download these slides!

On the Strong Minimalist Thesis:

Towards Efficient Computation and Perception

Sandiway Fong

Dept. of Linguistics

University of Arizona

sandiway.arizona.edu

sandiway.arizona.edu/kyotoDec2024.pdf ☛

sandiway AT arizona DOT edu

Kyoto University Language Sciences Lecture Series, Dec 8th 2024

Talk Outline

- involves some math, some computer science and syntax
- Don't worry:
  - I'll explain everything.
  - Please interrupt and ask questions!

Topics

- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language
  - simplicity of **I-Language**
  - **Merge, Minimal Search** and **operative complexity**
  - The **slow brain**
  - **Evolution**
- **Part 2: Parsing**
  - from E-Language to I-Language
  - describe a parser
  - *Merge operative complexity tamed*?

What is the Strong Minimalist Thesis (SMT)?

- a theory design guideline (Chomsky 2024)
- **SMT**: **Language** satisfies Einstein's *Miracle Creed*

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

...nstein published an essa...
...*American* magazine. Reflecting on re...
...king back over his distinguished career, Einste...
...ng philosophical:

Time and again the passion for understanding has led to the illu...

perienced reality, but that the totality of all sensory experience can be "comprehended" on the basis of a conceptual system built on premises of great simplicity. The skeptic will say that this is a "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)

can be "comprehended" on the basis of a conceptual system b...
...n premises of great simplicity. The skeptic will say that th...
...acle creed." Admittedly so, but it is a miracle creed...
...ne out to an amazing extent by the devel...
...1950, 342)

intro (McDonough 2022)

What does it mean for I-Language?

- "The Strong Minimalist Thesis (SMT) holds that language too may satisfy the miracle creed **at its core**." (Chomsky 2024)
- At the core: **I-Language**
  - **I = internal**: the expressions computed by Merge
  - could be a **well-formed thought** but not (directly) externalizable

well-formed thought but not externalizable

[pg.39, (Chomsky 2013)]

$\{C_Q$, {INFL, {{eagles, $\{C_{rel}$, {INFL, {~~eagles~~, $\{v_\theta$, fly}}}}}, $\{v_\theta$, swim}}}}

can

- Eagles that fly swim
- Eagles that fly *can* swim **?** (*turn into a question: front modal verb*)
- *Can* eagles that fly *swim*? C_Q: question about *swim* (not *fly*)

$\{C_Q$, {INFL, {{eagles, $\{C_{rel}$, {INFL, {can, {~~eagles~~, $\{v_\theta$, fly}}}}}}, $\{v_\theta$, swim}}}}

- Eagles that *can fly* swim (*let's try turning it into a question*)
- **Can* eagles that *fly* swim? **well-formed thought** (no EXT)

> *"… that is a fine thought, but it cannot be expressed by [this sentence]."*

What does it mean for I-Language?

- "The Strong Minimalist Thesis (SMT) holds that language too may satisfy the **miracle creed** **at its core**." (Chomsky 2024)
- At the core: **I-Language**
  - internal: the expressions computed by Merge
  - could be a **well-formed thought** but not (directly) externalizable
  - not linearly ordered, see **Basic Property** (BP)
  - Eagles that *can* fly *can swim*
- **E-Language**:
  - Externalized I-Language (**EXT**), e.g. pronounced or signed or written
  - linear order imposed by the modality
  - word order and spellout parameterized *by particular (E-)language*

return to talk about this soon!

Miracle Creed: nature maximizing simplicity

Dialogue Concerning the Two Chief World Systems (**Galileo 1632)**

- "***nature*** (*which by general agreement does not act by means of many things when it can do so by means of few*)"
  - **Context**: general discussion about motion of the planets

- *Quaderni d'anatomia IV* **(Leonardo da Vinci)**:
  - "*Every action in **nature** takes place in the shortest way possible.*"
  - quoted in *Leonardo's Optics* (Argentieri, 1956)

SMT **optimal** solution:

- *Nature adapts/optimizes what it has to work with*

Topics

- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language

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*. (Chomsky 2021)

- **BP: no**, simplest rule actually available:
  - *build structure, then determine nearest*
  - **not acquired**: observed in children, as early as 30 months

Human toolkit: we have linear order operations!

- **Number Agreement**:

  a. the bombing$_{sg}$ of the cities$_{pl}$ **was**$_{sg}$ criminal [pg.9, (Chomsky 2021)]

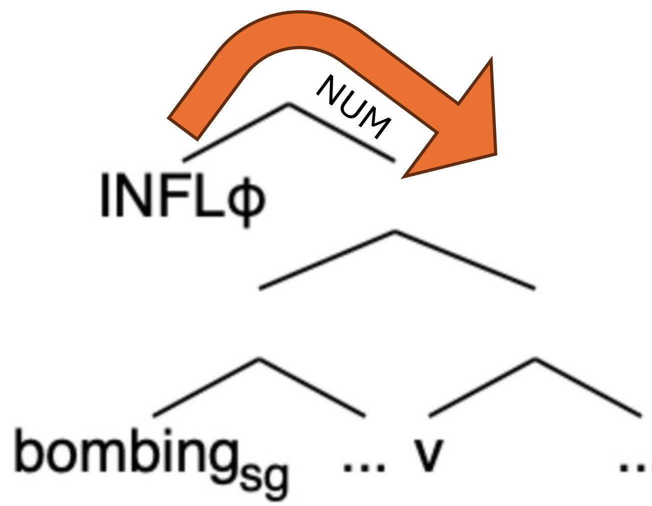
  b. * were$_{pl}$

  c. the bombings$_{pl}$ of the city$_{sg}$ **were**$_{pl}$ criminal

  d. * was$_{sg}$

Basic Property (BP) of Language

- first build structure:
  - *the bombing of the cities*
  - {bombing$_{\text{the,[sg]}}$, (of) {cities$_{\text{the,[pl]}}$}}
- *then do* **(Minimal) Search**:
  - e.g. search for NUM
  - **Ans**: [sg]



Basic Property (BP) of Language

[pg.9, (Chomsky 2021)]

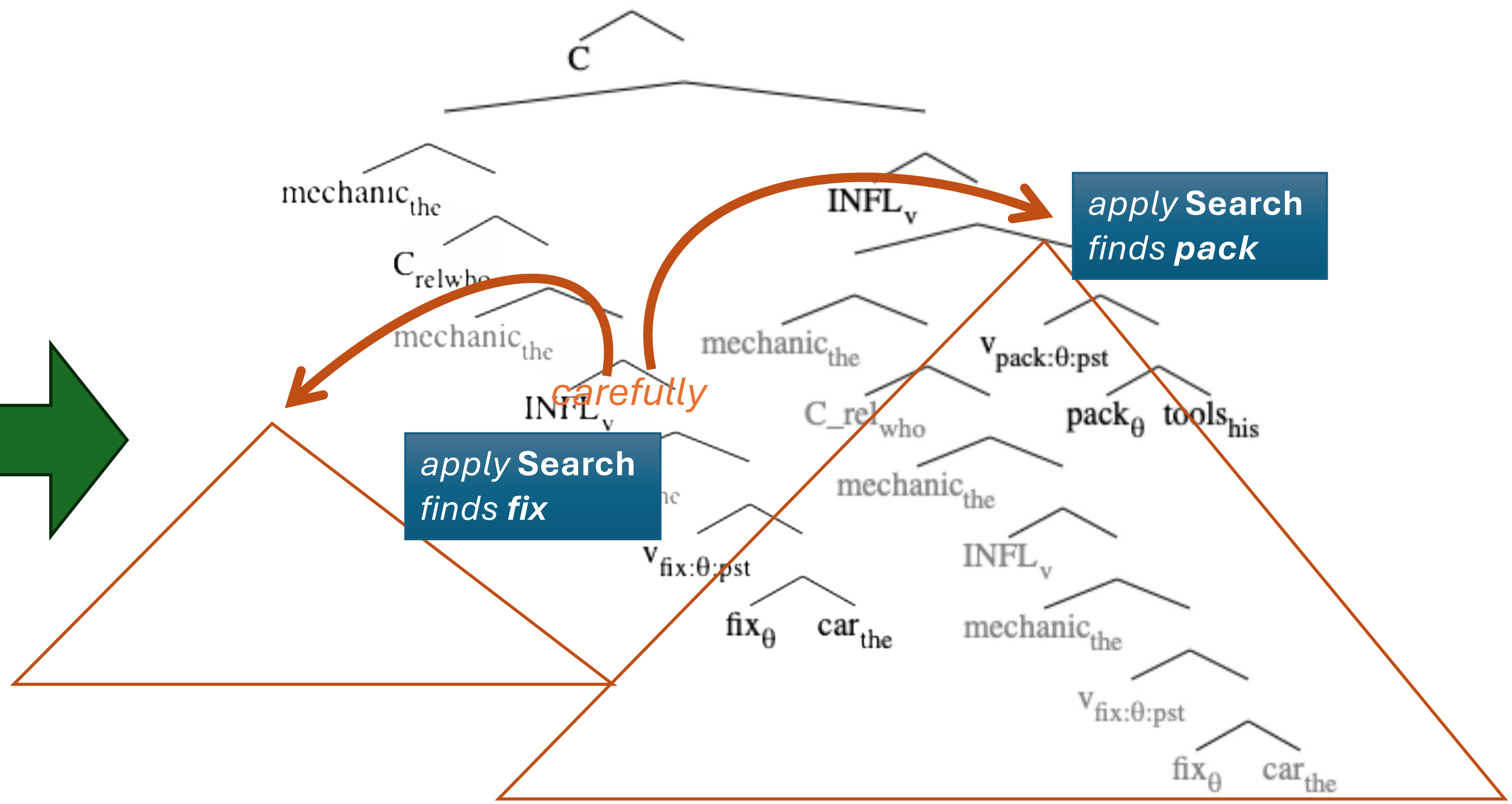
- "adverb *carefully* seeks a verb [to modify], but it cannot use the **simplest computation**: pick the **linearly closest** verb."

- **Construal**:
  - Below: [...] *marks linearly closest verb to the adverb*

- the mechanic who *fixed* the car *carefully* [*packed*] his tools ⟸ ANS: [*pack*] or *fix*
- *Carefully*, the mechanic who [*fixed*] the car *packed* his tools ⟸ ANS: *pack*
- the mechanic who *fixed* the car [*packed*] his tools *carefully* ⟸ ANS: [*pack*] ✅
- the mechanic who *carefully* [*fixed*] the car *packed* his tools ⟸ ANS: [*fix*] ✅

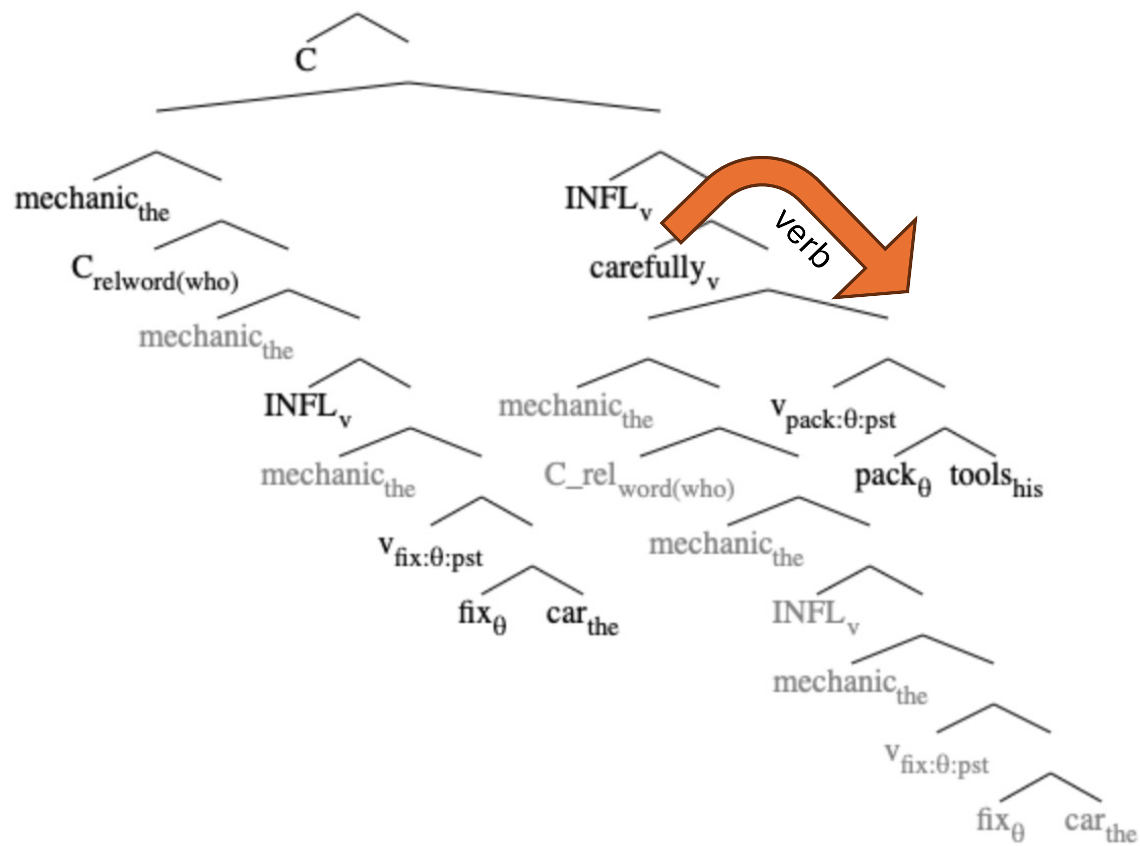
Why? the mechanic who *fixed* the car *carefully packed* his tools

Part 2: a **parser** *must* compute syntactic structure to understand *carefully*



Why? the mechanic who *fixed* the car *carefully packed* his tools

- *carefully* initiates a Search
- **Search** locates the relevant **term** (*a verb*)
- Search is **minimal**
- Simplest **structural** computation



Topics

- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language
  - simplicity of **I-Language**
  - **Merge, Minimal Search** and **operative complexity**

Merge

- 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**?*

we'll be talking about this very soon!

