Discovering Pronoun Categories using Discourse Information

Naho Orita (naho@umd.edu)
Rebecca McKeown (rmckeown@umd.edu)
Naomi H. Feldman (nhf@umd.edu)
Jeffrey Lidz (jlidz@umd.edu)
Department of Linguistics, 1401 Marie Mount Hall, University of Maryland, College Park, MD 20742 USA

Jordan Boyd-Graber (jbg@umiacs.umd.edu)
College of Information Studies, South Hornbake, University of Maryland, College Park, MD 20742 USA
University of Maryland Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 USA

Abstract

Interpretation of a pronoun is driven by properties of syntactic distribution. Consequently, acquiring the meaning and the distribution are intertwined. In order to learn that a pronoun is reflexive, learners need to know which entity the pronoun refers to in a sentence, but in order to infer its referent they need to know that the pronoun is reflexive. This study examines whether discourse information is the information source that the learner might use to acquire grammatical categories of pronouns. Experimental results demonstrate that adults can use discourse information to accurately guess the referents of pronouns. Simulations show that a Bayesian model using guesses from the experiment as an estimate of the discourse information successfully categorizes English pronouns into categories corresponding to reflexives and non-reflexives. Together, these results suggest that knowing which entities are likely to be referred to in the discourse can help learners acquire grammatical categories of pronouns.

Keywords: language acquisition; Bayesian modeling

English speakers know that the sentence in (1) means that Alice saw Alice in the mirror and the sentence in (2) means that Alice saw someone else in the mirror.

(1) Alice saw herself in the mirror.
(2) Alice saw her in the mirror.

These interpretations reflect adults' knowledge that reflexives like *herself* require different syntactic relations with their antecedents than non-reflexives like *her*. Evidence shows that children acquiring various languages have knowledge of the grammatical distributions of pronouns (Jakubowicz, 1984; Crain & McKee, 1985, among many). However, it is not yet known *how* children acquire this knowledge.

In English, the distribution of pronouns is governed by two constraints on the pronoun-antecedent relation: locality and c-command. Locality refers to the domain of the syntactic relation between the pronoun and its antecedent. Reflexive pronouns must have their antecedents in the local domain, corresponding approximately the same clause in English (Chomsky, 1973). The second constraint is that reflexive pronouns must be c-commanded by their antecedent (Reinhart, 1976). In the sentence 'Alice's sister saw herself', English speakers know that the antecedent of *herself* is not *Alice*, but *Alice's sister*. That is, when the hierarchical structure of the sentence is represented as a tree in Figure 1, the reflexive *herself* is contained in the sister node of its antecedent *Alice's sister*. Non-reflexive pronouns appear in exactly those

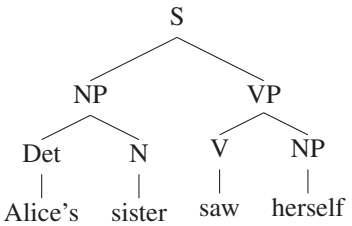

Figure 1: Syntactic tree showing a c-command relationship between the antecedent *Alice's sister* and the pronoun *herself*.

contexts where reflexives do not appear: approximately contexts in which the antecedent is either non-local, not in a c-commanding position, or both (Chomsky, 1973). This means that the relationship between the grammatical positions of antecedents and pronouns, as characterized by locality and c-command, defines the distribution of grammatical categories of pronouns in English.

Critically, these grammatical constraints concern the relationship between a pronoun and its antecedent. This means that syntactic knowledge alone is insufficient for acquiring pronouns, because it cannot be applied without knowing the intended antecedent of a pronoun. In order to learn that *herself* is reflexive, learners need to interpret the sentence in (1) as 'Alice saw Alice', recognizing that *Alice* and *herself* corefer to the same entity. However, in order to interpret the meaning of the sentence, they need to use the knowledge that *herself* is a reflexive pronoun, whereas *her* is a non-reflexive pronoun. This circularity poses a potentially difficult problem for children acquiring language.

In this paper we show that discourse information can help learners categorize pronouns into appropriate distributional classes. If learners use discourse information to predict that the pronoun *herself* in (1) is likely to refer to *Alice* and the pronoun *her* in (2) is likely to refer to someone else, this provides information that can help them categorize these pronouns into different classes. We examine (i) to what extent discourse context is informative for determining the referent of a pronoun and (ii) whether this estimate of a pronoun's reference is sufficient for learning to classify pronouns.

