Evidence of a predictive coding hierarchy in the human brain listening to speech

Charlotte Caucheteux, Alexandre Gramfort & Jean-Rémi King

Nature Human Behaviour, 2023

Caucheteux, Gramfort & King (2023)

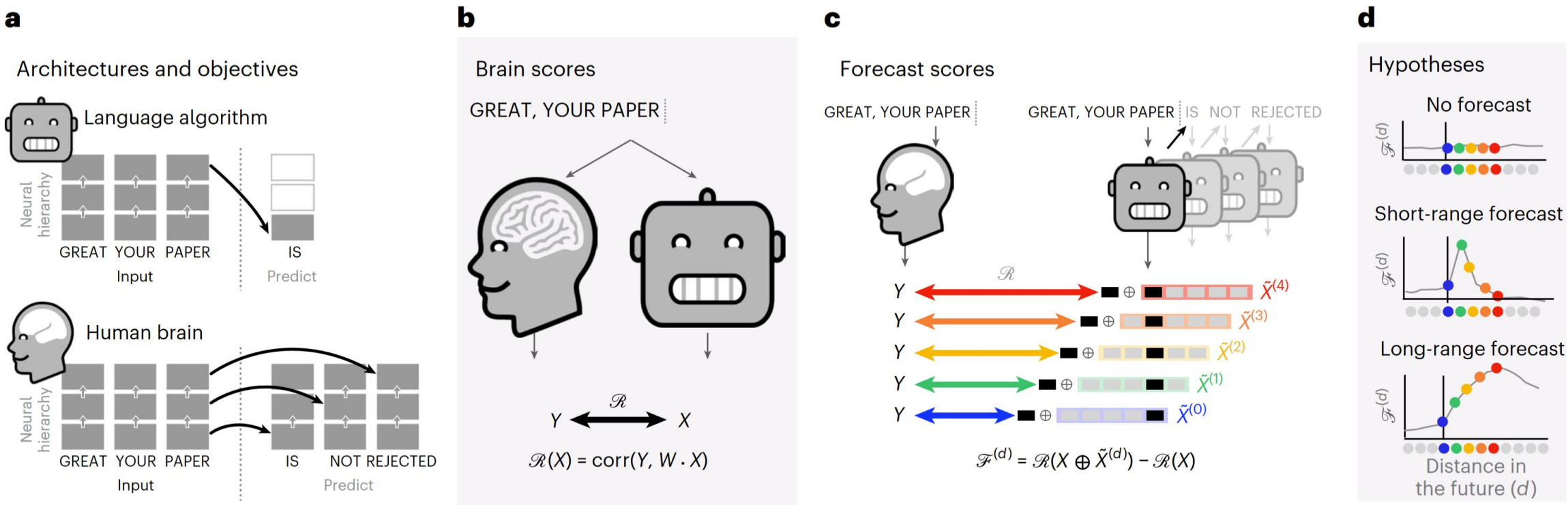
Introduction

- functional magnetic resonance imaging brain signals of 304 participants listening to short stories

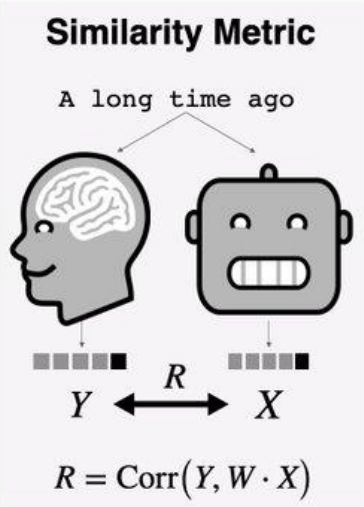
Hypotheses:

