The Role of Context Gating in Predictive Sentence Processing

Yasemin Gokcen, David Noelle, and Rachel Ryskin
Department of Cognitive and Information Sciences
University of California, Merced

Want to keep in touch?
Email: ygokcen@ucmerced.edu
or scan for my LinkedIn!

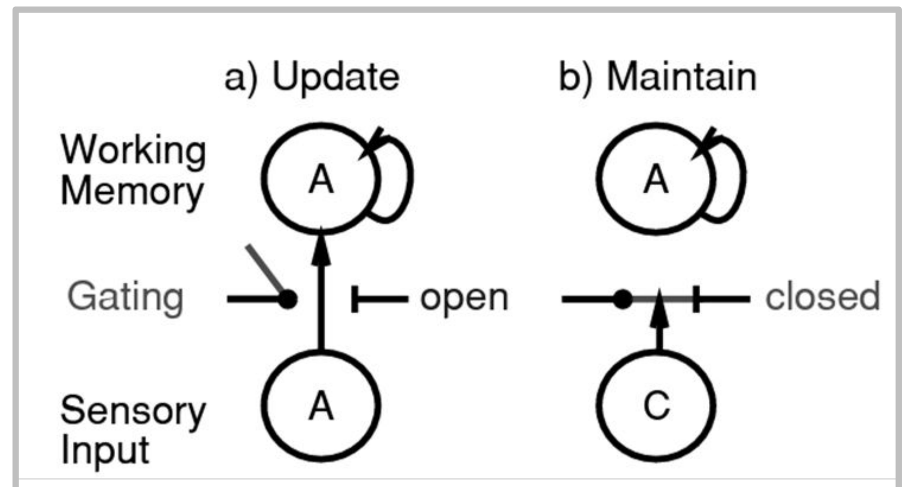
BACKGROUND

Prediction from context during language processing

- Language comprehension involves predicting upcoming input from context (Ryskin & Nieuwland, 2023).
- Surprisal represents the predictability of a word given context (Hale, 2001; Levy, 2008).
- The memory representation of the context is noisy (Hahn et al., 2022).

Working memory & gating

- Models of working memory (WM) mechanisms associated with prefrontal cortex (PFC) use *gating* to maintain and update task-relevant context (O'Reilly et al., 2002).



a) Update
b) Maintain
Working Memory
Gating
Sensory Input
A
A
A
C
open
closed

O'Reilly & Frank, 2006

- Memory is *updated when gate opens* to *relevant information* and maintained when gate closes to irrelevant information.
- Reinforcement learning makes gating decisions.
- The anatomy and physiology of the PFC are thought to support this mechanism (Noelle, 2012; Kriete et al., 2013).

QUESTION

Are the gating mechanisms of prefrontal cortex used to maintain context for prediction during language processing?

GENERAL APPROACH AND METHODOLOGY

- Examine language models that vary in their use of gating.
- Compare them with regard to performance and ability to predict neural measures of real-time language processing.

Language Models:

- Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models (van Schjindel & Linzen, 2018)
- **LSTM has PFC-like gating mechanisms**, RNN does not (Hochreiter & Schmidhuber, 1997)
- Trained on the Wikitext-2 dataset, tested on Natural Stories corpus (Futrell et al., 2021).
- Trained 5 replications of each model with different random initializations to ensure robustness.

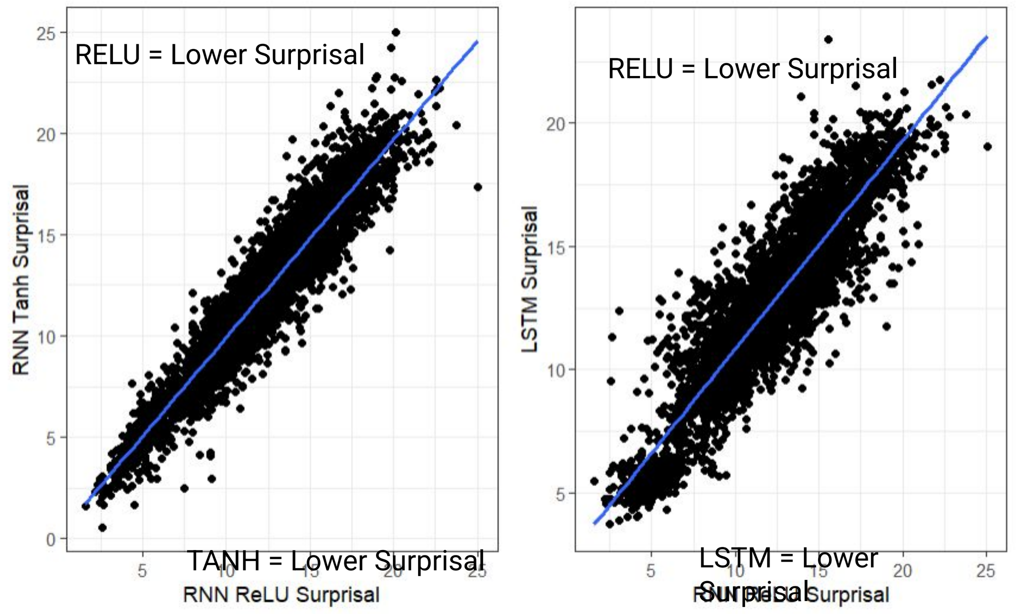
Neural activity:

- EEG passive listening task
 - Natural Stories corpus (Futrell et al., 2021)
 - 10 stories/narratives (~5 min each) featuring low frequency words/structures
 - 32 participants (19 female, mean age = 21.65 years old)
- Analysis
 - LMMs predicting average EEG amplitude from 16 centro-posterior channels in 200 ms windows

```
MeanAmp ~ Frequency + PreviousFrequency + CosineSimilarity +
PreviousCosineSimilarity + GPT Surprisal + PreviousSurprisal + WordDuration +
SentencePosition + StoryPosition +(1|Story)+ 1 + CosineSimilarity + GPT
Surprisal ||Subject)
```

A METRIC OF GATING UTILITY FOR PREDICTION

- High correlation (r = .949) in word-by-word surprisal when comparing between RNN models (Tanh vs. ReLUactivation functions) and slightly lower between LSTM and RNN (r = .893).
- **Gating utility metric: Word-by-word residual from a linear model predicting LSTM surprisal from RNN surprisal (negative value means LSTM has lower surprisal than RNN)**



RELU = Lower Surprisal
TANH = Lower Surprisal
RNN Tanh Surprisal
RNN ReLU Surprisal
RELU = Lower Surprisal
LSTM = Lower Surprisal
LSTM Surprisal
RNN ReLU Surprisal

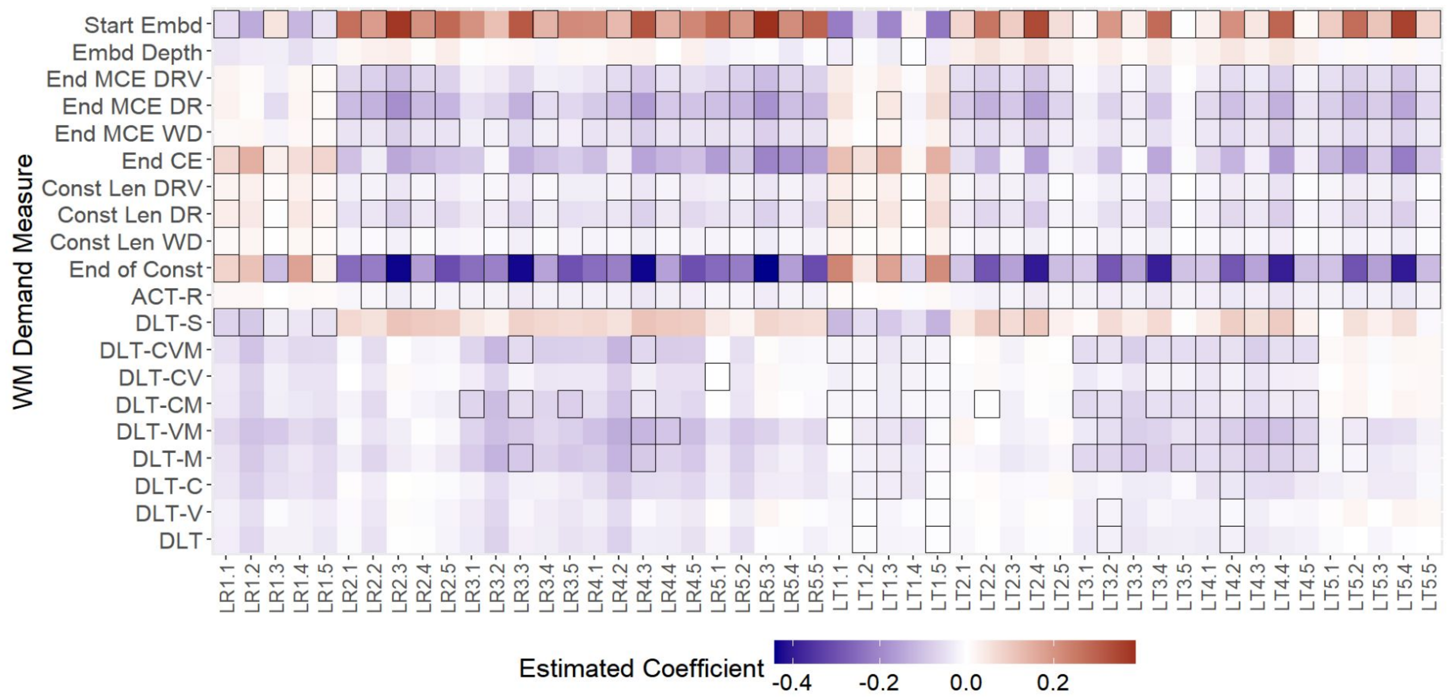
One dot = one word appearance

Words for which LSTM surprisal is lower than RNN surprisal may indicate sentence contexts where gating plays an important role due to WM demands.

