

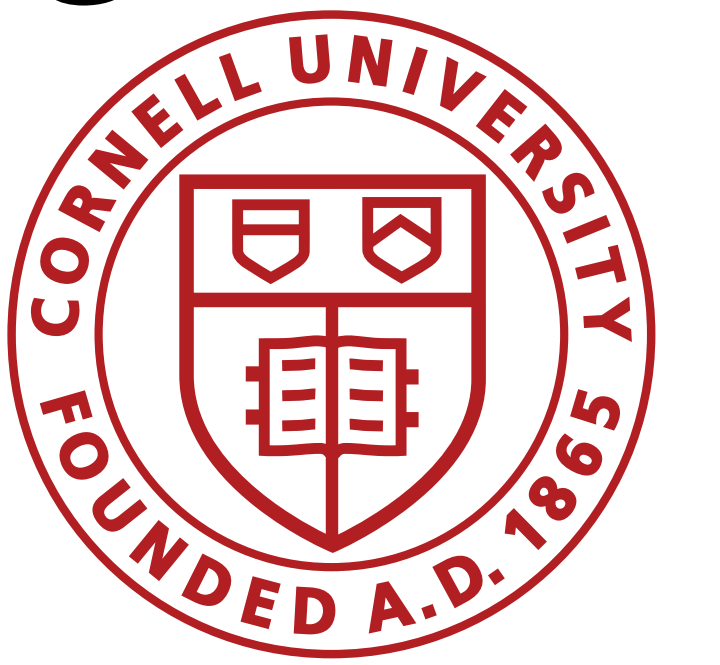
# Interaction with Context During Recurrent Neural Network Sentence Processing

Forrest Davis and Marten van Schijndel

Cornell University

fd252@cornell.edu

42nd Annual Meeting of the Cognitive Science Society (CogSci 2020), July 29–August 1, 2020



## Background

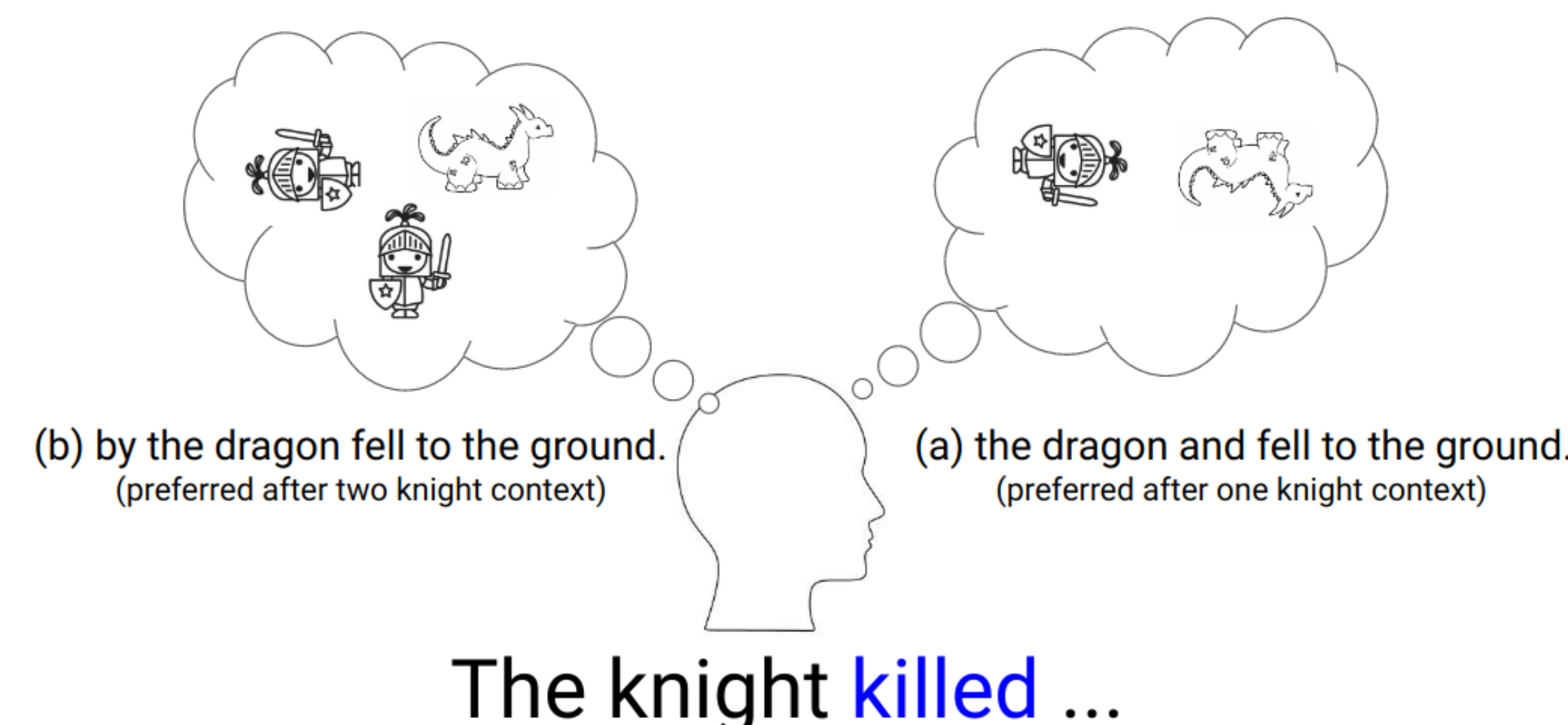
The knight killed ...