... a bit later

Simplest Merge

(1) External (EM)

{...} just notation, but we don't really use mathematical set theory

- X Y ⇒ {X, Y}
- X = {.. {..Y..}..} ⇒ {**Y**, {.. {..Y..}..}}, Y a **sub-term** of X
- assume all this happens in a **Workspace** (WS) **without replacement**

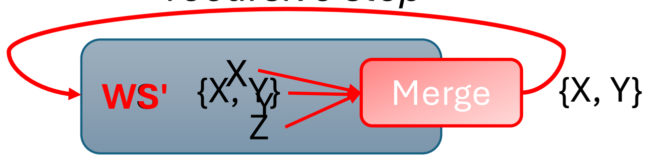
o/w can circumvent c-command

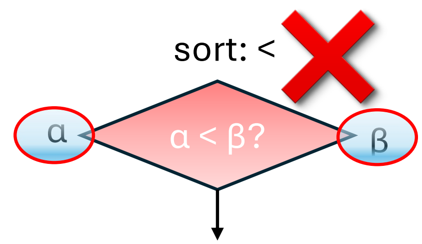
(2) Internal (IM)

Operative Complexity

- Adopt simplest **recursive** formal device
  - i.e. *Merge feeds Merge* in the Workspace (WS)
  - not a one-time operation*, cf. Conjoin* (Progovac 2015)

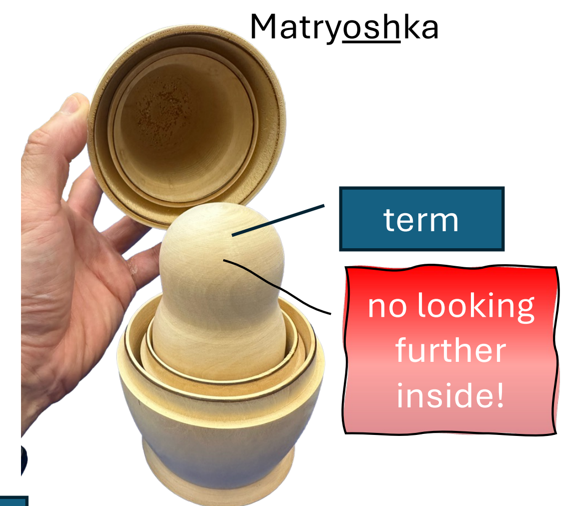
recursive step





Z ∈ WS inactive in the last round

Computation converges: one syntactic object

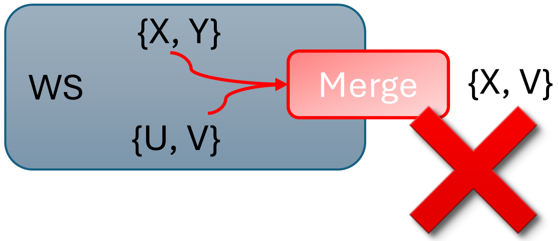


3rd Factor: all operations obey this

- (Minimal) **Search**:
  - *look in the WS or internally for a* ***term****, 1st thing you find, have to stop*

Minimal Search (MS)

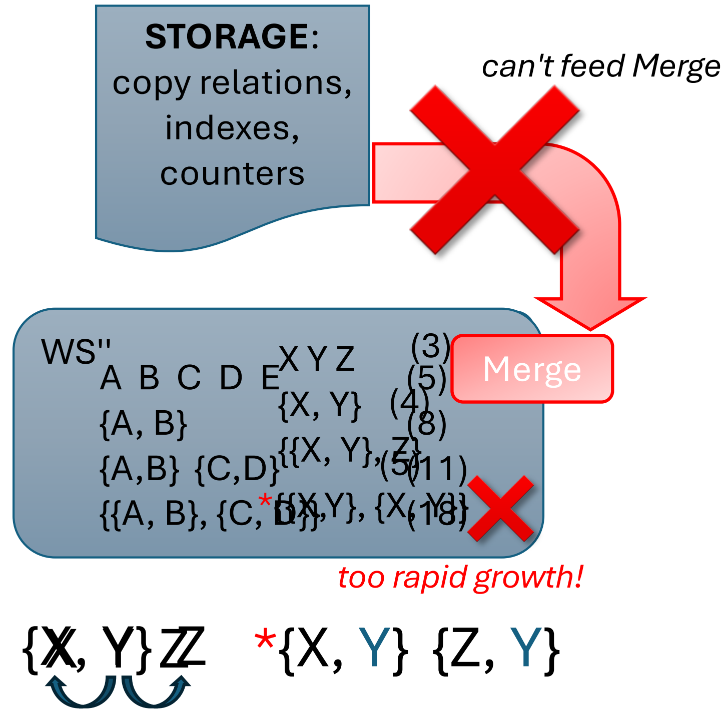
- (Chomsky p.c.):
  - **We assume that Merge like other operations observes it.**
  - **That's why only members of WS, not their terms, are eligible for [External Merge].**



- Chomsky (p.c.):
  - **Right now I don't see any reason why any operation should be exempt from MS. If so, MS can include structural identity checking -- which is its basic intuitive content.**

Merge is limited

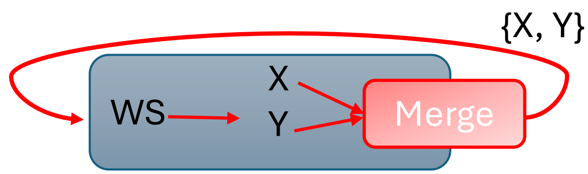
- Markovian assumption:
  - no storage/counter memory
  - no WS history: WS' cannot see WS *or earlier*
    - *too powerful: can build anything*
  - minimize WS complexity: Minimal Yield (MY)
    - *growth can be in terms of WS item + term access*
- **Simplest (recursive) Merge**
  - *no further elaboration permitted*
  - no parallel Merge
  - no sideways Merge
  - no 3 objects at a time
  - no splicing/tuck-in operations
  - *etc*.



no explicit ban needed: *violates WS Minimal Search*

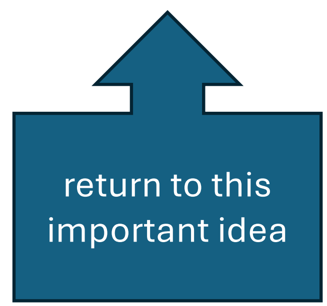
but see FormSet (Chomsky 2021; 2024)
John, Bill, my friends, the actor who won the Oscar …
John arrived and met Bill
also (Fong & Oishi *fc*.)
the politician is greedy, a cad and a charlatan

Operative Complexity



- **Question**: now, is *simplest* Merge efficient enough for biology?
- Actually, it has horrible combinatorics
  - *not feasible for biology,*
  - *not feasible for computers*

- **Answer**: Merge has Language Specific Constraints (LSCs)
  - I-Language Merge could be feasible



Topics

- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language
  - simplicity of **I-Language**
  - **Merge, Minimal Search** and **operative complexity**
  - The **slow brain**
  - **Evolution**

Evolution: *modern humans*

*Language, the ultimate symbolic mental function, it is **virtually impossible to conceive of thought** as we know it in its absence.*
(Tattersall 2006)

"if we are seeking a single cultural releasing factor that opened the way to **symbolic cognition**, the invention of language is the most obvious candidate." (Tattersall 2006)

until the emergence of behaviorally modern *H. sapiens*: **in general, technological innovations have been sporadic and rare**. The most-striking evidence for a distinct cognitive contrast between modern humans and all their predecessors, however, comes from Europe. *H. sapiens* came late to this continent and brought a new kind of stone tool based on striking long thin "blades" from a carefully prepared long core. **In short order these Europeans, the so-called Cro-Magnons, left a dazzling variety of symbolic works of prehistoric art.**

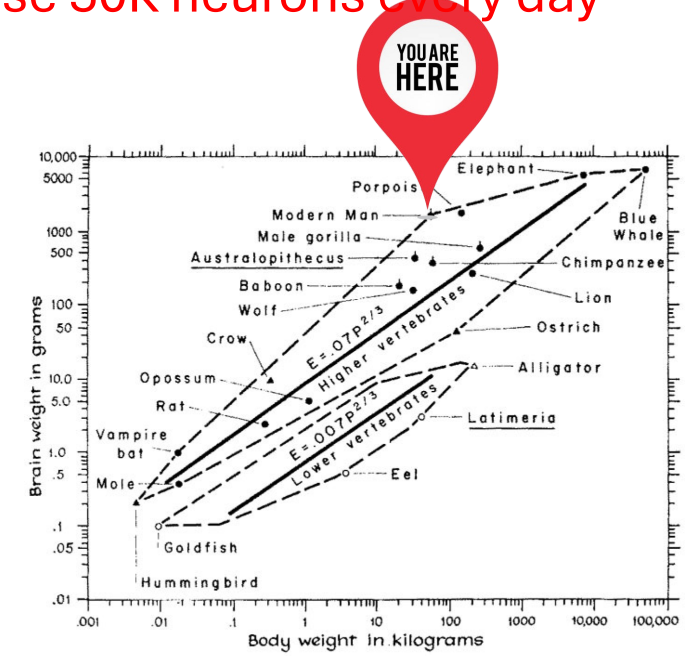
rare in mammals, but not unknown for primates

Are we special? Allometric scaling

Brain: 86 billion neurons

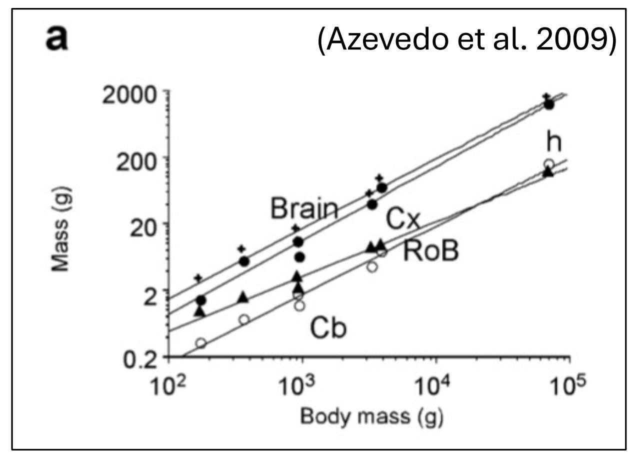
Great!

- we lose 50K neurons every day



Nah, we're not outliers!

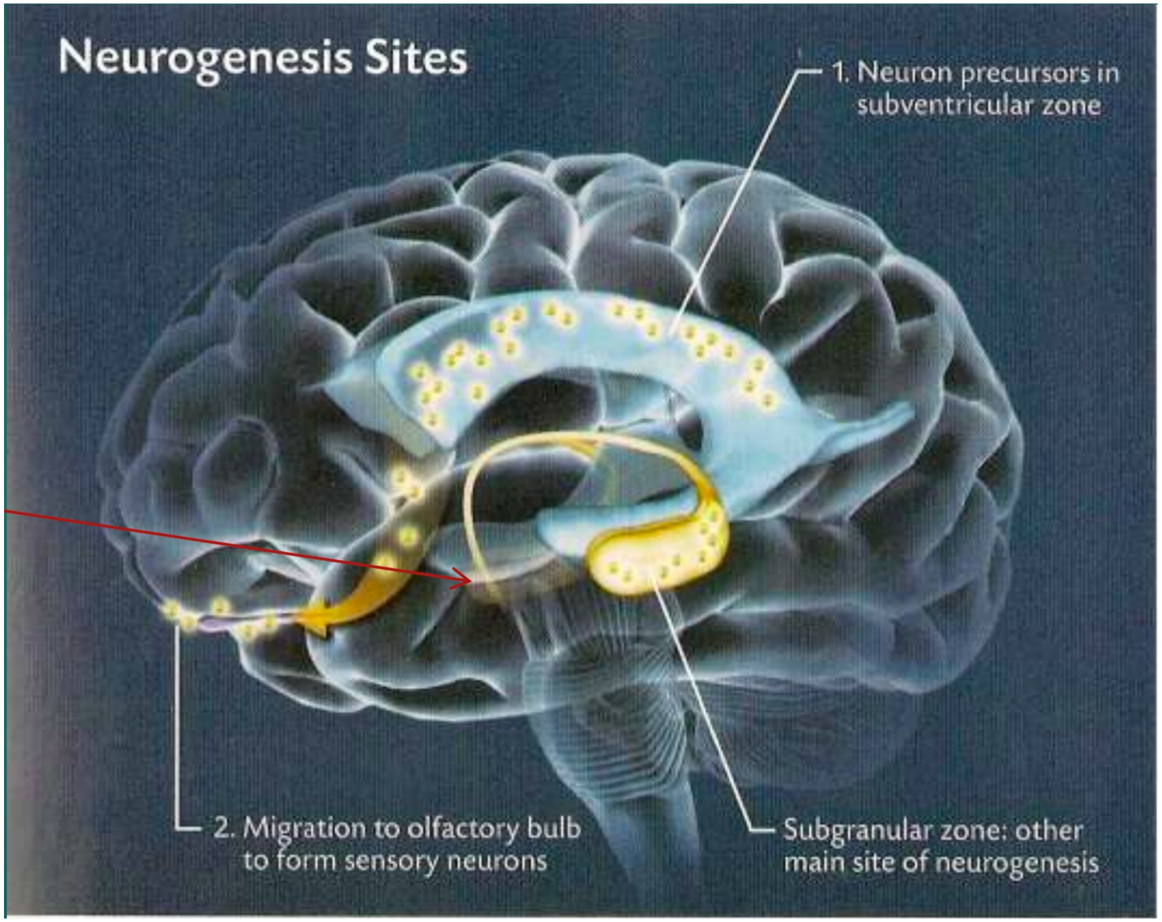
- Primate brain scaling: *uniquely human*?



Human Brain Development

Vella (2016):

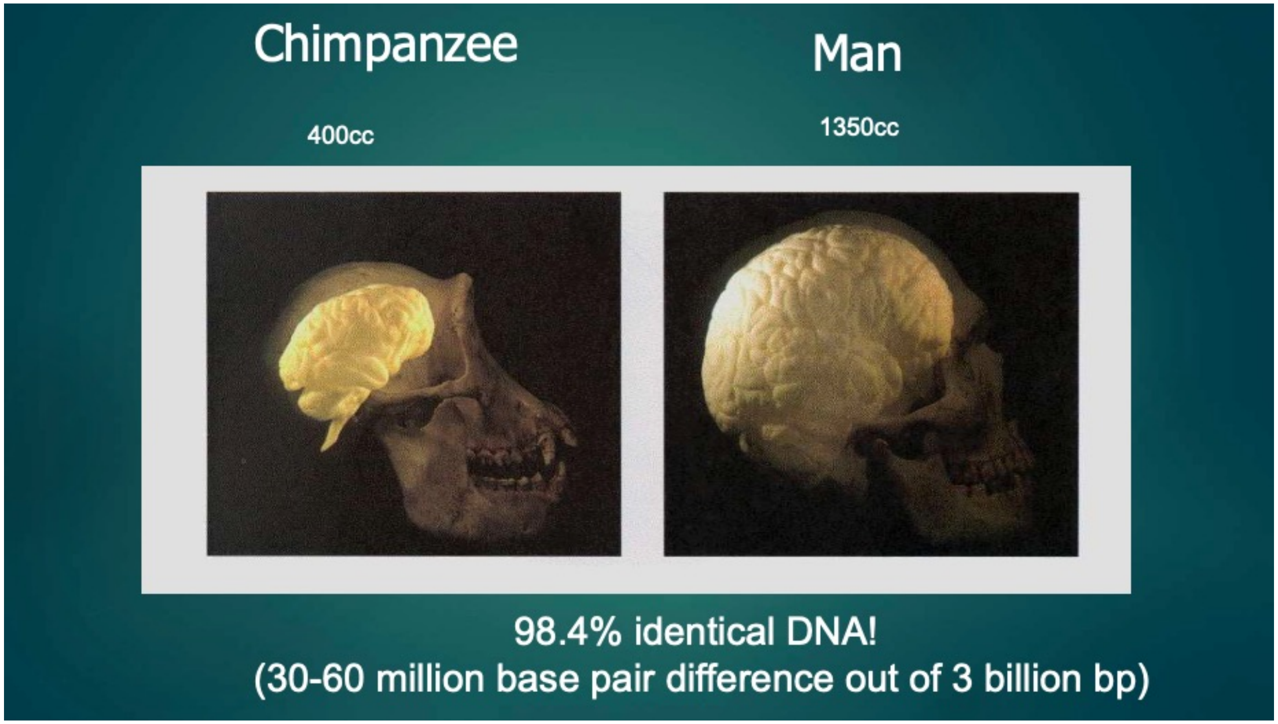
- **Perinatal neuron cell death**: Infant primates may have up to **twice the adult number of neurons**.
- Great Adolescent Pruning: **Age 5-21**
  - **Heavy synaptic pruning**: circuits are sculpted from the brain by **pruning** away cells and synapses.
  - Mechanisms: Programmed cell death (apoptosis), passive loss due to lack of stimulation, learning.
- 1.4K new neurons a day



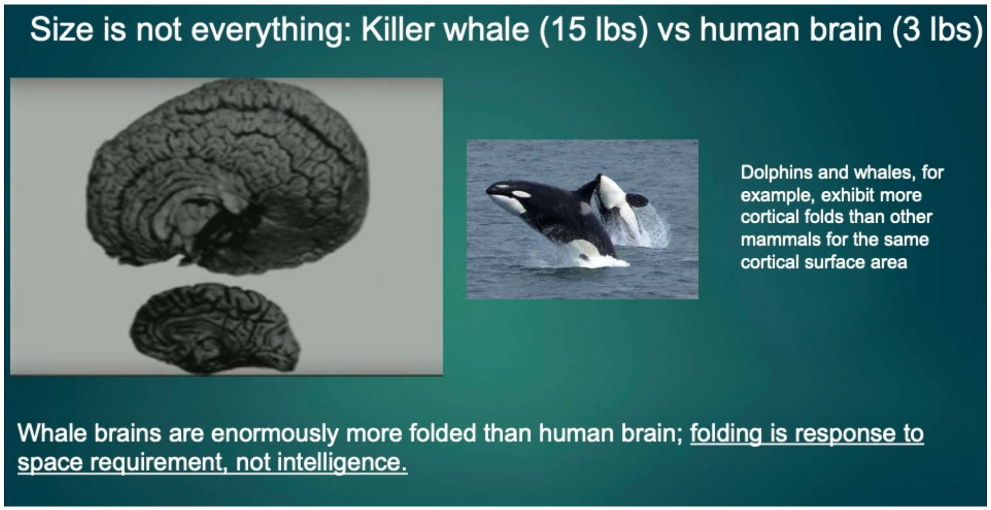
Primates

(Vella 2016)

- *Animals with large brains are rare*
- Energy cost is high (20W)
- Longer gestation
- More wiring means slower brain unless reorganized



Absolute brain size



Vella (2016)

[pg145. Darwin (1871)]

- *no one supposes that the* ***intellect*** *of any two animals or of any two men can be accurately gauged by the cubic contents of their skulls.*

Special, yes, but ...

