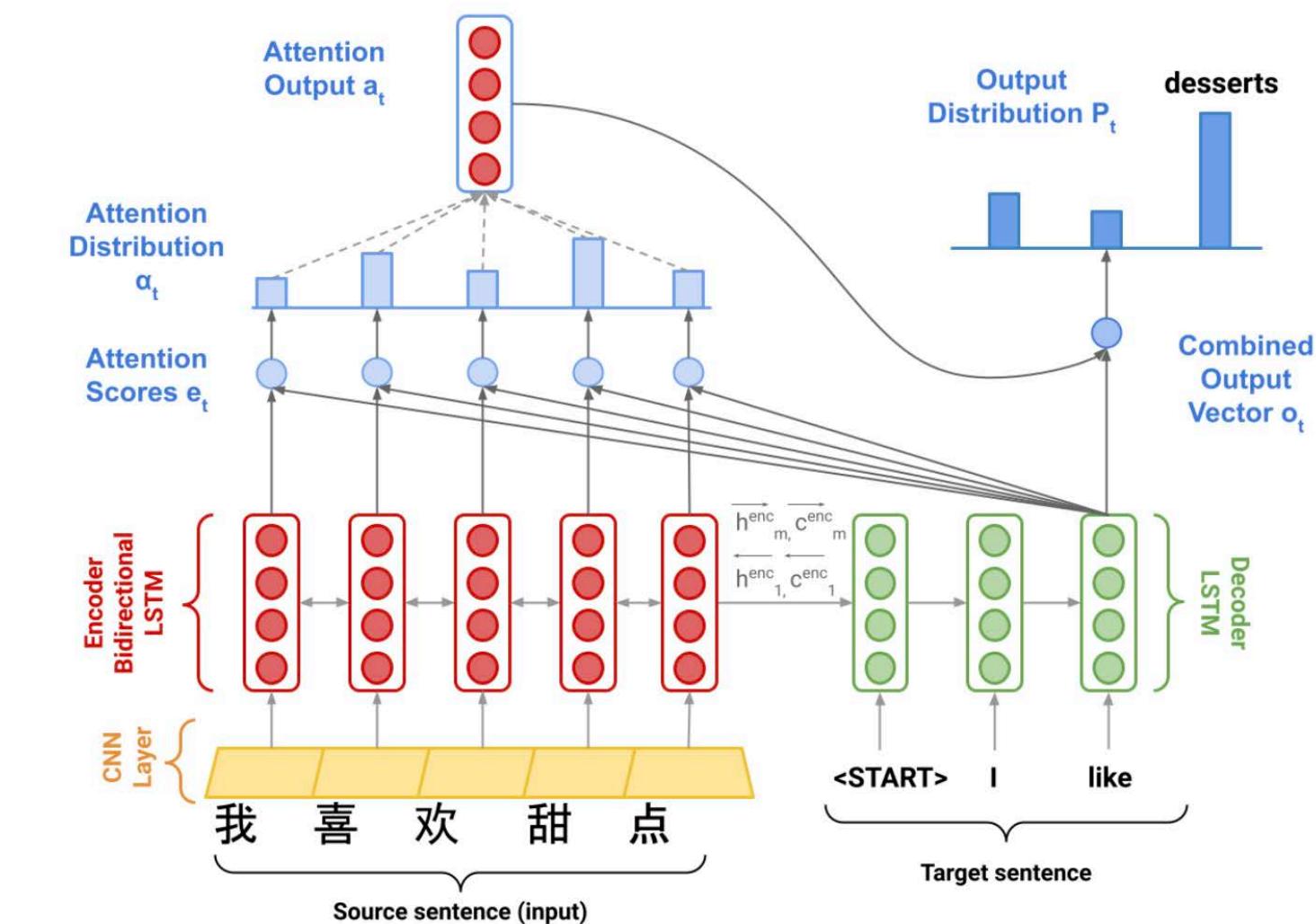
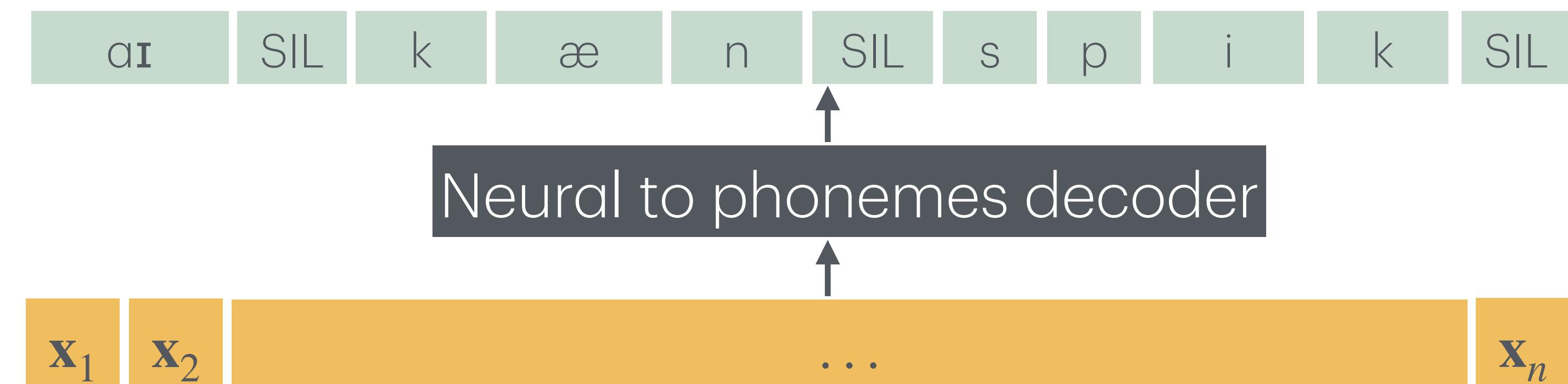
# Problem definition

- Neural feature inputs:  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \quad \mathbf{x}_i \in \mathbb{R}^{d \times 1}$
- Words outputs:  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\} \quad \mathbf{y}_i \in \mathbb{R}^{V \times 1}$



# Neural to phonemes decoder

- A Seq2Seq problem
  - Encoder-decoder models allow arbitrary alignment between inputs and outputs.
  - Neural to phonemes decoding only needs monotonic alignment.



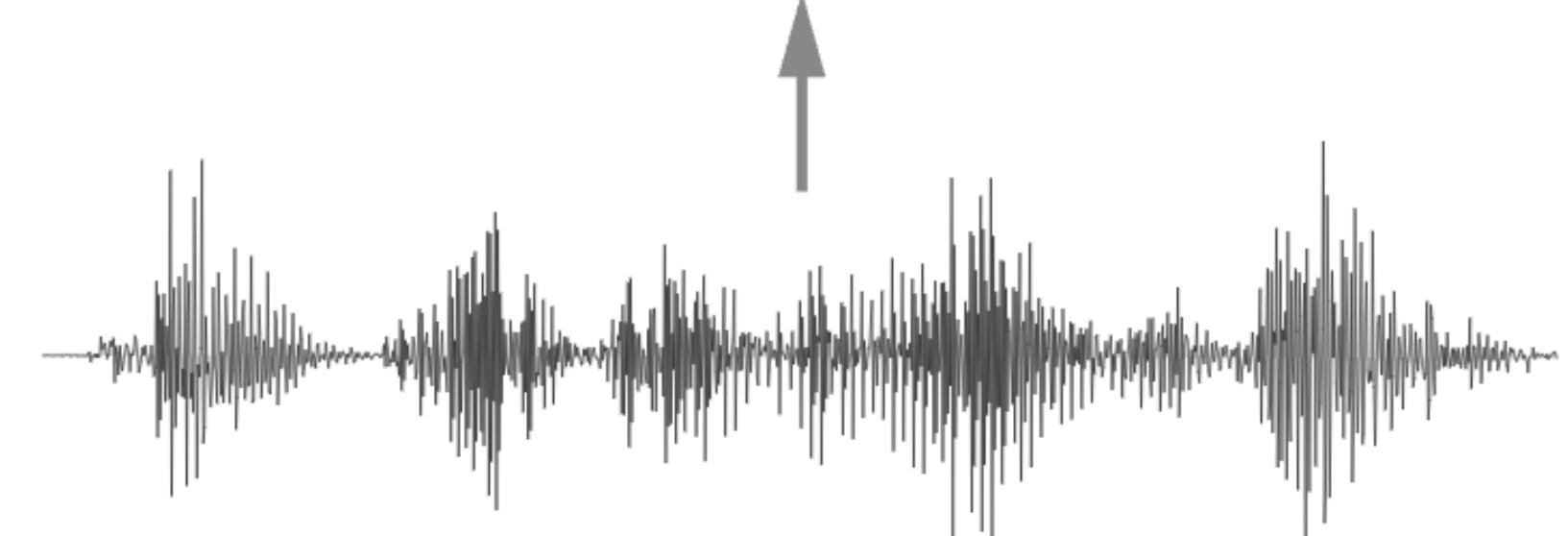
# Sequence modeling with Connectionist Temporal Classification (CTC)

