

Parsing and the Strong Minimalist Thesis



Sandiway Fong
Dept. of Linguistics
University of Arizona
sandiday.arizona.edu
sandiday AT arizona EDU





Strong Minimalist Thesis (SMT)

• Generative Grammar

- *idea: a formal system with derivation rules that generate sentences*

- *John ate a sandwich / *Sandwich a ate John*

(pg.113, Chomsky, 1956)

- **PSG:** ordered phrases, $S \rightarrow NP, VP$. $VP \rightarrow V, NP$. etc.

- Later on: efficient parsing algorithms invented (for artificial languages)

- **TG:** shuffling units around makes for simpler descriptions

- *the man -past have -en be -ing take the book*

(pg.120)

- *the man have-past be-en take-ing the book*

- *the man had been taking the book*

- $NP_1 \quad aux \quad V \quad NP_2 \quad \Rightarrow \quad NP_2 \quad aux \quad be \quad en \quad V \quad by \quad NP_1$

- $[the \textit{man}] \textit{-past eat [the food]} \Rightarrow [the \textit{food}] \textit{-past be -en eat by [the man]}$

(pg.121)

- *the shooting of the hunters (ambig.) / (unambig.) the growling of lions / the raising of flowers*

(pg. 123)

- No known efficient parsing algorithm

- **FSA:** even simpler system, more complicated descriptions: linear concatenation

- *plenty of evidence against this*

- *if S_1 then S_2 / either S_3 or S_4 / the man who said that S_5 is arriving today*

(pg. 115)

- Efficient parsing algorithms

- Chomsky (1956). Three models for the Description of Language. *IEEE Transactions on Information Theory*.

Transformations

Context-Free
grammar:

(20) $\Sigma : \# \hat{\text{Sentence}} \#$
 $F: \text{Sentence} \rightarrow \text{NP} \hat{\text{VP}}$
 $\text{VP} \rightarrow \text{Verb} \hat{\text{NP}}$

(32) (i) $\text{Verb} \rightarrow \text{Auxiliary} \hat{\text{V}}$
(ii) $\text{V} \rightarrow \text{take, eat, ...}$
(iii) $\text{Auxiliary} \rightarrow \text{C(M)}(\text{have} \hat{\text{en}}) (\text{be} \hat{\text{ing}})$
 $(\text{be} \hat{\text{en}})$
(iv) $\text{M} \rightarrow \text{will, can, shall, may, must}$
(v) $\text{C} \rightarrow \text{past, present}$

There is one major limitation on the independence of the elements introduced in (32). If we choose an intransitive verb (e.g., "come," "occur," etc.) as V in (32), we cannot select $\text{be} \hat{\text{en}}$ as an auxiliary. We cannot have such phrases as "John has been come," "John is occurred," and the like.

This complex and unique behavior of $\text{be} \hat{\text{en}}$ suggests that it would be desirable to exclude it from (32) and to introduce passives into the grammar in some other way.



Take this
out!

Transformations



(38) If S is a sentence of the form NP₁-Auxiliary-V-NP₂, then the corresponding string of the form NP₂-Auxiliary[^]be[^]en-V-by[^]NP₁ is also a sentence.

NP₁ aux V NP₂ ⇒ NP₂ aux be en V by NP₁

from
slide 1!

The advantages of this analysis of passives are unmistakable. Since the element be[^]en has been dropped from (32) it is no longer necessary to qualify (32) with the complex of restrictions discussed above.



A Tradeoff

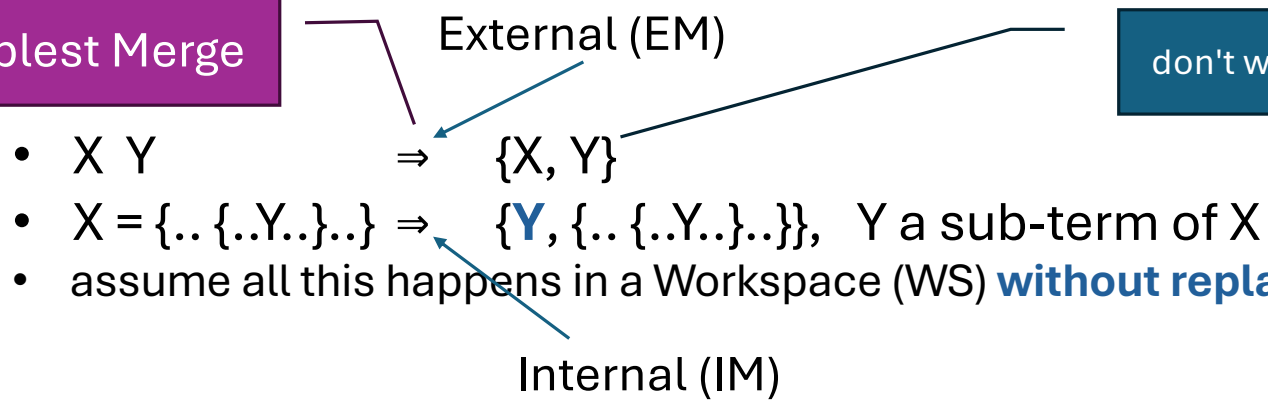
- Simple mechanisms
 - Complex descriptions of language phenomena
- Complex mechanisms
 - Simpler descriptions, e.g. using Transformations
- *1950s computational efficiency not yet in the picture ...*
 - computer stack (BURY/UNBURY), Turing (1945)
 - Cocke–Younger–Kasami (CKY) algorithm, Sakai (1961)
 - LR(k) parsing, (D)PDA discovery, Knuth (1965)
 - Earley algorithm, Earley (1968)



Strong Minimalist Thesis (SMT)

- SMT says
 - simplicity of mechanism is needed (*evolutionary plausibility*)
 - computational efficiency is needed (*slow wetware*)
 - simplicity of description is possible (Einstein's *Miracle Creed*)
- What is that simple mechanism?
 - Ask *what's the simplest (formal) device that permits phrases?*

Simplest Merge



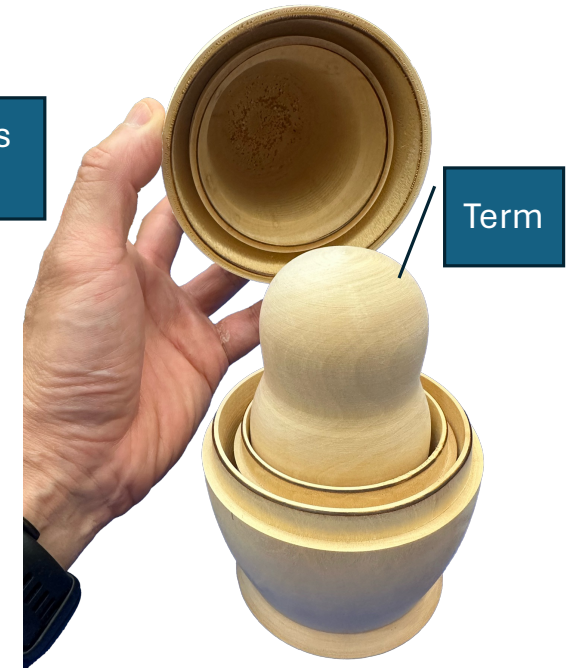
don't worry: we don't use set theory

o/w would circumvent c-command

Operative Complexity

3rd Factor: all operations on structure obey this

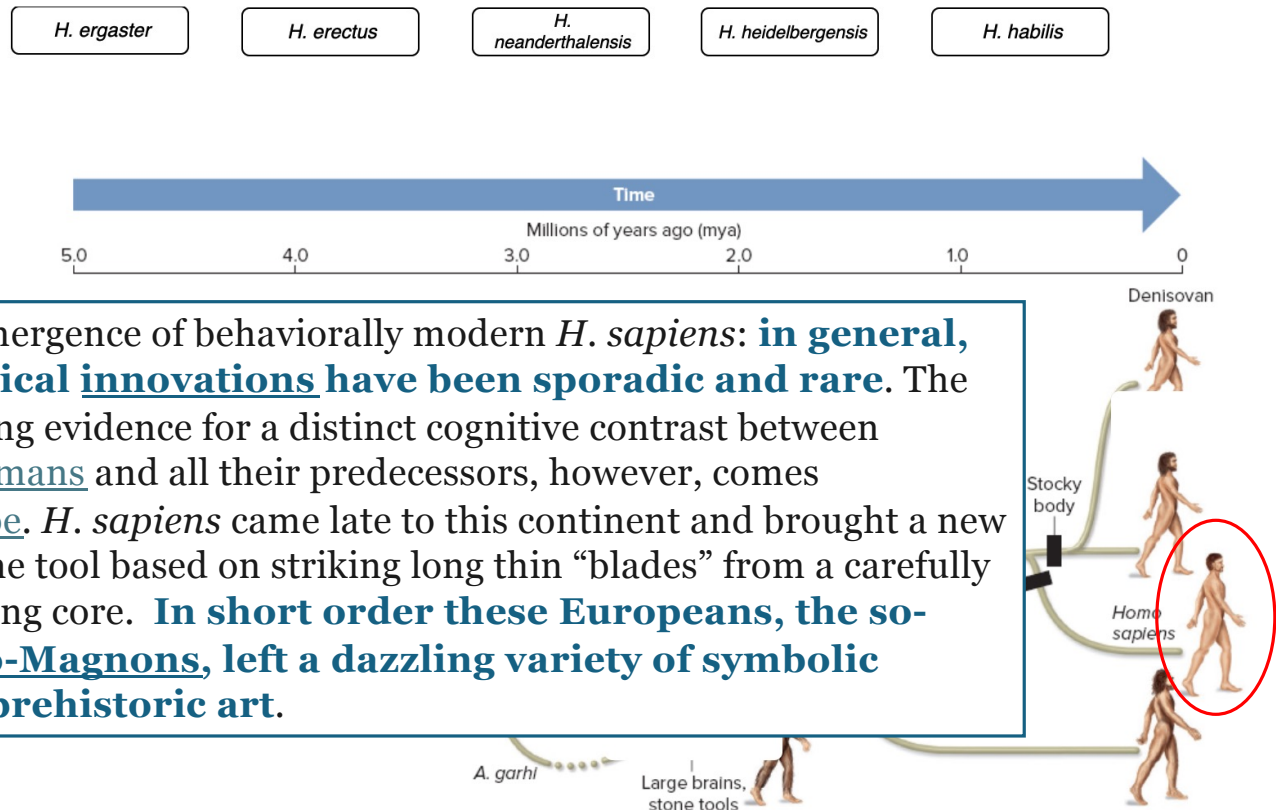
- Adopt simplest **recursive** formal device
 - i.e. *Merge feeds Merge* in the Workspace (WS)
- (Minimal) **Search**:
 - *look in the WS for something, 1st thing you find, **have to stop***
- (Chomsky p.c.):
 - **We assume that Merge like other operations observes it.**
 - **That's why only members of WS, not their terms, are eligible for EM.**
- Markovian assumption:
 - e.g. no memory, no (access to) WS history
 - no elaboration permitted: no 3 objects, no parallel/no sideways/no splicing etc.
- **Ask**: is Simplest Merge efficient?
 - has nasty combinatorics
- **Answer**: Merge has Language (organ) Specific Constraints (LSCs)
 - **Language Merge** is pre-wired for Theta Theory (*predicates and arguments*)



Return to this important idea

Strong Minimalist Thesis (SMT)

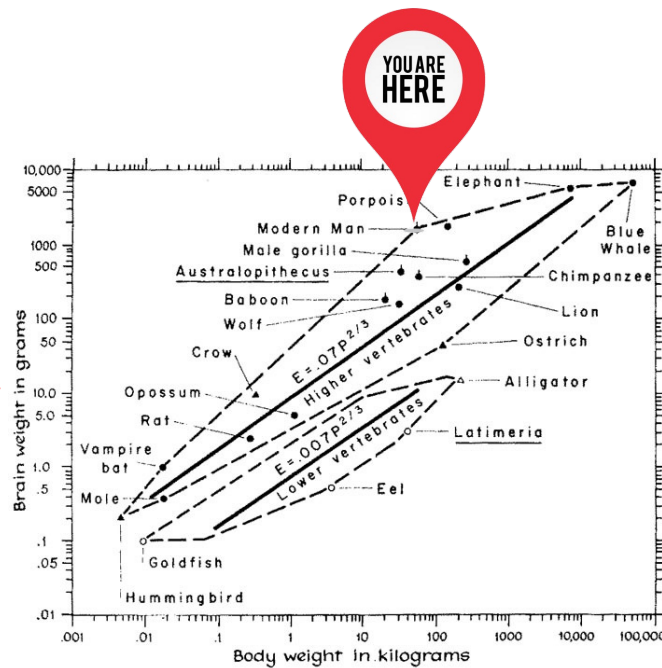
Tell brain and ev
story ...