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

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

Y'all noticed the 20% better weather forecasts, right? 😀

Topics

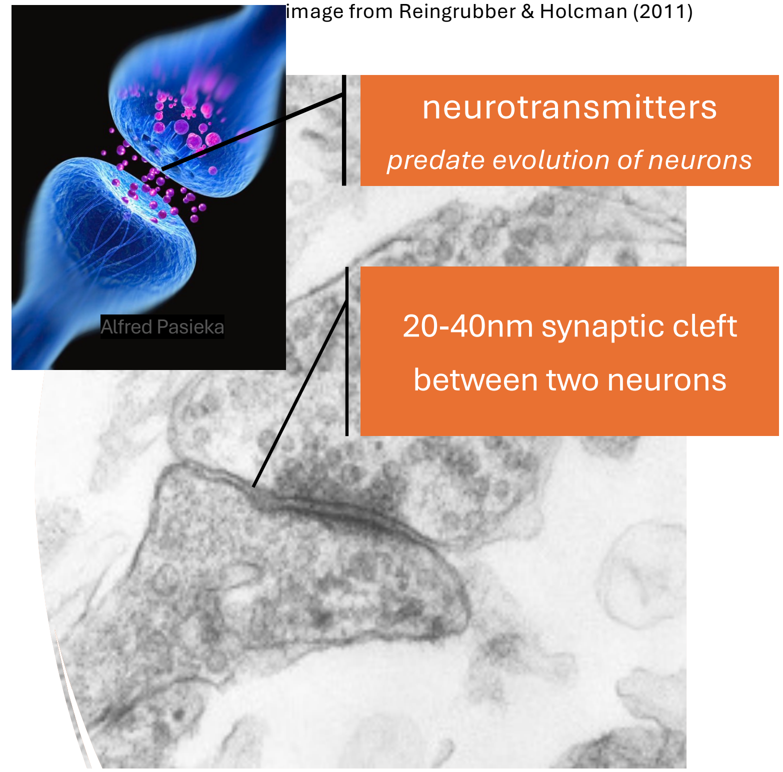
- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language
  - simplicity of **I-Language**
  - **Merge, Minimal Search** and **operative complexity**
  - The **slow brain**
  - **Evolution**

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

image from Reingrubber & Holcman (2011)




140 RENATI DES-CARTES

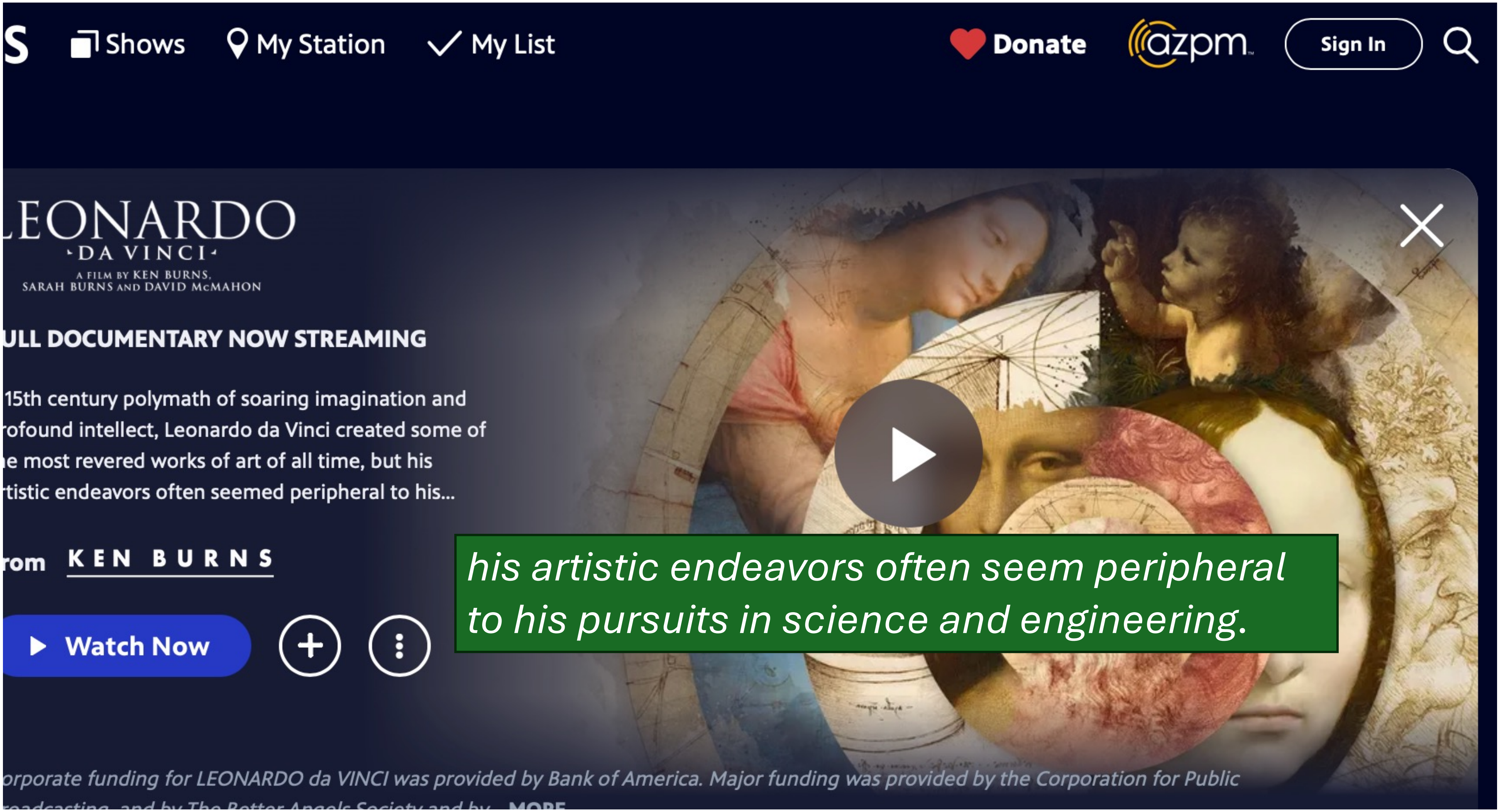
dulæ, quibus obverſus eſſe poteſt tubus 8, ſic reſpondere omnibus locis ad quæ brachium 7 converti poteſt, ut non alia de cauſa brachium illud ſit converſum ad objectum B, quam quia tubus ille reſpicit glandulæ punctum *b*. Quod-

ſi ſpiritus mutantes curſum ſuum, hunc tubum ad aliud glandulæ punctum convertant, puta verſus *c*, filamenta 8, 7, quæ

Brain: earlier theories

- *De Homine*
  - (Descartes 1662)
  - H: pineal gland
  - hydraulic muscle control
- Leonardo da Vinci
  - ventricles (brain)
  - *imprensiva*
  - *senso comune*
  - *memoria*

  (Pevsner 2019)


S
Shows
My Station
My List
Donate
azpm
Sign In
LEONARDO
DA VINCI
A FILM BY KEN BURNS,
SARAH BURNS AND DAVID McMAHON
ULL DOCUMENTARY NOW STREAMING
15th century polymath of soaring imagination and
rofound intellect, Leonardo da Vinci created some of
e most revered works of art of all time, but his
tistic endeavors often seemed peripheral to his...
rom KEN BURNS
Watch Now
his artistic endeavors often seem peripheral
to his pursuits in science and engineering.
orporate funding for LEONARDO da VINCI was provided by Bank of America. Major funding was provided by the Corporation for Public

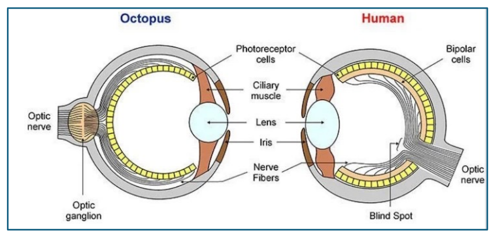
Evolution is slow, Language is recent

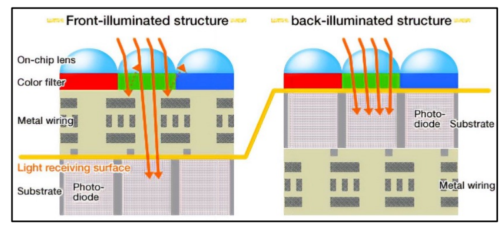
Land & Fernald (1992), *Animal Eyes* Land & Nilsson (2012)

- 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*
- *[Many parallels between Language and the visual system … not discussed here]*

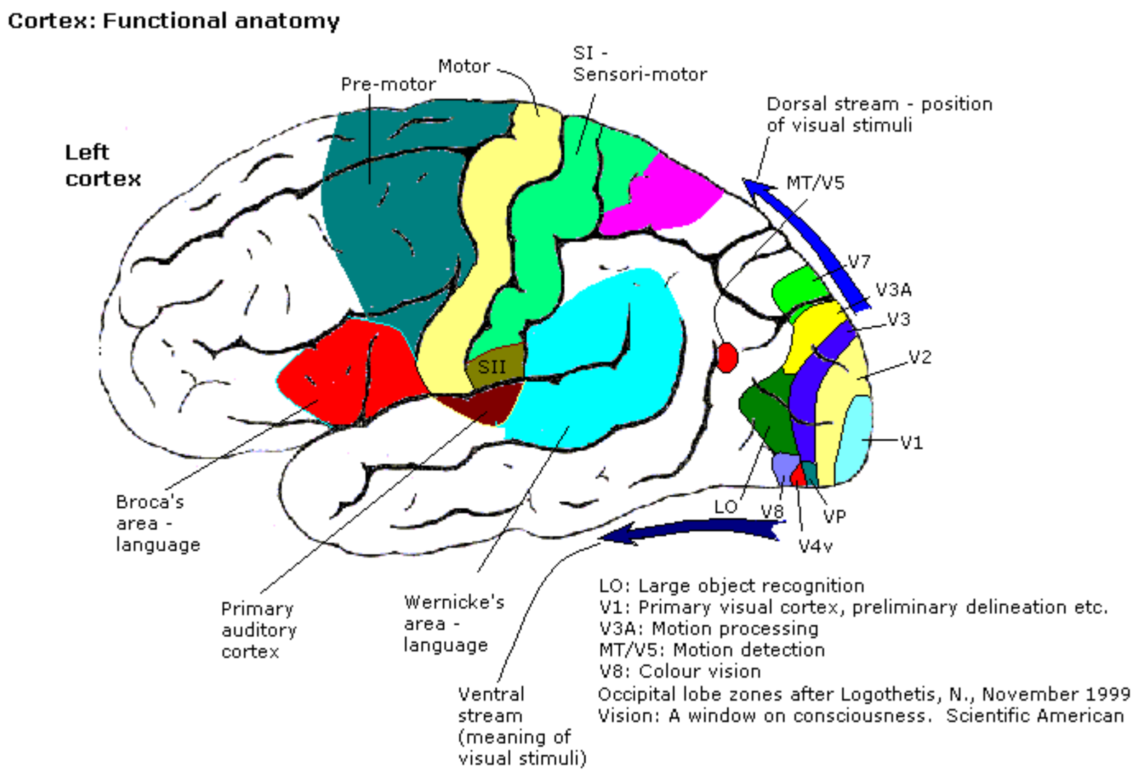




nevsemi.com

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

Vision: more area, more evolved than Language?



Vision developed much earlier: **Nature** had time to evolve it.

- 50% of the cortex
- V1 primary visual cortex: retinotopic map
- V2 neurons build upon the basic features detected in V1, extracting more complex visual attributes such as texture, depth, and color

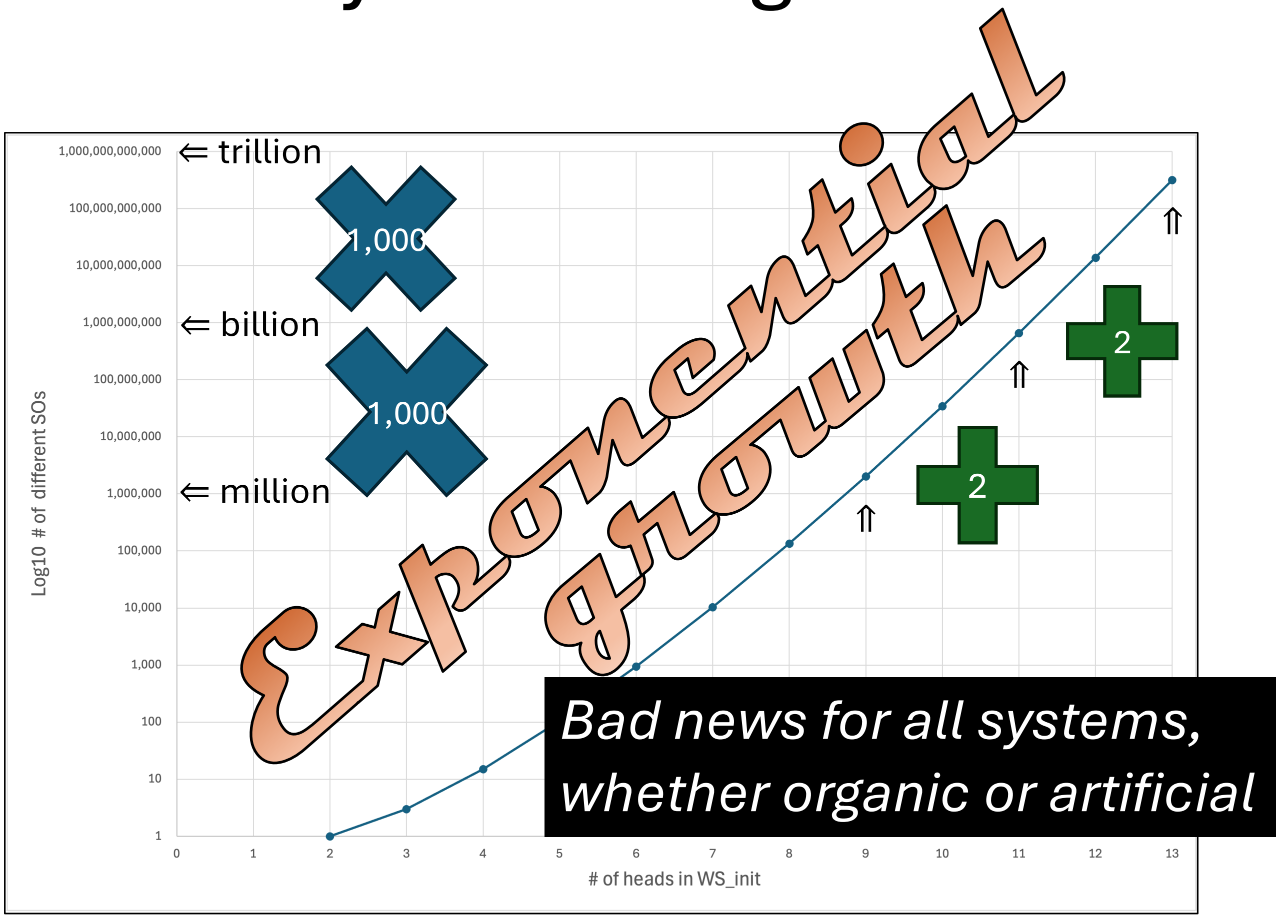
Topics

- **Part 1: Strong Minimalist Thesis** (SMT)
  - **Basic Property (BP)** of Language
  - simplicity of **I-Language**
  - **Merge, Minimal Search** and **operative complexity**
  - The **slow brain**
  - **Evolution**
- **Part 2: Parsing**
  - from E-Language to I-Language
  - describe a parser
  - *Merge operative complexity tamed*?

