

# Error Sensitivity of Humans and Language Models in Incremental Language Processing

Siyuan Song<sup>1</sup>, Thomas Hikaru Clark<sup>2</sup>

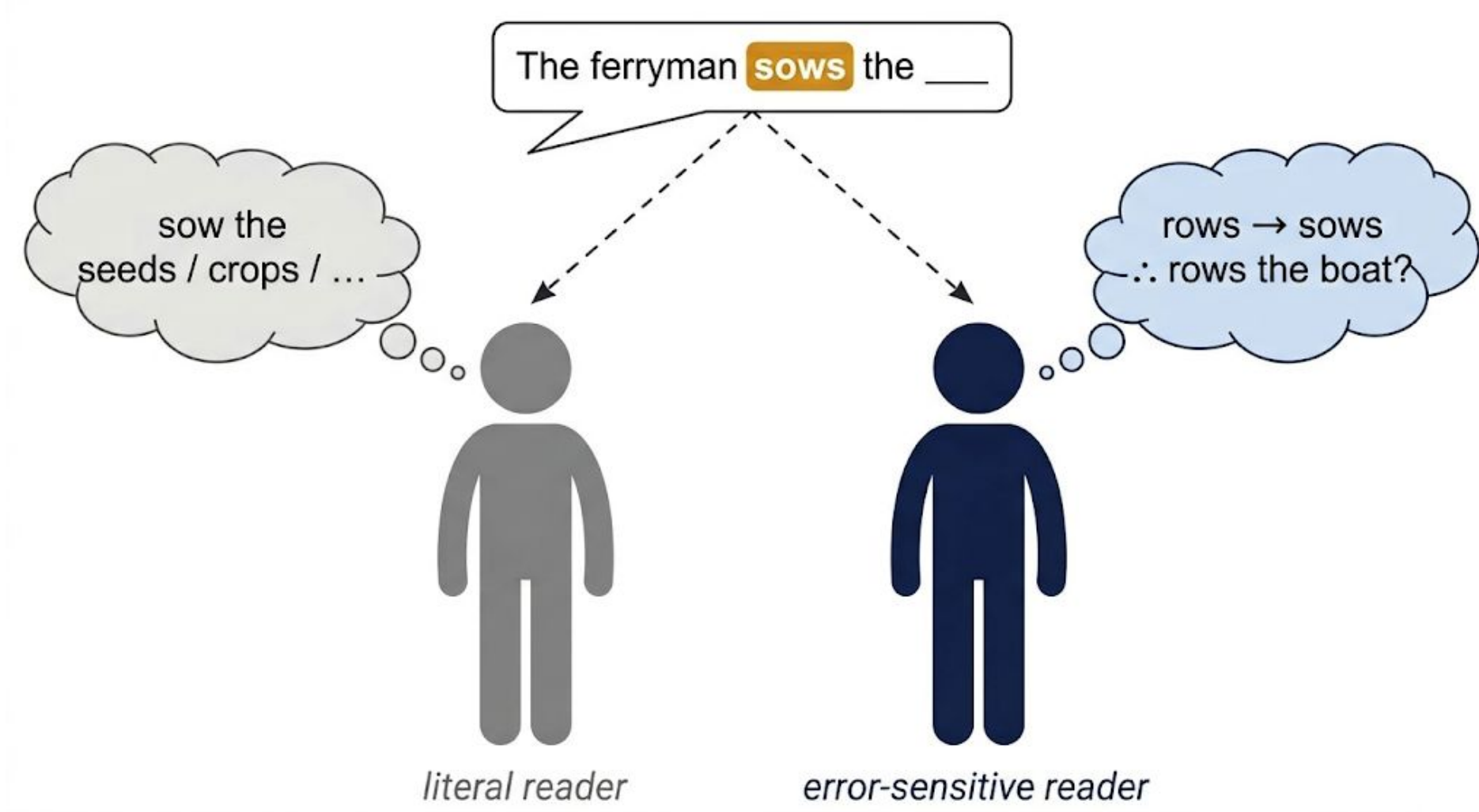
<sup>1</sup>Department of Linguistics, The University of Texas at Austin <sup>2</sup>Department of Brain and Cognitive Sciences, MIT



## Background

- Previous research on noisy-channel language processing [1,2,3] has typically focused on processing at the location of the error (e.g. longer RTs for typos).
- LLMs have some error-correcting abilities but also have a strongly incremental bias [4, 5].