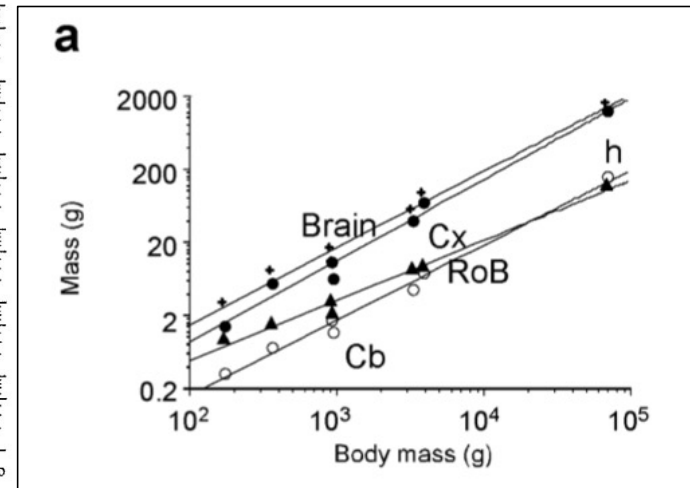
We special? Allometric scaling

Brain: 86 billion neurons

- perinatal neuron cell death: lose 50%
- development: heavy synaptic pruning
- we lose 50K neurons every day



- Primate scaling: *uniquely human? nah* (Azevedo et al. 2009)



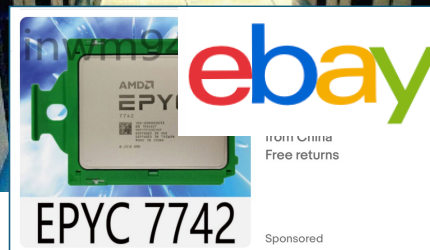
Yes, we special

- But it's not in the raw hardware, i.e. just adding more neurons
 - neuroanatomical differences: humans vs. nonhuman primates exist, e.g. *Broca's area*
 - for example, supercomputer below is just a scaled-up (AMD CPU) PC
 - recently upgraded in speed by 20% (Aug 2023)

National Weather Service: NOAA supercomputers *Dogwood* (VA) and *Cactus*



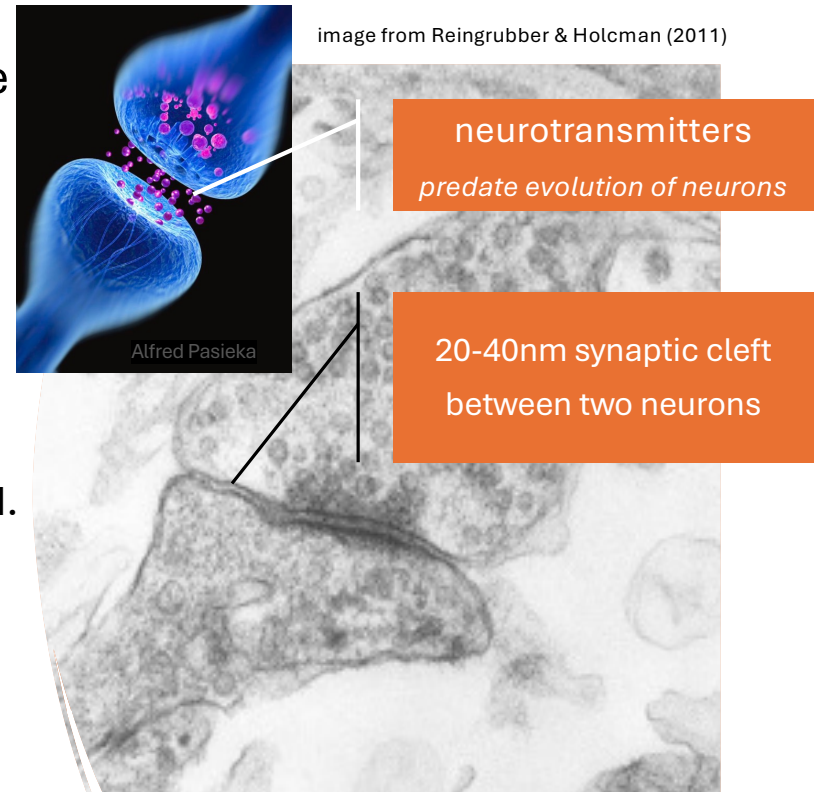
*You all noticed the
20% better
weather forecasts,
right? :-)*



Fact: Brain is slow, efficiency is important

Computational efficiency (and **bandwidth**) are important considerations for all **organic systems**:

- our sensory apparatus can generate vast amounts of data (*sensor mismatch*)
- a slow (*chemical*) brain limits what can be analyzed
- ***The War of Soups and Sparks*** (Valenstein, 2005) 19th century belief that neurons were electrically connected. *Neurophysiologists believed only electrical transmission is fast enough to activate skeletal muscles.* Mid-20th century: brain is chemical.
- neuron communication uses 50% of energy
- we (selectively) throw out/ignore almost all of the signal

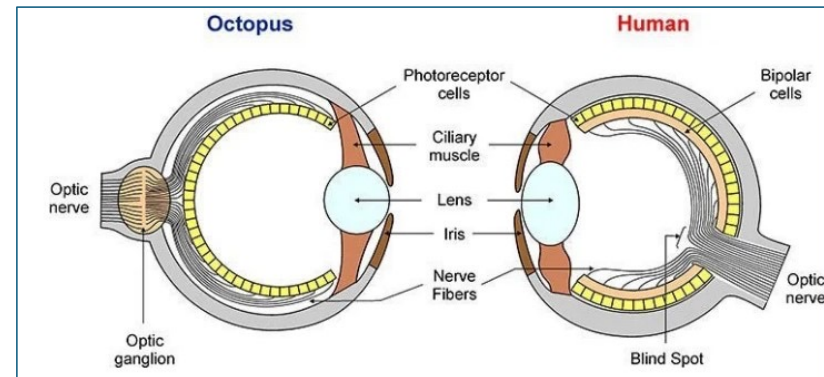


Evolution is slow

Land & Fernald (1992), *Animal Eyes* Land & Nilsson (2012)

- From the first opsin to high-resolution vision took about **170 million years** and was largely completed by the onset of the Cambrian, about 530 mya.
 - stage 1: receptors (evolved 40-65 times)
 - stage 2: optics (10 different systems)
- Most of the types of eye that we recognize today arose in a brief period during the Cambrian, about 530 million years ago.
- First brain cells (700 mya),
- First nervous system (500 mya, Cambrian). Jellyfish: eyes but no brain.
- First human-like brain (3-4 mya)
- Modern brain (1-0.2 mya)

- SMT optimal solution: *Nature adapts/optimizes what it has to work with*



- "camera eye" (cf. compound eye)
- octopus: color-blind, but can re-generate eyes
- we lost superior tetrachromatic vision 100 mya

Einstein's Miracle Creed

from McDonough (2022)

Time and again the passion for understanding has led to the illusion that man is able to comprehend the objective world rationally, by pure thought, without any empirical foundations—in short, by metaphysics. I believe that every true theorist is a kind of tamed metaphysicist. . . . The metaphysicist believes that the logically simple is also the real. The tamed metaphysicist believes that not all that is logical is experienced reality, but that the totality of reality can be “comprehended” on the basis of simple concepts on premises of great simplicity. This is the “miracle creed.” Admittedly so, but it is a miracle creed which has been borne out to an amazing extent by the development of science. (Einstein 1950, 342)

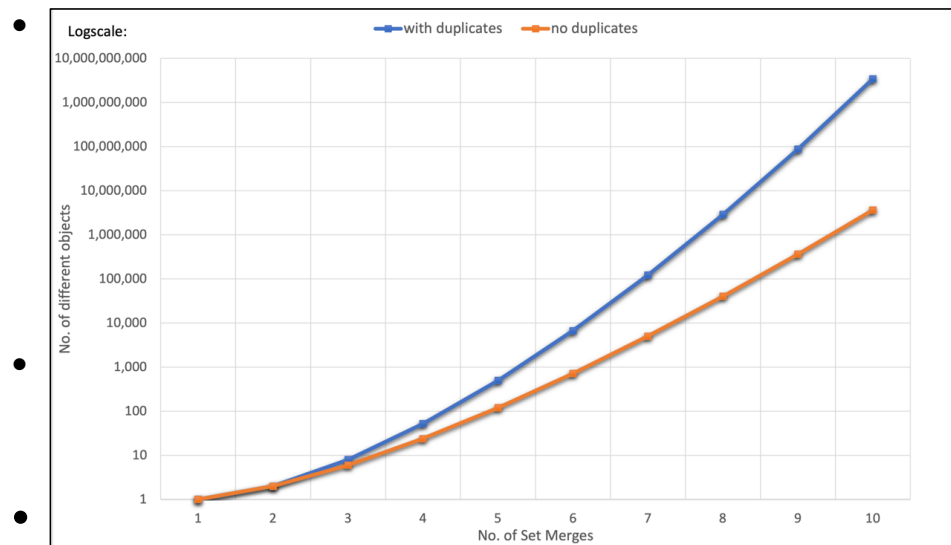
LLMs: "largest models typically have 100 billion parameters" GPT-4 1,760 billion (Wikipedia)

Merge: LSC

- Recall there are simpler things than Simplest Merge
 - e.g. linear concatenation (Chomsky 1956 Model #1)
 - e.g. **non-recursive** (so-called) Core-Merge, Fujita (2014) ← *same as concat-2?*
 - e.g. Conjoin, Progovac (2015)
 - e.g. recursive Internal Merge (IM) only (*as it's more efficient*)
 - **NO**: *can't work on its own, it can't get started*
 - e.g. recursive External Merge (EM) only
 - hmm, can build θ -configurations now, but all languages exhibit displacement
- *Posit them as earlier stages of evolution?*
 - *only necessary if you believe Merge cannot emerge in "one step"*

Merge: LSC

- Merge has nasty combinatorics



Merge) vs. EM (External Merge).
determines how to interpret identical

**can't observe theta theory (and duality
preferred derivation.**

operations (SBO) are **thought-related**,

- (External) Merge builds θ -configurations efficiently
- *i.e as early and quickly as possible*

Strong Minimalist Thesis (SMT)

- This is not a technical talk (*unfortunately*)
 - there are ways to **computationally tame** Merge (under SMT)
 - *i.e. without requiring the invention of new mechanisms/algorithms*
 - LSC: θ -theory
 - *arguments and predicates that seek arguments*
 - **Operative complexity**: Language Merge hugely better than Merge
 - **Phases** (*aka* WS partitioning)
 - for internal thought generation, NOT for parsing (*that's sorta okay*)
 - **Workspace Balancing**
 - wrt. θ -seekers and θ -relevant WS items (the problem of *unpronounced items*)



