Constituent-Constrained Prediction during Speech Comprehension

Jiajie Zou[1,4] (jiajie.zou@esi-frankfurt.de); **David Poeppel**[2,3] (dp101@nyu.edu); **Nai Ding**[1] (ding_nai@zju.edu.cn)

[1]Zhejiang University, China; [2]New York University, USA; [3]Max Planck Institute for Empirical Aesthetics, Germany; [4]Ernst Strüngmann Institute for Neuroscience, Germany

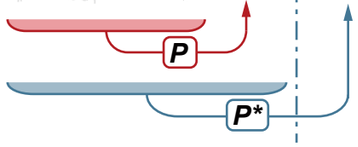
Poster & More

Scan me

Introduction

Continuously predicting upcoming word has been hypothesized as the central computational mechanism in human language comprehension[1,2]. **We investigate whether word prediction is additionally constrained by the major constituent structure.**

[*the boy with red shorts*] [*kicked the ball*]



P and P* are different prediction mechanisms. P* is less precise than P

Stimuli & Recording

Stimuli: Each trial contains ten ***Mandarin Chinese*** sentences. To isolate potential constituent boundary effects, we

- ***removed prosody*** *by independently synthesizing each word*[3].
- ***eliminated statistical cues*** *by including both statistically typical and atypical sentences.*

***Exp 1*: Sentences of variable duration**

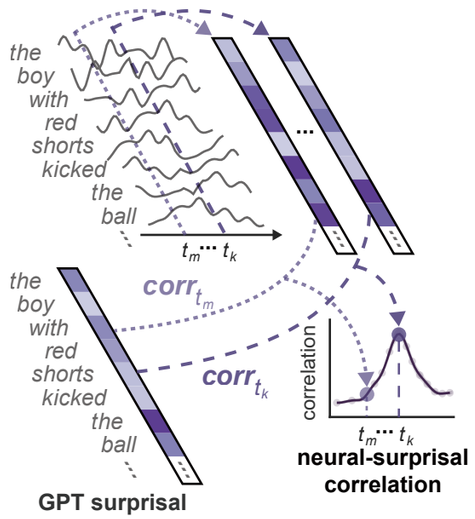
***Exp 2*: Blocks of sentences of fixed duration**

***Exp 3*: Mixed trials of sentences and disconnected phrases**

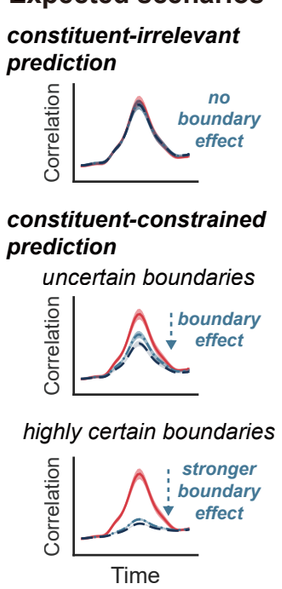
• non-initial ■ phrase-initial ★ sentence-initial

MEG Analyses & Hypotheses

Evaluation of prediction

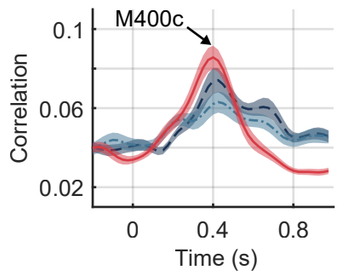


Expected scenarios

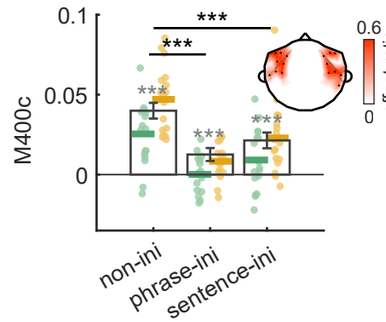


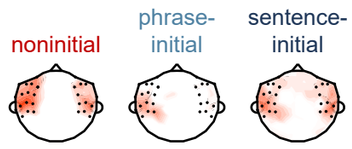
Exp 1: Prediction & Constituent Boundary Effect