Computational Complexity of Merge

- Merge as a mathematical abstraction
  - formal complexity of Merge raises issues
- Merge as applied to I-Language

Summary: Free Merge is mathematically bad



Heads	#SOs
2	1
3	3
4	15
5	105
6	945
7	10,395
8	135,135
9	2,027,025
10	34,459,425
11	654,729,075
12	13,749,310,575
13	316,234,143,225

*distinct syntactic objects for External Merge

Merge Combinatorics

Consider External Merge only

- and only those cases that **converge** on a single Syntactic Object (SO)

- Given WS_{init} =
  - h_1 h_2 converge on: $\{h_1, h_2\}$ (1 case, order unimportant)
  - h_1 h_2 h_3 converge on 3 cases:
    - $\{\{h_1, h_2\}, h_3\}$
    - $\{\{h_1, h_3\}, h_2\}$
    - $\{\{h_2, h_3\}, h_1\}$
  - h_1 h_2 h_3 h_4 converge on 15 cases:
    - $\{\{\{h_1, h_2\}, h_3\}, h_4\}$ $\{\{\{h_1, h_2\}, h_4\}, h_3\}$ $\{\{h_1, h_2\}, \{h_3, h_4\}\}$
    - $\{\{\{h_1, h_3\}, h_2\}, h_4\}$ $\{\{\{h_1, h_3\}, h_4\}, h_2\}$ $\{\{h_1, h_3\}, \{h_2, h_4\}\}$
    - $\{\{\{h_1, h_4\}, h_2\}, h_3\}$ $\{\{\{h_1, h_4\}, h_3\}, h_2\}$
    - $\{\{\{h_2, h_3\}, h_1\}, h_4\}$ $\{\{\{h_2, h_3\}, h_4\}, h_1\}$ $\{\{h_2, h_3\}, \{h_1, h_4\}\}$
    - $\{\{\{h_2, h_4\}, h_1\}, h_3\}$ $\{\{\{h_2, h_4\}, h_3\}, h_1\}$
    - $\{\{\{h_3, h_4\}, h_1\}, h_2\}$ $\{\{\{h_3, h_4\}, h_2\}, h_1\}$

Merge Combinatorics

Merge

- Given WS_{init} = h_1 h_2 h_3 h_4 h_5, converge on **105** cases:
  - Let #c(WS) = # *convergent* cases for WS.
  - **Example**: if |WS| = 3, #c(|WS|=3) = 3, e.g. *α β γ* ⇒ ① {{*α*, *β*}, *γ*} ② {{*α*, *γ*}, *β*} ③ {{*β*, *γ*}, *α*}
  - Let also redundant pairs be marked in red, e.g. h_2 h_3 means *{h_2 , h_3}
  - {h_1, h_2} h_3 h_4 h_5 = #c(|WS|=4) =15 cases
  - {h_1, h_3} h_2 h_4 h_5 = 15 cases
  - {h_1, h_4} h_2 h_3 h_5 = 15 – #c({h_1, h_4} {h_2, h_3} h_5) = 15 –
  - {h_1, h_5} h_2 h_3 h_4 = 15 – $_3C_2$ x #c({h_1, h_5} {h_2, h_3} h_4) =
  - {h_2, h_3} h_1 h_4 h_5 = 15
  - {h_2, h_4} h_1 h_3 h_5 = 12
  - {h_2, h_5} h_1 h_3 h_4 = 6
  - {h_3, h_4} h_1 h_2 h_5 = 12
  - {h_3, h_5} h_1 h_2 h_4 = 6
  - {h_4, h_5} h_1 h_2 h_3 = 6 cases
  - (15 + 15 + 12 + 6) + (15 + 12

to be explained below

we know this from the previous slide

- Given WS_{init} =
  - h_1 h_2 … converge on: {h_1, h_2} (1 case,

red: h_2 h_3, i.e. *must block Merge* h_2 …

3 items {h_1, h_4} {h_2, h_3} h_5 *produce 3 objects*

 # objects = 3

Why? It's r…

It will be in…

Don't want duplicates…

$_3C_2$ *= 3 x # convergent objects from* {h_1, h_4} {h_i, h_j} h_k

i,j,k drawn from 2,3,4.

…rge on: {h_1, h_2}

…erge on 3 cases:

red: h_2 h_3 h_4, i.e. *block redundant Merge of any* h_2 h_3 h_4 *pair*

Independently generated by {h_2, h_3} {h_2, h_4} {h_3, h_4} lines.

- {{h_2, h_3}, h_1}

i.e. anything containing the …

viz. any convergent object from {h_1, h_4} {h_2, h_3} h_5

…, h_3}, h_1}

- {{{h_3, h_4}, h_1}, h_2} {{{h_3, h_4}, h_2}, h_1}

Merge Combinatorics: $WS_{init} = h_1\ h_2\ h_3\ h_4\ h_5$

a simple computer program verifies

105

1. {{{{h1, h2}, h3}, h4}, h5}
2. {{{{h1, h2}, h3}, h5}, h4}
3. {{h4, h5}, {{h1, h2}, h3}}
4. {{{{h1, h2}, h4}, h5}, h3}
5. {{{{h1, h2}, h4}, h3}, h5}
6. {{h5, h3}, {{h1, h2}, h4}}
7. {{{{h1, h2}, h5}, h3}, h4}
8. {{{{h1, h2}, h5}, h4}, h3}
9. {{{h3, h4}, h5}, {h1, h2}}
10. {{{h3, h4}, {h1, h2}}, h5}
11. {{h5, {h1, h2}}, {h3, h4}}
12. {{{h3, h5}, {h1, h2}}, h4}
13. {{{h3, h5}, h4}, {h1, h2}}
14. {{{h4, h5}, {h1, h2}}, h3}
15. {{{h4, h5}, h3}, {h1, h2}}
16. {{{{h1, h3}, h4}, h5}, h2}
17. {{{{h1, h3}, h4}, h2}, h5}
18. {{h5, h2}, {{h1, h3}, h4}}
19. {{{{h1, h3}, h5}, h2}, h4}
20. {{{{h1, h3}, h5}, h4}, h2}
21. {{{h4, h5}, h2}, {h1, h3}}
22. {{{h4, h5}, {h1, h3}}, h2}
23. {{{{h1, h3}, h2}, h4}, h5}
24. {{{{h1, h3}, h2}, h5}, h4}
25. {{h4, h5}, {{h1, h3}, h2}}
26. {{{h4, h2}, h5}, {h1, h3}}
27. {{{h4, h2}, {h1, h3}}, h5}
28. {{h5, {h1, h3}}, {h4, h2}}
29. {{{h5, h2}, {h1, h3}}, h4}
30. {{{h5, h2}, h4}, {h1, h3}}
31. {{{{h1, h4}, h5}, h2}, h3}
32. {{{{h1, h4}, h5}, h3}, h2}
33. {{{{h1, h4}, h2}, h5}, h3}
34. {{{{h1, h4}, h2}, h3}, h5}
35. {{h5, h3}, {{h1, h4}, h2}}
36. {{{{h1, h4}, h3}, h5}, h2}
37. {{{{h1, h4}, h3}, h2}, h5}
38. {{h5, h2}, {{h1, h4}, h3}}
39. {{{h5, h2}, {h1, h4}}, h3}
40. {{{h5, h2}, h3}, {h1, h4}}
41. {{{h5, h3}, {h1, h4}}, h2}
42. {{{h5, h3}, h2}, {h1, h4}}
43. {{{{h1, h5}, h2}, h3}, h4}
44. {{{{h1, h5}, h2}, h4}, h3}
45. {{{{h1, h5}, h3}, h2}, h4}
46. {{{{h1, h5}, h3}, h4}, h2}
47. {{{{h1, h5}, h4}, h2}, h3}
48. {{{{h1, h5}, h4}, h3}, h2}
49. {{{{h2, h3}, h4}, h5}, h1}
50. {{{{h2, h3}, h4}, h1}, h5}
51. {{h5, h1}, {{h2, h3}, h4}}
52. {{{{h2, h3}, h5}, h1}, h4}
53. {{{{h2, h3}, h5}, h4}, h1}
54. {{{h4, h5}, h1}, {h2, h3}}
55. {{{h4, h5}, {h2, h3}}, h1}
56. {{{{h2, h3}, h1}, h4}, h5}
57. {{{{h2, h3}, h1}, h5}, h4}
58. {{h4, h5}, {{h2, h3}, h1}}
59. {{{h4, h1}, h5}, {h2, h3}}
60. {{{h4, h1}, {h2, h3}}, h5}
61. {{h5, {h2, h3}}, {h4, h1}}
62. {{{h5, h1}, {h2, h3}}, h4}
63. {{{h5, h1}, h4}, {h2, h3}}
64. {{{{h2, h4}, h5}, h1}, h3}
65. {{{{h2, h4}, h5}, h3}, h1}
66. {{{{h2, h4}, h1}, h5}, h3}
67. {{{{h2, h4}, h1}, h3}, h5}
68. {{h5, h3}, {{h2, h4}, h1}}
69. {{{{h2, h4}, h3}, h5}, h1}
70. {{{{h2, h4}, h3}, h1}, h5}
71. {{h5, h1}, {{h2, h4}, h3}}
72. {{{h5, h1}, {h2, h4}}, h3}
73. {{{h5, h1}, h3}, {h2, h4}}
74. {{{h5, h3}, {h2, h4}}, h1}
75. {{{h5, h3}, h1}, {h2, h4}}
76. {{{{h2, h5}, h1}, h3}, h4}
77. {{{{h2, h5}, h1}, h4}, h3}
78. {{{{h2, h5}, h3}, h1}, h4}
79. {{{{h2, h5}, h3}, h4}, h1}
80. {{{{h2, h5}, h4}, h1}, h3}
81. {{{{h2, h5}, h4}, h3}, h1}
82. {{{{h3, h4}, h5}, h1}, h2}
83. {{{{h3, h4}, h5}, h2}, h1}
84. {{{{h3, h4}, h1}, h5}, h2}
85. {{{{h3, h4}, h1}, h2}, h5}
86. {{h5, h2}, {{h3, h4}, h1}}
87. {{{{h3, h4}, h2}, h5}, h1}
88. {{{{h3, h4}, h2}, h1}, h5}
89. {{h5, h1}, {{h3, h4}, h2}}
90. {{{h5, h1}, {h3, h4}}, h2}
91. {{{h5, h1}, h2}, {h3, h4}}
92. {{{h5, h2}, {h3, h4}}, h1}
93. {{{h5, h2}, h1}, {h3, h4}}
94. {{{{h3, h5}, h1}, h2}, h4}
95. {{{{h3, h5}, h1}, h4}, h2}
96. {{{{h3, h5}, h2}, h1}, h4}
97. {{{{h3, h5}, h2}, h4}, h1}
98. {{{{h3, h5}, h4}, h1}, h2}
99. {{{{h3, h5}, h4}, h2}, h1}
100. {{{{h4, h5}, h1}, h2}, h3}
101. {{{{h4, h5}, h1}, h3}, h2}
102. {{{{h4, h5}, h2}, h1}, h3}
103. {{{{h4, h5}, h2}, h3}, h1}
104. {{{{h4, h5}, h3}, h1}, h2}
105. {{{{h4, h5}, h3}, h2}, h1}

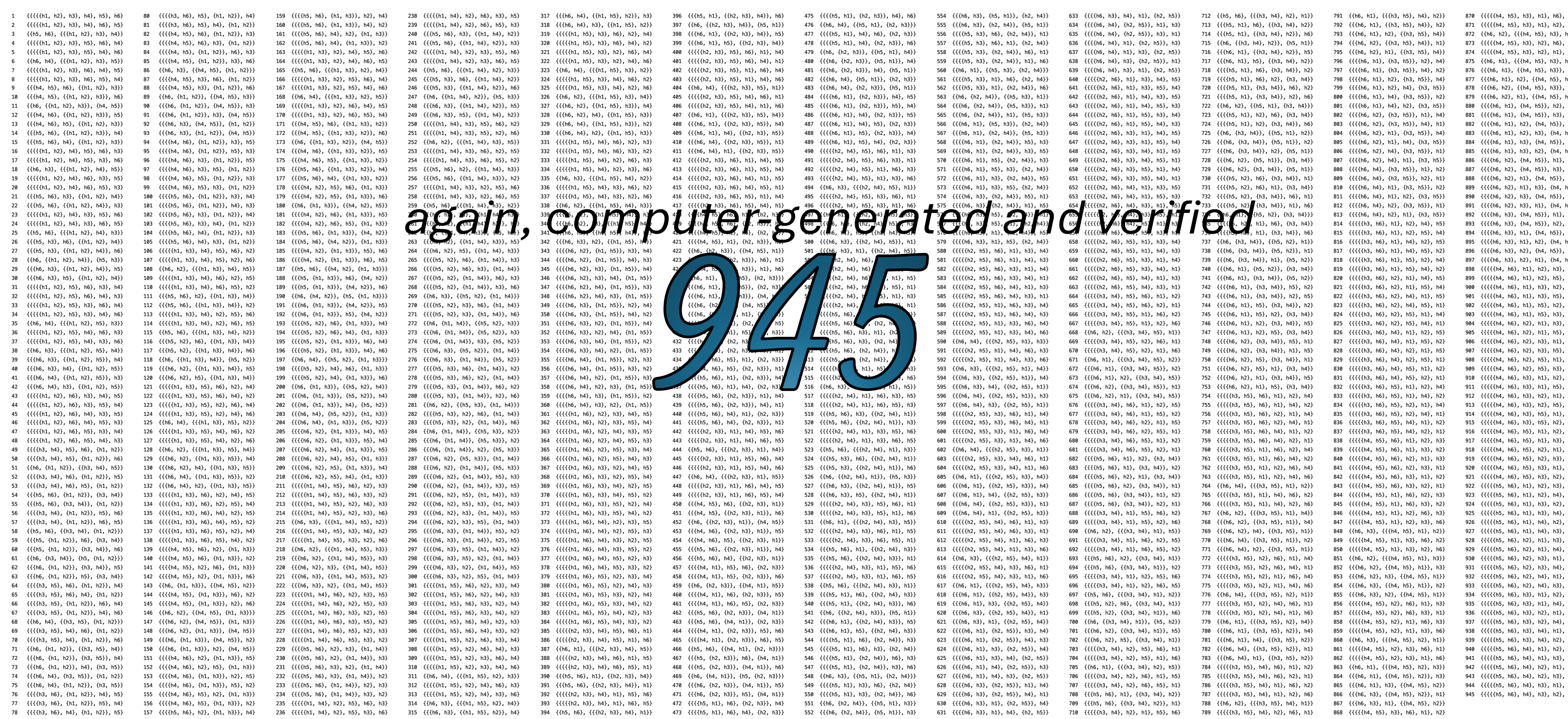
Merge Combinatorics: $WS_{init} = h_1 h_2 h_3 h_4 h_5 h_6$

- Given $WS_{init} = h_1 h_2 h_3 h_4 h_5 h_6$, converge on **945** cases:
  - $\{h_1, h_2\} + h_3 h_4 h_5 h_6$ = #c(|WS|=5) = 105
  - $\{h_1, h_3\} + h_2 h_4 h_5 h_6$ = #c(|WS|=5) = 105
  - $\{h_1, h_4\} + h_2 h_3 h_5 h_6$ = #c(|WS|=5) – #c($\{h_2, h_3\} h_{1,4} h_5 h_6$) = 105 – 15 = 90
  - $\{h_1, h_5\} + h_2 h_3 h_4 h_6$ = #c(|WS|=5) – 3 x #c($\{h_2, h_3\} h_{1,5} h_4 h_6$) = 105 – 3 x15 = 60
  - $\{h_1, h_6\} + h_2 h_3 h_4 h_5$ = 4! (*each h_2~h_5 must be singly Merged to* $\{h_1, h_6\}$)
  - $\{h_2, h_3\} + h_1 h_4 h_5 h_6$ = #c(|WS|=5) = 105
  - $\{h_2, h_4\} + h_1 h_2 h_5 h_6$ = #c(|WS|=5) – #c(|WS|=4) = 90
  - $\{h_2, h_5\} + h_1 h_2 h_3 h_6$ = #c(|WS|=5) – 3 x #c(|WS|=4) = 60
  - $\{h_2, h_6\} + h_1 h_3 h_4 h_5$ = 24
  - $\{h_3, h_4\} + h_1 h_2 h_5 h_6$ = 90
  - $\{h_3, h_5\} + h_1 h_2 h_4 h_6$ = 60
  - $\{h_3, h_6\} + h_1 h_2 h_4 h_5$ = 24
  - $\{h_4, h_5\} + h_1 h_2 h_3 h_6$ = 60
  - $\{h_4, h_6\} + h_1 h_2 h_3 h_5$= 24
  - $\{h_5, h_6\} + h_1 h_2 h_3 h_4$= 24
  - Total: (105+105 + 90 + 60+ 24) + (105 + 90 +60 + 24) + (90 + 60 + 24) + (60 + 24) + 24
  - = 3 x 105 + 3 x 90 + 4 x 60 + 5 x 24 = **945**

Pair (y,x)	2	3	4	5	6
1	105	105	90	60	24
2		105	90	60	24
3			90	60	24
4				60	24
5					24

Total: 945

Merge Combinatorics: $WS_{init} = h_1\ h_2\ h_3\ h_4\ h_5\ h_6$



Merge Combinatorics: $WS_{init} = h_1\ h_2\ h_3\ h_4\ h_5\ h_6\ h_7$

Top row of table (*transposed*), *n*=8

- $h_{1,2}\ h_3 \sim h_7$ = #c(|WS|=6) = 945
- $h_{1,3}\ h_2\ h_4 \sim h_7$ = 945
- $h_{1,4}\ h_2\ h_3\ h_5 \sim h_7$ = 945 – #c(|WS|=5) = 945 – 105 = 840
- $h_{1,5}\ h_2 \sim h_4\ h_6\ h_7$ = 945 – 3 x #c(|WS|=5) = 630
- $h_{1,6}\ h_2 \sim h_5\ h_7$ =945 – ${}_4C_2$ x #c(|WS|=5) + ${}_4C_2$ x #c(|WS|=4) / 2 = 945 - 6 x 105 + 6 x 15 / 2 = 360
- $h_{1,7}\ h_2 \sim h_6$ = 5! = 120

Why?

