Immature representation or immature deployment?
Modeling child pronoun resolution

Hannah Forsythe
University of California, Irvine
SSPB 2211, Irvine, CA 92617
hforsyth@uci.edu

Lisa Pearl
University of California, Irvine
SSPB 2219, Irvine, CA 92617
lpearl@uci.edu

Abstract

Children acquiring Spanish interpret subject pronouns differently from adults, initially relying on pragmatic cues instead of morphosyntactic cues that are more statistically reliable. Following Gagliardi et al. (2017), we use Bayesian cognitive modeling to explore the sources of this non-adult-like behavior, investigating whether it is more likely due to (i) noise in children's representation of the probability that some cues favor certain antecedents, or (ii) noise in children's deployment of otherwise adult-like probabilities. Results favor noisy deployment as the source of children's non-adult-like pronoun resolution.

1 Intro

When children produce and interpret language differently from adults, the underlying cause can be unclear: do children have an immature representation of the target language, or do they simply deploy that representation in an immature way? One way to get at this question is to design behavioral tasks that facilitate deployment (e.g., lowering processing demands, improving task pragmatics, using more sensitive behavioral measures), with the idea that any non-adult-like behavior that remains after deployment effects have been controlled for is likely due to representational issues. However, it can be difficult to know for sure that deployment effects have been completely controlled for in any given experiment.

Here, we show how cognitive modeling can be used to more directly target representational versus deployment explanations of children's non-adult-like behavior. Following the approach of Gagliardi et al. (2017), we use Bayesian models on a case study of Spanish-speaking children's pronoun resolution to explore whether their non-adult-like use of pronominal cues is better modeled as noise in (i) the representation of the information these cues provide, or (ii) the deployment of that information during interpretation. Results suggest that noisy deployment is more likely to underlie children's non-adult-like behavior in this case. We also discuss implications for both the development of pronoun knowledge, and the investigation of linguistic development more generally.

2 Non-adult-like pronoun resolution by children acquiring Spanish

In Spanish, the interpretation of subject pronouns in context depends on many constraints, but some constraints are stronger than others. For instance, subject pronouns can be probabilistically biased by cues such as an accompanying discourse **CONN**ective or by the speaker's choice of pronominal **FORM**. In (1), temporal connective *y después* ('and then') favors the subject antecedent (*la maestra*: 'the teacher') more strongly than causal *porque* ('because'), and use of the null subject (*pro*) favors the subject antecedent more strongly than the overt pronoun (*ella*). In contrast to these probabilistic cues, subject pronouns can be categorically disambiguated by the accompanying verb's agreement **MOR**phology. In (2), -3S indicates the singular subject while -3P indicates the plural object.

(1) La maestra saluda a la niña, (y después /
the teacher waves at the girls, (and then /
porque) Ø /ella sale
because) *pro*/she leaves

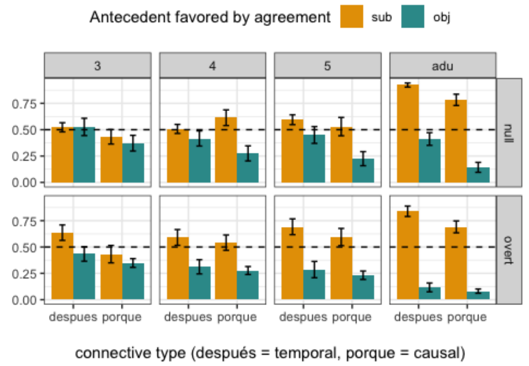
'The teacher waves at the girl, (and then/because) PRONOUN leave(s).'

(2) La maestra$_i$ saluda a las niñas$_k$, y Ø
the teacher waves at the girls, and *pro*
sale$_i$ / salen$_k$
leave-3S / leave-3P

'The teacher waves at the girls, and PRONOUN leave(s).'

To determine how Spanish-speaking adults and children use these three cues to interpret pronouns in context, Forsythe (accepted) used a forced-choice picture selection task with 47 adults, and 98 preschoolers. Participants listened to sentences like (1) and (2) and indicated their interpretation of the pronoun by choosing an illustration depicting either the subject interpretation (e.g., the teacher waving) or the non-subject interpretation (e.g., girls waving). Cues were fully crossed, systematically aligning and pitting each type against the other two. Figure 1 shows how often participants chose the subject interpretation.

Figure 1: How often children ages 3-5 and adults favor the subject interpretation of a pronoun in context, given different cues: connectives (*después, porque*), pronoun form (*null*, *overt*), and agreement morphology (agreeing with *sub*ject or *obj*ect).



