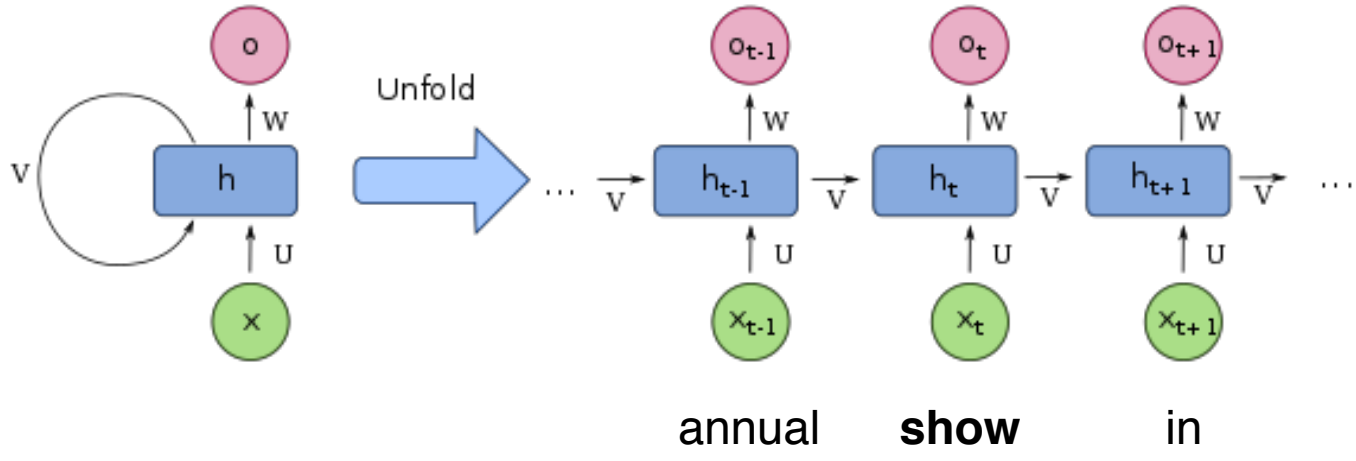
Language models capture anaphoric coreference to a large extent

Gemma Boleda
Universitat Pompeu Fabra / ICREA
joint work with Ionut Sorodoc and Kristina Gulordava

DLBCN 2019
Barcelona, December 19-20 2019

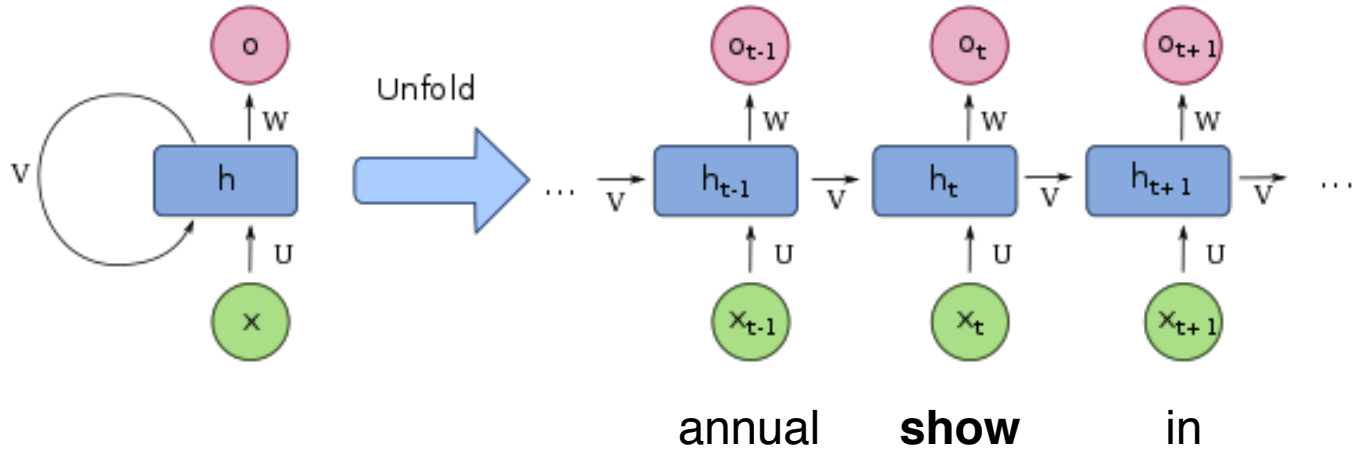
Horizon 2020
European Union funding
for Research & Innovation

Can LSTM-based language models do reference?