the quick brown fox

*The quick brown fox*

**Handwriting recognition:** The input can be  $(x, y)$  coordinates of a pen stroke or pixels in an image.

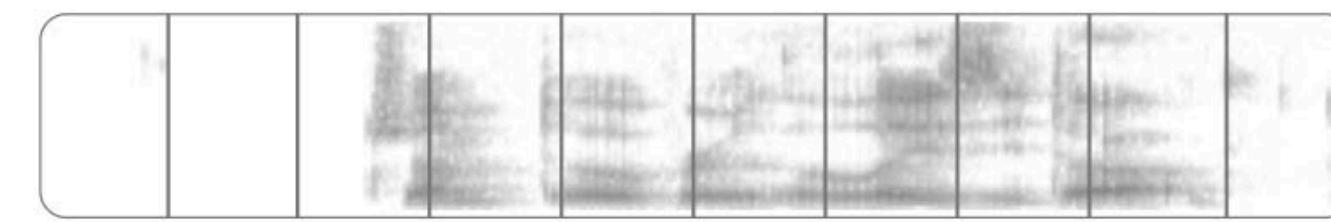
jumps over the lazy dog



**Speech recognition:** The input can be a spectrogram or some other frequency based feature extractor.

# Sequence modeling with CTC

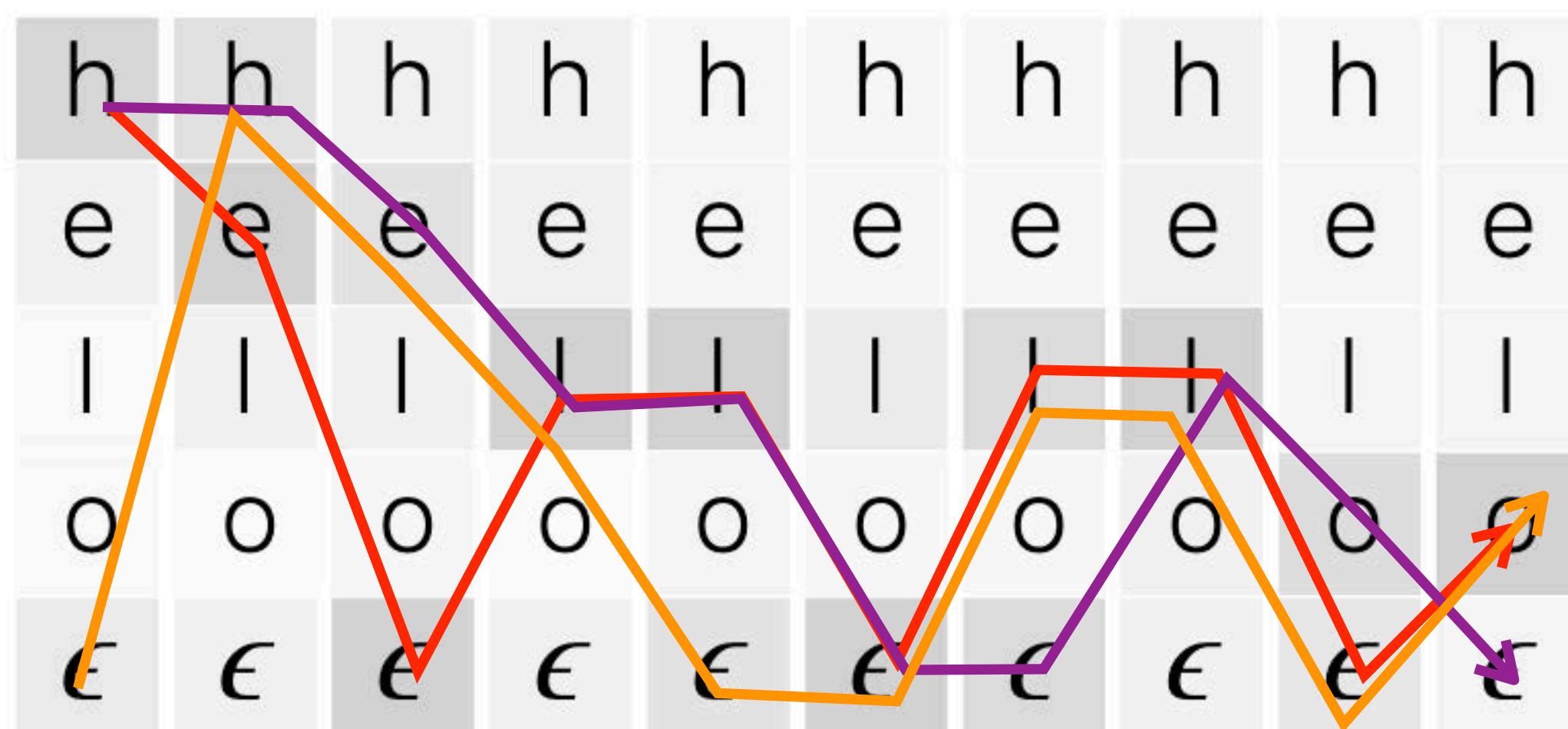
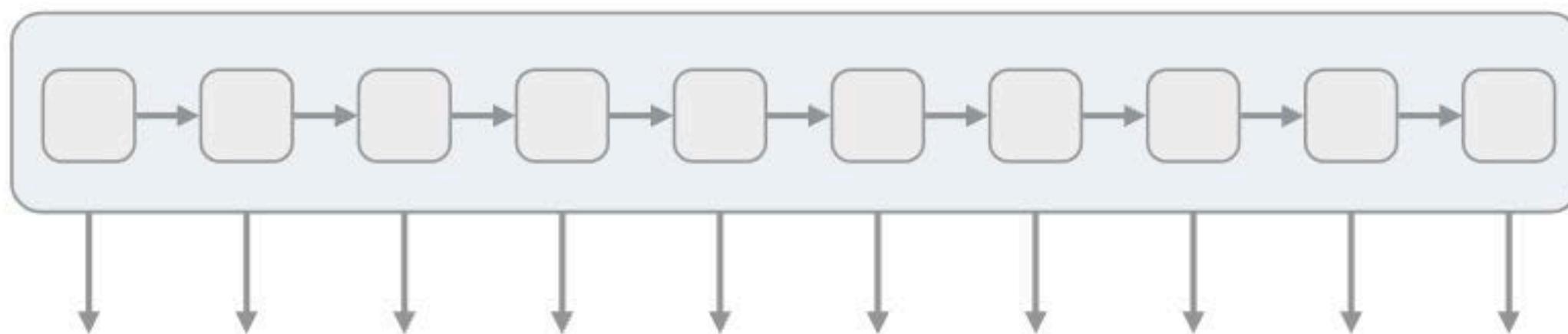
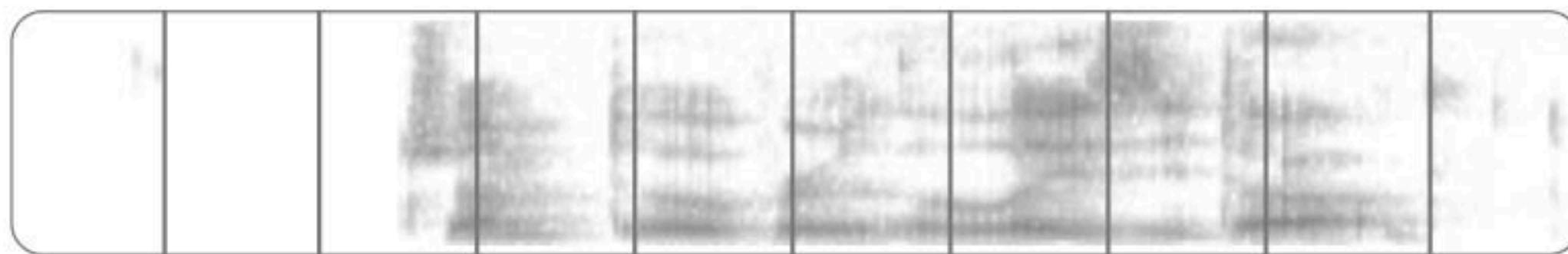
Inputs



CTC outputs post-processing

h h e ε ε l l l ε l l o

# CTC training



...  
ε h e l ε ε l l ε o  
h e l l ε l l ε o  
h h e l l ε ε l o o

$$p(Y | X) =$$

$$\sum_{A \in \mathcal{A}_{X,Y}}$$

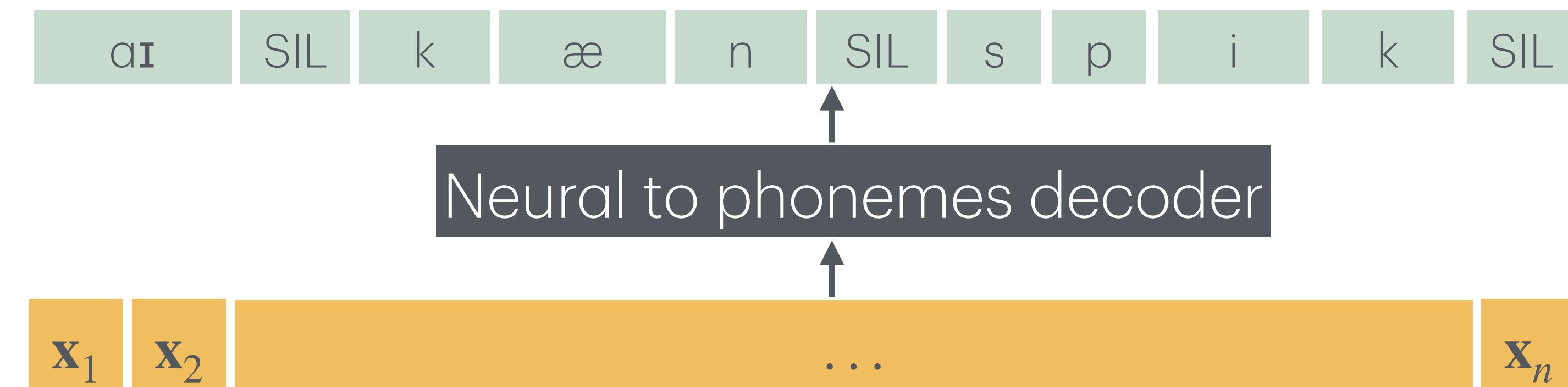
The CTC conditional probability

marginalizes over the set of valid alignments

$$\prod_{t=1}^T p_t(a_t | X)$$

computing the probability for a single alignment step-by-step.

