Bayesian modelling of clause-level constructional knowledge for Korean-speaking preschool children

Gyu-Ho Shin (University of Hawai'i at Mānoa) & Seongmin Mun (Université Paris Ouest Nanterre La Défense)

**BACKGROUND**

- Usage-based constructionist approaches
  - Language development as interactions between frequency and domain-general learning capacities (e.g., Goldberg, 2019; Tomasello, 2003)
  - Q: how do we appropriately represent developmental trajectories involving clusters of form-function pairings (i.e., constructions)?

- Bayesian-inference-based simulation
  - Assumption: human learning involves one's updated beliefs based on previous experience
  - Studies focused mostly on English (e.g., Alshahi & Stevenson, 2008; Barak et al., 2016; Perfors et al., 2011)
  - Q: to what extent are the implications of computational simulations generalisable across languages?

- Active transitives & suffixal passives in Korean
  - Korean: SOV language with overt case-marking
  - Clause-level constructions expressing a transitive event

**BAYESIAN SIMULATION**

- **Input composition**
  - All constructional patterns expressing a transitive event found in ChILDES (MacWhinney, 2000)
    | Type                        | Example                  | Frequency |
    |--------------------------------|--------------------------|-----------|
    | Suffixal passive, actor        | suffix NOM-ACC catch     | 1,757     |
    | Suffixal passive, undergoer    | suffix NOM-ACC catch     | 1,492     |
    | Suffixal passive, undergoer    | suffix NOM-ACC catch     | 1,492     |
    | Suffixal passive, no ACC       | suffix NOM-ACC catch     | 1,492     |
    | Suffixal passive, no NOM       | suffix NOM-ACC catch     | 1,492     |
    | Scrambled suffixal passive, no ACC | suffix NOM-ACC catch     | 1,492     |
    | Scrambled suffixal passive, no NOM | suffix NOM-ACC catch     | 1,492     |

- **Model training**
  - Frequency of constructional patterns in caregiver input
    - Initial priors for learning
  - Learning algorithm (adapted from Alshahi & Stevenson, 2008)
    - A new input xC corresponds to an existing construction xC, ranging over the indices of all the constructions in the model, with the maximum probability given xC:
    - Posterior probability is proportional to multiplication of conditional probabilities associated with xC and the prior of xC
    - Laplace smoothing to prevent the probability from converging upon zero
  - Two types of probability information
    - Constructual probability: probabilities of individual patterns
    - Transitional probability: conditional probabilities of constructional components within each pattern
      - Given language-specific properties in Korean, how a Korean learner formulates knowledge about active transitives and suffixal passives?

**RESULTS & DISCUSSION**

- **By-pattern posterior probabilities**
  - Dominance of several patterns over the others

<table>
<thead>
<tr>
<th>Type</th>
<th>Caregiver input (#)</th>
<th>Child production (#)</th>
<th>Posterior probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canonical active transitive</td>
<td>1,757</td>
<td>0.344</td>
<td>0.001</td>
</tr>
<tr>
<td>Scrambled active transitive</td>
<td>0.002</td>
<td>0.005</td>
<td>0.036</td>
</tr>
<tr>
<td>Canonical suffixal passive</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Scrambled suffixal passive</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

- **Inconsistency between simulation and child production**

**REFERENCES**


**DISCUSSION**

Q: How do we appropriately represent developmental trajectories involving clusters of form-function pairings (i.e., constructions)?

A: Bayesian inference-based simulation

**MODEL PERFORMANCE**

- Posterior probabilities of constructional patterns at every learning phase (one to 30)
  - as a proxy for the degree of clustering for these constructions

**Together, our findings…**

- support the idea that clause-level constructional knowledge grows through an interplay between input properties and domain-general learning capacities
- adds to cross-linguistic evidence for the effectiveness of Bayesian modelling on representing human learning

**ACKNOWLEDGEMENT**

Presented at the UPMER Conference 2021 (30th June 2020)

**Corresponding Author:** Gyu-Ho Shin, gyuho.shin@hawaii.edu