

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Pronouns Predict Verb Meanings in Child-Directed Speech

Permalink

<https://escholarship.org/uc/item/6002c0s3>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN

1069-7977

Authors

Laakso, Aarre
Smith, Linda

Publication Date

2004

Peer reviewed

Pronouns Predict Verb Meanings in Child-Directed Speech

Aarre Laakso (alaakso@indiana.edu)

Linda Smith (smith4@indiana.edu)

Department of Psychology, 1101 E. 10th Street
Bloomington, IN 47405 USA

Abstract

Do statistical regularities between pronouns and verbs help children learn verb meanings? This question is addressed by an analysis of child directed speech. The results show that there are statistical regularities in the co-occurrences of pronouns and verbs that could be used to cue verbs that describe physical motion, psychological states, and features such as transitivity. The learnability of these regularities is demonstrated in a simulation study.

Introduction

It is well known that learning the meanings of verbs is a difficult task for young children. It is also well known that pronouns make up a substantial proportion of the nouns that children hear. The distributional relations between pronouns and verbs thus could play a role in early verb learning.

There are several reasons why verb meanings are difficult for children to learn. Whereas parents label objects (relatively) often, they rarely label events or relations. There are no observable referents for many verbs, such as psychological state verbs like *look*, *think*, *want*, *believe*, and *know*. Even verbs that refer to observable actions present ambiguities—for example, when does the opening of a door begin? Further, the aspect of an action that is relevant is ambiguous, and could be, for example, the manner or the path. Finally, verb meaning often depends on taking a particular perspective on a scene; consider the difference between “giving” and “receiving.” In brief, meaning maps between verbs and the world are not transparent. Accordingly, many have suggested that word-word relations are particularly important to learning verbs (see, for example Gleitman, 1990; Gleitman & Gillette, 1995). Here we examine how statistical relations between pronouns and verbs in parental speech might help children learn the meanings of the verbs.

Pronouns are, by far, the most common syntactic subjects and objects in adult speech to children. Most syntactic subjects in spontaneous spoken adult discourse in general are pronouns (Chafe, 1994), and English-speaking mothers often begin with a high-frequency pronoun when speaking to their children, with *you* and *I* occurring most frequently (Valian, 1991). The sheer frequency of pronouns suggests that pronouns—and their statistical co-occurrences with verbs—may be developmentally very powerful.

Consistent with this idea, Childers & Tomasello (2001) suggested that children acquire lexically specific frames such as “I do it” as a way into learning syntactic frames. Cameron-Faulkner, Lieven, & Tomasello (2003) also

observed that parents use the inanimate pronoun *it* far more frequently as the subject of an intransitive sentence than of a transitive one. As Cameron-Faulkner et al. note, this suggests that intransitive sentences are used more often than transitives for talking about inanimate objects. It also suggests that the use of the inanimate pronoun might serve as a cue to some semantic aspects of the verb.

Pronouns may also help learners partition verbs that express psychological attitudes toward events and states of affairs into two rough categories. Verbs that express deontic status, such as goals, purposes or intentions (*try to*), volitions or desires (*want to*), and compulsions (*have to*) tend to take infinitival complements, whereas verbs that express epistemic status, such as perceptions (*see that*), beliefs (*think that*), and knowledge (*know that*) tend to take sentential (propositional) complements (Tomasello, 2003). In the ecology of early childhood, parents tend to be the ones who *know* whereas children tend to be the ones who *need*. All this suggests the potential value of examining the distributional relations among pronouns and verbs in language to young children.

Experiment 1

The first experiment consisted of a corpus analysis to demonstrate patterns of co-occurrence between pronouns and verbs in the child’s input.

Method

Parental utterances from the CHILDES database (MacWhinney, 2000) were coded for syntactic categories, then subjected to clustering and statistical analysis. The target children represented in the transcripts were aged approximately 1;4 – 6;1.