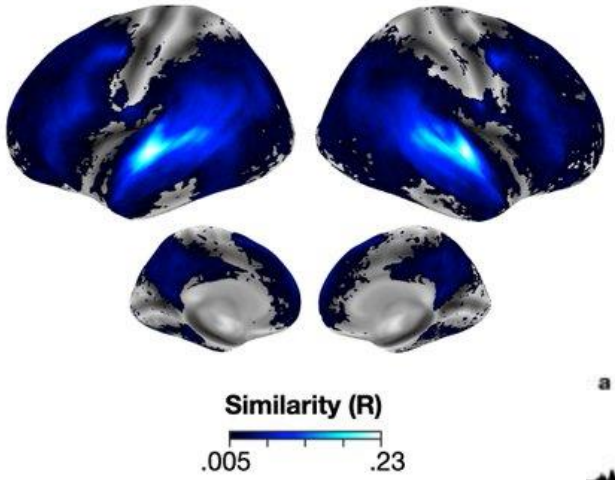
1. Activations of modern language models linearly map onto the brain responses to speech
2. Enhancing these algorithms with predictions that span multiple timescales improves this brain mapping
3. Predictions are organized hierarchically: frontoparietal cortices predict higher-level, longer-range and more contextual representations than temporal cortices

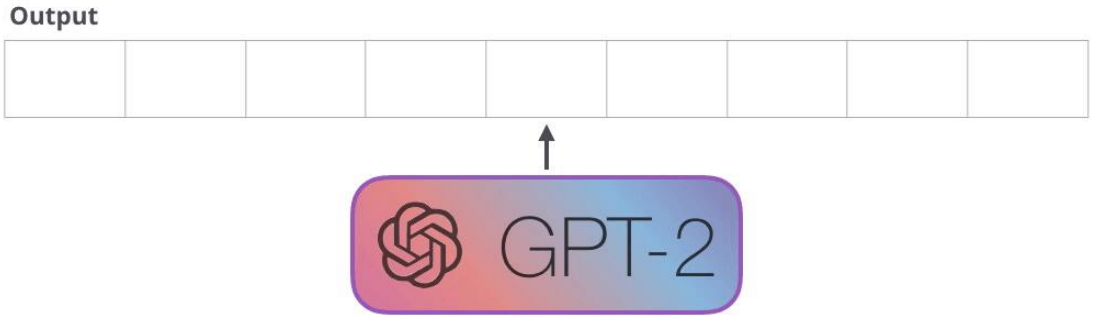
Experimental approach

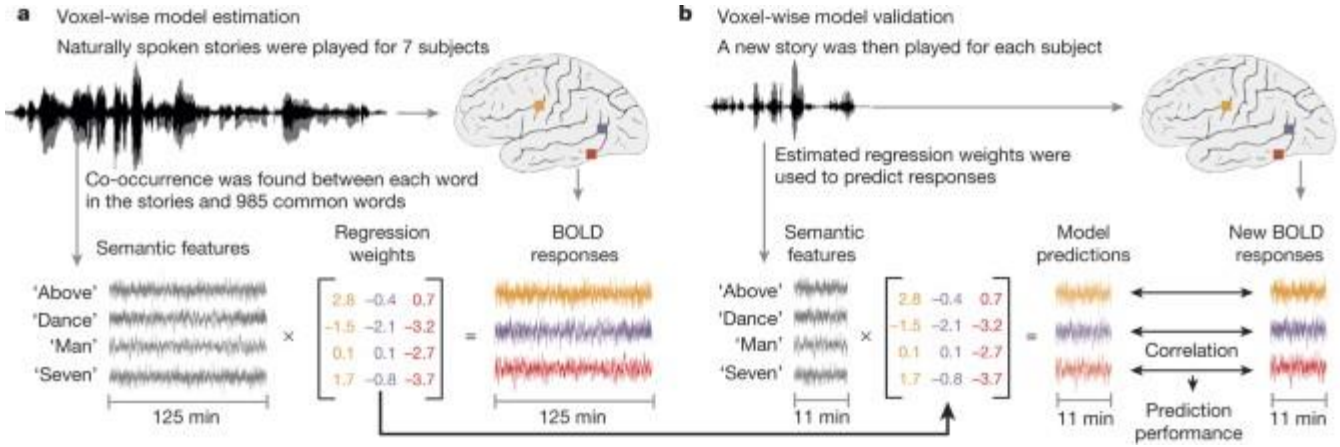


Brain score

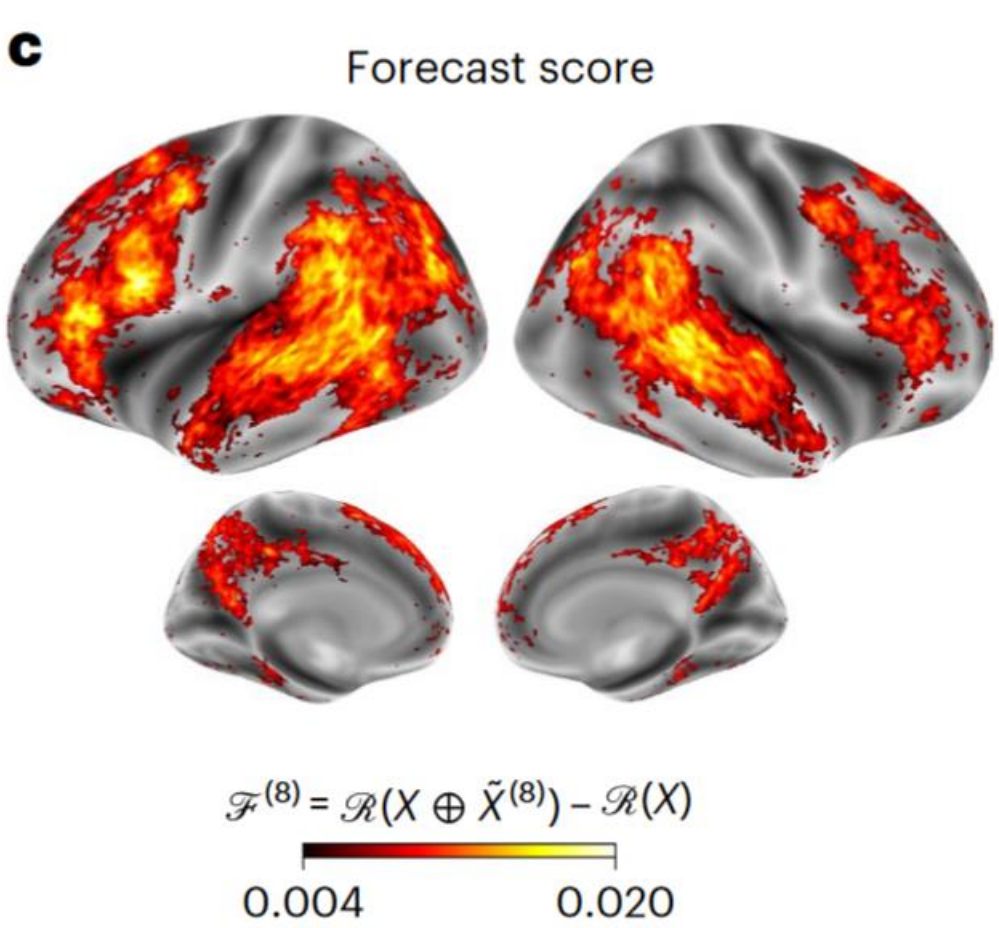

Similarity Metric