**Question:** Do comprehenders show sensitivity to possible errors in their expectations about upcoming words?



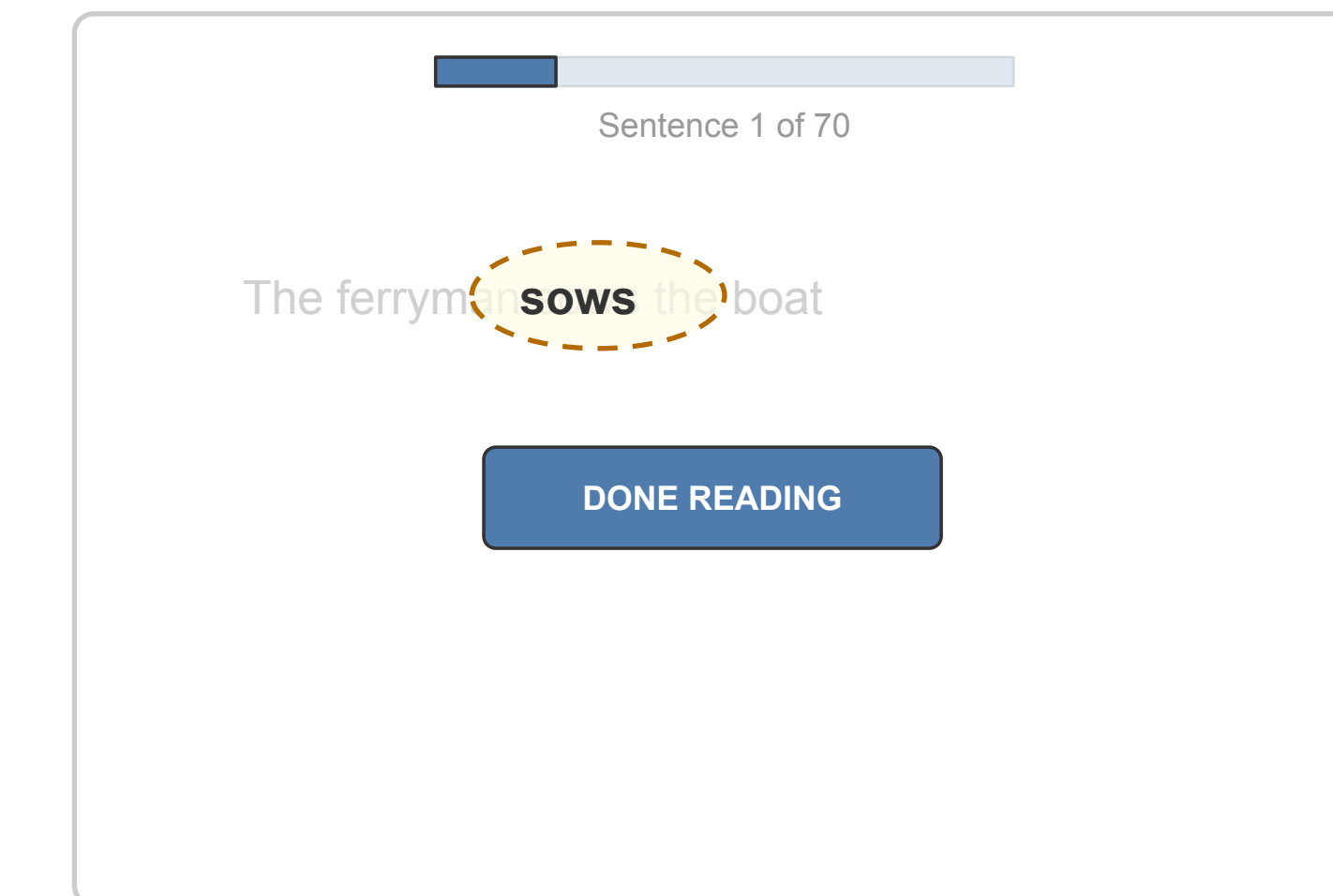
## Experiment Setup

**Materials** 3\*3\*5 design; [context] [subject] [verb] [object]

[context]	[subject]	[verb]	[object]