- (a) the dragon and fell to the ground.
- (b) by the dragon fell to the ground.

- In *isolation*, (b) can lead to a reading time slow down, known as a **garden path effect**

## Context alleviates garden-path effects

Humans use *discourse context* to **accommodate** conversational implicatures of the relative clause in (b) (Altmann and Steedman, 1988; Spivey-Knowlton et al, 1993; Trueswell and Tanenhaus, 1991)



- Referent contexts (i.e. contexts with two referents) alleviate garden-path effects
- Temporal contexts (i.e. future contexts) alleviate garden path effects
- Information structure and definiteness (i.e. discourse old definites) alleviate garden path effects

## Neural Network Language Models

- 10 RNN LMs trained on 80 million words of Wikipedia text
- 5 trained on data shuffled by sentence removing discourse cues (e.g., Gulordava et al. 2018); 5 on sentences in their original order
- Measured information-theoretical surprisal (Hale, 2001; Levy, 2008)

$$\text{surprisal}(w_i) = -\log P(w_i|w_1...w_{i-1})$$

## Research Question

How much pragmatic knowledge is contained in the linguistic signal? Can human-like garden path alleviation be learned from linguistic data alone, using recurrent neural network (RNN) language models (LMs)?

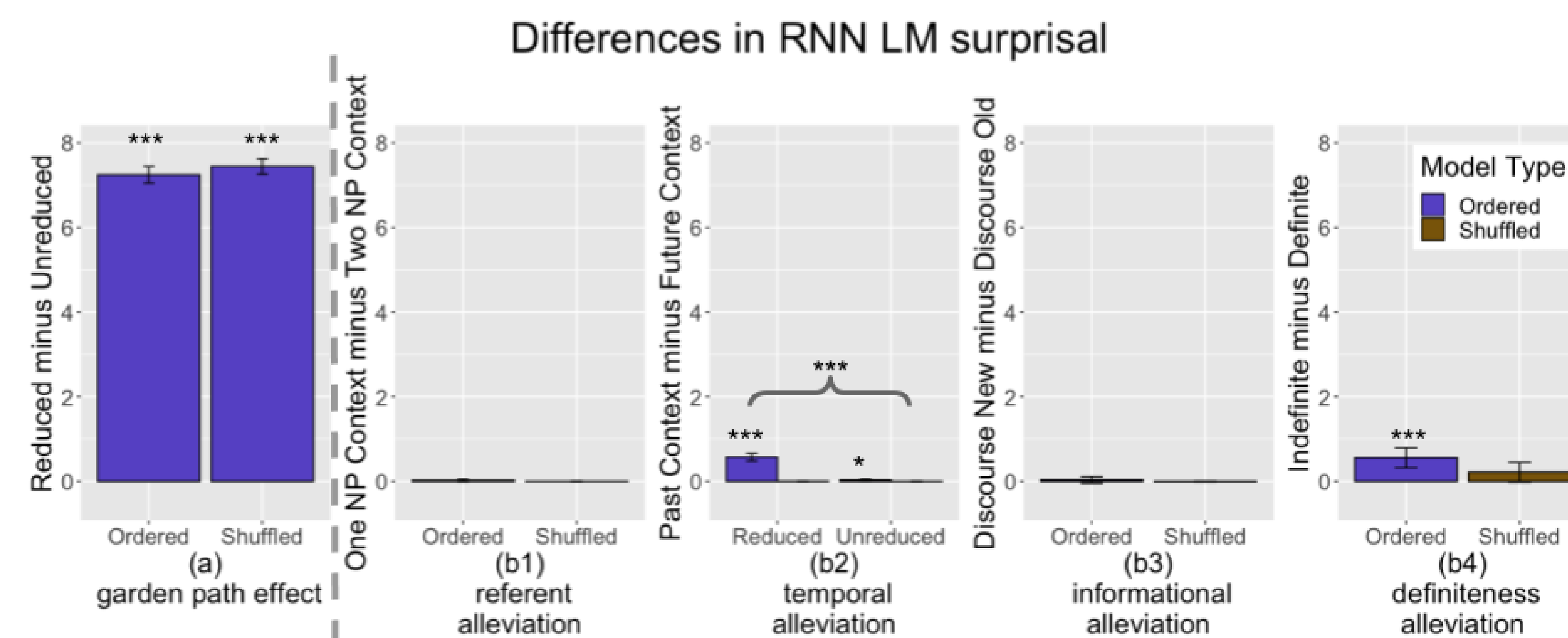


Figure 1: Differences between RNN LM surprisals. (a) between the full relative clause region in a reduced versus unreduced relative clause (garden path effect). (b1-4) when the critical region is preceded by a garden-path supporting context ((b1) contexts with a single referent, (b2) past contexts, (b3) discourse old contexts, and (b4) definite) and when preceded by a garden-path alleviating context ((b1) contexts with two referents, (b2) future contexts, (b3) discourse new contexts, and (b4) indefinite).

## Referent Stimuli

- Context**
  - 1NP - A knight and his squire were attacking a dragon. With its breath of fire, the dragon killed the knight but not the squire.
  - 2NP - Two knights were attacking a dragon. With its breath of fire, the dragon killed one of the knights but not the other.
