Broad Context Language Modeling as Reading Comprehension

Zewei Chu Hai Wang **Kevin Gimpel** David McAllester

LAMBADA: Word prediction requiring a broad discourse context

(Paperno et al., 2016)

he bent down and searched the large container, trying to find anything else hidden in it other than the _____

LAMBADA: Word prediction requiring a broad discourse context

(Paperno et al., 2016)

he turned to one of the cops beside him. “search the entire coffin.” the man nodded and bustled forward towards the coffin.

he bent down and searched the large container, trying to find anything else hidden in it other than the _____

LAMBADA: Word prediction requiring a broad discourse context (Paperno et al., 2016)

he turned to one of the cops beside him. "search the entire coffin." the man nodded and bustled forward towards the coffin.

he bent down and searched the large container, trying to find anything else hidden in it other than the **body**

- only answerable given discourse context!

LAMBADA: Word prediction requiring a broad discourse context (Paperno et al., 2016)

he turned to one of the cops beside him. "search the entire coffin." the man nodded and bustled forward towards the coffin.

he bent down and searched the large container, trying to find anything else hidden in it other than the **body**

- in >80% of instances, answer is in context (though not in this example)

Overview

- We view LAMBADA as reading comprehension and apply off-the-shelf models
- We improve state-of-the-art from 7.3% to 49%
- Manual analysis shows a variety of phenomena:
  - Easy ones solved by comprehension models
  - Hard ones need more semantics, coreference, external knowledge

CNN/Daily Mail Comprehension Tasks
(Hermann et al., 2015)

Document:

actress @entity1 has entered a rehab facility for her addictions , a spokesman said . " @entity1 has valiantly battled substance abuse over the years and whenever she has needed to seek treatment she has done so , " said spokesman @entity5 … @entity1 won an @entity15 in 1973 for her performance in " cabaret . " …

Question:

XXXXX won an @entity15 for her performance in " cabaret "

CNN/Daily Mail Comprehension Tasks
(Hermann et al., 2015)

Document:

actress @entity1 has entered a rehab facility for her addictions , a spokesman said . " @entity1 has valiantly battled substance abuse over the years and whenever she has needed to seek treatment she has done so , " said spokesman @entity5 … **@entity1 won an @entity15 in 1973 for her performance in " cabaret . "** …

Question:

XXXXX won an @entity15 for her performance in " cabaret "

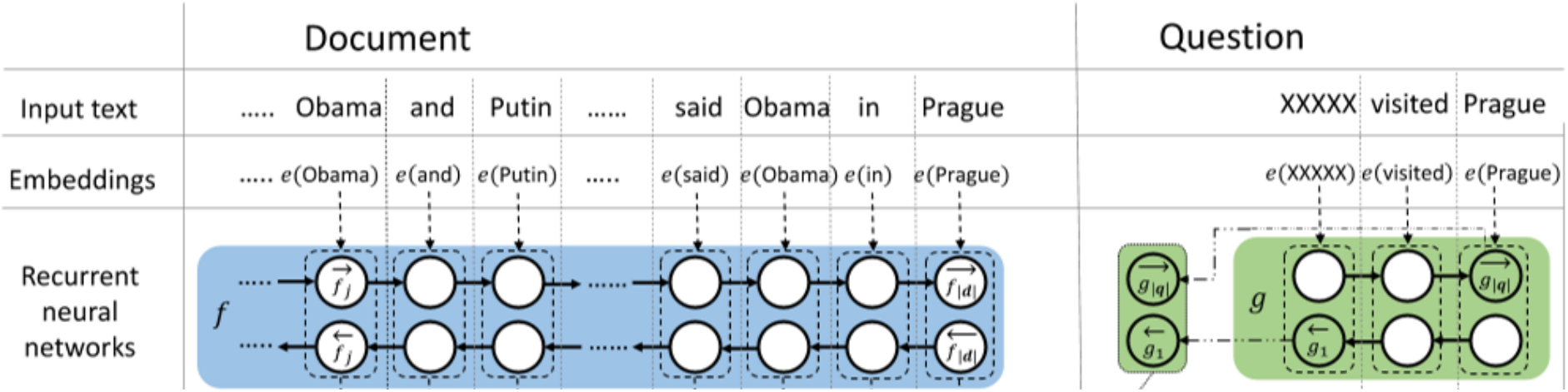
Neural Reading Comprehension Models

- lots of recent activity here!
  - Hermann et al. (2015)
  - Hill et al. (2016)
  - Chen et al. (2016)
  - Kadlec et al. (2016)
  - Dhingra et al. (2016)
  - *inter alia*
- we will describe the Attention Sum Reader (Kadlec et al., 2016) because it is simple and works well

Attention Sum Reader
(Kadlec et al., 2016)

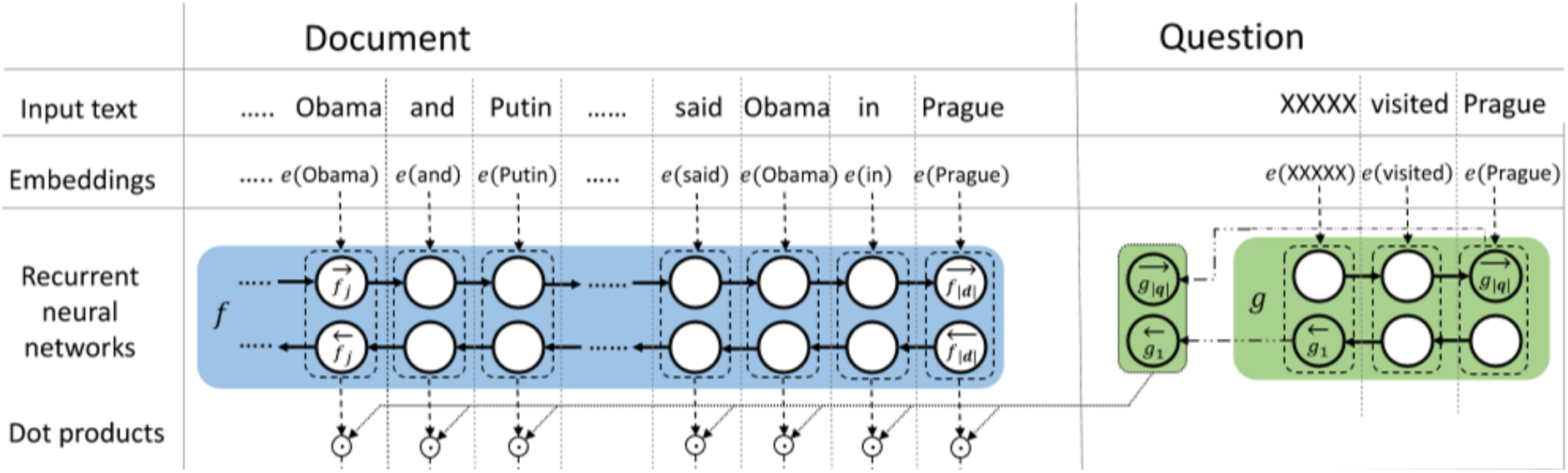
	Document								Question		
Input text	.....	Obama	and	Putin	.....	said	Obama	in	Prague	XXXXX	visited	Prague

