Generative Linguistics and Generative AI

Sandiway Fong

University of Arizona

Slides 1–44/183

Some material here presented earlier *at Kyoto University* (2024), *Osaka Kyoiku University* (2023)

Colloquium, Dept. of Linguistics. 3pm Oct 10 2025
Rm: Comm 311
https://arizona.zoom.us/j/83224247918

All That Sparkles Is AI

Generative AI: *three topics*

- The impact
- *vs*. Generative Linguistics
- What's under the hood?

LLM (*Large Language Model*):

- **L** = many parameters (***knowledge***)
- **LM** = **L**anguage **M**odel

LM *formally*: *probability distribution on a word sequence* (Markov 1907; 1913)

pre-training (***GPT***)*: unsupervised,* good at next **token** prediction (*better than us*)

think: **autocompleter**

RLHF: **R**einforcement **L**earning (**RL**) with **H**uman **F**eedback (**HF**): *reward loss f.*

AI's 'Oppenheimer moment'

(N. Robbins-Early, *Guardian,* Jul 14 2024)

- Intelligence
  - **AI-infused audio tool** helped assassinate *Ibrahim Biari* (+ 125 civilians) in Oct 2023 (*NYT*, Apr 25 2025)
  - Special **Arabic-language LLM** trained on intercepts *etc*.
- Dawn of autonomous weapon systems (AWS):
  - Communications can't be jammed. On-board AI to recognize/track targets.
  - **Example:** Russia's Shahed MS001drone uses a ($250?) Nvidia AI module

AI's 'Oppenheimer moment'

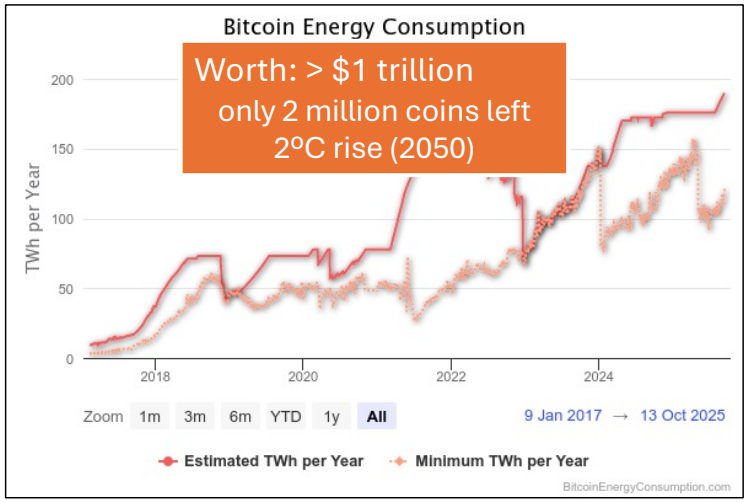
(N. Robbins-Early, *Guardian,* Jul 14 2024)

- Big Tech: Google, Microsoft, *etc*.:
  - when **DeepMind** was sold to **Google**, one of the conditions was that the AI would not be used for military purposes. But last January, that condition was removed.
  - now all [AI] companies provide technology to the **US Department of Defense**. And a few days ago, **Google** signed an important agreement with Israel to provide artificial intelligence systems to the armed forces.

    (R. Luna in a *Corriere Della Sera* interview with Geoffrey Hinton, 9/25/ 2025.)
  - rise of **Sovereign AIs** (*local language*; *national security*; *DeepSeek*)

Environmental Impact

T =10^{12}



- Power consumption: Thailand
- Carbon footprint: Belgium
- Water use: Switzerland
- *disproportionate impact on communities*

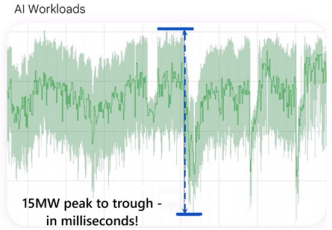
- AI rivals Bitcoin now
  - Query cost: *ChatGPT vs. Google* (10x)
  - **Microsoft AI**: *Three Mile Island* restart

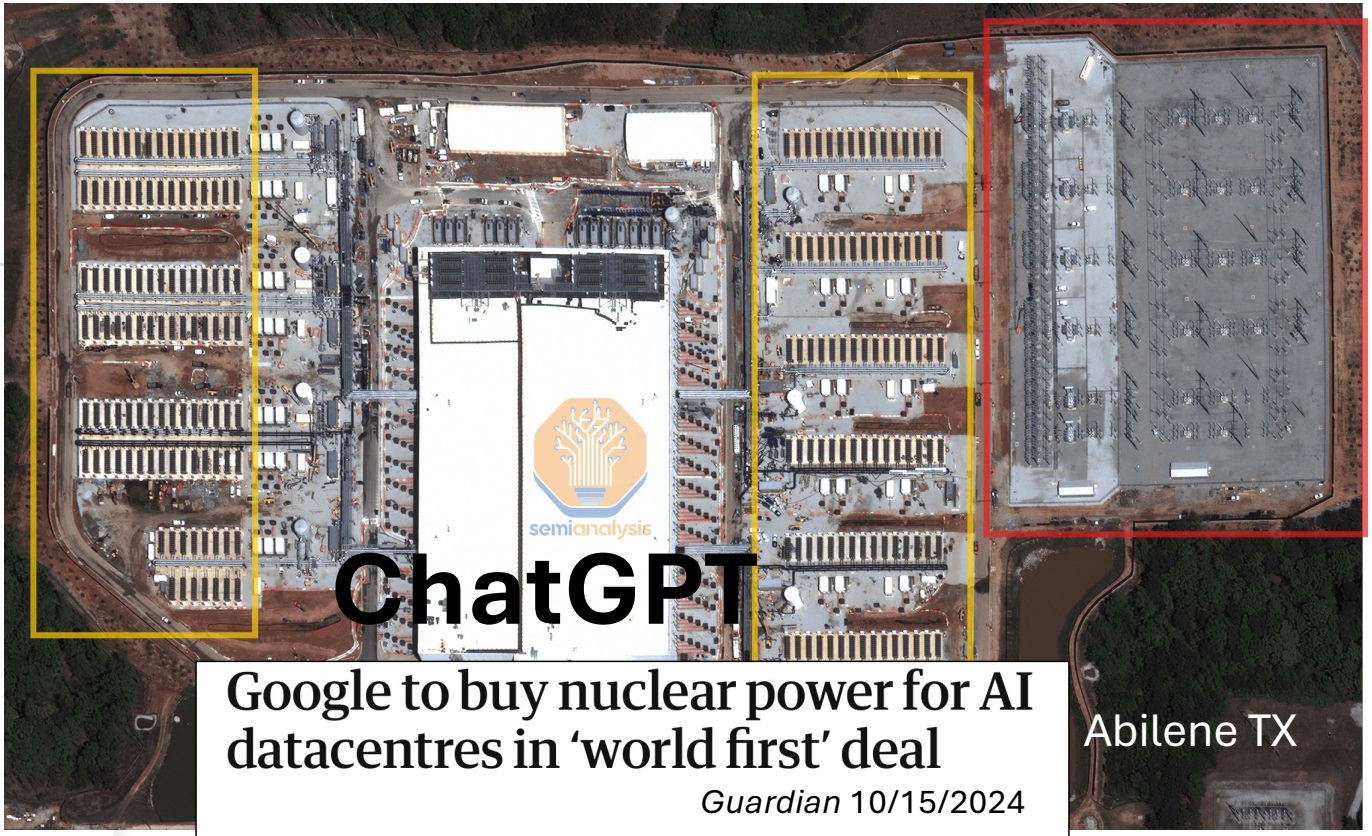
A unit of Pennsylvania's Three Mile Island nuclear plant will be restarted as part of a new energy-sharing agreement with Microsoft, which plans to use it to power the data centers it operates as part of its push into artificial intelligence.

Sources: *including* (Sapra *et al.* 2024), Digiconomist, *aka* A. de Vries-Gao (2025)

OpenAI Stargate datacenter

- ~400MW nameplate (2025), GW (mid-2026)
- 210 chillers
- on-site substation
- projected water use?