The paper is organized as follows. Our next section describes a behavioral experiment that measures the discourse information available to listeners. The following section

presents Bayesian modeling results showing that this discourse information can help bootstrap grammatical knowledge of pronoun categories. Finally, the last section addresses open questions and implications.

Experiment 1: Human Simulation

To test to what extent language contexts are informative about the referents of pronouns, we used a variant of the human simulation paradigm (Gillette, Gleitman, Gleitman, & Lederer, 1999). In this paradigm, adult participants guess the identity of a missing word on the basis of linguistic and/or situational data. Because participants know there is a word, but not what it is, they are simulating what it is like to be a language learner who hears a word but does not know its meaning. A goal of the human simulation paradigm is to see what can be inferred about the meaning of a word based on information present either in the linguistic input or in the scene. Past experiments using the human simulation paradigm have examined the degree to which adults (Gillette et al., 1999; Kako, 2005) or older children (Piccin & Waxman, 2007) can guess identities of common nouns and/or verbs.

In our experiment, adult participants were shown text excerpts of conversations between adults and children. Their task was to guess the identity of a word or phrase that had been blanked out, which was either a reflexive pronoun, a non-reflexive pronoun or a lexical noun phrase. The goal was to determine whether conversational context provides sufficient information for adults to guess what is being referred to. If so, this would provide evidence in favor of the idea that language learners can determine the referents of pronouns they do not yet know based on conversational context.

Methods

Participants Participants were 40 undergraduates at the University of Maryland, College Park (11 men, 29 women). All were native English speakers and all were at least 18 years old. Participants were enrolled in introductory linguistics courses and received course credit for their participation.

Materials Text excerpts of real recorded conversations between adults and young children were taken from the ENG-USA section of the CHILDES database (MacWhinney, 2000). In each excerpt, one line was bolded. This bolded line had a noun phrase (NP) that had been deleted and replaced with a blank. The deleted noun phrase always came from an adult utterance. There were 12 lines of dialogue before the bolded line and six lines afterwards. Every deleted noun phrase was the object of one of five verbs: *hurt*, *see*, *help*, *dry*, and *cover*. Using the same verbs in all contexts allowed us to factor out any possible contribution of verb knowledge to determining which pronoun was intended. The deleted noun phrases belonged to one of three categories: 25 were reflexive pronouns (4 tokens of *myself*, 1 token of *ourselves*, 7 tokens of *himself*, 10 tokens of *yourself*, and 3 tokens of *themselves*), 25 were non-reflexive pronouns (4 tokens of *me*, 1 token of *us*, 7 tokens of *him*, 10 tokens of *you*, 3 tokens of *them*), and 25 were lexical NPs (names – including Mommy or Daddy – and definite descriptions). This led to a total of 75 test items. Within the test items, frequencies of corresponding non-reflexive and reflexive pronouns were matched (e.g., *me* was matched in frequency with *myself*, etc).

The dialogues were chosen randomly from all adult utterances in CHILDES that used the relevant verbs with the relevant type of NP object, with the exception that we threw out utterances that were direct repetitions of a previous line or that were well-known quotations. Finally, the materials were chosen to balance, as much as possible, the person of the pronoun object of the verb (though due to an imbalance in the available CHILDES data we were still left with more second-person objects than first or third person). In addition to the lines of dialogue, each item in the experiment provided a list of participants in the conversation and the age of the child in the conversation. No information was given about the situation or context in which the conversation took place.

After each excerpt, participants were given a list of 15 choices for what NP could have gone in the blank. The choices always included the same five reflexive pronouns (yourself, myself, ourselves, himself, themselves) and non-reflexive pronouns (you, me, us, him, them). They also included five lexical NPs which would have been prominent in the conversation: e.g., the names of the participants (including Mommy or Daddy) and prominent people or objects mentioned in the conversational excerpts. If the actual sentence contained a lexical NP then this lexical NP was one of the five lexical NPs provided. The NPs were presented in alphabetical order.