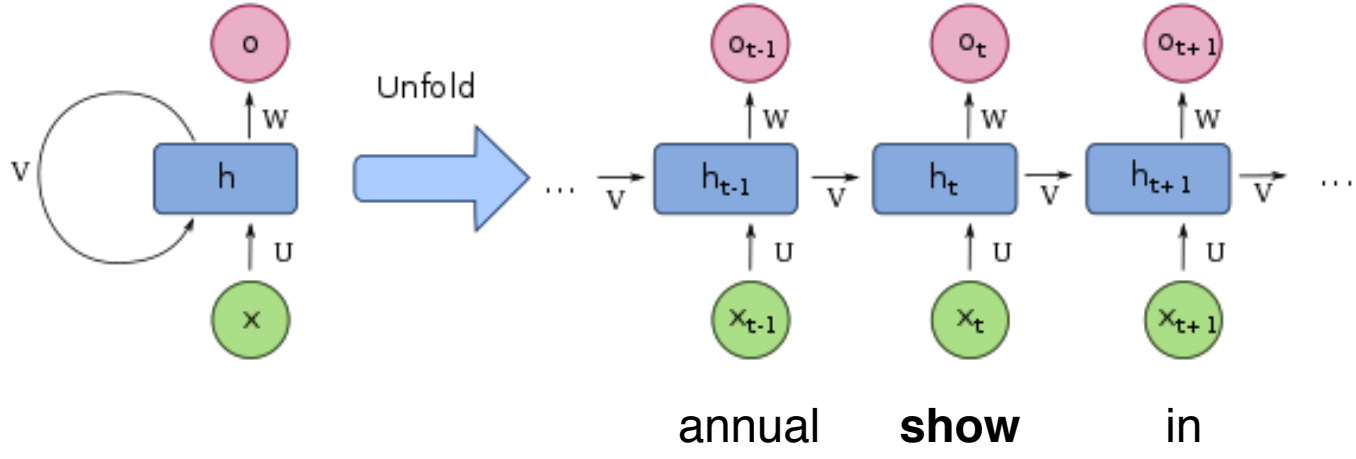
- language model: annual show in _

Can LSTM-based language models do reference?



- language model: annual show in _
- LSTM: standard recurrent model for NLP (still?)

Can LSTM-based language models do reference?



- language model: annual show in _
- LSTM: standard recurrent model for NLP (still?)
- they have been shown to capture hierarchical aspects of syntax (Linzen et al. 2016; Gulordava et al. 2018)

→ can they capture reference?

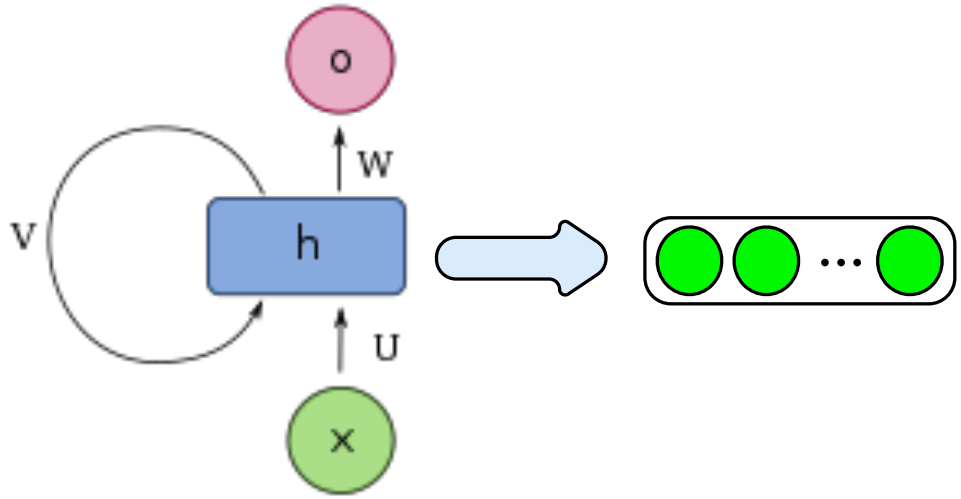
Graph by François Deloche - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=60109157