Google to buy nuclear power for AI datacentres in 'world first' deal

Guardian 10/15/2024

Tech company orders six or seven small nuclear reactors from California's Kairos Power

Locally

"[Pima] County officials have said they are under [NDAs] that kept them from naming the company."

https://www.tucsonaz.gov/Government/Office-of-the-City-Manager/Project-Blue-Information

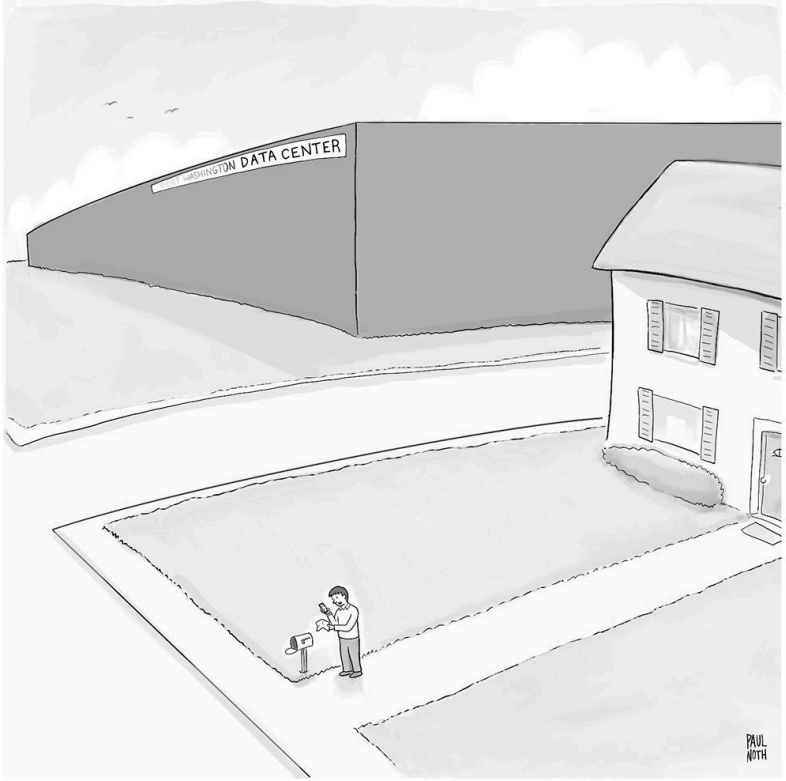
Wholesale cost of electricity

Every day, wholesale electricity prices are measured in real time by Locational Marginal Pricing (LMP) points on the power grid, called nodes. Bloomberg analyzed 25,000 LMP nodes since 2020.

Home > Home Energy & Utilities

The AI Data Center Boom Is Driving Up Electricity Costs, Research Shows

The price of electricity, juiced by demand from power-hungry data centers, is being passed on to residential customers.



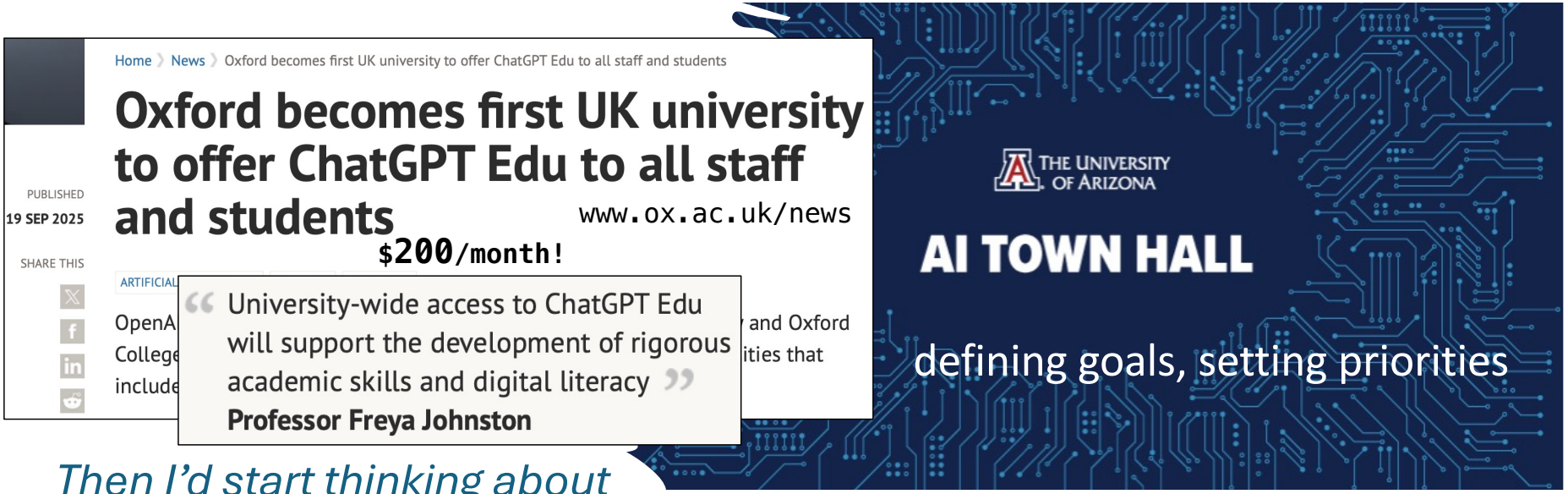
"ChatGPT, why is my electric bill so high?"

NIMBY: *where to put them?*

Big Tech AI push on campus

free ($7) smoothies!

Sep 9 2025



Then I'd start thinking about … what [university] support there is for integration of AI… Is my course, and is the university as a whole, on the front foot with regards to the use of AI?

(Sally Weale, 9/13/2025, Guardian)

NextGenAI OpenAI is committing $50M in research grants, compute fu
ccess to support students, educators and researchers advancing the frontiers
Other founding NextGenAI partners are Caltech, the California State U
e University, the University of Georgia, Harvard University, Howard U
Institute of Technology, the University of Michigan, the University
Ohio State University, the University of Oxford, Sciences Po, Texas
l as Boston's Children's Hospital and the Boston Public Library.

70 Years of Artificial Intelligence (AI) Promises

- **Neural Network AI** hype:
  - *Perceptrons* (Rosenblatt 1958) — AI Winter #1: (Minsky & Papert 1969)
- **Expert Systems** hype:
  - 1985 $1 billion, 2/3rds Fortune 500 — AI Winter #2: 1990s
- **Neural Network AI** hype round 2:
  - 1986 backprop — AI Winter #3?: results plateaued
  - rise of *Statistical NLP*: training datasets
- **Neural Network AI** hype round 3:
  - 2017 *Transformer* invented, 2019 ChatGPT — *peak AI or bubble when?*
  - OpenAI: 700 million weekly users
  - Training: WWW + synthetic (GPT-5). *Superintelligence era has begun* (Altman, 2025)
  - *The end of disease? I think that's within reach. Maybe within the next decade or so, I don't see why not*. (*60 Minutes*. Hassabis, Google/DeepMind, 2025)
  - 2024 Nobel Prizes in physics & chemistry went to AI
  - **AGI** next 5-10 years — *run out of data? Keep scaling up!*

Questions: Hype = Bubble?

- dot-com bubble / 2008 credit crunch / several cryptocurrency bubbles / 2022 NFT bubble

2027 AI bubble?

- *Speculation rules the world. It didn't used to. But from the 1980s through to 2008, something changed. Investors realised that they could get far more return from hype than from any kind of legitimate business.*
- *In fact, nearly half of the world's private investment is being funnelled into AI, and **AI speculation** is the main driving force behind the S&P 500's recent growth.*

(Will Lockett, *Medium* 9/14/2025)

Questions: Hype = Bubble?

