

Recurrent Neural Networks Always Learn English-Like Relative Clause Attachment

Forrest Davis and Marten van Schijndel
ACL 2020, Virtually from Ithaca
July 6-8th, 2020

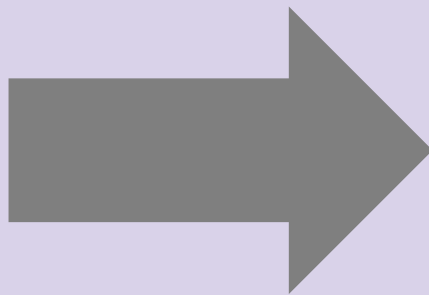
TL;DR

- Human comprehension biases differ by language
 - English vs. Spanish
- Language modeling objective with RNNs, Spanish comprehension bias **not** learned
- Raw text (production data) **systematically lacks** comprehension signal

Raw Text



WIKIPEDIA
The Free Encyclopedia



Understanding

- Question-Answering
- Natural Language Inference
- Topic Modeling
- Text Summarization

Aspects of **knowledge**
and natural language **meaning**
may not be learnable using
current techniques

Can human-like language
comprehension biases be
acquired from production
data?

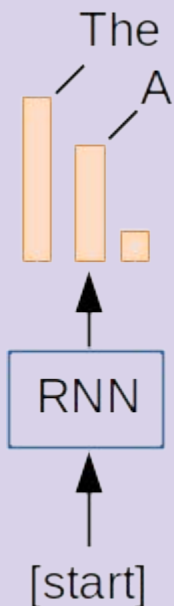
No

Davis and van Schijndel, 2020

Neural Network Language Modeling

RNN Word Prediction

[start] The man saw the dog [end]



Output:
Probability
distribution
over next word

Input:
Current word

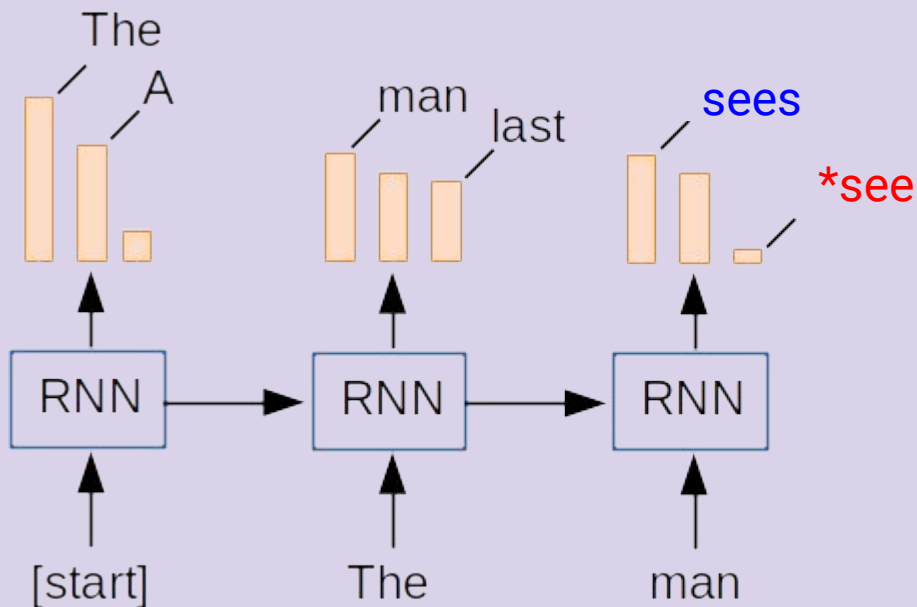
How do we probe RNN LMs for syntactic knowledge?