# What neural network to use?

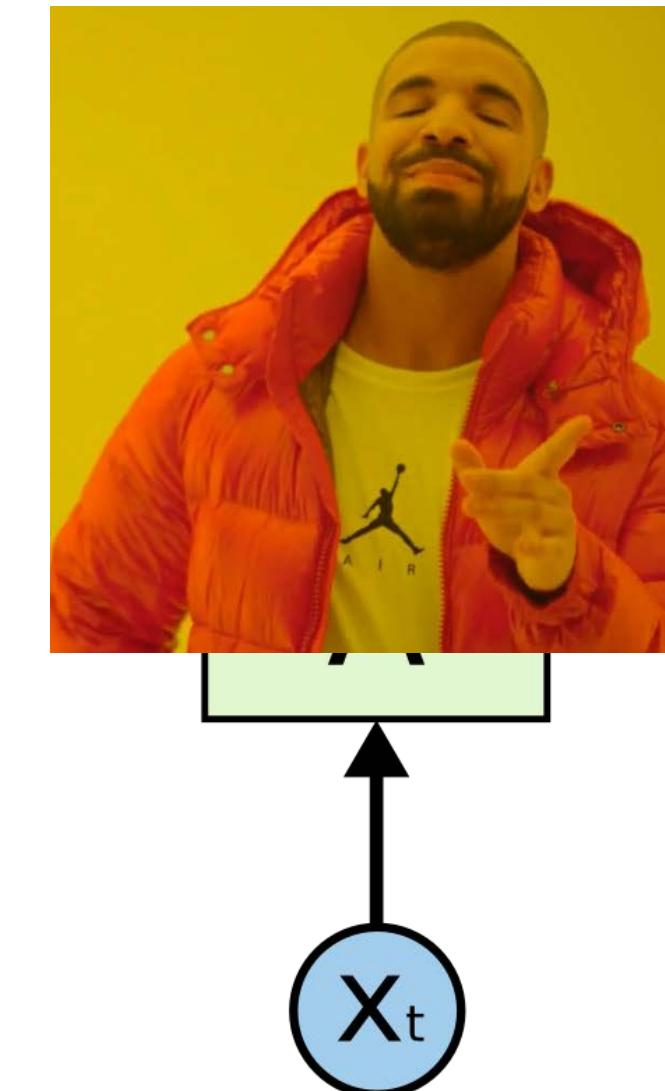


# What neural network to use?

Transformer



RNN



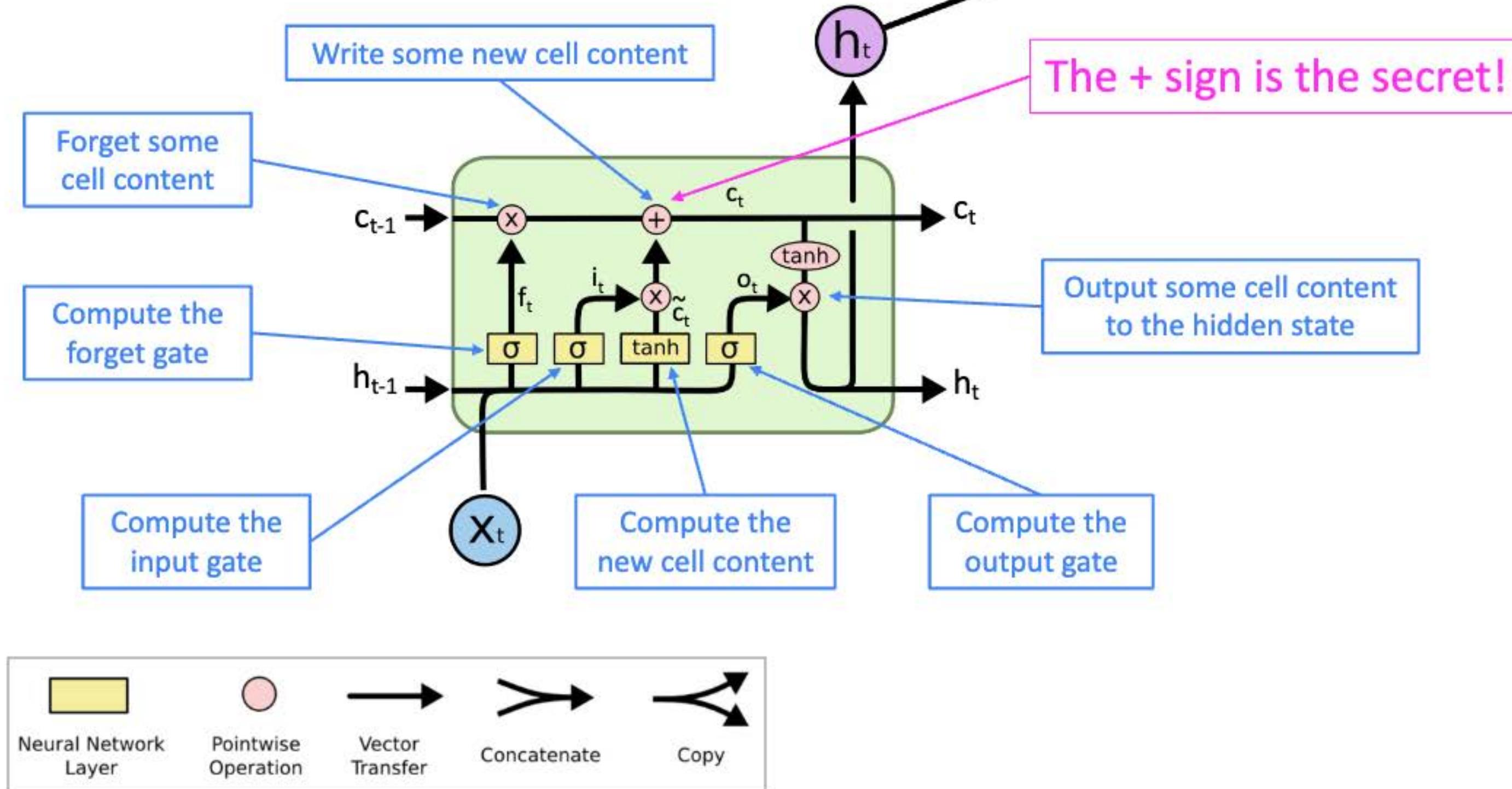
- 🤚 Large datasets
- 🤚 Long-range dependency

- 🤚 Small datasets
- 🤚 Short-range dependency
- 🤚 Efficient real-time processing

# LSTM

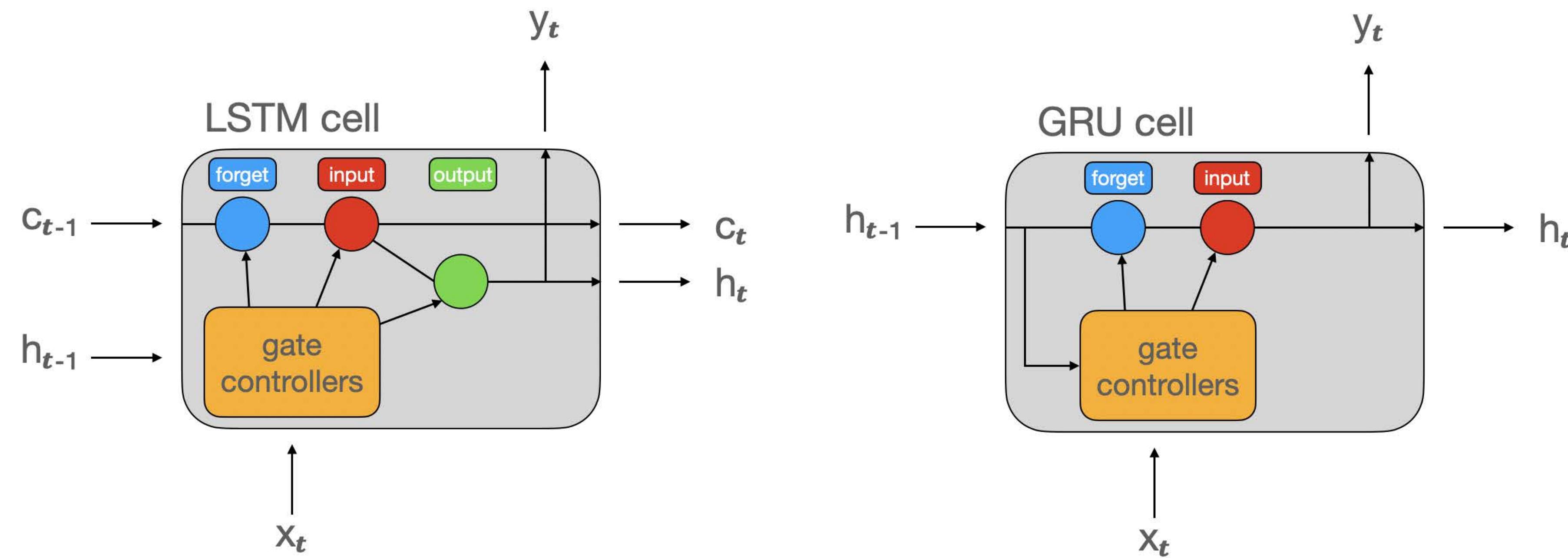
## Long Short-Term Memory (LSTM)

You can think of the LSTM equations visually like this:



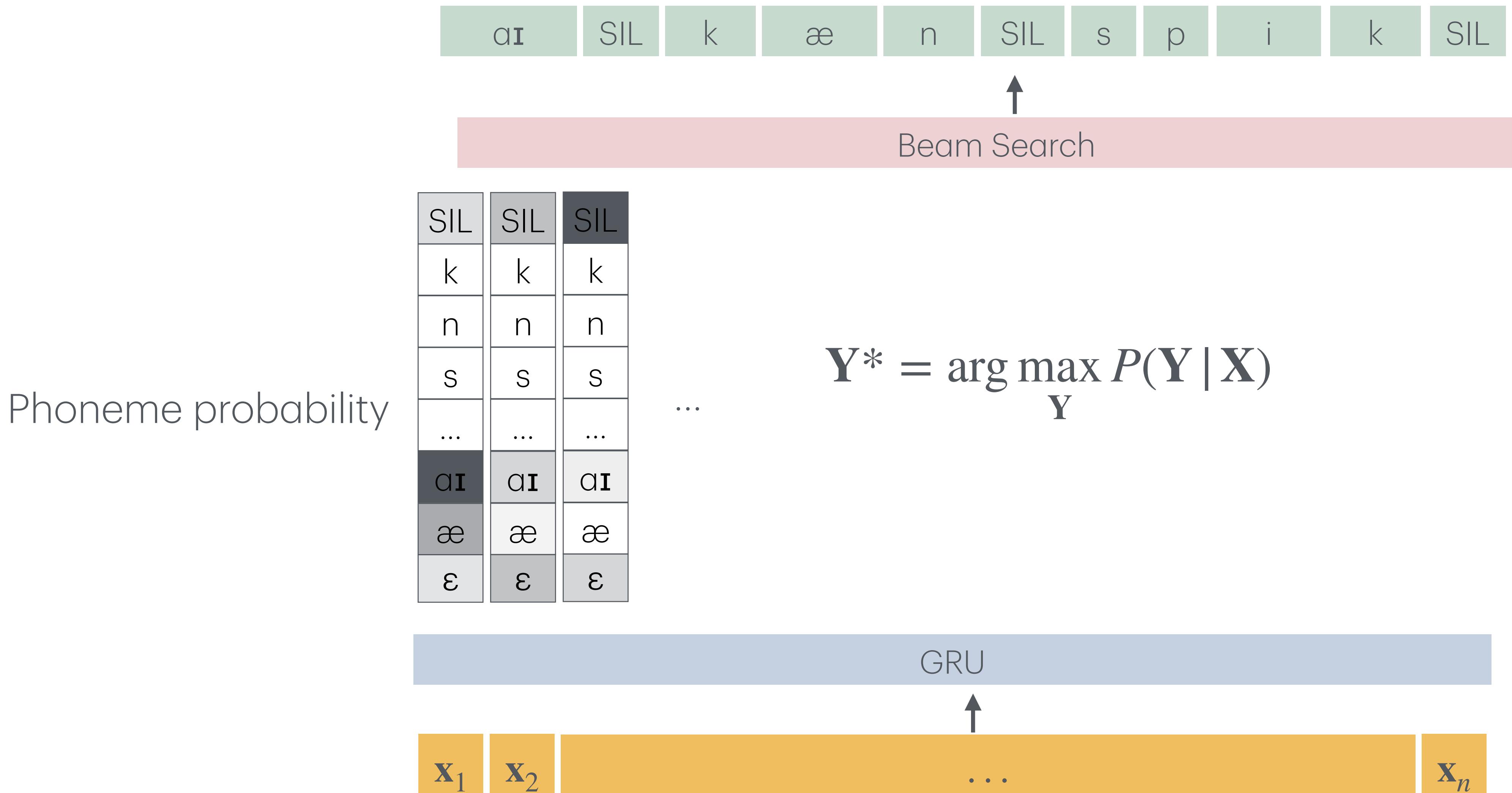
Source: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