- *for cor* *of*
  trainin
- *we find* *any*
  further

022)

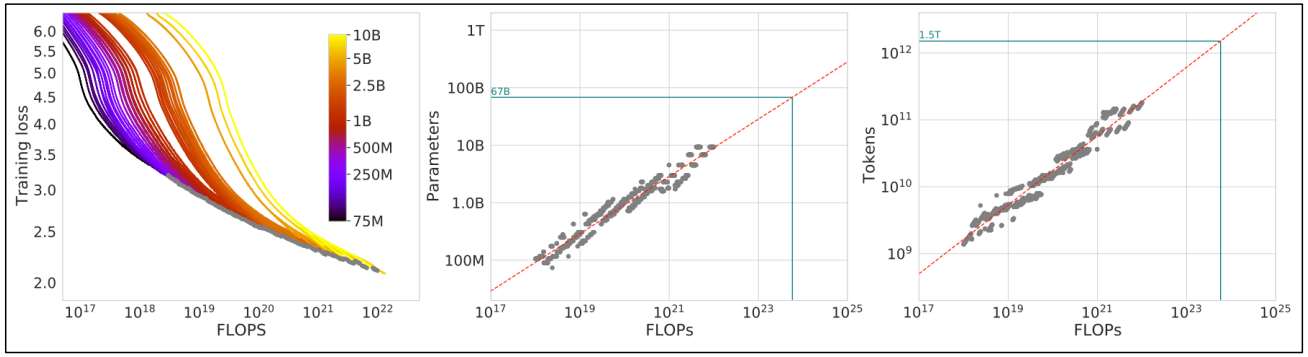
- *No Fre*
- *Floridi*

DeepMind

Training Compute-Optimal Large Language Models

Jordan Hoffmann*, Sebastian Borgeaud*, Arthur Mensch*, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katie Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Jack W. Rae, Oriol Vinyals and Laurent Sifre*

*Equal contributions



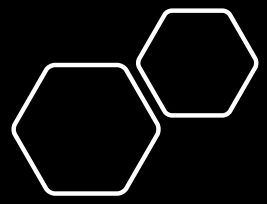
AI Companies Reportedly Struggling to Improve Latest Models

Summarize

Tim Hardwick • Wednesday November 13, 2024 5:30 am PST

Leading artificial intelligence companies including OpenAI, Google, and Anthropic are facing "diminishing returns" from their costly efforts to build newer AI models, according to a new *Bloomberg* report. The stumbling blocks appear to be growing in size as Apple continues a phased rollout of its own AI features through Apple Intelligence.

"The AGI bubble is bursting a little bit," said Margaret Mitchell, chief ethics scientist at AI startup Hugging Face.



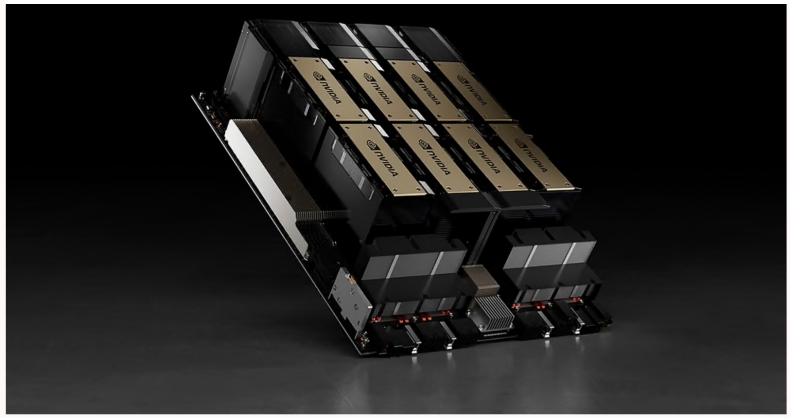
difficulty in finding "new, untapped sources of high-quality, human-made training data"

Expensive training

tae kim @firstadopter · Aug 16, 2023

Replying to @firstadopter

Raymond James estimates it costs Nvidia $3,320 to make a H100, which is then sold to customers for $25,000 to $30,000.



The Nvidia HGX H100 chip, designed for generative AI, is being bought in huge quantities by companies such as Microsoft for $30,000 each. Photograph: AP

- *Not just the electricity*
- *As I write, I see a report that next year Microsoft plans to buy 150,000 Nvidia chips – at $30,000 a pop.* ***It's a kind of madness.***

(J. Naughton, *The Observer*. 12/30/2023)

$4.5 billion

- (2025) Nvidia B200 192GB SXM (**1kW TDP**): $45,000–$50,000

"***OpenAI*** *currently massively subsidizing the cost of AI for the consumer."*

"***[NVIDIA]*** *acts as a financial backstop to the entire AI supply chain"*

“10 gigawatts to power the next era of intelligence.”

NVIDIA to put $100 billion into OpenAI from 2026

- Strategic partnership enables OpenAI to build and deploy at least 10 gigawatts of AI datacenters with NVIDIA systems representing millions of GPUs for OpenAI’s next-generation AI infrastructure.
- To support the partnership, NVIDIA intends to invest up to $100 billion in OpenAI progressively as each gigawatt is deployed.
- The first gigawatt of NVIDIA systems will be deployed in the second half of 2026 on NVIDIA’s Vera Rubin platform.

San Francisco and Santa Clara—September 22, 2025—NVIDIA and OpenAI today announced a letter of intent for a landmark strategic partnership to deploy at least 10 gigawatts of NVIDIA systems for OpenAI’s next-generation AI infrastructure to train and run its next generation of models on the path to deploying superintelligence. To support this deployment including datacenter and power capacity, NVIDIA intends to invest up to $100 billion in OpenAI as the new NVIDIA systems are deployed. The first phase is targeted to come online in the second half of 2026 using NVIDIA’s Vera Rubin platform.

Ensure Quality: the shadow (human) workforce

- **Gig workers** (*up to Ph.D level*):
  - censor output (**guardrails**)
  - correcting mistakes
  - better answers
  - *… her task was to enter details on **chemotherapy options for bladder cancer,** which haunted her because she wasn't an expert on the subject.*
  - "I just want people to know that AI is being sold as this tech magic – that's why there's a **little sparkle symbol** next to an AI response," said Sawyer. "But it's not. It's built on the backs of overworked, underpaid human beings."

    (Varsha Bansal, Guardian, 9/11/2025)

- **Outsourcing Trauma to the Developing World**
  - ***OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic*** (2/14/2022, *Time*)
  - [Open AI] said labeling harmful images was "a necessary step" in making its AI tools safer.
  - *ChatGPT and other generative models are not magic – they rely on massive supply chains of human labor and scraped data, much of which is unattributed and used without consent*
  - ***content moderators** for **Facebook**, whose jobs involved viewing images and videos of executions, rape and child abuse for as little as $1.50 per hour.*

Before we look inside the LLM

Factual Fidelity

The problem is, they have absolutely no commitment to the truth.

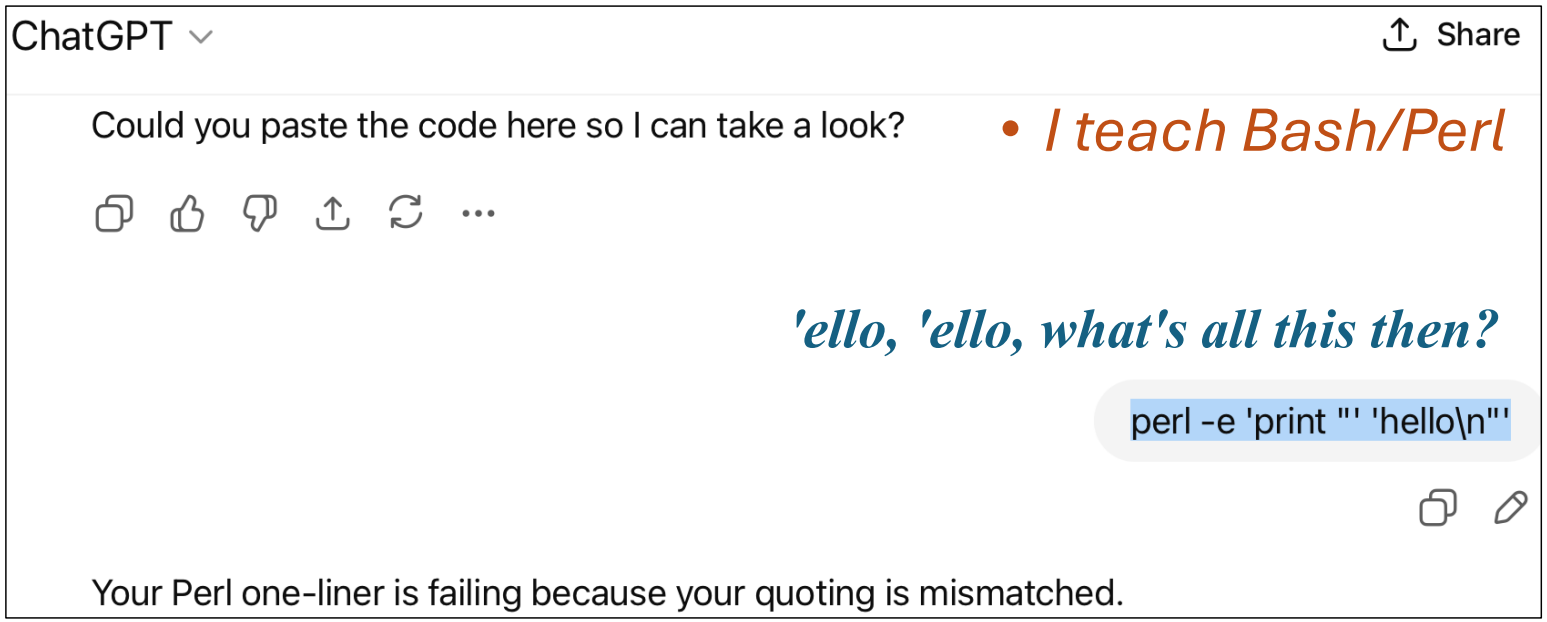
(Emily Bell, *Guardian*, Mar 3 2023)

AI language models are notorious bullshitters, often presenting falsehoods as facts. They are excellent at predicting the next word in a sentence, but they have no knowledge of what the sentence actually means. That makes it incredibly dangerous to combine them with search, where it's crucial to get the facts straight.