Procedure Participants were given an hour in a quiet room to complete the experiment. Test items were presented on paper, one per page. Participants were instructed to read the dialogues, which were real conversations between adults and children, and pick the word or phrase (from the list of 15 choices) they thought belonged in the blank. Participants wrote answers on a separate answer sheet. The test items were presented in random order. Twenty participants received the first 38 test items, and the remaining twenty participants received the remaining 37 test items.

Results and Discussion

Overall, participants were highly accurate at guessing the correct word from a list of 15 choices. The first row in Table 1 breaks up guesses of the correct word by syntactic category of the NP (reflexive pronouns, non-reflexive pronouns, or lexical NPs). Individual participants chose the correct NP out of 15 choices an average of 63.8% of the time. This ranged from 32.4% for the least accurate participant to 84.2% for the most accurate participant, with a standard deviation of 10.6%, and was significantly better than chance ($t(39) = 34.19$, $p < 0.0001$). These results show that adults can usually guess the identity of a missing NP given only a small amount of linguistic context.

However, these results underestimate participants' ability

	Lexical NP	Non-reflexive	Reflexive
% correct word	61.75	70.25	64.25
% plausibly correct word	66.75	81.25	68

Table 1: Percentage of correct answers and answers with a plausibly correct referent in Experiment 1

	Lexical NP	Non-reflexive	Reflexive
% Lexical NP guesses	**71.8**	23.4	15.8
% Non-reflexive guesses	23.2	**73.4**	16
% Reflexive guesses	5	3.2	**68.2**

Table 2: Confusion matrix obtained in Experiment 1

to guess what is being referred to. The second row in Table 1 shows guesses of a plausibly correct word, a word that plausibly had the same intended referent as the correct word (for instance, a pronoun with the same gender/number features as the name that had actually been used, or vice versa). These results show that adults are good at guessing which entity is referred to given a context, irrespective of grammatical knowledge relevant to pronouns.

Table 2 breaks up the results by syntactic category of the NP. Participants' guesses were usually of the same category that the actual word had been. Importantly, adults usually guessed correctly whether the missing word had been a reflexive pronoun—when the word actually had been reflexive, participants guessed a reflexive 68.2% of the time. When the word had been a lexical NP or a non-reflexive pronoun, they almost never guessed that it had been a reflexive.

This task parallels that of a child identifying an unfamiliar word. Of course, the parallel is not complete. In some ways, adult participants were provided with less information than the children they were meant to simulate: they only received a small excerpt of the conversation and did not receive any visual information. In other ways, the participants had more data: they already knew the meanings of all of the other words in the conversation, they had full syntactic and discourse knowledge where children might only have partial knowledge (e.g., Arnold, Brown-Schmidt, & Trueswell, 2007), and they were limited to 15 choices of possible meaning. Furthermore, choosing an answer in this experiment was not subject to any time pressures, whereas in actual acquisition processing speed could potentially impact the learner's ability to use the discourse context as an information source. However, to the extent that the adult simulation reflects the prior information presented in the discourse, it provides an estimate of the information that children might have access to. Where adults (who already know the distribution of reflexives) can guess that a missing word is reflexive, a child might be able to guess that a missing word co-refers with a specific NP. Together with syntactic knowledge of locality and c-command, this should provide learners with useful information for acquiring grammatical categories of pronouns. To explore this possibility, we formalize a Bayesian model that learns to categorize pronouns.

Experiment 2: Bayesian Model

In this section, we develop a Bayesian model that integrates the discourse information measured in Experiment 1. This model investigates whether the information in discourse could be sufficient to learn the grammatical categories of English pronouns in principle (a computational-level model; Marr, 1982). The model discovers:

1. how many pronoun categories there are in a language
2. the distribution of pronouns in each category
3. which syntactic position of an antecedent is associated with each pronoun category

This ideal learner is assumed to have (a) discourse knowledge that helps define the distribution of the potential antecedents, (b) syntactic knowledge relevant to pronoun categories (details follow), and (c) lexical knowledge that is sufficient for distinguishing pronouns from lexical noun phrases. Other linguistic information relevant to pronouns, such as gender and number, is not represented in our model; we ask simply whether our ideal learner can acquire two categories corresponding to reflexive and non-reflexive pronouns.