# Gated Recurrent Units (GRU)

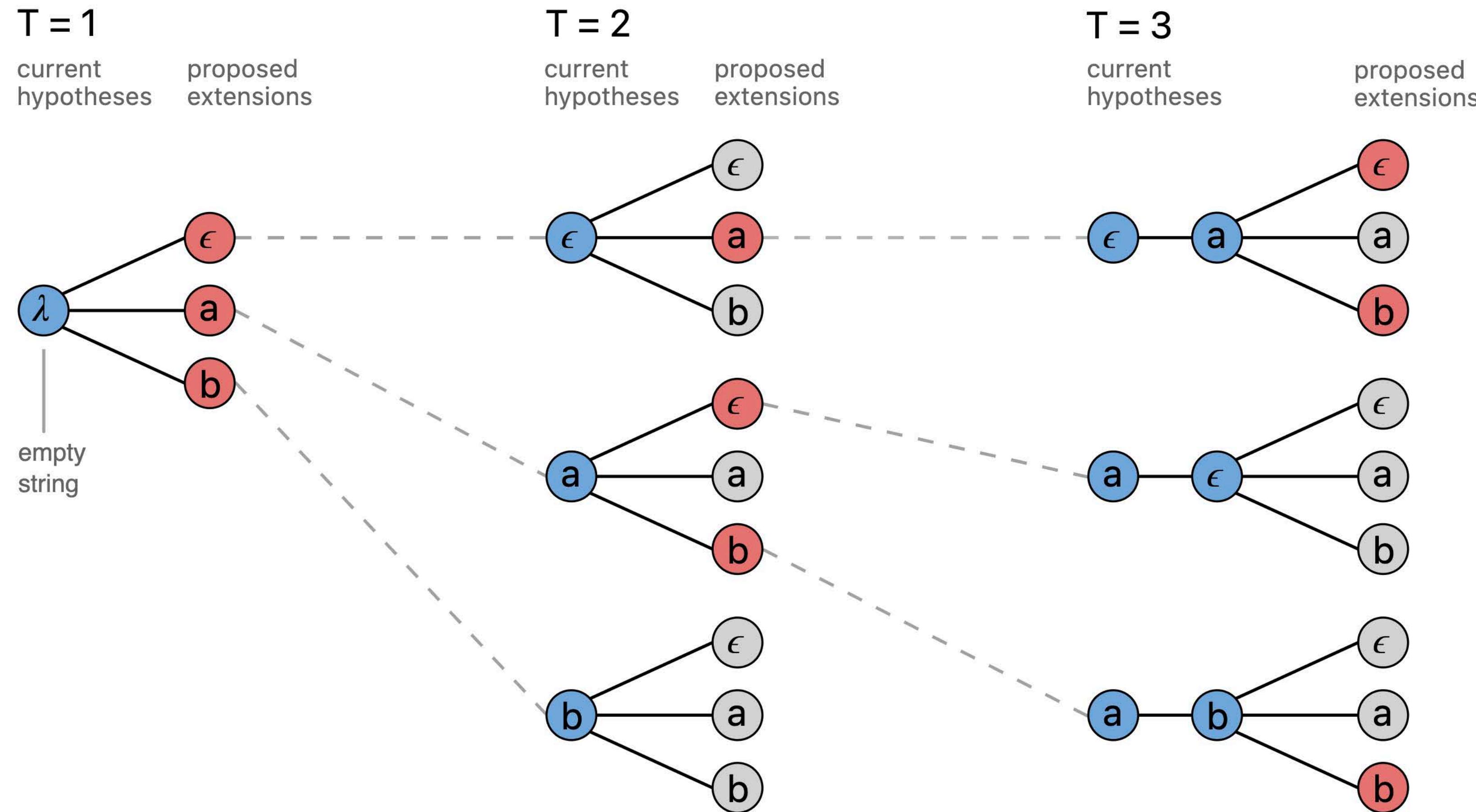


Source

# CTC inference

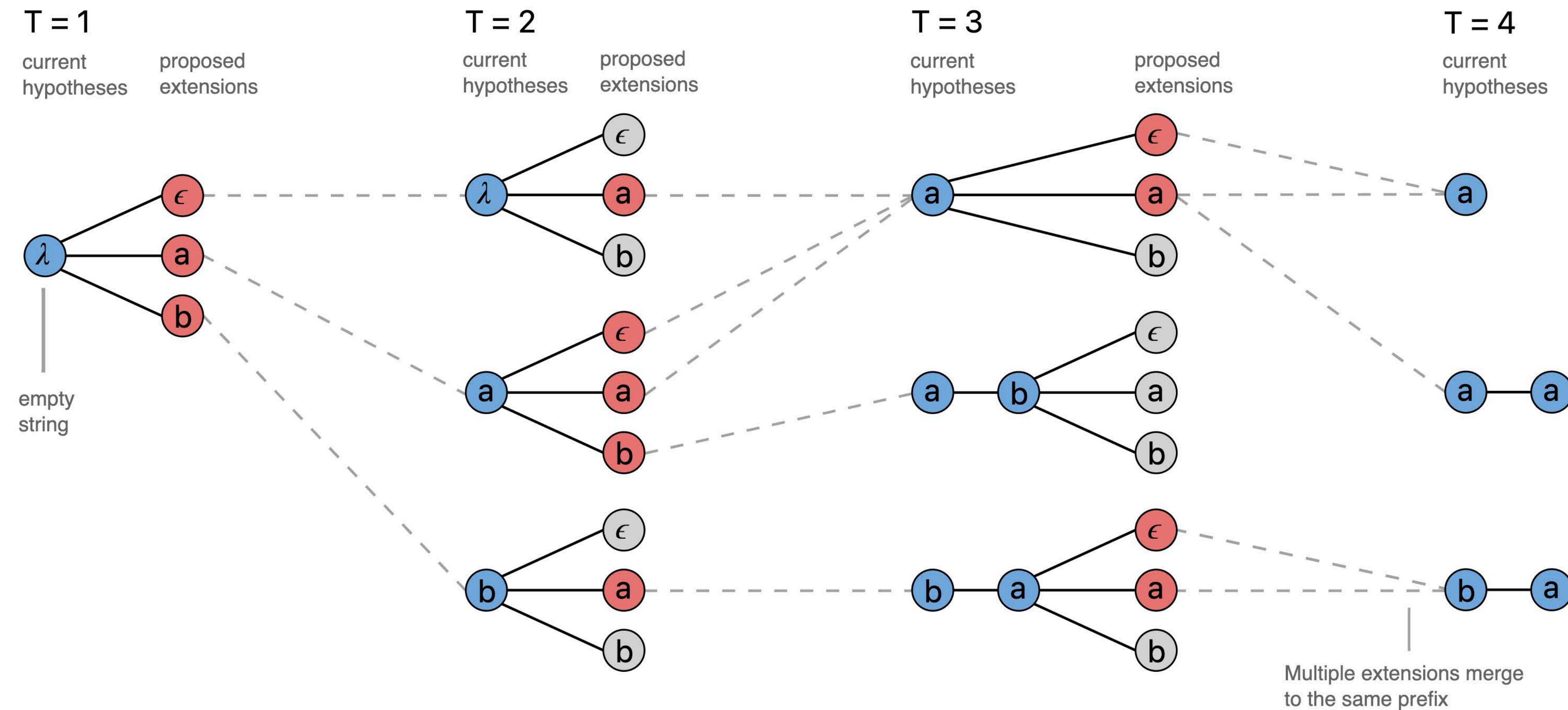


# CTC beam search



A standard beam search algorithm with an alphabet of  $\{\epsilon, a, b\}$  and a beam size of three.

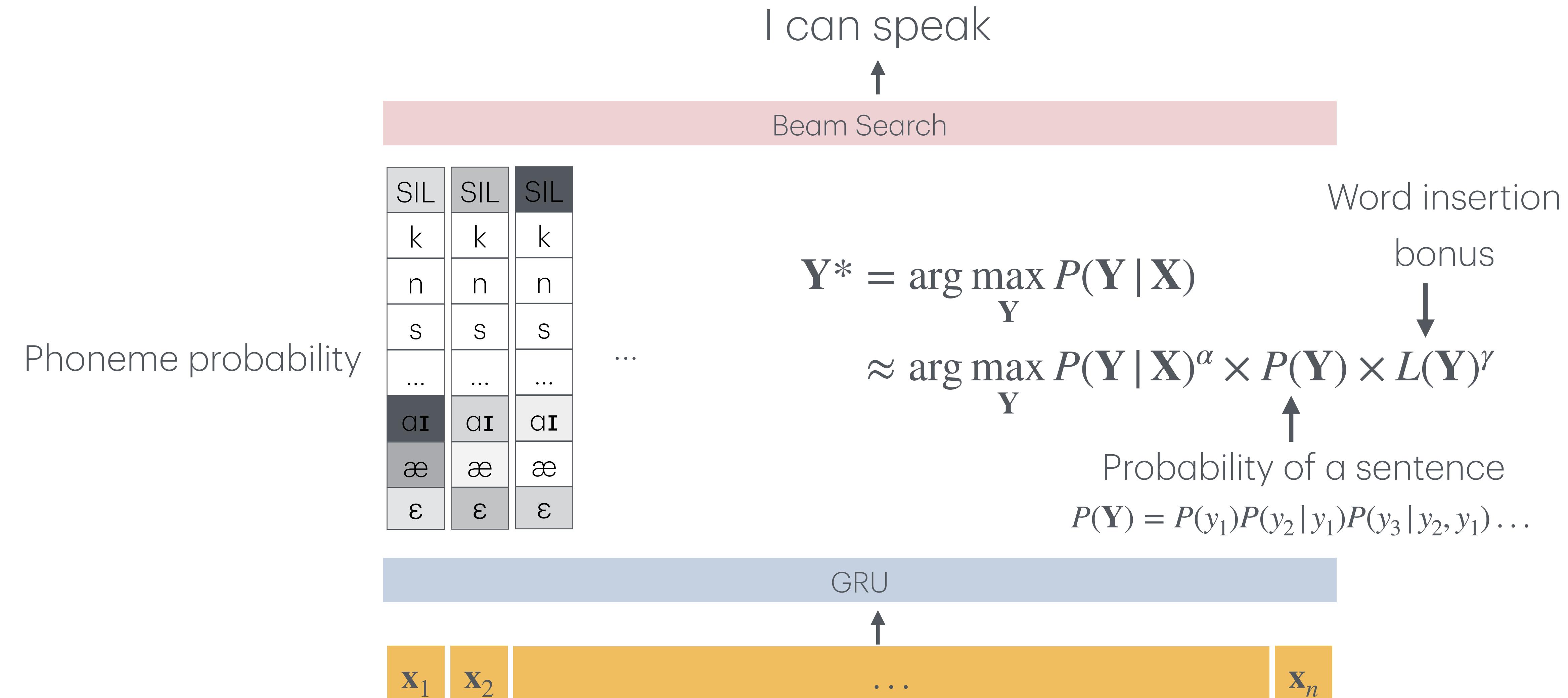
# CTC beam search



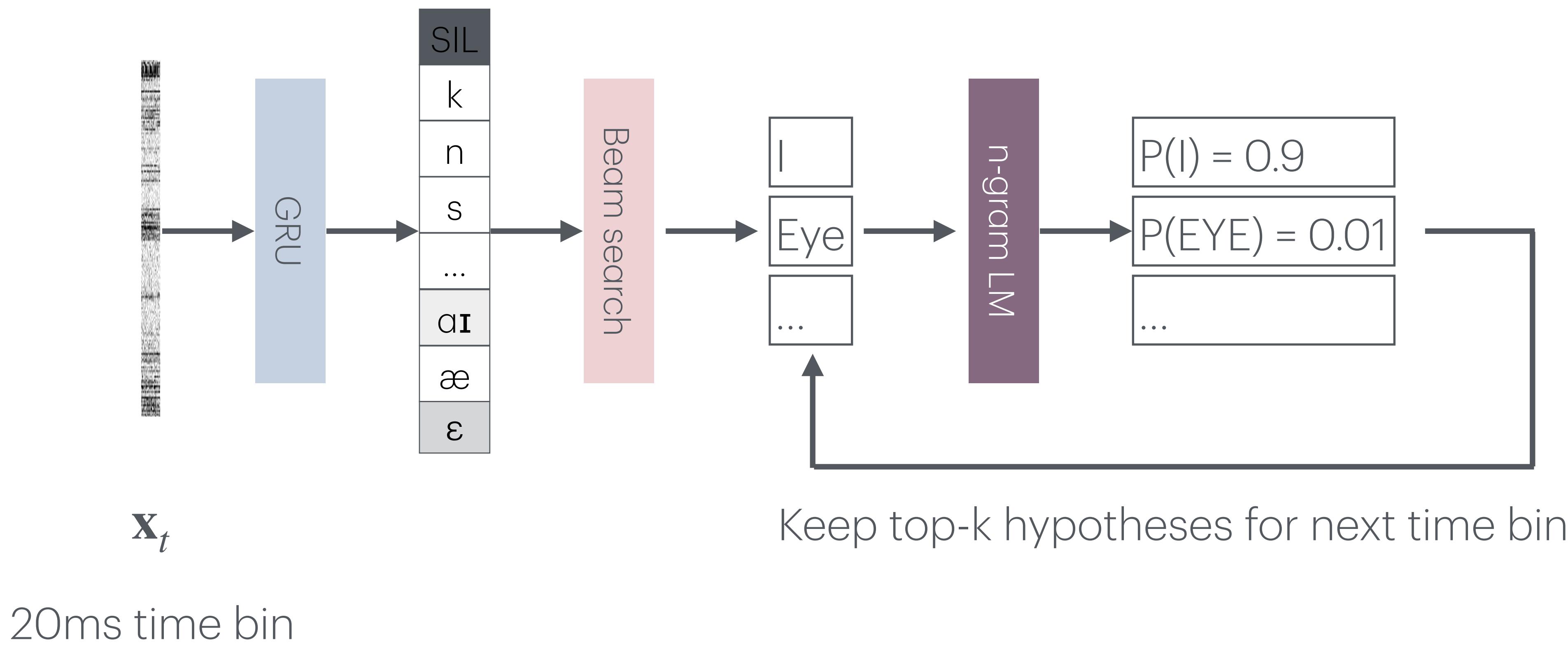
The CTC beam search algorithm with an output alphabet  $\{\epsilon, a, b\}$  and a beam size of three.