Attention Sum Reader
(Kadlec et al., 2016)



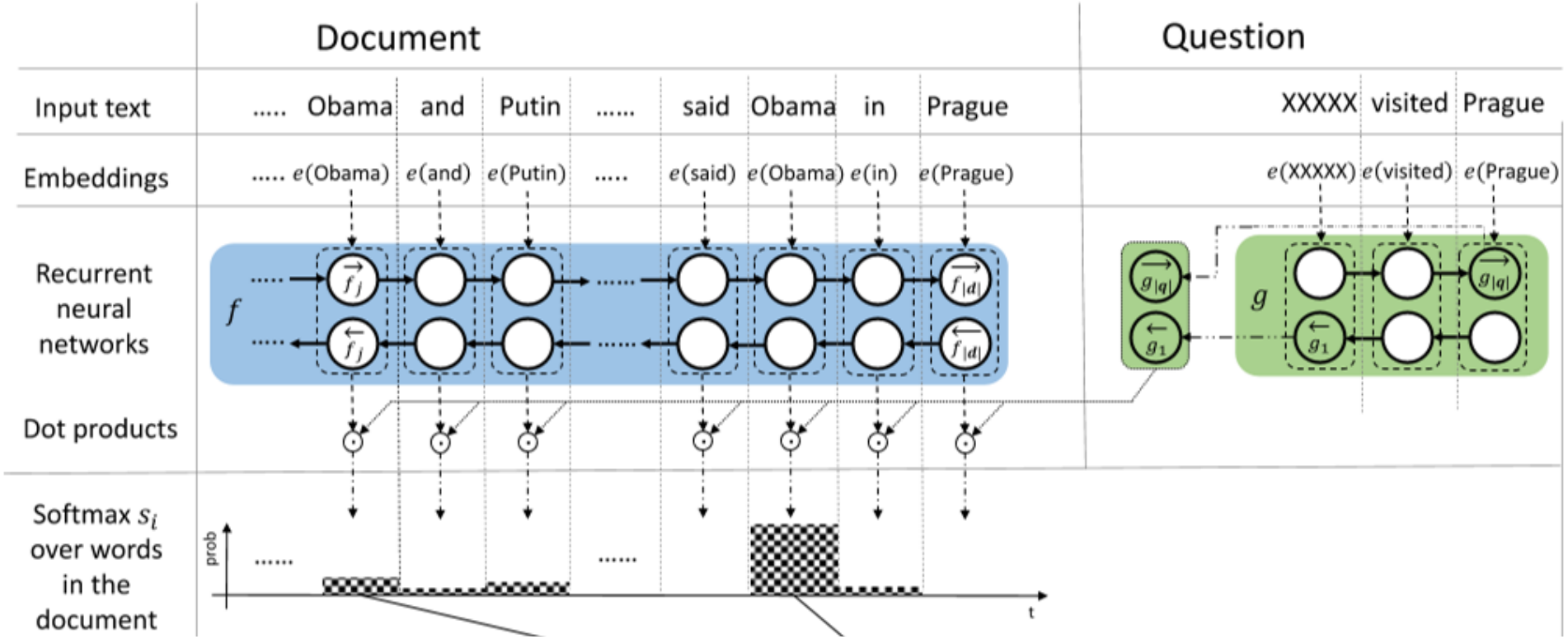
- Encode document using bidirectional RNN
- Encode question using another bidirectional RNN

Attention Sum Reader
(Kadlec et al., 2016)