Local relations: Pronominal anaphora

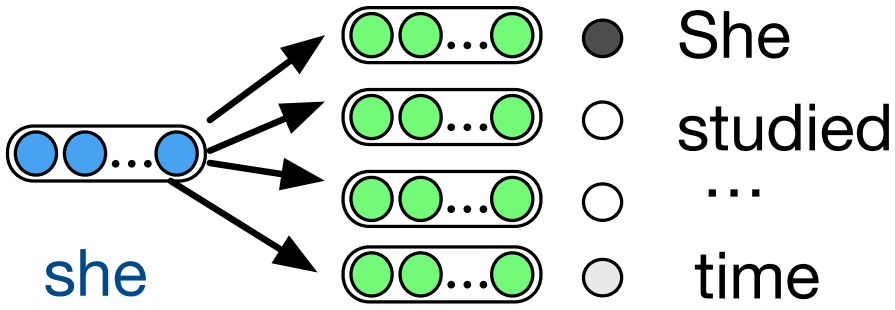
...he$_1$ was elected to be president of the People's Republic of China, and chairman of the$_2$ Central$_2$ Military$_2$ Commission$_2$. Yeping$_3$ Wang$_3$ was born in Shanghai in 1926. She$_3$ studied in Shanghai Foreign Language College, and started working in 1949. For a long time, she$_3$...

- Task: predict which element the target (**she**) refers to
- Correct if it points to any element in the coreference chain

Method



She studied … For a long time, she



Results

Model	Accuracy	
closest gold entity	56.1	
closest same-form token	61.3	
	unsup.	sup.
LSTM	41.7	64.8
Transformer	48.5	**75.9**

What does it learn?