- In $h_2 \sim h_5$, ${}_4C_2$ = 6, e.g. pick redundant pair, e.g. $\{h_2, h_3\}$, to cancel all combinations with this, but will over-cancel as included $\{h_2, h_3\}\ h_4\ h_5$ will generate also $\{h_2, h_3\}\ \{h_4, h_5\}$.
- Double pair is symmetrically available from $\{h_4, h_5\}\ h_2\ h_3$.
- With the double pair |WS|=4, e.g. $h_{1,6}\{h_2, h_3\}\{h_4, h_5\}\ h_7$.
- ∴ as correction, add back in half of those.

Pair (y,x)	2	3	4	5	6	7
1	945	945	840	630	360	120
2		945	840	630	360	120
3			840	630	360	120
4				630	360	120
5					360	120
6						120
Total						10,395

Merge Combinatorics: $WS_{init} = h_1\ h_2\ h_3\ h_4\ h_5\ h_6\ h_7\ h_8$

Top row of table (*transposed*), *n*=8, *k*=x-axis pair:

8	135135						
Pair (y,x)	2	3	4	5	6	7	8
1	10395	10395	9450	7560	5040	2520	720
2		10395	9450	7560	5040	2520	720
3			9450	7560	5040	2520	720
4				7560	5040	2520	720
5					5040	2520	720
6						2520	720
7							720

- $\{h_1, h_2\}$ #c(|WS|=n-1) = 10395
- $\{h_1, h_3\}$ #c(|WS|=n-1)
- $\{h_1, h_4\}$ #c(|WS|=n-1) - #c(|WS|=n-2) h_2, h_3
- $\{h_1, h_5\}$ #c(|WS|=n-1) – ${}_{k-2}C_2$ #c(|WS|=n-2) h_2~h_4
- $\{h_1, h_6\}$ #c(|WS|=n-1) – ${}_{k-2}C_2$ #c(|WS|=n-2) h_2~h_5 + ${}_{k-2}C_2$ #c(|WS|=n-3)/2 $\{h_2, h_3\}\{h_4, h_5\}$
- $\{h_1, h_7\}$ #c(|WS|=n-1) – ${}_{k-2}C_2$ #c(|WS|=n-2) h_2~h_6 + ${}_{k-2}C_2$ #c(|WS|=n-3) ${}_{k-4}C_2$ /2 $\{h_2, h_3\}\{h_4, h_5\}$ h_6
- $\{h_1, h_8\}$ (n-2)! = 720 h_2~h_7

Why is ${}_{k-4}C_2$ /2 correct? Each ${}_{k-2}C_2$ pick of $\{h_2, h_3\}$ can generate 3 (k=7) additional redundant pairs from h_4~h_6. These pairs are $\{h_4, h_5\}$, $\{h_4, h_6\}$ and $\{h_5, h_6\}$. And each are also generated by symmetry, e.g. $\{h_2, h_3\}\{h_4, h_5\}$ h_6 is also generated as $\{h_4, h_5\}\{h_2, h_3\}$ h_6 by picking $\{h_4, h_5\}$ first instead.
Why is #c(|WS|=5) correct? We have three pairs with the double redundancy, e.g. $\{h_1, h_7\}$, $\{h_2, h_3\}$ and $\{h_4, h_5\}$ with leftover h_6 plus h_8. Total 5 items in the WS.

Merge Combinatorics: $WS_{init} = h_1\ h_2\ h_3\ h_4\ h_5\ h_6\ h_7\ h_8\ h_9$

9	2027025							
Pair (y,x)	2	3	4	5	6	7	8	9
1	135135	135135	124740	103950	75600	45360	20160	5040
2		135135	124740	103950	75600	45360	20160	5040
3			124740	103950	75600	45360	20160	5040
4				103950	75600	45360	20160	5040
5					75600	45360	20160	5040
6						45360	20160	5040
7							20160	5040
8								5040

Top row (transposed), *n*=9, *k*=x-axis pair:

- $\{h_1, h_2\}$ #c(|WS|=n-1) = 135135
- $\{h_1, h_3\}$ #c(|WS|=n-1)
- $\{h_1, h_4\}$ #c(|WS|=n-1) – #c(|WS|=n-2) h_2, h_3
- $\{h_1, h_5\}$ #c(|WS|=n-1) – ${}_{k-2}C_2$ #c(|WS|=n-2) $h_2 \sim h_4$
- $\{h_1, h_6\}$ – ${}_{k-2}C_2$ #c(|WS|=n-2) $h_2 \sim h_5$ + ${}_{k-2}C_2$ #c(|WS|=n-3)/2 $\{h_2, h_3\}\{h_4, h_5\}$
- $\{h_1, h_7\}$ – ${}_{k-2}C_2$ #c(|WS|=n-2) $h_2 \sim h_6$ + ${}_{k-2}C_2$ #c(|WS|=n-3) ${}_{k-4}C_2$/2 $\{h_2, h_3\}\{h_4, h_5\}$
- $\{h_1, h_8\}$ – ${}_{k-2}C_2$ #c(|WS|=n-2) $h_2 \sim h_7$ + ${}_{k-2}C_2$ #c(|WS|=n-3) ${}_{k-4}C_2$/2 $\{h_2, h_3\}\{h_4, h_5\}$ – ${}_{k-4}C_2$ #c(|WS|=n-4) $\{h_2, h_3\}\{h_2, h_3\}\{h_6, h_7\}$
- $\{h_1, h_9\}$ (n-2)! = 5040 $h_2 \sim h_7$

correct overcounting

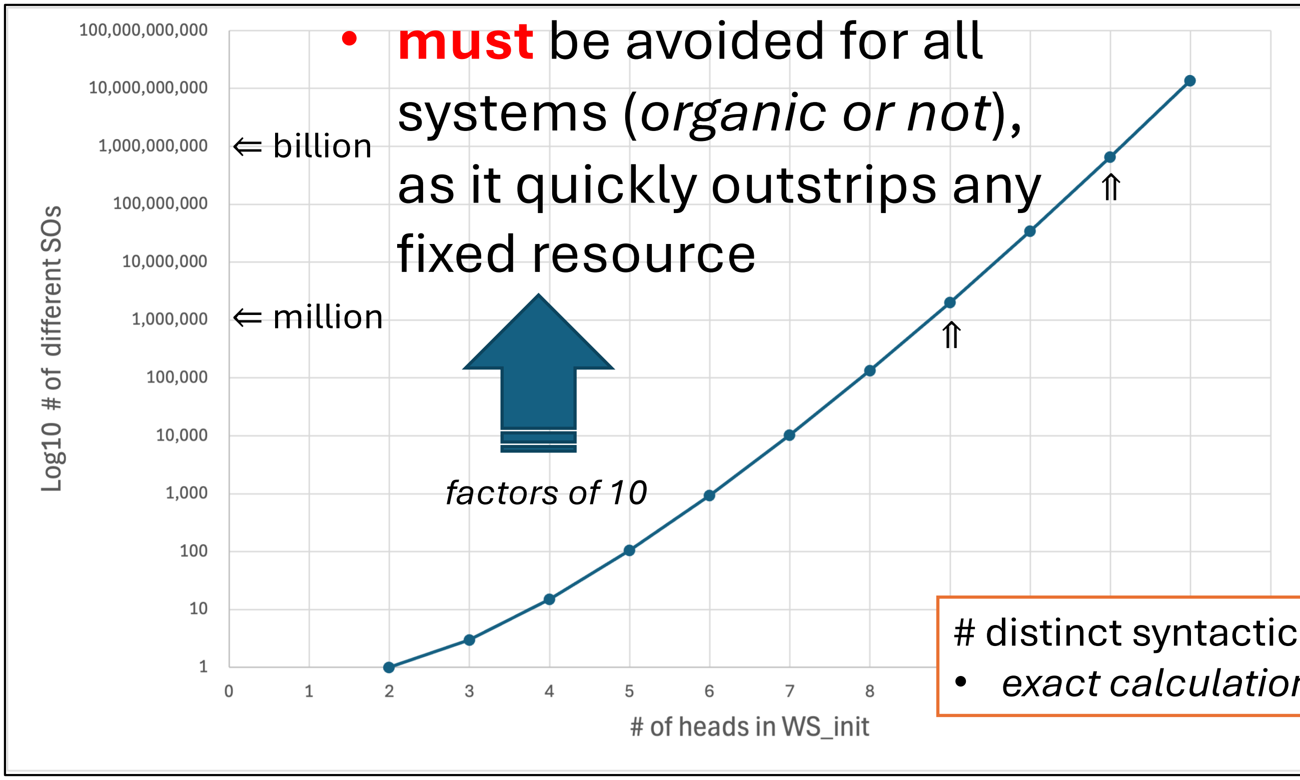
correct overcounting correction

correct correction to overcounting correction

Merge Combinatorics: upper triangles

	A	B	C	D	E	F	G	H	I
1	9	**2027025**							
2	Pair (y,x)	2	3	4	5	6	7	8	9
3	1	135135	135135	124740	103950	75600	45360	20160	5040
4	2		135135	124740	103950	75600	45360	20160	5040
5	3			124740	103950	75600	45360	20160	5040
6	4				103950	75600	45360	20160	5040
7	5					75600	45360	20160	5040
8	6						45360	20160	5040
9	7							20160	5040
10	8								5040

Merge Combinatorics: Summary



Heads	#SOs
2	1
3	3
4	15
5	105
6	945
7	10,395
8	135,135
9	2,027,025
10	34,459,425
11	654,729,075
12	13,749,310,575
13	316,234,143,225

distinct syntactic objects for External Merge

- *exact calculation, verified by computer*

Computational Complexity of Merge

- Merge as a mathematical abstraction
  - not feasible, e.g. as a generate-and-test model
  - biologically implausible
  - in fact, implausible for any real computational system
- Merge as applied to I-Language

Computational Complexity of Merge

- Merge as a mathematical abstraction
  - not feasible, e.g. as a generate-and-test model
- Merge as applied to I-Language
  - Language Organ Specific constraints
    - limit the complexity of Merge
  - **LSC**, e.g. (Chomsky 2021)
  - Theta theory (θ-roles and predicate heads)
  - functional section (verbal projection: INFL, v)
  - other 3rd Factor considerations, e.g. Nature & computational limits and optimization

I-Language Merge: θ-driven

- Chomsky (p.c.):
  - **Theta positions are detectable everywhere**
  - Conversation goes:
    - Well, there are no marking for IM (Internal Merge) vs. EM (External Merge).
    - INT reads the computed structure and determines how to interpret identical inscriptions.
    - **That's true, but it doesn't mean that IM can't observe theta theory (and duality ...), crashing and hence cancelling the preferred derivation.**
- (Chomsky 2024):
  - [T] All relations and structure-building operations (SBO) are **thought-related**, with semantic properties interpreted at CI.
- Merge is θ-aware & θ-driven:
  - (External) Merge builds θ-configurations efficiently
  - *i.e as early and quickly as possible*

I-Language Merge: selection-driven

[pg.132, (Chomsky 2000)]

- (53) *Properties of the probe/selector a must be satisfied before new elements of the lexical subarray are accessed to drive further operations*.

- Example:
  - head INFL triggers (Internal) Search for a θ-relevant item
    - *pronounced at its left edge as the surface subject in English*
    - {INFL$_{\phi}$, {v$_{pres}$, {arrive, train$_{a}$}}} ⇒ {train$_{a}${INFL$_{\phi}$, {v$_{pres}$, {arrive, ~~train~~$_{a}$}}}}
    - {INFL$_{\phi}$, {John, {v$_{past}$, {meet, Mary}}}} ⇒ {John, {INFL$_{\phi}$, {~~John~~, {v$_{past}$, {meet, Mary}}}}}
  - [Interesting question: *there*-insertion]

Part 2: Parsing

- From E-Language to I-Language
- **Why should we study this?**
  - *Well, we've been analyzing examples from in Part 1 ...*
  - *can* eagles that fly *swim*?
  - *the *bombing*$_{sg}$ of the cities$_{pl}$ ***were***$_{sg}$ criminal
  - the mechanic who *fixed* the car *carefully packed* his tools

What you've been doing is **parsing**!

Communication and Thought

- **Language** organ is designed to construct thoughts efficiently
- **Language** is not designed for efficient communication

too bad. Nature doesn't care. [pg.11, (Chomsky 2021)]

- *If that makes expressions harder to process and even makes some thoughts impossible to express without circumlocution, too bad. Nature doesn't care.* [pg.11, (Chomsky 2021)]
  - **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)]

a current research topic for me!

It makes no sense to say that *some system evolved for X*
"the spine evolved for keeping us upright," or "language evolved for communication"

Perception and Parsing

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

Aside: Phrase structure parsing

- Computer Science:
  - computer stack (BURY/UNBURY), Turing (1945)
  - Cocke–Younger–Kasami (CKY) algorithm, Sa 96
  - LR(*k*) parsing, (D)PDA discovery, Knuth (1
  - Earley's algorithm for Context-Free Gra ar (1968)

- **Transformational Grammar**
  - *no viable algorithm exists*

Parsing vs. Internal Thought

- Operative Complexity less for Internal Thought
  - *Language is optimized for thought, not communication*
- No **Phases**
  - Chomsky *MI* (2000) assumes WSs 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]

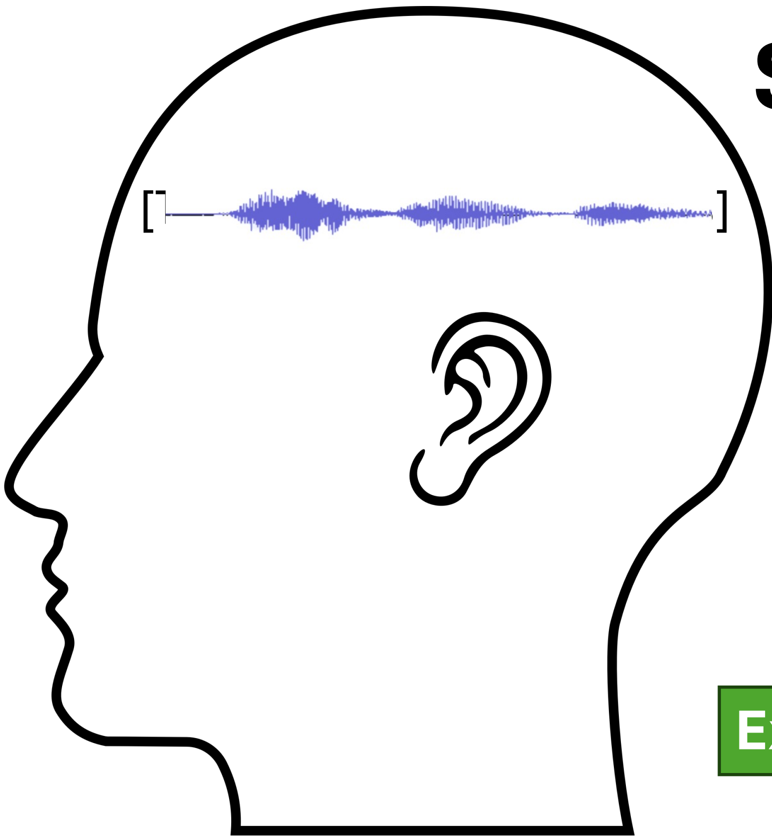
(27) a. $P_1 = [_{CP}$ that global warming must be taken seriously]
b. $P_2 = [_{CP}$ that glaciers are receding]
c. $P_3 = [_{vP}$ [the demonstration P_2 [show P_1]]]

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

Sub-arrays reduce operative complexity

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 2021)]
    - *see solutions in the SMT Parser ...*
  - **EXT**: *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*."