Regarding (b) above, this ideal learner is assumed to already know locality and c-command before learning pronoun categories, and is further assumed to know that these are relevant for categorizing pronouns. Thus, the learner is able to identify the syntactic position of each potential antecedent. The model distinguishes four syntactic positions based on the knowledge of locality and c-command; [+local,+c-command], [+local,-c-command], [-local,+c-command], and [-local,-c-command]. In English, if an antecedent is in a syntactic position described by [+local,+c-command], that pronoun must be a reflexive pronoun. If the potential antecedent is elsewhere, that pronoun must be a non-reflexive pronoun. However, the learner does not know in advance which syntactic position is associated with which pronoun category, and needs to acquire this knowledge from the input. We return to this issue of prior syntactic knowledge in the Discussion.

Generative Model

Our model assumes the following generative process. For each pronoun, an antecedent in one of the four syntactic positions described above is chosen given prior discourse knowledge ($\mathcal{D}$). Then a pronoun category is chosen based on the syntactic position of the antecedent, and a pronoun is generated from the chosen pronoun category.

Figure 2 illustrates this process with a graphical model.[1] Each antecedent category distribution θ_j is a random variable

[1] This model is a nonparametric extension to the author-topic model (Rosen-Zvi, Griffiths, Steyvers, & Smyth, 2004) that allows for an infinite number of categories (called topics in their model).

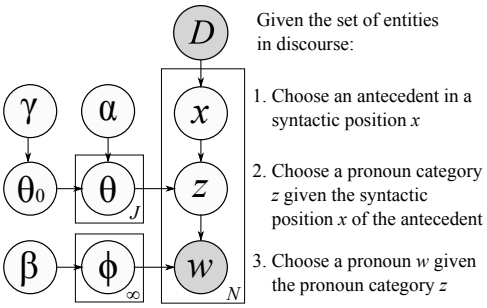

Figure 2: Graphical Model

that encodes the distribution over pronoun categories favored by an antecedent in syntactic position j. For example, the category *reflexive* would have high probability in the distribution $\theta_{[+\text{local},+\text{c-command}]}$. (Here we use the category name *reflexive* for exposition, but the model does not associate any labels with the pronoun categories it recovers.) Each category word distribution ϕ_k is a distribution over words that encodes the probability distribution over pronouns in pronoun category k. For example, pronouns such as *herself* and *myself* would have high probabilities in the distribution $\phi_{\text{reflexive}}$. In addition to learning this distribution, our model learns the number of pronoun categories needed to describe the data. For each pronoun in the corpus, an antecedent in a syntactic position x is assumed to be sampled from a distribution we refer to as discourse knowledge $\mathcal{D}$ (see the next section for the details). A pronoun category z is then sampled from the multinomial distribution with parameter θ associated with the syntactic position x of the antecedent and a pronoun w is sampled from a multinomial distribution with parameter ϕ associated with pronoun category z.

To learn the number of pronoun categories based on the observed data, we use the hierarchical Dirichlet process (Teh, Jordan, Beal, & Blei, 2006). The advantage of this model is that it allows the model to flexibly learn how many categories are present in the data, while still allowing categories to appear across multiple grammatical contexts. The summary of the generative process follows.

1. Draw a distribution over pronoun categories $\theta_0 \sim \text{GEM}(\gamma)$, where GEM is the Griffiths, Engen, McCosky distribution (Pitman, 2002).
2. For each antecedent syntactic position $j = 1 \ldots 4$, draw a pronoun category distribution $\theta_j \sim \text{DP}(\alpha, \theta_0)$.
3. For each pronoun category $k = 1 \ldots \infty$, draw a distribution over tokens $\phi_k \sim \text{Dir}(\beta)$.
4. For each pronoun in the corpus $n = 1 \ldots N$
   (a) Draw an antecedent syntactic position from the discourse knowledge $x_n \sim \mathcal{D}$
   (b) Draw a pronoun category $z_n \sim \text{Mult}(\theta_{x_n})$.
   (c) Draw a word $w_n \sim \text{Mult}(\phi_{z_n})$.

Prior Knowledge

The observed discourse knowledge $\mathcal{D}$ defines a prior distribution over potential antecedents in the discourse. Recall that our ideal learner maps each antecedent in the discourse deterministically to its syntactic position (defined in terms of locality and c-command), and in this way $\mathcal{D}$ defines a distribution over syntactic positions x for each pronoun's antecedent.