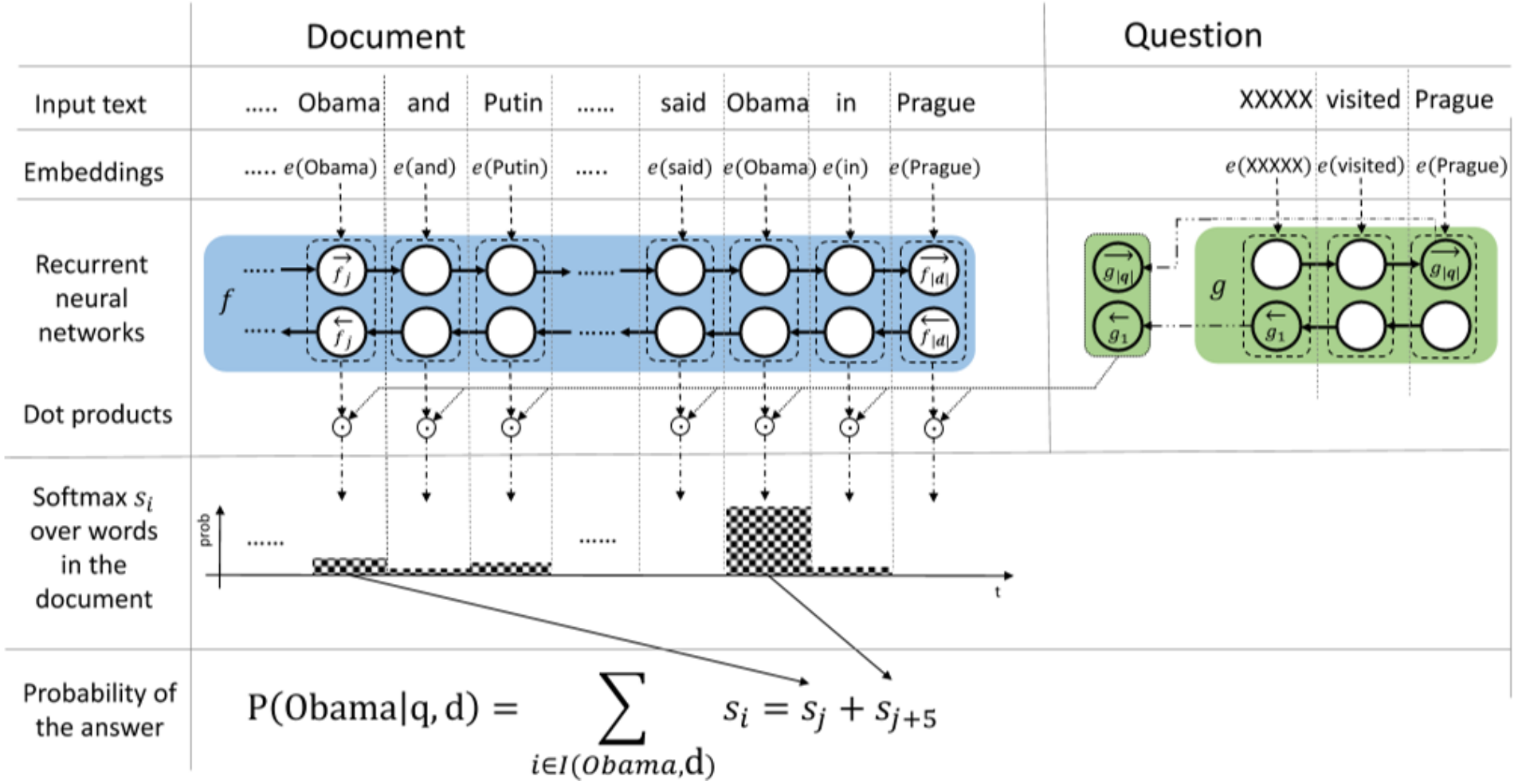
- Compute attention over positions of document using question representation

Attention Sum Reader
(Kadlec et al., 2016)

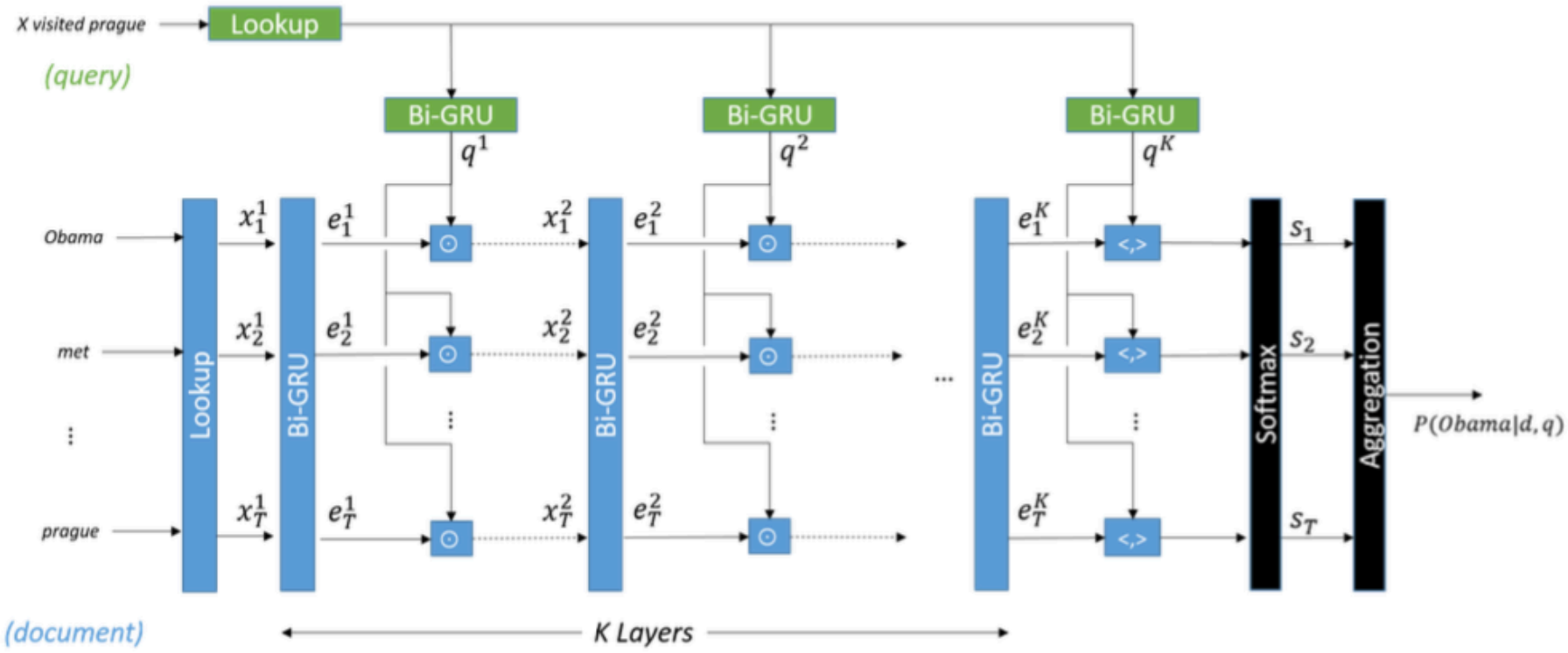


- Normalize over positions of document

Attention Sum Reader
(Kadlec et al., 2016)



Gated Attention Reader
(Dhingra et al., 2016)



Training Data?