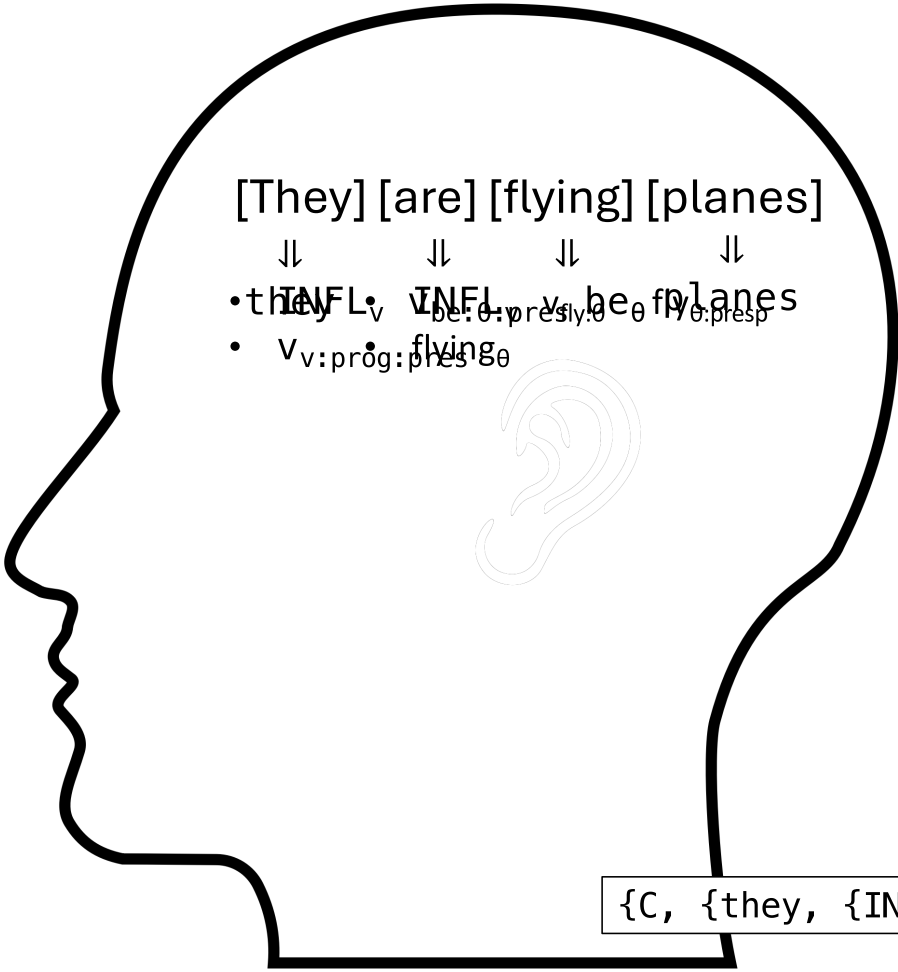
SMT Parser: *how it works*

[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."

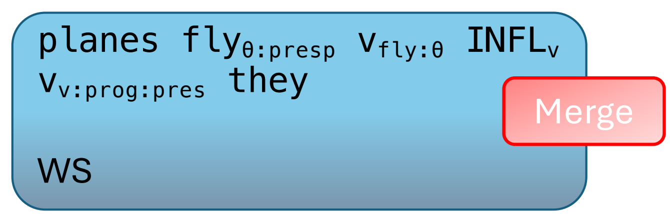
- ***they*** *– are – flying planes*
- ***they*** *– are flying – planes*

Examples: `sandiway.arizona.edu/smtparser`



How it works

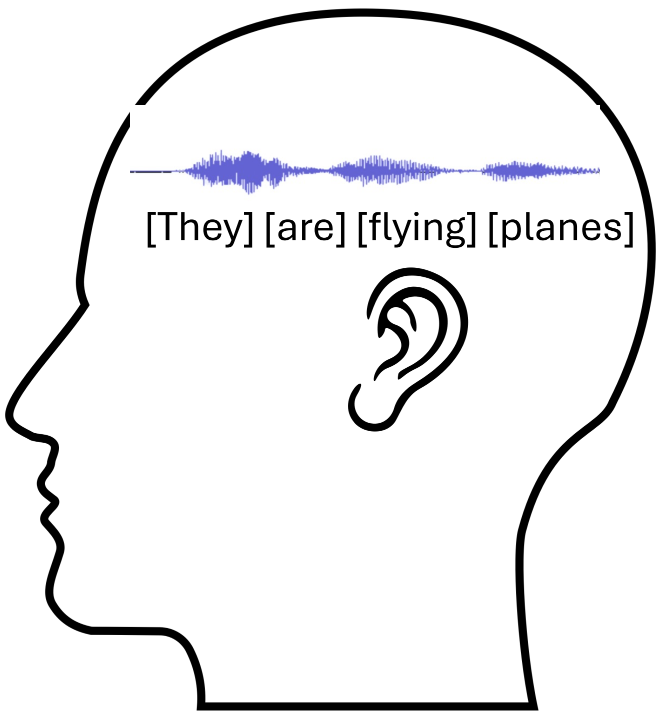
- **Parsing**:
  - recognize a word from the input signal
  - look it up in **LEX**
  - heads go in an Initial Workspace (WS_{init})



{C, {they, {INFL$_{v}$, {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}$_{7}$

- Merge fires!

How it works



Two workspaces (WS_{init})

1. `planes` $\texttt{fly}_{\theta:presp}$ $\texttt{v}_{fly:\theta}$ $\texttt{INFL}_{v}$ $\texttt{v}_{v:prog:pres}$ `they`
2. `planes` $\texttt{flying}_{\theta}$ $\texttt{be}_{\theta}$ $\texttt{v}_{be:\theta:pres}$ $\texttt{INFL}_{v}$ `they`

- *could be more …*

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

- **Derivation:**

1. planes $\text{fly}_{\theta:\text{presp}}$ $\text{v}_{\text{fly}:\theta}$ INFL_{v} $\text{v}_{\text{v:prog:pres}}$ they
2. {$\text{fly}_{\theta:\text{presp}}$, planes} $\text{v}_{\text{fly}:\theta}$ INFL_{v} $\text{v}_{\text{v:prog:pres}}$ they
3. {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, planes}} INFL_{v} $\text{v}_{\text{v:prog:pres}}$ they
4. {they, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, planes}}} INFL_{v} $\text{v}_{\text{v:prog:pres}}$
5. {$\text{v}_{\text{v:prog:pres}}$, {they, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, planes}}}} INFL_{v}
6. {INFL_{v}, {$\text{v}_{\text{v:prog:pres}}$, {**they**, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, planes}}}}}
7. {**they**, {INFL_{v}, {$\text{v}_{\text{v:prog:pres}}$, {they, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, planes}}}}}}
8. {**C**, {they, {INFL_{v}, {$\text{v}_{\text{v:prog:pres}}$, {they, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, $\text{planes}\}_7$
9. {C, {they, {INFL_{v}, {$\text{v}_{\text{v:prog:pres}}$, {they, {$\text{v}_{\text{fly}:\theta}$, {$\text{fly}_{\theta:\text{presp}}$, $\text{planes}\}_7$
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:**

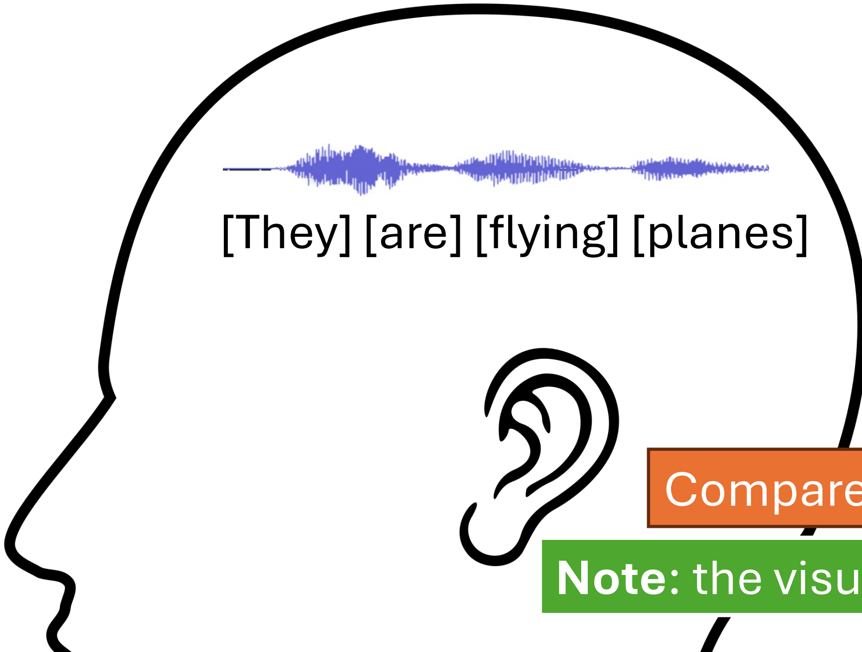
1. planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_{v}$ v$_{v:prog:pres}$ they
2. {fly$_{\theta:presp}$, planes} v$_{fly:\theta}$ INFL$_{v}$ v$_{v:prog:pres}$ they
3. {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}} INFL$_{v}$ v$_{v:prog:pres}$ they
4. {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}}} INFL$_{v}$ v$_{v:prog:pres}$
5. {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}}}} INFL$_{v}$
6. {INFL$_{v}$, {v$_{v:prog:pres}$, {**they**, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}}}}}
7. {**they**, {INFL$_{v}$, {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}}}}}}
8. {C, {they, {INFL$_{v}$, {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}$_{7}$
9. {C, {they, {INFL$_{v}$, {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}$_{7}$
10. they 3pl pres. be flying planes
11. they are flying planes

θ-configuration

Merge output: *converged*

FormCopy

Linear Spellout



How it works

[pg.118, Chomsky (1956)]

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

Compare the EXT output with what you originally heard

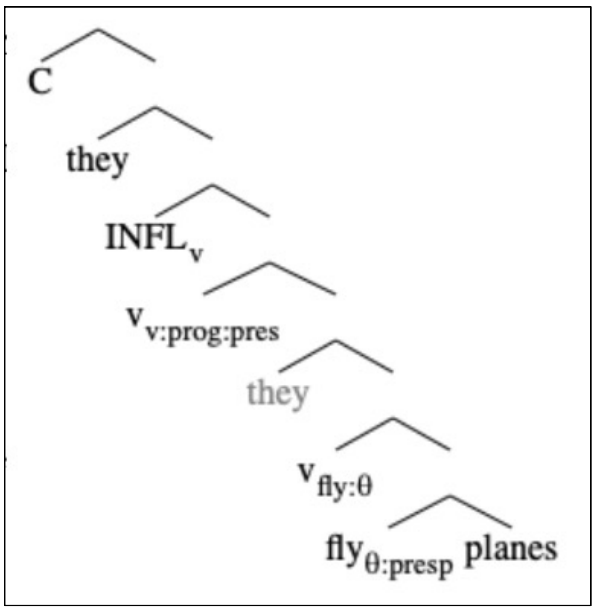
Note: the visual system can also exhibit parsing ambiguity

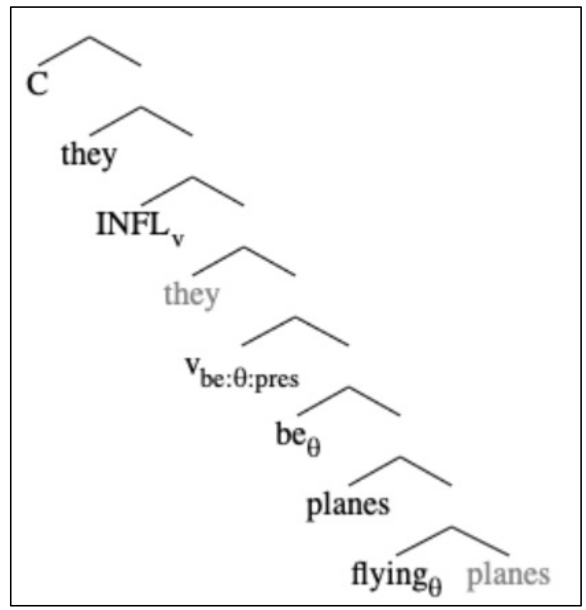
1.	{C,	{they,	{INFL$_v$,	{v$_{v:prog:pres}$,	{they,	{v$_{fly:\theta}$,	{fly$_{\theta:presp}$,	planes}$_7$
		they	3pl	pres. be			flying	planes
		they		are			flying	planes

2.	{C,	{they,	{INFL$_v$,	{they,	{v$_{be:\theta:pres}$,	{be$_\theta$,	{{flying$_\theta$,	planes},	planes}$_8$
		they	3pl		pres.	be	flying		planes
		they				are	flying		planes

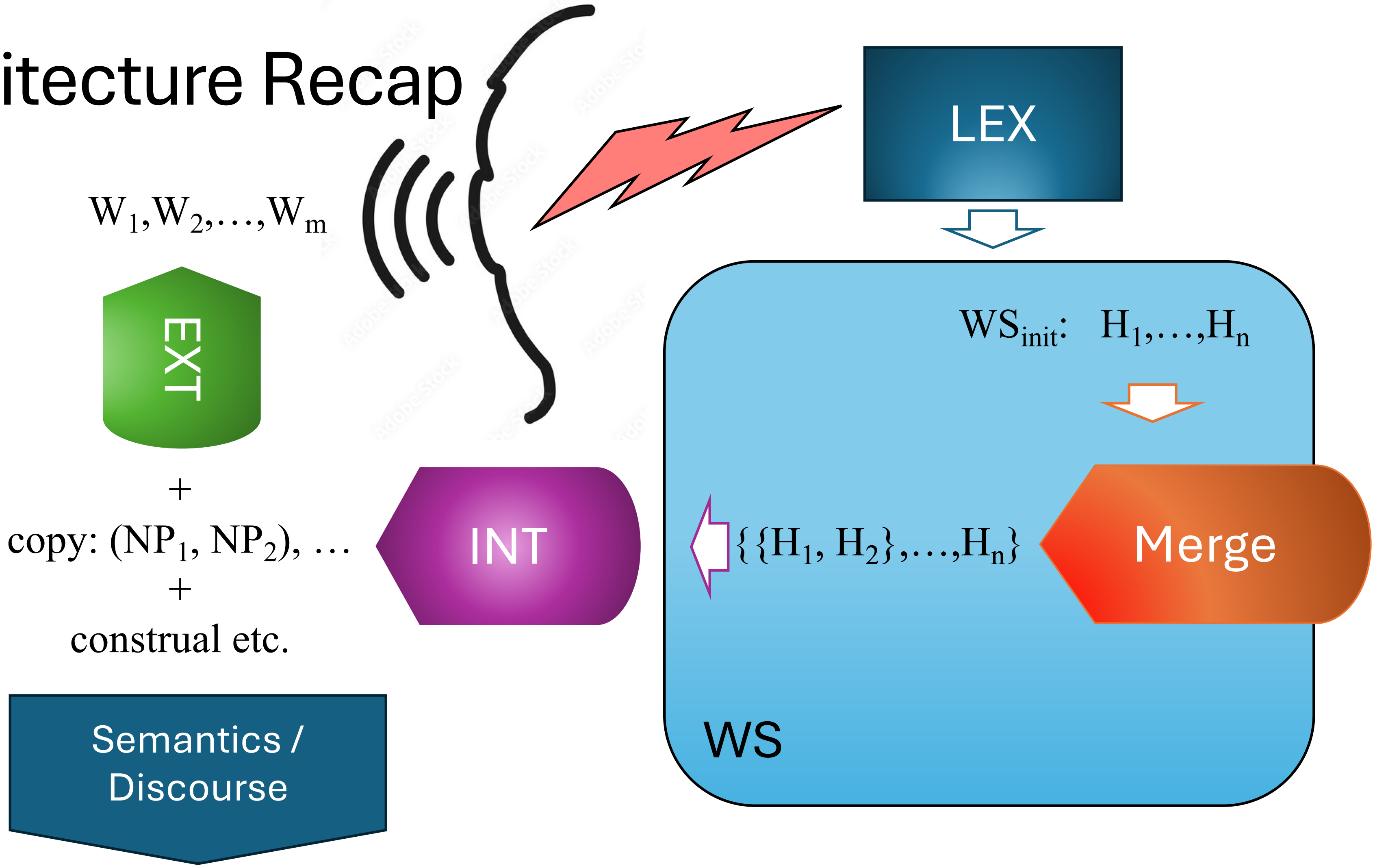
SMT Parser

1. {C, {they, {INFL$_{v}$, {v$_{v:prog:pres}$, {they, {v$_{fly:\theta}$, {fly$_{\theta:presp}$, planes}}}}}}}
2. {C, {they, {INFL$_{v}$, {they, {v$_{be:\theta:pres}$, {be$_{\theta}$, {planes, {flying$_{\theta}$, planes}}}}}}}}