Rather than specify a parametric form for this prior distribution, we estimate it directly from participants' responses from Experiment 1. In one experimental item, for example, participants guessed the identity of the missing word in the sentence "You drying __ off?". Nine out of 20 participants guessed that the missing word is *yourself*, six out of 20 guessed *him*, three out of 20 guessed *me*, and two out of 20 guessed *Seth*. Under the assumption that experimental participants have sampled their responses from a shared prior distribution over entities in the discourse, these guesses provide an estimate of participants' beliefs about how likely each entity is to be referred to in the discourse.

Where participants chose *yourself*, the antecedent of this pronoun is *you*, which is a local and c-commanding antecedent. Where participants chose *him* and *me*, the antecedents could be in any of the remaining three syntactic positions, but in this particular dialogue the only potential antecedents for non-reflexives are neither local nor c-commanding.[2] We ignored responses in which participants chose lexical NPs (here *Seth*) based on the assumption that learners distinguish pronouns from lexical NPs. We then normalized each count by the total number of pronoun guesses. The resulting prior distribution over syntactic positions for antecedents in this example is $p(x_{[+local,+c\text{-}command]}|\mathcal{D}) = 0.5$ and $p(x_{[-local,-c\text{-}command]}|\mathcal{D}) = 0.5$. In this way the results from Experiment 1 provide us with an informative prior distribution regarding which entities are likely to be referred to in the discourse, and through simulations we can test whether this prior knowledge helps an ideal learner acquire pronoun categories[3].

Inference

Given this generative process, we can use Bayesian inference to recover the learner's beliefs about pronoun categories. We use the Gibbs sampling algorithm from Rosen-Zvi et al. (2004) to estimate latent variables: the antecedent-category parameter θ, the category-word parameter ϕ, the antecedent's syntactic position x, and the pronoun category z. The assignments of x and z for a particular token are sampled as a block, conditioned on everything else, so that in each iteration we compute the conditional distribution $p(x_i, z_i | w_i, \mathbf{x}_{-i}, \mathbf{z}_{-i})$ where $\mathbf{x}_{-i}$ and $\mathbf{z}_{-i}$ denote all syntactic position and category

[2] In cases of non-reflexive guesses where potential antecedents appeared in multiple syntactic positions, we assumed the prior probability for each syntactic position to be proportional to the number of potential antecedents in that position.

[3] This prior distribution differs from the distribution seen in Tables 1 and 2 because it is based on individual experimental items rather than on aggregated data.

assignments not including the *i*th pronoun. This is proportional to

$$p(w_i|x_i, z_i, \mathbf{x}_{-i}, \mathbf{z}_{-i}) \cdot p(z_i|x_i, \mathbf{x}_{-i}, \mathbf{z}_{-i}) \cdot p(x_i|\mathcal{D})$$

where the first term is the likelihood function from Rosen-Zvi et al. (2004), the second is defined by the hierarchical Dirichlet process (Teh et al., 2006), and the third is estimated directly from participants' responses in Experiment 1.

Simulations

In order to test the effectiveness of discourse information for the categorization of pronouns, our simulations compare three models: a Baseline model, a Discourse model, and a Strong syntax model. The Baseline model has information about locality and c-command, but it lacks information about which entities are likely to be referred to in the discourse. It assumes that potential antecedents are sampled uniformly, so that $p(x_i|\mathcal{D})$ is defined by counting the number of discourse entities that appear in each syntactic position. The Discourse model is identical to the Baseline model, but it contains the adult-like discourse knowledge estimated in Experiment 1, as described above. Comparing the performance of the Discourse model to the Baseline model allows us to quantify the degree to which discourse information helps an ideal learner acquire pronoun categories.

The Strong syntax model is similar to the Baseline model in that it assumes that potential antecedents are sampled uniformly, but it additionally incorporates built-in knowledge of the grammatical constraints on reflexive and non-reflexive pronouns in English. This model knows there are two grammatical categories of pronouns. Furthermore, it knows that pronouns that have local c-commanding antecedents are reflexive pronouns and that pronouns that do not have local c-commanding antecedents are non-reflexive pronouns (i.e., the antecedent-category parameter θ is observed). Thus, the model only needs to learn the distribution of each category over pronouns. Comparing this Strong syntax model to the Baseline model allows us to examine whether this type of strong prior syntactic knowledge is sufficient to help learners categorize pronouns.