(*MIT Technology Review*, Feb 14 2023)

How about understanding?

- Explainable AI: *chain-of-thought* ...



A simple computer language

2. Use single quotes for the Perl code, escape the single quote inside:

bash Copy code

```
perl -e 'print "\'hello";'
```

3. Concatenate to avoid escaping:

bash

```
perl -e 'print "'"'"'"hello";'
```

Bash shell (macOS, Linux):

- manual: http://www.gnu.org/software/bash/manual/

3.1.2.2 Single Quotes

Enclosing characters in single quotes (‘'’) preserves the literal value of each character within the quotes. A single quote may not occur between single quotes, even when preceded by a backslash.

want this (@a is an array):

```
@a=('a', 'b', 'c')
```

but we can't write:

1. ' … ' *fine if no ' inside*
2. ' … ' … ' … ' *doesn't work*
3. ' … \' … \' …' *cannot work*

- *recall ChatGPT doesn't understand this*

(This is a shell trick: `'` [end single quotes], `"'"` inserts a literal single quote, then `'` starts single quotes again.)

Generative AI vs. *Generative Linguistics*

formal system for generating language expressions

Biological basis for the study of language; brain is a biochemical [computer]: Lenneberg (1967; 1969)

Piantadosi vs. Chomsky

- **Modern language models refute Chomsky's approach to language** (Piantadosi, *lingbuzz* 2023; 2024)
- *The rise and success of [**LLMs**] undermines virtually every strong claim for the innateness of language that has been proposed by **generative linguistics**.*
- *Modern machine learning has subverted and bypassed the entire theoretical framework of Chomsky's approach, including its core claims to particular insights, principles, structures, and processes.*

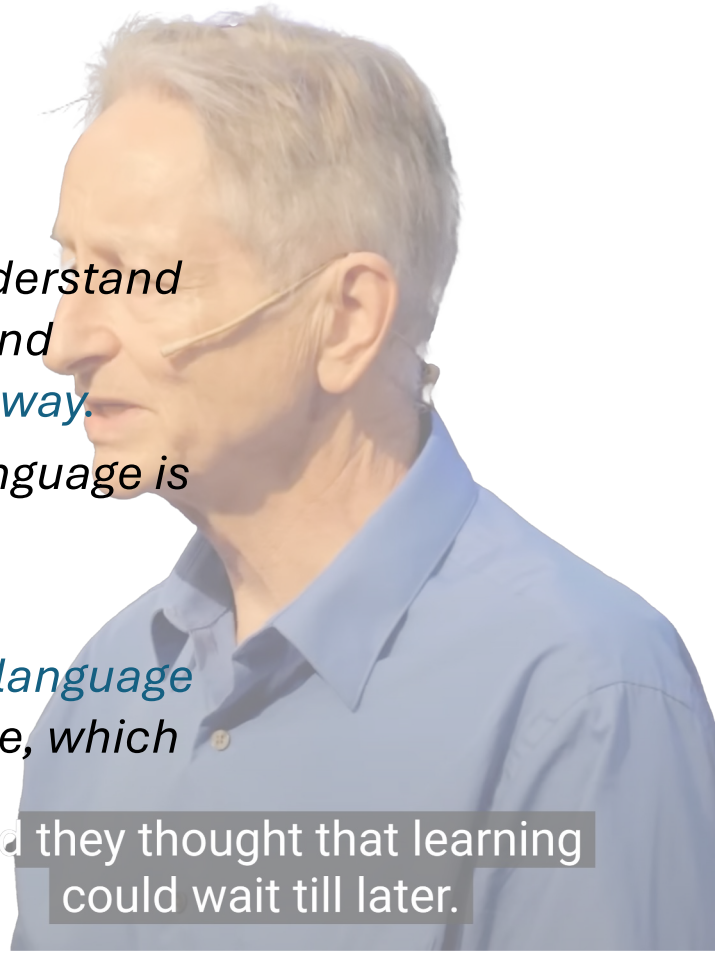
sandiway.arizona.edu: Fong (2025) IJL Volume 37, Issue 1, pp. 59-74.

Futrell and Mahowald (2025)

- How Linguistics Learned to Stop Worrying and Love the Language Models, *Behavioral & Brain Sciences (BBS)* (2025)
  - *Formal linguistics has not presented an alternative model with the demonstrated practical language-learning capabilities of neural models.*
  - *The direction of developments in machine learning suggests that the gap between human and machine learning is more likely to be closed not through built-in domain-specific formal restrictions, but rather through more powerful domain-general learning algorithms.*
  - *Finding these internal mechanisms is not a prerequisite for the reality of linguistic structure — the mechanisms may be irreducibly complex.*
  - *It is rare in the history of science for a scientific theory to turn out as disconnected from a corresponding engineering application as formal generative linguistics has turned out to be for language models.*