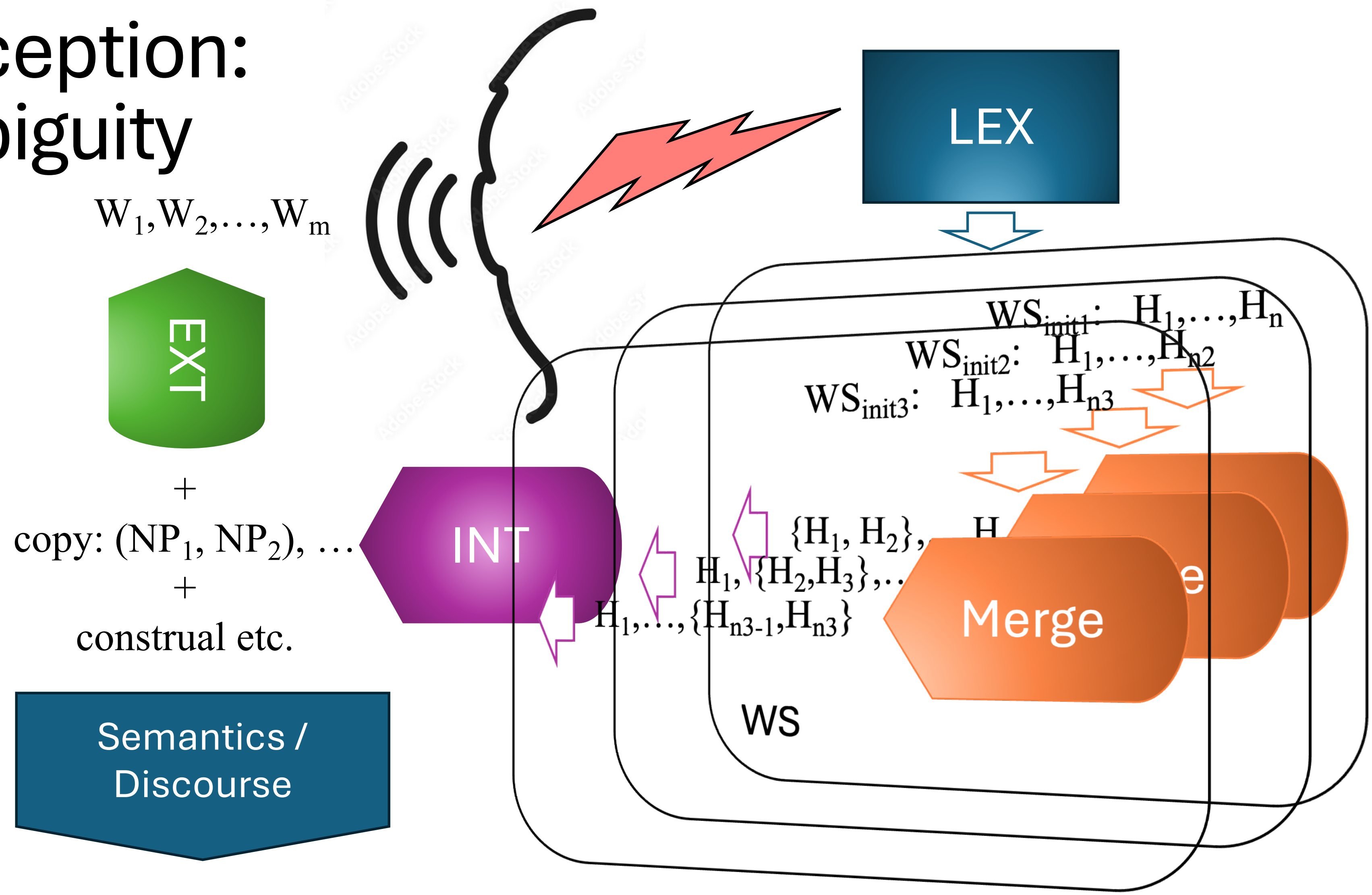




Architecture Recap



Perception: ambiguity



Jokes

- Many jokes are based on the human parser reflexively computing 2 parses
- 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*

SMT Parser

sandiway.arizona.edu/smtparser/flying_planes.html

Hand-built LEX

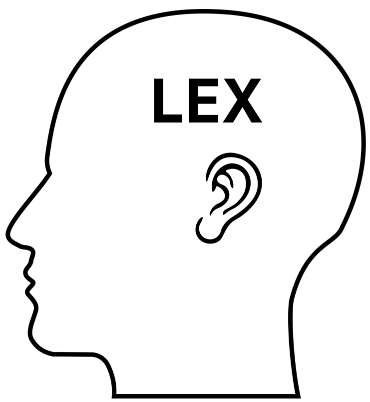
Words: they are flying planes
- **Initial WS 1:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 2:** planes flying$_\theta$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 3:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 4:** planes flying$_\theta$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 5:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{v:prog:pres}$ they
- **Initial WS 6:** planes flying$_\theta$ v$_{v:prog:pres}$ they
- **Initial WS 7:** planes flying$_\theta$ v$_{v:pass:pres}$ they

WordNet LEX (nltk)

Words: they are flying planes
- **Initial WS 1:** planes flying are they
- **Initial WS 2:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying are they
- **Initial WS 3:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ are they
- **Initial WS 4:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ are they
- **Initial WS 5:** planes flying$_\theta$ are they
- **Initial WS 6:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying$_\theta$ are they
- **Initial WS 7:** planes flying v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 8:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 9:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 10:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 11:** planes flying$_\theta$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 12:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying$_\theta$ v$_{pred:pres}$ INFL$_v$ they
- **Initial WS 13:** planes flying be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 14:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 15:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 16:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 17:** planes flying$_\theta$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 18:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ flying$_\theta$ be$_\theta$ v$_{be:\theta:pres}$ INFL$_v$ they
- **Initial WS 19:** planes flying v$_{v:prog:pres}$ they
- **Initial WS 20:** planes fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{v:prog:pres}$ they
- **Initial WS 21:** plane$_\theta$ v$_{plane:\theta:pres}$ INFL$_{v:3sg}$ fly$_{\theta:presp}$ v$_{fly:\theta}$ INFL$_v$ v$_{v:prog:pres}$ they
- **Initial WS 22:** planes flying$_\theta$ v$_{v:prog:pres}$ they
- **Initial WS 23:** planes flying v$_{v:pass:pres}$ they
- **Initial WS 24:** planes flying$_\theta$ v$_{v:pass:pres}$ they

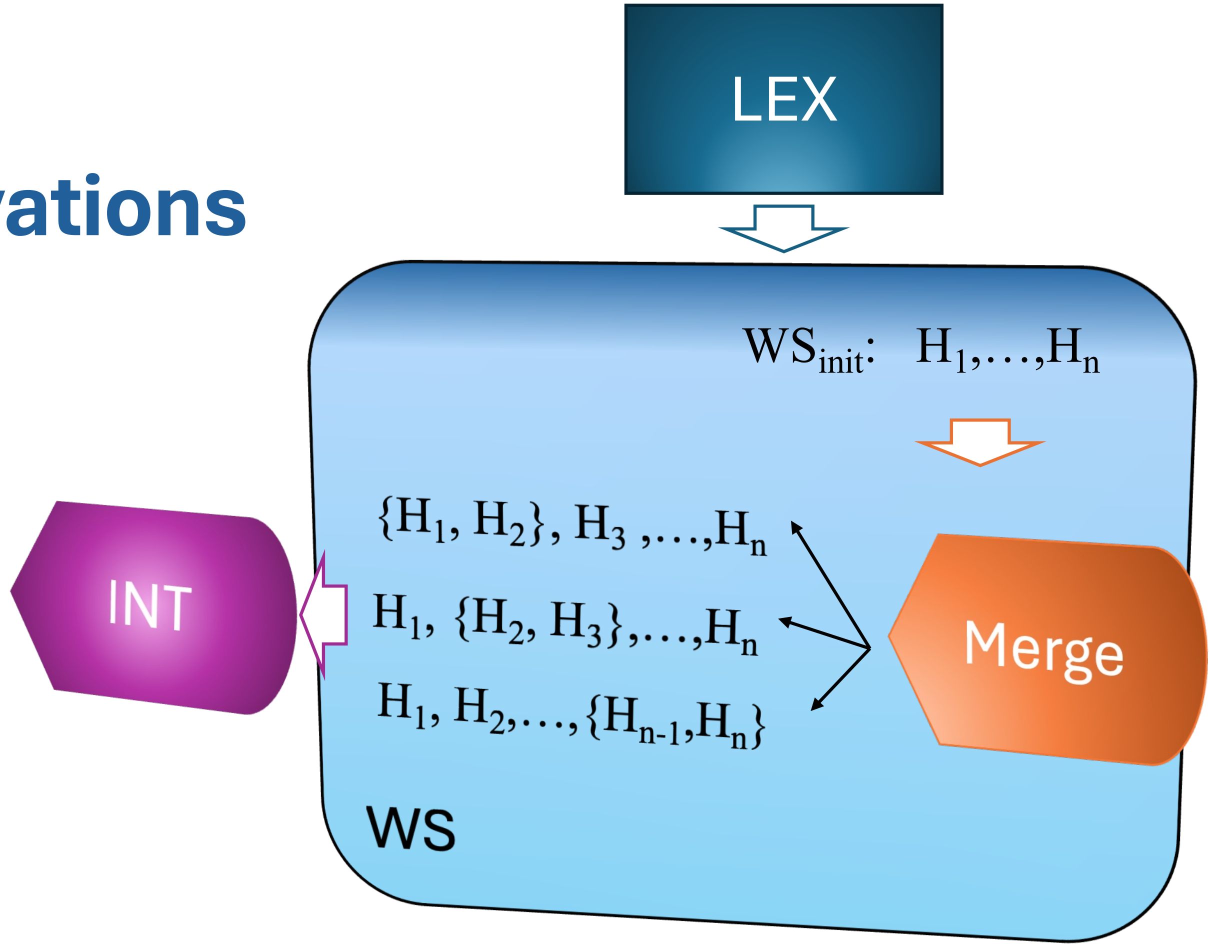
how many entries come to mind?

context, experience



One WS_{init}

multiple derivations



SMT Parser

Recall example:

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

Question: *two parses from one WS$_{init}$ or two?*

just one WS$_{init}$:

tools$_{his}$ pack$_{\theta}$ v$_{pack:\theta:pst}$ INFL$_{v}$ carefully$_{v}$ car$_{the}$ fix$_{\theta}$ v$_{fix:\theta:pst}$ INFL$_{v}$ C$_{relword(who)}$ mechanic$_{the}$

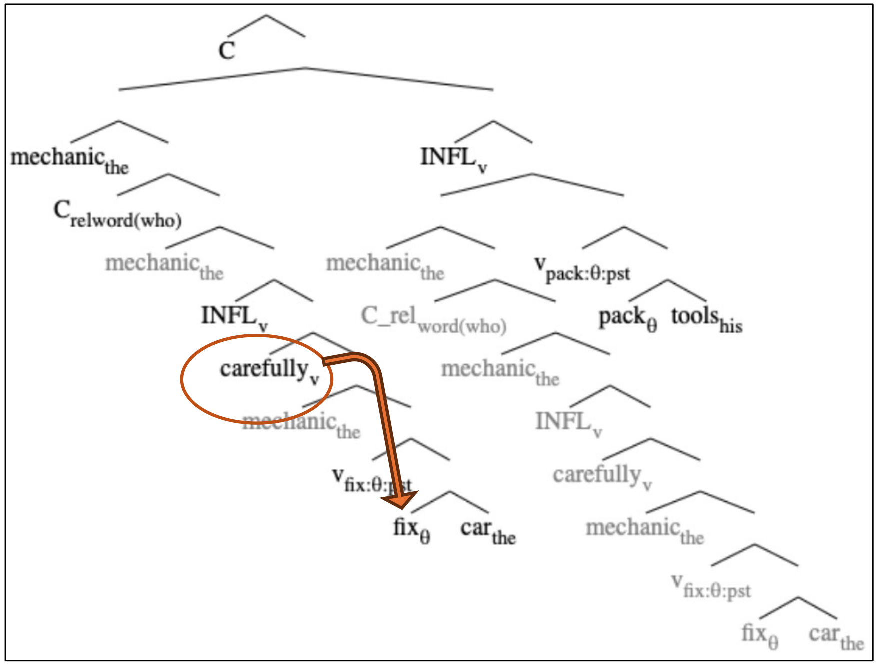
Parses:

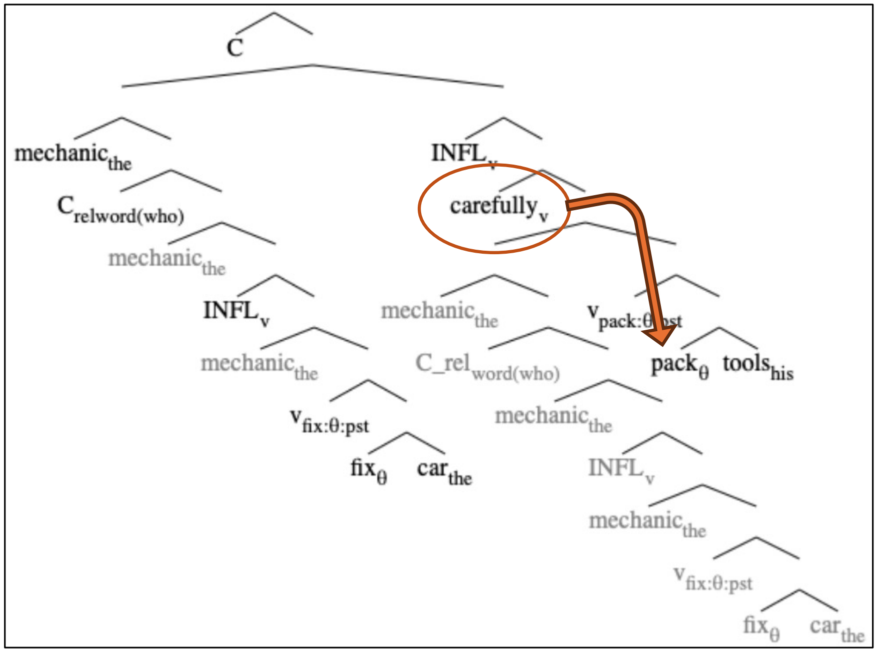
1. {C, {{mechanic$_{the}$, {C$_{relword(who)}$, {mechanic$_{the}$, {INFL$_{v}$, {**carefully**$_{v}$, {mechanic$_{the}$, {v$_{fix:\theta:pst}$, {**fix**$_{\theta}$, car$_{the}$}}}}}}}}, {INFL$_{v}$, {{mechanic$_{the}$,{C_rel$_{word(who)}$,{mechanic$_{the}$,{INFL$_{v}$,{carefully$_{v}$,{mechanic$_{the}$,{v$_{fix:\theta:pst}$,{fix$_{\theta}$,car$_{the}$}}}}}}}}, {v$_{pack:\theta:pst}$, {pack$_{\theta}$, tools$_{his}$}}}}}

2. {C, {{mechanic$_{the}$, {C$_{relword(who)}$, {mechanic$_{the}$, {INFL$_{v}$, {mechanic$_{the}$, {v$_{fix:\theta:pst}$, {fix$_{\theta}$, car$_{the}$}}}}}}}, {INFL$_{v}$, {**carefully**$_{v}$, {{mechanic$_{the}$,{C_rel$_{word(who)}$,{mechanic$_{the}$,{INFL$_{v}$,{mechanic$_{the}$,{v$_{fix:\theta:pst}$,{fix$_{\theta}$,car$_{the}$}}}}}}}, {v$_{pack:\theta:pst}$, {**pack**$_{\theta}$, tools$_{his}$}}}}}}}

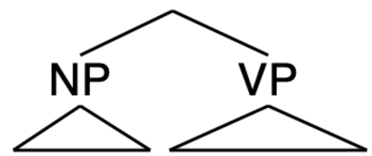
SMT Parser

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





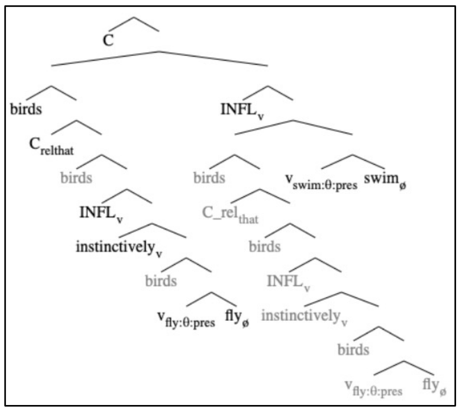
Parallelism

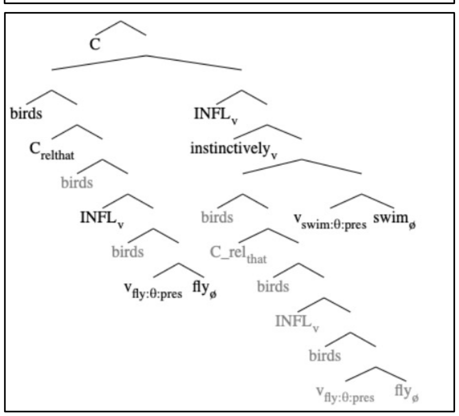


Repetitions exist in I-language because derivation is in parallel. Thus in an NP–VP structure, NP and VP are generated in parallel, with no interaction, and they might draw independently from the lexicon yielding structurally identical objects that are not copies, as in *John saw John*, with two independent occurrences of *John*. This is not a logical necessity. Evolution might have taken a different course, taking all identical inscriptions to be copies.