GATING UTILITY IS CORRELATED WITH WM DEMANDS

WM demand measures from 3 frameworks (Shain et al., 2022):

- **DLT**: Dependency Locality Theory (Gibson, 2000)
 - Greater distance between words in a dependency incurs greater WM cost.
- **ACT-R**: Adaptive Character of Thought—Rational (Anderson et al., 2004)
 - Provides a measure of WM retrieval costs and does not consider storage/maintenance.
- **LCP**: Left-Corner Parsing (Lewis & Vasishth, 2005)
 - WM demand increases as simultaneous processing of embeddings and constituents within a sentence increases.



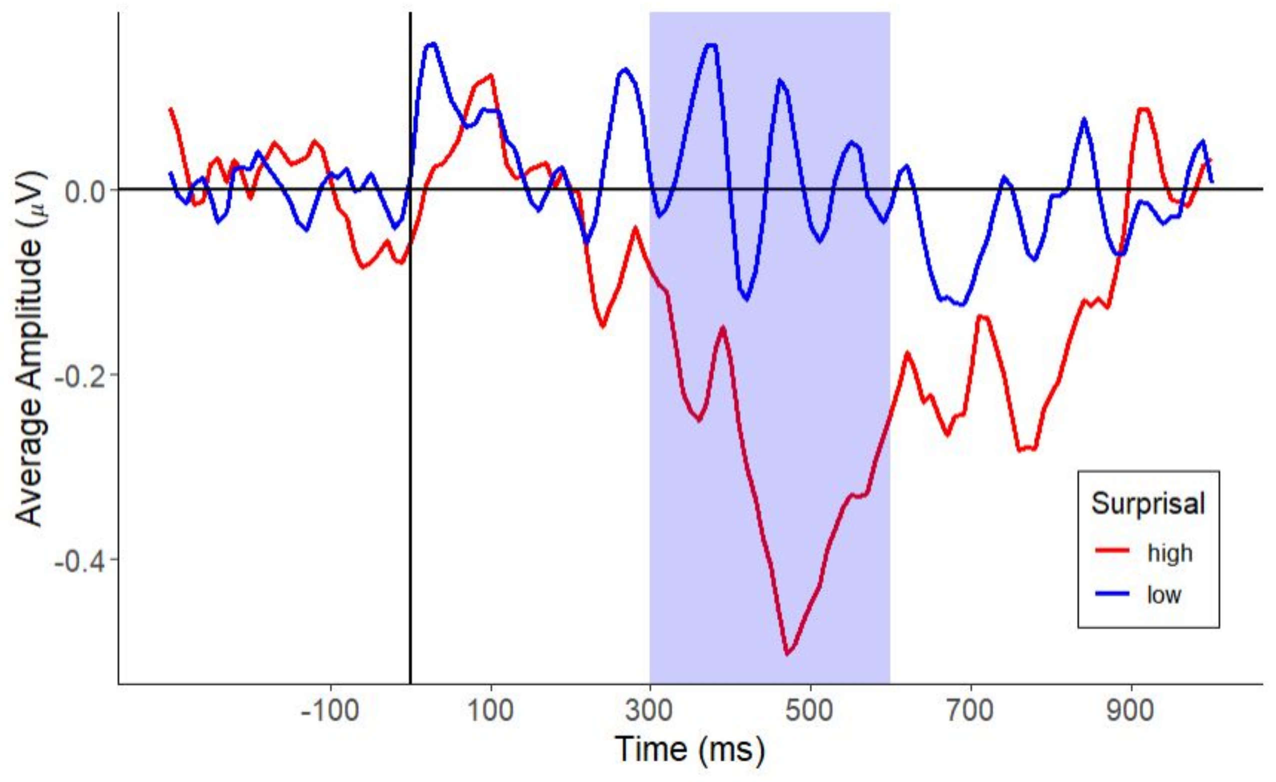
WM Demand Measure
Start Embd
Embd Depth
End MCE DRV
End MCE DR
End MCE WD
End CE
Const Len DRV
Const Len DR
Const Len WD
End of Const
ACT-R
DLT-S
DLT-CVM
DLT-CV
DLT-CM
DLT-VM
DLT-M
DLT-C
DLT-V
DLT
Estimated Coefficient
-0.4
-0.2
0.0
0.2

- Each cell represents the β coefficient from a linear model predicting gating utility from a WM demand measure.
- Columns = model pair from which gating utility is obtained (e.g., LR2.3 = LSTM 2 ~ RNN ReLU 3)
- Black border = statistically significant after Bonferroni correction
- Most measures of WM demand are negatively associated with the residual measure of gating utility.