Figures for Transformer; same picture for LSTM

Syntax

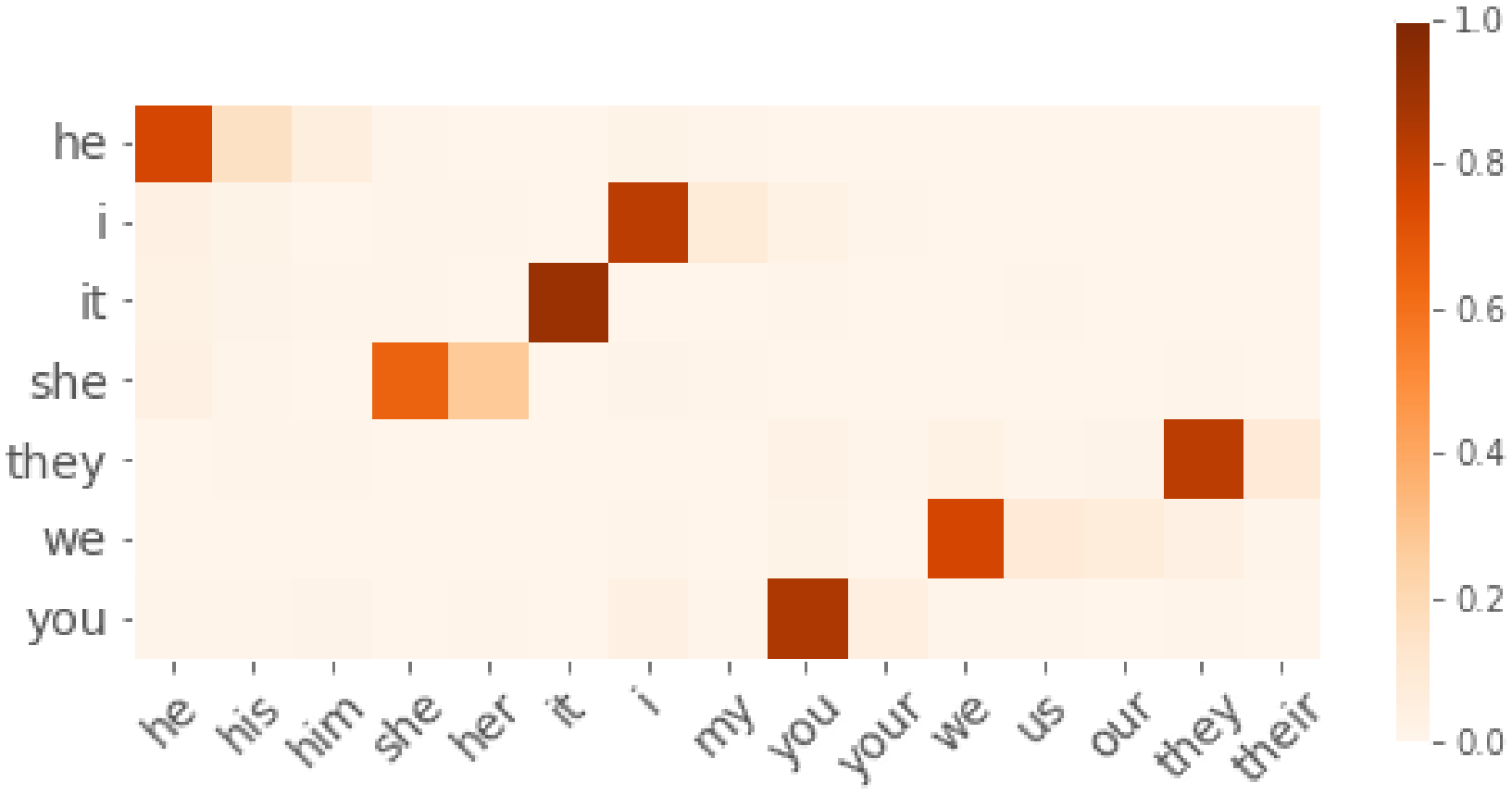
- “Pronouns refer to nominal elements”

