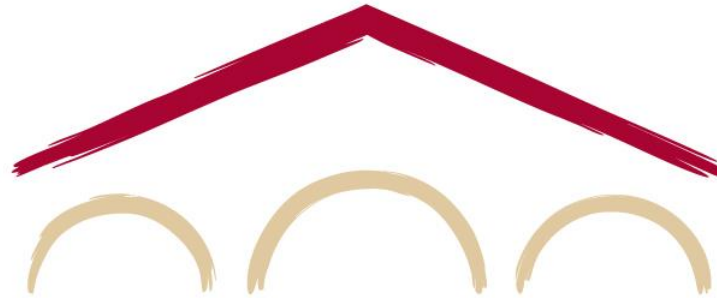# Natural Language Processing with Deep Learning
CS224N/Ling284

Isabel Papadimitriou

Lecture 14: Insights between NLP and Linguistics
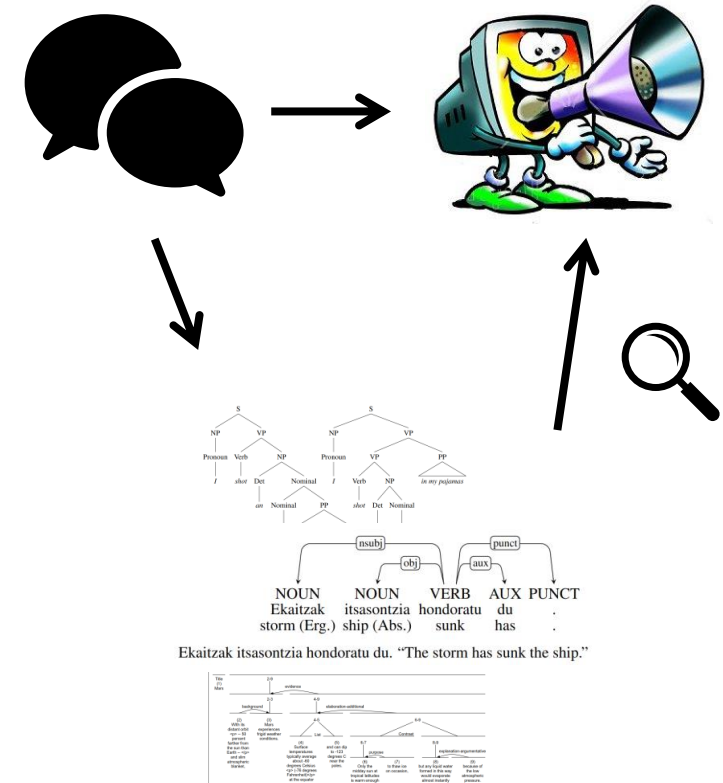
Reminders

- We're in the project part of the class!

- You should have a mentor/grader assigned

  - You can go to any OH, but your mentor will probably know most about your project

- Project milestone is due March 2nd. That's **next Thursday!**

Large language models: a paradigm shift for the role of linguistics

Before:
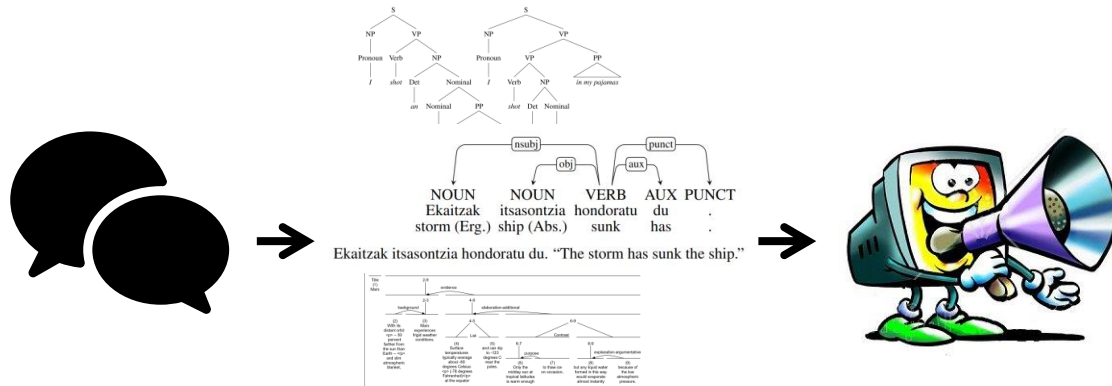
Now:

In this lecture, we'll:

- Learn some more linguistics – language is an amazing thing!

- And discuss some questions (hopefully reach some conclusions!):

  - Where does linguistics fit in, for today's NLP?

  - What does NLP have to gain from knowing about and analyzing human language?

Lecture Plan

1. Structure in human language

2. Linguistic structure in NLP

3. Going beyond pure structure (in linguistics and deep learning)

4. Multilinguality in NLP

Lecture Plan

1. **Structure in human language**

2. Linguistic structure in NLP

3. Going beyond pure structure (in linguistics and deep learning)

4. Multilinguality in NLP

Human language

- All humans have language, and no other animal communication is similar

- Language can be manipulated to say infinite things

How Deaf Children in Nicaragua Created a New Language

It happened on the playground.

BY SHOSHI PARKS • JULY 13, 2018

- But the brain is finite – some sort of set of rules

- Can manipulate these rules to say **anything**

- We can talk about things that don't exist, that can't exist, things that are totally abstract, and we can express subtle differences between similar things

[Hockett, 1960; Chomsky, 1957]

(Recap) There's structure underlying language

Isabel broke the window
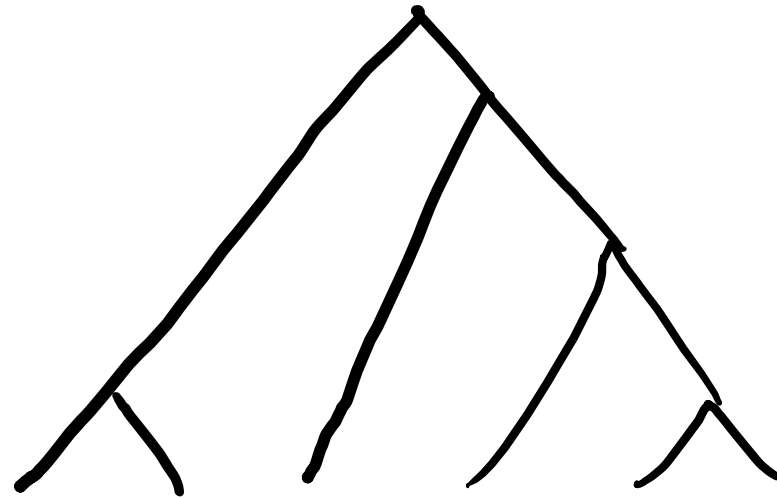The window was broken by Isabel

The cat is batting the toy
The toy is being batted by the cat

The plid yorbed the plof
The plof was yorbed by the plid

- We have some knowledge of structure that's separate from the words we use and the things we say

The cat sat on the mat

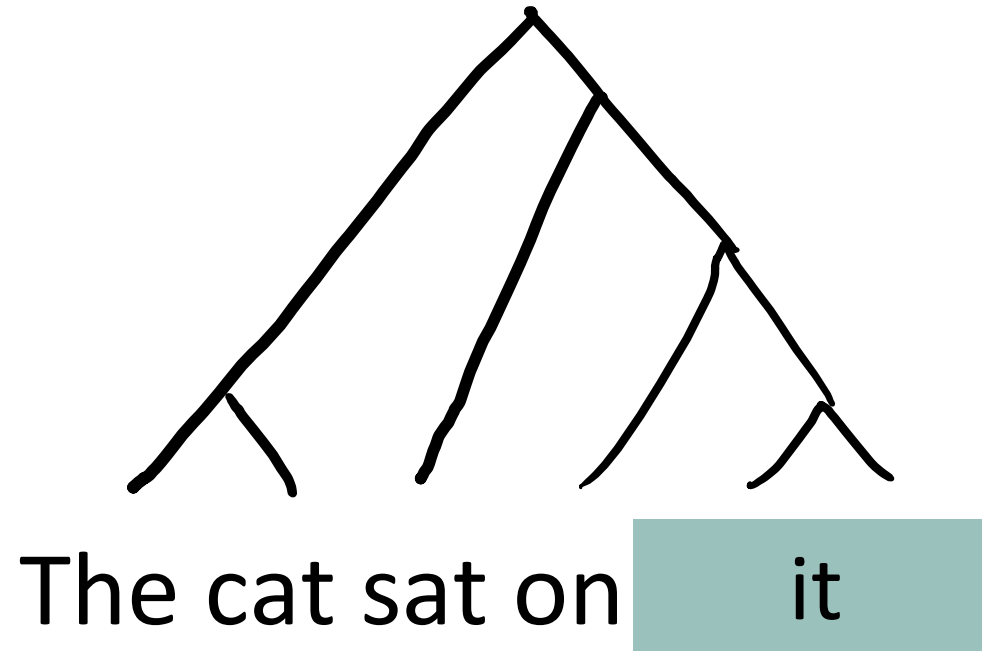# Structure dictates how we can use language
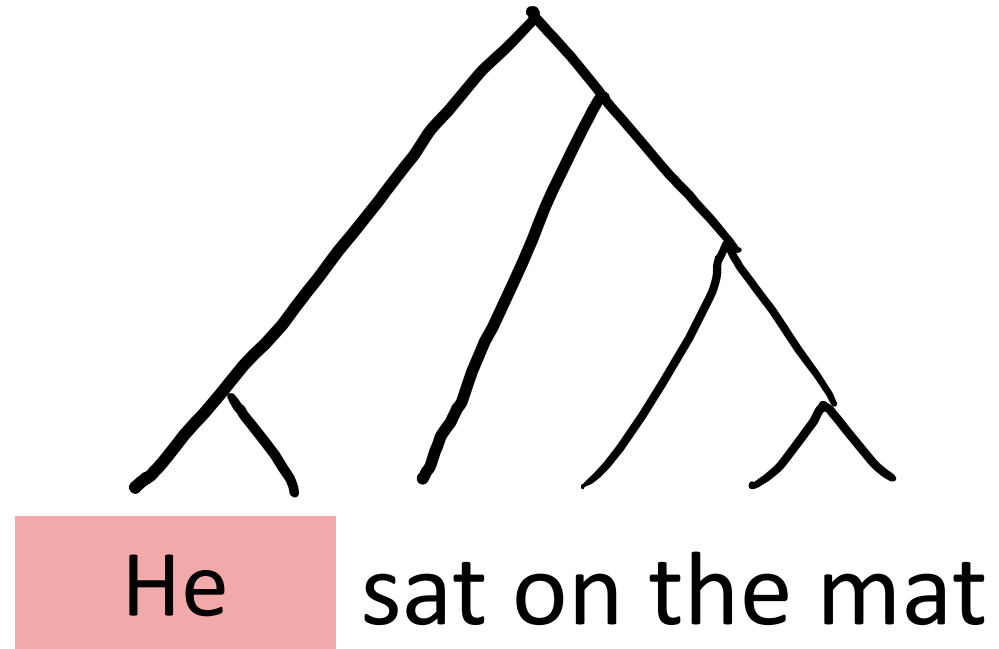
He sat on the mat