/	he (neutral)	rows	clean
on the river	the ferryman	sows	neighbor real-word
in the field	the farmer	plants	non-neighbor real-word
		aows	neighbor non-word
		hamnts	non-neighbor non-word

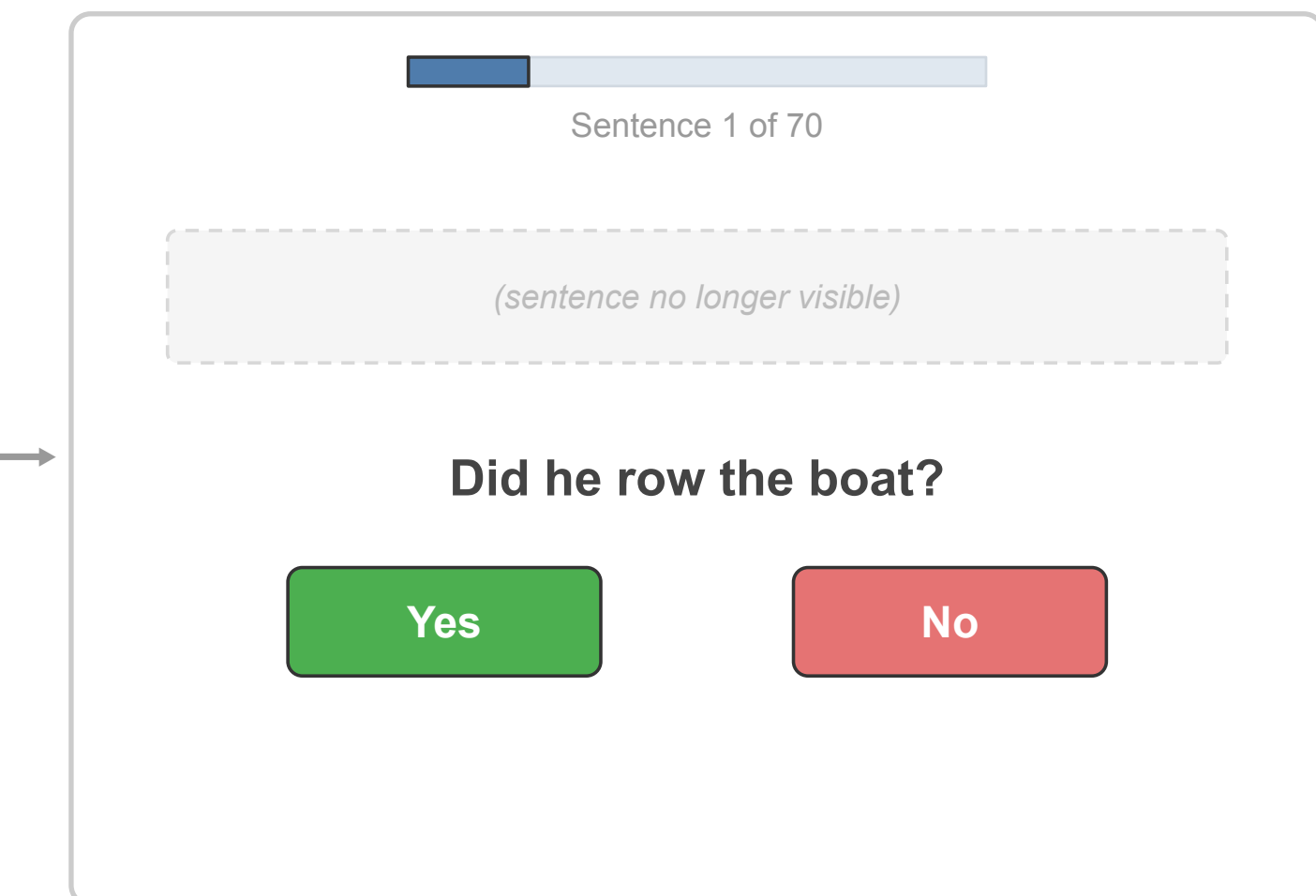
**Mouse-Tracking Experiment (MoTR [6])**