A. Hannun 2017

# CTC inference with language models

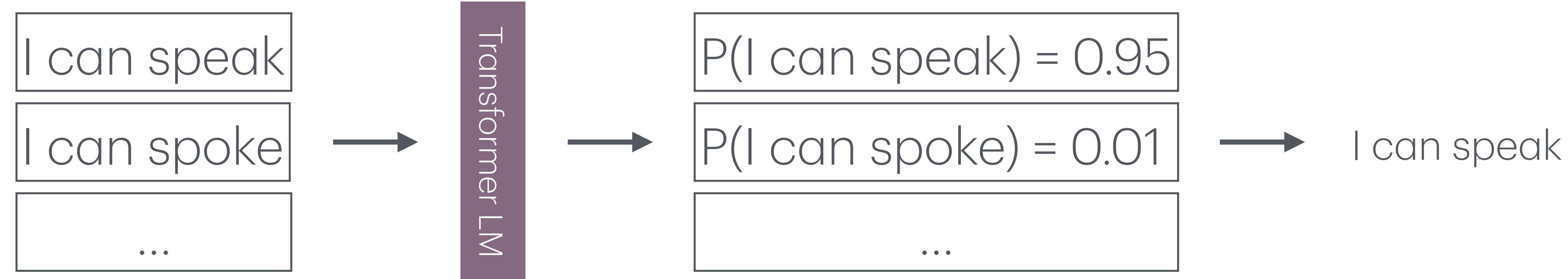


# Integrating language models in real-time decoding



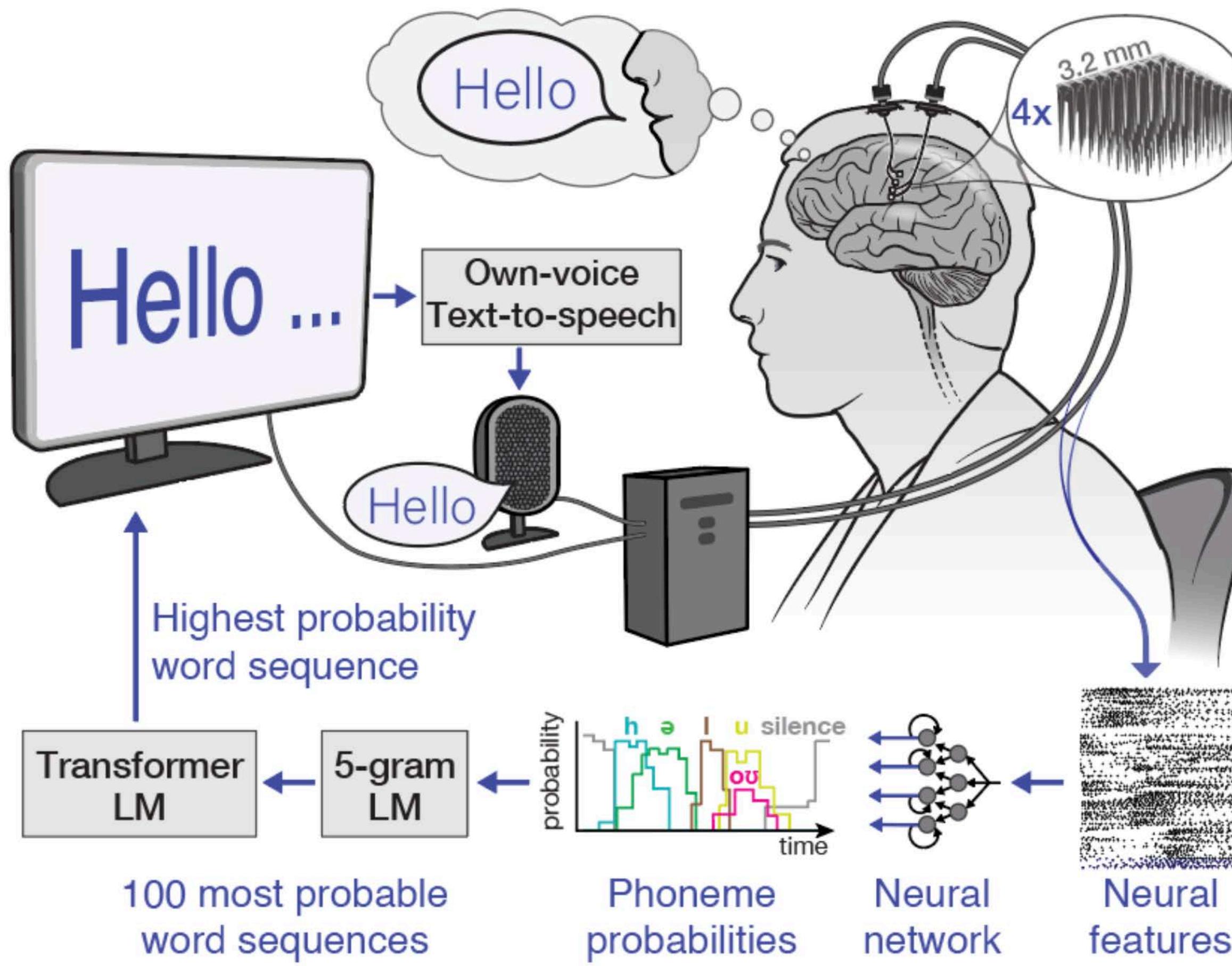
Keep top-k hypotheses for next time bin

# Transformer LM for 2nd pass rescoring



$n$ -best hypotheses

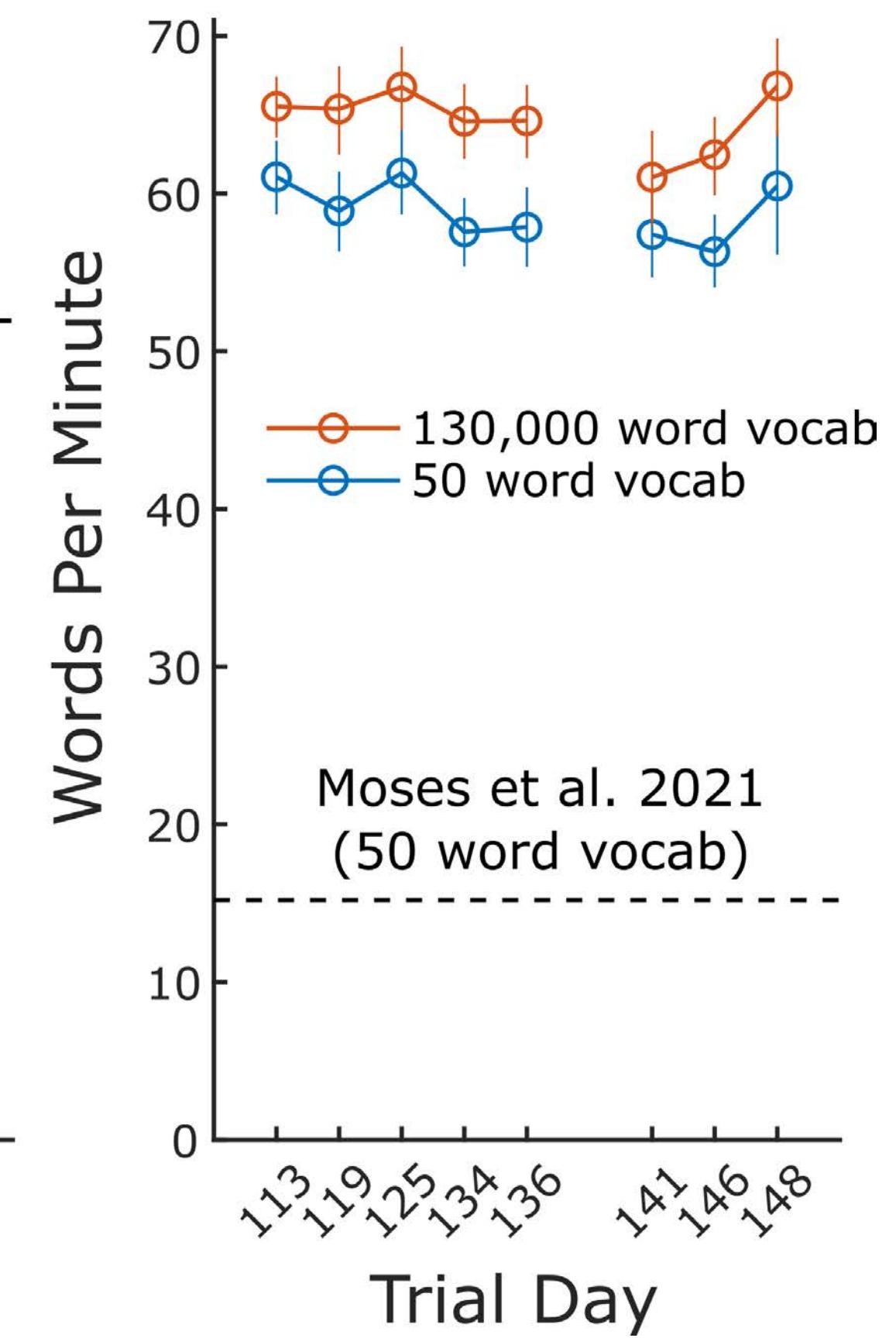
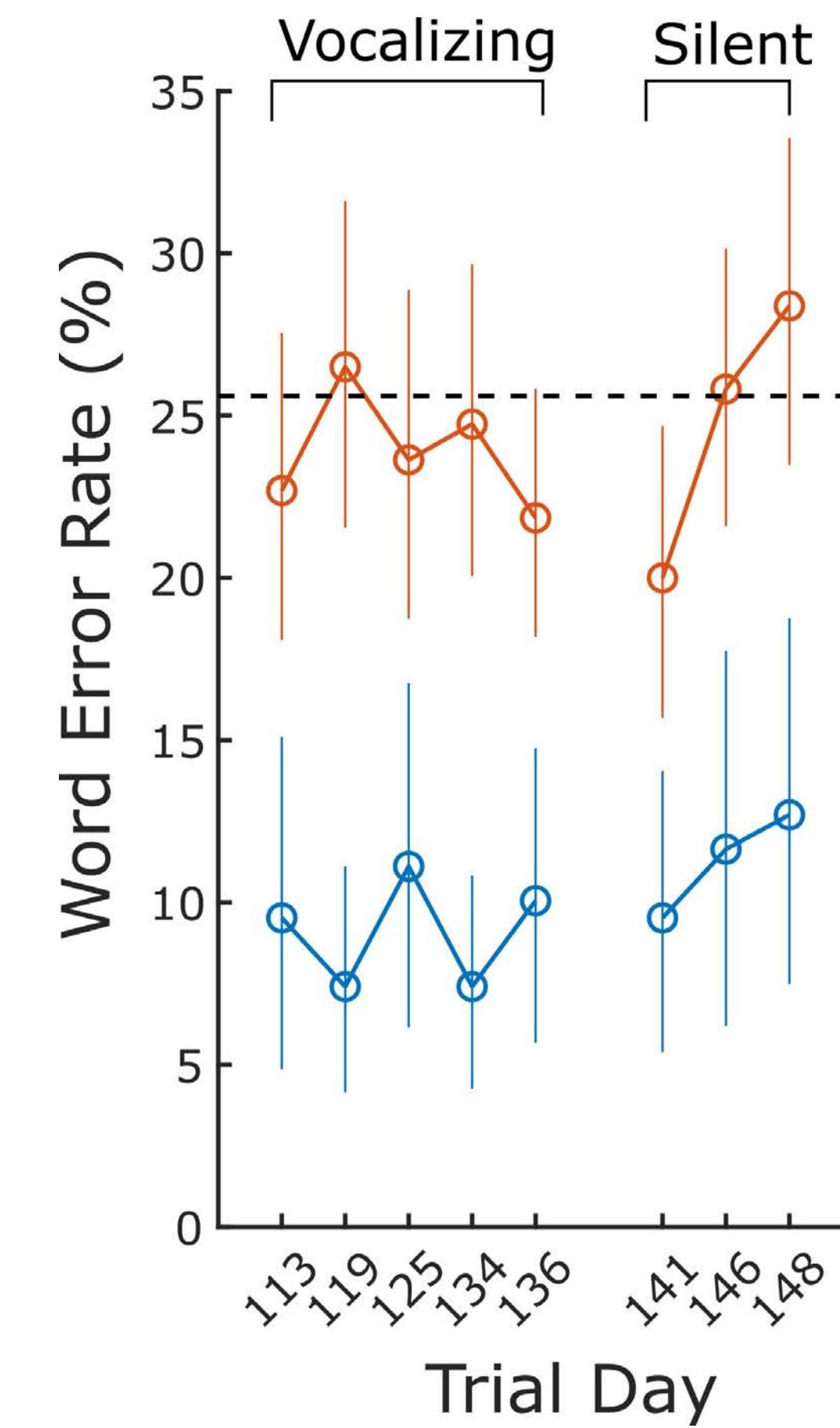
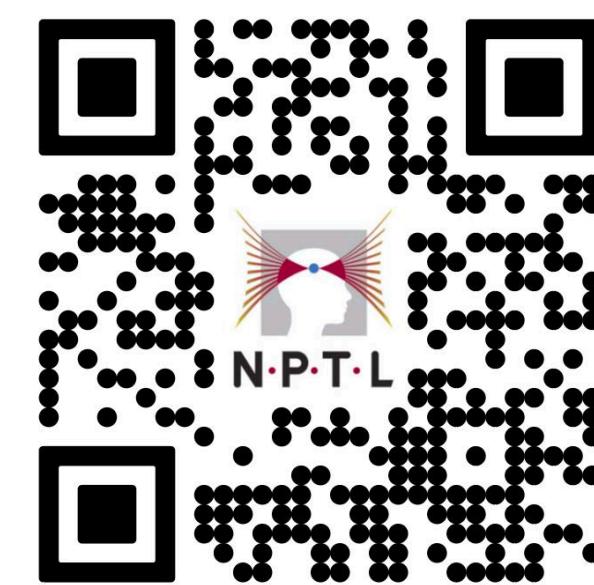
# Putting everything together



# Evaluation

**Word error rate:** normalized edit distance between predicted words and ground truth words.