- Target**
  - Reduced - The knight killed by the dragon fell to the ground with a thud.
  - Unreduced - The knight who was killed by the dragon fell to the ground with a thud.

from Spivey-Knowlton et al (1993)

## Temporal Stimuli

- Context**
  - Past - Several students were sitting together taking an exam in a large lecture hall earlier today. A proctor noticed one of the students cheating.
  - Future - Several students will be sitting together taking an exam in a large lecture hall later today. A proctor will notice one of the students cheating.
- Target**
  - Reduced - The student spotted by the proctor received/will receive a warning.
  - Unreduced - The student who was spotted by the proctor received/will receive a warning.

from Trueswell and Tanenhaus (1991)

## Conclusion

- Both training conditions (i.e. with or without discourse contexts) exhibit garden path effects
- Temporal contexts and definiteness reduce the garden-path effect for only the RNN LMs trained on ordered data
- Aspects of multi-sentence discourse effects on sentence processing can be acquired from linguistic data alone

## Discussion

- RNN LMs learn some pragmatic uses of tense (i.e. discourse about the future is likely to continue referring to the future) and definiteness (i.e. indefinites are more likely than definites to be subjects)
- Points to two alternatives: 1) some kinds of garden path alleviation may not require pragmatic inference, or 2) some relevant aspects of pragmatic reasoning are learnable from just language data.

## Information Structure and Definiteness Stimuli

- Indefinite+New**

It was a fun evening. Even though the band left (.) **a party** went on for at least another two hours.
- Definite+New**

It was a fun evening. Even though the band left (.) **the party** went on for at least another two hours.
- Definite+Old**

A party was organized this evening. Even though the band left (.) **the party** went on for at least another two hours.

from Grodner et al. (2003);  
contexts based on Besserman and Kaiser (2016)