(Chomsky 2021)

SMT Parser

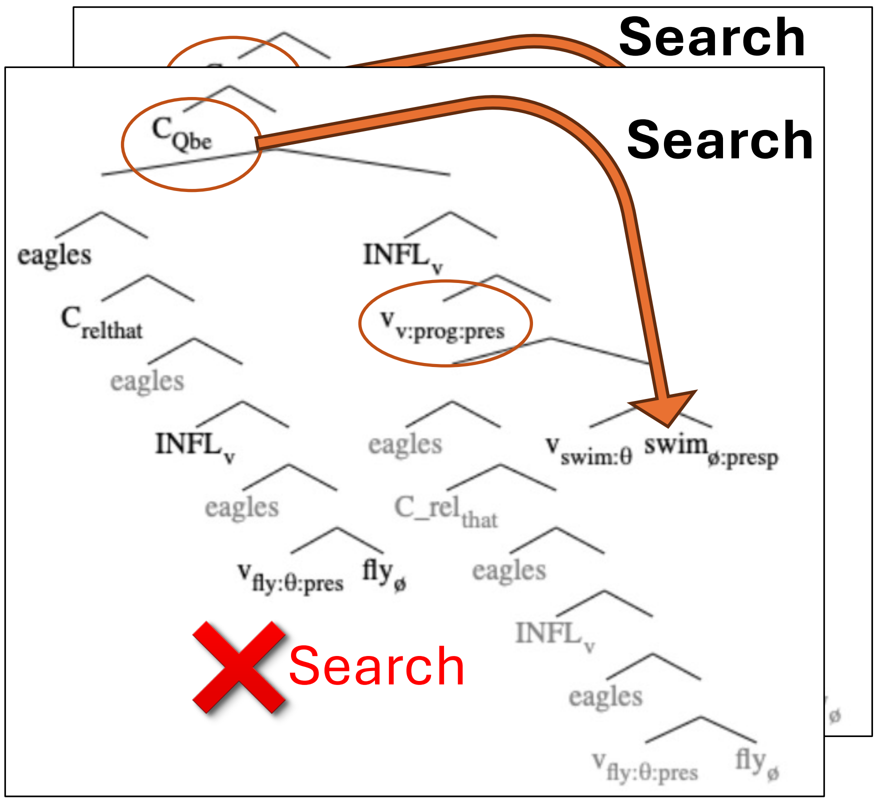




[pgs.8,103,117 (Berwick & Chomsky 2016)]

- 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*

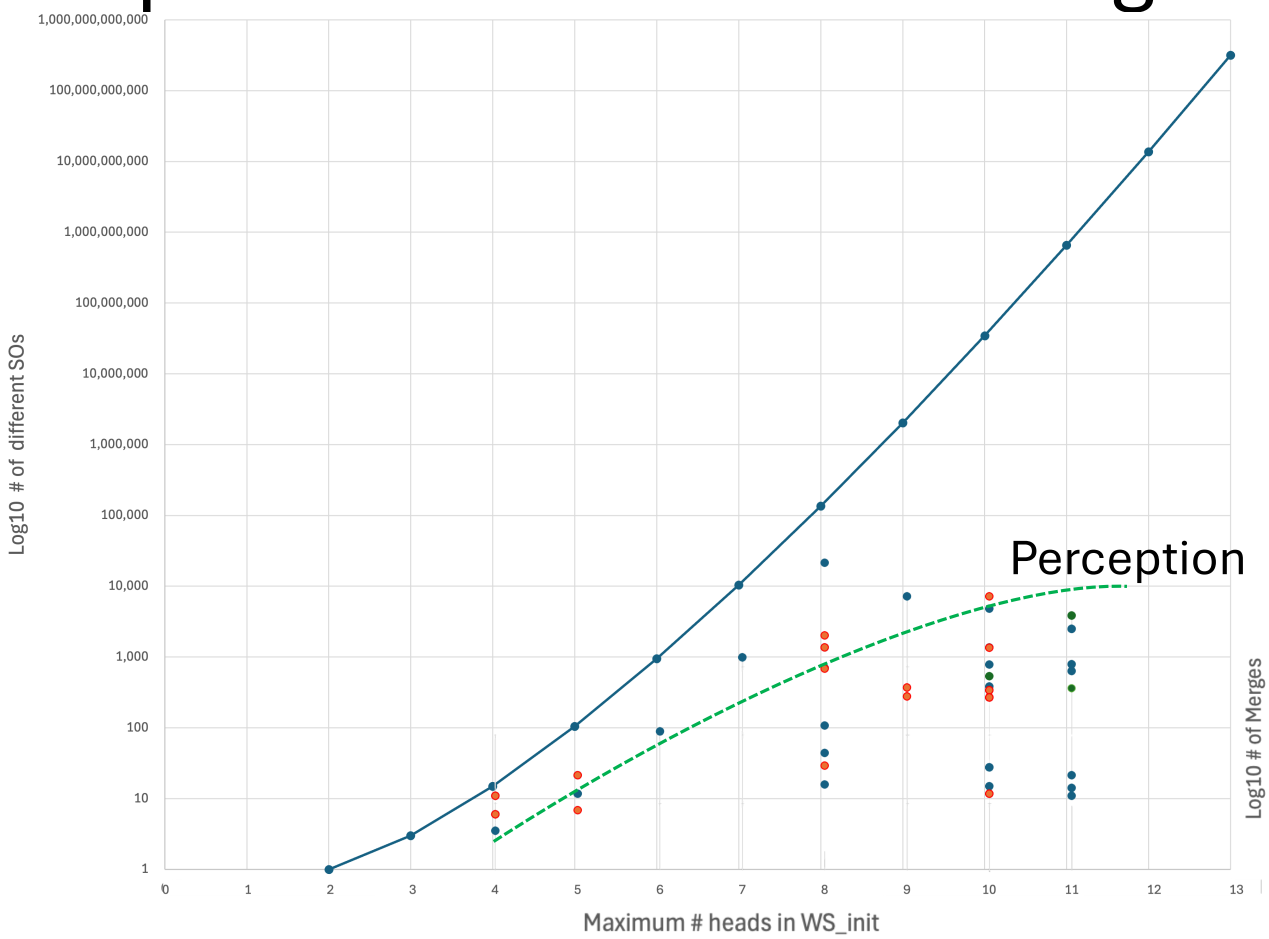
SMT Parser: interrogative C_Q probe



[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*]."

Computation: did we tame Merge?



Data:
61 examples on
`sandiway.arizona.edu/smtparser/`

Computation: did we tame Merge?

- **Operative complexity**:
  - I-Language Merge hugely better than Merge (*even for Perception*)
  - multiple WS_{init} for Parsing, single for Internal Thought
  - **Phases** (*aka* WS partitioning)
    - for Internal Thought, NOT for parsing (not described: *head clustering*)
- **Workspace Balancing**
  - wrt. θ-*seekers* and θ-*relevant* WS items
    - the problem of ***unpronounced items***

Repetitions and WS_{init}

- **Chomsky example**:
  - *the man who saw many people didn't see many people*
- Suppose we **minimize WS size**, create:
  `{v`$_{pst}$`, {see, many people}}  man`$_{the}$`  INFL  C`$_{rel}$` INFL  Neg  C`
- Construct relative clause, e.g.
  - `{man`$_{the}$`, {C`$_{rel}$`, {~~man~~`$_{the}$`, {INFL, {~~man~~`$_{the}$`, {v`$_{pst}$`, {see, many people}}}}}}`
- **Now stuck!**
  - would need to invent a new operation to *deep fish* and copy out `{v`$_{pst}$`, …}` ***SMT**
  - *Markovian assumption: *no reach back into Merge history*
  - ***Duality**: *only EM can introduce a theta role-bearing item*
  - computer science: table parsed phrases ***SMT**

Repetitions and Workspace θ-Balancing

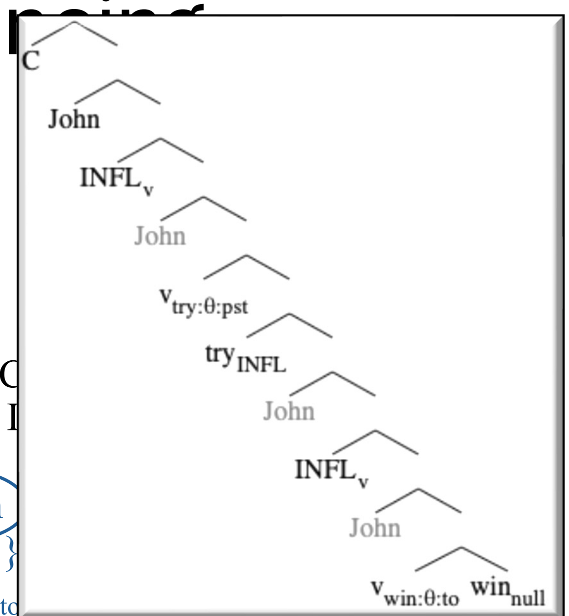
- Theta Theory informs and drives WS convergence:
  - for a derivation to converge, the number of θ-seekers and θ-relevant items must converge and balance out, i.e. arguments and θ-seekers must match up (with nothing left over in the WS).
- Example:
  - John wants to win
  - {C, {John, {$INFL_{v:\theta}$, {~~John~~, {$v_{want:\theta}$, {$want_{INFL}$, {~~John~~, {$INFL_{v:\theta}$, {~~John~~, {$v_{win:\theta}$, win}}}}}}}}}
- (Inner Thought) balanced WS_{init}:
  - $INFL_v$ $v_{win:\theta}$ win $INFL_v$ $v_{want:\theta}$ want 2×John
- (Perception) unbalanced WS_{init}:
  - C $INFL_{v:\theta}$ $v_{want:\theta}$ $want_{INFL}$ EA $INFL_{v:\theta}$ $v_{win:\theta}$ win (θ-*seekers*: $v_{want:\theta}$+ $v_{win:\theta}$; θ-*relevant*: EA)

Replicate Existing θ-relevant item

Repetitions and Workspace θ-Balan

- **Derivation** (*Perception*):

1. **WS$_{init}$:** win_{null} $v_{win:\theta:to}$ $INFL_v$ try_{INFL} $v_{try:\theta:pst}$ $INFL_v$ John C
2. {$v_{win:\theta:to}$, win_{null}} $INFL_v$ try_{INFL} $v_{try:\theta:pst}$ $INFL_v$ John C
3. {John, {$v_{win:\theta:to}$, win_{null}}} $INFL_v$ try_{INFL} $v_{try:\theta:pst}$ $INFL_v$ C
4. John {John, {$v_{win:\theta:to}$, win_{null}}} $INFL_v$ try_{INFL} $v_{try:\theta:pst}$ $INFL_v$ C
5. θ-balance $_v$, {John, {$v_{win:\theta:to}$, win_{null}}}}} John try_{INFL} $v_{try:\theta:pst}$ I
6. {try_{INFL} INFL$_v$ triggers Internal Search for θ-relevant term $_{y:\theta:pst}$ John
7. {$v_{try:\theta:pst}$,
8. {John, {$v_{try:\theta:pst}$, {try_{INFL}, {John, {$INFL_v$, {John, {$v_{win:\theta:to}$, win_{null}}}}}}}}
9. {John, {$INFL_v$, {John, {$v_{try:\theta:pst}$, {try_{INFL}, {John, {$INFL_v$, {John, {$v_{win:\theta:to}$
10. {C, {Jo INFL$_v$ triggers Internal Search for θ-relevant term n, {$v_{win:\theta:to}$, win_{null}}}}}}}}}}}



INT/EX Converged on a single syntactic object

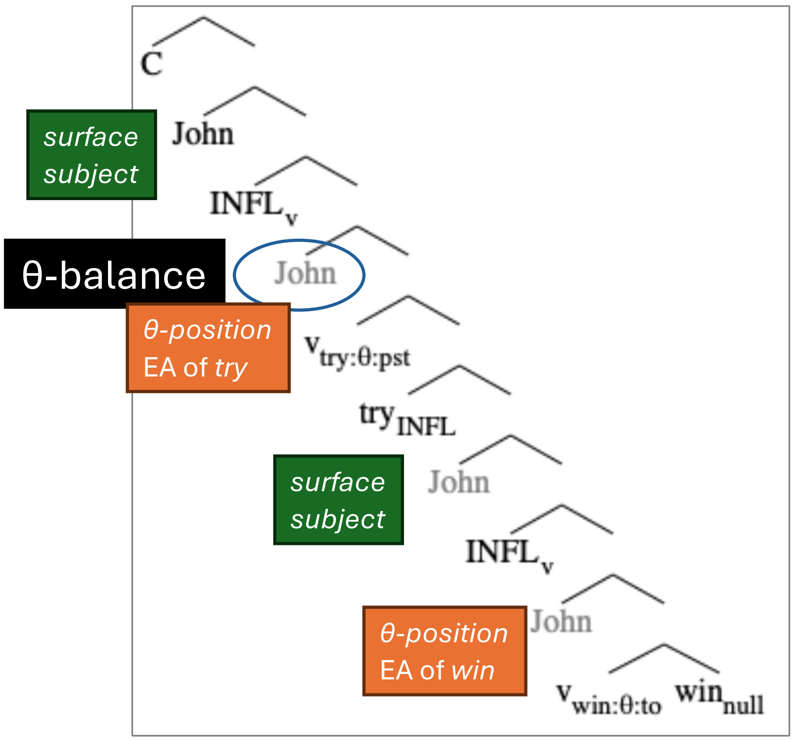
{C, {John, {$INFL_v$, {~~John~~, {$v_{try:\theta:pst}$, {try_{INFL}, {~~John~~, {$INFL_v$, {~~John~~, {$v_{win:\theta:to}$, win_{null}}}}}}}}}}

Spellout: identify copy relations: identical inscription c-command (*Phase limited*)

John tried to win

Amalgamation

Repetitions and Workspace θ-Balancing



- Introduced for Perception only
  - Inner Thought comes θ-balanced
- 4 positions for *John*
  - only one is pronounced
  - cf. *John saw John* / **John saw*

Words: John tried to win
▶ **Initial WS 1:** win$_{null}$ v$_{win:\theta:pres}$ INFL$_v$ try$_{INFL}$ v$_{try:\theta:pst}$ INFL$_v$ John
▶ **Initial WS 2:** win$_{null}$ v$_{win:\theta:to}$ INFL$_v$ try$_{INFL}$ v$_{try:\theta:pst}$ INFL$_v$ John

sandiway.arizona.edu/smtparser/try_win.html