Structure dictates how we can use language

The cat sat on it

Structure dictates how we can use language

The cat sat there

The cat did so
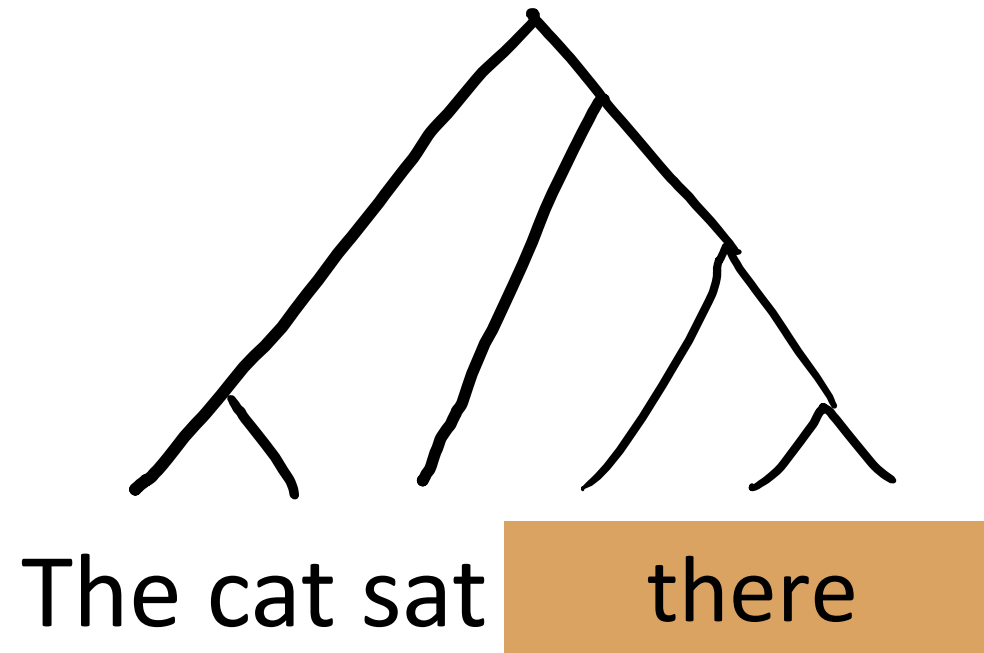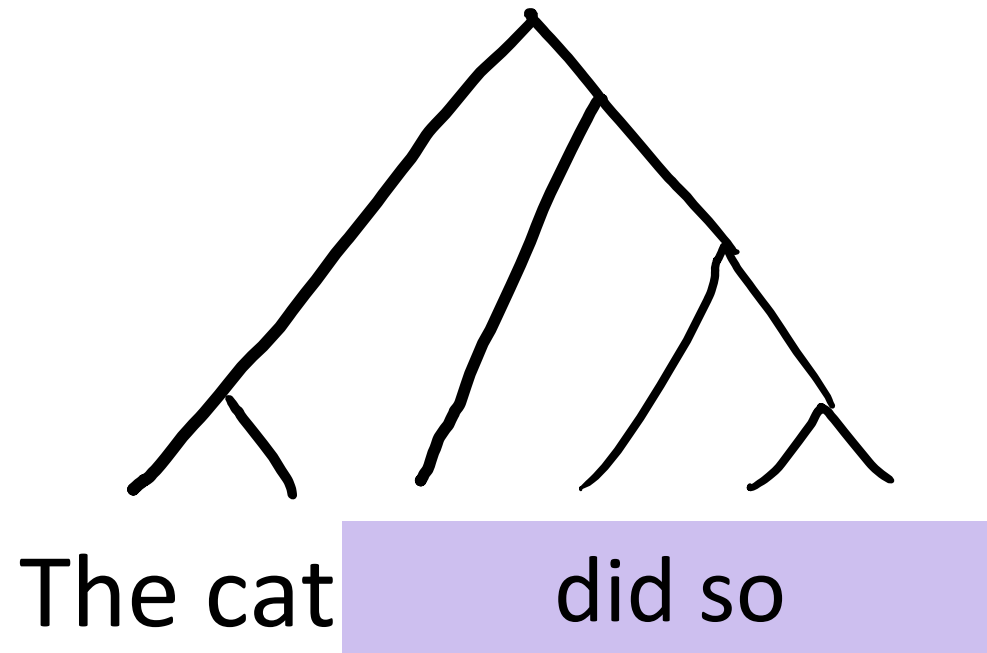
Structure dictates how we can use language
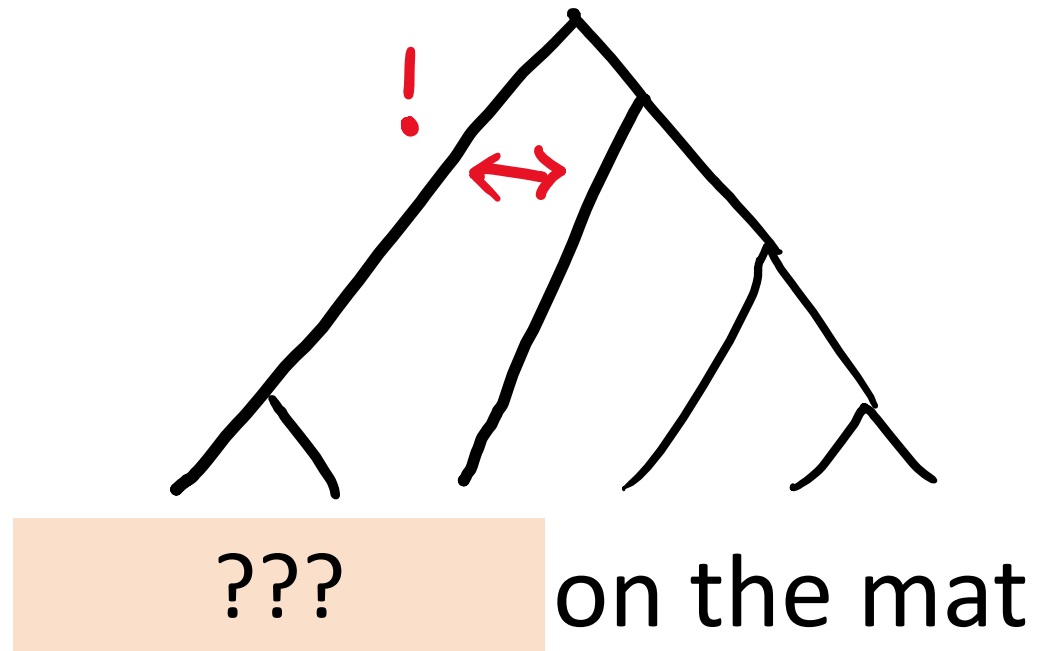
We implicitly know complex rules about structure

- What can we pull out to make a question?

Leon is <u>a doctor</u>   <u>What</u> is Leon?

My cat likes <u>tuna</u>   <u>What</u> does my cat like?

Leon is a doctor and <u>an activist</u>   ✖ <u>What</u> is Leon a doctor and?

We implicitly know complex rules about structure

- When can we move the object to the end?

I dictated <u>the letter</u> to my secretary

I dictated <u>the letter that I had been procrastinating writing for weeks and weeks</u> to my secretary

I dictated to my secretary <u>the letter that I had been procrastinating writing for weeks and weeks</u>

✘ I dictated to my secretary <u>the letter</u>

Not the rules you learned in school!

- A community of speakers (eg, Standard American English speakers) share a rough consensus of their implicit rules.

- **A grammar**: an attempt to describe all these rules

- What we are taught as "rules of grammar" often have other purposes than describing the English language

- When they say...

  - Never start a sentence with 'And'

  - It's incorrect to say "I don't want nothing"

  - "Focus your thoughts and sound formal for this high school essay"

  - "The dialect with the most power in the US does not do negation in this way"
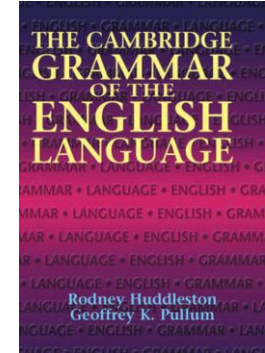
Grammaticality

- A community of speakers (eg, Standard American English speakers) share a rough consensus of their implicit rules.
- All the utterances we can generate from these rules are **grammatical**.
  - If we cannot produce an utterance using these rules, it's **ungrammatical**

Example

- Subject, Verb, and Object appear in **SVO order**
- **Subject pronouns** (I/she/he/they) have to be subjects, object pronouns (me/her/him/them) have to be objects

  - ✔ "I love her"   ✘ "Me love she"

18

Grammaticality

- A community of speakers (eg, Standard American English speakers) share a rough consensus of their implicit rules.

- All the utterances we can generate from these rules are **grammatical**.

  - If we cannot produce an utterance using these rules, it's **ungrammatical**

Example

- Subject, Verb, and Object appear in **SVO order**

- **Subject pronouns** (I/she/he/they) have to be subjects, object pronouns (me/her/him/them) have to be objects

- ✖ "Me a cupcake ate"

- The meaning is clear

- But our rules of grammaticality **don't seem to cut us much slack**

Grammaticality

- A community of speakers (eg, Standard American English speakers) share a rough consensus of their implicit rules.

- All the utterances we can generate from these rules are **grammatical**.

  - If we cannot produce an utterance using these rules, it's **ungrammatical**

Example

- Sentences can be grammatical without any meaning

  - ✔ "Colorless green ideas sleep furiously"
  - ✘ "Colorless green ideas sleeps furious"

Grammaticality

- A community of speakers (eg, Standard American English speakers) share a rough consensus of their implicit rules.

- All the utterances we can generate from these rules are **grammatical**.
  - If we cannot produce an utterance using these rules, it's **ungrammatical**

- But people don't fully agree: everyone has their own idiolect, grammaticality is graded

Example

- Not everyone is as strict for some wh- constraints

  - ⸮"I saw who Emma doubted reports that we had captured in the nationwide FBI manhunt"

[Hofmeister and Sag 2011]

Why do we even need rules?

- Grammaticality rules accept useless utterances
  - "Colorless green ideas sleep furiously"
- And block out perfectly communicative utterances
  - "Me cupcake ate"
- … why would this be a fundamental facet of human intelligence?

- Remember, a basic fact about language is that *we can say anything*
- If we ignore the rules because we know what is probably intended…
  - Then we would be limiting possibilities!

- In my kitchen horror novel where the ingredients become sentient, I want to say
  "the onion chopped the chef".

Language is Compositional

- A set of rules that define grammaticality
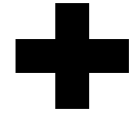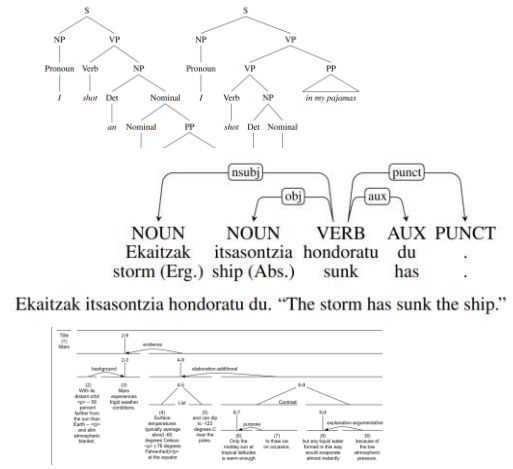
- A lexicon of words that relate to the world we want to talk about

+
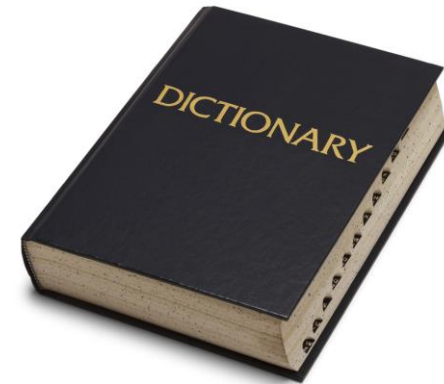
Anything we want to say!