→ Refers to element in a chain: 92.6%

→ even when it doesn’t, refers to a nominal element 82% of the time

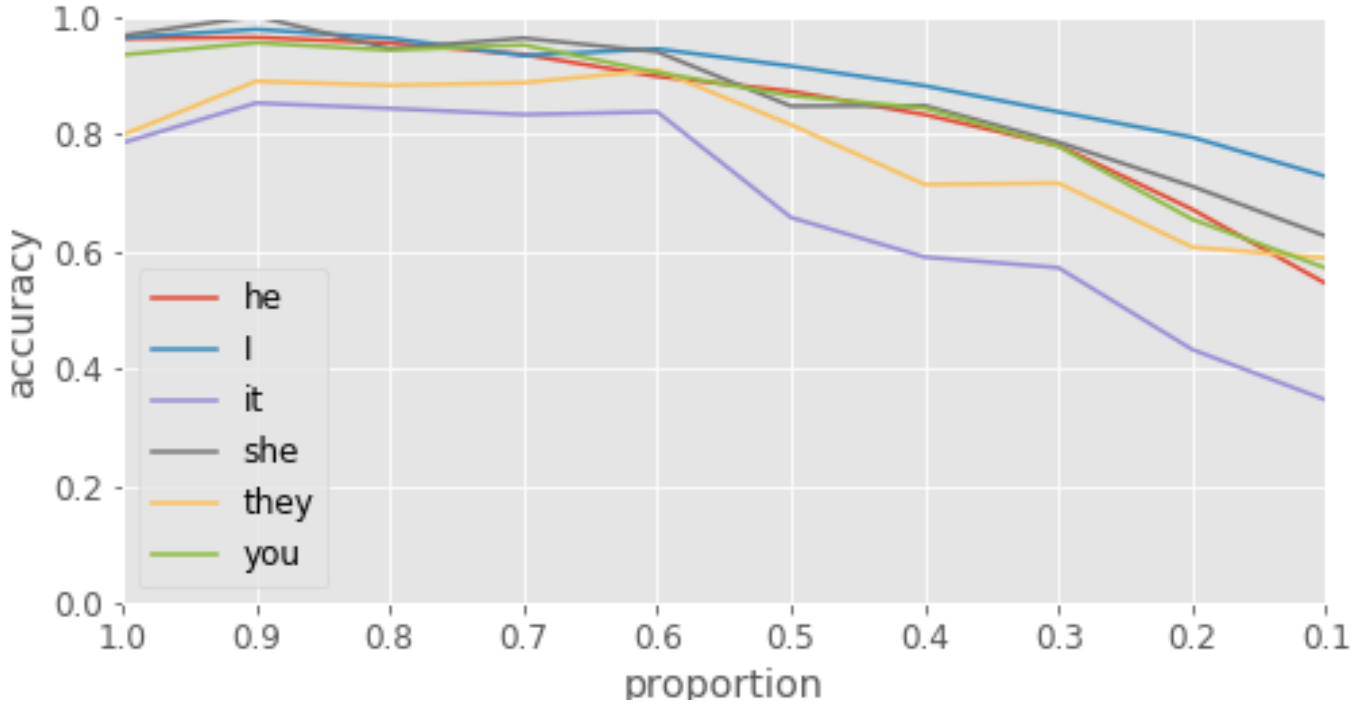
- “Pronouns agree in gender and number”
- Across sentences

Syntax: Pronoun agreement



But, what about semantics (reference)?

She studied in Shanghai Foreign Language College , and started working in 1949 . <eos> For a long time , **she**



distractors confuse the model, but they do not fool it completely

Distinguishing entities - local

She studied in Shanghai Foreign Language College , and started working in 1949 . <eos> For a long time , **she**

	Accuracy
No distractor	81.8
Distractor(s)	73.8
Same gender	48.6
Same number	65.3
Same pronoun	49.1

relies too much on morphosyntactic features and recency
but baseline 15.7-25.5!

Interim conclusion

language models are able to distinguish mentions of different entities to some extent, although they are far worse at this than at capturing morphosyntactic features