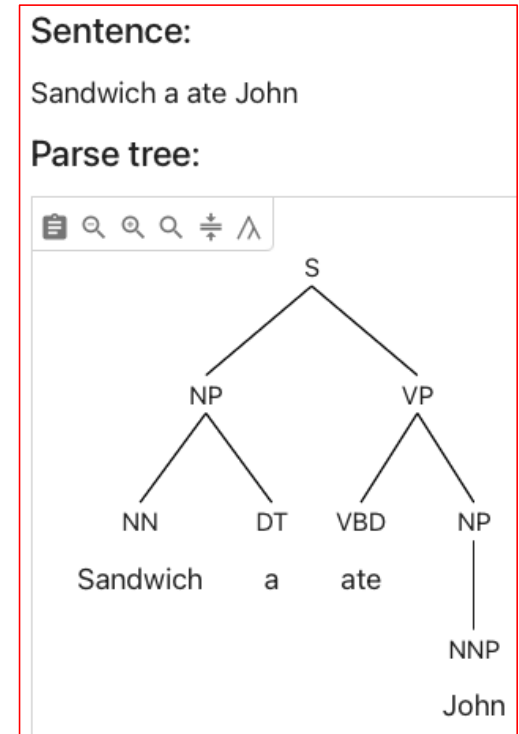
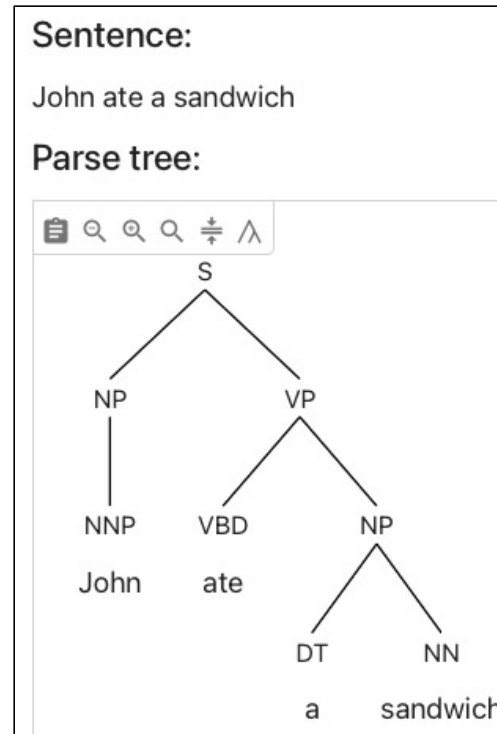
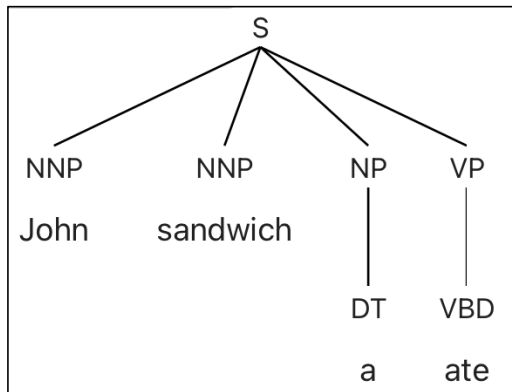
Berkeley Neural Parser

parser.kitaev.io

Grammaticality

- *John ate a sandwich /*
**Sandwich a ate John*

(pg.113, Chomsky, 1956)

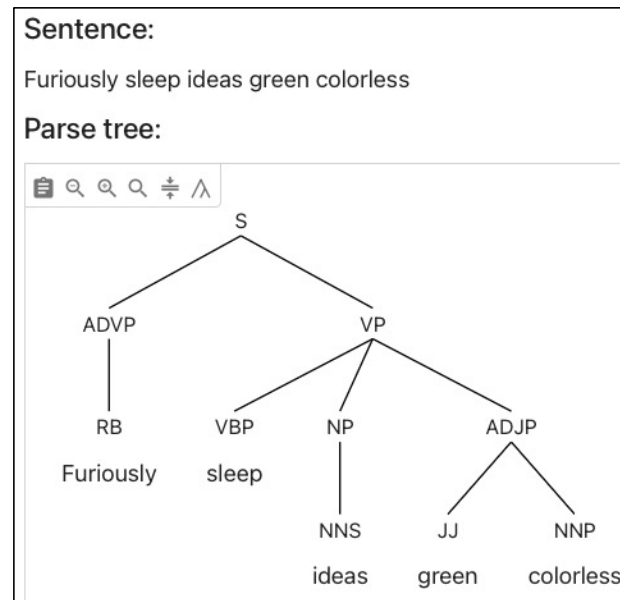
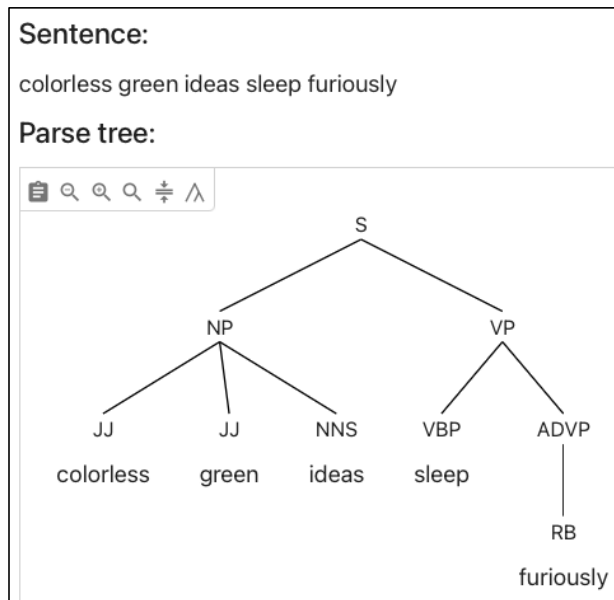




Berkeley Neural Parser

parser.kitaev.io

Colorless green ideas sleep furiously / **Furiously sleep ideas green colorless*
(pg.116, Chomsky, 1956)



Statistical systems
can accept nearly
anything, including
unnatural languages.
*Is that a feature or
a bug?*

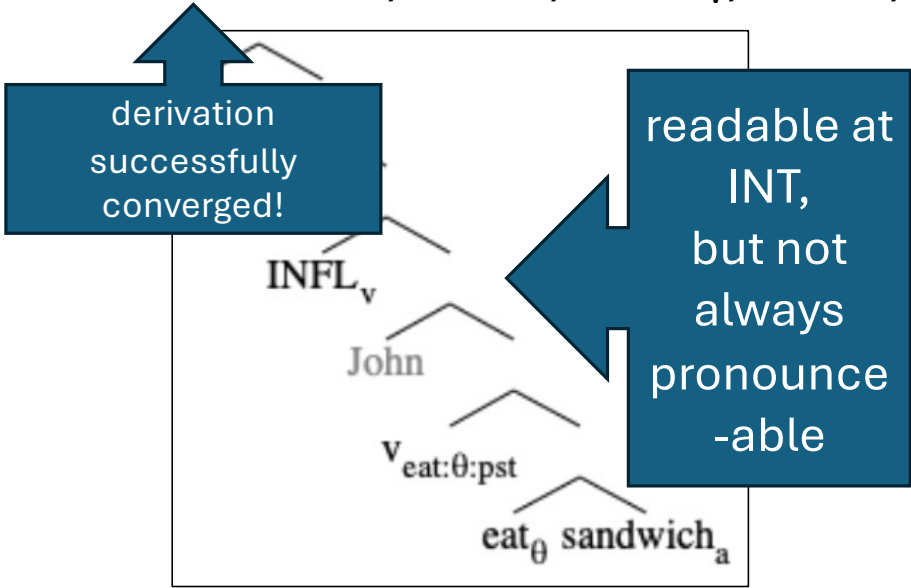
SMT Parser

Words: John ate a sandwich

Initial WS 1: sandwich_a eat_θ v_{eat:θ:pst} INFL_v John

ONE OBJECT: {C, {John, {INFL_v, {John, {v_{eat:θ:pst}, {eat_θ, sandwich_a}}}}}}

a sandwich = sandwich_a



Words: sandwich a ate John

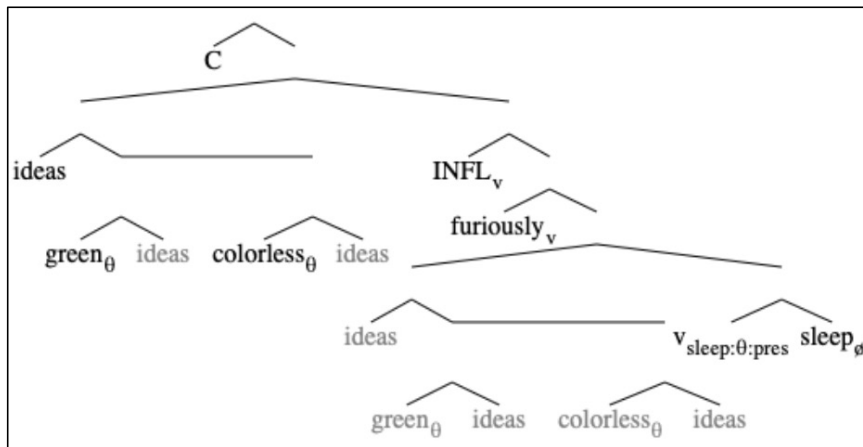
- ▼ Initial WS 1: John_a eat_θ v_{eat:θ:pst} INFL_v sandwich
- ▶ WS 1: {eat_θ, John_a} v_{eat:θ:pst} INFL_v sandwich
- INT/EXT: $\hat{\wedge}$ {C, {sandwich, {INFL_v, {sandwich, {v_{eat:θ:pst}, {eat_θ, John_a}}}}}}
- ▶ Initial Spellout: sandwich 3sg pst eat a John
- Error:** inconsistent spellout
- ▶ WS 1: {eat_θ, John_a} v_{eat:θ:pst} INFL_v sandwich
- ▶ WS 1: {eat_θ, sandwich} John_a v_{eat:θ:pst} INFL_v
- INT/EXT: $\hat{\wedge}$ {C, {John_a, {INFL_v, {John_a, {v_{eat:θ:pst}, {eat_θ, sandwich}}}}}}
- ▶ Initial Spellout: a John 3sg pst eat sandwich
- Error:** inconsistent spellout
- ▶ WS 1: {eat_θ, sandwich} John_a v_{eat:θ:pst} INFL_v
- ▶ WS 1: {v_{eat:θ:pst}, eat_θ} John_a INFL_v sandwich
- ▶ WS 1: {eat_θ, v_{eat:θ:pst}} John_a INFL_v sandwich

SMT Parser

Words: colorless green ideas sleep furiously

Initial WS 1: furiously_v sleep_∅ v_{sleep:∅:pres} INFL_v ideas green_∅ colorless_∅
{C, {{ideas, {{green_∅, ideas}, {colorless_∅, ideas}}}, {INFL_v, {furiously_v,
{{ideas, {{green_∅, ideas}, {colorless_∅, ideas}}}}, {v_{sleep:∅:pres}, sleep_∅}}}}}}

