3.1 Syntactic:
Structure Dependence

Maria Johnson
Prof. Lisa Pearl · LSCI 151C

The key setup

Pearl 2022: **"poverty of the stimulus** is just another way of saying that there's an **induction problem** for the child…
children nonetheless seem to reliably resolve that ambiguity, and
end up with the right answer despite the induction problem."

Today's case:

(a) *Can the penguin [who is on the iceberg] t_{can} find a fish ?*

(b) *The penguin [who is on the iceberg] can find a fish .*

To go from (b) → (a): which auxiliary moves, and why?

💬 *Idea from last class: the mind imposes hierarchical organization on a sequential signal or input. This is why structure-dependence even exists as a concept!*

The correct hypothesis & hypothesis space

✅ The correct hypothesis

"Move the **main clause** auxiliary"

Can [the penguin who is on the iceberg] find a fish?

Structure-dependent — relies on the hierarchical notion of "main clause".

The **overhypothesis**:
"Use structure-dependent rules"
(Scholz & Pullum 2006)

❌ Competing (wrong) hypotheses

Structure-independent rules that work for most child-directed speech:

"Move the first auxiliary"
"Move the last auxiliary"
"Move the 4th-from-end auxiliary"

→ *All compatible with simple sentences.*
→ *The data don't rule them out.*

💬 *Why would a structure-independent rule feel like a reasonable guess to a learner?*

The available data (or so we initially thought)

Initial assumption: Only **complex yes/no questions** were informative, and only **unambiguous** ones counted.
There were very few of those in child-directed speech (Pullum & Scholz 2002; Legate & Yang 2002).

What children actually hear:

"Is the cat happy?" → *compatible with structure-dependent rule AND "move first auxiliary"*

"Can the dog run?" → *compatible with structure-dependent rule AND "move first auxiliary"*

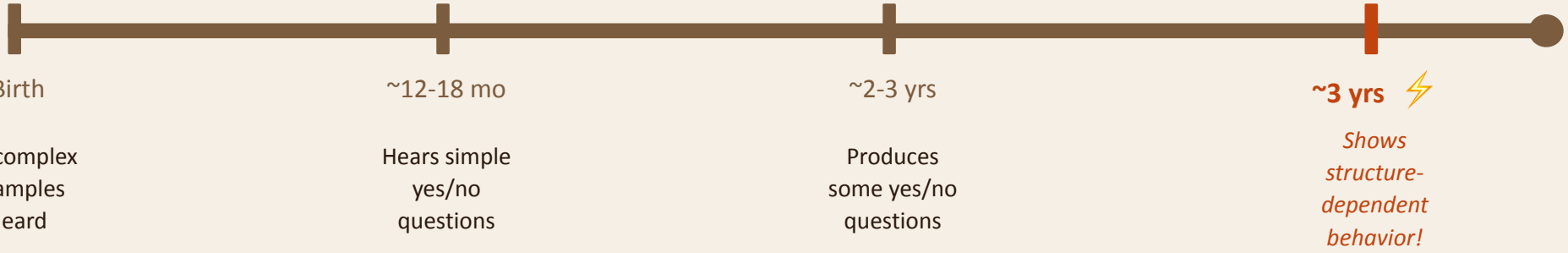
"Is the penguin who is happy dancing?" → *still ambiguous — "move first auxiliary" also gets this right*

"Can the penguin who is on the iceberg find a fish?" → **finally disambiguates! But children hear this type very rarely.**

"The data are compatible with more than one hypothesis — they don't pinpoint the correct hypothesis on their own." (Pearl 2021)

Age of acquisition for constrained generalizations

Crain & Nakayama (1987): English children as young as **3 years old** already know the rule must be structure-dependent. They
almost never produce structure-independent errors — even though unambiguous input is very rare.



The puzzle: Children succeed **quickly** and with very **little unambiguous input.**
They must have some kind of innate bias helping them — but what kind?
Language-specific (linguistic nativist) or **domain-general** (non-linguistic nativist)?

"…this structure-dependent knowledge seemed very much to be an instance of overcoming poverty of the stimulus." (Pearl 2021)

So…

what kind of bias do children need?

Language-specific (linguistic nativist)

or

Domain-general (non-linguistic nativist)

The investigations try to answer this.

What did the models test?

Approach	Basic idea	What it shows
Surface patterns	Can word-sequence statistics alone solve it?	**Not reliably — poverty of the stimulus persists**
Simplicity / Bayesian bias	Can preferring simpler, reusable structure help?	**Yes — a non-linguistic bias is sufficient**
Syntax + meaning together	Does connecting syntax to meaning guide the learner?	**Helpful — but some of the required biases may be linguistic**
Neural networks	Can the right behavior emerge from the right architecture?	**Sometimes — but we can't see inside. The bias question stays open.**

💬 *Do any of these approaches feel like they genuinely solve the problem — or do they all just push the question back one level?*

How to interpret these investigations

Pearl 2022: "Because some of the useful innate biases…seem to be linguistic, the **linguistic nativist viewpoint is supported**.
[But] if all the linguistic biases turn out to be derived from other non-linguistic biases, then the **non-linguistic nativist viewpoint would be supported**."

🌍 Probably non-linguistic

- View rules as part of the whole system
- **Simplicity bias** — prefer compact, reusable reps
- **Bayesian inference** — balance complexity vs. data fit

🧠 Possibly linguistic (open!)

- Which **structural building blocks** to use
- **The tight link between syntax and meaning**

→ If derivable from general biases → non-linguistic
→ If not → linguistic nativism

💬 ***Kevin:*** *"Is simplicity bias domain-general, or does its specific application to hierarchical syntax suggest an underlying linguistic template?"*

Thank you!