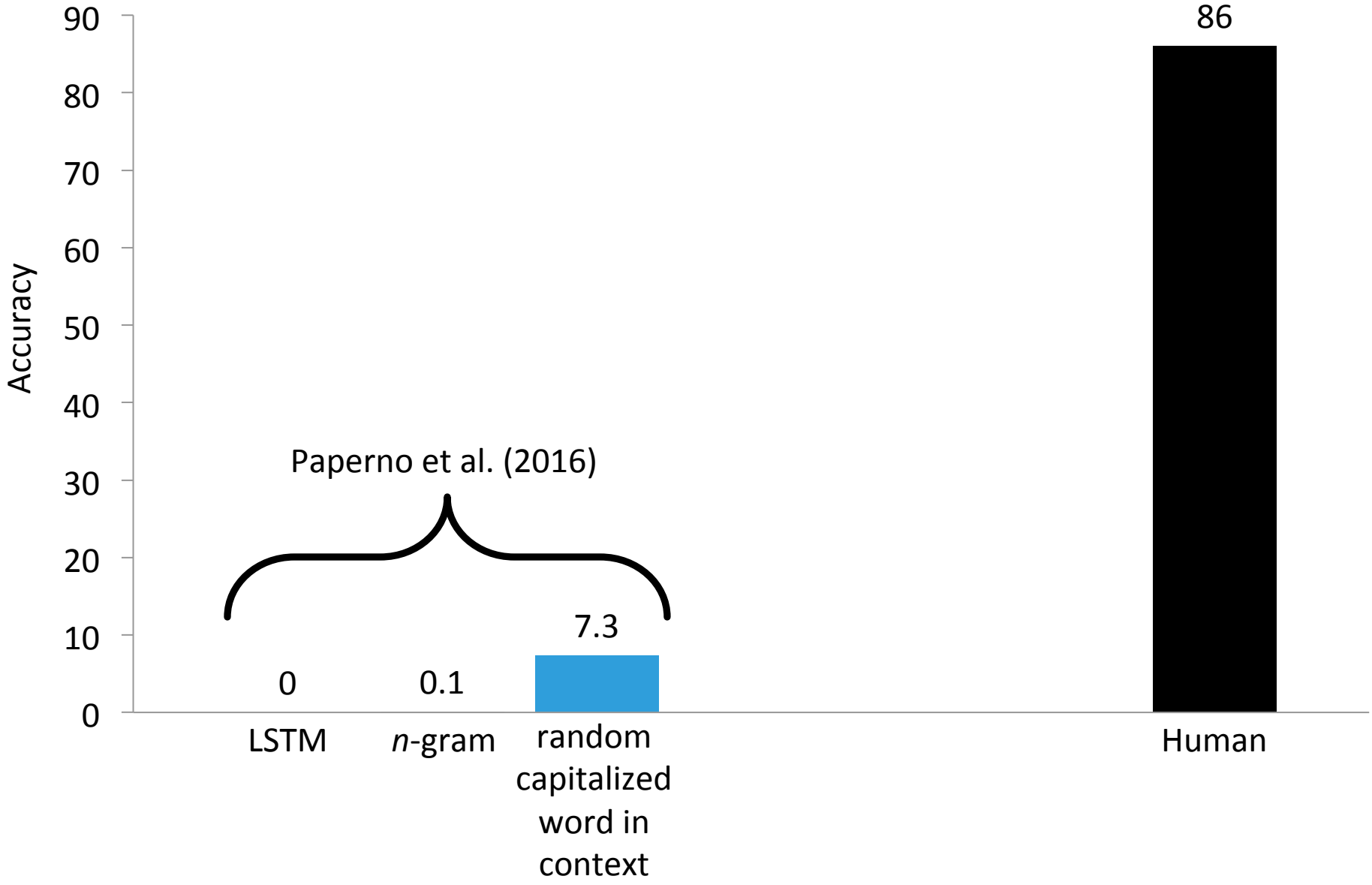
- we need data to train these comprehension models