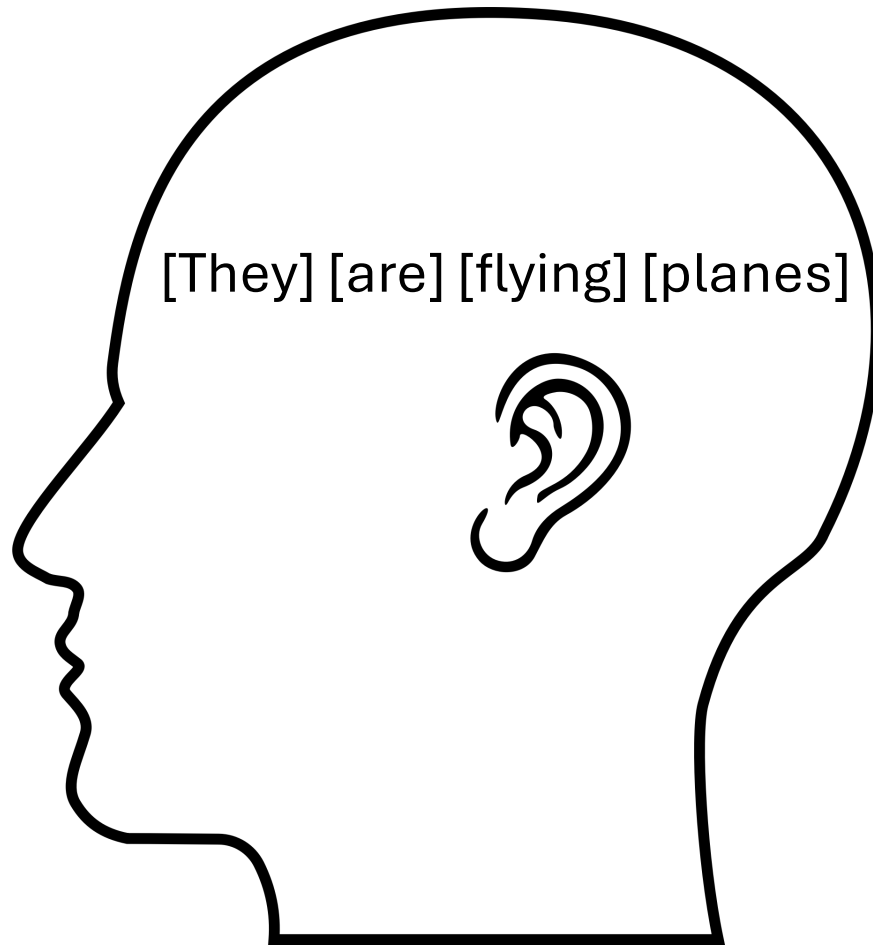
Initial WS 2: furiously_v sleep_∅ v_{sleep:∅} INFL_v ideas green_∅ colorless_∅



Technical aside

Uses FormSet as well as Merge
e.g. *John arrived and met Bill*

Chomsky GK

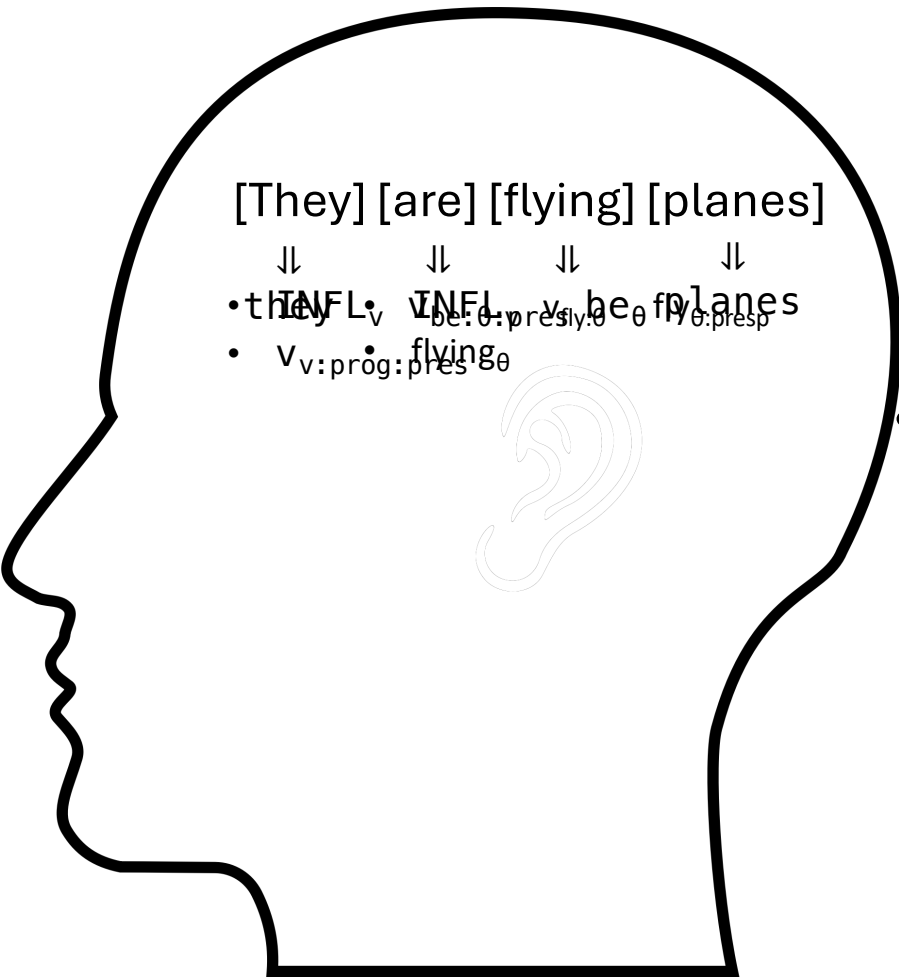


How it works

[pg.118, Chomsky (1956)]

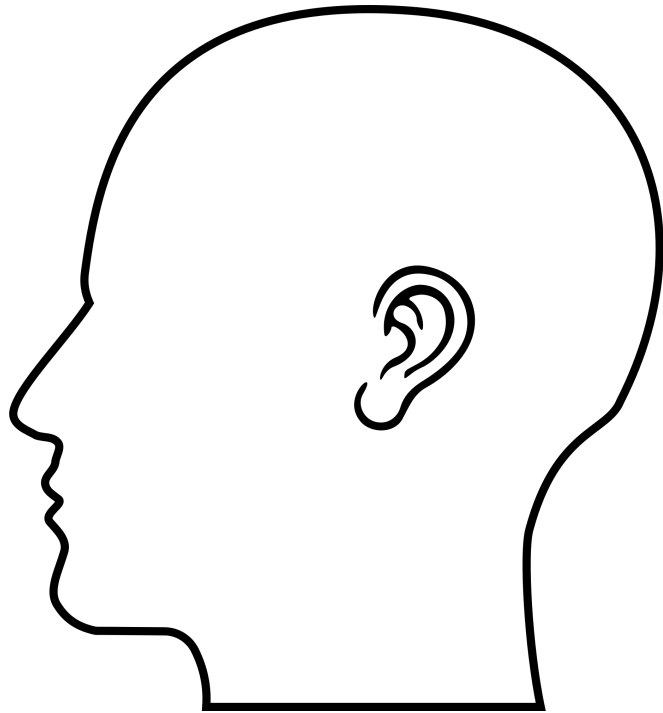
this sentence will have two phrase structures assigned to it; it can be analyzed as "they - are - flying planes" or "they - are flying - planes." And in fact, this sentence is ambiguous in just this way; we can understand it as meaning that "those specks on the horizon - are - flying planes" or "those pilots - are flying - planes."

- they - are - flying planes
- they - are flying - planes



How it works

- Parsing:
 - recognize a word from the input signal
 - look it up in LEX
 - heads go in an Initial Workspace (WS)
 - Merge fires!



[They] [are] [flying] [planes]

How it works

Workspaces (WS):

- planes fly_{θ:presp} V_{fly:θ} INFL_v V_{v:prog:pres} they
- planes flying_θ be_θ V_{be:θ:pres} INFL_v they
- could be more ...

[They] [are] [flying] [planes]

• Derivation:

1. planes fly_{θ:presp} v_{fly:θ} INFL_v v_{v:prog:pres} they
2. {fly_{θ:presp}, planes} v_{fly:θ} INFL_v v_{v:prog:pres} they
3. {v_{fly:θ}, {fly_{θ:presp}, planes}} INFL_v v_{v:prog:pres} they
4. {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}} INFL_v v_{v:prog:pres}
5. {v_{v:prog:pres}, {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}}}} INFL_v
6. {INFL_v, {v_{v:prog:pres}, {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}}}}}}
7. {they, {INFL_v, {v_{v:prog:pres}, {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}}}}}}}}
8. {C, {they, {INFL_v, {v_{v:prog:pres}, {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}}}}}}}}
9. {C, {they, {INFL_v, {v_{v:prog:pres}, {they, {v_{fly:θ}, {fly_{θ:presp}, planes}}}}}}}}}}
10. they 3pl pres. be flying planes
11. they are flying planes

θ-configuration

Merge output:
converged

FormCopy

Linear
Spellout

[They] [are] [flying] [planes]

• Derivation:

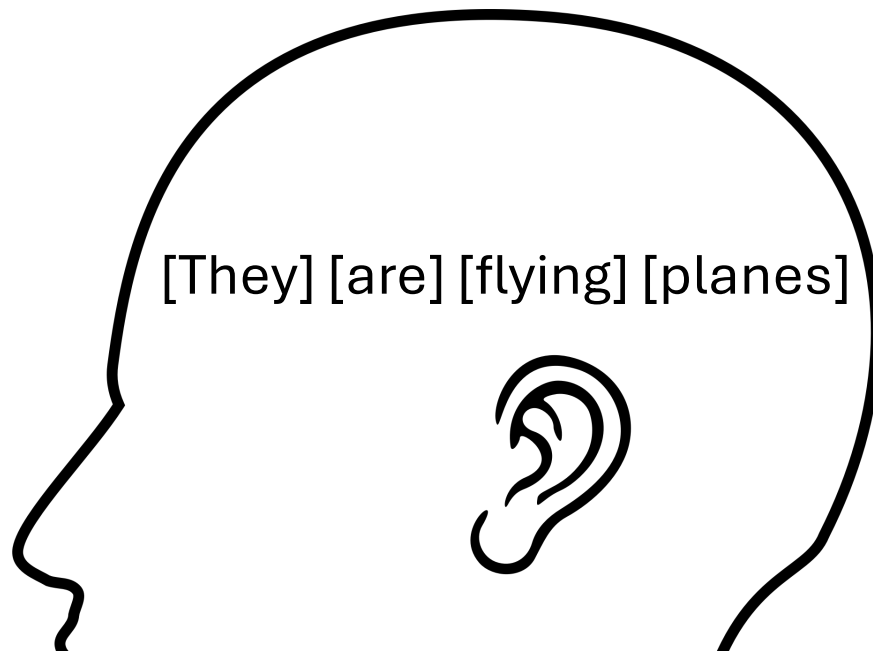
1. planes flying_θ be_θ v_{be:θ:pres} INFL_v they
2. {flying_θ, planes} be_θ v_{be:θ:pres} INFL_v they
3. {{flying_θ, planes}, planes} be_θ v_{be:θ:pres} INFL_v they
4. {be_θ, {{flying_θ, planes}, planes}} v_{be:θ:pres} INFL_v they
5. {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}} INFL_v they
6. {they, {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}} INFL_v
7. {INFL_v, {they, {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}}}}
8. {they, {INFL_v, {they, {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}}}}
9. {C, {they, {INFL_v, {they, {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}}}}
10. {C, {they, {INFL_v, {they, {v_{be:θ:pres}, {be_θ, {{flying_θ, planes}, planes}}}}}}

θ-configuration

Merge output:
converged

Linear
Spellout

they	3pl	pres.	be	flying	planes
they			are	flying	planes



How it works

[pg.118, Chomsky (1956)]

- they – are – flying planes
- they – are flying – planes

- | | | | | | | | |
|----|--|------|-----|----------|-----|--------|--------|
| 1. | {C, {they, {INFL _v , {v _{v:prog:pres} , {they, {v _{fly:θ} , {fly _{θ:presp} , planes}} ₇ | they | 3pl | pres. be | are | flying | planes |
| | they | | | | are | flying | planes |
| 2. | {C, {they, {INFL _v , {they, {v _{be:θ:pres} , {be _θ , {{flying _θ , planes}, planes}} ₈ | they | 3pl | pres. | be | flying | planes |
| | they | | | | are | flying | planes |