Lecture Plan

1. Structure in human language

    Language is made up of syntactic rules that combine with each other and with the lexicon to form limitless combinations

2. **Linguistic structure in NLP**

3. Going beyond pure structure (in linguistics and deep learning)

4. Multilinguality in NLP

Linguistic Structure in NLP

- Linguistic structure in humans
  - There is a *system* for producing language, that can be described by discrete rules

- Do NLP systems work like that?

- They definitely used to!

Before self-supervised learning

- The way to approach doing NLP was through understanding the human language system, and trying to imitate it

- **Example:** Parsing

  - I want my sentiment analysis system to classify this movie review correctly

    - "My uncultured roommate hated this movie, but I absolutely loved it"

  - How would we do this?

  - We might have some semantic representation of some key words like "hate" and "uncultured", but how does everything relate?

How do humans structure this string of words?

- Many linguists might tell you something like this:
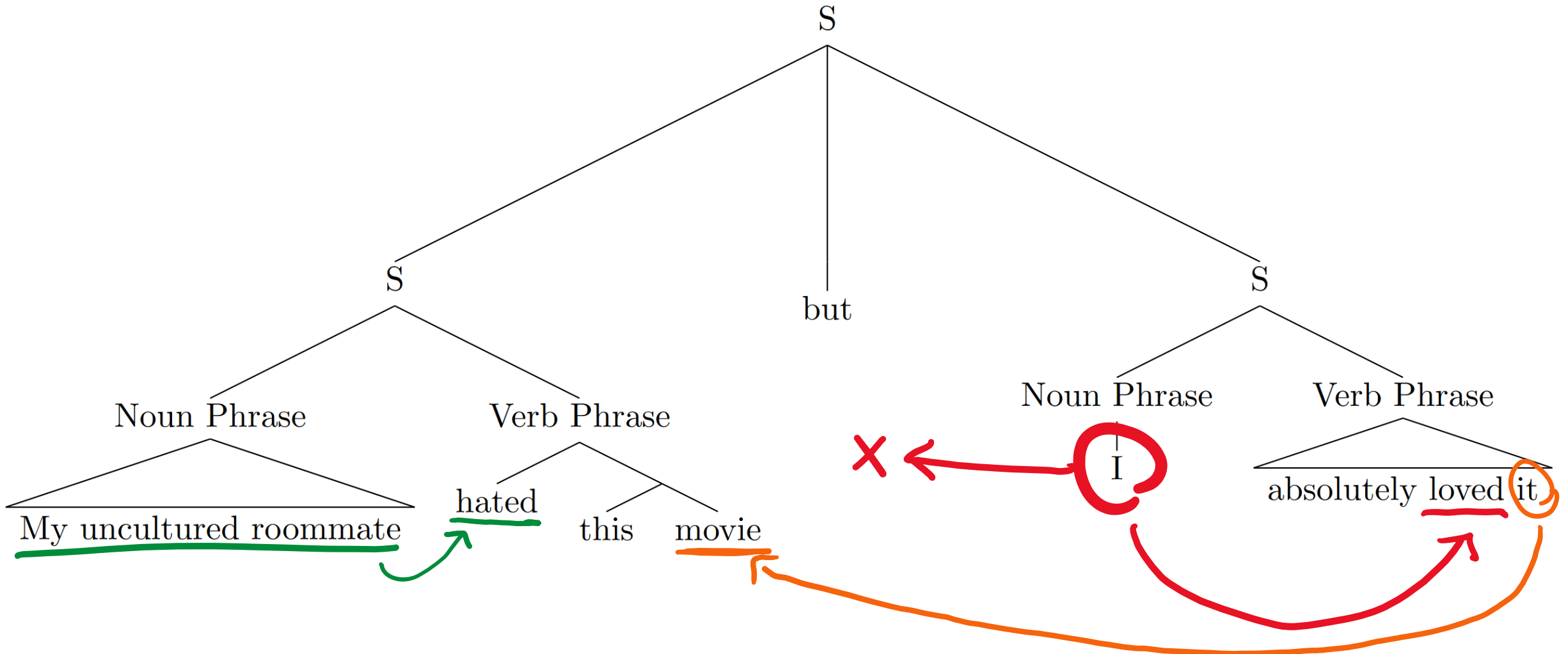
Linguistic theory helped NLP reverse-engineer language

Pipelines

Syntax
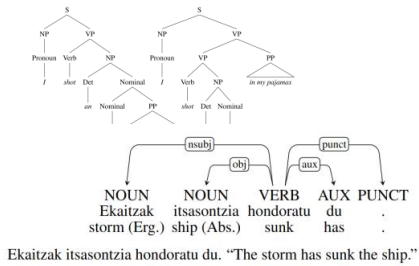
Semantics

Discourse

Input
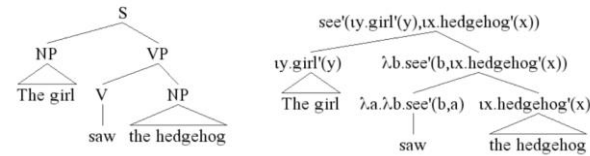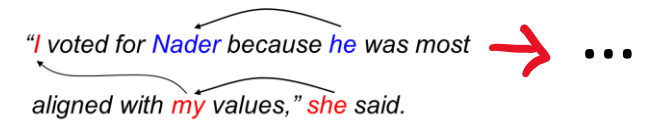
Now, language models just seem to catch on to a lot of these things!

IS Hi! Whaddya think of this movie review? "My uncultured roommate hated this movie, but I absolutely loved it"

As an AI language model, I do not have the ability to form opinions or beliefs. However, based on the given movie review, it seems that the person who wrote it had a positive experience watching the movie, while their roommate did not enjoy it. It's important to remember that everyone has their own tastes and preferences when it comes to movies and other forms of entertainment, and it's okay to have differing opinions. It's always good to approach a movie with an open mind and form your own opinion rather than solely relying on someone else's viewpoint.

29

Linguistic Structure in NLP

- Linguistic structure in humans

  - There is a *system* for producing language, that can be described by discrete rules

- Do NLP systems work like that?

- They definitely used to!

- Now, NLP works better than it ever has before – and we're not constraining our systems to know any syntax

- What about structure in modern language models?

Testing structural knowledge in large language models

Training data

Rules about language

Testing structure: novel words in old structures

- Humans easily integrate new words into our old syntactic structures

- Shows that syntactic structures exist independently of the words they have appeared with

- Do language models have this ability?

Jabberwocky
BY LEWIS CARROLL

'Twas brillig, and the slithy toves
 Did gyre and gimble in the wabe:
All mimsy were the borogoves,
 And the mome raths outgrabe.

"Beware the Jabberwock, my son!
 The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
 The frumious Bandersnatch!"

He took his vorpal sword in hand;
 Long time the manxome foe he sought—
So rested he by the Tumtum tree
 And stood awhile in thought.

And, as in uffish thought he stood,
 The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
 And burbled as it came!

One, two! One, two! And through and through
 The vorpal blade went snicker-snack!
He left it dead, and with its head
 He went galumphing back.

"And hast thou slain the Jabberwock?
 Come to my arms, my beamish boy!
O frabjous day! Callooh! Callay!"
 He chortled in his joy.

The COGS benchmark: New word-structure combinations

Task: semantic interpretation
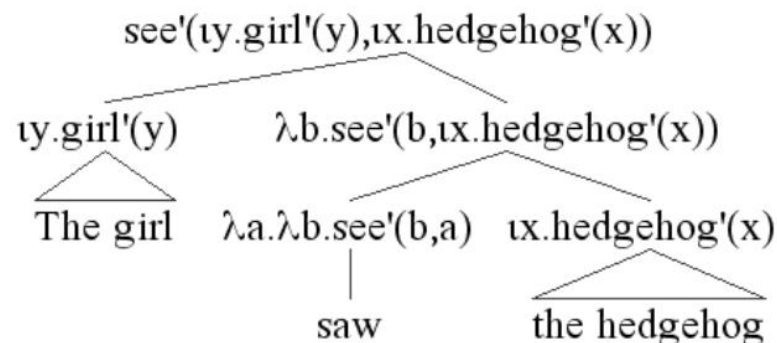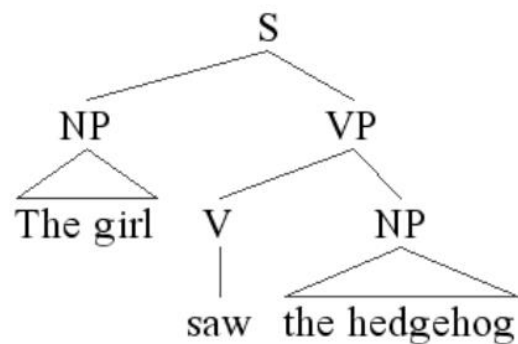
Training and test sets have distinct words and structures in different roles

[Kim and Linzen 2020]
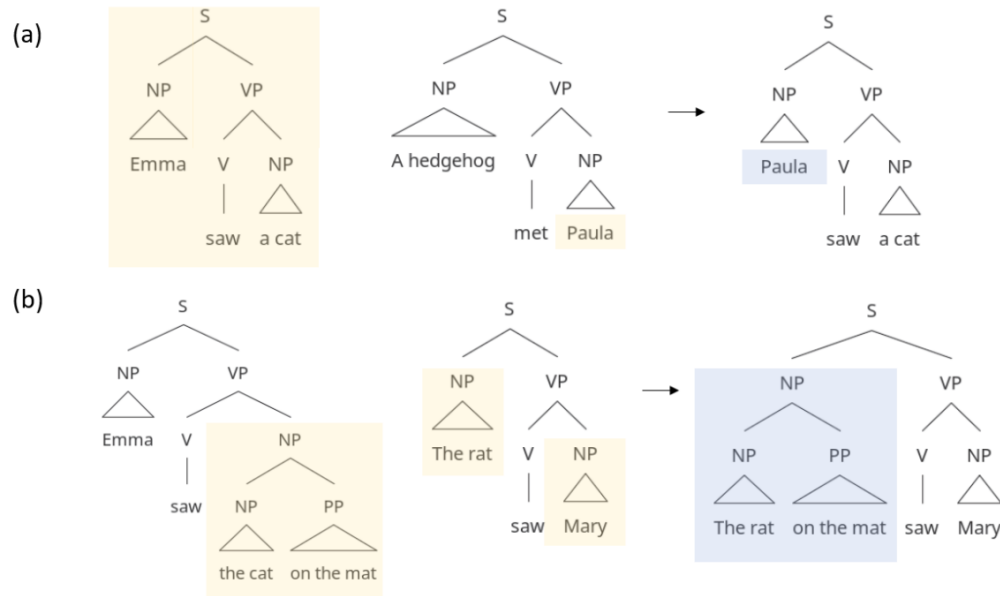
Large language models haven't aced this (at least T5)

	model	average
mar.	marian_defr_scr	62.7 ± 0.5
	marian_defr	83.4 ± 0.1
t5-base	t5_base_scr	32.3 ± 2.2
	t5_base	83.3 ± 0.1
	mt5_base	83.4 ± 0.1
	ct5_base	82.6 ± 0.1
	pt5_base	16.1 ± 2.3
t5-3b	t5_3b_scr	15.5 ± 0.6
	t5_3b	84.1 ± 0.2
	mt5_xl	84.6 ± 0.1
	pt5_xl	0.0 ± 0.0

Table 2: Exact match accuracies on the COGS generalization set. Numbers in small font are standard errors over at least three independent runs for each model. Breakdown of accuracy into different conditions can be found in the output files provided on the accompanying github repository.