$$WER(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{distance(\mathbf{Y}, \hat{\mathbf{Y}})}{length(\mathbf{Y})}$$



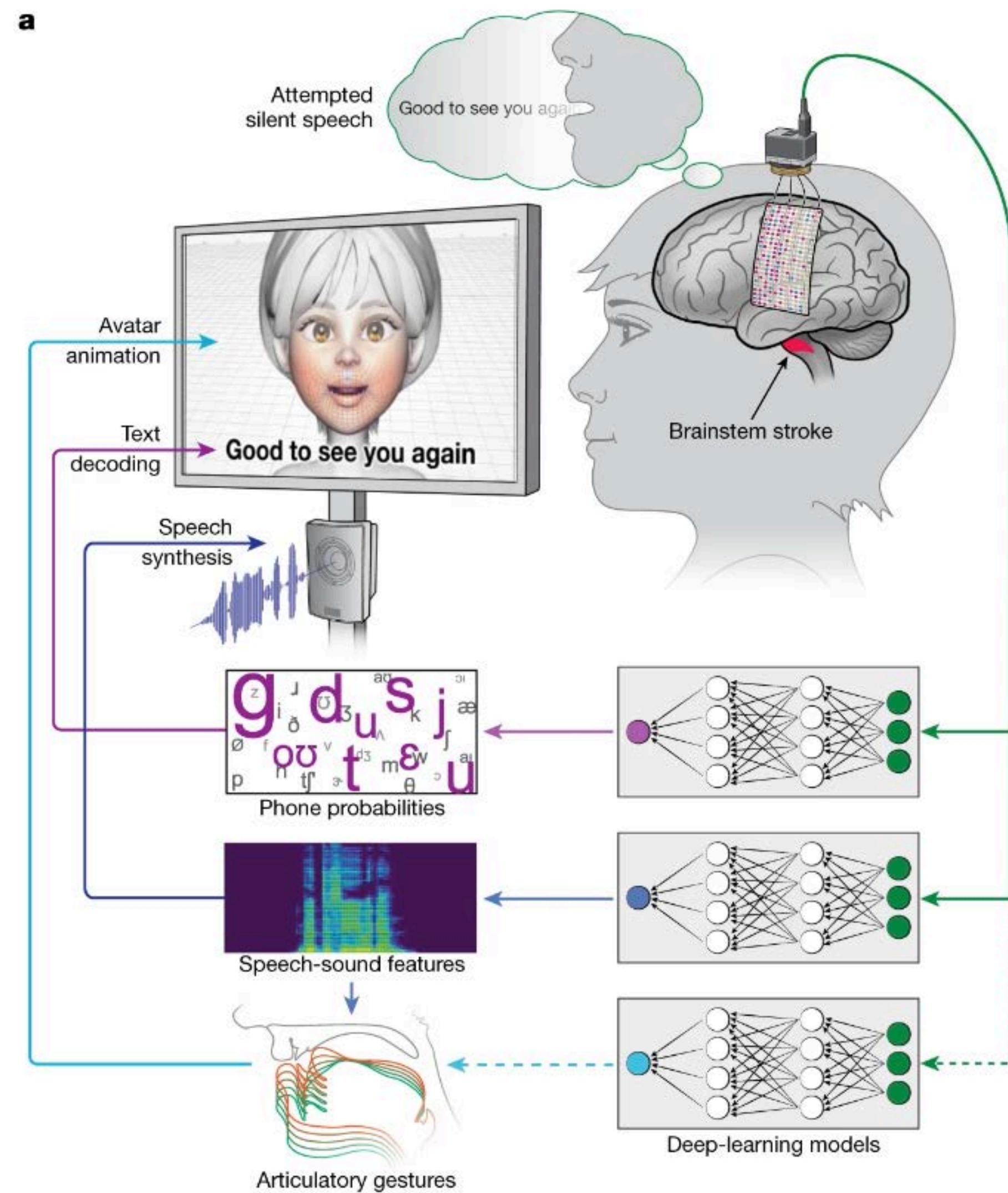
# What T12 said

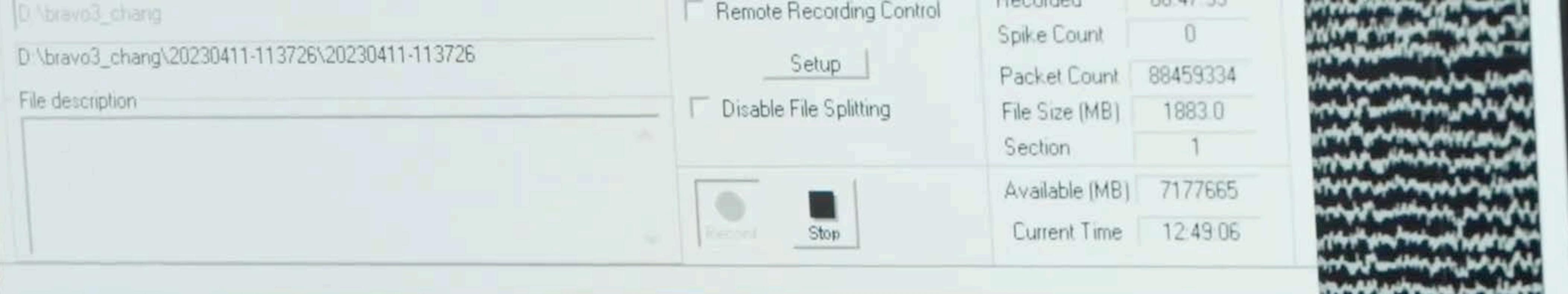
“So many years of not being able to communicate and then suddenly the people in the room got what I said. I don’t remember what I exactly said after the prescribed script finished, but it had to be along the lines of ‘Holy shit, it worked, I’m so happy, and you guys did it.’”