Step 1: Reading



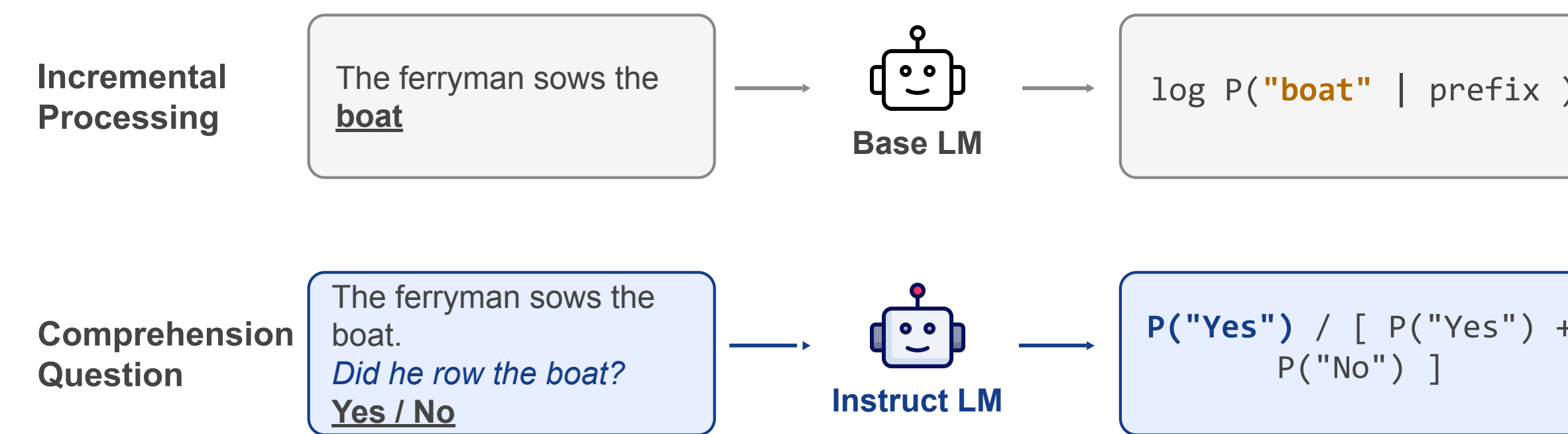
Measure: gaze duration on each word (from mouse tracking)