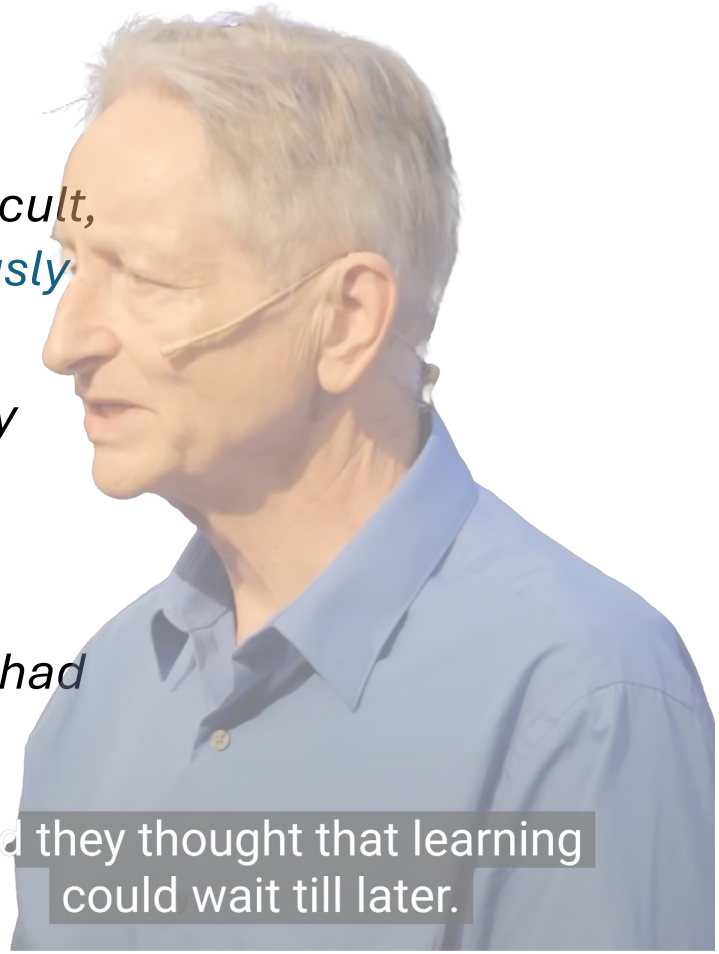
Hinton vs. Chomsky

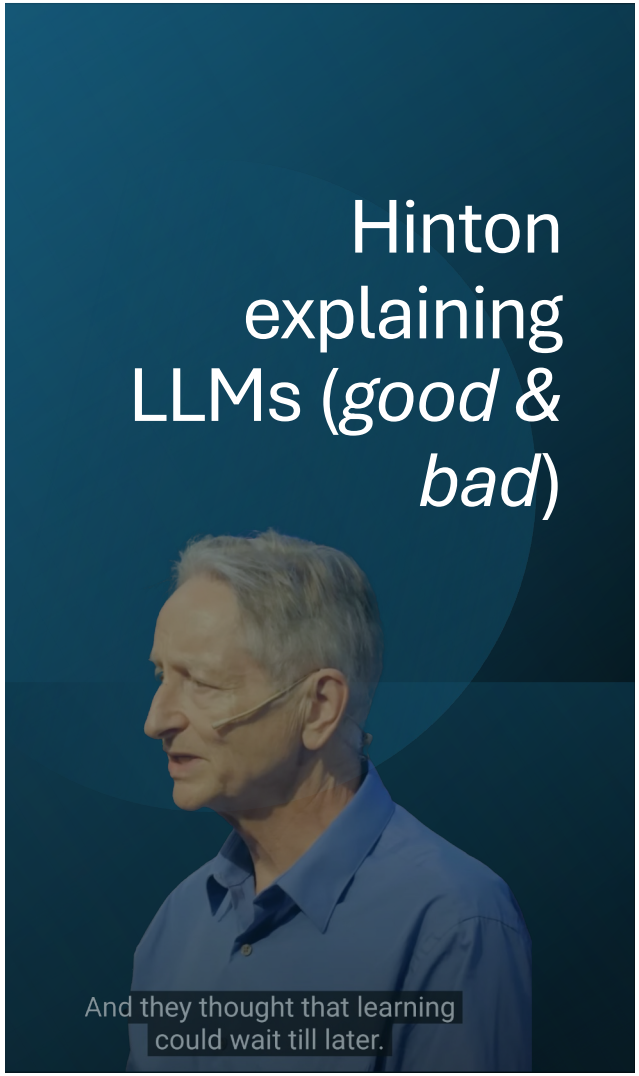
- Hinton *Royal Institution* lecture (2025):
  - *That's what understanding is when you understand language and when these [LLMs] understand language. We understand in just the same way.*
  - *And that's a much better model of what language is than anything the linguists ever had.*
  - *The linguists of course hate it.*
  - *[Linguists] also thought that knowledge of language was innate, knowledge of syntax was innate, which is just stupid.*



Hinton vs. Chomsky

- *It's the sign of a cult, that in order to join the cult, you have to believe something that is obviously silly, like language isn't learned.*
- *There's a guy called Chomsky, when they say things like, "These things don't understand anything, they're just a statistical trick," they don't actually have a model of what understanding is, because they never really had a model of what understanding was.*





- How it works
  - ***Attention:*** *You can't just associate a feature vector with [a word] that captures the meaning directly. You have to hedge your bets. And then as you go up through layers of the network, you disambiguate it using influences from nearby things.*
  - ***Backprop:*** *You're trying to predict the next word.* [...] *it's all done with features and then interactions, when you get it wrong, you back-propagate information that learns all these features and then interactions.*
  - ***That's how language works for us, and it's how it works for these large language models. We are basically the same.***

Learning and Training

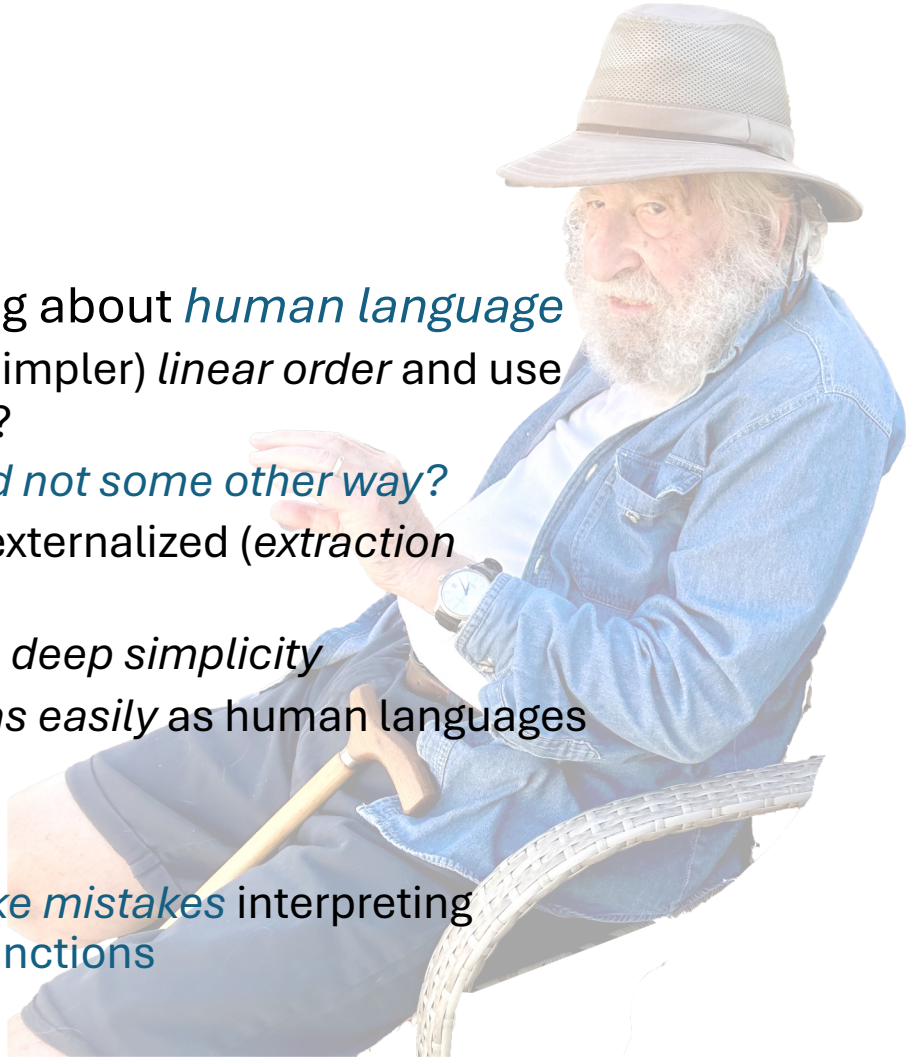
- Human brain: **20W** (*oft-quoted figure*)
  - training: *auto, rapid acquisition* (despite *Poverty of Stimulus*)
  - *data: 4,000 ~ 35,000 words/day @ 12 months* (Brusche *et al*. 2020)
  - *vocabulary: age 1: 50, 3: 1,000, 5: 10,000 words* (Shipley & McAfee 2015)
  - *structure dependence*: 18 months (Chomsky cites Shi *et al*. 2020)
- Meta's Llama-3.1-405B LLM
  - *405B parameters*
  - *training: 31 million hours on H100-80GB* ~3,540 years
  - *TDP:* ***350W***

Chomsky

Even if LLMs work, they tell us nothing about *human language*

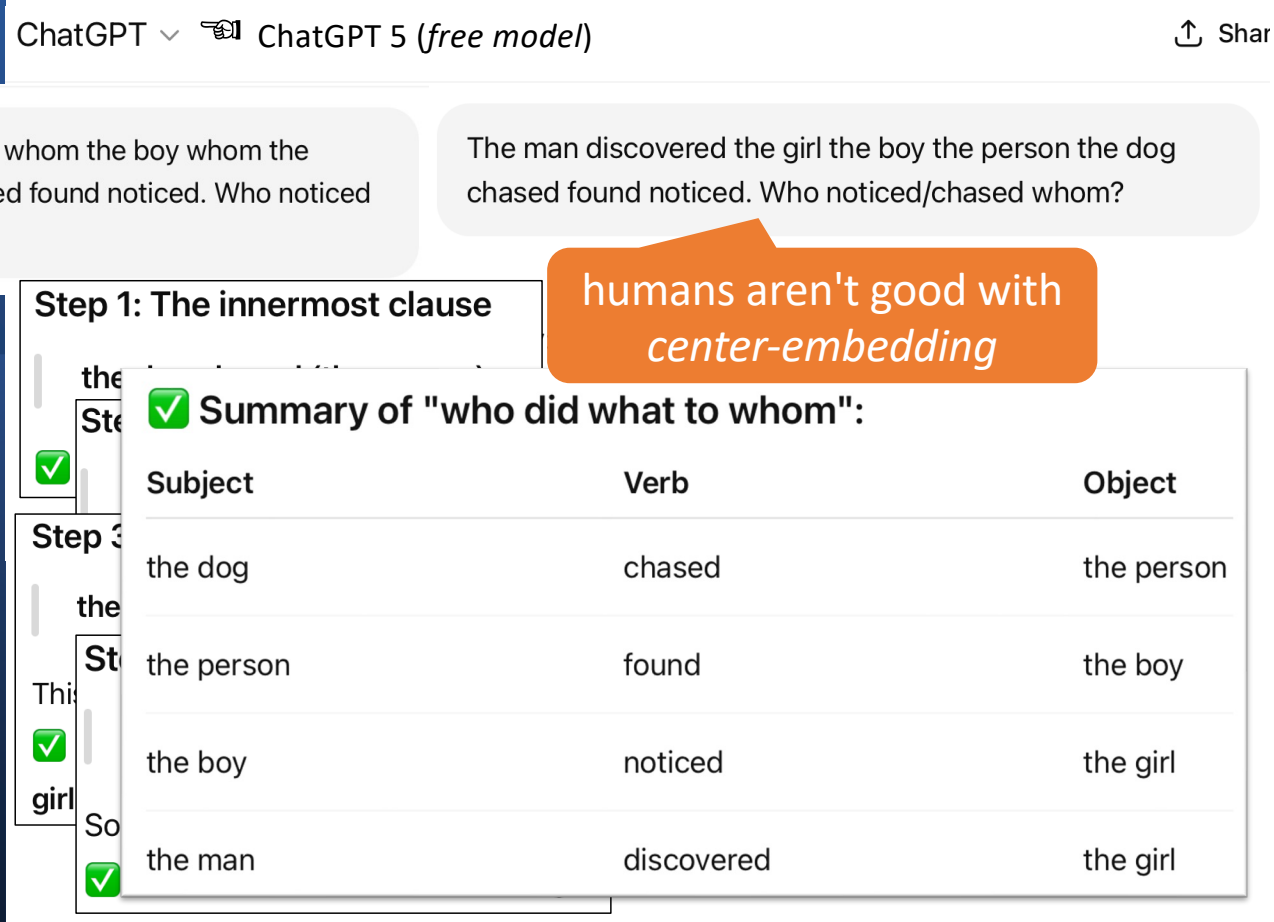
- Why does human language ignore (simpler) *linear order* and use *structural relations* (*Basic Property*)?
- *Why is human language this way and not some other way?*
- Fine thoughts that can't be directly externalized (*extraction asymmetries*)
- Believes in Einstein's *Miracle Creed*: *deep simplicity*
- LLMs learn artificial languages *just as easily* as human languages (*but humans can't, e.g. Moro's work*)

