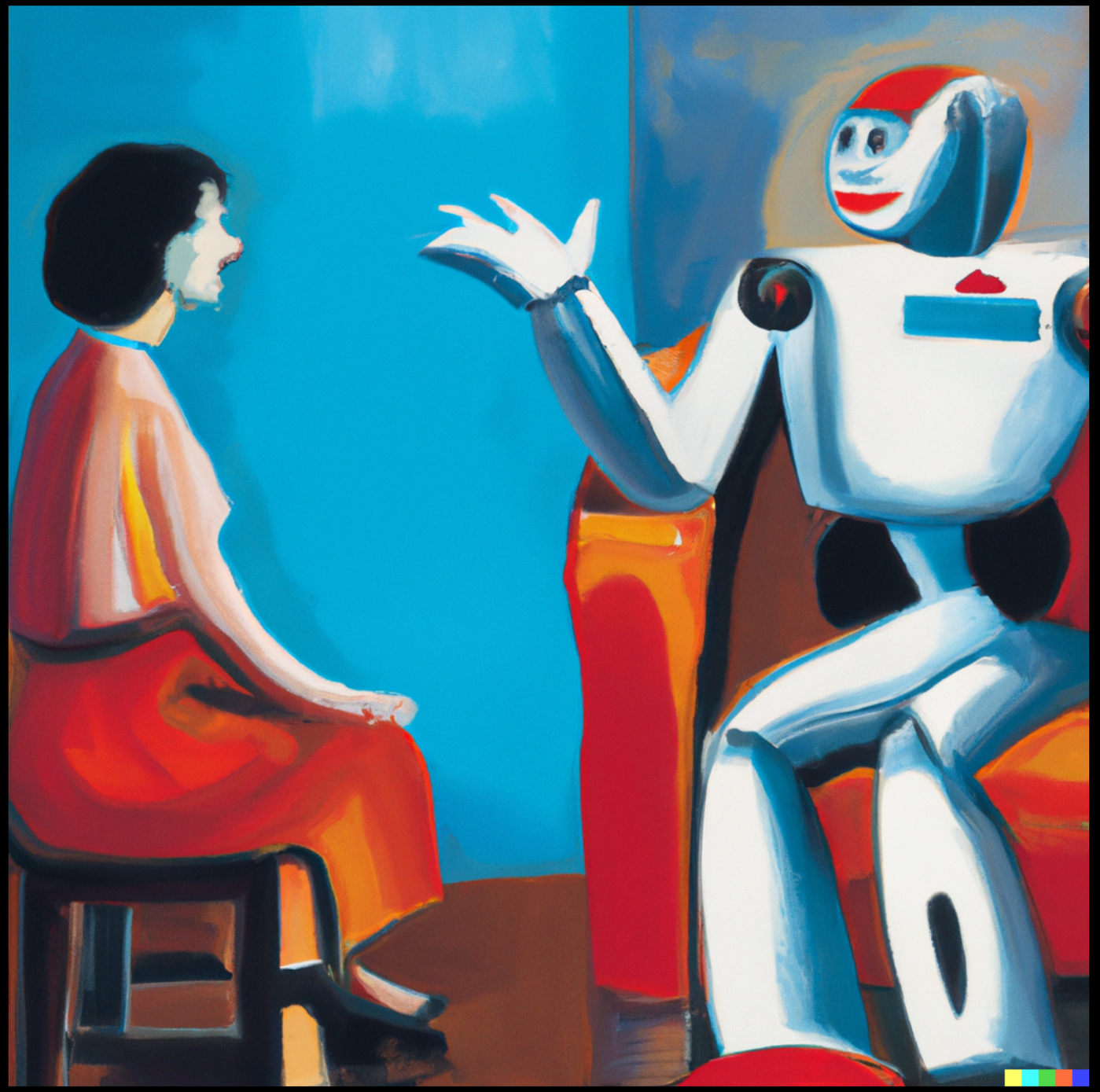
ILFC Seminar

Dissociating Formal and Functional Linguistic Competence in Large Language Models

Anna Ivanova, MIT

September 13, 2023



The rise of AI psychology

Theory of Mind May Have Spontaneously Emerged in Large Language Models

Michal Kosinski

Assessing Working Memory Capacity of ChatGPT

Dongyu Gong[1,2] (dongyu.gong@psy.ox.ac.uk)

Identifying and Manipulating the Personality Traits of Language Models

Graham Caron and **Shashank Srivastava**

Communication

The Political Biases of ChatGPT

David Rozado

AI Psychometrics: Using psychometric inventories to obtain psychological profiles of large language models

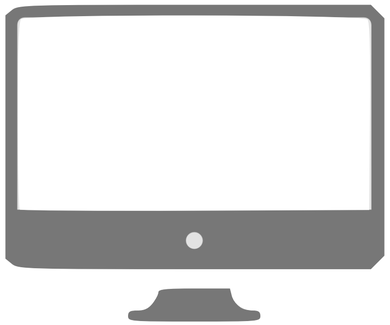
Max Pellert, Clemens M. Lechner, Claudia Wagner, Beatrice Rammstedt & Markus Strohmaier

Can AI language models replace human participants?

Danica Dillion,[1] Niket Tandon,[2] Yuling Gu,[2] and Kurt Gray[1,*,@]

The Turing test

My name is Tom



My name is Tom

When evaluating LLM capabilities,
we should dissociate language and thought.

Roadmap

- Introduction
- Formal vs. functional linguistic competence
- Formal competence: grammar
- Functional competence
  - math
  - world knowledge
- Case study: event knowledge
- Toward better models

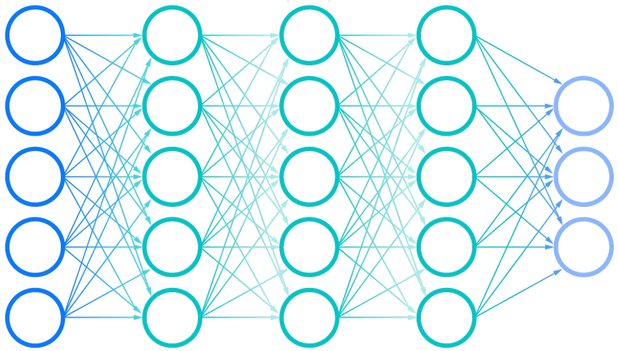
Roadmap

- **Introduction**
- Formal vs. functional linguistic competence
- Formal competence: grammar
- Functional competence
  - math
  - world knowledge
- Case study: event knowledge
- Toward better models

Large Language Models

Large language models today are…

neural networks



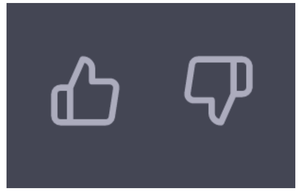
trained on large amounts of text

on the word-in-context prediction task

The fox chased the XXXX

The fox XXXX the rabbit.

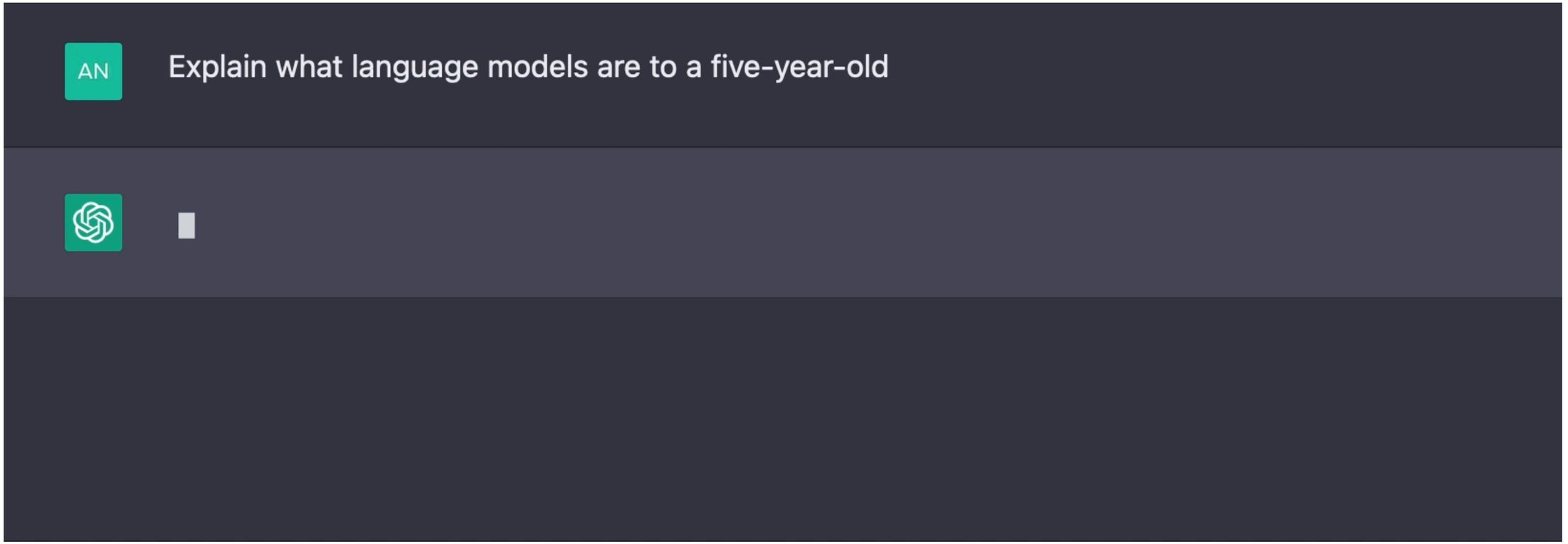
and sometimes fine-tuned on additional tasks



Large Language Models

These models are becoming very good at generating paragraphs of text in response to a prompt.

Example: ChatGPT (released in December 2022)

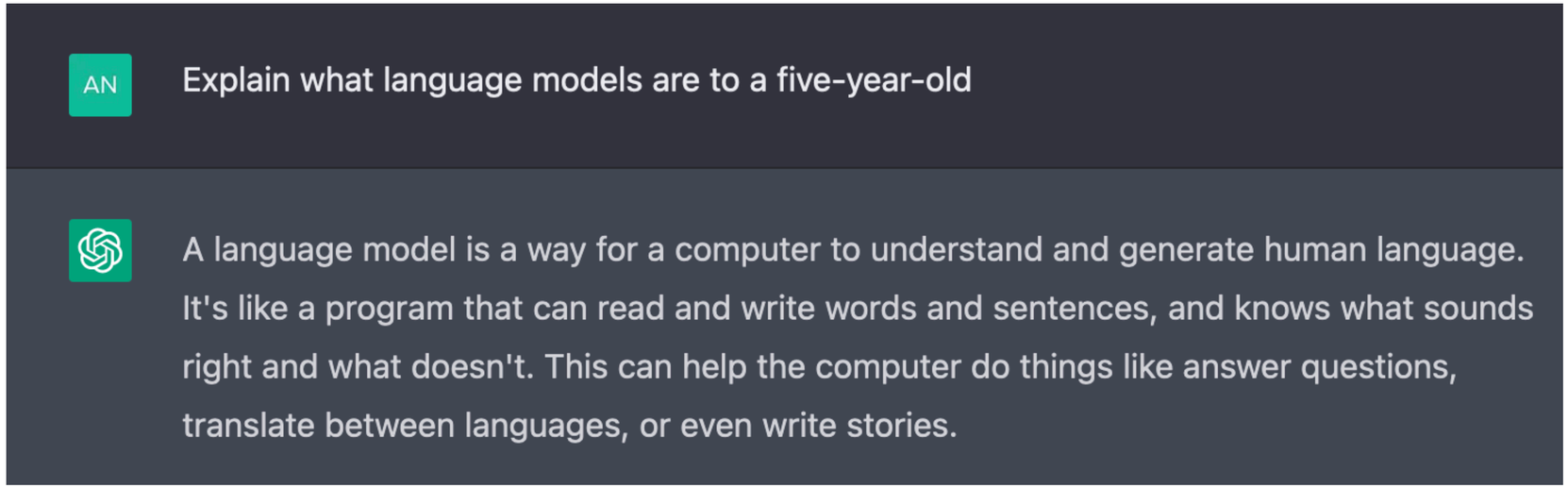


Large Language Models

These models are becoming very good at generating paragraphs of text in response to a prompt.