Step 2: Comprehension Question

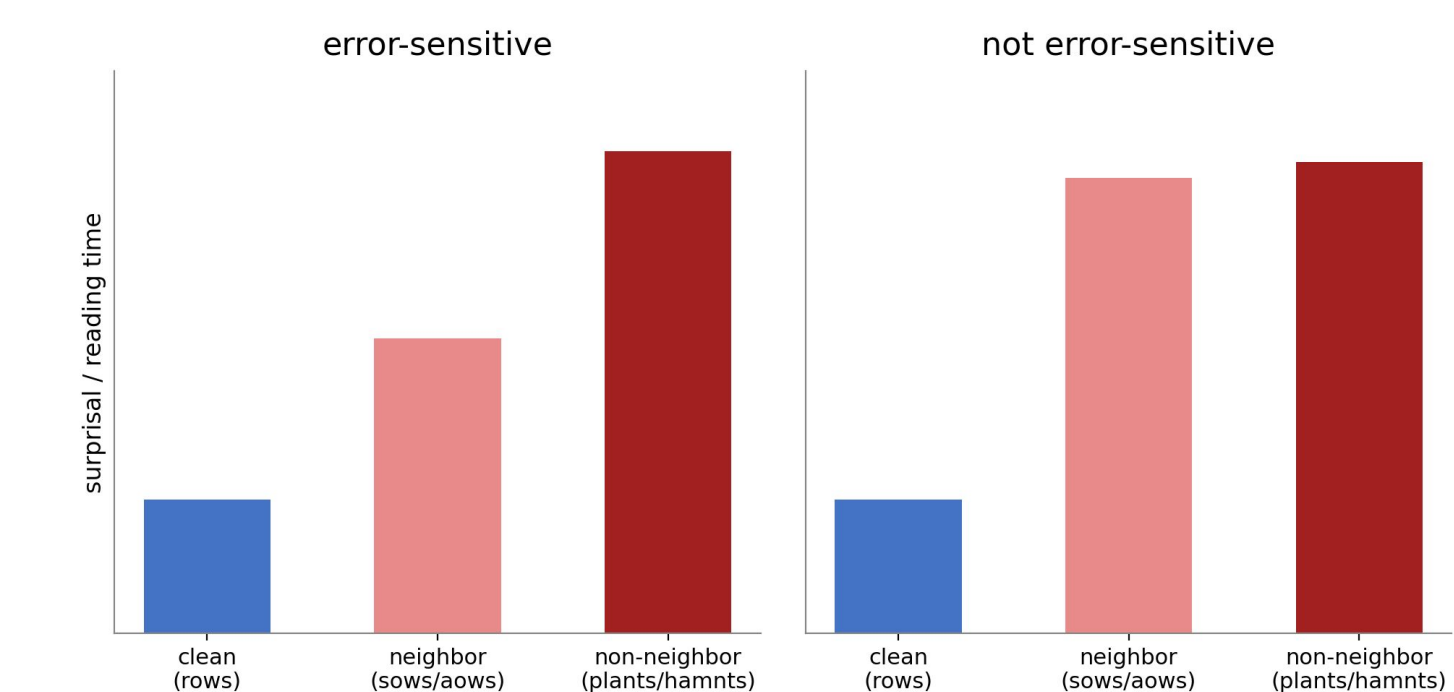


Measure: proportion of "Yes" responses

**Experiments on LLMs**

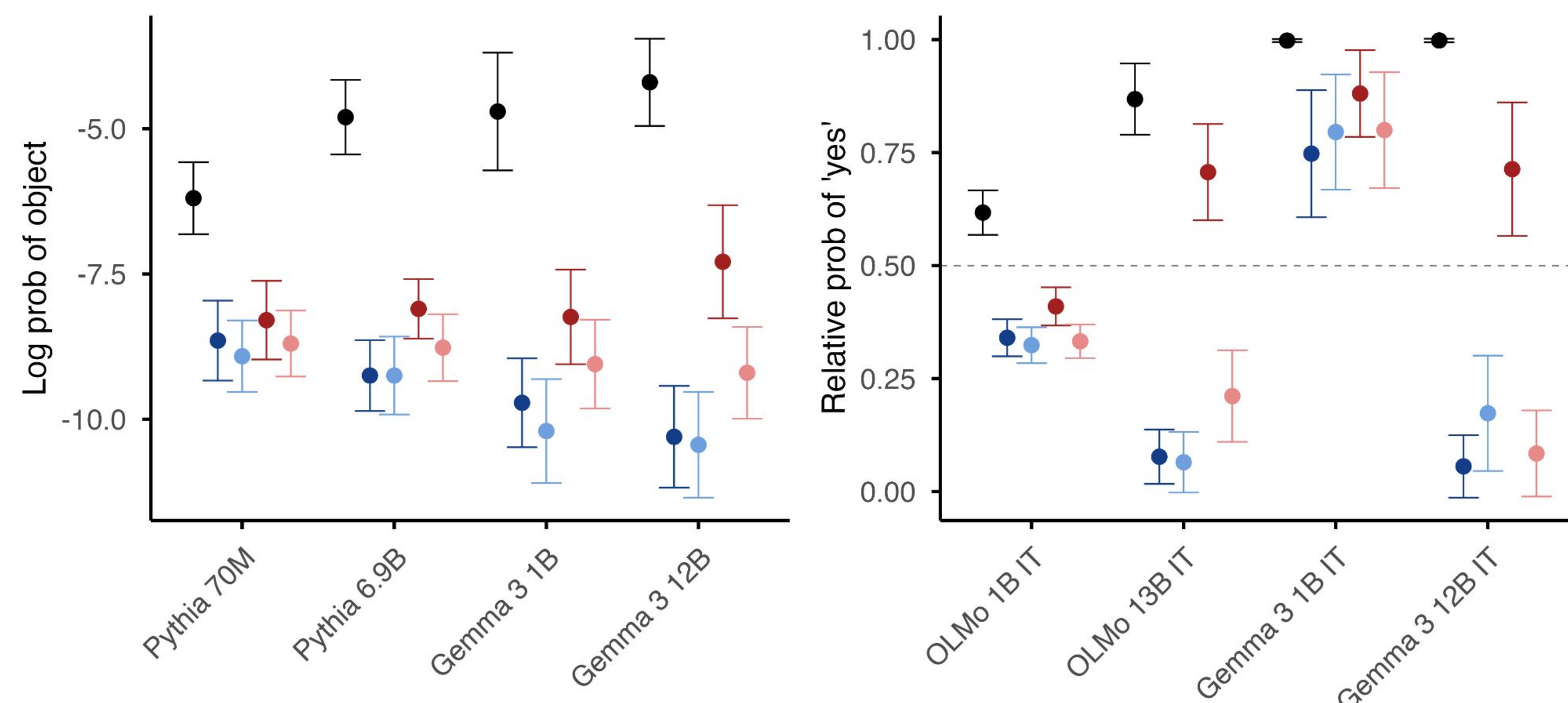


Error-sensitive readers should show less surprise at the object word following **neighbor errors** than **non-neighbor errors**, as correctable errors allow inference of the intended verb.