- More generally:
  - Statistical/artificial NN systems *make mistakes* interpreting sentences; lack grammaticality distinctions



Do LLMs work like humans?

adapted from my 438/548 class slides (HW9)
10/9/2025



Homework 9 Extras

2️⃣ Labeled Dependencies (relations)

```scss
discovered(man, girl)
noticed(boy, girl)
found(person, boy)
chased(dog, person)
disliked(cat, dog)
```

You can think of this as a cascade of predicates:

ChatGPT

Sh

csharp

Copy code

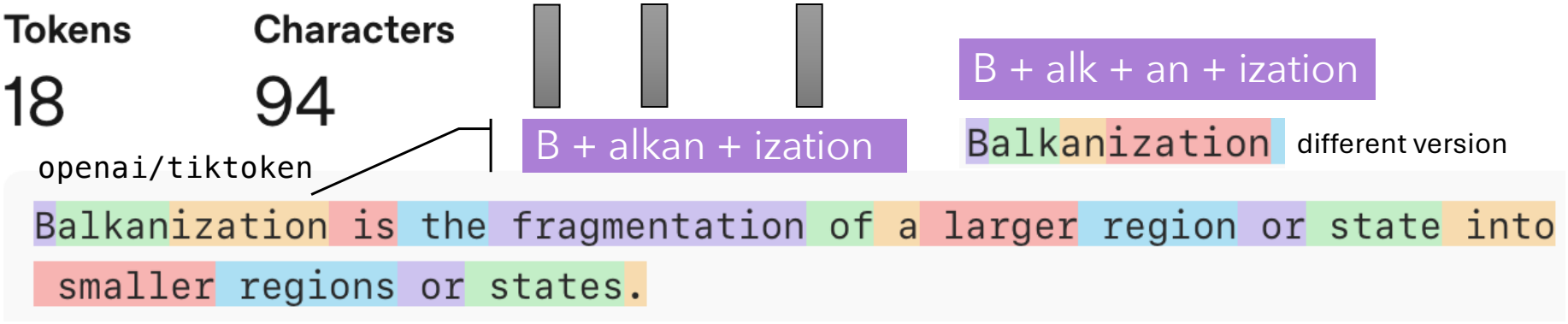
```
[S
  [NP The man]
  [VP discovered
    [NP the girl
      [RelClause
        (the boy
          [RelClause
            (the person
              [RelClause
                (the dog
                  [RelClause
                    (the cat disliked)
                  ] chased
                ) found
              ] noticed
            )
        )
      ]
    ]
  ]
]
```

The man discovered the girl the boy the person the dog the cat disliked chased found noticed. Who noticed/chased/dislked whom?

A quick peek under the hood

Embeddings: vectors for tokens

- *How does an LLM know the meaning of unknown words?*
- **Answer**: it can't
- **Tokenization**: sub-word (~4 characters) (OpenAI's **BPE**)



- **Note**: not a *bag-of-words* scheme: a unique position vector is added to each token vector; e.g. GPT-2: 1024 token context size

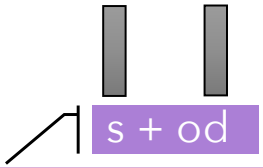
(*from my 438/538 class slides 9/9/2025*)

Language: *infinite employment of finite means*

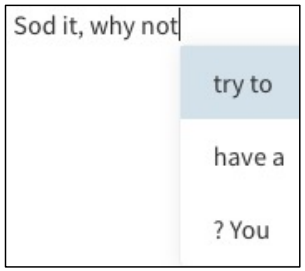
(von Humboldt, cited by Chomsky)

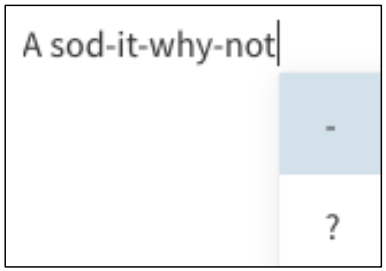
- *Galileo expressed his amazement at what is, in fact, an astonishing fact; with a finite number of symbols one can construct in the mind an infinite number of linguistically formulated thoughts and can even go on to reveal to others who have no access to our minds their innermost workings.*

2023 Keio Lecture 2 (00:36) (Chomsky)



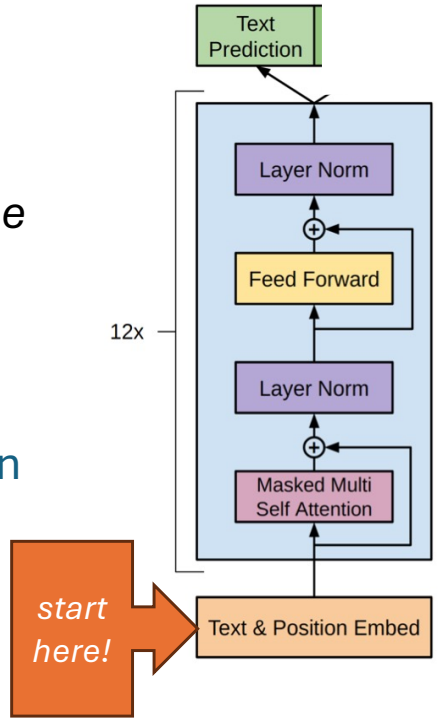
- **Sod it, why not**? (*take a chance*)
- *He chested the ball down, swivelled and cracked a* ***sod-it-why-not*** *shot that took a slight deflection off Evans and beat the diving Onana at the near post.* (Guardian 9/3/2023)





Generative Pre-Training Transformer (GPT)

- Transformer (encoder/decoder)
  - *Attention is all you need*. Vaswani et al. (2017)
  - training is fully parallel: *enormous speed-up*
  - generation is still sequential: *one token at a time*
- **GPT**:
  - *Improving Language Understanding by Generative Pre-Training*. Radford et al. (2018)
  - unsupervised pre-training: next token prediction
  - multi-layer Transformer decoder:
    - multi-headed masked self-attention + feedforward layer



Subject Verb Agreement and Attention

Example (Chomsky 2021):