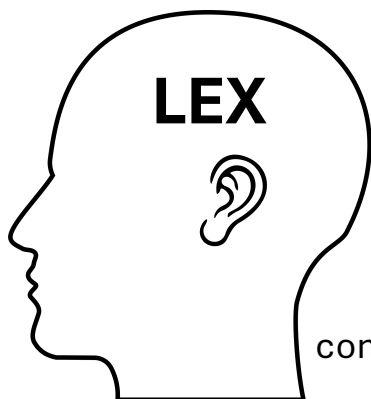
SMT Parser

sandway.arizona.edu/smtparser

Hand-built LEX

Words: they are flying planes

- ▶ **Initial WS 1:** planes fly0:pres vfly:0 INFL_v vpred:pres INFL_v they
- ▶ **Initial WS 2:** planes flying0 vpred:pres INFL_v they
- ▶ **Initial WS 3:** planes fly0:pres vfly:0 INFL_v be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 4:** planes flying0 be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 5:** planes fly0:pres vfly:0 INFL_v vv:prog:pres they
- ▶ **Initial WS 6:** planes flying0 vv:prog:pres they
- ▶ **Initial WS 7:** planes flying0 vv:pass:pres they



*how many
entries
come to
mind?*

context, experience

WordNet LEX (nltk)

Words: they are flying planes

- ▶ **Initial WS 1:** planes flying are they
- ▶ **Initial WS 2:** plane0 vplane:0:pres INFL_v:3sg flying are they
- ▶ **Initial WS 3:** planes fly0:pres vfly:0 INFL_v are they
- ▶ **Initial WS 4:** plane0 vplane:0:pres INFL_v:3sg fly0:pres vfly:0 INFL_v are they
- ▶ **Initial WS 5:** planes flying0 are they
- ▶ **Initial WS 6:** plane0 vplane:0:pres INFL_v:3sg flying0 are they
- ▶ **Initial WS 7:** planes flying vpred:pres INFL_v they
- ▶ **Initial WS 8:** plane0 vplane:0:pres INFL_v:3sg flying vpred:pres INFL_v they
- ▶ **Initial WS 9:** planes fly0:pres vfly:0 INFL_v vpred:pres INFL_v they
- ▶ **Initial WS 10:** plane0 vplane:0:pres INFL_v:3sg fly0:pres vfly:0 INFL_v vpred:pres INFL_v they
- ▶ **Initial WS 11:** planes flying0 vpred:pres INFL_v they
- ▶ **Initial WS 12:** plane0 vplane:0:pres INFL_v:3sg flying0 vpred:pres INFL_v they
- ▶ **Initial WS 13:** planes flying be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 14:** plane0 vplane:0:pres INFL_v:3sg flying be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 15:** planes fly0:pres vfly:0 INFL_v be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 16:** plane0 vplane:0:pres INFL_v:3sg fly0:pres vfly:0 INFL_v be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 17:** planes flying0 be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 18:** plane0 vplane:0:pres INFL_v:3sg flying0 be0 vbe:0:pres INFL_v they
- ▶ **Initial WS 19:** planes flying vv:prog:pres they
- ▶ **Initial WS 20:** planes fly0:pres vfly:0 INFL_v vv:prog:pres they
- ▶ **Initial WS 21:** plane0 vplane:0:pres INFL_v:3sg fly0:pres vfly:0 INFL_v vv:prog:pres they
- ▶ **Initial WS 22:** planes flying0 vv:prog:pres they
- ▶ **Initial WS 23:** planes flying vv:pass:pres they
- ▶ **Initial WS 24:** planes flying0 vv:pass:pres they

Jokes: multiple parses

- Many jokes are based on the human parser reflexively computing more than one parse
- Examples:
 - **As I handed my dad his 50th birthday card, he looked at me with tears in his eyes and said,**
 - **"You know, one would've been enough."**

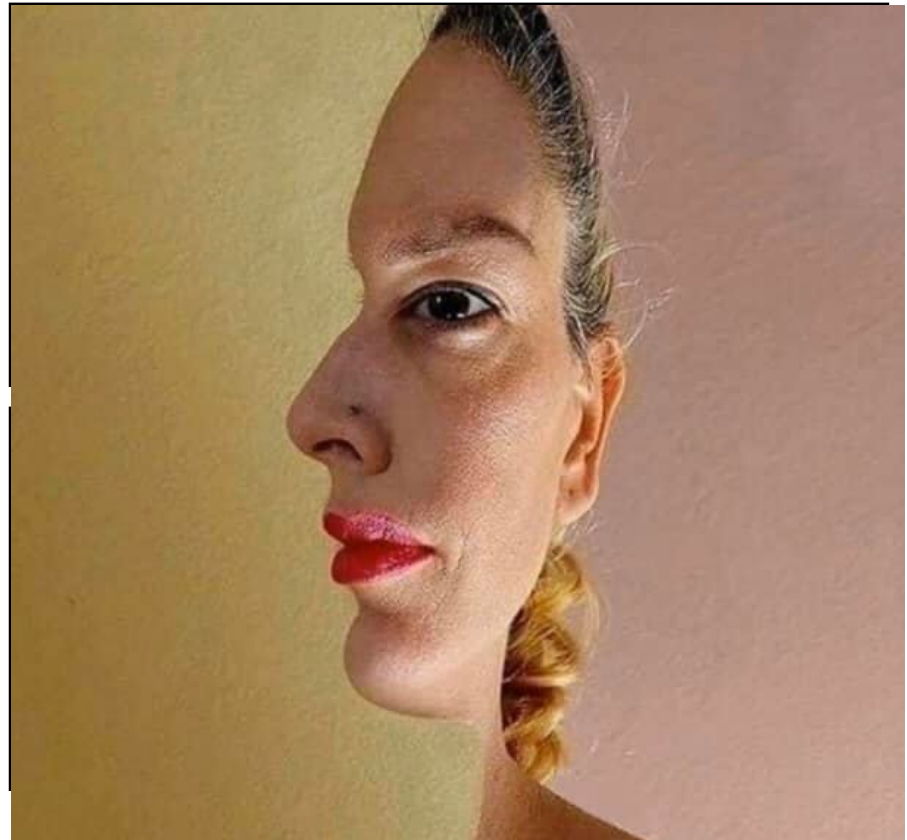
- *on a bicycle*





Pam: Is it okay to eat hot dogs with hands?
Sam: No. Hot dogs don't have hands!

Note: the visual system can also exhibit parsing ambiguity



SA

You

Is it okay to eat hot dogs with hands? No. Hot dogs don't have hands! What's funny about this?

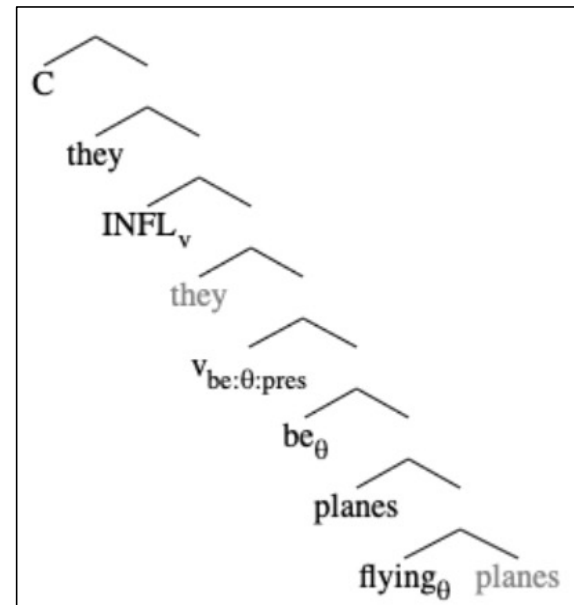
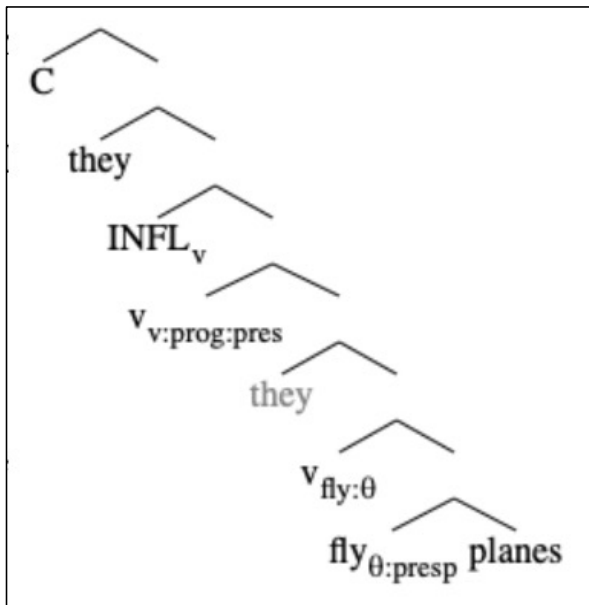


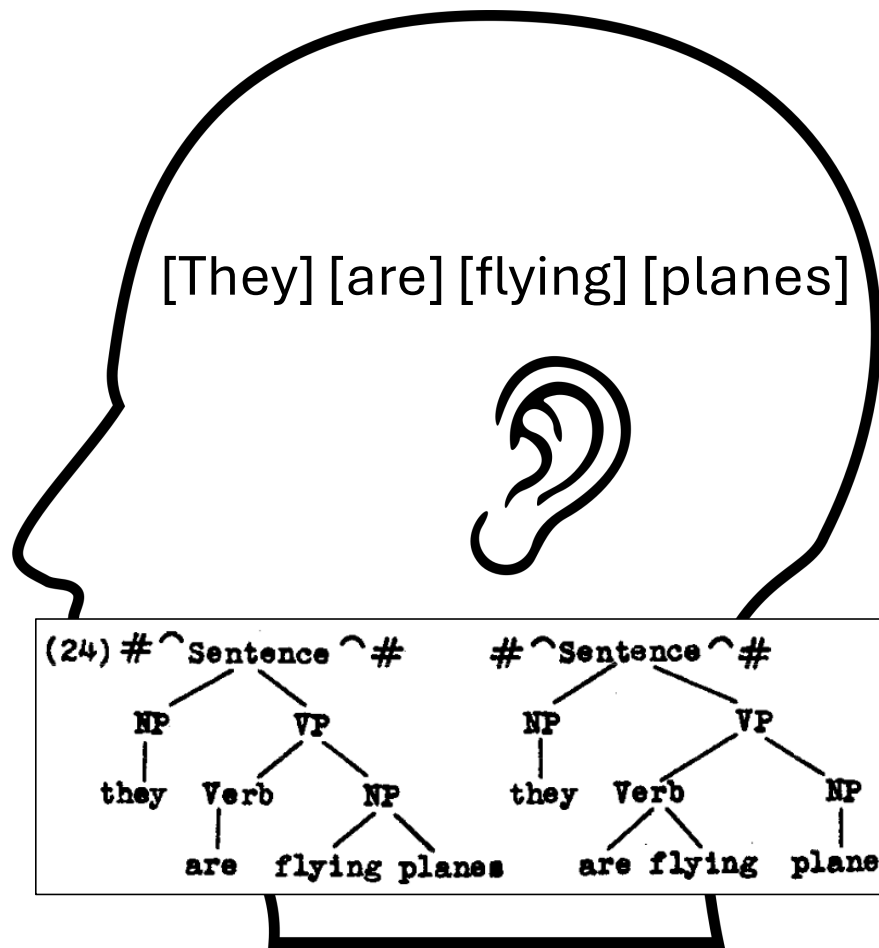
ChatGPT

The humor in the response "No. Hot dogs don't have hands!" comes from the unexpected twist in the interpretation of the question. Instead of addressing whether it's socially acceptable for humans to eat hot dogs with their hands, the response takes the question literally and humorously points out that hot dogs themselves don't possess hands to eat with. This unexpected and literal interpretation creates a humorous moment by subverting the original question in a clever and playful way.

