Predicting long context effects using surprisal

Georgia-Ann Carter
University of Edinburgh, Edinburgh, United Kingdom

Frank Keller
University of Edinburgh, Edinburgh, United Kingdom

Paul Hoffman
University of Edinburgh, Edinburgh, United Kingdom

Abstract

We know that context influences the facilitation of language comprehension. Previous research has shown that discourse coherence influences this contextual facilitation, with comprehenders making stronger predictions about upcoming words when reading highly coherent narratives. However, it is unclear whether the predictions made by Large Language Models (LLMs) exhibit similar discourse-level influences. As such, we investigate whether surprisal values from LLMs reflect longer context effects. We calculated word-level surprisal values (as a measure of prediction strength) for passages that vary in coherence. We used these to predict human reading times for the same passages collected from 289 participants. We found that surprisal only predicted reading times early in the target sentence, and that GPT-2's surprisal values were not influenced by discourse coherence, in contrast to human reading data. This has implications on the use of Transformer-based LLMs in modelling human cognition.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*