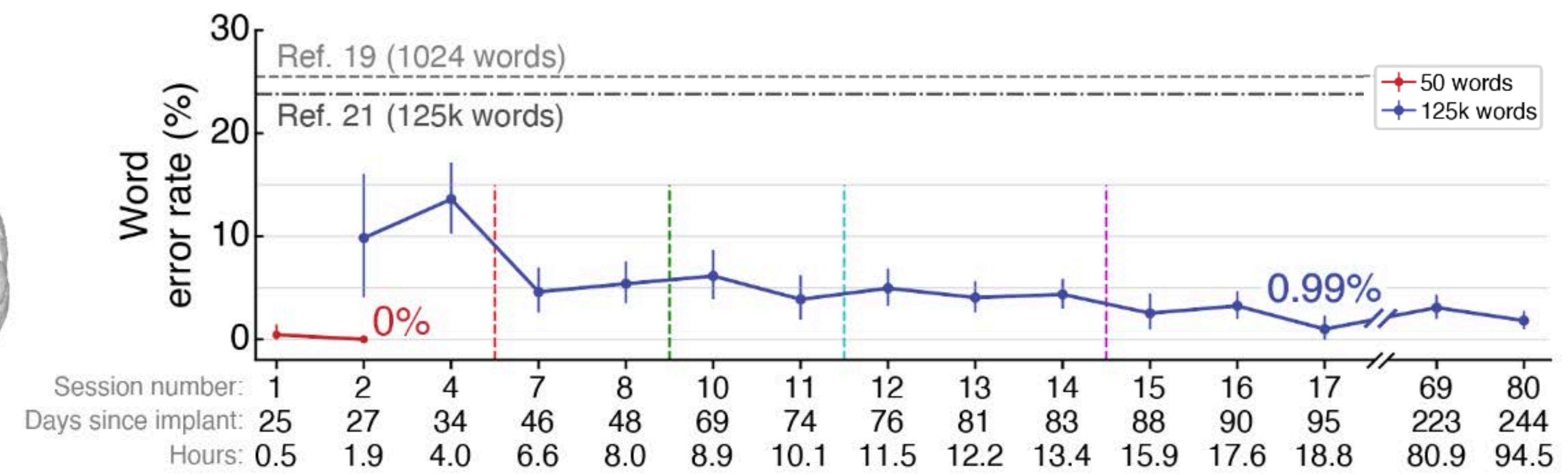
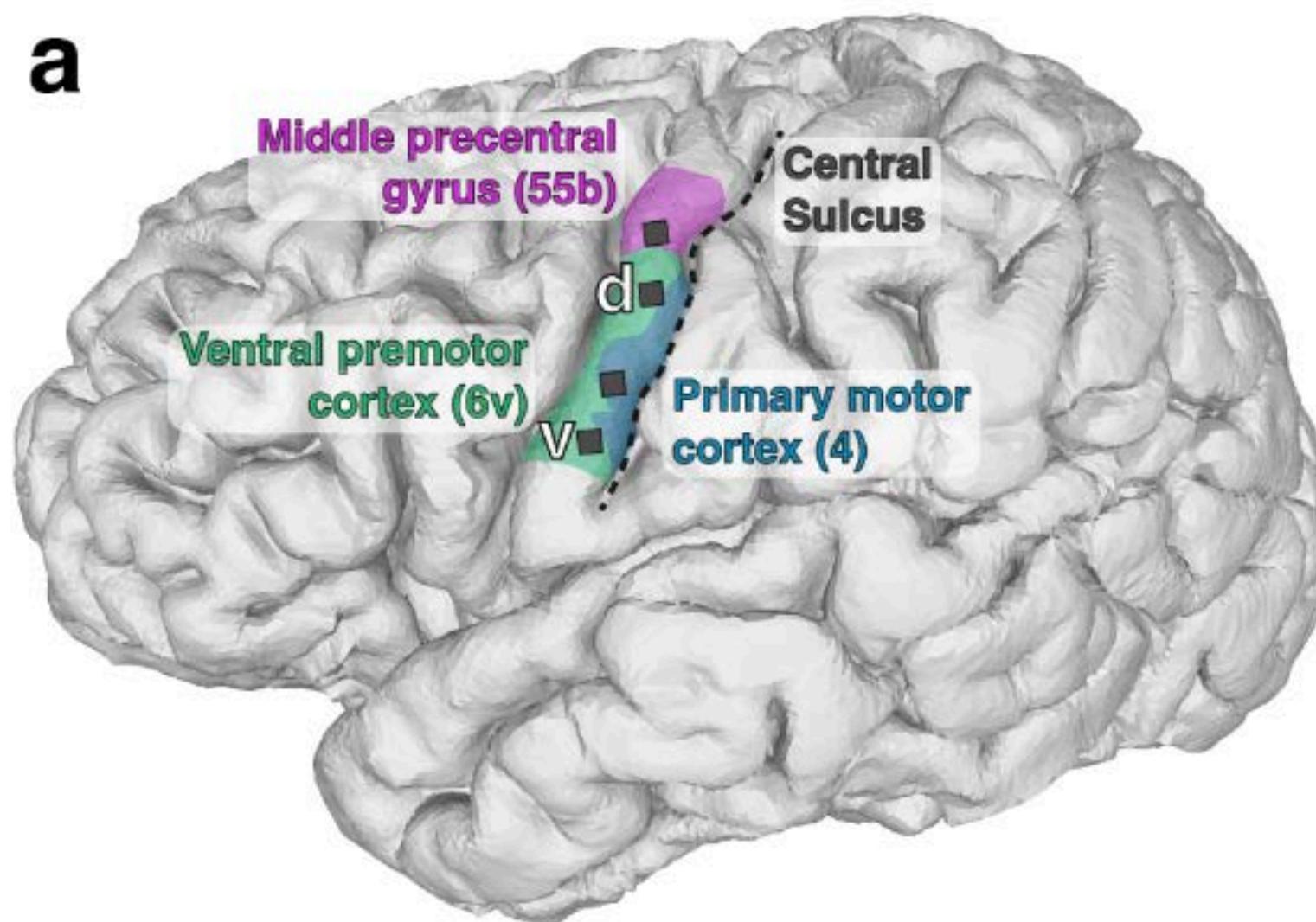
# Future of Speech BCIs