Linking Hypothesis

Models should assign higher probability to **grammatical** constructions than **ungrammatical** constructions

Exemplar of probing RNN LM representations

The man *see/sees



Previous work

Models prefer grammatical over ungrammatical constructions

- Subject-verb agreement
- Center-embedding
- Syntactic Islands

Linzen et al., 2016
Gulordava et al., 2018
Wilcox et al., 2019a,b

Previous work

This syntactic knowledge may be more brittle with inability to generalize

van Schijndel et al., 2019

Previous work

Models prefer grammatical over ungrammatical constructions

- Subject-verb agreement
- Center-embedding
- Syntactic Islands

This syntactic knowledge may be more brittle with inability to generalize

Linzen et al., 2016
Gulordava et al., 2018
Wilcox et al., 2019a,b

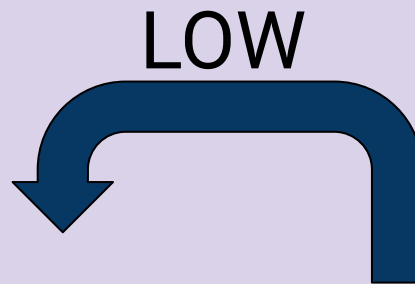
van Schijndel et al., 2019

Ambiguous Relative Clause Attachment

Ambiguous Relative Clause Attachment

John met the agent of the rocker *who is divorced*

Ambiguous Relative Clause Attachment



John met the agent of the rocker *who is divorced*

Ambiguous Relative Clause Attachment

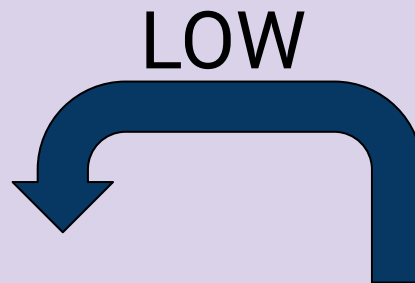
HIGH



John met the agent of the rocker *who is divorced*

Ambiguous Relative Clause Attachment

English speakers
have a preference
for **LOW**



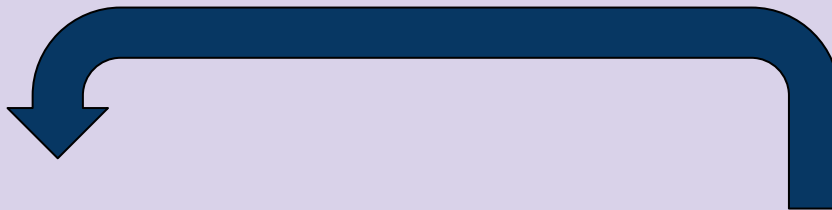
John met the agent of the rocker *who is divorced*

Carreiras and Clifton, 1993;
Frazier and Clifton, 1996;
Carreiras and Clifton, 1999;
Fernández, 2003

