Incorporating Context into Language Encoding Models for fMRI

Shailee Jain & Alexander G. Huth {shailee, huth}@cs.utexas.edu

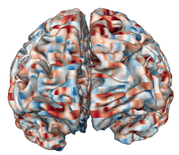
huthlab

THE UNIVERSITY OF TEXAS AT AUSTIN

How does the human brain process language?

'…I reached over and slowly undid my seatbelt…'

?



Stimulus: 11 natural, narrative stories from *The Moth Radio Hour*. (2 hours or ~23,800 words)

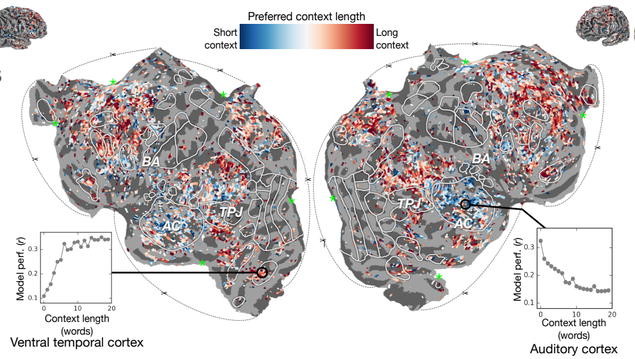
Response: BOLD recorded using 3T fMRI scanner. 6 subjects (2 female) Voxel size: 2.24x2.24x4.1 mm. TR: 2.0045 seconds.

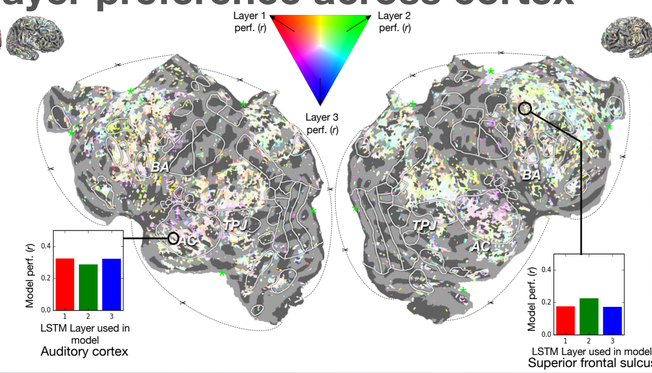
(Huth, 2016 & deHeer, 2017)

Context models reveal timescale & LSTM layer preference across cortex

Low-level areas prefer shorter context.

High-level semantic areas like long sequences.

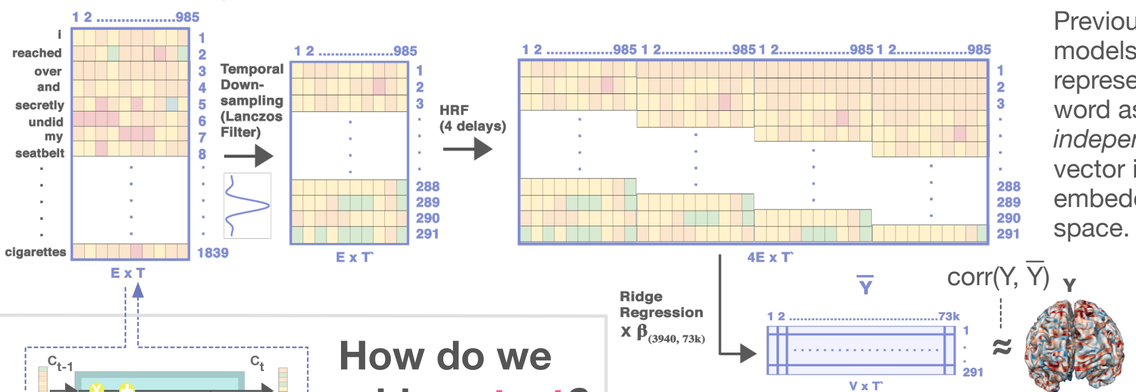




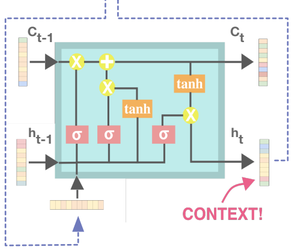
Low-level areas prefer LSTM layers 1&3.

High-level semantic areas prefer LSTM layer 2.

Embedding-space models



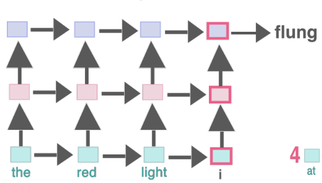
Previous models represent each word as an *independent* vector in an embedding space.



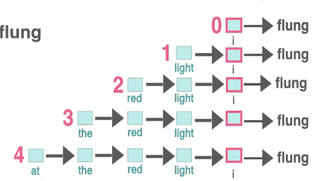
Pre-train LSTM LM on text corpus with 985-d representations.

How do we add context?

3 LSTM Layers



20 Context Lengths



Each word has 3x20 = 60 different context representations from LSTM LM.

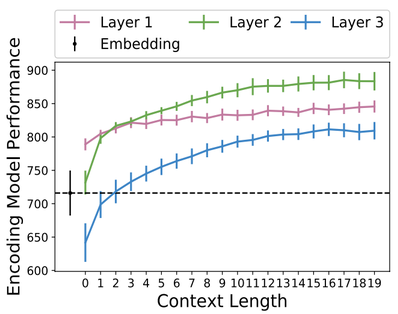
60 Contextual spaces!

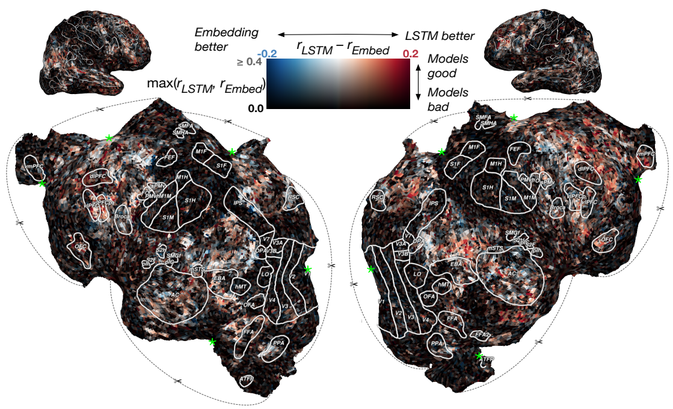
Does context help?

Contextual spaces beat embedding baseline!

LSTM Layers perform differently

Adding *more* context helps overall



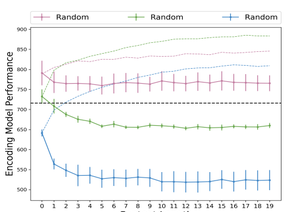


Is it really context?

Distorting context for each word leads to bad representations and poor encoding performance.

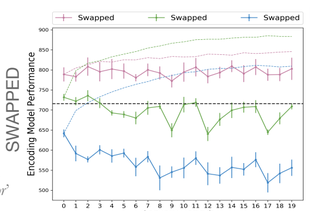
'When his foot hit the brake at the red light I flung open the *door*'

'Wind blue I camp laundry funny universe at woah yellow smoke and on my *door*'



'When his foot hit the brake at the red light I flung open the *door*'

'And I think oh god she lives here how awful we go to the *door*'



Video

Paper