Each model was trained on 50 dialogues from Experiment 1, 25 with reflexive and 25 with non-reflexive pronouns. For each dialogue, the model was provided with the pronoun, a prior distribution over possible antecedents for that pronoun, and the syntactic positions of those antecedents relative to the pronoun. Through the unsupervised learning procedure described above, the models recovered a distribution over categories associated with each syntactic position and a distribution over pronouns for each category.

Results For each model, we ran 10 independent Gibbs chains for 2000 iterations each. Hyperparameters α, β, and γ were fixed at 1.0, 0.01, and 0.001, respectively[4]. We computed pairwise F-scores using the final samples from each chain. The Baseline model consistently failed to learn the correct categories, achieving a mean pairwise F-score of 0.55 across the 10 sampling chains. In all 10 chains, the model learned 3-4 categories, where the correct number of categories is two. Table 3 shows the distribution over pronouns belonging to each category obtained at the 2000th iteration of the sampling run with the highest likelihood. The maximum likelihood estimate $p(word|category)$ gives the proportion of times each pronoun occurs in a category, based on a single sample from the posterior distribution over z and x.

Category 1		Category 2	
Word	p(word|category)	Word	p(word|category)
him	0.5	yourself	0.4
me	0.29	himself	0.28
them	0.21	myself	0.16
us	0.0	themselves	0.12
you	0.0	us	0.04
myself	0.0	me	0.0
yourself	0.0	you	0.0
himself	0.0	him	0.0
themselves	0.0	them	0.0
ourselves	0.0	ourselves	0.0

Category 3	
Word	p(word|category)
you	0.91
ourselves	0.09
me	0.0
us	0.0
him	0.0
them	0.0
myself	0.0
yourself	0.0
himself	0.0
themselves	0.0

Table 3: Baseline model results

The Discourse model performed much better than the Baseline model, achieving a mean pairwise F-score of 0.97 across the 10 sampling runs. In seven of the 10 runs, the model perfectly categorized English pronouns into two classes. In two additional runs, the model learned two categories, but the membership was not consistent. In the final run, the model learned three categories. Table 4 shows the pronouns belonging to each category, obtained at the 2000th iteration of the Gibbs sampling run which had the highest likelihood. The pronouns associated with each category are reflexive pronouns and non-reflexive pronouns, respectively. This model also learned that there are exactly two categories, as expected. These results indicate that discourse information can help an ideal learner categorize pronouns.

Although the Baseline model has prior knowledge of c-command and locality, it is still possible that the low performance in this model might result from insufficient syntactic knowledge. For this reason, we compare the Strong syntax model with the Baseline model to see whether even stronger prior syntactic knowledge is sufficient for categorizing pronouns. The mean F-score was 0.56 for this model. Table 5 shows the pronouns in each category, obtained at the 2000th iteration of a Gibbs sampling run which had the highest likelihood. The lack of improvement of the Strong syntax model over the Baseline model suggests that simply having strong prior syntactic knowledge is not sufficient for acquiring grammatical categories of pronouns.

[4]We chose the best parameter values based on multiple runs, but results were qualitatively consistent across a range of parameter values. The same parameter values were used for all three models.

Category 1		Category 2	
Word	p(word|category)	Word	p(word|category)
yourself	0.4	you	0.4
himself	0.28	him	0.28
myself	0.16	me	0.16
themselves	0.12	them	0.12
ourselves	0.04	us	0.04
you	0.0	yourself	0.0
him	0.0	himself	0.0
me	0.0	myself	0.0
them	0.0	themselves	0.0
us	0.0	ourselves	0.0

Table 4: Discourse model results

Category 1		Category 2	
Word	p(word|category)	Word	p(word|category)
yourself	0.29	you	0.63
him	0.21	me	0.25
himself	0.21	us	0.06
myself	0.12	ourselves	0.06
them	0.09	him	0.0
themselves	0.09	them	0.0
me	0.0	myself	0.0
us	0.0	yourself	0.0
you	0.0	himself	0.0
ourselves	0.0	themselves	0.0

Table 5: Strong syntax model results

These simulation results suggest that knowing which entities are likely to be referred to in the discourse can help learners acquire grammatical categories of pronouns. On the other hand, simply having strong prior knowledge about the grammatical distribution of pronouns is not sufficient to support the acquisition of pronoun categories.