A long time ago
R
Y
X
$R = \mathrm{Corr}(Y, W \cdot X)$

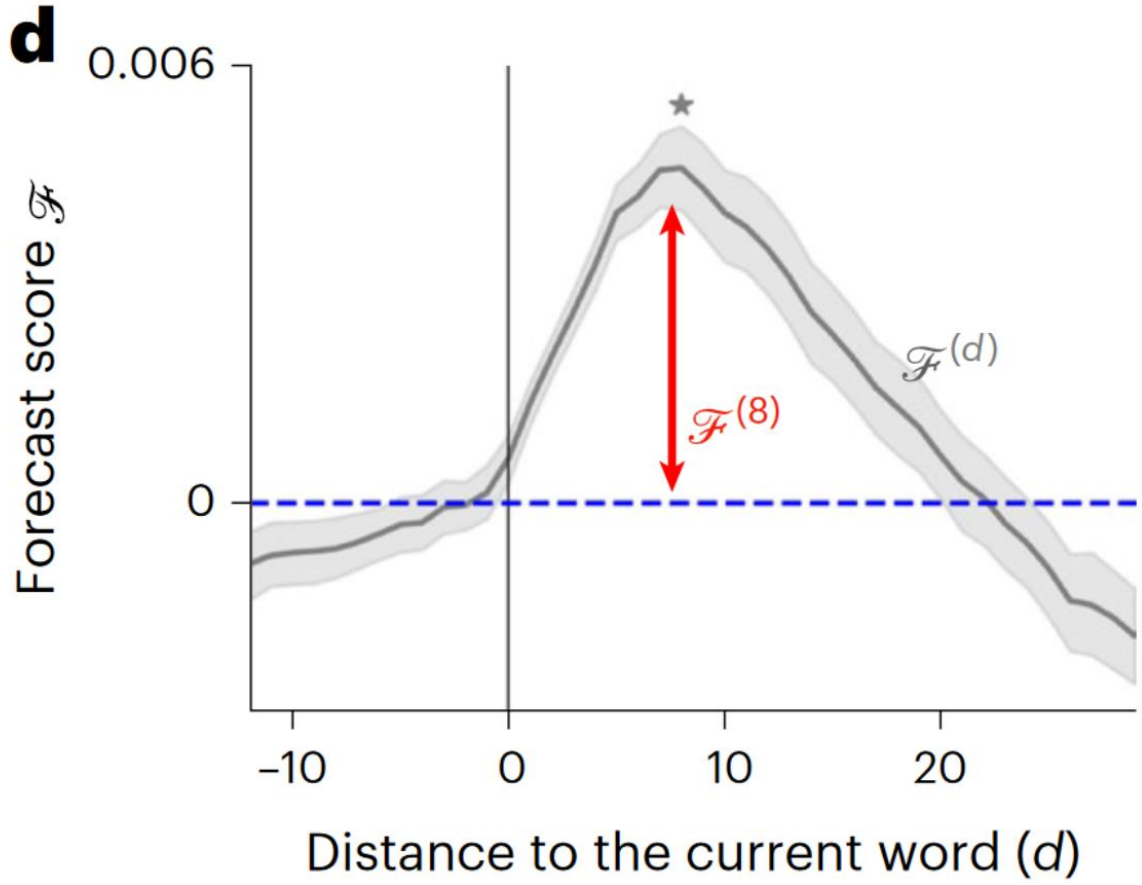

Similarity (R)
.005
.23


Output
GPT-2

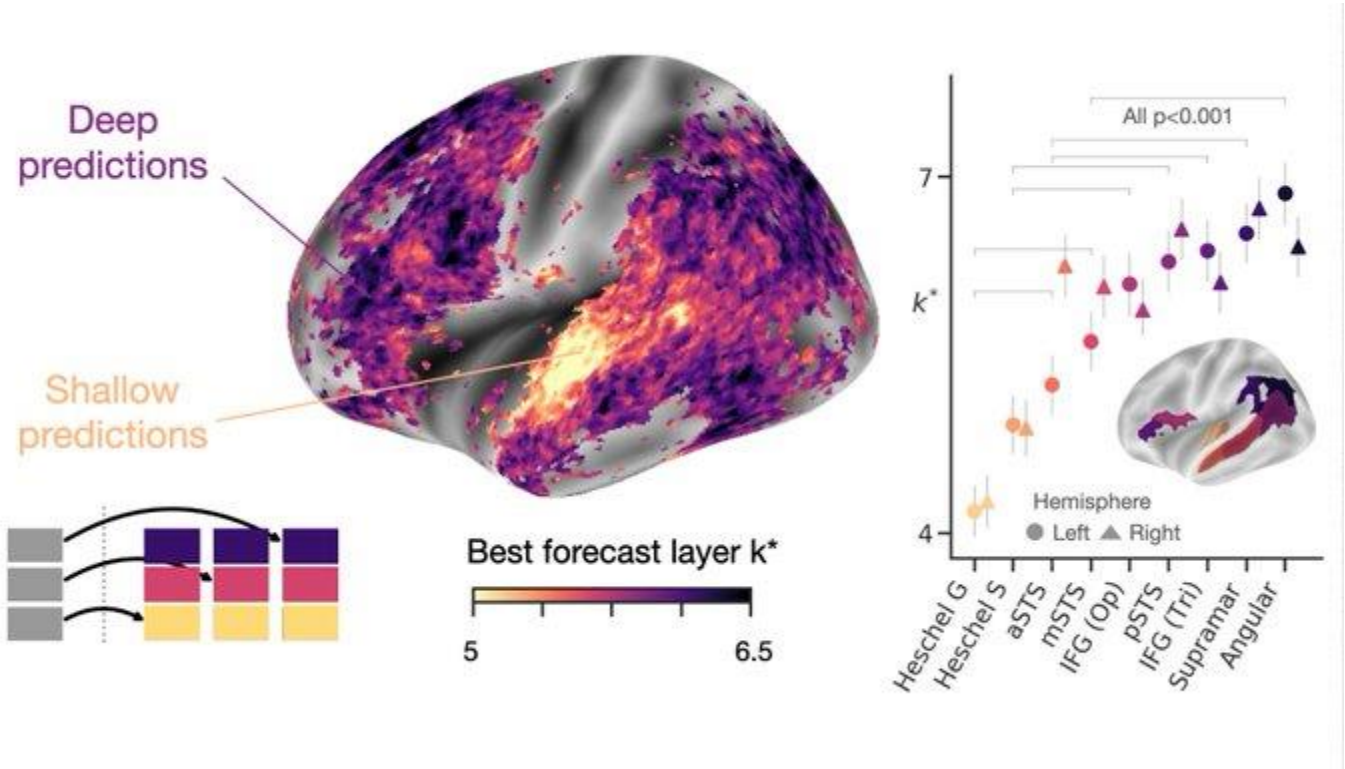

a
Voxel-wise model estimation
Naturally spoken stories were played for 7 subjects
Co-occurrence was found between each word in the stories and 985 common words
Semantic features
Regression weights
BOLD responses
'Above'
'Dance'
'Man'
'Seven'
125 min
125 min
b
Voxel-wise model validation
A new story was then played for each subject
Estimated regression weights were used to predict responses
Semantic features
Model predictions
New BOLD responses
11 min
11 min
Correlation
Prediction performance

Forecast score





Forecast depth



Discussion

„Brain activity is best explained by the activations of deep language algorithms enhanced with long-range and high-level predictions.“

1. the lateral, dorsolateral and inferior-frontal cortices and the supramarginal gyrus exhibited the longest forecast distances.
2. low-level predictions best model the superior temporal sulcus and gyrus, while high-level predictions best model the middle temporal, parietal and frontal areas. More generally, our results support the idea that, unlike current language algorithms, the brain is not limited to predict word-level representations but rather predicts multiple levels of representations
3. Syntax may be explicitly represented in neural activity40,63,64, predicting high-level semantics may be at the core of long-form language processing

→ Results strengthen the role of hierarchical predictive coding in language processing and illustrate how the **synergy between neuroscience and artificial intelligence** can unravel the computational bases of human cognition.