Ambiguous Relative Clause Attachment

Spanish speakers
have a preference
for **HIGH**

HIGH



John met the agent of the rocker *who is divorced*

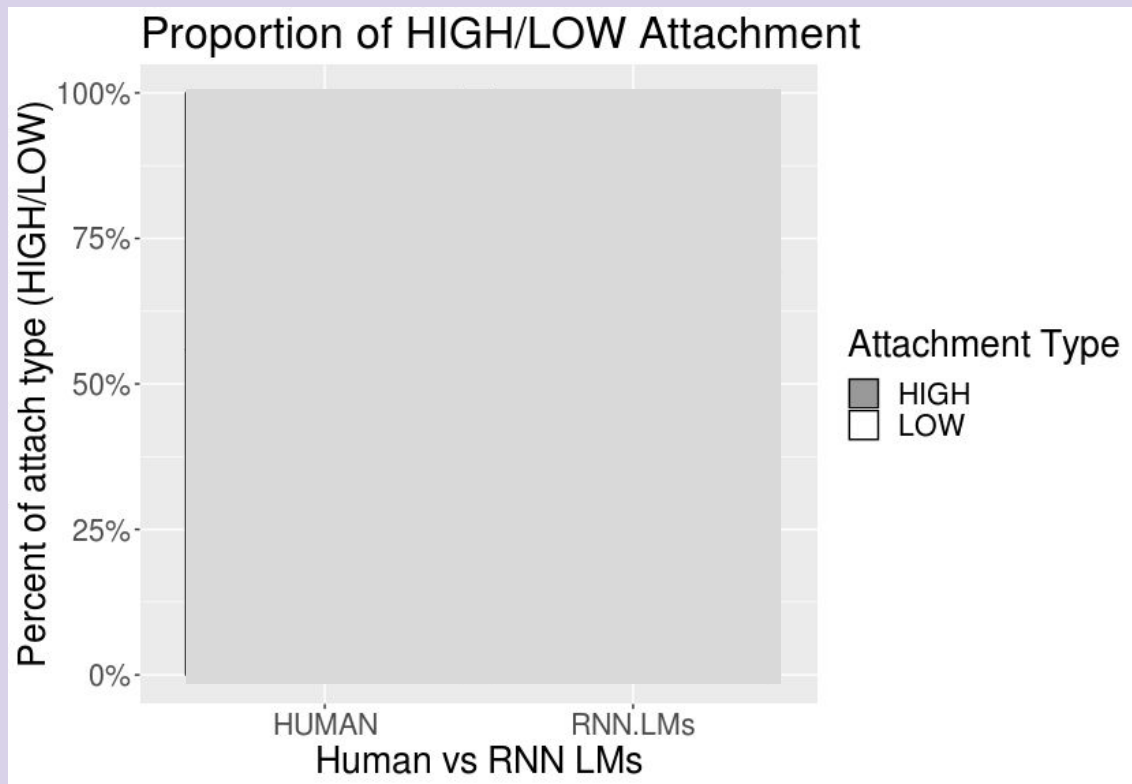
Carreiras and Clifton, 1993;
Carreiras and Clifton, 1999;
Fernández, 2003

Local (LOW)
Non-Local (HIGH)

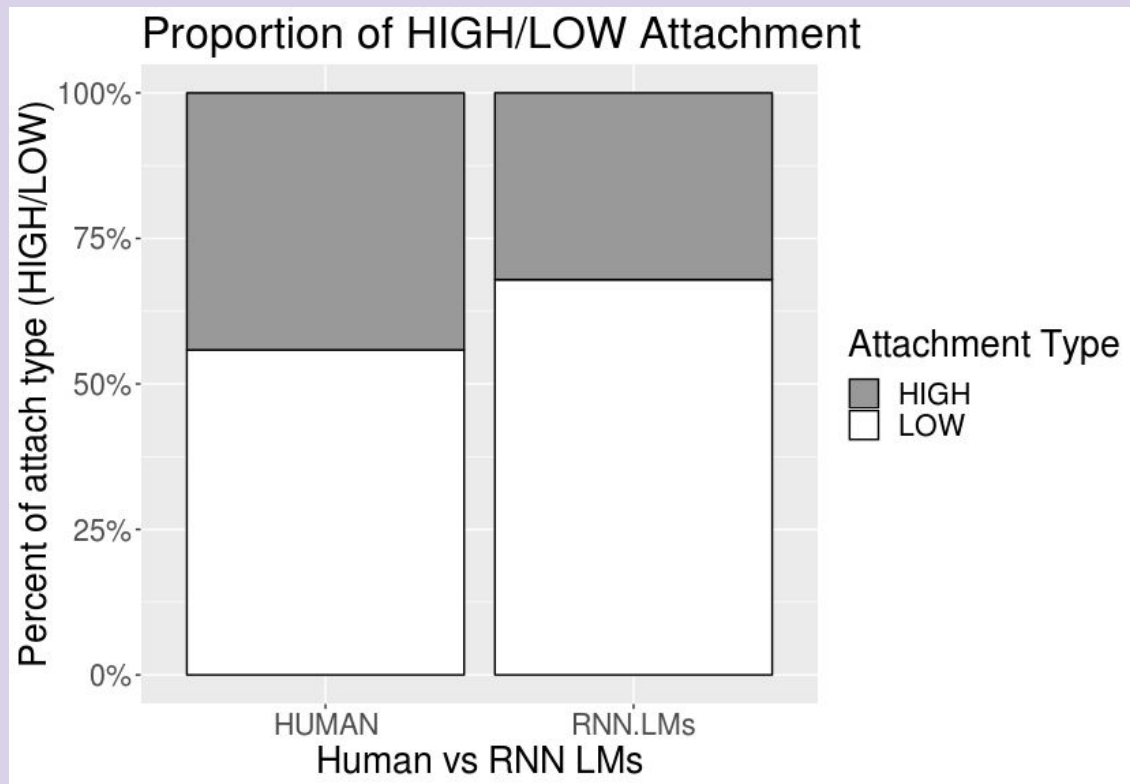
Afrikaans	Japanese
Arabic	Norwegian
Croatian	Persian
Danish	Polish
Dutch	B. Portuguese
English	Romanian
French	Russian
German	Spanish
Greek	Swedish
Italian	Thai