Exciting features:

- Novel sentences (not taken verbatim from the web)
- Grammatically correct
- (Seemingly) meaningful



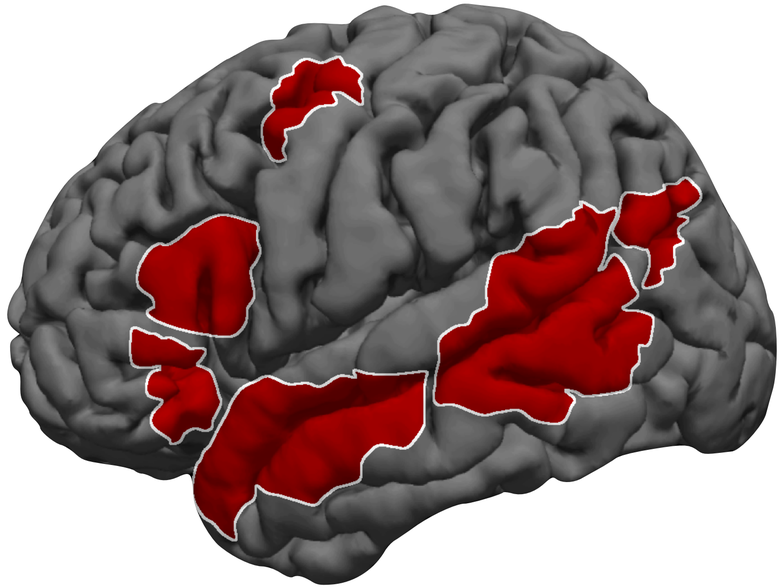
Roadmap

- Introduction
- **Formal vs. functional linguistic competence**
- Formal competence: grammar
- Functional competence
  - math
  - world knowledge
- Case study: event knowledge
- Toward better models

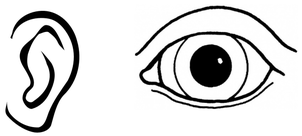
Language and the brain

Language processing in the brain takes place within a separate network.

Words, phrases, sentences



Listening and reading

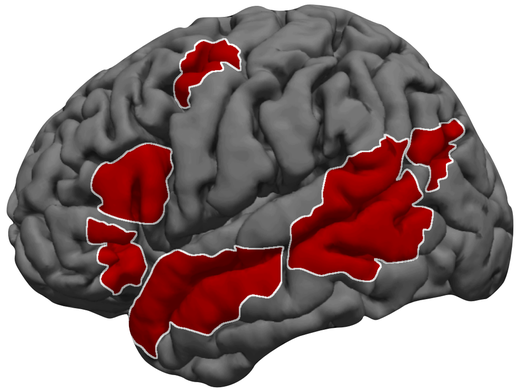


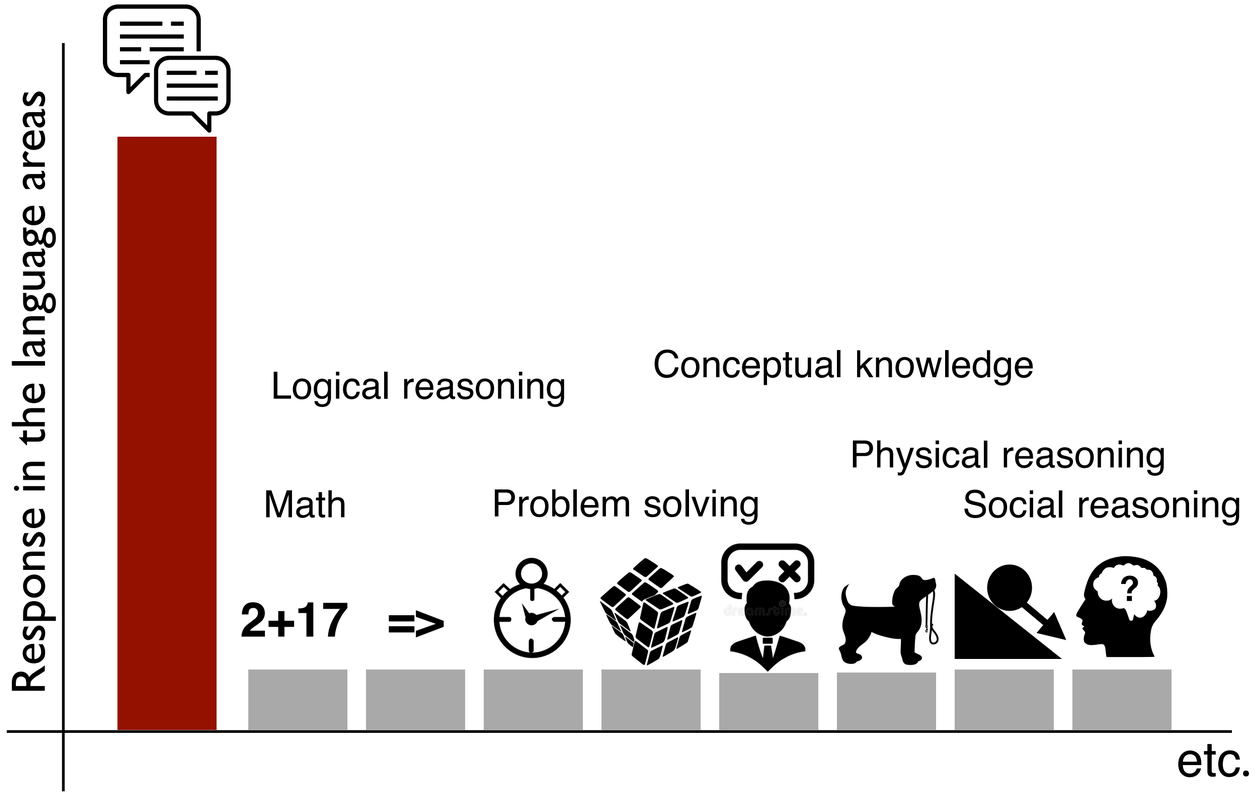
Speaking and writing

Fedorenko et al, 2010, 2011; Scott et al, 2017; Hu, Small et al, 2022; etc, etc

Language and the brain

Language areas **show little/no response** when we engage in diverse thought-related activities.

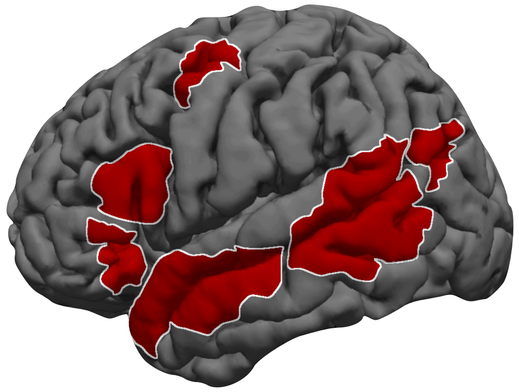




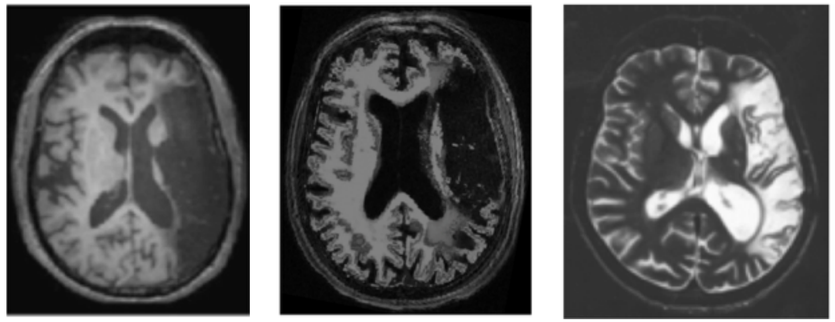
slide adapted from Ev Fedorenko; for a review, see Fedorenko & Varley, 2016

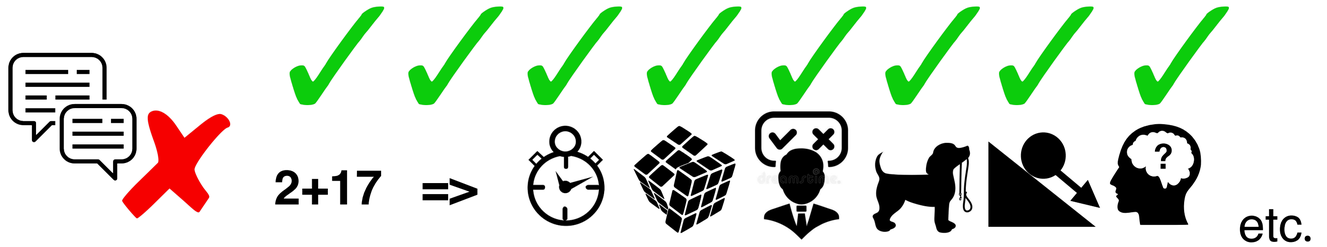
Language and the brain

Language areas can be **damaged** with little/no effect on thought-related activities.