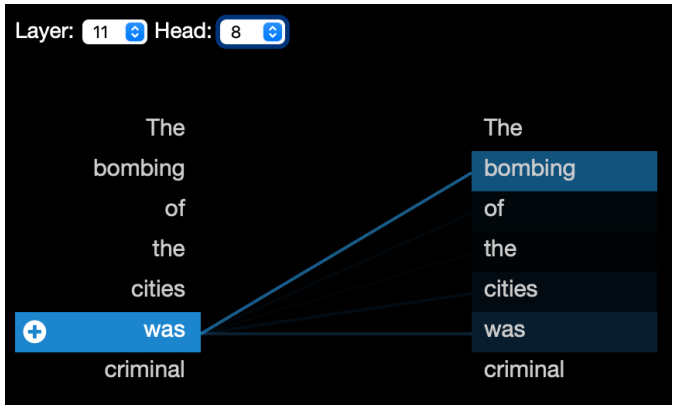
- The *bombing* of the *cities was* criminal

Adapted Examples:

- The *indiscriminate bombing*$_3$ of the *cities was*$_7$ criminal
- The *indiscriminate bombing*$_3$ of the *ancient and modern cities*$_9$ *was* criminal

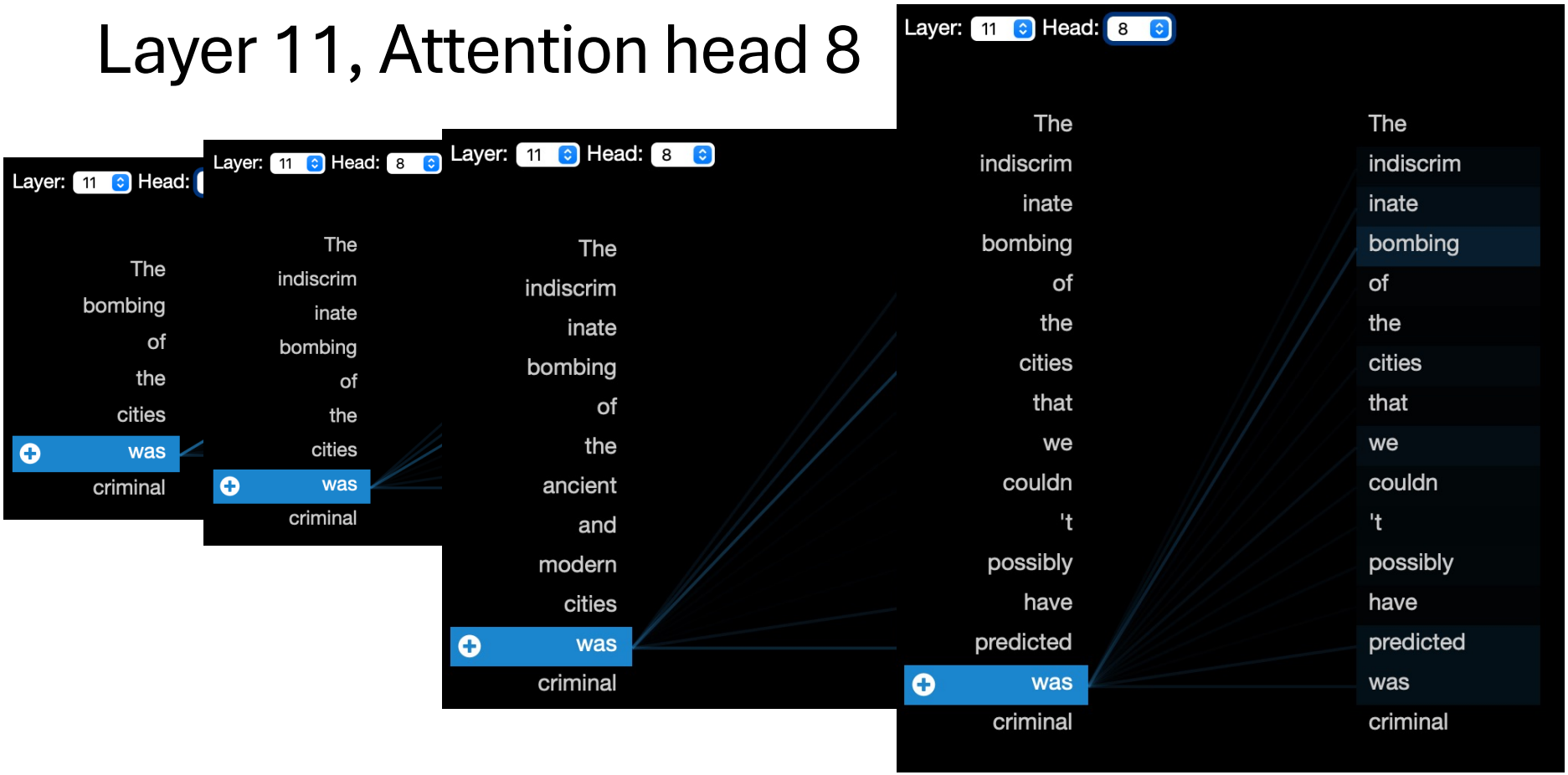
* The *indiscriminate bombing*$_3$ of the *ancient and modern cities*$_9$ *were* criminal

- GPT-2 small (12 levels; 12 heads)



- The *indiscriminate bombing*$_3$ of the *cities that we couldn't possibly have predicted was*$_{13}$ criminal

Layer 11, Attention head 8



Adverb-Verb Construal and Attention

- The mechanic who *carefully* ***fixed*** the car *packed* his tools
- The mechanic who ***fixed*** the car *carefully* ***packed*** his tools
- *Carefully*, the mechanic who *fixed* the car ***packed*** his tools

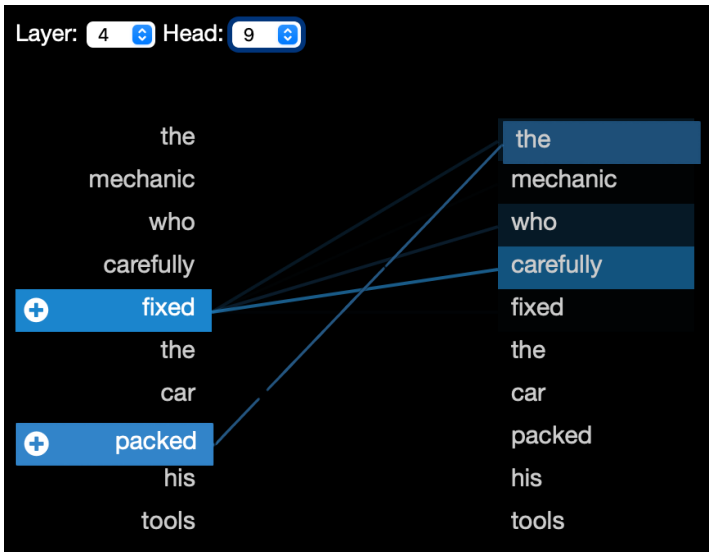
adapted from (Chomsky 2021)

Vig & Belinkov (2019)

5.3.2 Dependency Relations

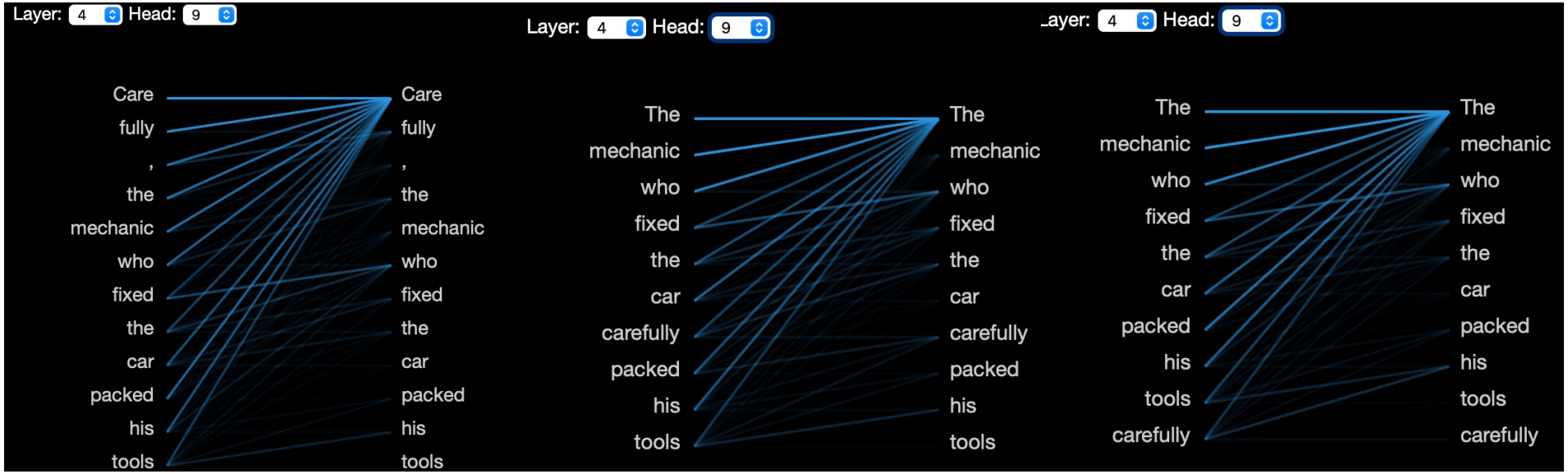
Figure 8 shows the dependency alignment scores (Eq. 4) broken out by layer. Attention aligns with dependency relations most strongly in the middle layers, consistent with recent syntactic probing analyses (Liu et al., 2019; Tenney et al., 2019).