[Orhan 2022]

Testing syntax in Jabberwocky sentences

- We can test a model's **latent space** to see if it encodes structural information [Hewitt and Manning 2019]

- Does this work when we introduce new words?
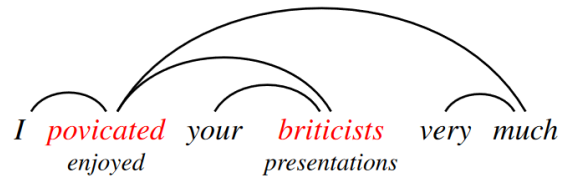
Figure 2: An unlabeled undirected parse from the EWT treebank, with Jabberwocky substitutions in red.
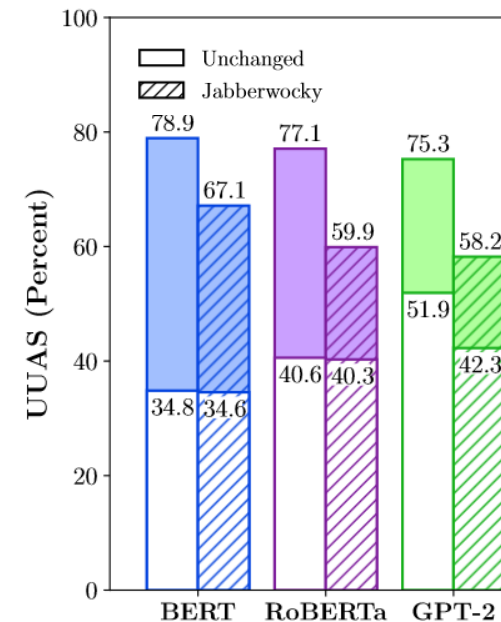
[Maudslay and Cotterell 2021]

Testing how syntactic structure maps on to meaning

- Syntax → Meaning

- In English, the syntax of **word order** gives us the "who did what to whom" meaning

  - A verbed B → A is the do-er, B is the patient

- We can test this in language models

... and then **A** talked to **B**

It's about time that **C** found the sneaky and wiley **D**...

... I would have never thought that the **E** would knock down the **F**.

→ Language Model

Latent

Space

More subject

More object

A C
E

F B

D

[Papadimitriou et al 2022]
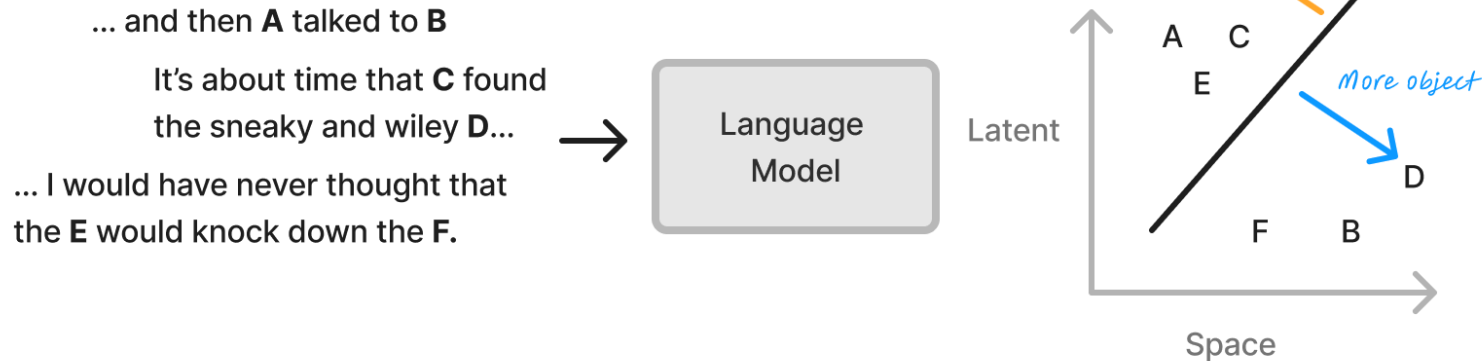
Testing how syntactic structure maps on to meaning

- Syntax → Meaning

- In English, the syntax of word order gives us the "who did what to whom" meaning

  - A verbed B → A is the do-er, B is the patient

- We can test this in language models

The **chef** chopped the **onion**,   The **onion** chopped the **chef**

Will these be classified differently?

[Papadimitriou et al 2022]

Testing how syntactic structure maps on to meaning

- Syntax → Meaning

- In English, the syntax of word order gives us the "who did what to whom" meaning

  - A verbed B → A is the do-er, B is the patient

- We can test this in language models

Progressions of the same words in original and swapped sentences

[Papadimitriou et al 2022]
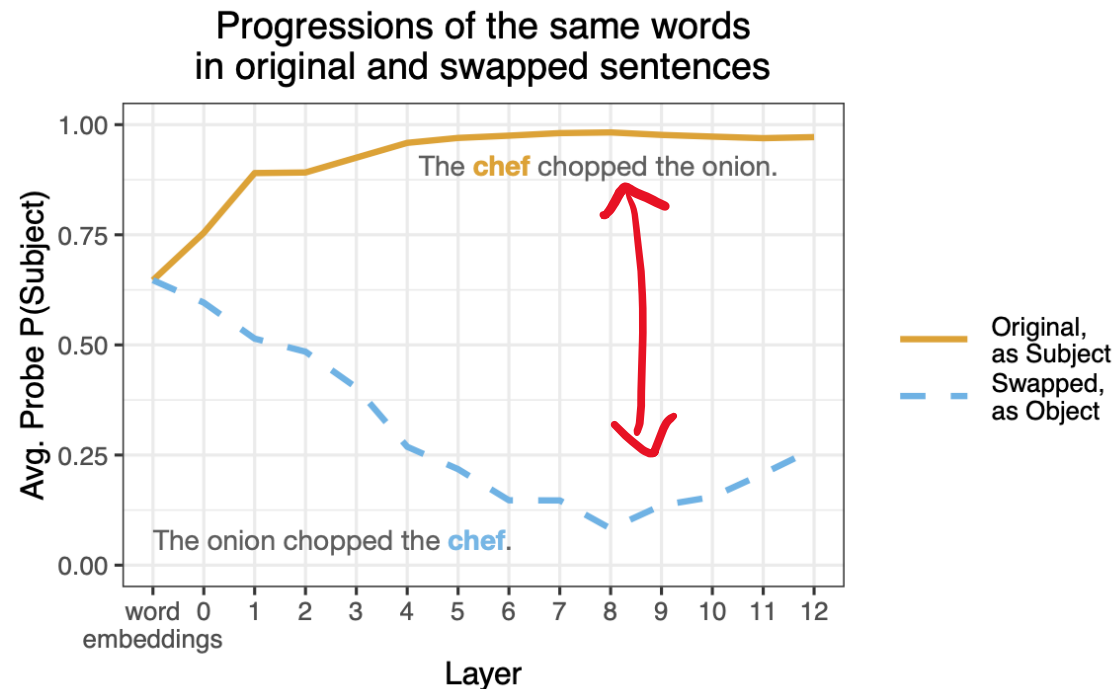
Structure in language models

- Language models aren't engineered around discrete, linguistic rules

- But the pretraining process isn't just a bunch of surface-level memorization

- How much do we get a discrete, rule-based system from large scale pretraining?

  - There's syntactic knowledge, but it's complicated

- But remember – there's no ground truth for how language works!

  - If we knew how to fully describe English with a bunch of discrete rules, we would just make an old-school NLP system and it would be amazing.

Lecture Plan

1. Structure in human language

2. Linguistic structure in NLP

3. **Going beyond pure structure (in linguistics and deep learning)**

4. Multilinguality in NLP

Meaning plays a role in linguistic structure

- There is a lot of rich information in words that affects the final structure of language

- The rich **semantics** of words is always playing a role in forming and applying the rules of language

  - I ate a cookie   I ate

  - I devoured a cookie   ✗ I devoured

  - Time/days/afternoon/harvests?/moons?/✗ trees **elapsed**

  - I **diagonalized** the matrix

That is how we train our models these days!

Start with rich semantic input!
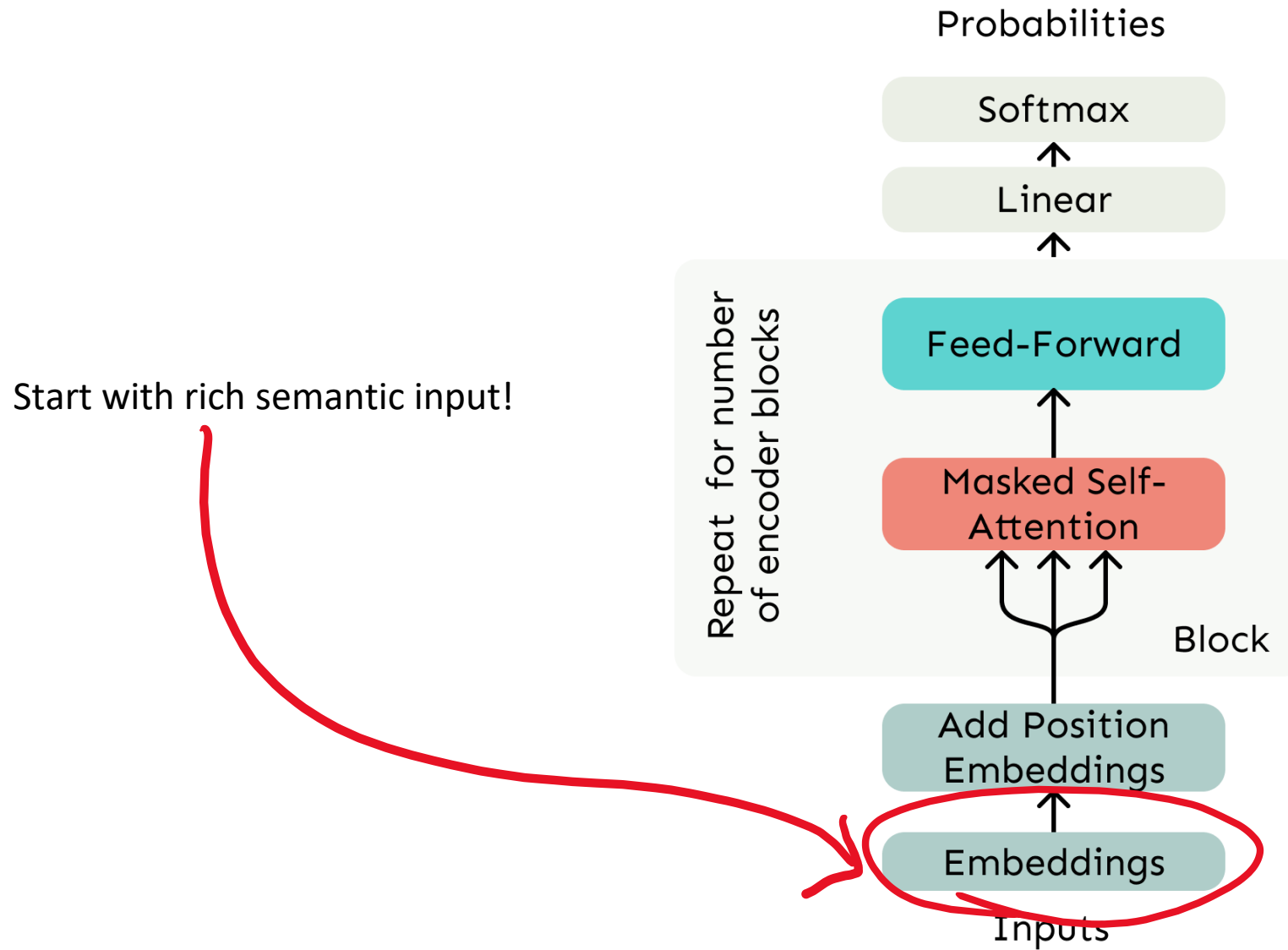
Example: differential object marking

- Structurally, anything can be an object

  - Remember, "The onion chopped the chef"

- But many languages have a special syntactic way of dealing with this

  - "Hey! Watch out! That's an object"

Spanish:

(1)   a. Encontré   un problema.
        I.found   a   problem
        'I found a problem'
      b. Encontré   a un superviviente.
        I.found   A a   survivor
        'I found a survivor'

Hindi:

(5) *wo-Ø*   *ek bakraa-Ø /ek bakre=ko*   *bec-taa*   *hae*
    He-NOM   one goat-NOM/one goat=ACC   sell-IPFV.SG.M.   be.PRS.3SG.
    'He sells a goat / the goat'

43

[Fábregas 2013;
de Hoop and Narashiman 2005]

Language models are also aware of these gradations

... and then **A** talked to **B**

It's about time that **C** found the sneaky and wiley **D**...