Sample patients' lesions:

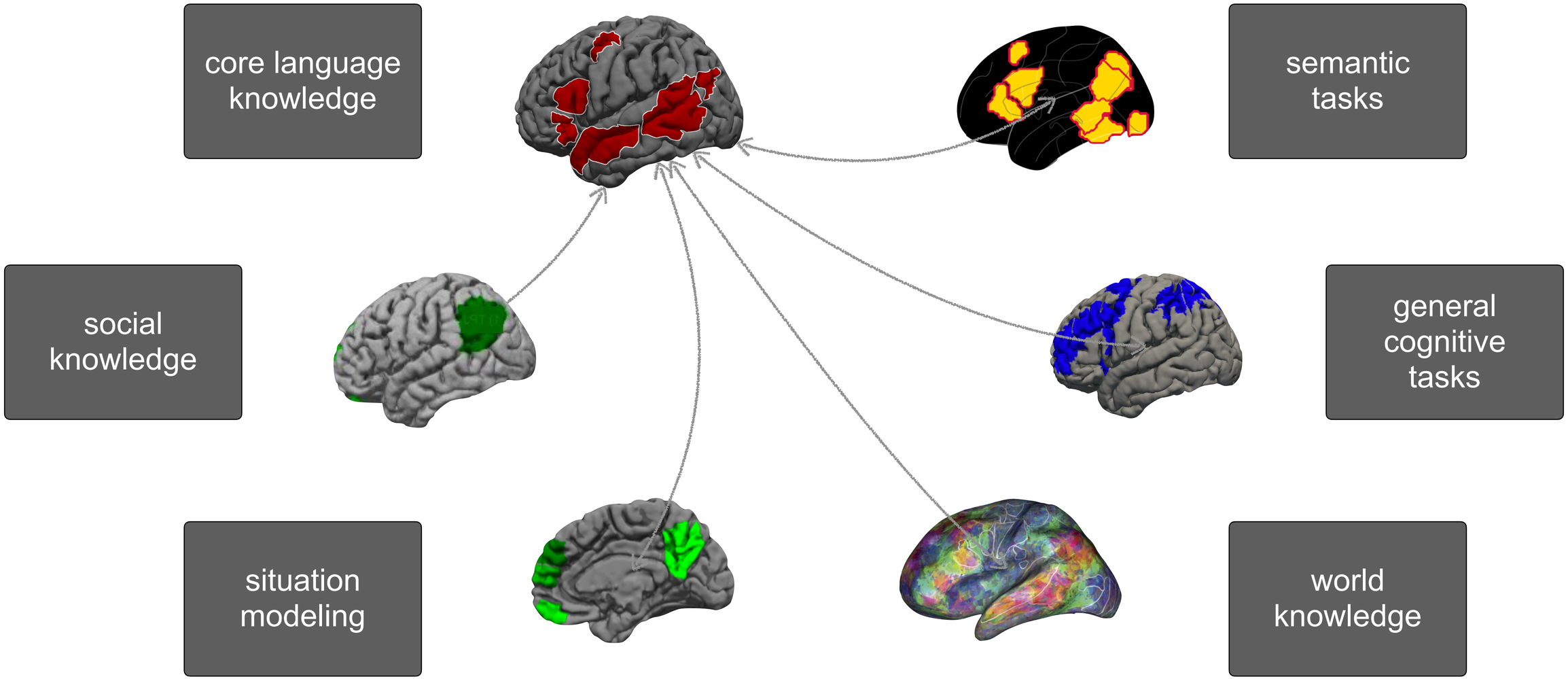




etc.

slide adapted from Ev Fedorenko; for a review, see Fedorenko & Varley, 2016

Formal and functional linguistic competence

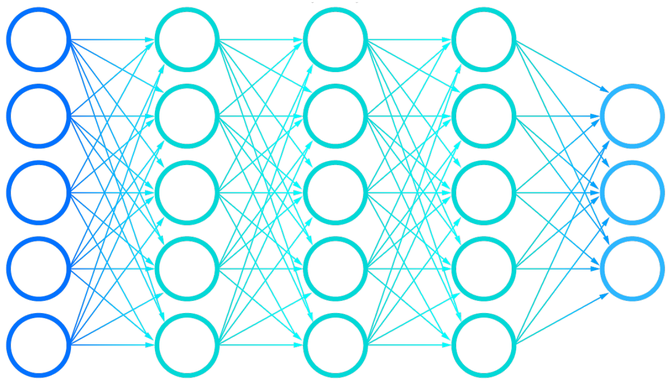


Mahowald, Ivanova et al, arXiv

Formal and functional linguistic competence

core language knowledge

semantic tasks



social knowledge

general cognitive tasks

situation modeling

world knowledge

Mahowald, Ivanova et al, arXiv

Formal and functional linguistic competence

FORMAL COMPETENCE
(language-specific)

core language knowledge

FUNCTIONAL COMPETENCE
(non-language-specific)

social knowledge

situation modeling

semantic tasks

general cognitive tasks

world knowledge

Mahowald, Ivanova et al, arXiv

Formal and functional linguistic competence

FORMAL COMPETENCE
(language-specific)

The keys to the cabinet **are** on the table.

FUNCTIONAL COMPETENCE
(non-language-specific)

Six birds were sitting on a tree. Three flew away, but then one came back. There are now **four** birds.

Roadmap

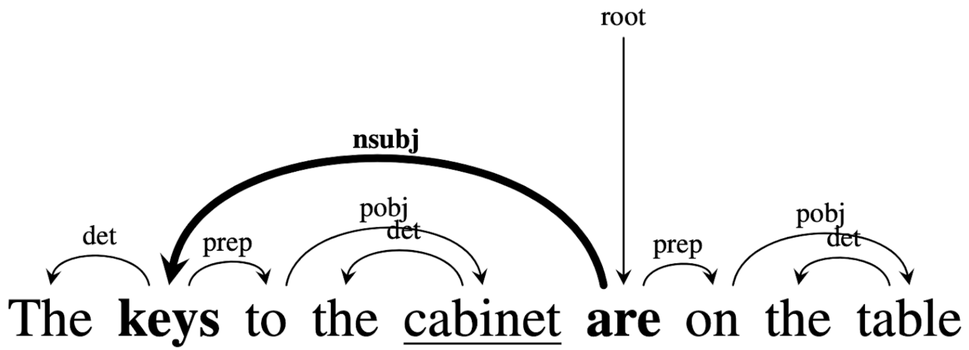
- Introduction
- Formal vs. functional linguistic competence
- **Formal competence: grammar**
- Functional competence
  - math
  - world knowledge
- Case study: event knowledge
- Toward better models

Formal competence in large language models

Statistical Language Modeling	Structure
Shannon (1940s and 1950s): ngrams, statistical language models, prediction engines	Chomsky (1957): “Despite the undeniable interest and importance of semantic and statistical studies of language, they appear to have no direct relevance to the problem of determining or characterizing the set of grammatical utterances.”
Rumelhart & McClelland (1986): “Thus the behavior of the model was lawful even though it contained no explicit rules.”	Pinker & Prince (1988): “We conclude that connectionists' claims about the dispensability of rules in explanations in the psychology of language must be rejected”
2010s and 2020s: Manning et al. (2020): “However, we demonstrate that modern deep contextual language models learn major aspects of [linguistics] structure, without any explicit supervision.”	Everaert et al. (2015) on structures, not strings: “Applying analytical or statistical tools to huge corpora of data in an effort to elucidate the intriguing properties of parasitic gaps will not work."
2023: Steve Piantadosi says “Modern language models refute Chomsky’s approach to language”	Roni Katzir (2023): “While LLMs are successful as engineering tools, we saw that they are very poor theories of human linguistic cognition.”

adapted from Kyle Mahowald

Formal competence in large language models



KY

Complete the sentence "Far beyond the houses, the person who has the cats that the dogs with the cute faces and fluffy tails chased"

is probably worried about the safety of their beloved feline pets and hoping they can escape unharmed.

adapted from Kyle Mahowald

Formal competence in large language models

Systematic evaluation:
BLiMP (the Benchmark of Linguistic Minimal Pairs for English; Warstadt et al, 2019)

Phenomenon	N	Acceptable Example	Unacceptable Example
Anaphor agr.	2	*Many girls insulted themselves.*	*Many girls insulted herself.*
Arg. structure	9	*Rose wasn't disturbing Mark.*	*Rose wasn't boasting Mark.*
Binding	7	*Carlos said that Lori helped him.*	*Carlos said that Lori helped himself.*
Control/raising	5	*There was bound to be a fish escaping.*	*There was unable to be a fish escaping.*
Det.-noun agr.	8	*Rachelle had bought that chair.*	*Rachelle had bought that chairs.*
Ellipsis	2	*Anne's doctor cleans one important book and Stacey cleans a few.*	*Anne's doctor cleans one book and Stacey cleans a few important.*
Filler-gap	7	*Brett knew what many waiters find.*	*Brett knew that many waiters find.*
Irregular forms	2	*Aaron broke the unicycle.*	*Aaron broken the unicycle.*
Island effects	8	*Whose hat should Tonya wear?*	*Whose should Tonya wear hat?*
NPI licensing	7	*The truck has clearly tipped over.*	*The truck has ever tipped over.*
Quantifiers	4	*No boy knew fewer than six guys.*	*No boy knew at most six guys.*
Subject-verb agr.	6	*These casseroles disgust Kayla.*	*These casseroles disgusts Kayla.*

Model	Overall
GPT-2	81.5
Human	88.6

davinci (175B)	**0.84**
GPT-NeoX (20B)	0.839
TNLG v2 (6.7B)	0.835
GPT-J (6B)	0.834

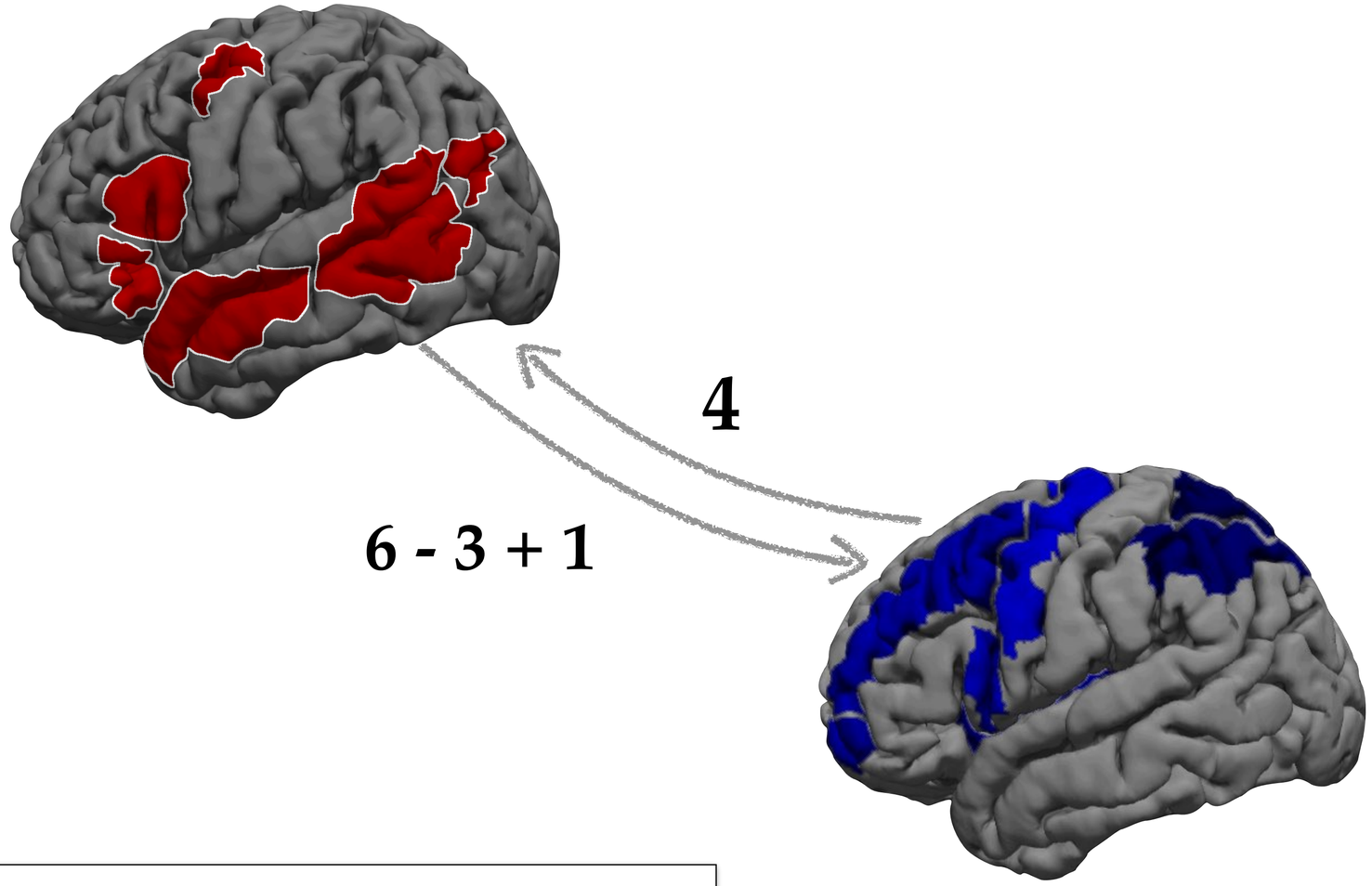
HELM

Formal competence in large language models

Conclusion: large language models have essentially mastered formal linguistic competence.

