Syntactic islands: Pearl 2023 Modeling Syn Acq 3.2.3

Kevin Guzman
Prof. Pearl
LSCI 151C

The Core Problem

Defining wh-Dependencies and Structural Constraints

Defining wh-Dependencies

Long-Distance Gaps

Relationships between a wh-word and its understood position. These can be theoretically infinite in length.These dependencies are **structure-dependent** rather than just linear.

The Learning Challenge

If language allows theoretically infinite length, how do children discover that certain structural "islands" are forbidden without explicit negative feedback?

The Four Forbidden Islands

Whether Island: "What do you wonder whether Jack bought?"

Complex NP Island: "What did you make the claim that Jack bought?"

Subject Island: "What do you think the joke about was hilarious?"

Adjunct Island: "What do you worry if Jack buys?"

> *"How do children... acquire knowledge about the structure-dependence of linguistic representations, yet still allow acquisition to proceed successfully from the data children typically encounter?"*
>
> — Pearl & Sprouse (2013), Page 30

The Trigram Model

Probabilistic Learning from Child-Directed Speech

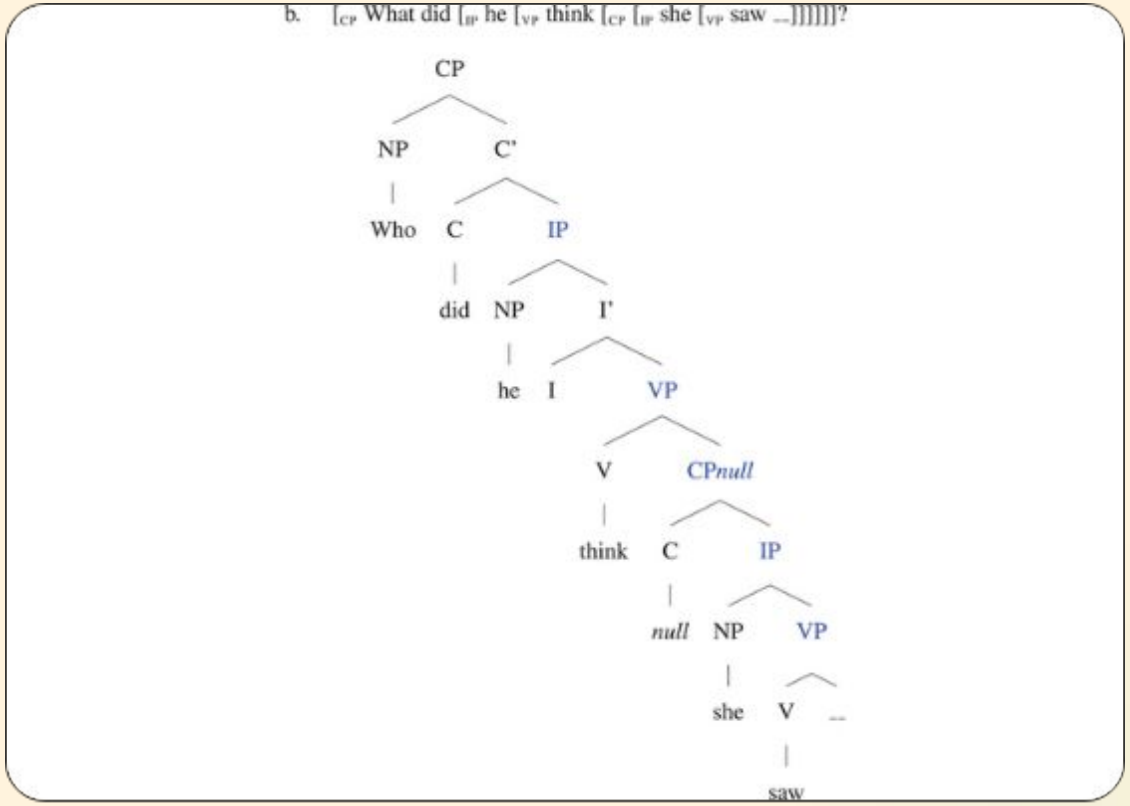
> *"The model tracks the frequencies of syntactic trigrams (triplets of phrase nodes) along a dependency path."*
>
> — *Conceptual Summary of the Learner Model*

Syntactic Trigrams

Container Nodes

The learner parses the sentence into a phrase structure tree and extracts a sequence of "container nodes" (e.g., CP, IP, VP) that dominate the gap.

The path is broken into **trigrams** (triplets of nodes). The learner tracks the frequency and probability of these overlapping segments.



The Probabilistic Strategy

1. Parse

Characterize the dependency as a sequence of phrase structure nodes.

2. Track

Identify and record the frequencies of container node trigrams.

3. Calculate

Multiply trigram probabilities to generate a total dependency probability.

How the Model "Judges" Grammar

Normal Sentence

Every 3-node chunk (trigram) is common in speech.

HIGH PROBABILITY

Result: Grammatical

A “Normal Sentence" is accepted as grammatical because every trigram is common in child-directed speech, which is the specific input used for the model.

Island Violation

Contains a chunk that the learner has never heard.

LOW PROBABILITY

Result: Ungrammatical

The Logic: Total Sentence Score = (Trigram A) × (Trigram B) × (Trigram C)

Cognitive Limitations

One hallmark of the syntax of human languages is the ability to have long-distance dependencies: relationships between two words in a sentence that are not adjacent to each other. Long-distance dependencies, such as the dependency between *what* and *stole* in (11a), can be potentially infinite in length.[6] However, there are specific syntactic structures that long-distance dependencies can't cross. These structures are known as *syntactic islands*. Four examples of syntactic islands are shown in (11b)-(11e) (Chomsky, 1965; Ross, 1967; Chomsky, 1973). During acquisition, children must infer the constraints on long-distance dependencies that allow them to recognize that the *wh*-dependencies in (11b)-(11e) are not allowed, while the *wh*-dependency in (11a) is fine.

(11) a. What does Jack think that Lily said that the goblins stole t_{what}?
b. *What do you wonder [whether Jack bought t_{what}]? (whether island)
c. *What did you make [the claim that Jack bought t_{what}]? (complex NP island)
d. *What do you think [the joke about t_{what}] was hilarious? (subject island)
e. *What do you worry [if Jack buys t_{what}]? (adjunct island)

[6] Of course, there's clearly an upper bound on the number of words and/or clauses that an English speaker can keep track of during language processing. However, this restriction appears to be based on the limited nature of human working memory capacity rather than an explicit structural restriction on the length of English *wh*-dependencies.

Is the difficulty of wh-dependencies a purely grammatical restriction, or a consequence of human processing limits? Pearl & Sprouse (2013) acknowledge that while they abstract away from memory for the model, real-time constraints impact how children encode and track these paths.

Footnote 6: Working Memory

Further Discussion Points

- **Nativism vs. Empiricism:** Does the model's success with "child-directed speech" prove island constraints are not innate?
- **Prior Knowledge:** The model requires the ability to parse hierarchical trees. Is this "prior knowledge" innate or learned?
- **Smoothing:** Is the model's use of a "smoothing constant" for unobserved items a biologically plausible mechanism?