- LAMBADA only includes dev/test sets

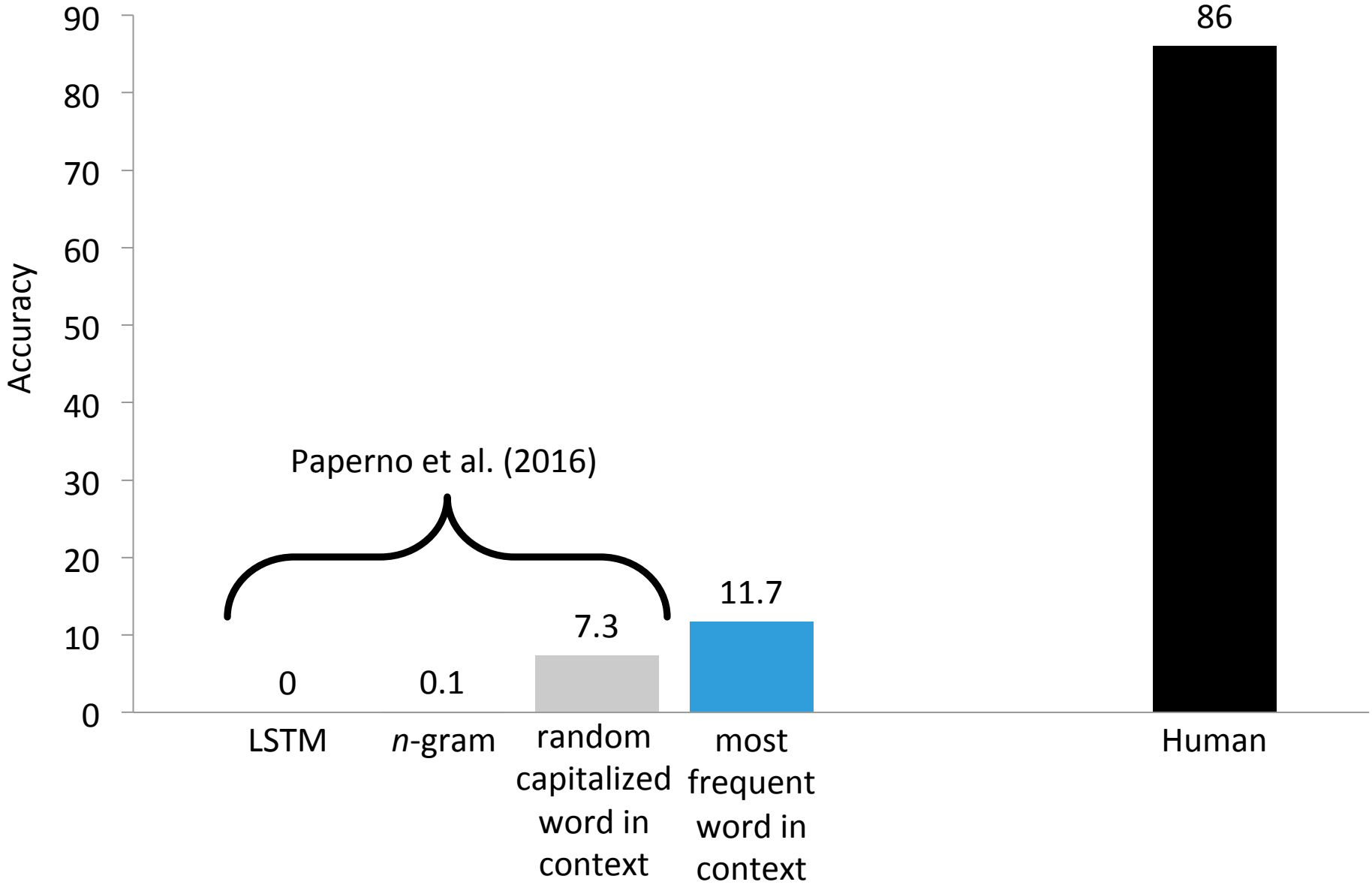
Automatic Training Data Creation

- observation: in >80% of LAMBADA, answer word is in context
- we automatically create training instances where answer word is in context
  - each instance has 4-5 sentences and >=50 words
  - total of 1.8 million instances for training
- training data is available (see my web page)