Roadmap

- Introduction
- Formal vs. functional linguistic competence
- Formal competence: grammar
- **Functional competence**
  - **math**
  - world knowledge
- Case study: event knowledge
- Toward better models



Six birds were sitting on a tree. Three flew away, but then one came back.

How many birds are there now?

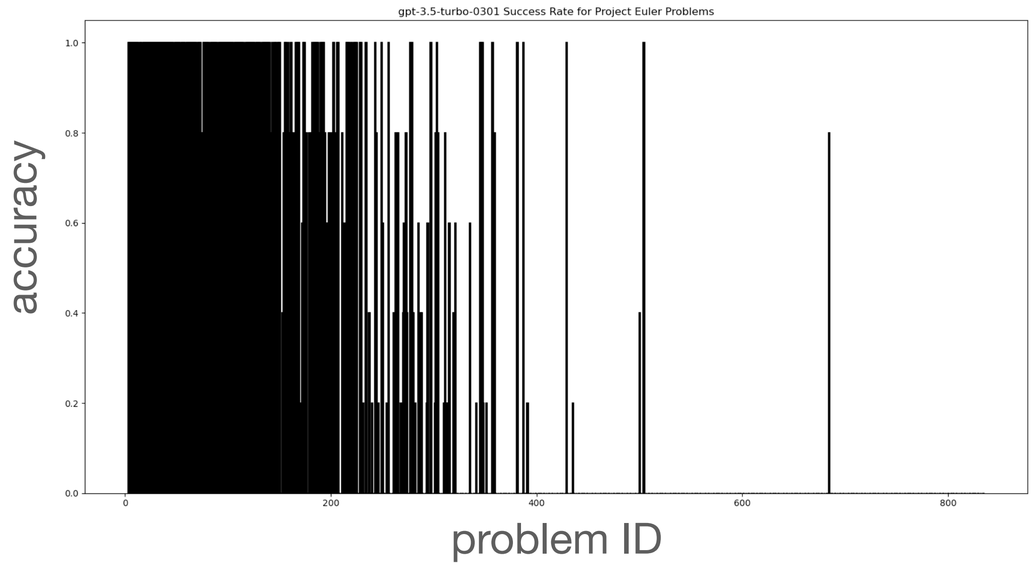
Large language models and math/reasoning

If behavior is identical, why should we care?

It’s not identical - the mechanism affects the system's ability to

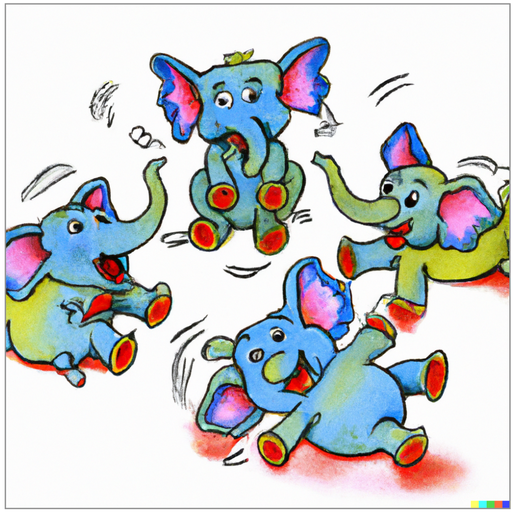
(a) generalize

Project Euler
(example from Chris Cundy)



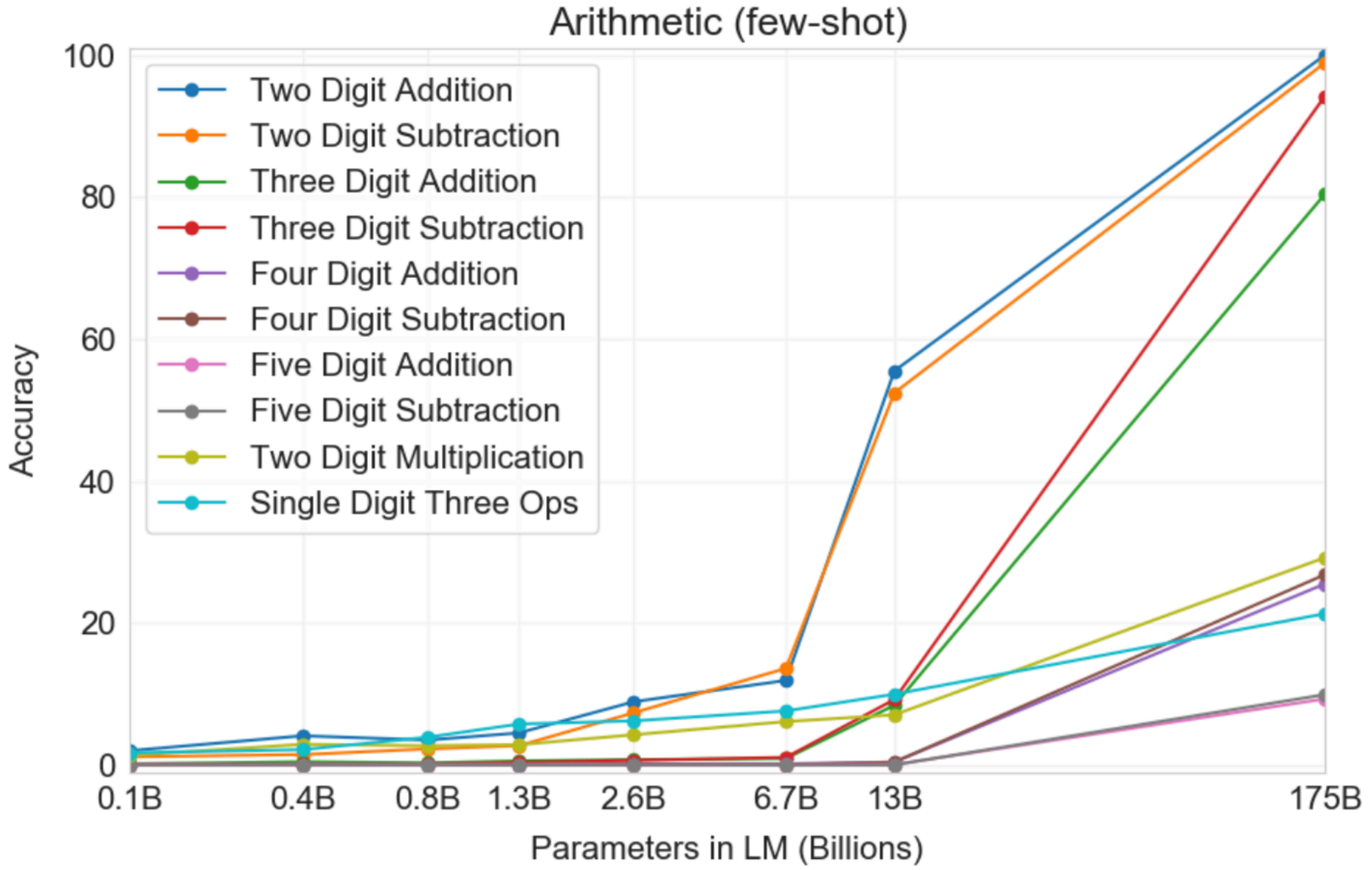
(b) integrate verbal and nonverbal input

DALL-E

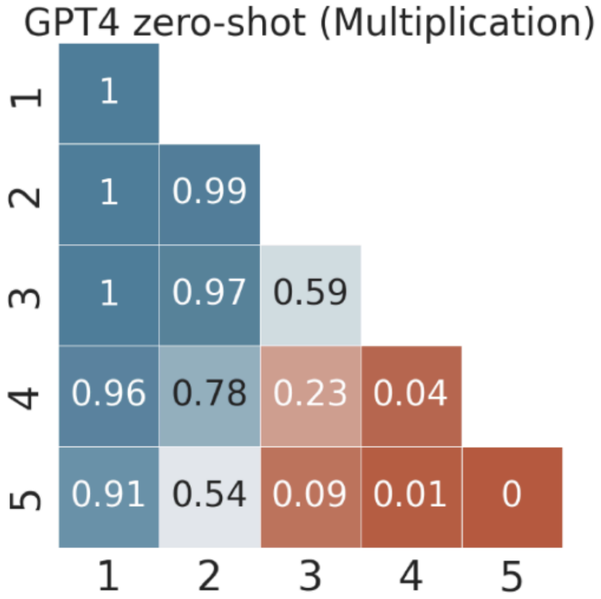


“five elephants playing, a cartoon"

Large language models and math/reasoning

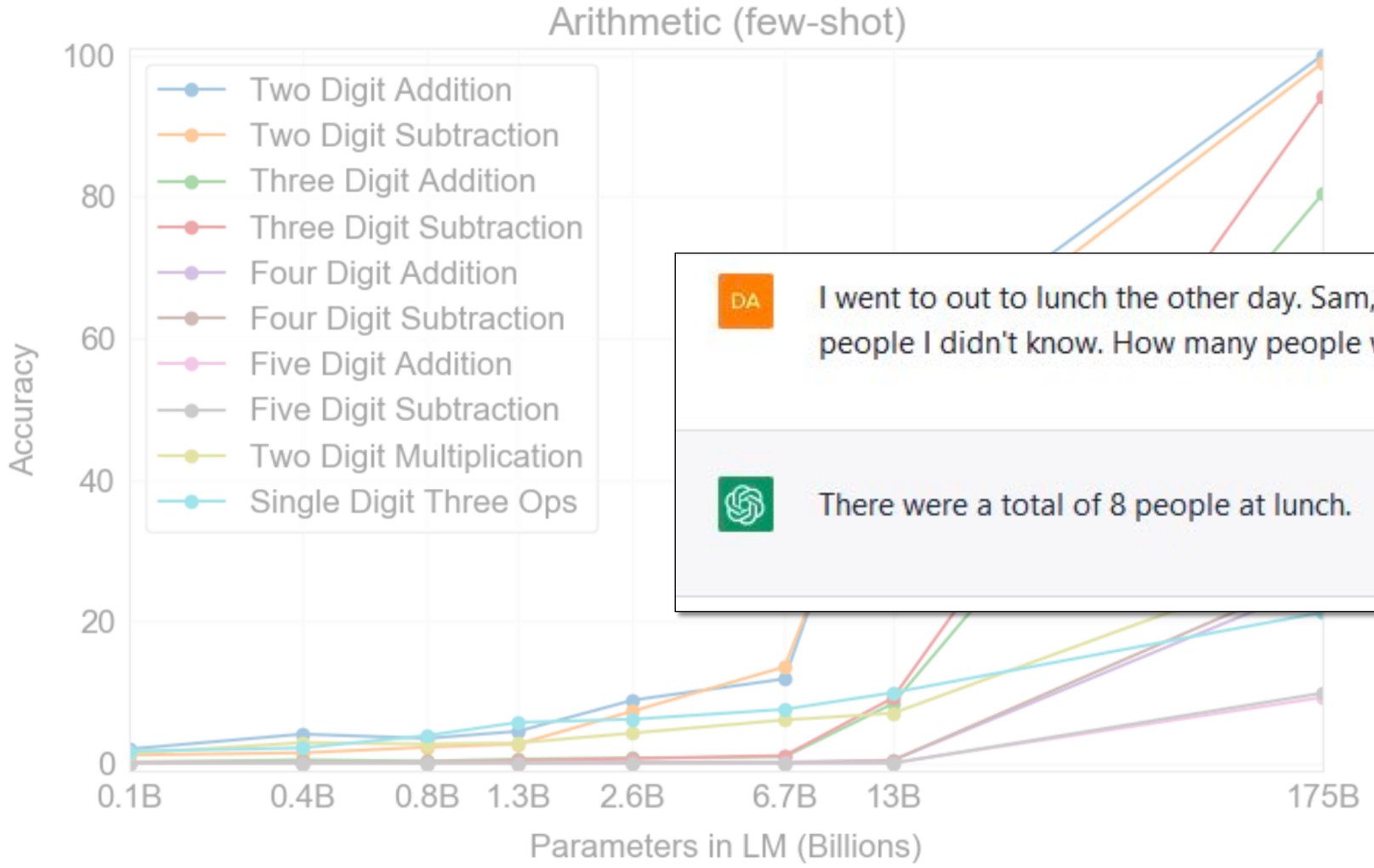


Brown et al (2020)