SMT Parser

- {C, {they, {INFL_v, {V_{v:prog:pres}, {they, {V_{fly:θ}, {fly_{θ:presp}, planes}}}}}}}
- {C, {they, {INFL_v, {they, {V_{be:θ:pres}, {be_θ, {planes, {flying_θ, planes}}}}}}}}}





How it works

[pg.118, Chomsky (1956)]

- they – are – flying planes
- they – are flying – planes

(20) $\Sigma : \# \hat{\text{Sentence}} \#$
 $F: \text{Sentence} \rightarrow \text{NP} \hat{\text{VP}}$
 $\text{VP} \rightarrow \text{Verb} \hat{\text{NP}}$

(23) $\text{Verb} \rightarrow \text{are} \hat{\text{flying}}$
 $\text{Verb} \rightarrow \text{are}$
 $\text{NP} \rightarrow \text{they}$
 $\text{NP} \rightarrow \text{planes}$
 $\text{NP} \rightarrow \text{flying} \hat{\text{planes}}$



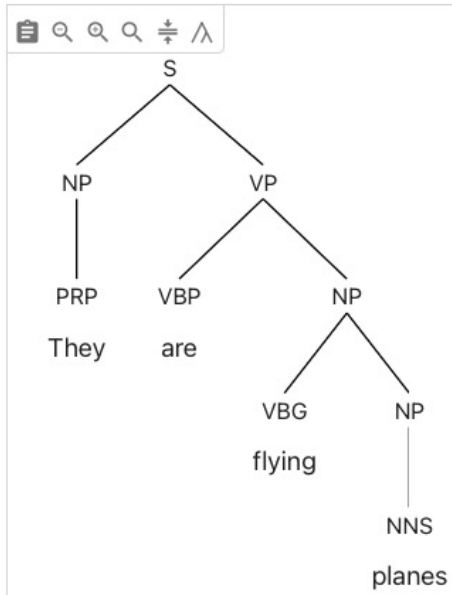
Berkeley Neural Parser

parser.kitaev.io

Sentence:

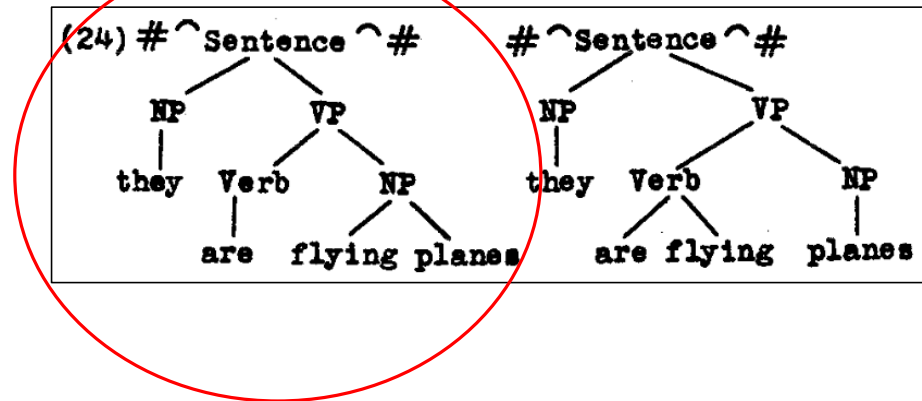
They are flying planes

Parse tree:



They are flying planes

(pg.118, Chomsky, 1956)





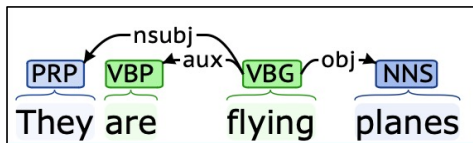
CoreNLP

corenlp.run
stanza.run

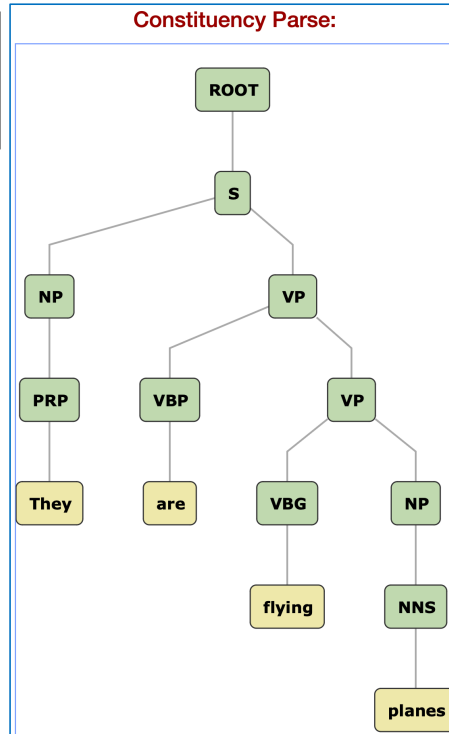


Stanza 1.6.0

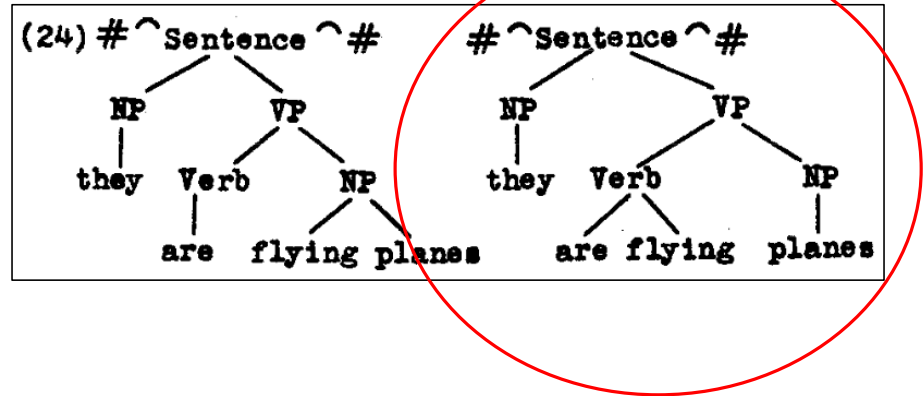
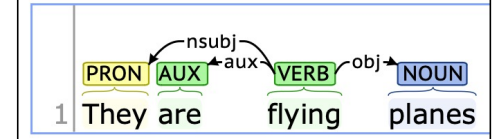
Enhanced++ Dependencies:



Constituency Parse:



Universal Dependencies:



More Structural Ambiguity

It's generally hard to avoid computing structural descriptions

PP attachment:

- *John saw the boy with a telescope*

(generally, *prepositions are hard for computer parsing!*)

Location-PP or not?

He decided on the boat

[pg.101 Chomsky *Aspects* (1965)]

- "he chose the boat"
- "he made his decision while on the boat"
- *He decided on the boat on the train*
- *On the train, he decided on the boat*

He decided on the boat



ChatGPT

"He decided on the boat" suggests someone made a choice while on a boat, or perhaps chose the boat itself for a specific purpose or journey.

Basic Property (BP) of Language

- simplest computational rule: *pick nearest (appropriate) word*

The simplest operation is certainly within the cognitive repertoire.
A child has no problem picking the first bead on a string.

- **BP: no**, simplest rule available:
 - *build structure, then determine nearest*
 - observed by children (as early as 30 months), not learned

- **Number Agreement:**

a. the bombing_{sg} of the cities_{pl} **was**_{sg} criminal

[pg.9, Chomsky GK (2021)]

b. * **were**_{pl}

c. the bombings_{pl} of the city_{sg} **were**_{pl} criminal

d. * **was**_{sg}

ChatGPT

ChatGPT 4o mini ▾

The bombing of the cities



The bombing of cities has been a significant and tragic aspect of warfare throughout history, especially during the 20th century. This tactic involves using aerial or missile bombardment to target cities, typically aiming to destroy infrastructure, disrupt civilian life, demoralize the population, and weaken an enemy's capacity for war. Here are some notable instances:

the bombing of the cities have



It seems like you're referring to a historical event involving the bombing of cities. Could you

Basic Property (BP) of Language

[pg.9, Chomsky GK (2021)]

- "adverb *carefully* seeks a verb [to modify], but it cannot use the simplest computation: pick the linearly closest verb."

- **Construal:**

- [...] *marks linearly closest verb*

- the mechanic who *fixed* the car *carefully* [*packed*] his tools

← [pack] or fix

- *Carefully*, the mechanic who [*fixed*] the car *packed* his tools

← pack

- the mechanic who *fixed* the car [*packed*] his tools *carefully*

← [pack]

- the mechanic who *carefully* [*fixed*] the car *packed* his tools

← [fix]

ChatGPT

the mechanic who fixed the car carefully packed his tools. What did he do carefully?

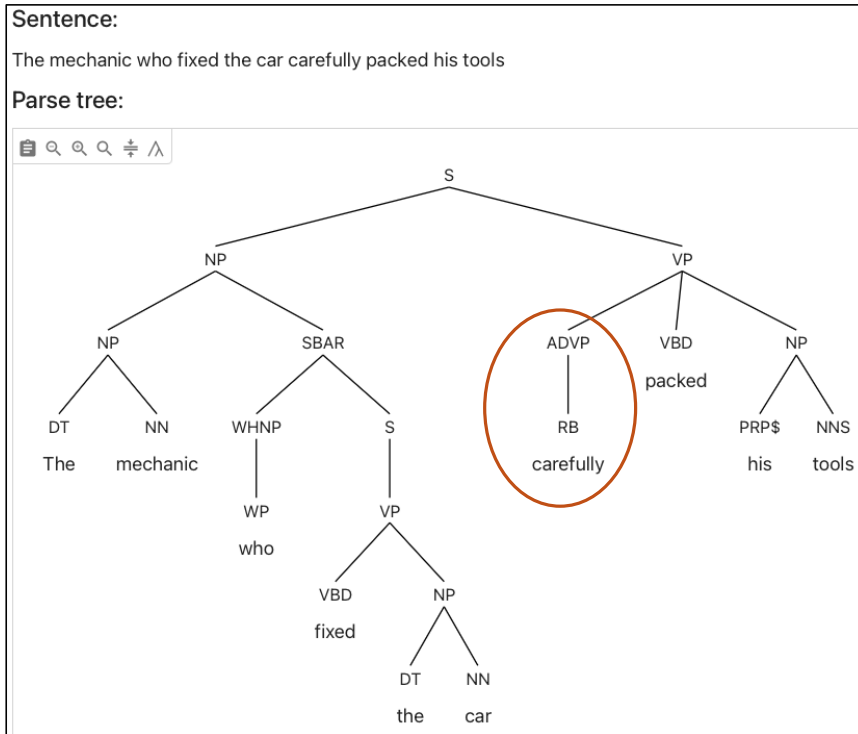


The mechanic carefully packed his tools. The action he did carefully is "packed."

Did he carefully fix the car?



Based on the sentence provided, it only states that the mechanic **carefully packed his tools**. It does not specify how he fixed the car, so we cannot conclude that he fixed the car carefully from this information alone.