... I would have never thought that the **E** would knock down the **F**.

→ Language Model

Latent

Space

More subject

More object

A   C
  E
        D
      F   B

[Papadimitriou et al 2021]
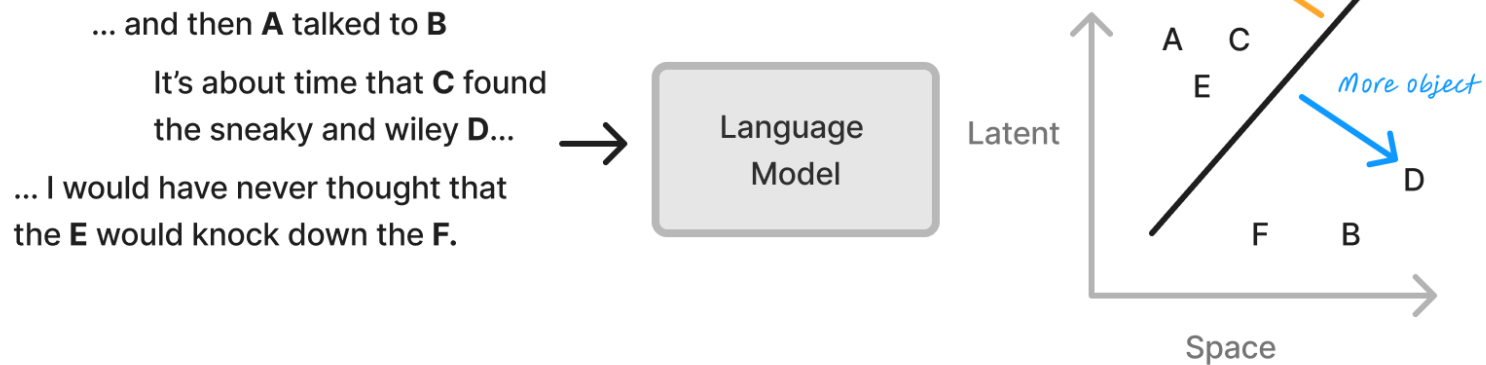
Language models are also aware of these gradations

Accuracy is high, but…
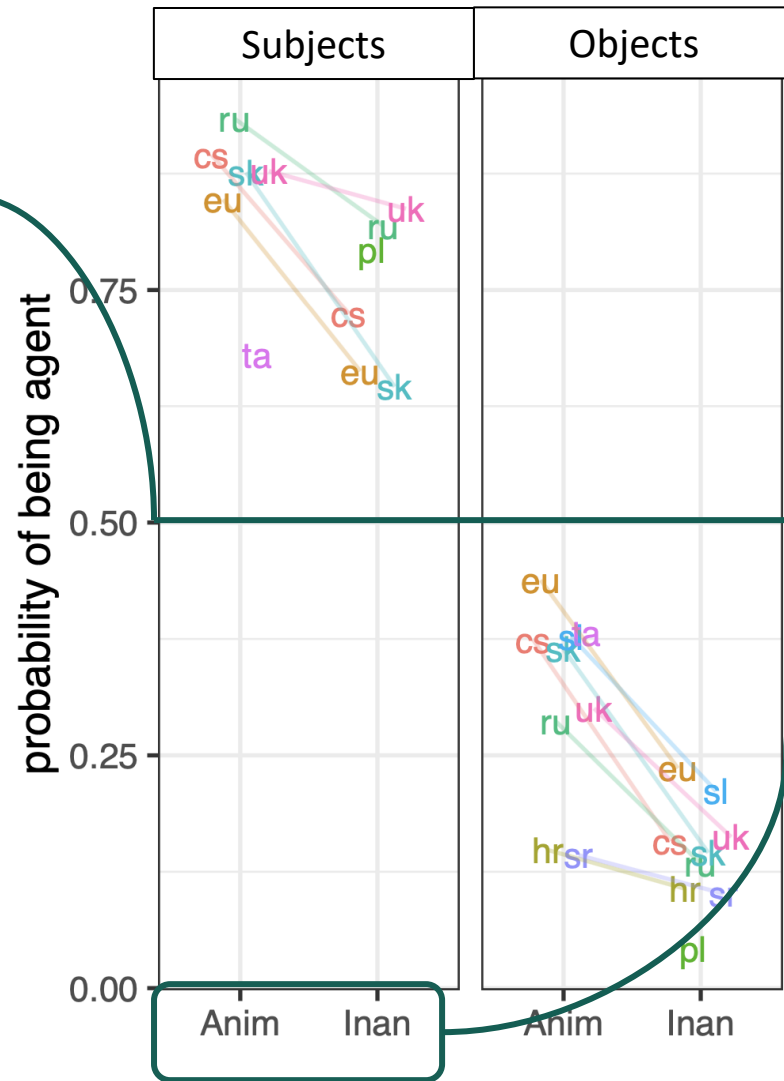
Animacy influences this grammatical distinction

[Papadimitriou et al 2021]

Example: maybe not all structure-word compositions are possible

- In many cases, if something seems too outlandish, we assume the more plausible interpretation

- Psycholinguistics experiments:

  - "The mother gave the daughter the candle", "The mother gave the candle to the daughter"

  - "The mother gave the candle the daughter"

- Of course, outlandish meanings are not impossible to express (nothing is!)

  - "The mother picked up her daughter, and handed her to the candle, who is sentient"

- Marking less plausible things more prominently is a pervasive feature of grammar

[Gibson et al 2013]

Meaning can't always be composed from individual words

- Language is full of idioms

  - And not just canned wisdoms like "don't count your chickens before they hatch"

- We're constantly using constructions that we couldn't get from just a syntactic + semantic parse

  - "I wouldn't <u>put it past</u> him", "They're <u>getting to me</u> these days", "That won't <u>go down well</u> with the boss"…

- And even mixed constructions that can compositionally take arguments!

  - "He won't X, let alone Y", "She slept the afternoon away", "The bigger they are, the more expensive they are", "That travesty of a theory"

[Croft 2020, Jackendoff 2011; Tomasello 2006]

Testing constructions in language models

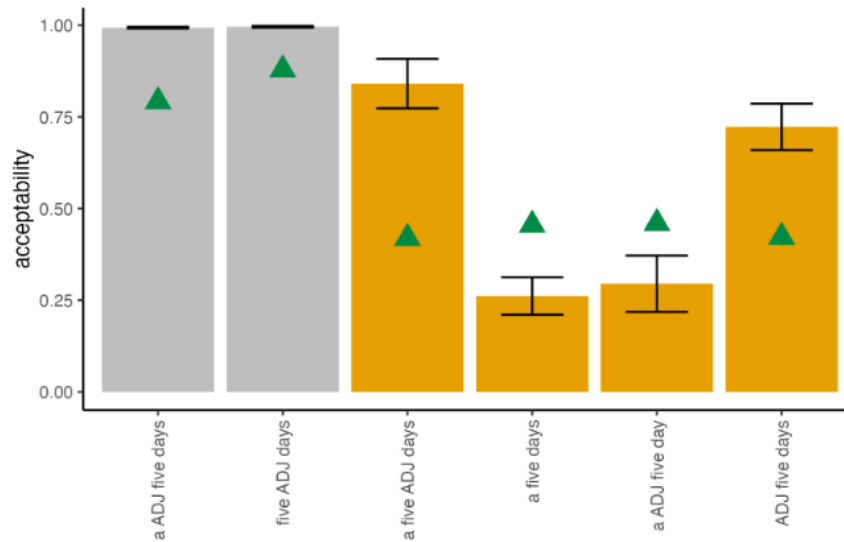"A beautiful five days in Austin" construction

"The X-er the Y-er" construction

Figure 1: GPT-3 acceptability judgments (bars), compared to human ratings (green triangles) on a matched set of sentences.

Sentence	Label
" The higher up the nicer ! "	Positive
She thinks the more water she drinks the better her skin looks .	Positive
It becomes an obsession lightly because the more fish you catch the higher your adrenaline flows .	Positive
It is worth noting , however , that the more specific you are the better .	Positive
In other words , the more videos you make the greater your audience reach .	Positive
Subtract the smaller from the larger . "	Negative
The way the older guys help out the younger guys is fantastic .	Negative
In this procedure the lower lip is pulled ventrally to expose the lower incisors .	Negative
The 5th bedroom is on the lower floor with easy access to the lower bath .	Negative
Note the distinctive bend of the larger vein adjacent to the smaller vein at the top .	Negative

Table 8: Examples of corpus data

48

[Mahowald 2023; Weissweiler et al 2022]

The meaning of words is sensitive and influenced by context

Frame	Sense
1. break the vase	shatter
2. break the computer	render inoperable
3. break the news	reveal
4. break the silence	interrupt
5. break the record	surpass
6. break the code	decipher
7. break the law	violate
8. break the horse	tame
9. break a $10 bill	make change
10. break the fall	lessen
11. the weather broke	changed
12. the day broke	began

(a) Uses without particles/predicates.

Frame	Sense
13. break off the engagement	end
14. break out	begin
15. break out in hives	get
16. break into the building	intrude
17. break down the problem	analyze
18. break down the proteins	decompose
19. break in	enter
20. break in	interrupt
21. break free	escape
22. break even	profit = loss
23. break forth	emerge
24. break to the right	turn

(b) Uses with particles/predicates.

Table 2: Senses for *break*. A comprehensive account of senses may not be possible (Section 5.3).

[Petersen and Potts 2022]

Fine-grained lexical semantics in language models

Figure 2: t-SNE of break with RoBERTa-large, layer 1
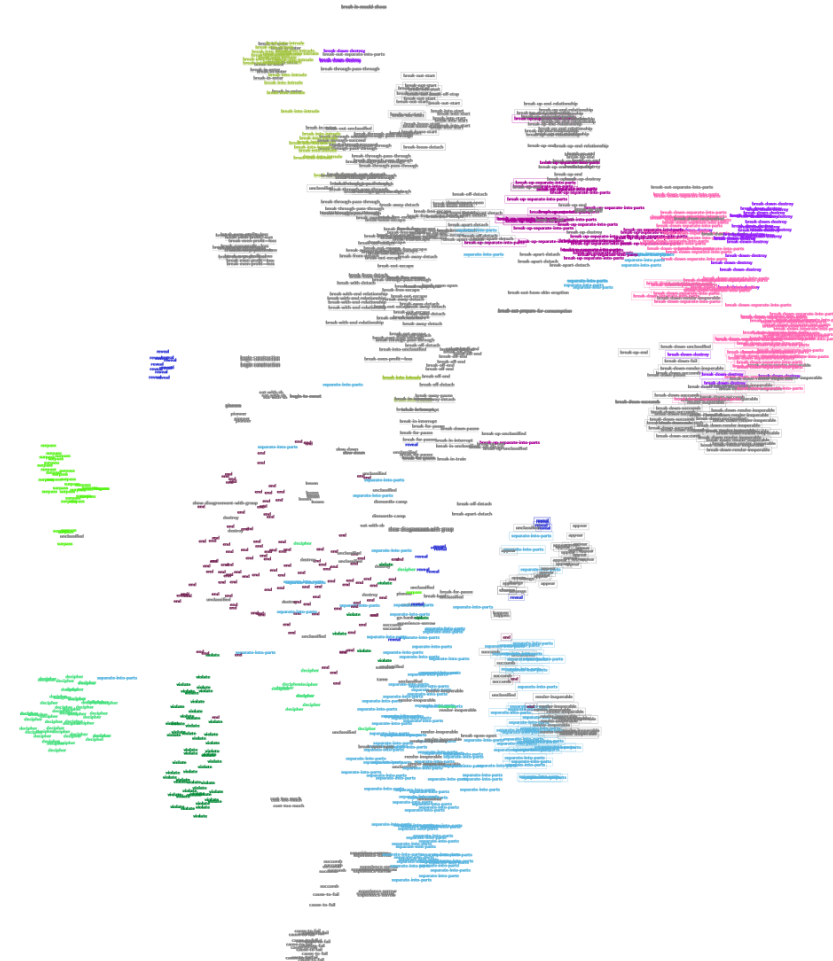
Figure 3: t-SNE of break with RoBERTa-large, layer 24

[Petersen and Potts 2022]

A big question in NLP: how to strike the balance?