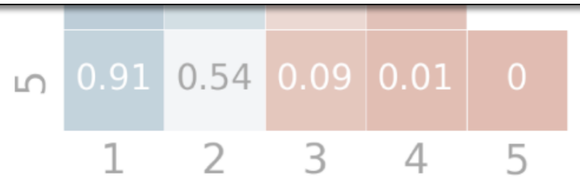
Dziri et al (2023)

Large language models and math/reasoning



Brown et al (2020)

Gary Marcus, 2023:



Dziri et al (2023)

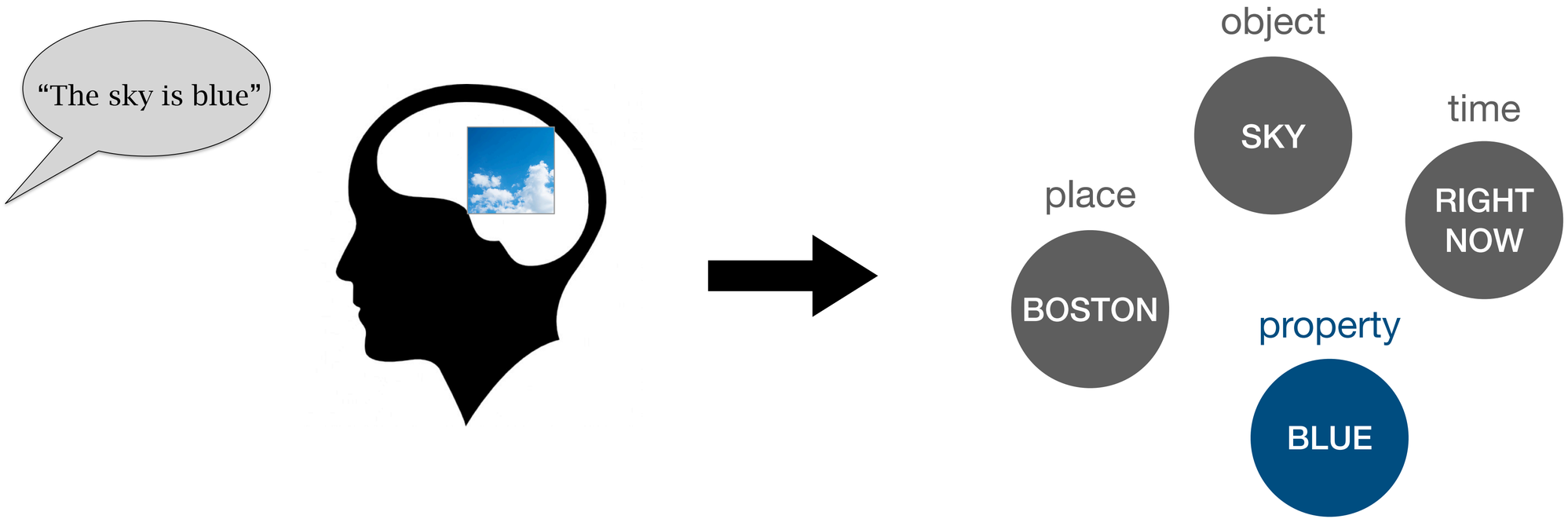
Large language models and math/reasoning

Conclusion: large language models diverge substantially from humans in their mathematical reasoning mechanisms

Roadmap

- Introduction
- Formal vs. functional linguistic competence
- Formal competence: grammar
- **Functional competence**
  - math
  - **world knowledge**
- Case study: event knowledge
- Toward better models

Large language models and world knowledge

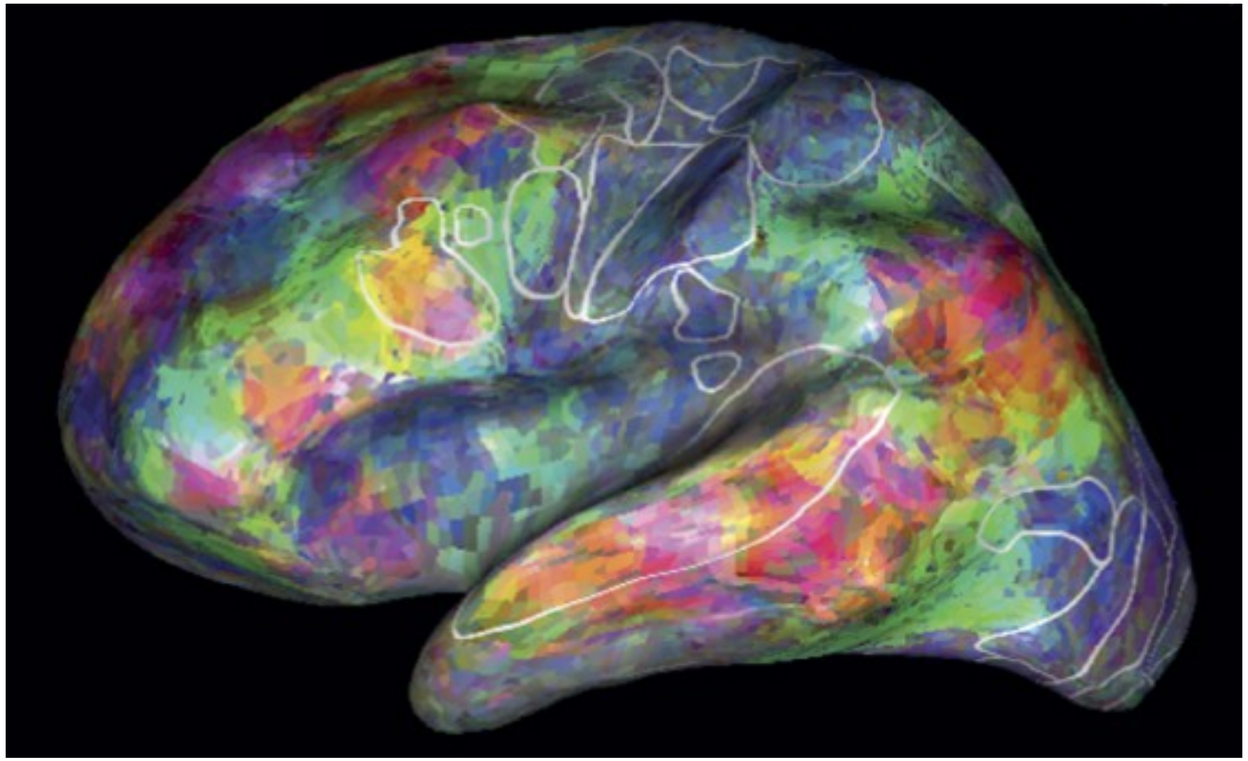


<u>More than language:</u>

- embedded in context (place and time are inferred)
- the same effect can be achieved without language input (e.g. by looking out the window)
- preserved in the face of language impairment

Large language models and world knowledge

Huth et al., 2016



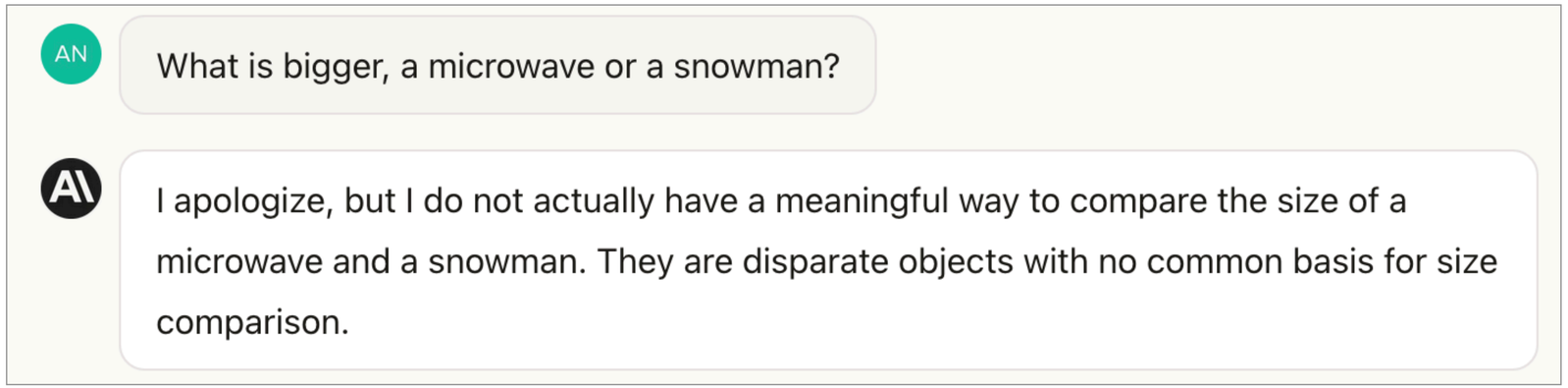
Large language models and world knowledge

If behavior is identical, why should we care?