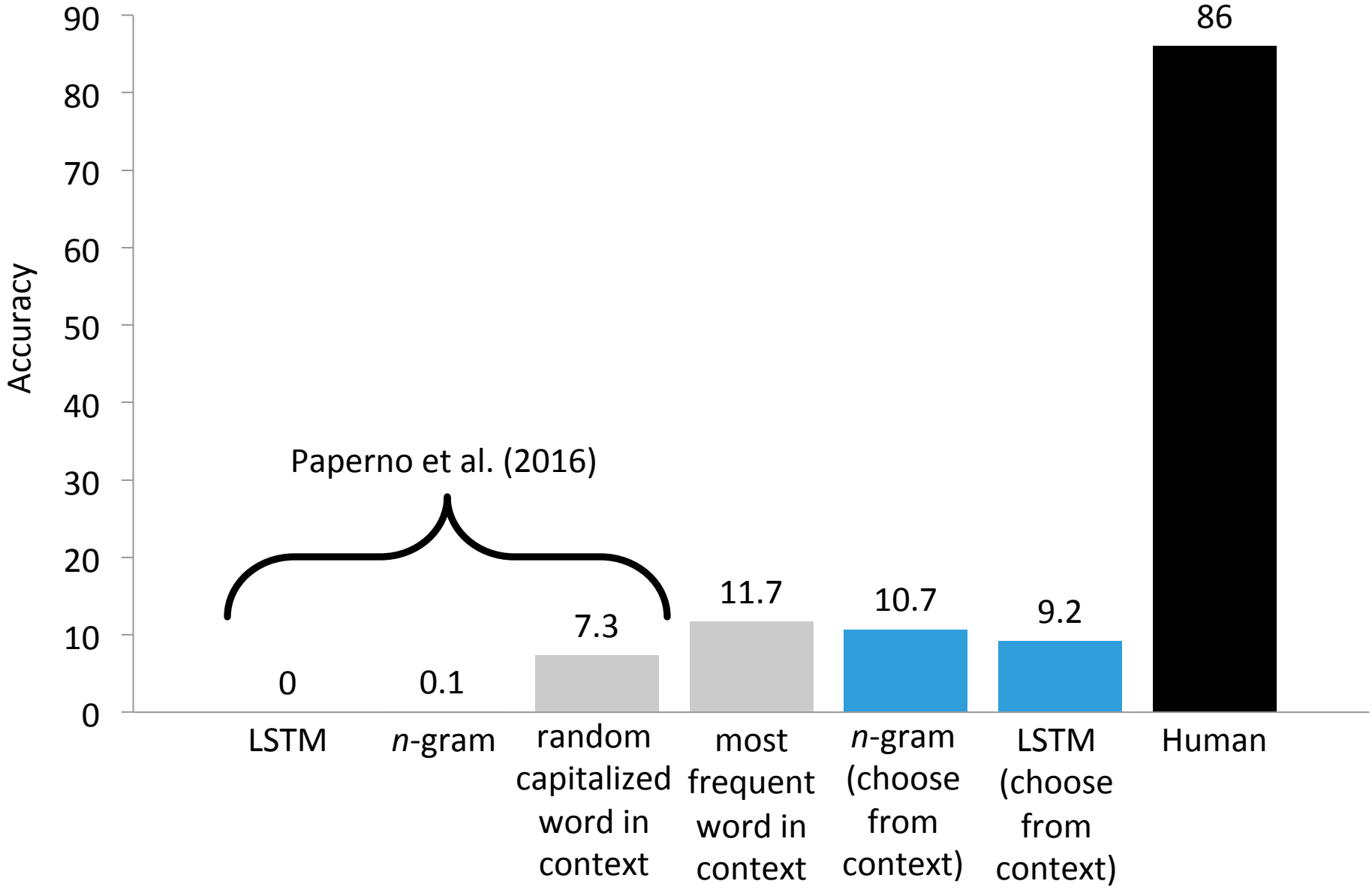
Results: LAMBADA Test Set



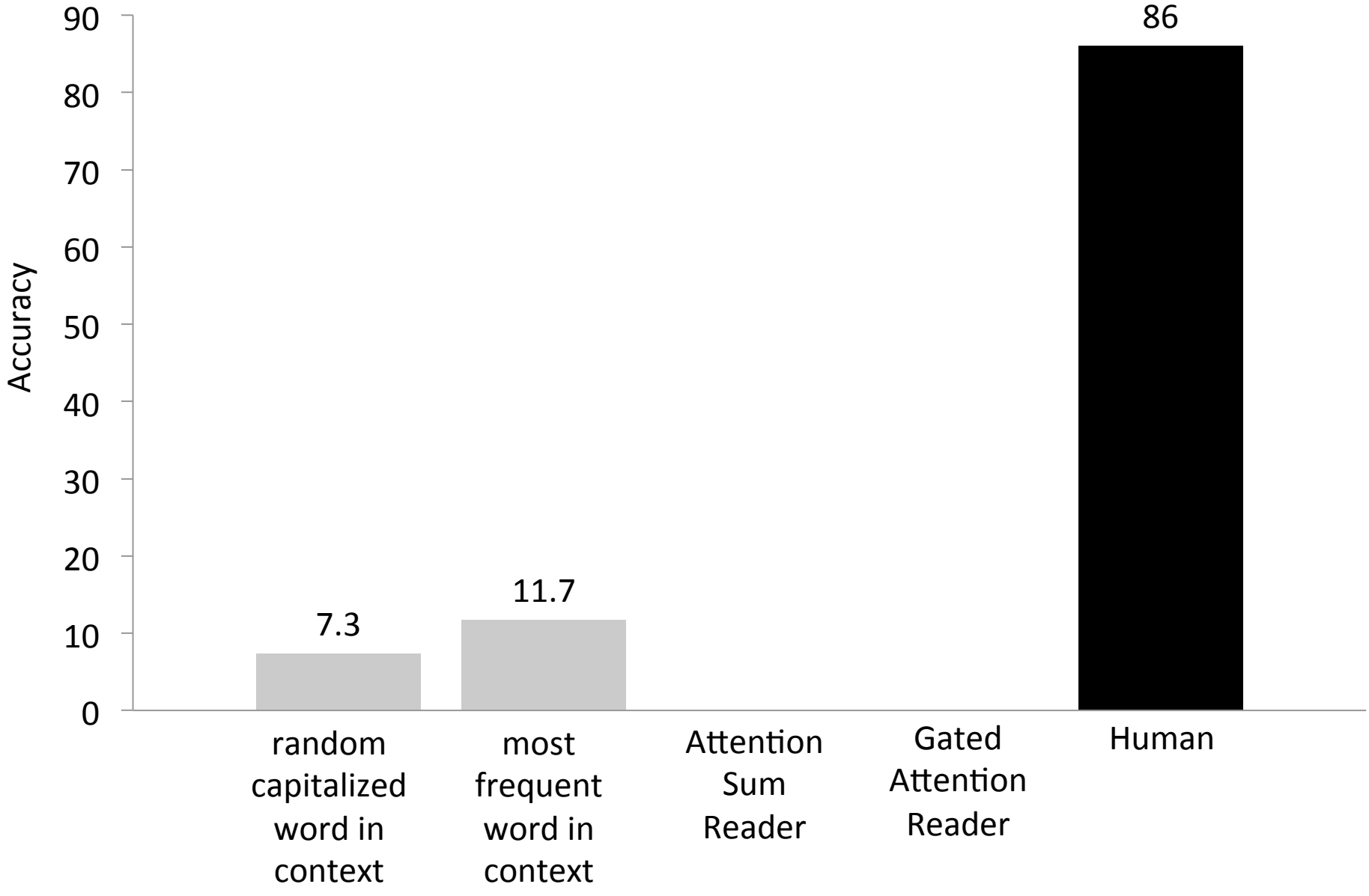
Results: LAMBADA Test Set



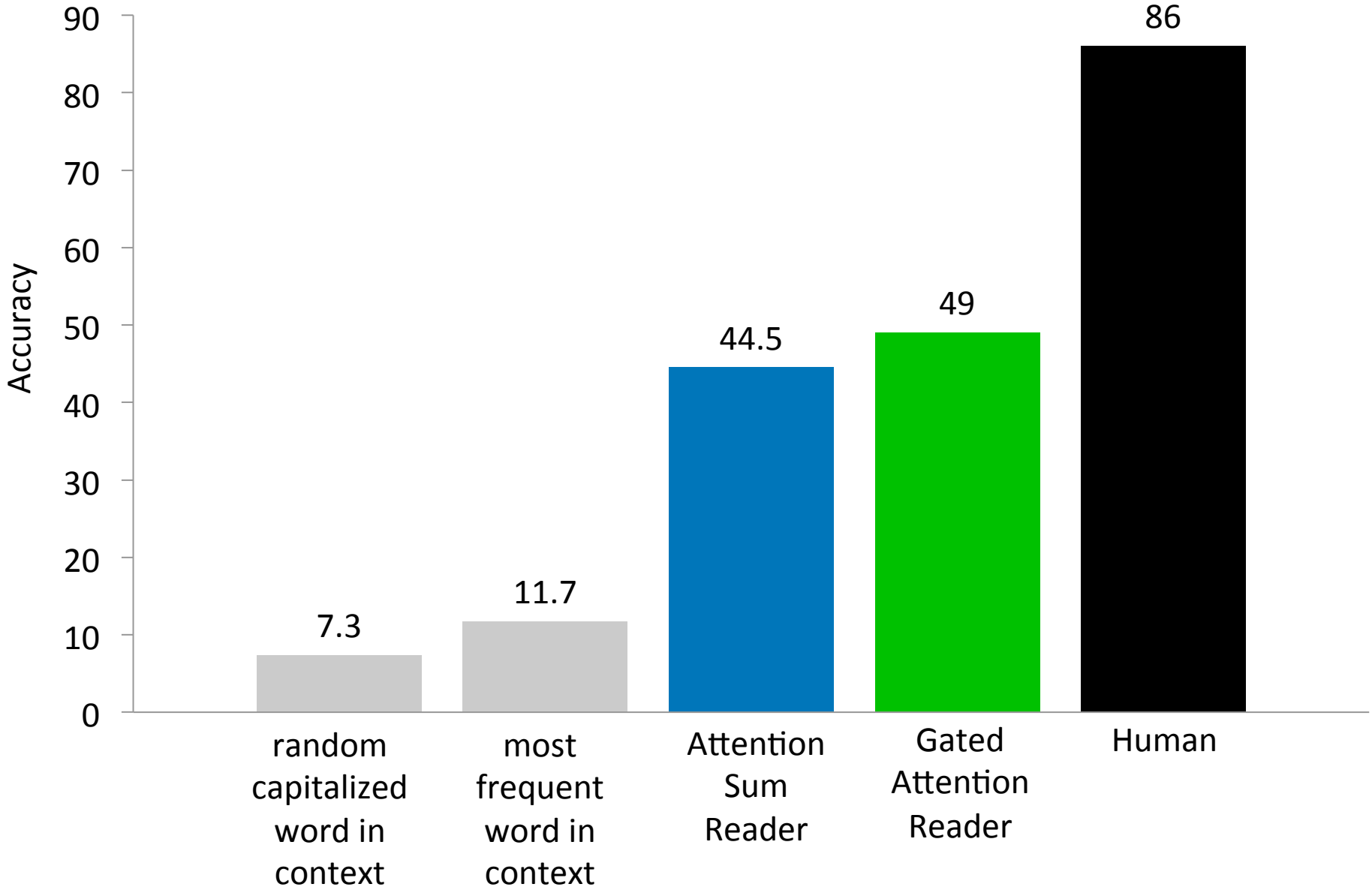
Results: LAMBADA Test Set



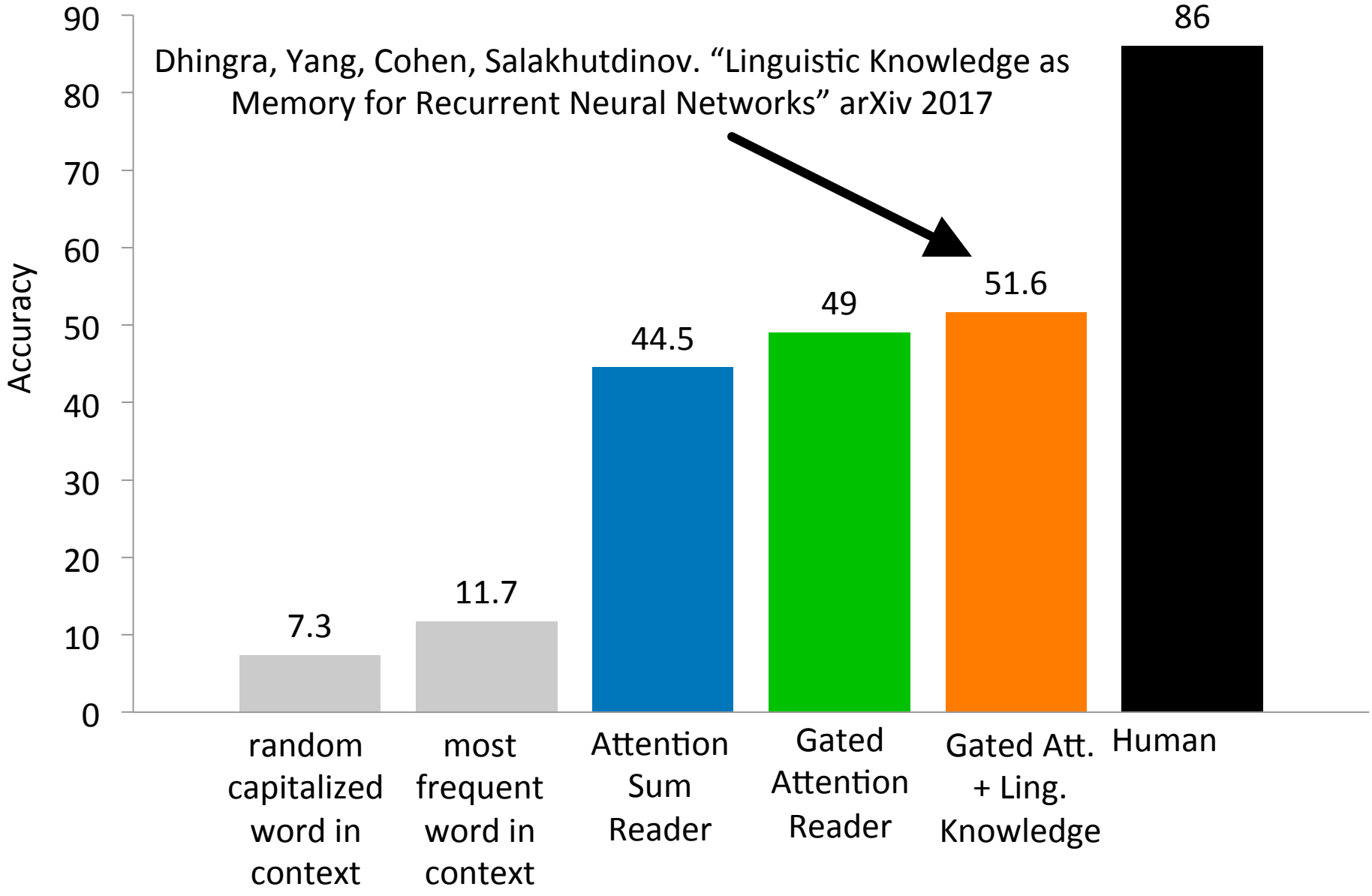
Results: Neural Readers on LAMBADA



Results: Neural Readers on LAMBADA



Results: Neural Readers on LAMBADA



Manual Analysis

- 100 LAMBADA instances
- 86 of 100 correct
- labeled instances with types of understanding needed to solve them

Example Label: Simple Speaker Tracking

… it meant that dog and cat could contribute to the conversation.