"While language is full of both broad generalizations and item-specific properties, linguists have been dazzled by the quest for general patterns. Of course, the abstract structures and categories of language *are* fascinating. But I would submit that what is even more fascinating is the way that these general structures arise from and interact with the more specific items of language use, producing a highly conventional set of general and specific structures that allow the expression of both conventional and novel ideas"

Joan Bybee, *Frequency of Use and the Organization of Language* (2006)

A big question in NLP: how to strike the balance?

- Language is characterized by the fact that it's an amazingly abstract system

  - And we want our models to capture that

- But meaning is so rich and multifaceted

  - High-dimensional spaces are much better at capturing these specificities subtleties than any rules we could come up with

- Where do deep learning models stand now, between surface-level memorization and abstraction?

  - This is what a lot of analysis and interpretability work is trying to understand

  - Stay tuned for the analysis lectures!

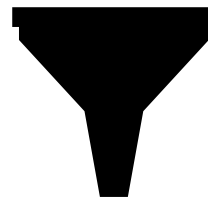- Remember: this is not even a solved question for humans!

Lecture Plan

1. Structure in human language

2. Linguistic structure in NLP

3. Going beyond pure structure (in linguistics and deep learning)

4. **Multilinguality in NLP**

So far, we've been talking about English

- Languages are so diverse!

- There are ~7,000 languages in the world

Multilingual Language Models

One transformer language model

Multilingual Language Models

- Multilingual language models let us **share parameters**

  - between high-resource languages and low-resource languages (~100s in total)

- Pretraining and transfer learning have brought so much unexpected success to NLP

- We get great linguistic capability and generality that we don't ask for in training

- Will this self-supervised learning paradigm deliver **between** languages?

Concerns for Multilinguality in NLP

- Languages are so diverse!

- There are ~7,000 languages in the world

- We can look at multilingual NLP through two lenses

- Languages are also remarkably diverse – **Lanugage typology**

  - Does multilingual NLP capture the specific differences of different languages?

- But languages are similar to each other in many ways – **Language universals**

  - Does multilingual NLP capture the parallel structure between languages?

Language diversity: evidentiality

- "In about a quarter of the world's languages, **every statement must specify the type of source on which it is based**"

Examples in Tariana

1.1 Juse   irida   di-manika-**ka**
José   football   3sgnf-play-REC.P.VIS
'José has played football (we saw it)'

1.2 Juse   irida   di-manika-**mahka**
José   football   3sgnf-play-REC.P.NONVIS
'José has played football (we heard it)'

1.3 Juse   irida   di-manika-**nihka**
José   football   3sgnf-play-REC.P.INFR
'José has played football (we infer it from visual evidence)'

1.4 Juse   irida   di-manika-**sika**
José   football   3sgnf-play-REC.P.ASSUM
'José has played football (we assume this on the basis of what we already know)'

1.5 Juse   irida   di-manika-**pidaka**
José   football   3sgnf-play-REC.P.REP
'José has played football (we were told)'
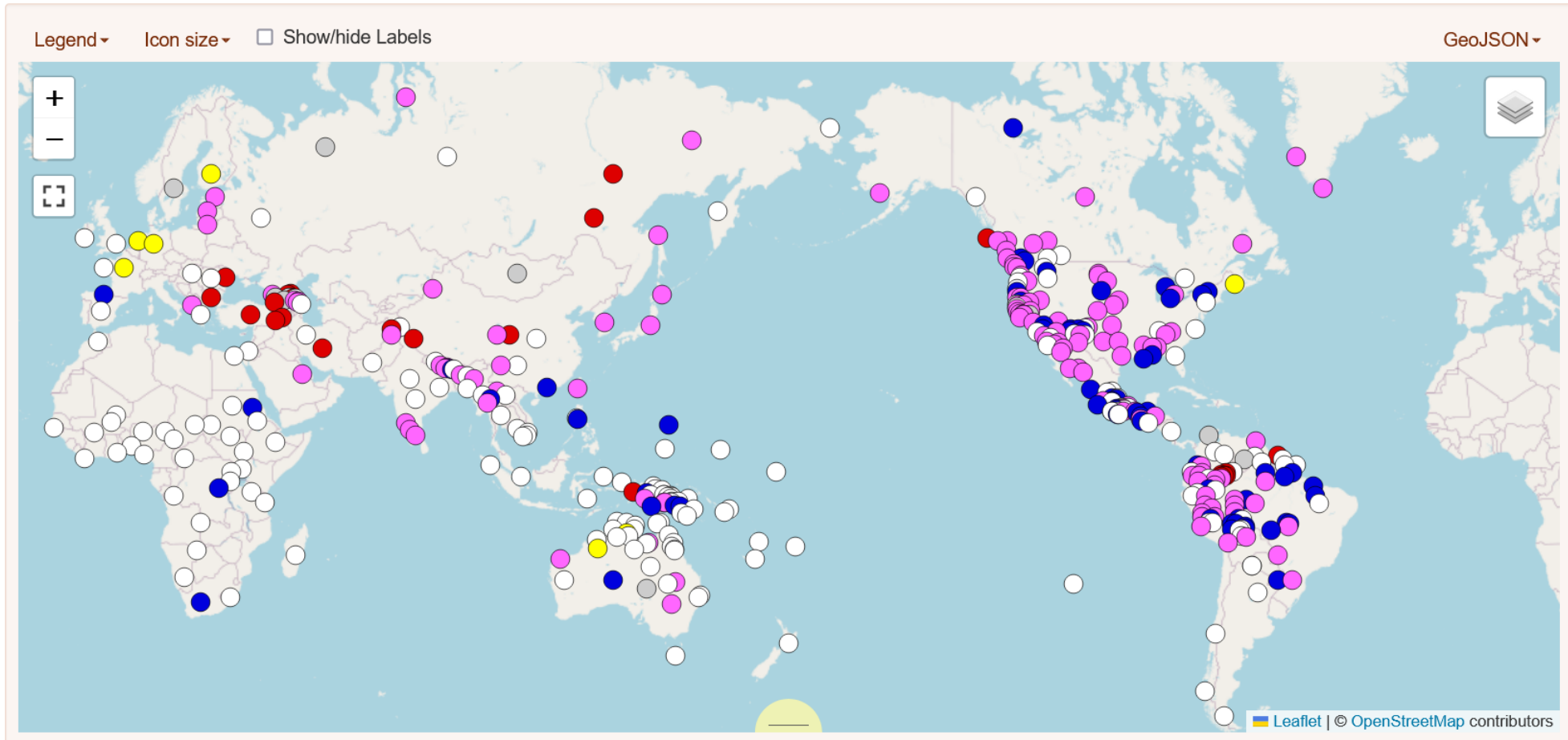
[Aikhenvald 2004]

Feature 78A: Coding of Evidentiality

This feature is described in the text of chapter 78 [Coding of Evidentiality] by Ferdinand de Haan [cite]

You may combine this feature with another one. Start typing the feature name or number in the field below.

[× 78A: Coding of Evidentiality] [Submit]

Values

○	No grammatical evidentials	181
● (pink)	Verbal affix or clitic	131
● (red)	Part of the tense system	24
● (blue)	Separate particle	65
● (yellow)	Modal morpheme	7
● (grey)	Mixed	10

Legend▾   Icon size▾   ☐ Show/hide Labels   GeoJSON▾

Leaflet | © OpenStreetMap contributors

Feature 81A: Order of Subject, Object and Verb

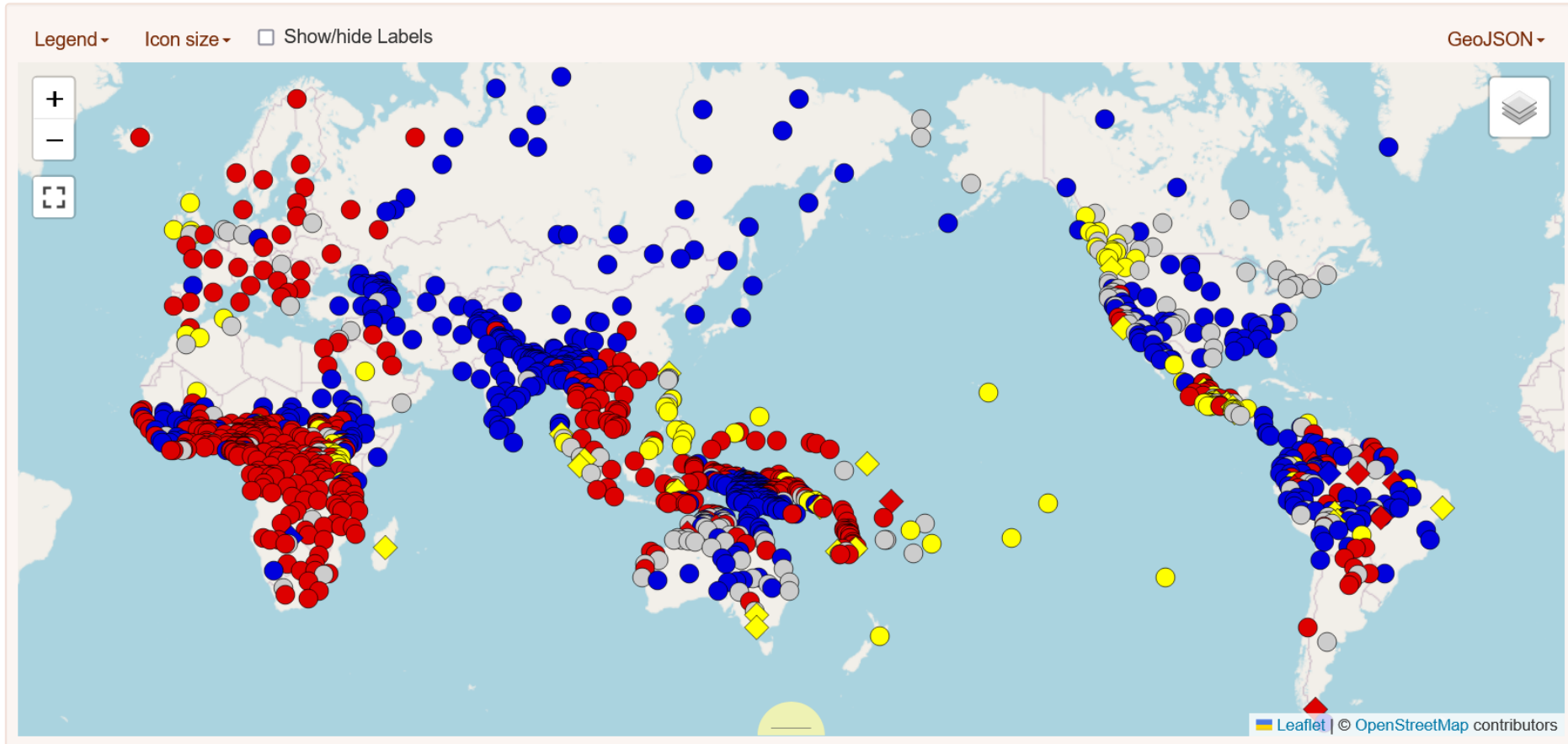This feature is described in the text of chapter 81 [Order of Subject, Object and Verb] by Matthew S. Dryer [cite]

You may combine this feature with another one. Start typing the feature name or number in the field below.