It’s not identical - the mechanism affects the system's ability to

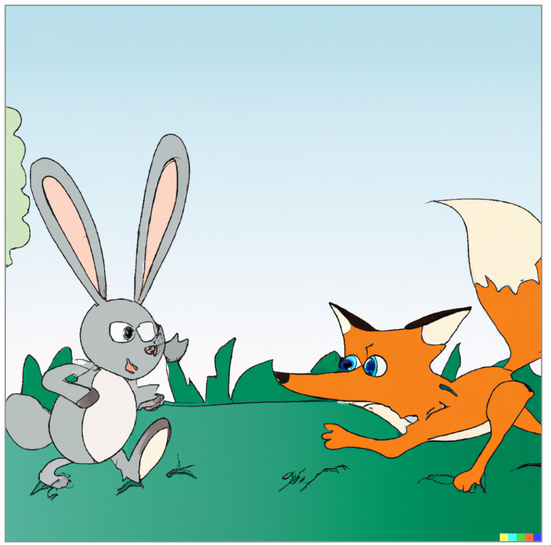
(a) generalize

(b) integrate verbal and nonverbal input



see also Ettinger, 2020; Kassner & Schütze, 2020; Talmor et al, 2020

DALL-E



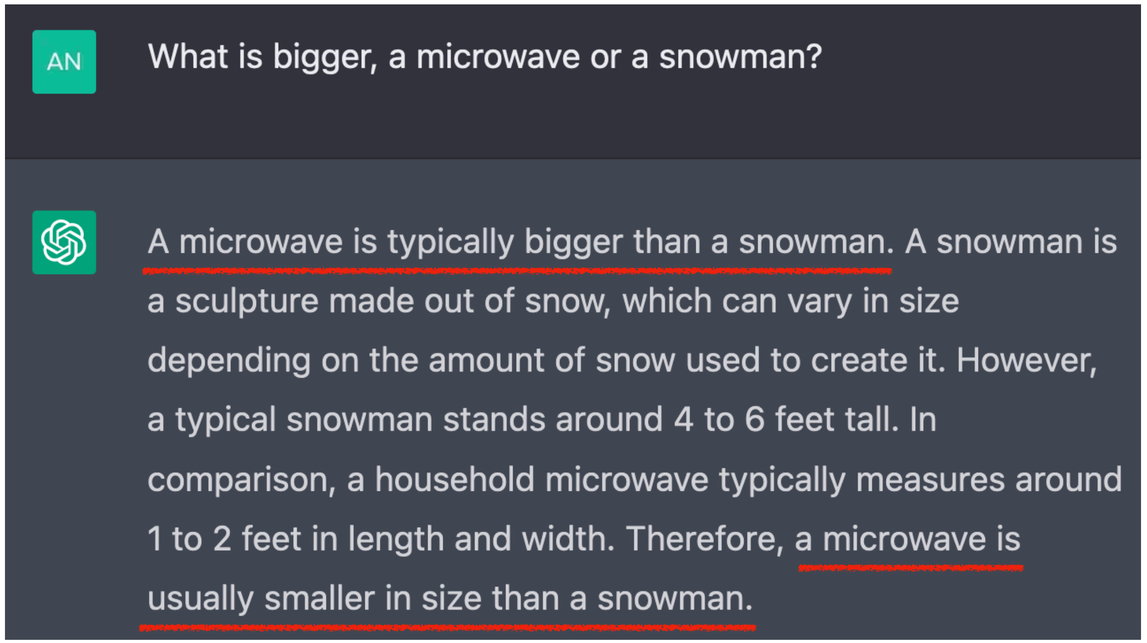
“a fox is chasing a rabbit, a cartoon"

Large language models and world knowledge

If behavior is identical, why should we care?

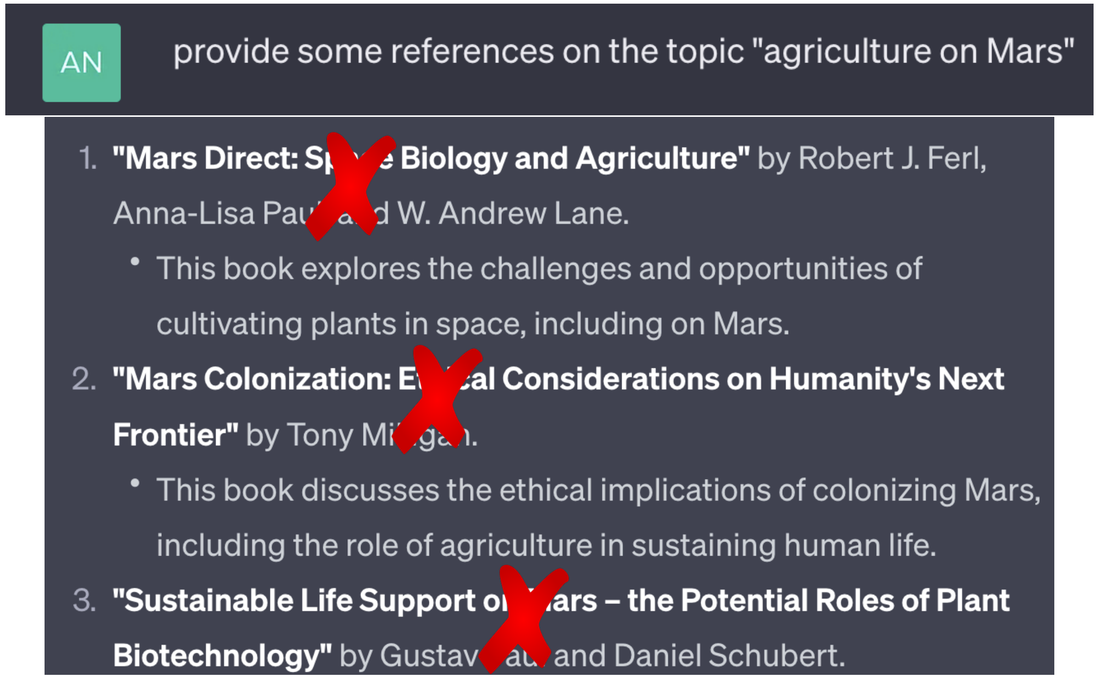
It’s not identical - the mechanism affects the system's ability to

(c) be consistent



see also Elazar et al, 2021; Ribeiro et al, 2019; Ravichander et al, 2020, etc

(d) be factually accurate



see also Liu et al., 2023

Large language models and world knowledge

Conclusion: large language models diverge substantially from humans in their world knowledge representations and use

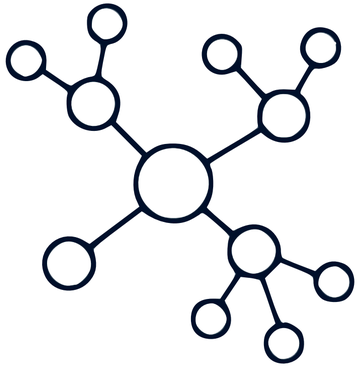
Roadmap

- Introduction
- Formal vs. functional linguistic competence
- Formal competence: grammar
- Functional competence
  - math
  - world knowledge
- **Case study: event knowledge**
- Toward better models

Language and event knowledge

Generalized Event Knowledge (GEK; McRae & Matsuki 2009)

- storage of **templates** of **common events** observed in the world



Single words activate GEK

arrest ⇄ cop

arrest ⇄ criminal

Words rapidly **combine** to cue specific concepts in GEK

The **journalist** checked the **spelling**.

The **mechanic** checked the **brakes**.

Does GEK rely on language processing in humans?

slide adapted from Carina Kauf

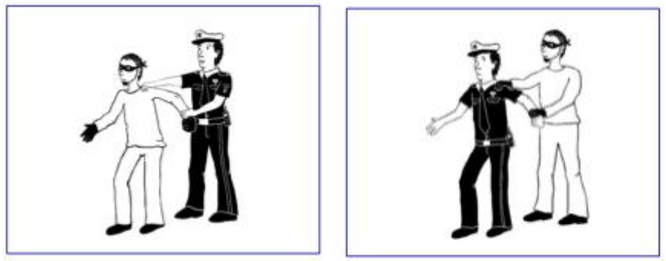
Language and event knowledge

SENTENCES

The cop is arresting the criminal.

The criminal is arresting the cop.

PICTURES



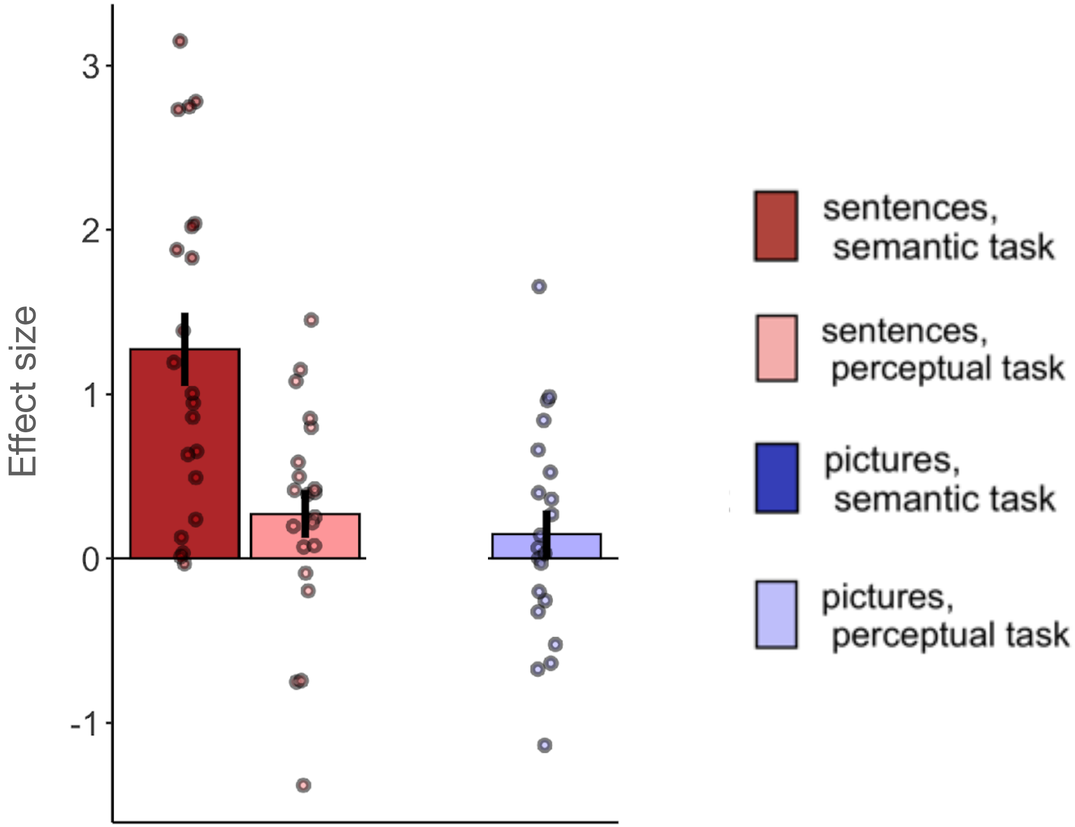
TASKS

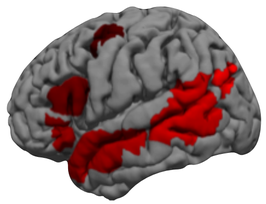
SEMANTIC

(plausible or implausible?)

PERCEPTUAL

(moving left or right?)