Adults favor the expected interpretation (subject vs. non-subject) on the basis of agreement morphology, but interestingly, this preference is not completely categorical: it is modulated by the cues of connective and pronominal form, which probabilistically bias the interpretation towards or away from the subject antecedent. That is, adults rely on all three cues when interpreting pronouns. Children's behavior shows qualitatively different patterns, with three-year-olds relying only on connectives, four-year-olds relying only on morphology, and five-year-olds relying on both morphology and connectives. Importantly, it is unclear from these results whether children's non-adult-like pronoun interpretation behavior is due to an immature representation of the information that these cues carry (e.g., three-year-olds only have an adult-like representation of connectives) or to an immature ability to deploy the representations they have (e.g., three-year-olds have an adult-like representation of morphology, but fail to access it correctly in the moment). This is where cognitive modeling can help.

3 Modeling child pronoun resolution behavior

This pronoun interpretation behavior serves as our modeling target: a successful model will match children's behavioral patterns in each experimental condition as closely as possible. The model's input will be the same input that children acquiring Spanish use when learning how each of these cues predicts pronoun antecedents. Table 1 shows the rate of reference to the preceding subject antecedent and to singular antecedents, for different cue types, based on samples drawn from a corpus of 54,757 child-directed Spanish utterances.

Table 1: Rates of reference to different antecedent types in the presence of different CONNectives, pronoun FORMs, and agreement MORphology in child-directed Spanish.

cue	value	subject antecedent
CONN	*después*	(29/54) 54%
	porque	(52/149) 35%
FORM	null	(1,093/2,376) 46%
	overt	(64/291) 22%
		singular antecedent
MOR	singular	(5,655/5,662) 99.9%
	plural	(9/1,336) 0.7%

All the cues in these child-directed speech samples appear to follow the patterns we expect from adult behavior: connectives and pronominal form are more probabilistic cues, while agreement morphology is fairly categorical. This input pattern makes it surprising that younger children initially don't rely on agreement morphology.

To probe the underlying source of this immature behavior, we follow Gagliardi et al. (2017), who model linguistic immaturity as noise–either noise in the modeled child's representation of the information a given cue provides, or noise in the ability to reliably use that information in novel situations, such as an experimental task. Here, we ask whether children's non-adult-like interpretation of Spanish pronouns is best captured as noise in the representation of the information provided by cues from connectives, pronominal form, and agreement morphology, or as noise in how this information is accessed during the experiment.

3.1 Baseline model

We model children's reasoning process, which combines the information provided by cues in the child's input with the child's prior about the pronoun's most likely antecedent, using Bayesian inference as in (1). Bayesian inference is often used for cognitive development modeling, as it can capture human behavior very well (e.g., Perfors et al. (2011) Pearl and Mis (2016)).

The modeled child calculates the probability of a potential pronoun antecedent α (e.g., the teacher) given a particular combination of cues extracted from the pronoun and its utterance (e.g., f_{MOR}:sg, f_{CONN}: *después*, f_{FORM}: null), which corresponds to the posterior $P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}})$. This posterior is calculated by considering two probabilities extracted from the input: (i) the likelihood of each cue's value, given that type of antecedent ($P(f_{\text{CUE\_VAL}}|\alpha)$) and (ii) the prior probability of referring to this type of antecedent ($P(\alpha)$).

$$P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}}) \propto P(f_{\text{MOR}}|\alpha) \cdot P(f_{\text{CONN}}|\alpha) \cdot P(f_{\text{FORM}}|\alpha) \cdot P(\alpha) \quad (1)$$

This version of the modeled child makes optimal use of the cues as they appear in the input and will therefore rely most heavily on the most reliable cues, such as morphology–in clear contrast to what we observe in children. To model a child with either immature representations of cue information or immature deployment of cue information, we introduce noise into this optimal model.

3.2 A noisy representation model

The noisy representation model we implement encodes the idea that children behave differently from adults because they have an immature representation of one or more pronominal cues, which is caused by noisily extracting cue information from the input. For example, suppose the link between singular and plural surface agreement and underlying number features is immature. This would prevent the child from accurately tracking how the number semantics of a pronoun's accompanying agreement marker predicts the number semantics of its antecedent (i.e., for the child, singular morphology might not categorically predict a singular antecedent). This in turn could flatten the dramatic difference between $P(\alpha\text{:sg}|f_{\text{MOR}}\text{:sg})$ and $P(\alpha\text{:sg}|f_{\text{MOR}}\text{:pl})$ that is evident from the Spanish-language input in Table 1. Whatever the cause, noisy encoding of cue information from the input will yield non-adult-like likelihood terms.