Discussion

This study examined the potential utility of discourse information as a cue to the acquisition of pronoun categories. We showed that discourse information can help adults accurately guess the identities of missing pronouns, and that a Bayesian model with prior knowledge of discourse information can accurately recover grammatical categories of pronouns without knowing in advance how many categories are present in a language. This supports a role for discourse information in helping learners acquire grammatical knowledge of pronoun categories and shows one way in which they can overcome the circularity problem inherent to language acquisition at the syntax-semantics interface.

While it is possible that hearing a few unambiguous sentences could also be sufficient for acquiring pronoun categories, our analysis shows that this type of unambiguous data may not be required. Instead, an ideal learner can achieve the same outcome by relying on the discourse information that is actually present in child-directed speech. The data used in our analysis were taken from CHILDES, and therefore provide a good characterization of input a child receives. However, one limitation of our work is that distributions of verbs and pronouns were balanced in our experimental stimuli, whereas they may not be balanced in the input. To ensure that the true distributions of verbs and pronouns support learning, it will be important to replicate our modeling results on more extensive corpora.

Our model assumed that learners have prior knowledge of the relevance of syntactic locality and c-command relations to the acquisition of pronouns, but we do not know the degree to which this parallels children's acquisition. Children appear to have acquired relevant locality constraints on pronouns by age five at the latest (Zukowski, McKeown, & Larsen, 2008), though we do not know when knowledge of the domains themselves becomes available to learners. Knowledge of c-command also appears to be available to children at this age or even earlier (Lidz & Musolino, 2002; Sutton, Fetters, & Lidz, 2012). However, cross-linguistically, locality and c-command are neither necessary nor sufficient for defining the distributions of grammatical categories of pronouns. Future modeling work will explore the potential role of discourse as an evidentiary source not only in discovering categories of pronouns, but also in determining which grammatical features are relevant for anaphoric dependencies.

Acknowledgments. We thank Viet-An Nguyen, Ke Zhai, Motoki Shiga, and members of the UMD Computational Psycholinguistics group for helpful comments and discussion. This research was supported by NSF IGERT 0801465 and NSF grant 1018625 (JBG).

References

Arnold, J., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order of mention in pronoun processing. *Language and Cognitive Processes*, *22*, 527–565.

Chomsky, N. (1973). Conditions on Transformations. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle.* New York: Holt, Rinehart & Winston.

Crain, S., & McKee, C. (1985). The acquisition of structural restrictions on anaphora. In S. Berman, J. Choe, & J. McDonough (Eds.), *Proceedings of NELS* (Vol. 15). Amherst, Mass: GLSA.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*, 135–176.

Jakubowicz, C. (1984). On markedness and the binding principles. In *Proceedings of NELS* (Vol. 14, pp. 154–182).

Kako, E. (2005). Information sources for noun learning. *Cognitive Science*, *29*, 223–260.

Lidz, J., & Musolino, J. (2002). Children's command of quantification. *Cognition*, *84*, 113–154.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. Third Edition.* (Tech. Rep.). Mahwah, NJ: Lawrence Erlbaum Associates.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Piccin, T. B., & Waxman, S. R. (2007). Why nouns trump verbs in word learning: New evidence from children and adults in the Human Simulation Paradigm. *Language Learning & Development*, *3(4)*, 295–323.

Pitman, J. (2002). Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition. *Combinatorics, Probability and Computing*, *11*, 501–514.

Reinhart, T. (1976). *The Syntactic Domain of Anaphora*. Unpublished doctoral dissertation, MIT.

Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P. (2004). The author-topic model for authors and documents. In *20th Conference on Uncertainty in Artificial Intelligence.*

Sutton, M., Fetters, M., & Lidz, J. (2012). Parsing for Principle C at 30 months. In *Proceedings of the 36th Boston University Conference on Language Development* (pp. 581–593).

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, *101*.

Zukowski, A., McKeown, R., & Larsen, J. (2008). A tough test of the locality requirement for reflexives. In *Proceedings of the 32nd Boston University Conference on Language Development* (pp. 586–597).