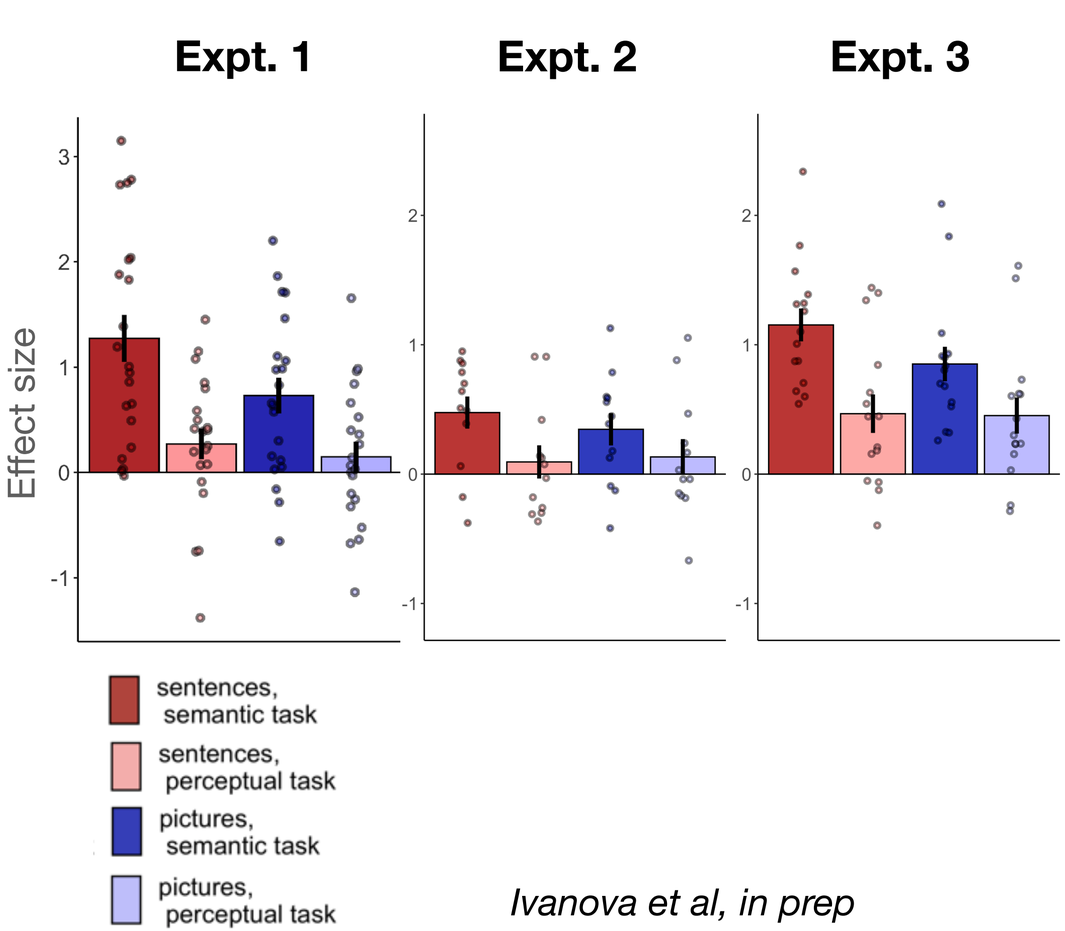




Ivanova et al, 2021

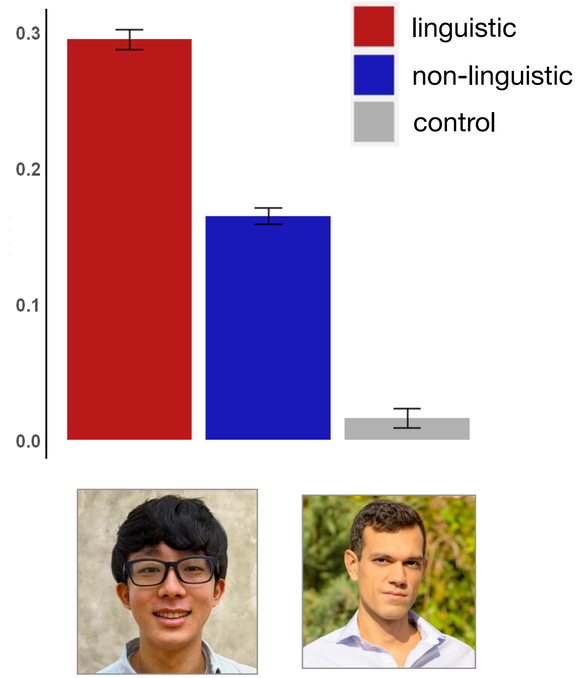
Language network & event semantics

The results generalize to other event semantics experiments, but not to object semantics.



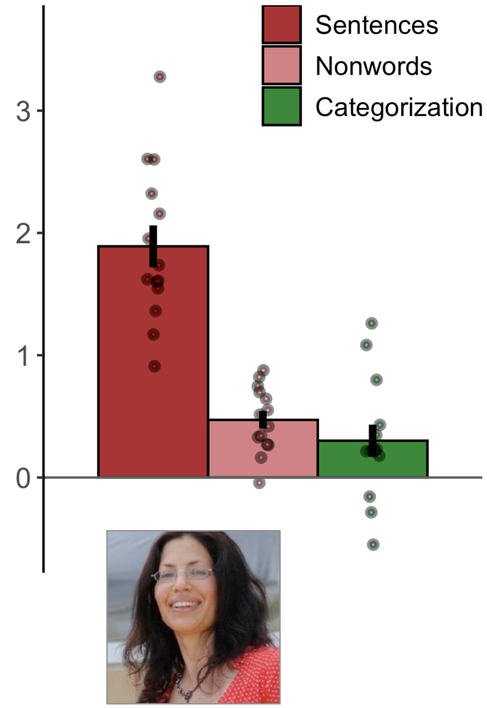
Ivanova et al, in prep

Naturalistic event viewing/listening



Sueoka, Paunov*, Ivanova et al, bioRxiv*

Object categorization



Benn, Ivanova* et al, 2023*

Language and event knowledge

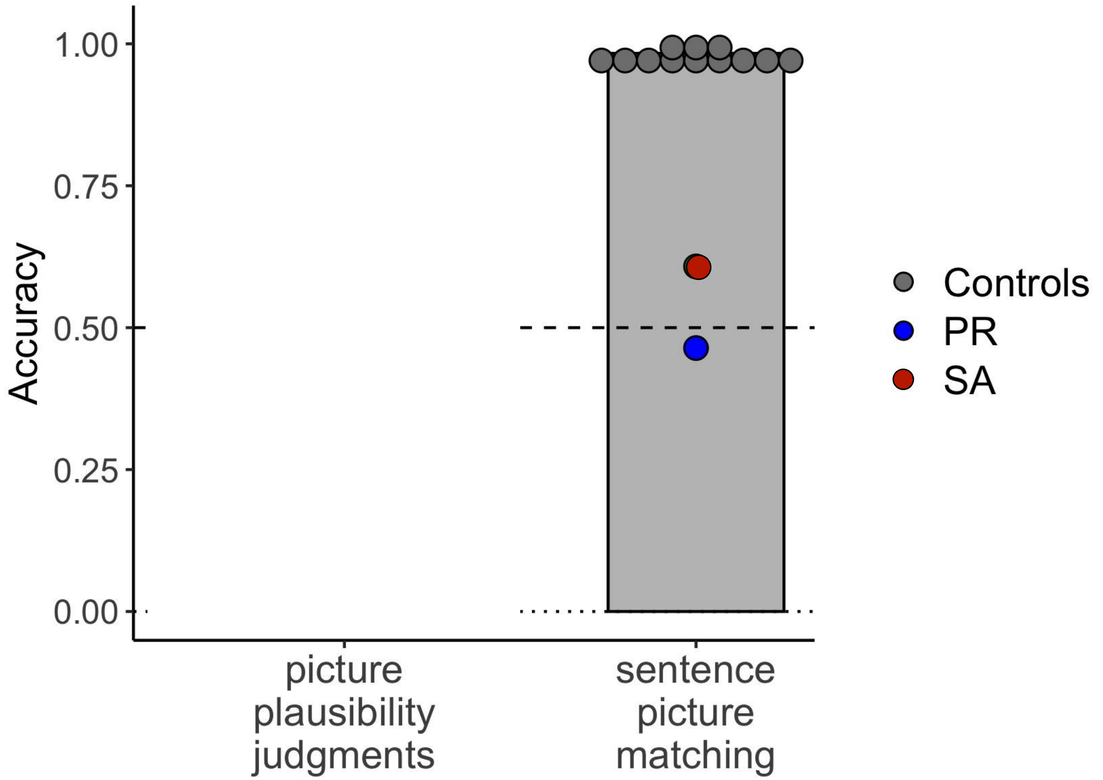
- Two participants with global aphasia (PR and SA)
- 12 age-matched controls

SENTENCES

The cop is arresting the criminal.

The criminal is arresting the cop.

PICTURES



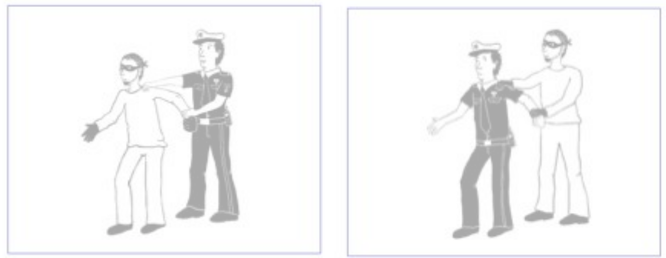
Ivanova et al, 2021

Language and event knowledge

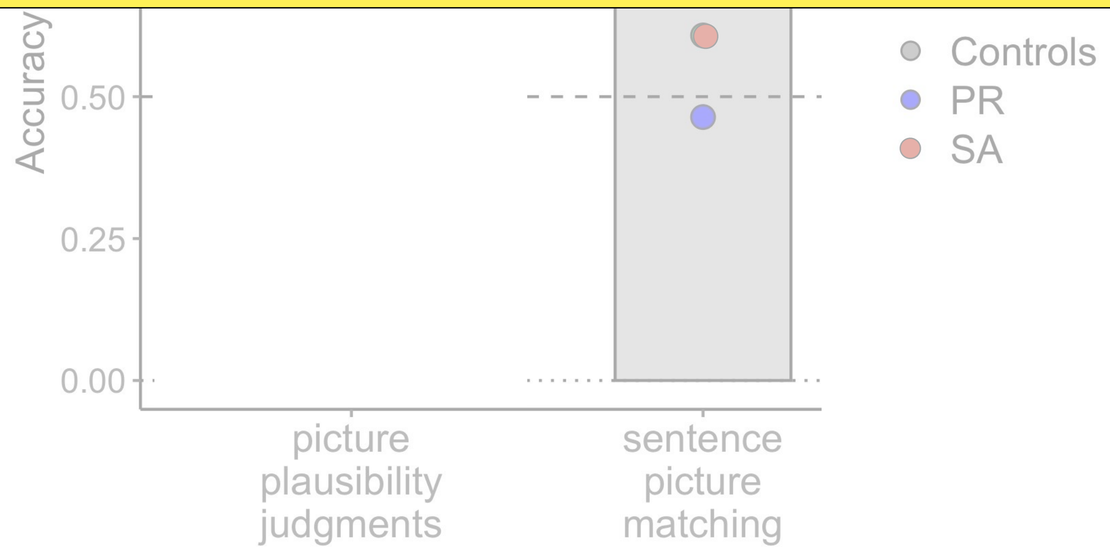
- Two participants with global aphasia (PR and SA)
- 12 age-matched controls

SENTENCES

PICTURES



The language network is **recruited but not required** for event semantics.