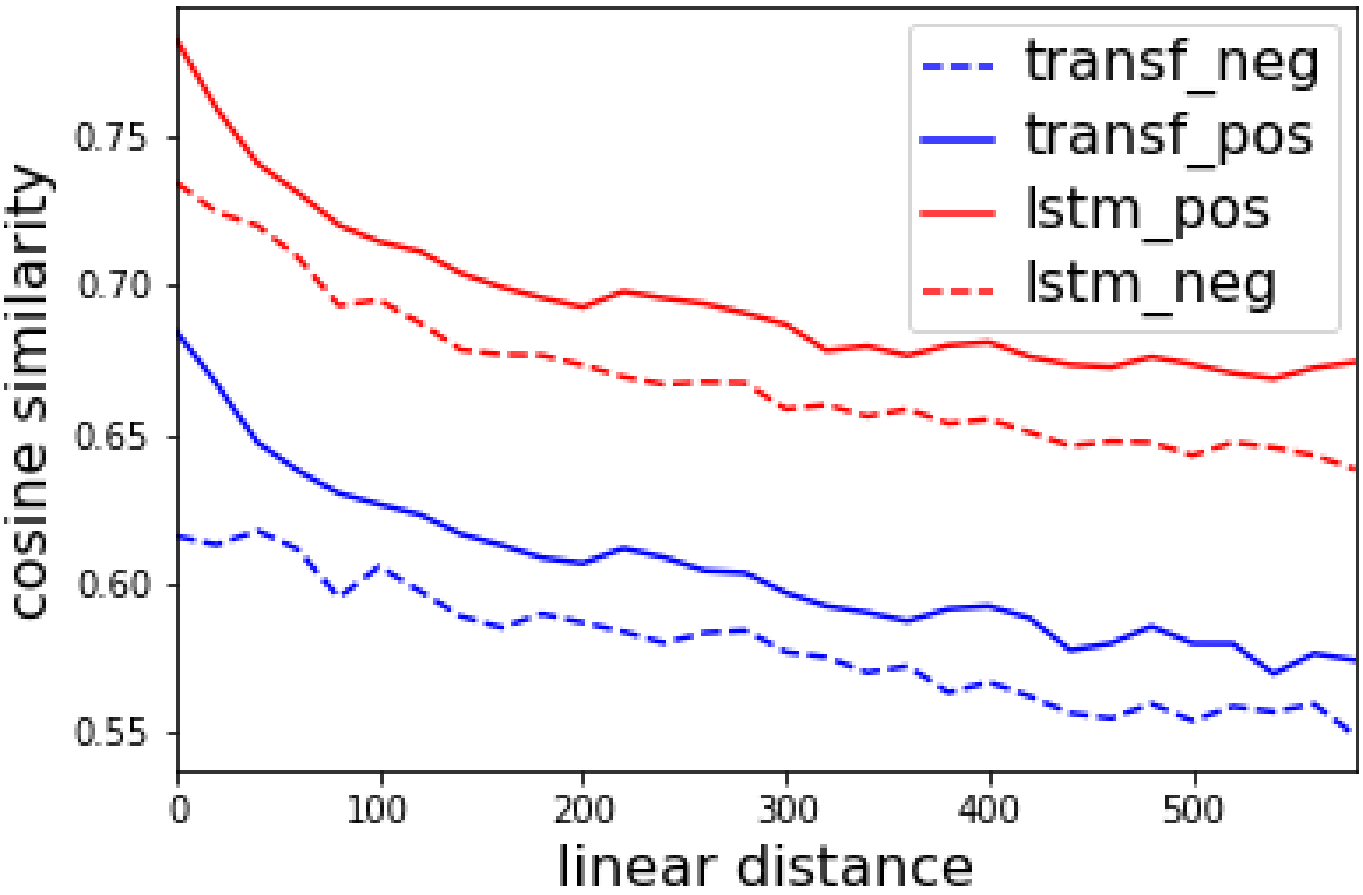
Distinguishing entities - global (document)

he was elected to be president of the People's Republic of China, and chairman of the Central Military Commission. Yeping Wang was born in Shanghai in 1926. She studied in Shanghai Foreign Language College, and started working in 1949. For a long time, she

- Take all pronominal mentions in a document
- Task: distinguish those that refer to the same entity from those that refer to different entities

