SPRING 2025, CompCog Models of LangAcq

PEARL 2023 MODELING SYNTACTIC ACQUISITION SECTION 3.2.2

Seojin Hwang April 28 Mon

Structure Dependence

- Linguistic representations are hierarchically organized.
- Rules operate over structure, not simple linear word order.

Example of Structure Dependence

- English complex yes/no question formation
 - a. Structure-dependent rules: "Move the auxiliary in the main clause to the front."
- (10) a. [_{CP} Can the penguin [_{CP} who is on the iceberg] t_{can} find a fish]?
 b. [_{CP} The penguin [_{CP} who is on the iceberg] can find a fish].
 - b. Structure-independent rules: "Move the last auxiliary."
 - → unable to account for full range of English complex yes/no questions

Children's Early Knowledge

• From a developmental standpoint:

Children seem to know that syntactic rules should be structuredependent.

- Crain & Nakayama (1987) showed that children as young as three use structure-dependent rules.
- From an input standpoint:

Children have little unambiguous input proving structuredependence.

→ How do they acquire this knowledge so quickly?

Two Pieces of Children's Structure-dependent Knowledge

(i) Linguistic representations are hierarchically structured.

(ii) Rules manipulating linguistic elements utilize this structure.

Hypotheses

1. Innate Bias Hypothesis:

Children have an innate bias to use hierarchical structure.

- how they represent language \rightarrow (i)
- how they hypothesize rules that manipulate linguistic elements \rightarrow (ii)

2. Learning Without Bias Hypothesis:

Children can converge on [correct] knowledge (especially, rules) w/o an innate bias for structure-dependence.

→ Structure-Independent Learning Strategy

Structure-Independent Learning Model

- Reali & Christiansen (2005) demonstrated that:
 - a. A probabilistic learning model using bigram & trigram word frequencies in English child-directed speech
 - b. Seems to distinguish grammatical vs ungrammatical forms of complex yes/no questions w/o a bias for structure.
 - c. Example:
 - Grammatical: Is the boy who is watching Mickey Mouse happy?
 - Ungrammatical: * Is the boy who watching Mickey Mouse is happy?

Critique of Structure-Independent Learning Model

- Kam et al. (2008) showed that:
 - a. Reali and Christiansen's model succeeded due to "lucky fluke."
 - 1) Used irrelevant bigram information.
 - 2) Failed on a wider range of complex yes/no questions in English.
 - b. "This potential structure-independent learning strategy does not in fact work (at least as implemented)."

Variant of the Innate Bias Hypothesis

- Perfors et al. (2011) proposed that:
 - a. Children's innate bias is to *allow* structure-dependent representations into hypothesis space in the first place, not to exclude others.
 - \rightarrow Less specific innate bias
 - b. Uses a computational-level hierarchal Bayesian model.

A Computational-Level Hierarchal Bayesian Model (Perfors et al., 2011)

- Leveraging English child-directed speech input (abstracted into syntactic category sequence)
 - a. The modeled learner found that hierarchical structures are preferable to other possible representations.
 - Evidence for structure-dependence comes not only from complex yes/no questions, but also from various other utterances.
 - → indirect positive evidence

A Computational-Level Hierarchal Bayesian Model (Perfors et al., 2011)

2. Parsing & Grammatical Judgements

- a. The modeled learner could parse grammatical complex yes/no questions but failed to parse ungrammatical ones.
- b. The learner could correctly judge grammaticality based on its acquired structure-dependent grammar.

 \rightarrow The model shows that it is *in principle* possible to have general innate knowledge favoring structure-dependence, allowing successful language acquisition from data.

Limitations of the Model (Perfors et al., 2011)

- Computational models show *in principle* feasibility.
 - Based on the assumption that children can do optimal inference to identify the optimal representational hypothesis.
 - b. They abstract away from children's cognitive limitations (e.g., memory, processing limitations) on the inference process.
- Future research needed to test under human-like cognitive limitations.

Conclusion

- (i) Linguistic representations are hierarchically structured.
- (ii) Rules manipulating linguistic elements utilize this structure.
- Children's structure-dependent knowledge may come from an innate bias or learning without bias.
 - Structure-independent models struggle to explain acquisition.
 - Hierarchical Bayesian models show *in principle* feasibility, but children's cognitive limitations remain a challenge.

SPRING 2025, CompCog Models of LangAcq

THANK YOU

Seojin Hwang April 28 Mon