# Multimodal Speech BCI





# An accurate speech BCI for personal use





100% CORRECT

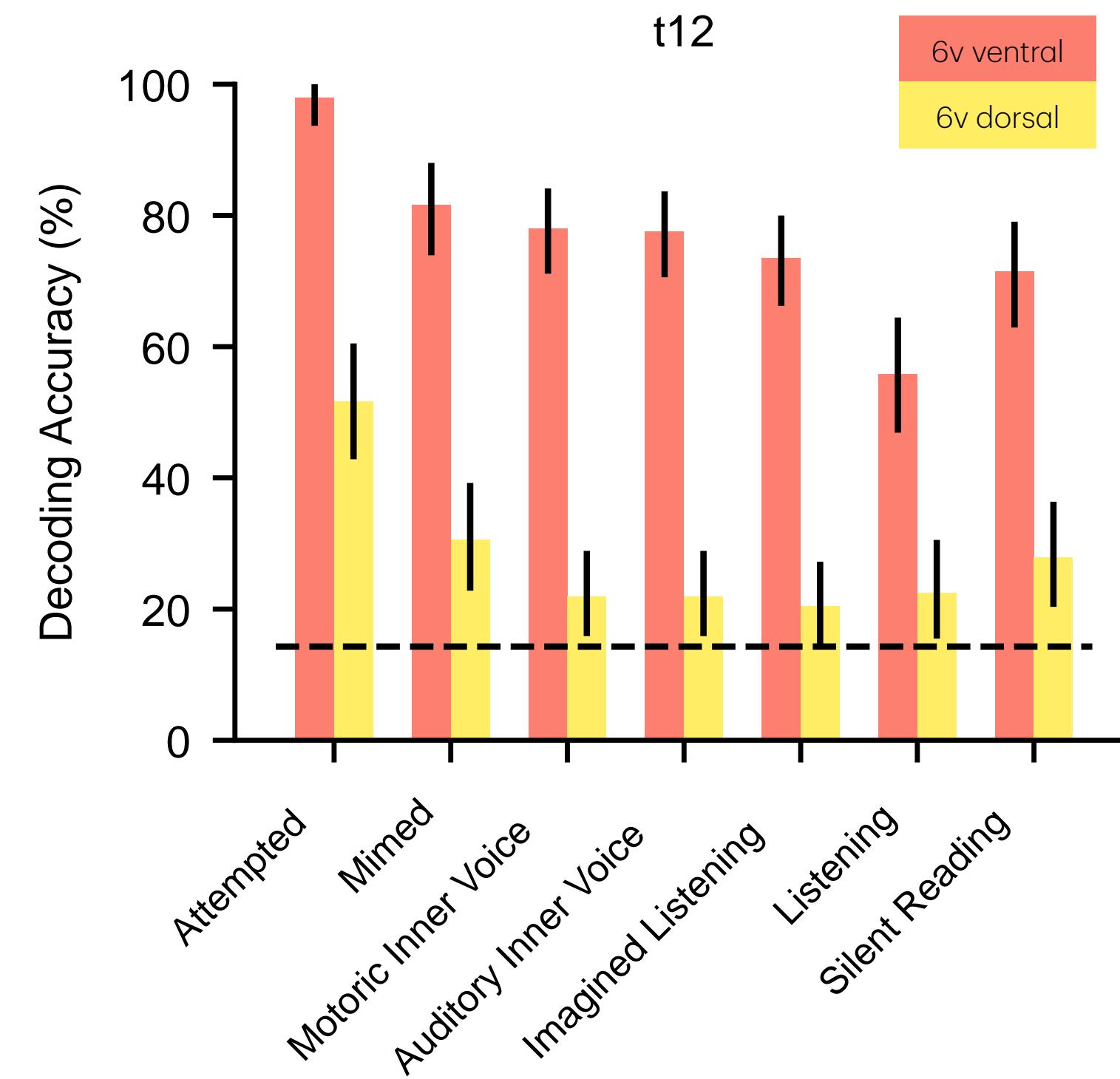
MOSTLY  
CORRECT

INCORRECT

DONE

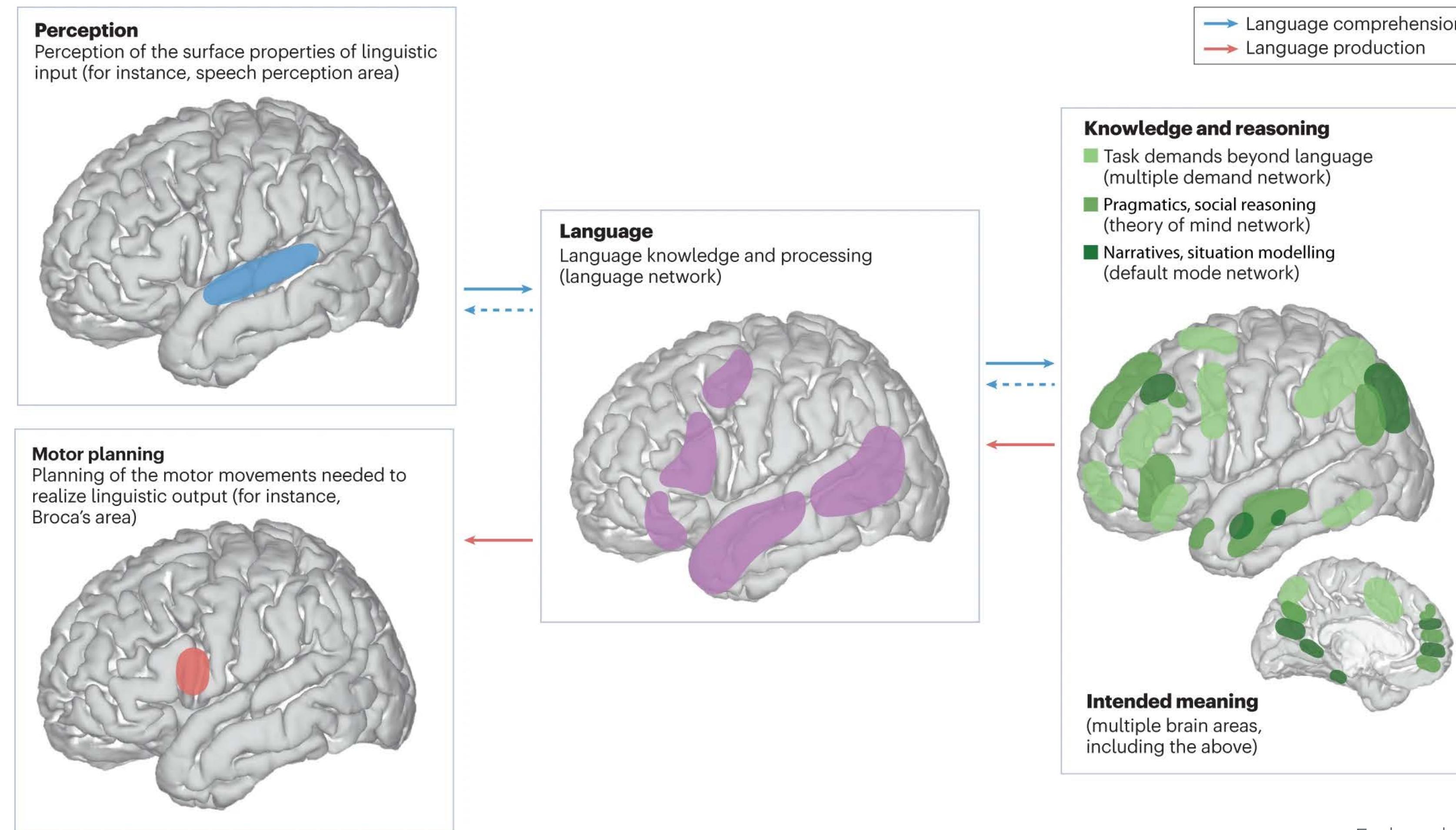
Activate spelling  
mode

# Restoring effortless and natural communication by decoding inner speech



Cope & Kalantzis / Transpositional Grammar • meaningpatterns.net/inner-speech

# Language processing in the brain



# BCIs raise new neuroethics considerations

- Should BCIs be allowed to read inner thoughts and memories?
  - Read out inner speech that would not naturally be enacted?
  - Read out memories that may otherwise be lost to Alzheimer's disease?
  - Read out subconscious fears to assist desensitization psychotherapy?

# BCIs raise new neuroethics considerations

- Should BCIs be allowed to enhance cognitive function beyond natural levels?
  - Move a robotic arm faster and more accurately than a native arm?
  - Purchase a memory to skip a grade of mathematics in high school?
- We are currently grappling with the same questions:
  - Steroids, stimulants, elective plastic surgery ...

# BCIs raise new neuroethics considerations

Although some of these ideas and questions may appear farfetched at present, as brain function and dysfunction continues to be revealed, BCI systems could build on these discoveries and create even more daunting ethical quandaries. But equally important is the immediate need to help people suffering from profound neurological disease and injury through restorative BMIs. In order to achieve the right balance it is imperative that we as physicians, scientists and engineers proceed in close conversation and partnership with ethicists, government oversight agencies, and patient advocacy groups.

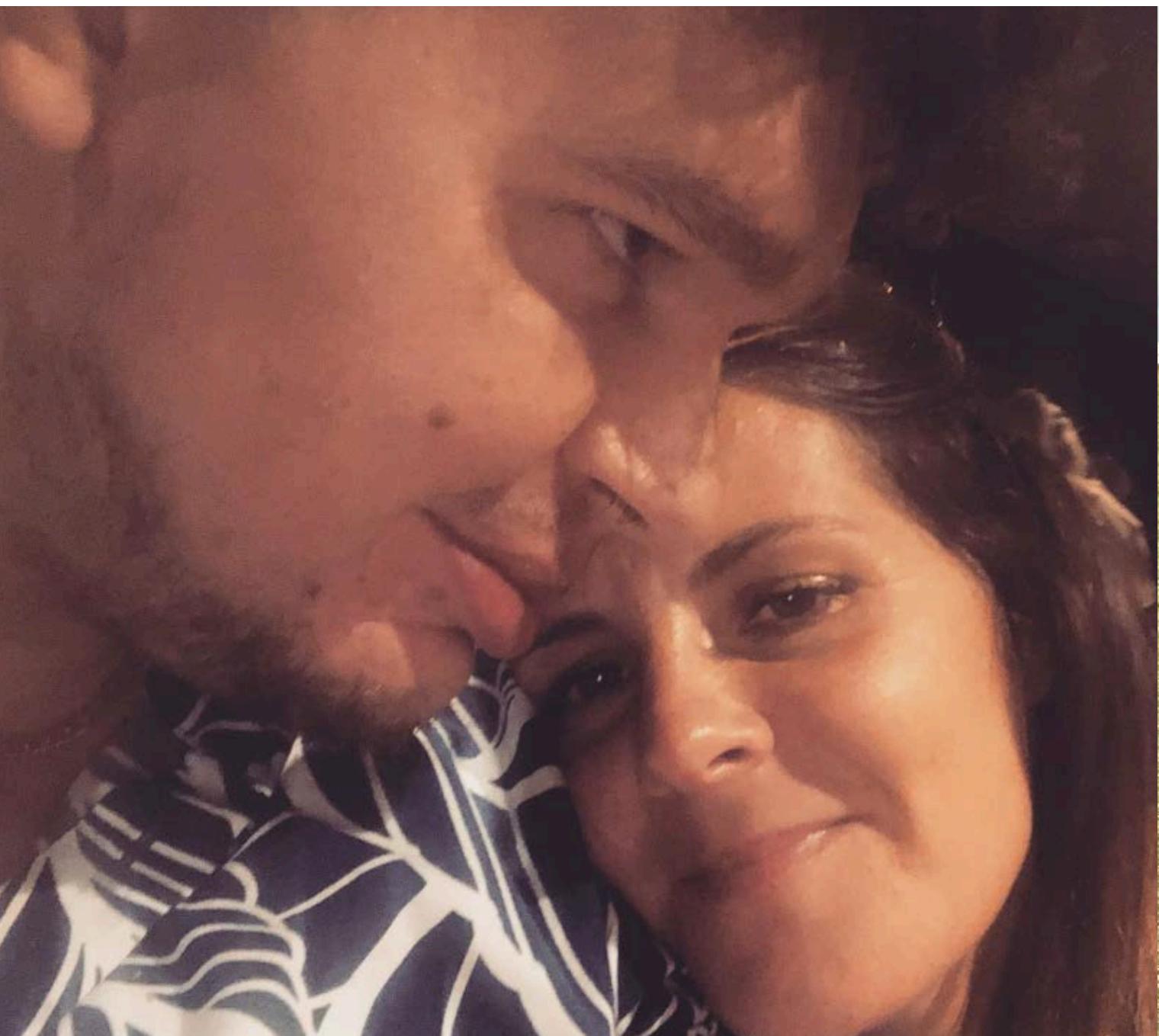
# Summary

- Recent advancements in AI and NLP, combined with years of neuroscience and neuroengineering research, show potential for restoring natural communication to people with speech impairments.
- We will soon have systems to assist people with communication disorders and paralysis.
- And understand better how the brain processes language!
- This brings hope to people like Howard and T12!



NEW HOPE FOR ALS PATIENTS





@howard\_wicks

# Acknowledgements



**Krishna Shenoy**

(9/3/1968 - 1/21/2023)

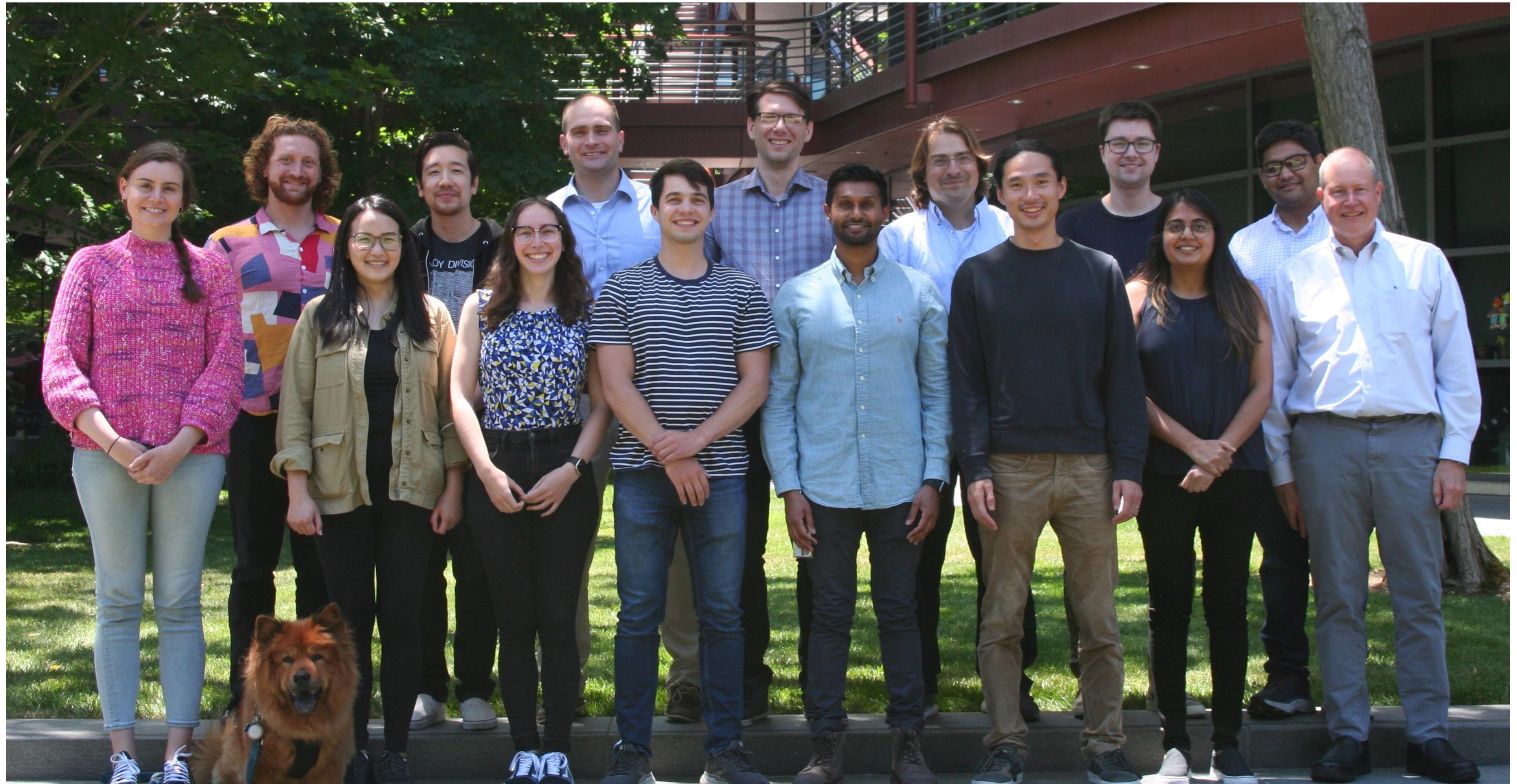


Howard Hughes  
Medical Institute

Participants  
T5, T6 and T12

## Funding

ALS Association  
Stanford: Bio-X, Wu Tsai  
Neuroscience, OPA



## Stanford Neural Prosthetics Translational Lab (NPTL)

NIH: NINDS, NIDCD, NICHD  
VA Rehab. R&D Service  
Mass. General Hospital ECOR

Larry and Pamela Garlick  
Samuel and Betsy Reeves  
John Gunn