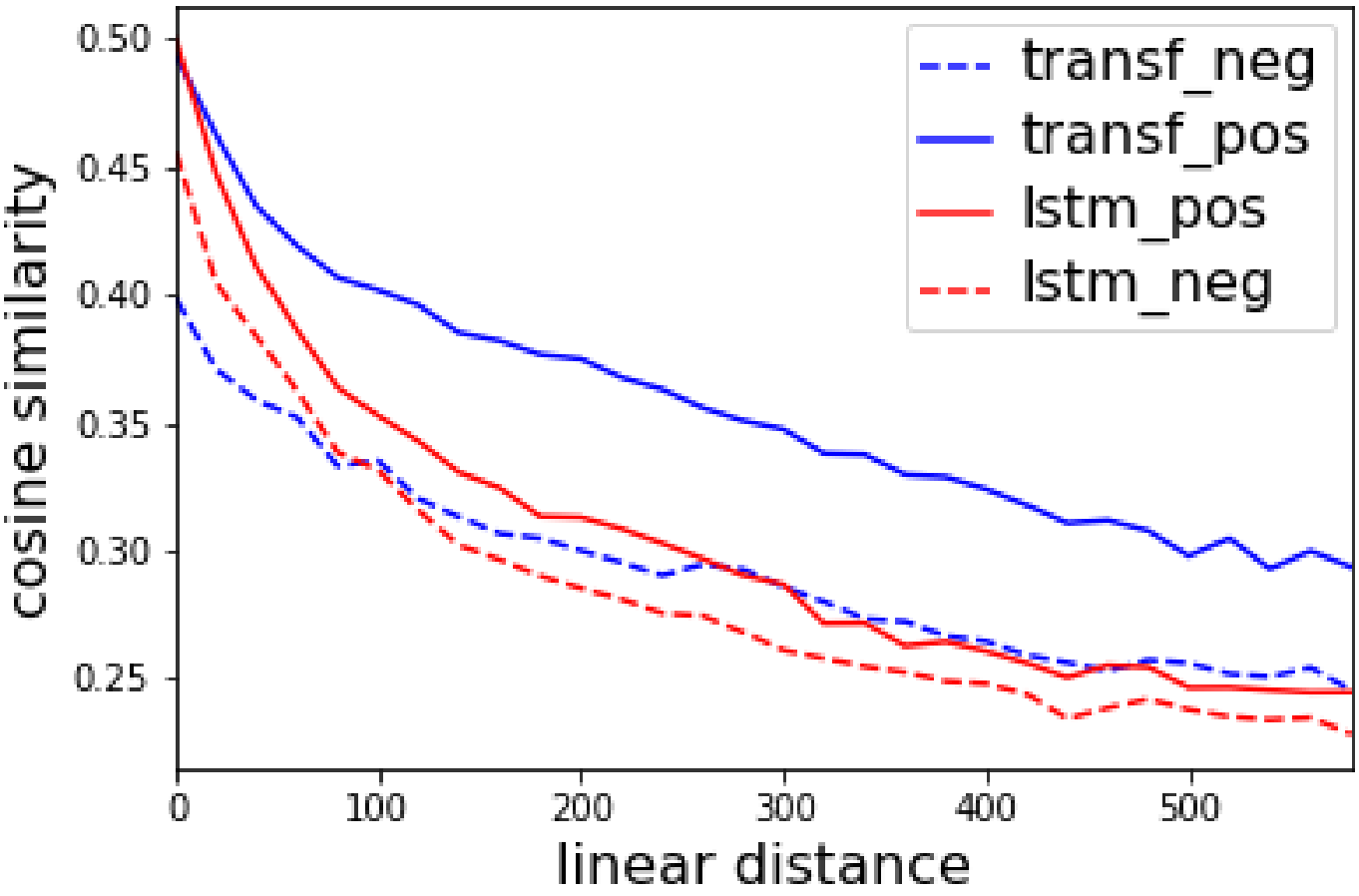
  only with the same pronoun pair (she-she)

Distinguishing entities - global: unsupervised



Linear distance in the discourse vs. cosine distance, for all the mention pairs with the same token pronoun.

Distinguishing entities - global: supervised



Linear distance in the discourse vs. cosine distance, for all the mention pairs with the same token pronoun.

Conclusion

language models

- learn syntax – grammatical constraints of anaphoric reference

Conclusion

language models

- learn syntax – grammatical constraints of anaphoric reference
- learn semantics (reference), to a surprising degree
- surprising?