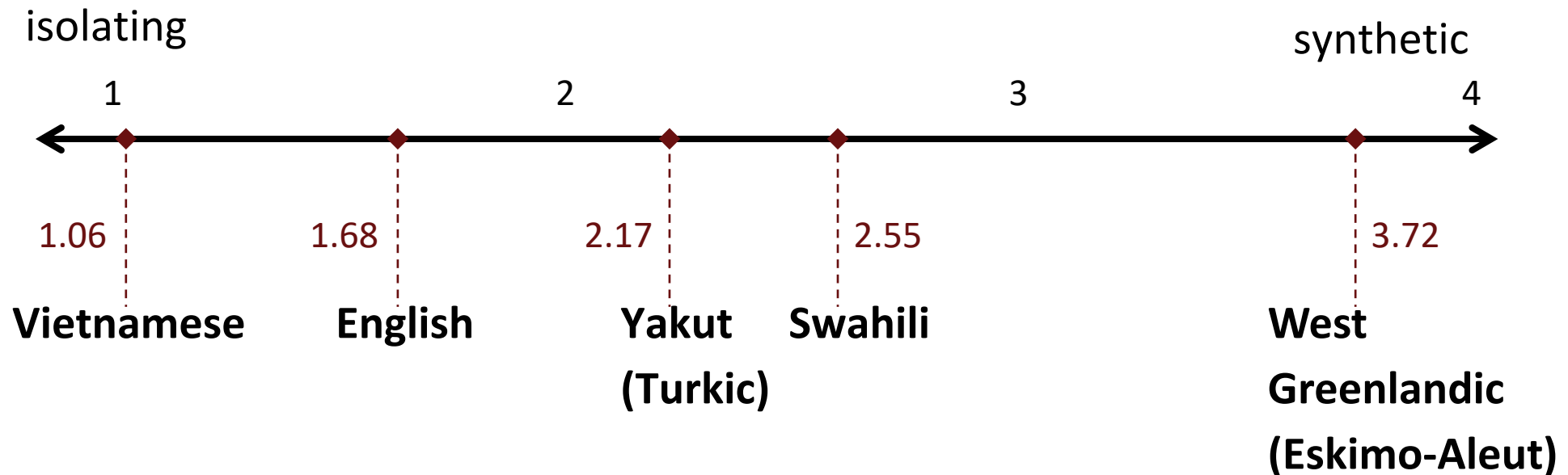× 81A: Order of Subject, Object and Verb

[Submit]

Values

●	SOV	564
●	SVO	488
●	VSO	95
◆	VOS	25
◆	OVS	11
◆	OSV	4
●	No dominant order	189

Legend ▾   Icon size ▾   ☐ Show/hide Labels

GeoJSON ▾

Leaflet | © OpenStreetMap contributors

Morphemes per word

[Joseph Greenberg. 1954. A Quantitative Approach to the Morphological Typology of Language. IJAL 26:3.]

isolating

synthetic

1   2   3   4

1.06   1.68   2.17   2.55   3.72

Vietnamese   English   Yakut   Swahili   West
(Turkic)   Greenlandic
(Eskimo-Aleut)

[Slide from Dan Jurafsky]

Many Morphemes per word: Yupik

tuntussuqatarniksaitengqiggtuq

tuntu	-ssur	-qatar	-ni	-ksaite	-ngqiggte	-uq
reindeer	-hunt	-FUT	-say	-NEG	-again	-3SG.IND

"He had not yet said again that he was going to hunt reindeer."

Except for the morpheme *tuntu* "reindeer", none of the other morphemes can appear in isolation.[a]

[https://en.wikipedia.org/wiki/Polysynthetic_language]

How are motion events described?

- In English, the **manner of motion** is usually expressed on the verb

  The bottle **floated into** the cave

- In Spanish, the **direction of motion** is usually expressed on the verb

  La botella **entró** a la cueva **flotando**

- Satelite-framing languages ("into", "on to", "off of" are satellites) and verb-framing languages

- Not an exclusive categorization!

我<u>跑出</u>了厨房。
Wǒ <u>pǎo chū</u> le chúfáng.
I run exit PFV kitchen
"I ran out of the kitchen."

[Talmy 1985, Chen and Guo 2009]

Language universals

- (We'll gloss over this a bit here) Universal Grammar in the Chomskyan sense
  - Is there an abstract structure that unites all languages?
  - This is a huge question in linguistics
  - Defining an abstraction where we can say: all languages are some version of it

- Tendencies exhibited across languages: most languages that do A also do B [Greenberg 1963]

- Languages all deal in similar types of relations, like subject, object, modifiers [Universal Dependencies Nivre et al 2016, Foley and van Valin 1984]

Language universals

- (We'll gloss over this a bit here) Universal Grammar in the Chomskyan sense

  - Is there an abstract structure that unites all languages?

  - This is a huge question in linguistics

  - Defining an abstraction where we can say: all languages are some version of it

- Tendencies exhibited across languages: most languages that do A also do B [Greenberg 1963]

- Languages all deal in similar types of relations, like subject, object, modifiers [Universal Dependencies Nivre et al 2016, Foley and van Valin 1984]

Ask me!

Subject – Object relations are encoded in parallel

… and then **A** talked to **B**

    It's about time that **C** found
    the sneaky and wiley **D**…

… I would have never thought that
the **E** would knock down the **F**.

→

Language Model

Latent

More subject

A   C

E

More object

D

F   B
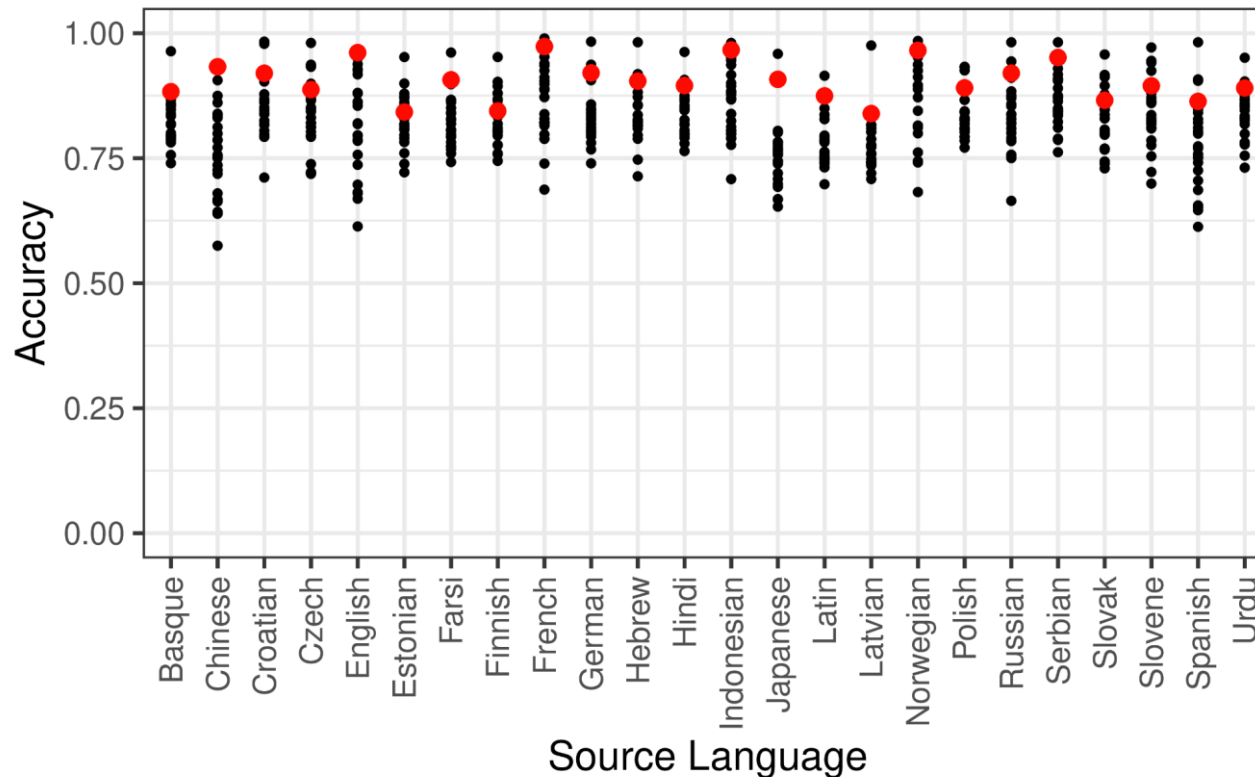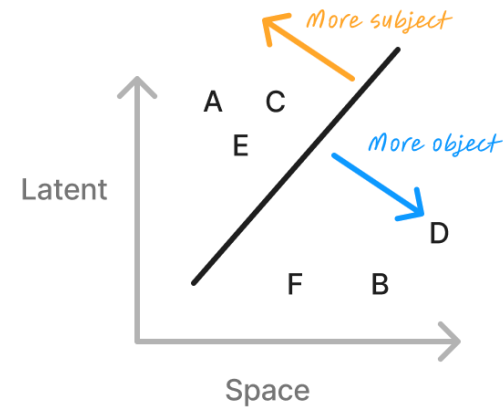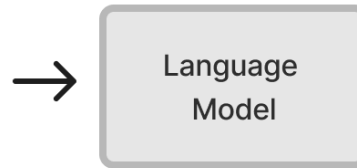
Space

Red dots are in-language accuracy, black dots are cross-language accuracy.

[Papadimitriou et al 2021]
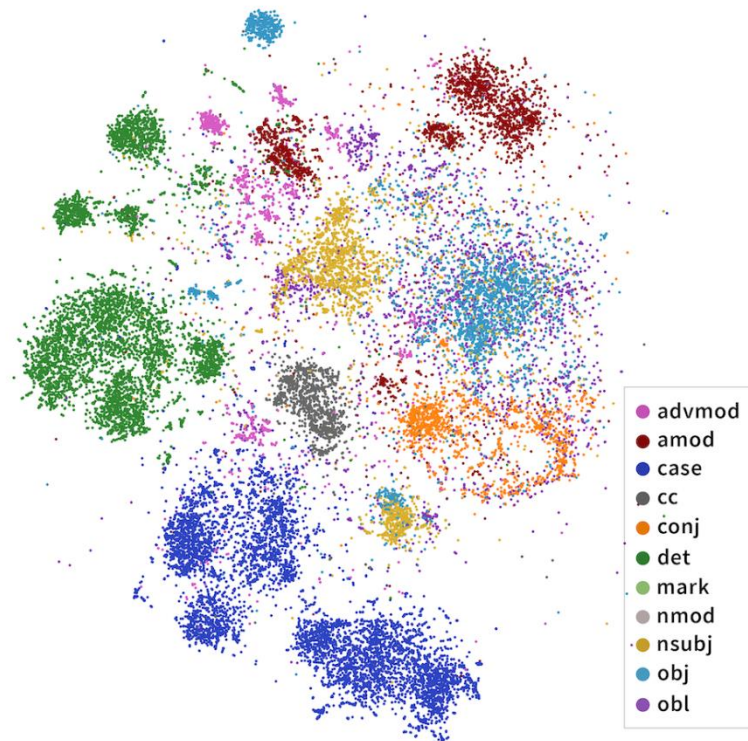
UD relations are encoded in parallel ways

Figure 5: t-SNE visualization of syntactic differences in Spanish projected into a holdout subspace (learned by a probe trained to recover syntax trees in languages other than Spanish). Despite never seeing a Spanish sentence during probe training, the subspace captures a surprisingly fine-grained view of Spanish dependencies.