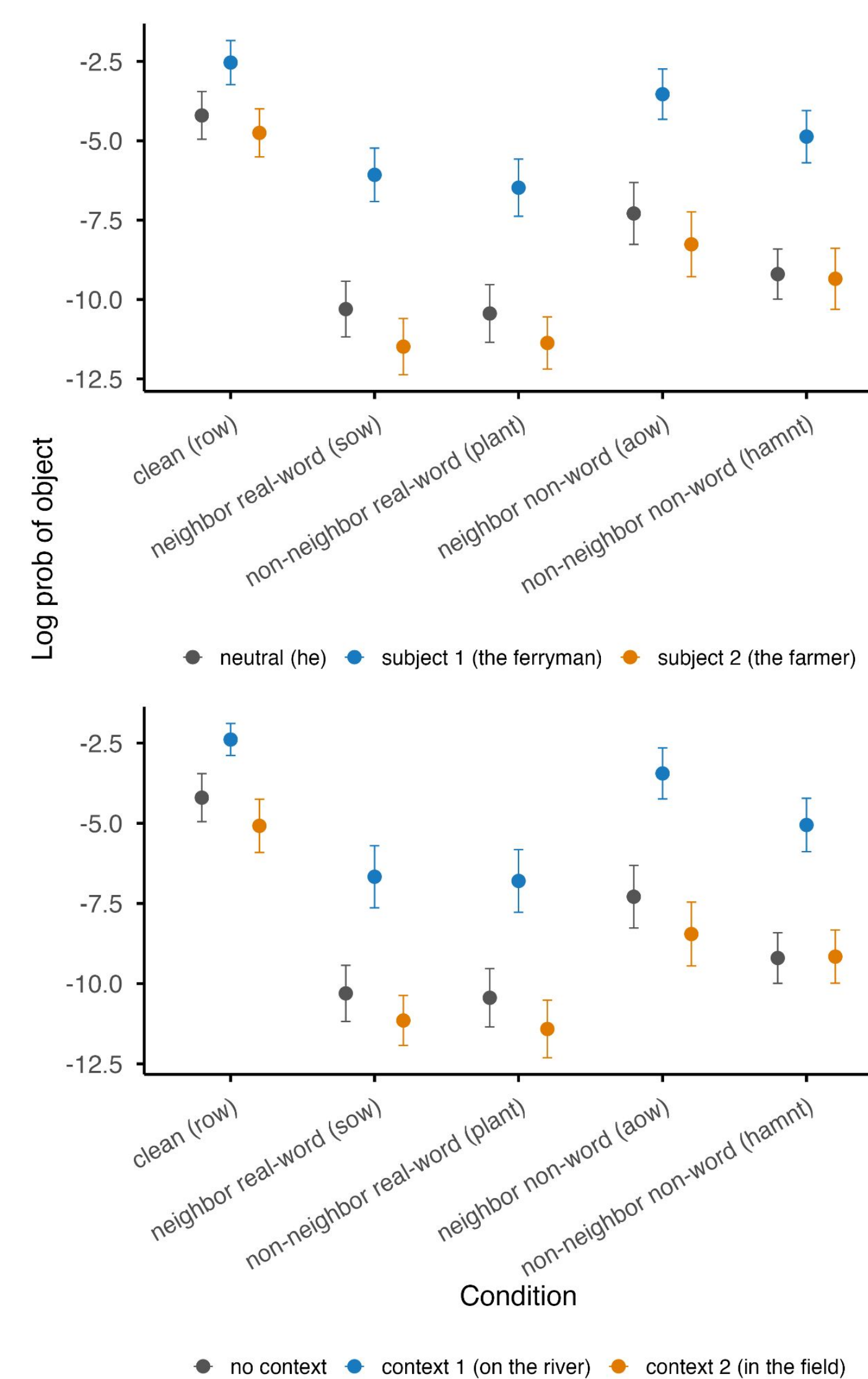


## Results

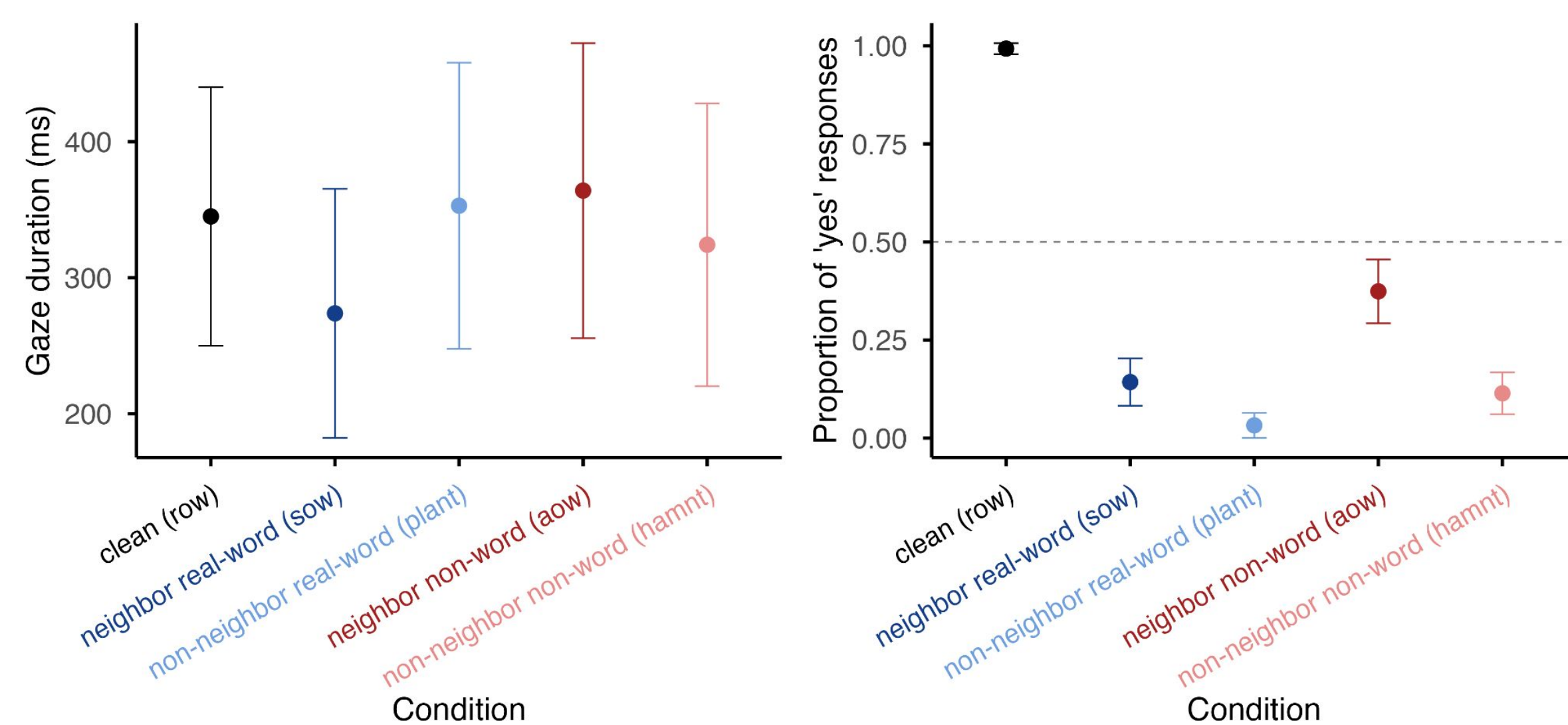
**1. LLMs are only sensitive to non-word errors**



**3. Subject and context modulate LM predictions as expected**



**2. Humans are not sensitive to errors in next word prediction**



## Key Takeaways

1. LMs show limited error sensitivity — only for non-word errors.
2. Humans show error sensitivity in comprehension but not in incremental processing.
3. Humans and LMs diverge in how they handle real-word errors.

## Future Work

1. LMs detect errors only when input is implausible. → Investigate mechanisms behind LM error processing and improve robustness to typos in human-AI interaction.
2. Humans show error sensitivity in comprehension but not incrementally. → Identify factors driving error-sensitive behavior and regression during reading.

[1] Levy, Roger. "A noisy-channel model of human sentence comprehension under uncertain input." Proceedings of the 2008 conference on empirical methods in natural language processing. 2008.  
 [2] Gibson, Edward, Leon Bergen, and Steven T. Piantadosi. "Rational integration of noisy evidence and prior semantic expectations in sentence interpretation." Proceedings of the National Academy of Sciences 110.20 (2013): 8051-8056.  
 [3] Ryskin, Rachel, et al. "An ERP index of real-time error correction within a noisy-channel framework of human communication." Neuropsychologia 158 (2021): 107855.  
 [4] Zhang, Yunxiang, et al. "Interpreting the robustness of neural NLP models to textual perturbations." Findings of the Association for Computational Linguistics: ACL 2022. 2022.  
 [5] Gan, Esther, et al. "Reasoning robustness of llms to adversarial typographical errors." Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024.  
 [6] Wilcox, Ethan Gottlieb, et al. "Mouse Tracking for Reading (MoTR): A new naturalistic incremental processing measurement tool." Journal of Memory and Language 138 (2024): 104534.  
 This cover has been designed using resources from [Flaticon.com](https://flaticon.com). Illustrative figures are created with the help of Claude and Gemini.