“hey, dog, ever been in one of these things?” asked cat.

“no.” replied dog sadly.

“you say that so glumly,” said ____

Example Label: Simple Speaker Tracking

… it meant that dog and cat could contribute to the conversation.

“hey, dog, ever been in one of these things?” asked **cat**.

“no.” replied dog sadly.

“you say that so glumly,” said ____

- Answerable by tracking who is speaking without understanding what they are saying
- 19 of 100
- these are also easy (16/19 correct)

Example Label 2: Coreference Resolution

instead, he danced, turned circles under the spider's body, and crouched close to the spider's abdomen to avoid its bite.

brian joined thomas. the two boys fought together, just as they had done in our prison, just as the spiders were doing. now both were stabbing and dodging, double-teaming against the _____

- 21 of 100
- these are hard (8/21 correct)

Example Label 3: External Knowledge

…he turned to one of the cops beside him. “search the entire coffin.”

the man nodded and bustled forward towards the coffin. he bent down and searched the large container, trying to find anything else hidden in it other than the _____

- Involves knowledge possessed by human readers but not contained in context
- 24 of 100
- these are also hard (5/24 correct)

Conclusions

- reading comprehension models improve state-of-the-art on LAMBADA from 7.3% to 49% (since improved further)
- but the last 50% are difficult!
- manual analysis reveals several categories of phenomena, including coreference and external knowledge

Thanks!

Example Label 1: Unambiguous Name Cue

“no worries; she's fine,” said sheila, opening the door for him to enter.

“oh, i can't stay. i just came to offer my services. if you need anything, don't hesitate to ask. i'm right across the street.”

“actually, we need someone to fix our plumbing,” said _____

Example Label 1: Unambiguous Name Cue

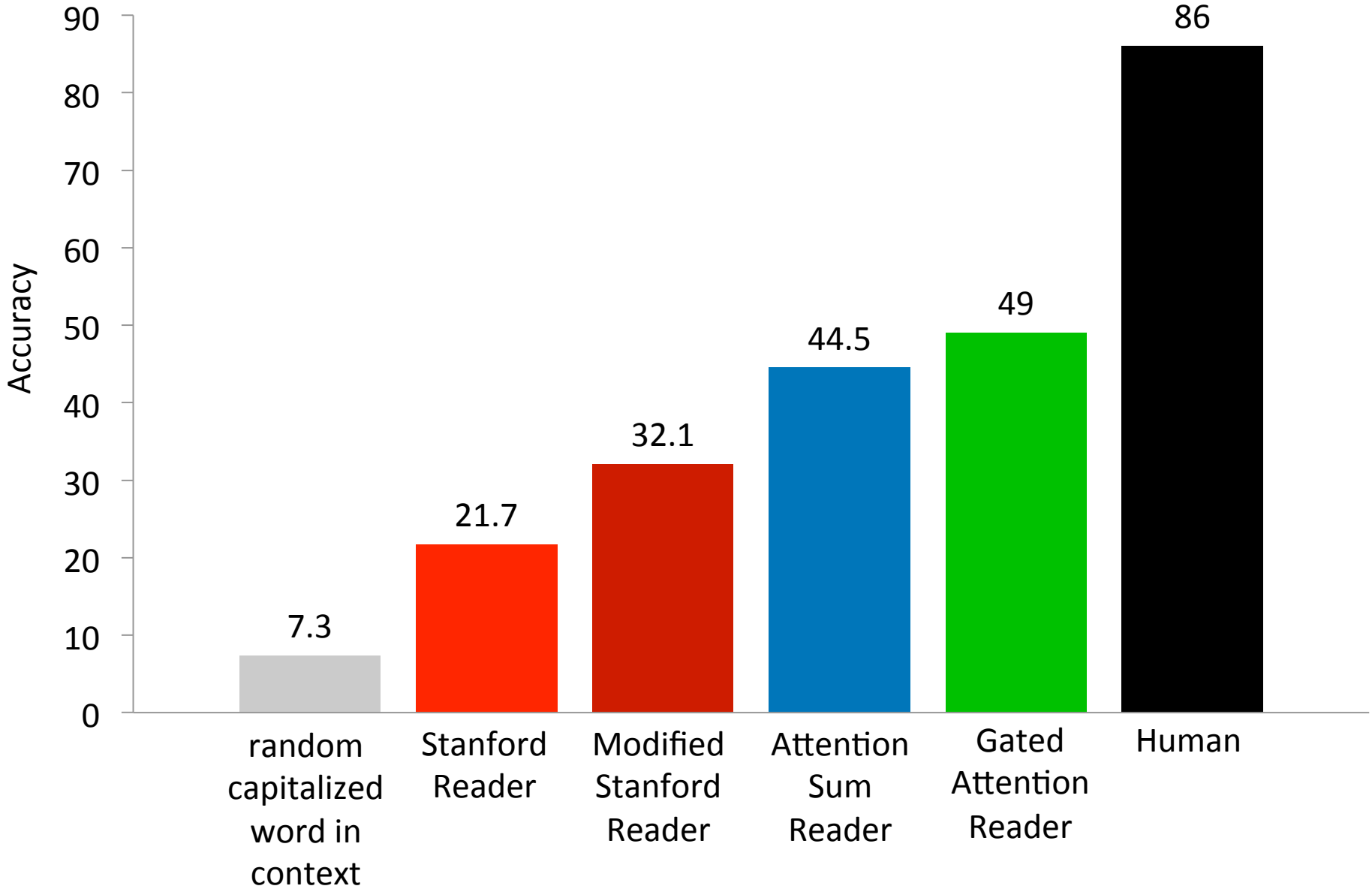
"no worries; she's fine," said **sheila**, opening the door for him to enter.

"oh, i can't stay. i just came to offer my services. if you need anything, don't hesitate to ask. i'm right across the street."

"actually, we need someone to fix our plumbing," said _____

- answer is clearly a name based on local cues, and context only contains a single name
- 9 of 100
- these are easy (8/9 correct by GA Reader)

Results: Neural Readers on LAMBADA



Results: Neural Readers on LAMBADA