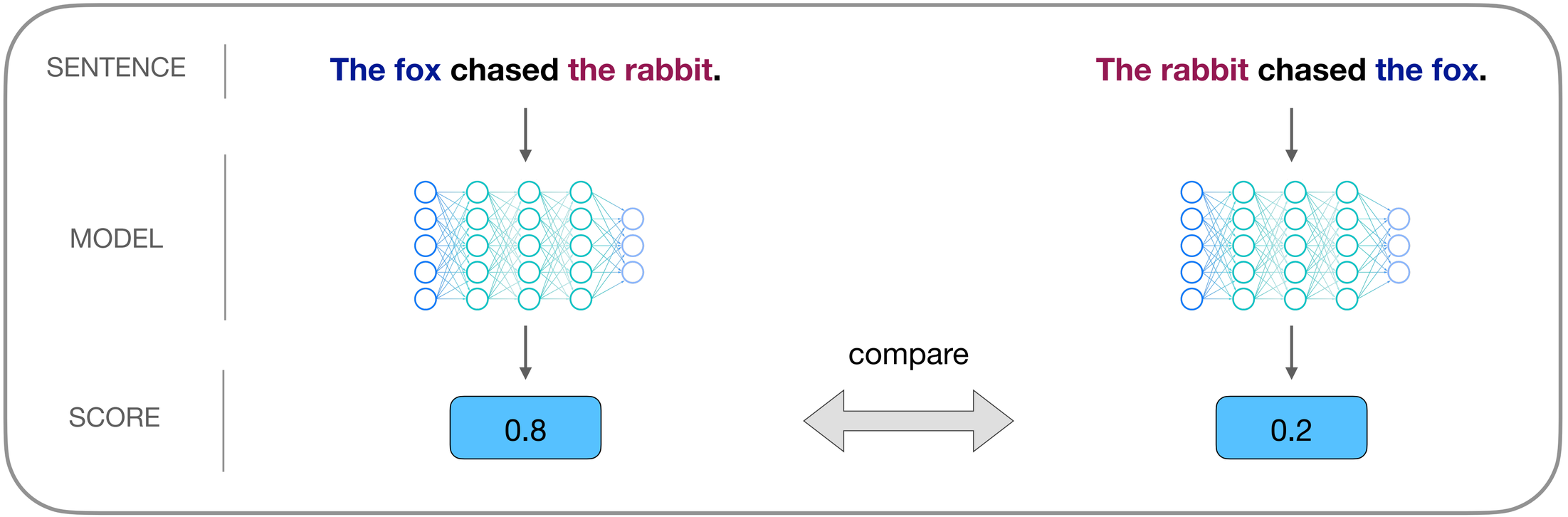
Ivanova et al, 2021

Language models and event knowledge

co-lead:
Carina Kauf

Does event knowledge naturally arise in language models?

Approach: minimal sentence pairs



Kauf, Ivanova* et al, arXiv*

Language models and event knowledge

Animate-Inanimate, impossible

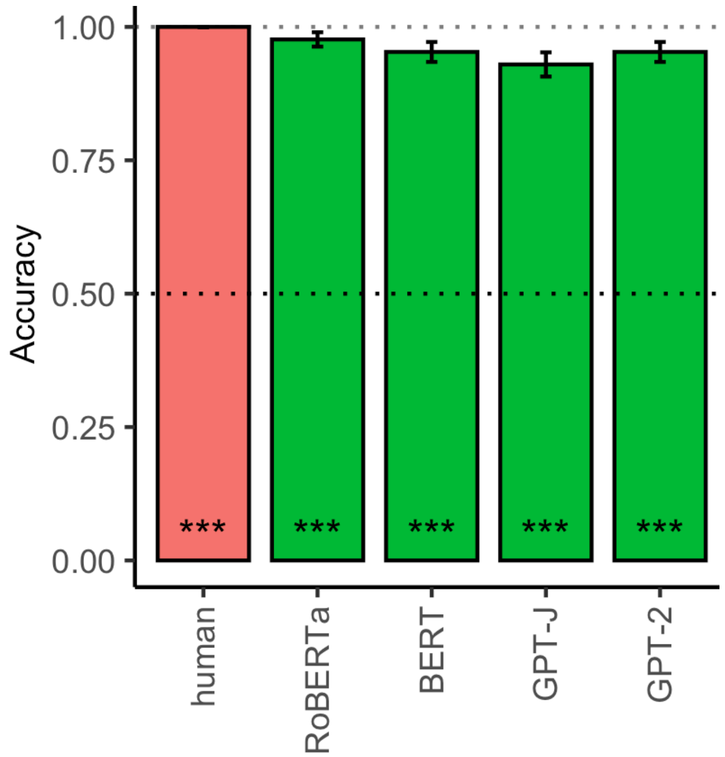
The teacher bought the laptop.

The laptop bought the teacher.

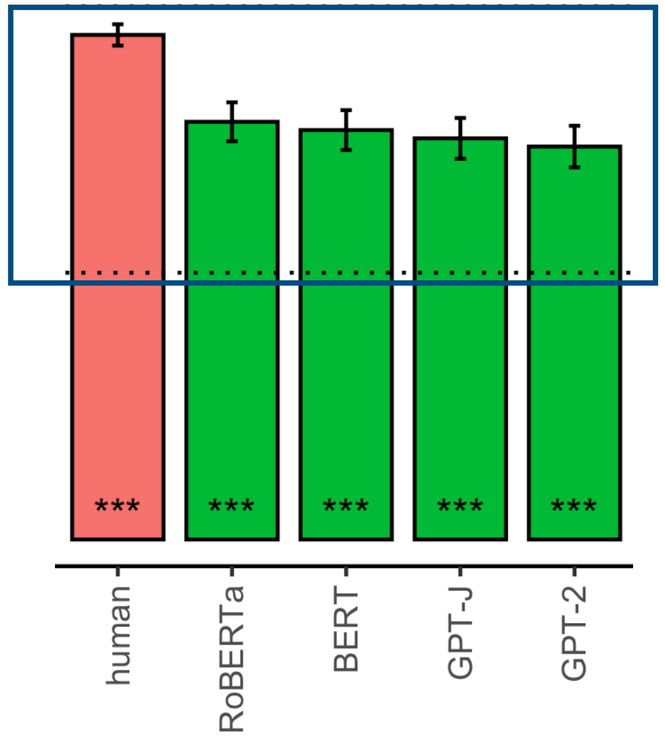
Animate-Animate, unlikely

The fox chased the rabbit.

The rabbit chased the fox.



"the gap between the impossible and the unlikely"



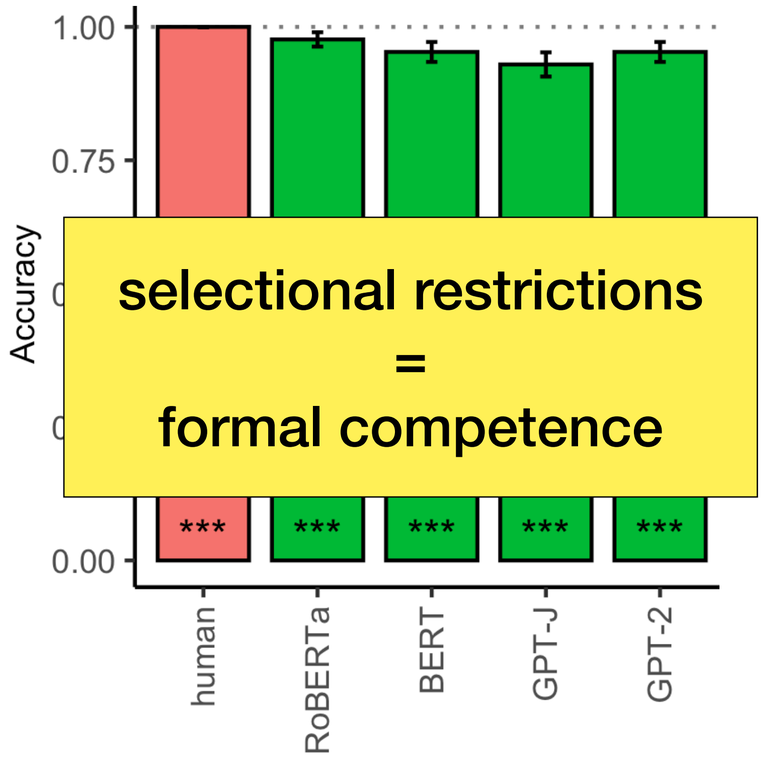
Kauf, Ivanova* et al, arXiv*

Language models and event knowledge

Animate-Inanimate, impossible

The teacher bought the laptop.

The laptop bought the teacher.

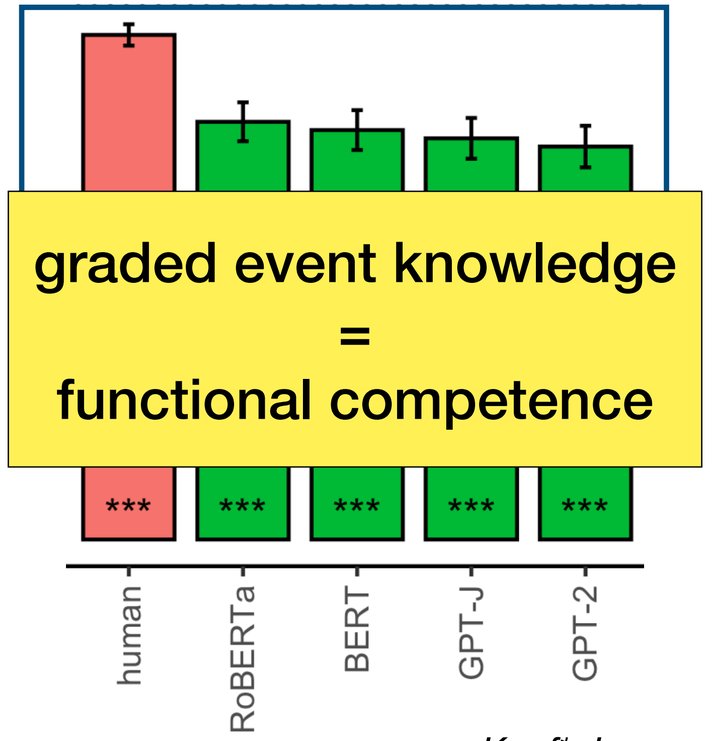


selectional restrictions
=
formal competence

Animate-Animate, unlikely

The fox chased the rabbit.

The rabbit chased the fox.



graded event knowledge
=
functional competence

Kauf, Ivanova* et al, arXiv*

Roadmap

- Introduction
- Formal vs. functional linguistic competence
- Formal competence: grammar
- Functional competence
  - math
  - world knowledge
- Case study: event knowledge
- **Toward better models**

Implications for future models

The formal/functional competence distinction has two implications:

1. **Modular models**
   - architectural modularity
   - emergent modularity

From Word Models to World Models:
Translating from Natural Language to the Probabilistic Language of Thought

Lionel Wong[1⋆], Gabriel Grand[1⋆], Alexander K. Lew[1], Noah D. Goodman[2], Vikash K. Mansinghka[1], Jacob Andreas[1], Joshua B. Tenenbaum[1]

⋆ *Equal contribution.*

Implications for future models

The formal/functional competence distinction has two implications:

1. **Modular models**
   - architectural modularity
   - emergent modularity
2. **Targeted benchmarks**
   - formal competence: BLiMP, SyntaxGym, etc.
   - functional competence: ?

Summary

- Formal competence = knowledge of linguistic rules and patterns
- Functional competence = non-language-specific skills required for real-life language use
- This distinction (grounded in neuroscience) helps to clarify the discourse around LLMs & to develop targeted assessments of their abilities.

Thanks to…

Kyle Mahowald

Carina Kauf

Ev Fedorenko

and all the other co-authors

brain+cognitive sciences

MIT Schwarzman College of Computing

MIT Quest for Intelligence

MCGOVERN INSTITUTE

Thank you for listening!

Email: annaiv@mit.edu

Twitter: @neuranna