Conclusion

language models

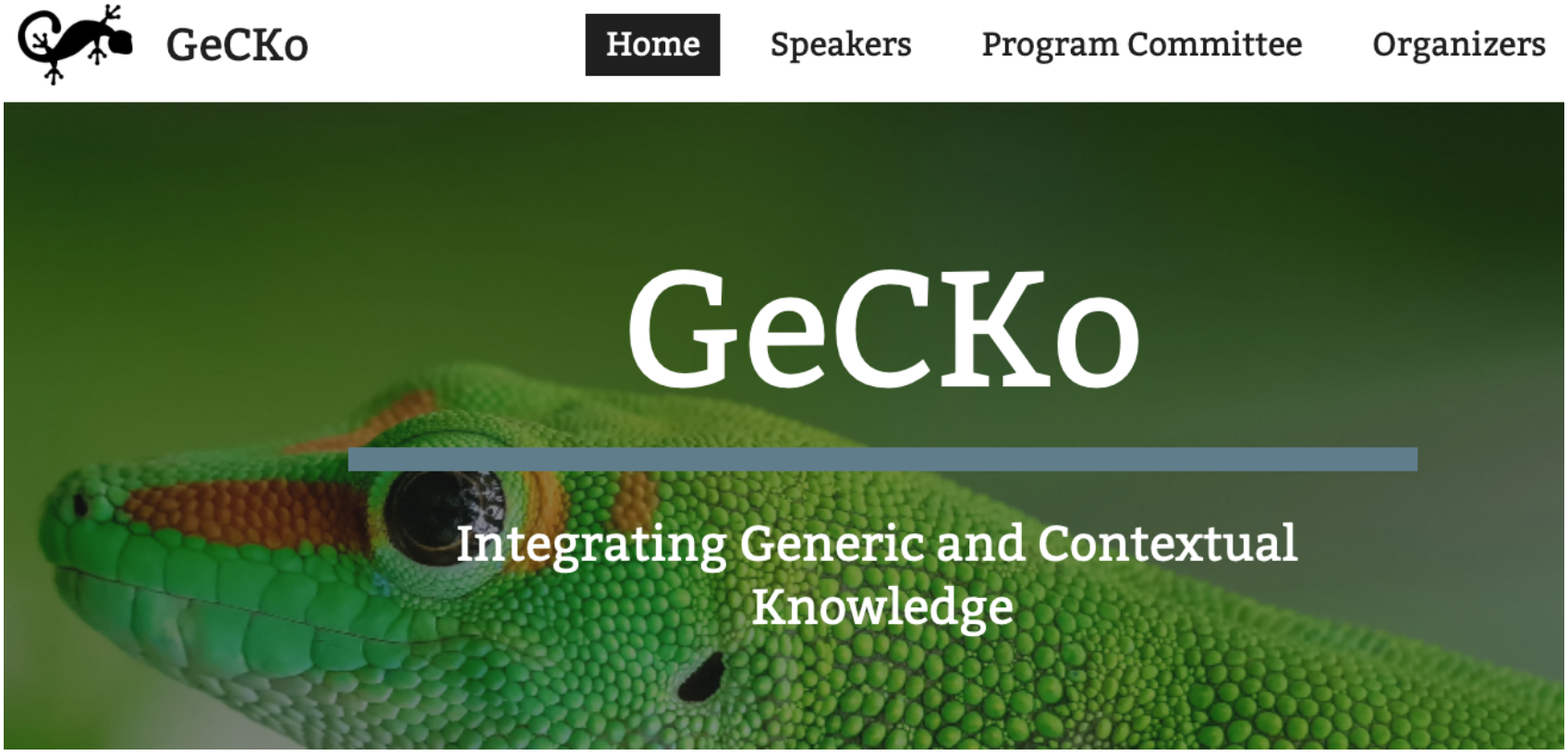
- learn syntax – grammatical constraints of anaphoric reference
- learn semantics (reference), to a surprising degree
- surprising?
- Entities not “there” in text
- . . . but neither is syntax

Conclusion

language models

- learn syntax – grammatical constraints of anaphoric reference
- learn semantics (reference), to a surprising degree
- surprising?
- Entities not “there” in text
- . . . but neither is syntax
- syntax: formal property of language. Can be acquired by observing co-occurrence patterns (think linguists).
- reference: depends on the relationship between language and the world.

Let’s continue the discussion!



May 18 2020, Universitat Pompeu Fabra

https://sites.google.com/view/gecko2020

Language models capture anaphoric coreference to a large extent

Gemma Boleda
Universitat Pompeu Fabra / ICREA
joint work with Ionut Sorodoc and Kristina Gulordava

DLBCN 2019
Barcelona, December 19-20 2019

amore
erc
European Research Council
Established by the European Commission
European Commission
Horizon 2020
European Union funding
for Research & Innovation
ICREA
upf. Universitat Pompeu Fabra Barcelona
Computational Linguistics and Linguistic Theory