Using developmental modeling to specify learning and representation of the passive in English children

Emma Nguyen[1] and Lisa Pearl[2]

University of Connecticut[1], University of California-Irvine[2]

UCONN UNIVERSITY OF CONNECTICUT

Computation of Language Laboratory — UC Irvine

Lexical Semantic Profiles and Age of Acquisition

Lexical profiles (combinations of lexical features) correlate well with the observed age of acquisition (AoA) of the English verbal *be*-passive (Maratsos et al. 1985; Nguyen & Pearl 2017)

*Eleanor **was hugged** by Edward.* < *Eleanor **was loved** by Edward.*

The semantic features defining these profiles were descriptive, proposed and defined to explain specific empirical results. We additionally propose the syntactically-motivated feature of transitivity.

8 lexical features of interest: 7 semantic (F1-F7: Stative, Volitional, Affectedness, Object-Experiencer, Subject-Experiencer, Agent-Patient, and Actionality), 1 syntactic (F8: Transitivity).

Profile	Verb	F1	F2	F3	F4	F5	F6	F7	F8	Obs AoA
1	carry	-	+	+	-	-	+	+	+	3yrs
2	annoy	+	+	+	+	-	-	-	+	3-4yrs
3	find	-	-	-	-	-	+	+	+	4-5yrs
4	forget	-	-	-	-	+	-	-	+	4-5yrs
5	hate	+	-	-	-	+	-	-	+	5yrs

Model Evaluation

Assumption: c_{+pass} priors < 0.50 are reasonable, indicating an initial bias against the passive structure (i.e., the passive operation takes some work).

Goal: Identify **which features can be heeded** in order to yield **passivization for all 5 observed profiles**, while still having a passivization prior < 0.5.

Lexical Features in Child-Directed Speech

If a modeled child is taking full advantage of the input, would she yield the observable passive acquisition behavior?

Profile	1	2	3	4	5
All	**0.21**	**0.43**	*0.81*	*0.98*	*0.99*

The child would have to be hugely biased in favor of the passive (the passive would have to be really easy) to passivize verbs from profiles 3-5.

Now what?

It's possible that children are filtering the input.

Input filtering is predicted to be necessary (Gagliardi et al. 2012, 2017) - children can't heed all the lexical feature information in their input at age five in order to passivize the way they do.

Simulating input filters over lexical features

Input filters can operate over any lexical feature. If a lexical feature is filtered out by the modeled child, that feature value is ignored during the likelihood calculation.

Approach: Use **Bayesian inference** to formalize the conditions under which **children's input** could yield **observed passive acquisition behavior**.

View learning as a **classification** problem:
Is a verb in the passivizable class (**+pass**) or the not-passivizable class (**-pass**), based on the lexical features associated with verbs that are +/-pass in the child's input?

$$P(c_{+pass} \mid v_{f_1}, \ldots v_{f_n}) \propto P(c_{+pass}) \cdot \prod_{f_i \in F} P(v_{f_i} \mid c_{+pass})$$

The probability of a verb being in the passivizable class, based on the the verb's lexical feature values

How difficult passivization is (near 0 = hard, near 1 = easy)

How likely these lexical feature values are for verbs in the passivizable class, based on the input

Model's target output: the five profiles corresponding to the 30 verbs experimentally attested for by age five (Nguyen & Pearl 2017).

We don't know what this should be *a priori* for children.

We estimate these values from a portion of the CHILDES Treebank (Pearl & Sprouse 2013), which includes the Brown Corpus (Brown 1973) [Adam, Eve, and Sarah] and the Valian corpus (Valian 1991): 113,024 child-directed speech utterances from 1;06 to 5;01 (62,772 tokens of 747 verbs (73% were passivizable))

How: We can compare the likelihoods of the lexical profile feature values for c_{+pass} and c_{-pass}. Passivizing a verb with a specific lexical profile ($P(c^{+pass} \mid v_{f1} \ldots v_{fn}) > 0.50$) results when c_{+pass}'s prior * likelihood > c_{-pass}'s prior * likelihood. From this, we can then calculate the prior $P(c_{+pass})$, using the **ratio of the likelihoods.**

$$l_{+pass} = \prod_{f_i \in F} P(v_{f_i} \mid c_{+pass}) \qquad l_{-pass} = \prod_{f_i \in F} P(v_{f_i} \mid c_{-pass})$$

What we do: Calculate ***what the prior would need to be*** in order for five-year-olds to **passivize the verbs they do**, assuming they were learning from lexical feature frequencies in their input.

$$P(c_{+pass}) \cdot l_{+pass} > P(c_{-pass}) \cdot l_{-pass}$$

$$P(c_{+pass}) \cdot l_{+pass} > (1 - P(c_{+pass})) \cdot l_{-pass}$$

$$\boxed{P(c_{+pass}) > \frac{l_{-pass}}{l_{+pass} + l_{-pass}}}$$

Table: Likelihood probabilities

v_{f_i}	$P(v_{f_i} \mid c_{+pass})$ 1	$P(v_{f_i} \mid c_{+pass})$ 0	$P(v_{f_i} \mid c_{-pass})$ 1	$P(v_{f_i} \mid c_{-pass})$ 0
ACTIONAL	0.92	0.08	0.89	0.11
STATIVE	0.07	0.93	0.09	0.91
VOLITIONAL	0.92	0.08	0.77	0.23
AFFECTED	0.85	0.15	0.53	0.47
OBJ-EXP	0.05	0.95	0.02	0.98
SUBJ-EXP	0.03	0.97	0.05	0.95
AGT-PAT	0.87	0.13	0.63	0.37
TRANS	0.94	0.06	0.71	0.29

Results of Input Filters

Profile	1	2	3	4	5
Trans	0.43	0.43	0.43	0.43	0.43
Trans & Obj-Exp	0.44	0.21	0.44	0.44	0.44

Passivization priors of feature combinations that still allow a match to five-year-old passivization behavior across the five verb profiles. **Only two feature combinations are viable (both of which involve transitivity)** - all other combinations lead to passivization priors > 0.5 in order to match observable passivization at age five.

Takeaway

The transitivity feature seems key, while the object-experiencer feature may also play a role.

Discussion

Five-year-olds are predicted to view the passive structure as somewhat - but not strongly - costly. The input filtering that successfully matches observed passivization behavior leads to priors that are usually only slightly less than 0.5. (*Note: The highest estimate is the smallest the prior could be and still allow passivization. If the passive prior is the same across verbs and verb profiles, the highest estimate for a combination is therefore the one that may be assumed to hold.*)

Future work can apply the same process to three- and four-year-old behavior to yield a quantified snapshot of the developmental trajectory of both the passive learning process and representation.

Selected References: Crain & Fodor. (1993). Competence and Performance.; Liter, Huelskamp, Weerakoon, & Munn. (2015). What drives the Maratsos Effect, agentivity or eventivity?; Maratsos, Fox, Becker, & Chalkley. (1985) Semantic restrictions on children's passives.; Messenger, Branigan, McLean, & Sorace. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming.; Nguyen. (2016). Felicity of the by-phrase is not enough: Re-evaluating O'Brien et al. (2006).; Pearl, & Sprouse. (2013). Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem.; Pinker, LeBeaux, & Frost. (1987). Productivity and constraints in the acquisition of the passive.