The noisy representation model in (2) flattens the distributions for each likelihood term using softmax ($e^{\sigma \cdot P}$), which is standardly used for this purpose to model decision-making tasks, including language tasks (e.g. Frank and Goodman (2012); Goodman and Stuhlmüller (2013); Scontras and Goodman (2017)). The level of noise associated with each cue type is controlled by the parameters σ_{MOR}, σ_{CONN}, and σ_{FORM}, with smaller values indicating a flatter distribution and greater values indicating a sharpened distribution.

$$P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}}) \propto e^{\sigma_{\text{MOR}} P(f_{\text{MOR}}|\alpha)} \cdot e^{\sigma_{\text{CONN}} P(f_{\text{CONN}}|\alpha)} \cdot e^{\sigma_{\text{FORM}} P(f_{\text{FORM}}|\alpha)} \cdot P(\alpha) \quad (2)$$

3.3 Two noisy deployment models

We also implement two noisy deployment models encoding the idea that children behave differently from adults because they immaturely access adult-like cue representations during the experimental task. Both models accurately encode the cue information from children's input but deploy this information inaccurately, either (i) occasionally deleting cue information (*noisy deletion*), or (ii) substituting accurate cue information with a default value (*noisy default*). Such deletion or substitution of cue information from experimental items could be caused by a variety of factors, including limited working memory capacity, background noise, inattention, and so on. Whatever the reason, the result is that the child inaccurately deploys this otherwise accurate cue information.

More specifically, both noisy deployment models rely on cue likelihoods ($P(f_{\text{CUE\_VAL}}|\alpha)$) accurately obtained from the input but access this information probabilistically via mixture modeling. The noisy deletion model (3) mixes the optimal model with models that delete one, two, or all three cues. In other words, when this modeled child is unable to deploy a given cue, she simply drops that cue's information.

$$\begin{aligned} &P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}}) \propto \\ &[(\beta_{\text{MOR}})(\beta_{\text{CONN}})(\beta_{\text{FORM}})\ P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}}) + \\ &(1-\beta_{\text{MOR}})(\beta_{\text{CONN}})(\beta_{\text{FOR}})\ P(\alpha|f_{\text{CONN}}, f_{\text{FORM}}) + \\ &... + \\ &(1-\beta_{\text{MOR}})(1-\beta_{\text{CONN}})(1-\beta_{\text{FORM}})] \times \\ &P(\alpha) \end{aligned} \quad (3)$$

The noisy default model (4) mixes the optimal model with models that substitute the cue's true value ([*acc*]) with a default ([*def*]), which we determined by sampling from the distribution of cue values in the child's input. In other words, when this modeled child is unable to deploy a given cue, she inserts a default value.

$$
\begin{aligned}
&P(\alpha|f_{\text{MOR}}, f_{\text{CONN}}, f_{\text{FORM}}) \propto \\
&\quad [(\beta_{\text{MOR}})(\beta_{\text{CONN}})(\beta_{\text{FORM}}) \\
&P(\alpha|f_{\text{MOR}} = [acc], f_{\text{CONN}} = [acc], f_{\text{FORM}} = [acc]) + \\
&\quad (1-\beta_{\text{MOR}})(\beta_{\text{CONN}})(\beta_{\text{FORM}}) \\
&P(\alpha|f_{\text{MOR}} = [def], f_{\text{CONN}} = [acc], f_{\text{FORM}} = [acc]) + \\
&\quad \ldots + \\
&\quad (1-\beta_{\text{MOR}})(1-\beta_{\text{CONN}})(1-\beta_{\text{FORM}}) \\
&P(\alpha|f_{\text{MOR}} = [def], f_{\text{CONN}} = [def], f_{\text{FORM}} = [def])] \times \\
&\quad P(\alpha)
\end{aligned} \tag{4}
$$

In both mixture models, the level of noise associated with each cue is determined by how much each sub-model contributes to the mix. Specifically, in (3) β_{MOR}, β_{CONN}, and β_{FORM} indicate the rate at which morphological, connective, and form cues are included, while in (4) they indicate the rate at which the accurate cue value is retained.