Figure 8: Proportion of attention that is aligned with dependency relations, aggregated by layer. The orange line shows the baseline proportion of token pairs that share a dependency relationship, independent of attention.



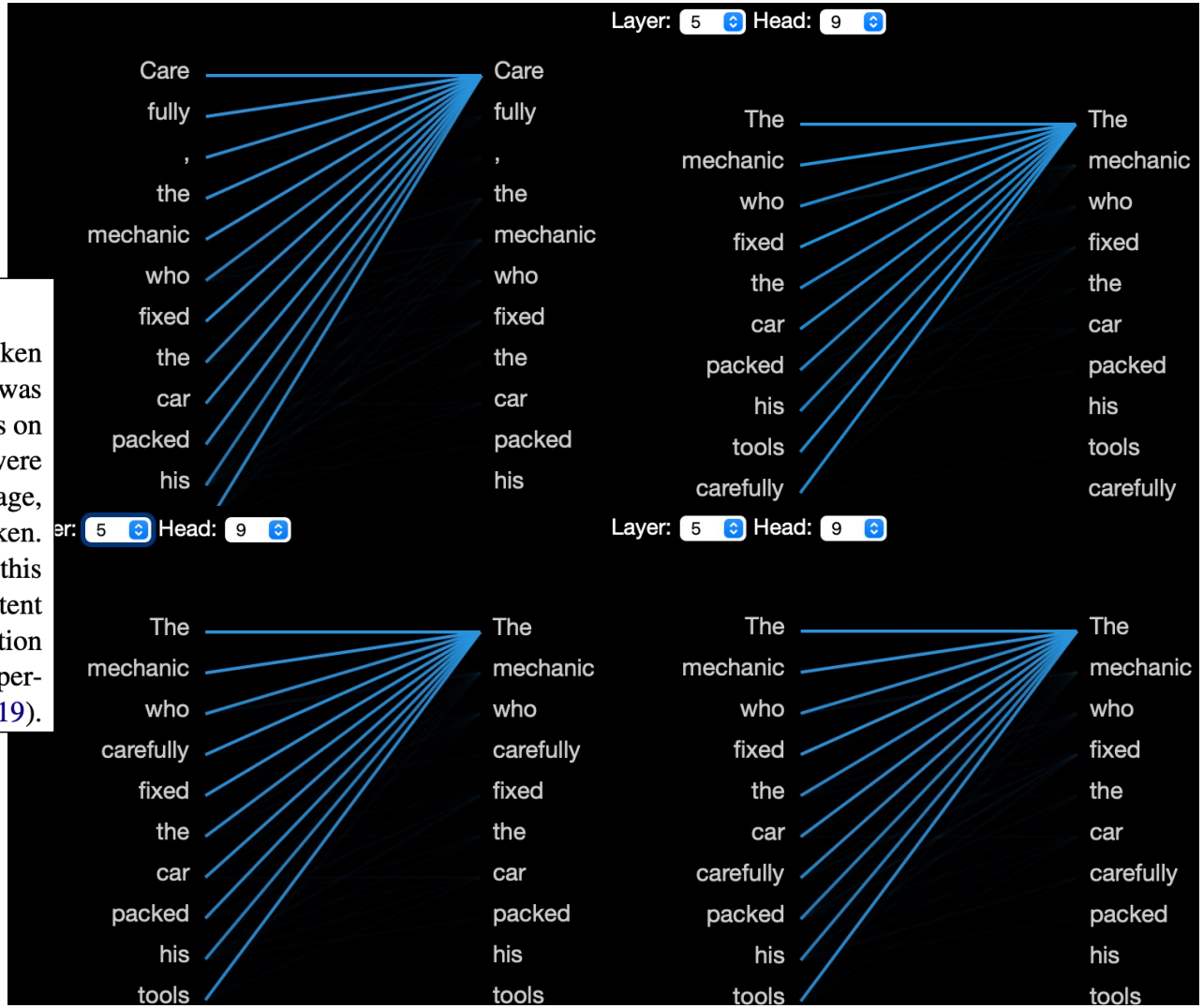
Adverb-verb construal

- Head 9: no sign of dependency relation



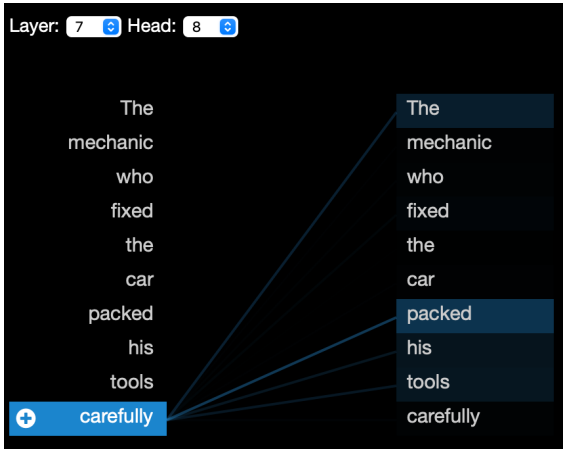
5.2.3 Filtering Null Attention

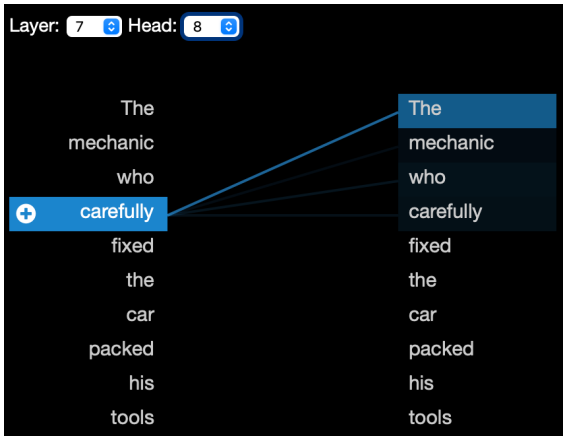
We excluded attention focused on the first token of each sentence from the analysis because it was not informative; other tokens appeared to focus on this token by default when no relevant tokens were found elsewhere in the sequence. On average, 57% of attention was directed to the first token. Some heads focused over 97% of attention to this token on average (Figure 5), which is consistent with recent work showing that individual attention heads may have little impact on overall model performance (Voita et al., 2019; Michel et al., 2019).

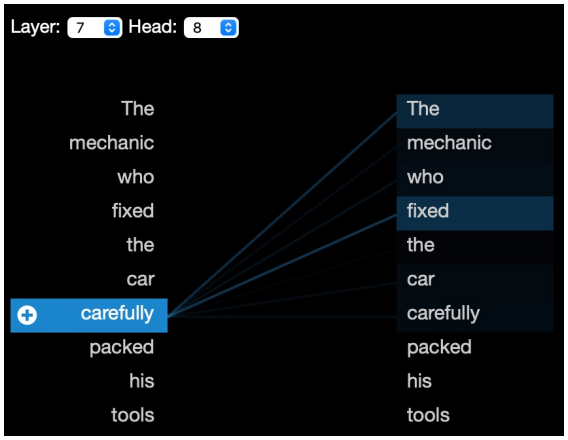


Adverb-verb construal

- Layer 7 Head 8







- Joke works because of two particular senses of *badly*

1. *so badly*: desperately/intensely/acutely
2. *badly*: poorly/incompetently/ineffectively

***Hinton on Attention** (Royal Institution Lecture):* *You can't just associate a feature vector with [a word] that captures the meaning directly. You have to hedge your bets. And then as you go up through layers of the network, you disambiguate it using influences from nearby things.*

†Sentence #	Single clause	+ Context
1. *desperately/poorly*	0.722/0.596	0.778/0.566
2. *desperately/poorly*	0.565/0.920	0.589/0.909

†*(prelim.) see next slide*

- *special thanks to Braa Oudeh (previous slide), Masayuki Oishi*