[Chi et al 2021]

But language specificity is also important!

- Does Multilingual BERT have an accent?

[Papadimitriou et al 2022]
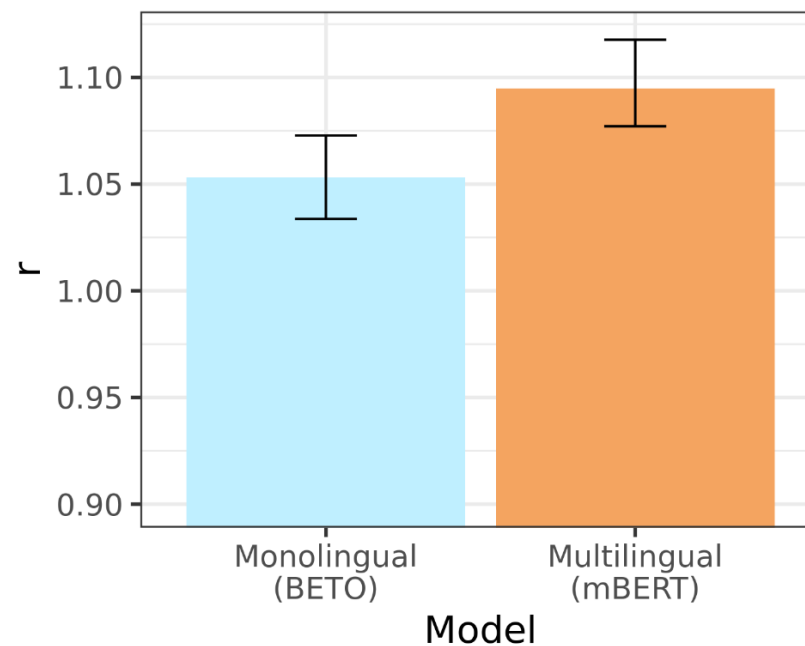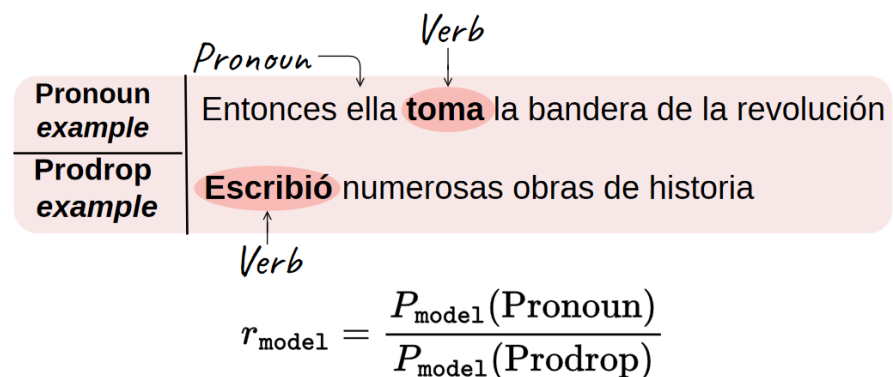
Data quality is very variable in multilingual corpora

20% of languages are less than 50% correct
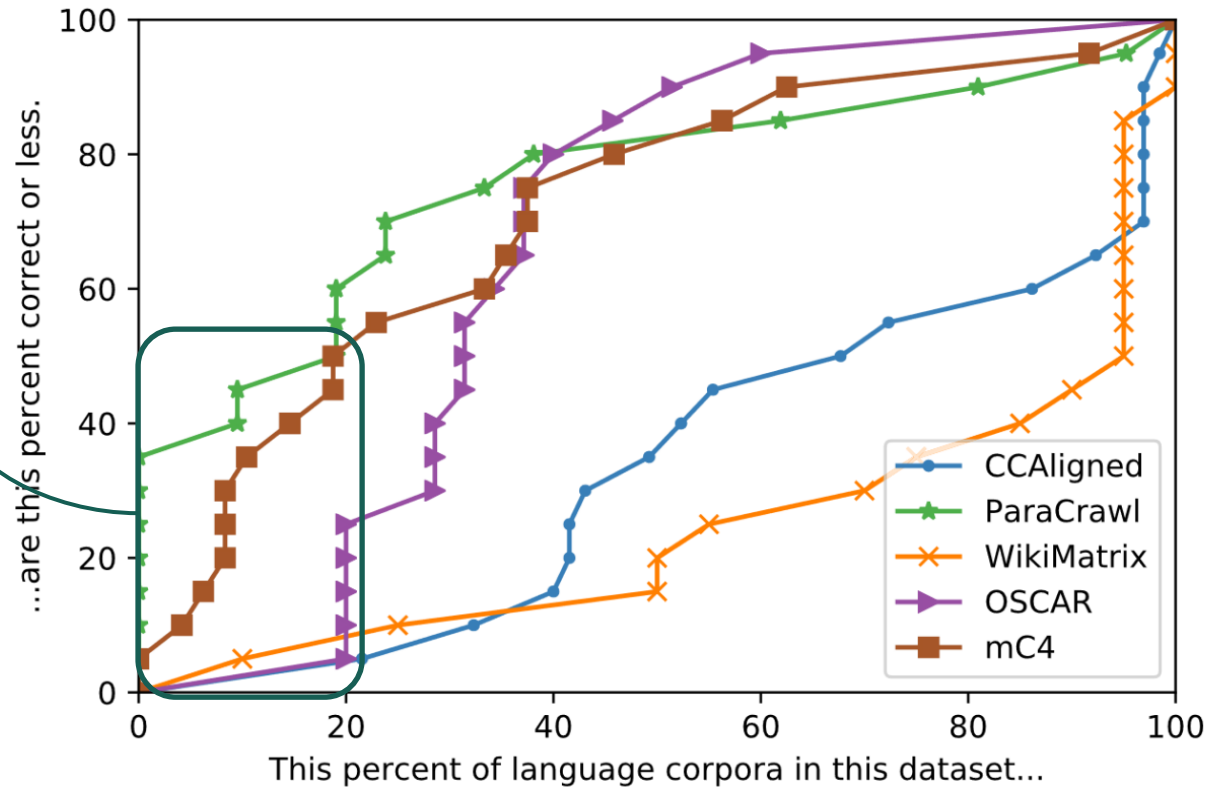

Figure 1: Fraction of languages in each dataset below a given quality threshold (percent correct).

[Kreutzer et al 2022; Nekoto et al 2020]

Less parameter-sharing in multilingual models

- AfriBERTa: pretrain a model only on low-resource languages
  - "Small Data? No Problem!"

- XLM-V: a larger vocabulary so each language is almost separate
  - Why should we do vocabulary sharing?
  - Use a huge vocabulary: 1 million words
  - Better performance on cross-lingual evaluation

[Ogueji et al 2021; Liang et al 2023]

Combining language-specificity and language-generality

Fine-tune some parameters for specific languages

2a. Sparse language fine-tuning

1. Pretrained model

2b. Sparse task fine-tuning
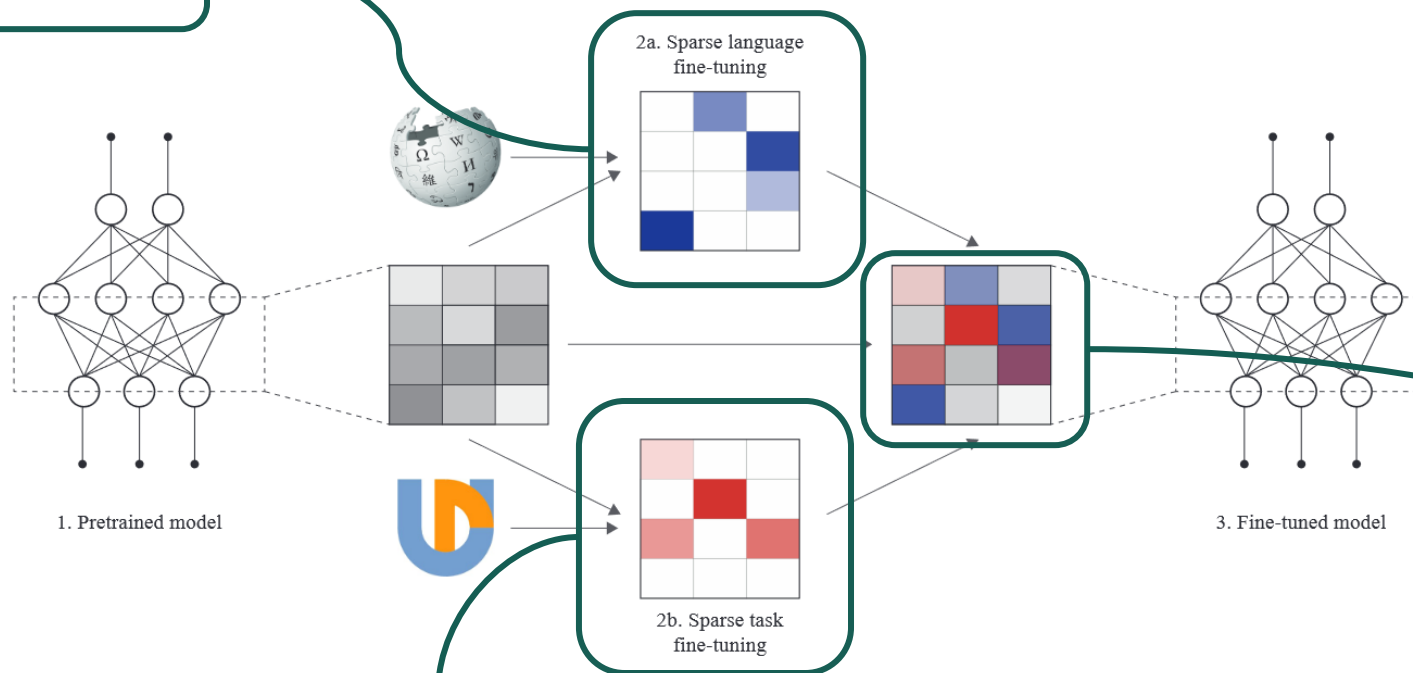
3. Fine-tuned model

And others for specific tasks

And we can just add the separate fine-tuned models!

[Ansell et al 2021]

Multilingual NLP

- We want to balance **language-generality** and good parameter sharing…

- With preserving language **diversity** and specificity

- How is multilingual NLP doing?

  - Other issues we've not mentioned: dialects, language continuums…

- How can deep learning work for low-resource languages?

- What are the ethics of working in NLP for low-resource languages?

Lecture Plan

1. Structure in human language

2. Linguistic structure in NLP

3. Going beyond pure structure (in linguistics and deep learning)

4. Multilinguality in NLP

Linguistics can be a tool to help us understand deep learning

- We can investigate what is going on in black box models

- The subtleties in linguistic analysis can help us understand what we want or expect from the models we work with

- NLP engineering does not reverse-engineer human language…

- …but linguistic insights still have a place in understanding large language models

- And in many more ways beyond what we've discussed here!

  - Language acquisition in babies and self-supervised learning

  - Language and other cognitive systems (vision, instructions, music)

  - Discourse, conversation, and communication

  - …