3.4 Results and discussion

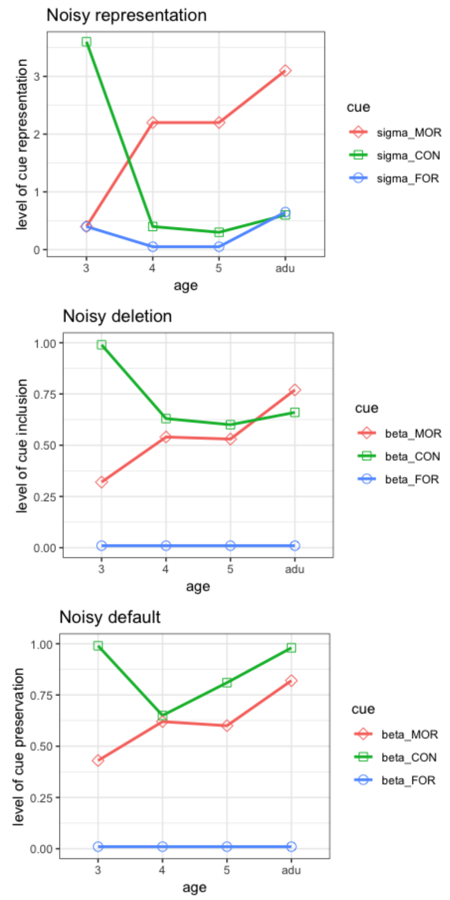
All three noisy models significantly outperform the baseline optimal model in capturing child behavior at age 3 (GLLR: all $\chi^2(3)>1641$, all $p<0.001$), 4 (all $\chi^2(3)>1451$, all $p<0.001$), and 5 (all $\chi^2(3)>976$, all $p<0.001$); notably, the same is true for adult behavior (all $\chi^2(3)>509$, all $p<0.001$). Among the three noisy models, the two noisy deployment models have lower overall MSEs and higher log likelihoods. This suggests that the noisy deployment models are consistently better at capturing the pronoun resolution behavior of the adults and children in this experiment and, moreover, that the underlying source of children's non-optimal pronoun interpretations is more likely to be an immature deployment of their otherwise mature representations of cue information.

In terms of the amount of noise associated with each cue, our models show similar developmental patterns (see Figure 2). First, there's a steady decrease in the noise associated with agreement morphology as children get older (i.e., the red lines show an increase in children's reliance on this cue). Second, the noise associated to connectives (green lines) is almost adult-like from age four on, while three-year-olds appear more sensitive than adults. Third, none of the best-fitting models indicate much use of pronominal form at all (i.e., blue lines close to 0)–including the best-fitting models for adults. This suggests that mostly ignoring pronominal form in this task is in fact adult-like.

4 Conclusion and future directions

Here we have shown how to use cognitive modeling to implement two different types of developmental theories about why children's interpretation of Spanish subject pronouns is non-adult-like. Our results suggest that immature deployment, rather than immature representation of cue information, is the more likely cause of children's behavior. In particular, younger children seem to inconsistently access their representations of how agreement morphology predicts the number of the pronoun's antecedent.

Figure 2: Best-fitting noise parameters for each model and age group. Larger values indicate less noise.



However, we do note that children show qualitatively different behavior in singular vs. plural conditions and that the noisy representation model is particularly bad at capturing this difference. This suggests that future work may improve model fit by using separate noise parameters for singular and plural morphology, rather than a single noise parameter for agreement morphology. This may result in a better quantitative fit to child behavior, especially for the noisy representation model.

More generally, our approach demonstrates how computational modeling can complement behavioral approaches to the investigation of language development, affording a clearer picture of what it is that changes as children grow into adults.

References

Hannah Forsythe. accepted. Resolving pronouns with multiple cues: Children use pragmatics before morphology. poster to be presented. In *44th annual Boston University Conference on Language Development (BUCLD44)*.

Michael C Frank and Noah D Goodman. 2012. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Annie Gagliardi, Naomi H Feldman, and Jeffrey Lidz. 2017. Modeling statistical insensitivity: Sources of suboptimal behavior. *Cognitive Science*, 41(1):188–217.

Noah D Goodman and Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science*, 5(1):173–184.

Lisa Pearl and Benjamin Mis. 2016. The role of indirect positive evidence in syntactic acquisition: A look at anaphoric one. *Language*, 92(1):1–30.

Amy Perfors, Joshua Tenenbaum, Thomas Griffiths, and Fei Xu. 2011. A tutorial introduction ot bayesian models of cognitive development. *Cognition*, 120(3):302–321.

Gregory Scontras and Noah D Goodman. 2017. Resolving uncertainty in plural predication. *Cognition*, 168:294311.