Probing RC attachment preferences in RNN LMs

RNN LMs seem to have a LOW bias



RNN LMs seem to have a LOW bias

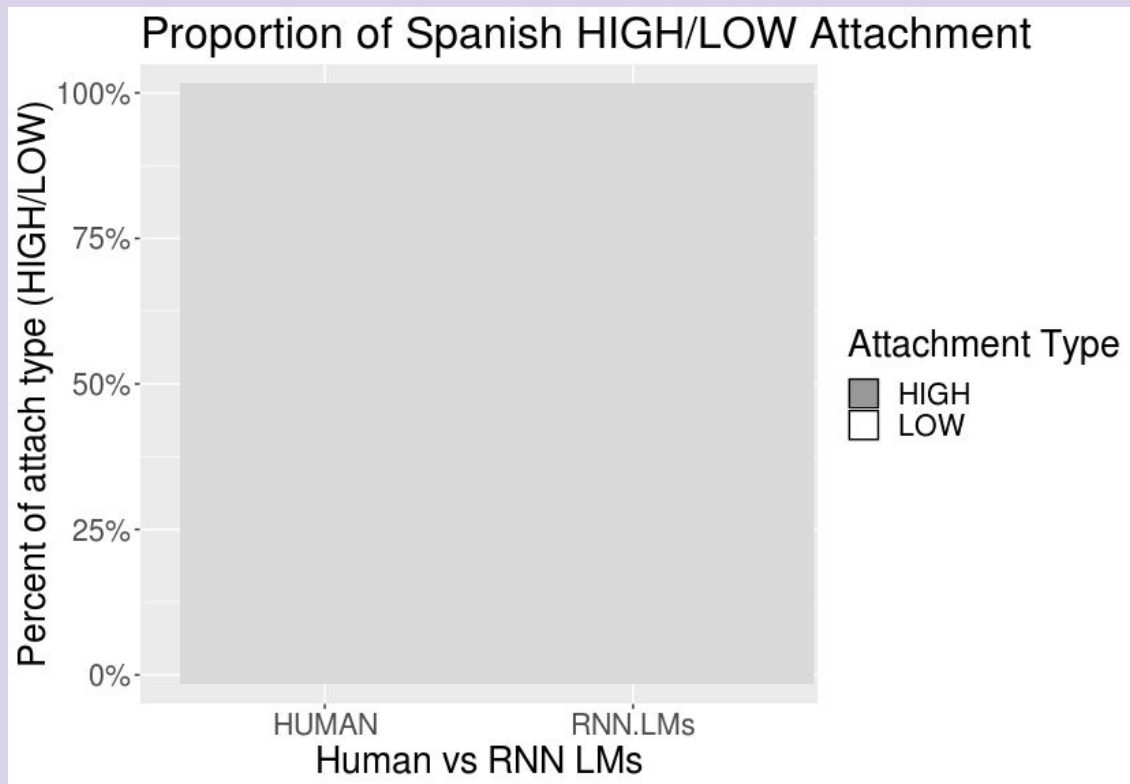


p-value < 0.00001
Bayes Factor > 100

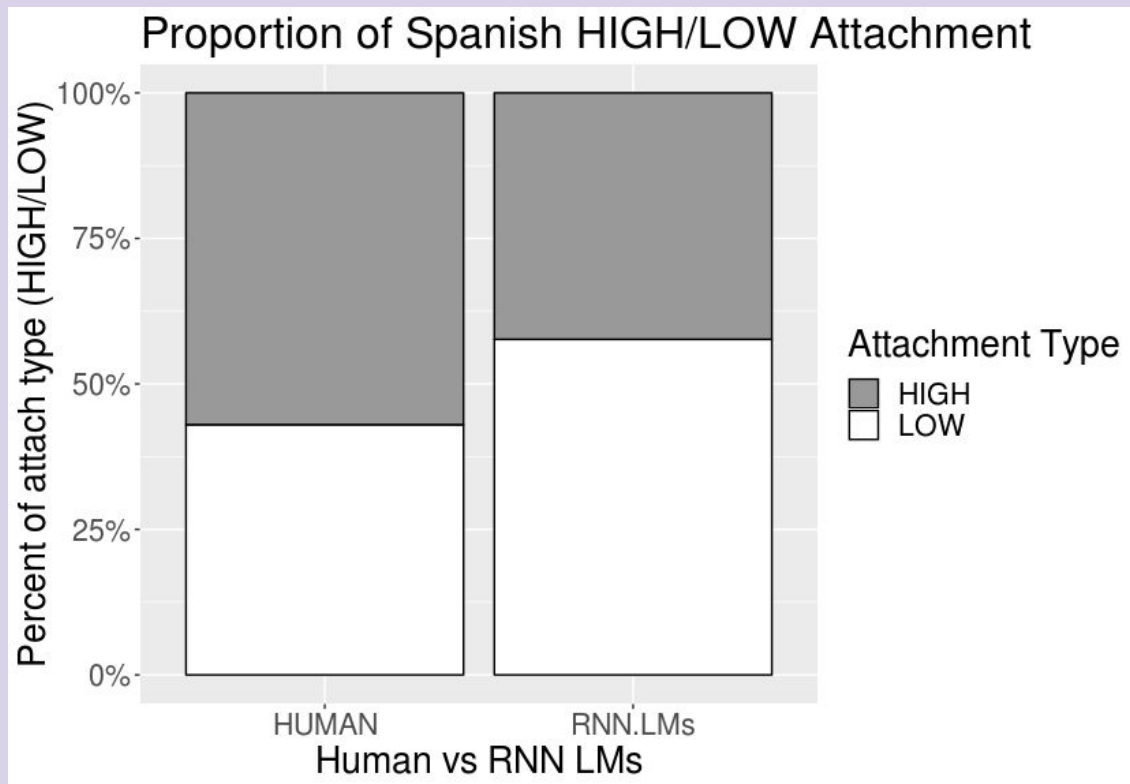
Interim Thoughts

Is this a preference for LOW?

Spanish Results



Spanish Results



p-value < 0.00001
Bayes Factor > 100

**Why can't the model
learn the Spanish
attachment
preference?**

RNN LMs can acquire HIGH or LOW bias when trained on synthetic data (may even have a slight HIGH bias)

Comprehension

≠

Production

Comprehension signal not in raw text data

Spanish Wikipedia (training corpus):

LOW 69% more frequent than HIGH

Spanish Newswire data:

LOW 21% more frequent than HIGH

Conclusion

Human linguistic comprehension is not learnable from text alone (production data)

Possible explanation for why more data doesn't improve model quality

Future Work

Explore additional attachment phenomena (prepositional phrase attachment)

Different model architectures (e.g., Transformers, RNNG, OrderedNeurons)

Human comprehension signal (e.g., reading times, fMRI, EEG)

Thanks!

Email - fd252@cornell.edu

Website - <https://conf.ling.cornell.edu/forrestdavis/>