As WM demand increases, LSTM models are more likely to outperform RNNs at next-word prediction, potentially due to the presence of gating which may be important for context memory.

SURPRISAL IS RELATED TO N400 ERP EFFECT

- Fit LMMs predicting word-by-word N400 amplitude (averaging over 300-600ms)
- Surprisal values from OpenAI's GPT-Davinci



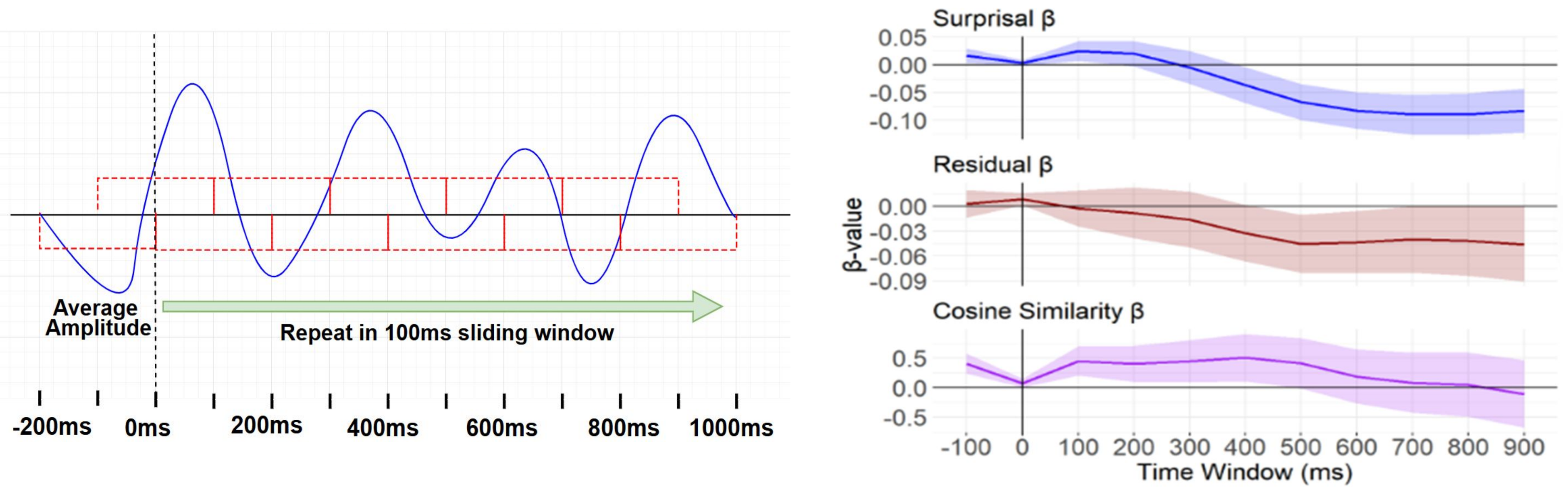
Average Amplitude (μV)
0.0
-0.2
-0.4
-100
100
300
500
700
900
Time (ms)
Surprisal
high
low

Surprisal and cosine similarity only significant predictors of N400

Words that are higher in surprisal elicit more negative N400, replicating prior findings in a naturalistic auditory dataset (Frank, Otten, Galli, & Vigliocco, 2015; Szewczyk & Federmeier, 2022; Michaelov & Bergen, 2024)

TIME COURSE OF EFFECT OF GATING ON NEURAL RESPONSE

- Fit LMMs with **Gating Utility** as added predictor across the word epoch using a sliding window



Average Amplitude
Repeat in 100ms sliding window
-200ms
0ms
200ms
400ms
600ms
800ms
1000ms
Surprisal β
Residual β
Cosine Similarity β
β-value
Time Window (ms)

- Surprisal effect becomes significantly negative around 300-500 ms, aligning with N400 time window.
- Gating utility effect on average amplitude becomes significantly negative starting in 400-600 ms window.
- Cosine similarity with preceding word has early positive effect on amplitude

DISCUSSION & FUTURE DIRECTIONS

- PFC-style gating mechanism may critically support prediction in context during sentence processing in humans.
- Recruitment of PFC-style gating reflected in neural response to word presentation in context.

- Future directions: Explore topography of gating utility effects on ERP and test overlap with P600 component (associated with cognitive control).