Boundaries weaken neural-surprisal correlation



Boundary effect is not confounded by statistical cues





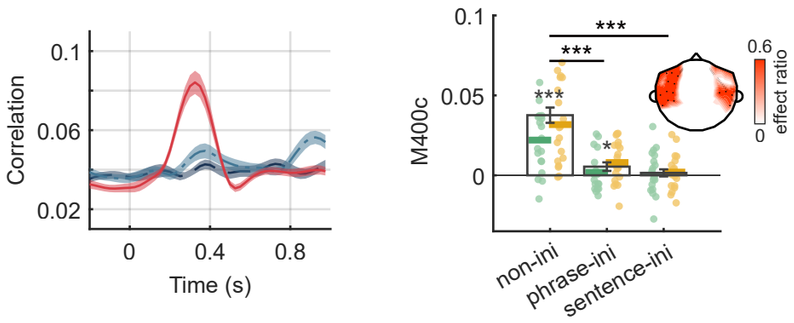
statistically **typical** sentences (*"many football teams **aimed** to win the title"*)

statistically **atypical** sentences (*"streams of tears **flowed** down the cheeks"*)

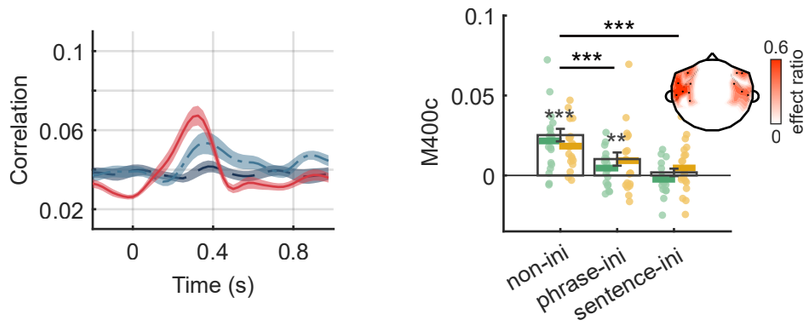
The brain more precisely predicts an upcoming word when the word is expected to continue a constituent.

Exps 2 & 3: Influence of Boundary Certainty

***Exp 2*: When boundary is more certain due to the fixed constituent duration, the boundary effect strengthens.**



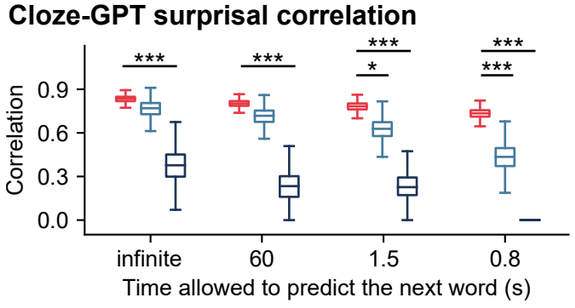
***Exp 3*: When boundary certainty is reduced by mixing non-sentence trials, the boundary effect weakens.**



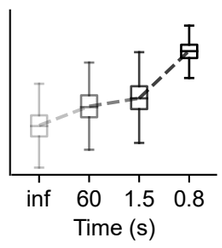
The constituent-boundary effect is stronger when constituent boundaries are more expected.

Behavioral Cloze Test & Influence of Time

Cloze-GPT surprisal correlation



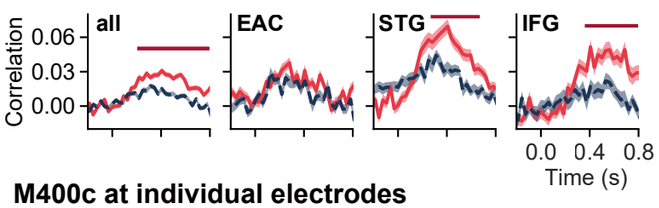
Progressive boundary effect



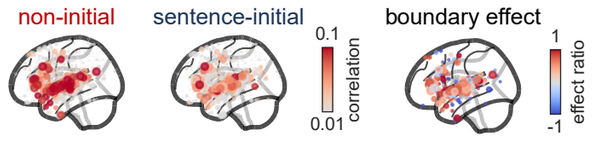
The constituent-boundary effect is observed at the behavioral level and results from limits on processing time

Boundary Effect for English Narratives (ECoG)[4]

Neural-surprisal correlation in ROIs



M400c at individual electrodes



The constituent-boundary effect is observed in both Mandarin and English, and the effect originates from STG and IFG

Conclusions

Constituent structure constrains and guides word prediction during real-time language comprehension. This constituent-boundary effect is not driven by prosodic or statistical cues, is stronger at more expected boundaries, and generalizes across Mandarin and English.

References

1. Schrimpf et al., 2021. *The neural architecture of language: Integrative modeling converges on predictive processing.*
2. Goldstein et al., 2022. *Shared computational principles for language processing in humans and deep language models.*
3. Ding et al., 2016. *Cortical tracking of hierarchical linguistic structures in connected speech.*
4. Zada et al., 2025. *The 'Podcast' ECoG dataset for modeling neural activity during natural language comprehension.*