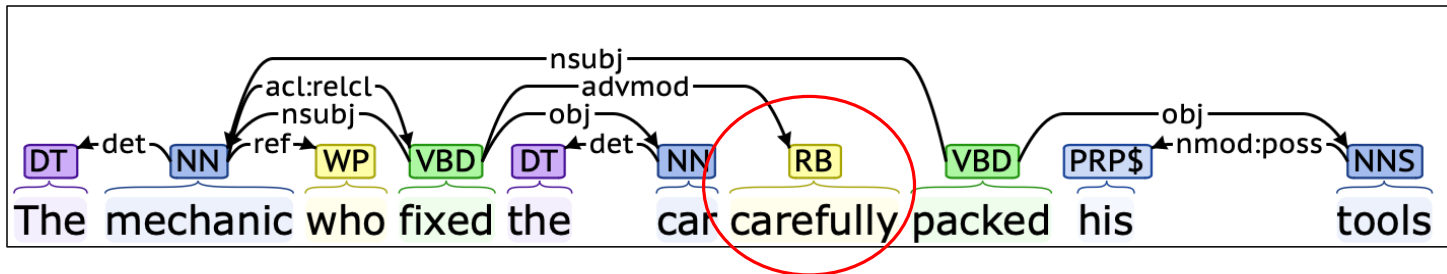


*The mechanic who fixed the car
carefully packed his tools*

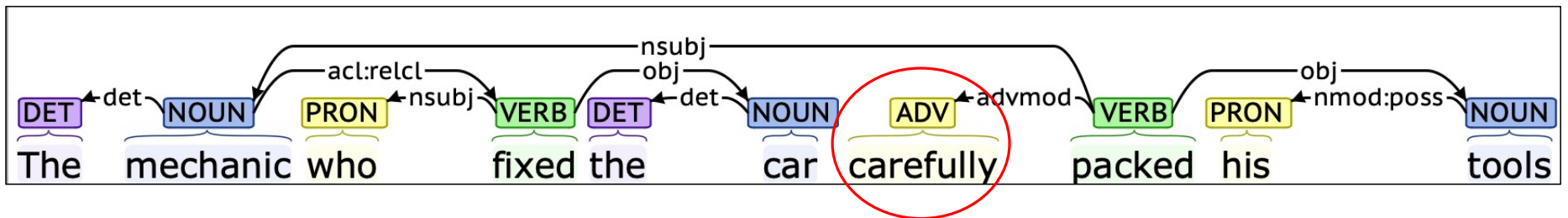
[Chomsky GK (2021), pg.9]



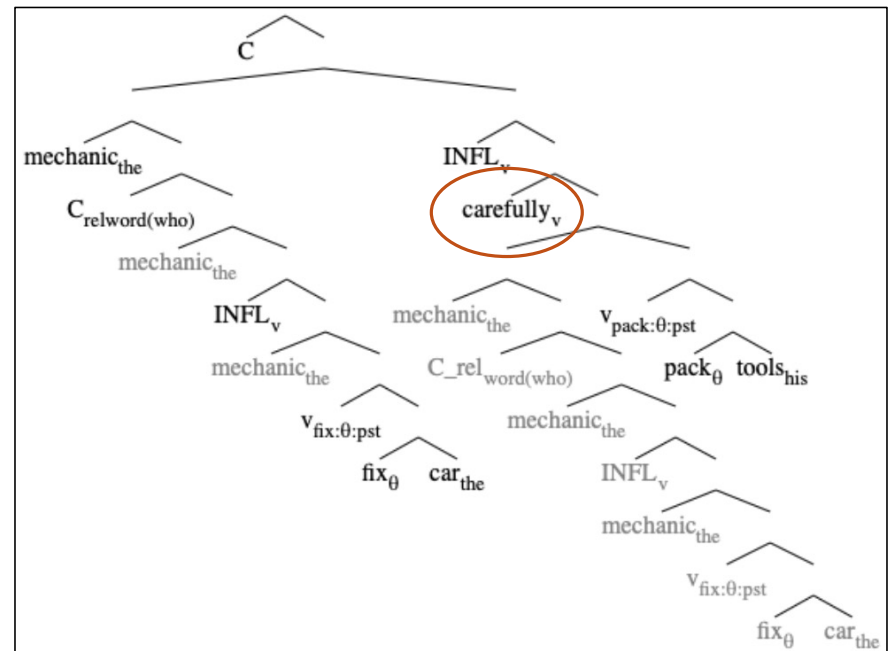
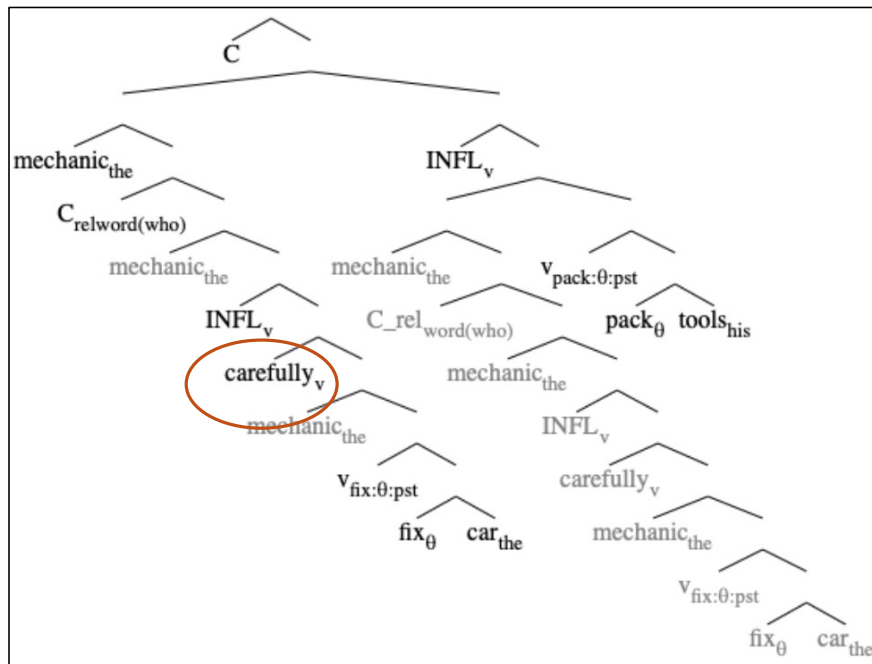
- `corenlp.run`



- `stanza.run`



SMT Parser



SMT Parser

One question: *two parses from one initial WS or two?*

Answer: just one.

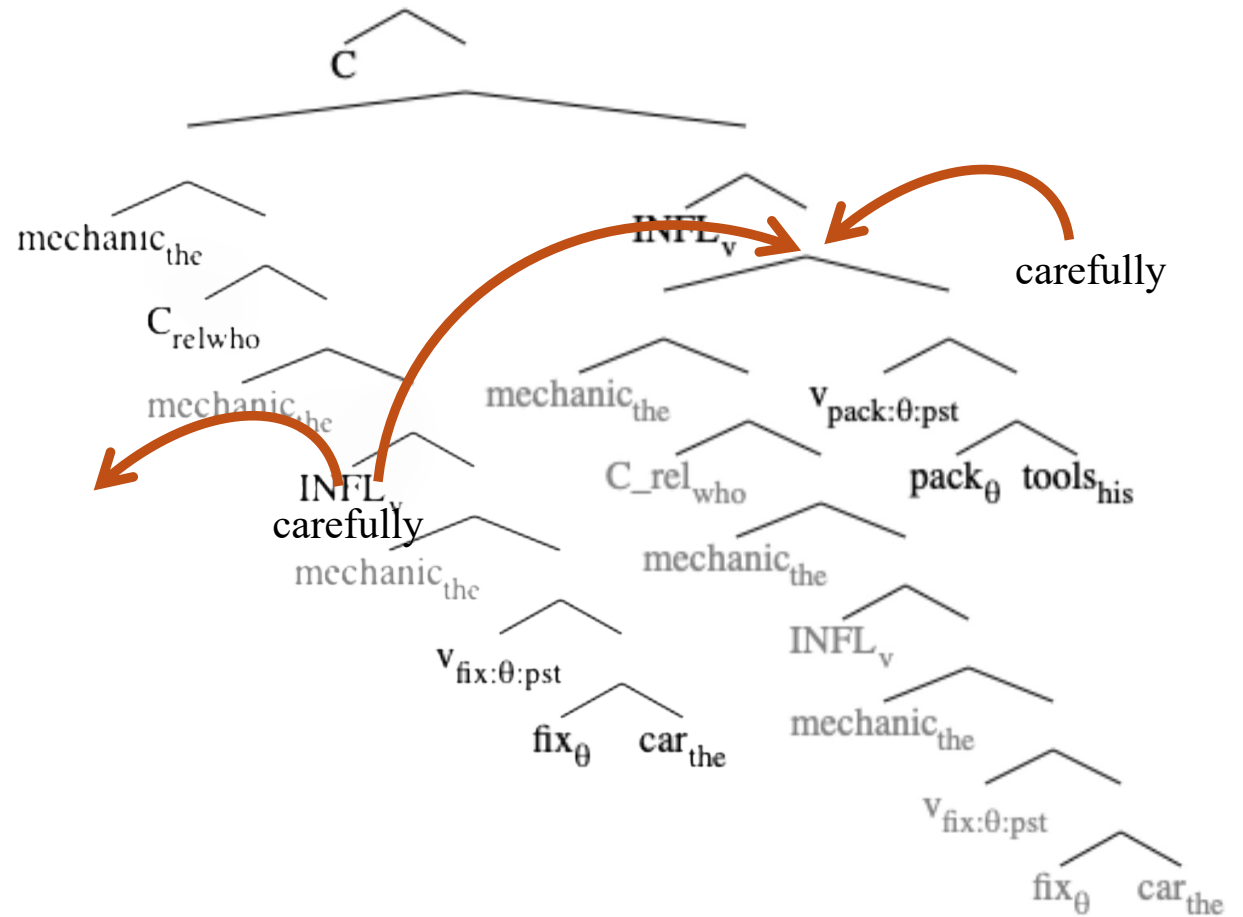
WS: tools_{his} pack_θ v_{pack:θ:pst} INFL_v carefully_v car_{the} fix_θ v_{fix:θ:pst} INFL_v
C_{relword(who)} mechanic_{the}

Parse:

1. {C, {{mechanic_{the}, {C_{relword(who)}, {mechanic_{the}, {INFL_v, {carefully_v, {mechanic_{the},
{v_{fix:θ:pst}, {fix_θ, car_{the}}}}}}}}}}, {INFL_v,
{{mechanic_{the}, {C_{relword(who)}, {mechanic_{the}, {INFL_v, {carefully_v, {mechanic_{the}, {v_{fix:θ:pst}
, {fix_θ, car_{the}}}}}}}}}}, {v_{pack:θ:pst}, {pack_θ, tools_{his}}}}}}}
2. {C, {{mechanic_{the}, {C_{relword(who)}, {mechanic_{the}, {INFL_v, {mechanic_{the}, {v_{fix:θ:pst},
{fix_θ, car_{the}}}}}}}}}}, {INFL_v, {carefully_v,
{{mechanic_{the}, {C_{relword(who)}, {mechanic_{the}, {INFL_v, {mechanic_{the}, {v_{fix:θ:pst}, {fix_θ, car_{th}
e}}}}}}}}}, {v_{pack:θ:pst}, {pack_θ, tools_{his}}}}}}}}}

SMT Parser

- Why?
- **Answer:** on the edge between the two chunks



SMT Parser

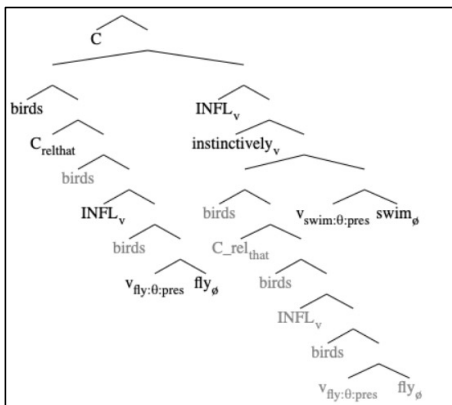
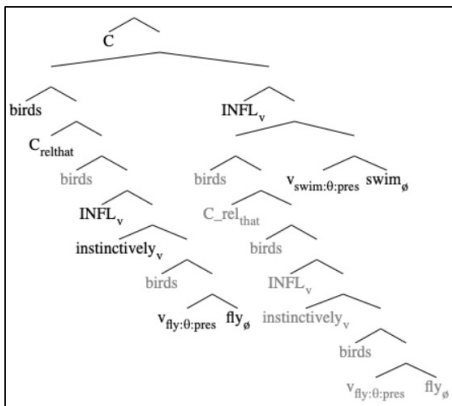
[pgs.8,103,117 Berwick & Chomsky (2016)]

- Similarly ambiguous sentences:

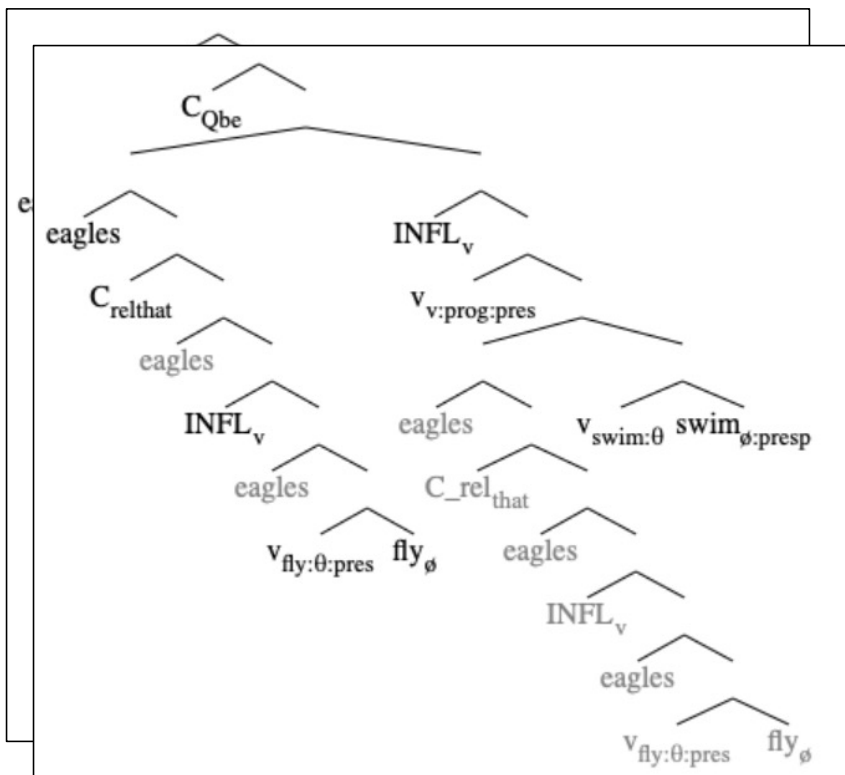
- *Birds that fly **instinctively** swim*
- *The desire to fly **instinctively** appeals to children*

- and unambiguous counterparts:

- ***Instinctively**, birds that fly swim*
- ***Instinctively**, the desire to fly appeals to children*



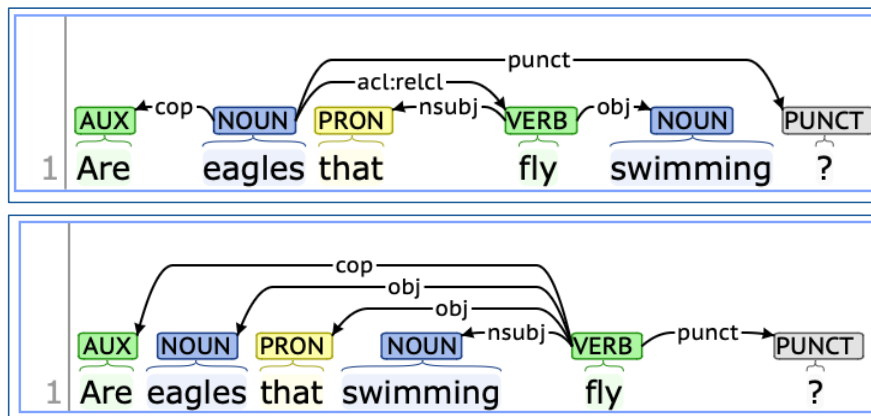
SMT Parser



[pg.39, Chomsky POP (2013)]

- *Can* eagles that fly swim?
 - "the question is about ability to swim, not to fly."
- *Are* eagles that fly swimming?
- **Are* eagles that swimming fly?
 - "... does not ask whether **it is the case that eagles that are swimming fly**. ... that is a fine thought, but it cannot be expressed by [this sentence]."

Stanza 1.6.0

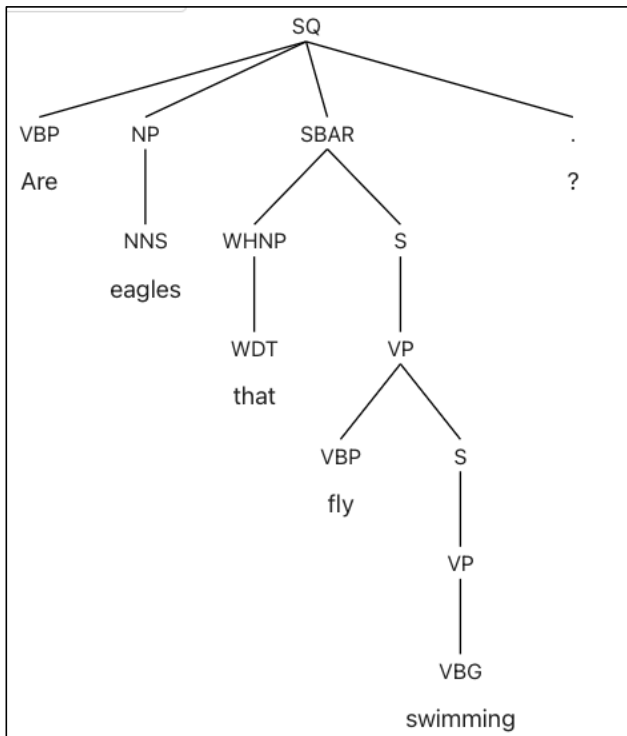


[pg.39, Chomsky POP (2013)]

- *Are* eagles that fly swimming?
- **Are* eagles that swimming fly?

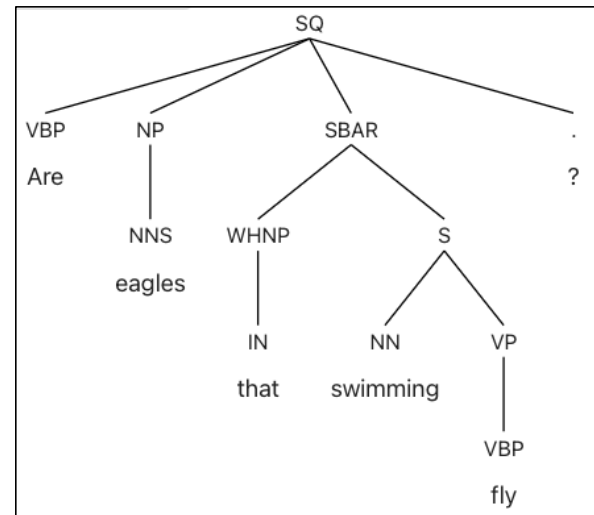


Berkeley Neural Parser



[pg.39, Chomsky POP (2013)]

- *Are* eagles that fly swimming?
- **Are* eagles that swimming fly?



Parsing not the same as Internal Thought

- Operative Complexity less for Internal Thought
 - *Language is optimized for thought, not communication*
- No Phases
 - Chomsky *MI* (2000) assumes WS's are pre-partitioned:

(26) the demonstration that glaciers are receding showed that global warming must be taken seriously

The prefinal phases of the derivation are the syntactic objects corresponding to (27a–c).⁵⁵

- (27) a. $P_1 = [_{CP} \text{ that global warming must be taken seriously}]$
b. $P_2 = [_{CP} \text{ that glaciers are receding}]$
c. $P_3 = [_{VP} [\text{the demonstration } P_2 [\text{show } P_1]]]$

For each new phase, a subarray provides the lexical material required and the operations proceed in the manner already sketched, with P_1/P_2

The SMT Parser Project

- *Isn't it a mystery that we can parse externalized language at all?*
 - No help from SMT
 - Only Merge builds thoughts (**BP**)
 - Not enough time for Nature to tinker with language
 - Not enough time to evolve new systems or mechanisms, e.g. a parsing algorithm

Basic Property (BP) of Language

[Chomsky GK (2021), pg.9]

- The puzzle is that from infancy on we ignore 100% of what we hear (linear order) and reflexively use only structures that we never hear but that our mind constructs, with non-trivial computations.
- The reason must be that linear order is simply not available to the I-language, the system that constructs thoughts.
- Why then does speech require linearization? The articulatory system cannot produce structures ...
 - The [sensory-motor] systems used for externalization have nothing to do with language; they were in place long before language emerged, and have not changed since.
 - Pure language is the internal system that generates thought.

Communication and Thought

- Language organ is designed to construct thoughts efficiently
- Language is not designed for efficient communication
- If that makes expressions hard to process and even makes some thoughts impossible to express without circumlocution, **too bad. Nature doesn't care.** [Chomsky GK (2021), pg.11]
- **Nature doesn't care.** [Chomsky GK (2021), pg.11]
 - EXT cannot have come before Merge.
 - The modern doctrine that language may have evolved from animal **communication** seems quite untenable. [pg.10, Chomsky GK (2021)]
- It makes no sense to say that *some system evolved for X*
 - “the spine evolved for keeping us upright,” or “language evolved for communication”

Communication and Thought

- **Communicative efficiency** is always sacrificed
 - The most serious cases involve deletion of copies in accord with computational efficiency, leading to some of the **hardest parsing problems**. [pg.10, fn.12, Chomsky GK (2021)]
 - *see solutions in the SMT Parser ...*
 - Externalization: *John or the men *is/*are in the room*
 - *... unproblematic for expression of thought if feature valuation kept to late insertion so that only the bare copula reaches the thought level (as in some spoken dialects)*
- Note that **statistical information is irrelevant to I-language** as a matter of principle, though as has always been assumed in the generative enterprise (see Chomsky 1957), it can be highly relevant to processing and acquisition.

Acknowledgments

- Masayuki Oishi, Jason Ginsburg, Noam Chomsky
- **Parser**
 - sandiway.arizona.edu/smtparser
- **Theory**
 - Chomsky 3M:
 - Three models for the Description of Language. *IEEE Transactions on Information Theory*. 1956.
 - Chomsky MI:
 - Minimalist Inquiries: The Framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Martin, R., D. Michaels & J. Uriagereka. MIT Press. 2000.
 - Chomsky POP:
 - Problems of Projection. *Lingua* 130:33-49. 2013.
 - Chomsky GK:
 - Minimalism: Where Are We Now, and Where Can We Hope to Go. *Gengo Kenkyu*, 160:1–42. 2021.
 - Chomsky MC:
 - The Miracle Creed and SMT. In Greco, M. & Mocci, D. (Eds.), *A Cartesian dream: A geometrical account of syntax: In honor of Andrea Moro*. *Rivista di Grammatica Generativa/Research in Generative Grammar*. (to appear)

APPENDIX

Appendix: TG

- Why is (1), but not (2) and (3), ambiguous?

1) *the shooting of the hunters*

2) *the growling of lions*

3) *the raising of flowers*

[pg.123, Chomsky 3M (1956)]

- Answer:

- *lions growl => the growling of lions*

- *they raise flowers => the raising of flowers*

- *hunters shoot*

- *they shoot the hunters*

Appendix: TG

[pg.192, Chomsky *Remarks* (1970)]

- *Tomatoes grow*
- *The growth of tomatoes* (unambiguous)
 - *tomatoes grow => the growth of tomatoes*
- *John grows tomatoes*
 - John [+cause] [_S *tomatoes grow*]
 - **John grows tomatoes => the growth of tomatoes*
- *The growing of tomatoes* (ambiguous)
 - *lions growl => the growling of lions*
 - *they raise flowers => the raising of flowers*

the corresponding derived nominal is excluded (though not, of course, the corresponding nominalization *the growing of tomatoes* — we return to nominalizations of this type on p. 214). Hence the lack of ambiguity offers empirical support for a combination of the lexicalist hypothesis with the causative analysis, though not for either of these assumptions taken in isolation.

Appendix: TG

[pg.214–215, Chomsky *Remarks* (1970)]

- *John grows tomatoes*
 - John [+cause, grow] tomatoes

In other words, we postulate that there is a feature [+cause] which can be assigned to certain verbs as a lexical property. Associated with this feature are certain redundancy rules which are, in this case, universal, hence not part of the grammar of English but rather among the principles by which any grammar is interpreted. These principles specify that an intransitive with the feature [+cause] becomes transitive and that its selectional features are systematically revised so that the former subject becomes the object.

To account for the distinction between (56c) and (57), we must restrict the feature [+cause] with respect to the feature that distinguishes derived nominals such as *growth* from forms such as *growing*, limiting it to the latter case.