Materials The following corpora were used: Bates, Bliss, Bloom 1970, Brown, Clark, Cornell, Demetras, Gleason, Hall, Higginson, Kuczaj, MacWhinney, Morisset, New England, Post, Sachs, Suppes, Tardiff, Valian, Van Houten, Van Kleeck and Warren-Leubecker.¹

Coding was performed using a custom web application that randomly selected transcripts, assigned them to coders, collected coding input, and stored it in a MySQL database. The application occasionally assigned the same transcript to all coders, in order to measure reliability. Five undergraduate coders were trained on the coding task and

¹ The full references for each corpus may be found in (MacWhinney, 2000).

the use of the system. Cluster analysis and other statistical analyses were performed using MATLAB and R.

Procedure Each coder was presented, in sequence, with each main tier line of each transcript she was assigned, together with several lines of context; the entire transcript was also available for viewing by clicking a link on the coding page. For each line, she indicated (a) whether the speaker was a parent, target child, or other; (b) whether the addressee was a parent, target child, or other; (c) the syntactic frames of up to 3 clauses in the utterance; (d) for each syntactic frame, up to 3 subjects, auxiliaries, verbs, direct objects, indirect objects and obliques. Nouns appearing in prepositional phrases were coded as obliques (with the exception of recipients indicated with “to”, which were coded as indirect objects). Object complements were indicated by coding the direct object of the matrix verb as “<clause>” and coding the constituents of the complement clause as the next clause associated with the utterance. This was intended both to simplify the coding scheme and to avoid attributing too much grammatical knowledge to the child—we do not assume that the child can convert an utterance into an accurate parse tree, only that she can identify verbs and the nouns that surround them.

The syntactic frames were: no verb, question, passive, copula, intransitive, transitive and ditransitive. The *no verb* frame included clauses—such as “Yes” or “OK”—with no main verb. The *question* frame included any clause using a question word—such as “Where did you go?”—or having inverted word order—such as “Did you go to the bank?”—but not merely a question mark—such as “You went to the bank?” The *passive* frame included clauses in the passive voice, such as “John was hit by the ball.” The *copula* frame included clauses with a copula (including *be*, *seem* and *become*) as the main verb, such as “John is angry.” The *intransitive* frame included clauses with no direct object, such as “John ran.” The *transitive* frame included clauses with a direct object (or an object complement) but no indirect object, such as “John hit the ball.” The *ditransitive* frame included clauses with an indirect object, such as “John gave Mary a kiss.”

In total, 59,977 utterances were coded from 123 transcripts. All of the coders coded 7 of those transcripts for the purpose of measuring reliability. Average inter-coder reliability (measured for each coder as the percentage of items coded exactly the same way they were coded by another coder) was 86.1%.²

We only considered parental child-directed speech (PCDS), defined as utterances where the speaker was a

parent and the addressee was a target child. A total of 24,286 PCDS utterances were coded, for a total of 28,733 clauses. More than a quarter (28.36%) of the PCDS clauses contained no verb at all; these were excluded from further analysis. Clauses that were questions (16.86%), passives (0.02%), and copulas (11.86%) were also excluded from further analysis. The analysis was conducted using only clauses that were intransitives (17.24% of total PCDS clauses), transitives (24.36%) or ditransitives (1.48%), a total of 12,377 clauses.

We formed 2 matrices from these clauses: a verbs-by-subjects matrix and a verbs-by-objects matrix. The verbs-by-subjects matrix contained only verbs used with an overt subject; its size was 621 verbs by 317 nouns (subjects). The verbs-by-objects matrix contained only verbs used with a direct object; its size was 524 verbs by 907 nouns (objects). Each cell of each matrix contained the proportion of times that verb was used with that noun (as subject or object) in a coded clause.

For the purposes of exploratory data analysis, we then performed 4 cluster analyses. First, we took the 50 nouns most commonly used as objects and clustered them according to their proximity in verb space, i.e., the space formed by considering each verb as a dimension. Each noun was placed along each dimension according to the proportion of times it was used with the corresponding verb. Hence, a noun never used as the object of a particular verb would be at 0, and a noun always used as the object of a particular verb would be at 1. Second, we clustered the 50 most common subject-nouns in verb space. Third, we took the 50 verbs most commonly used with objects and clustered them according to their proximity in noun space (defined analogously to verb space). Finally, we clustered the 50 most common verbs-with-subjects in noun space.

Results

We cannot show the cluster diagrams here due to space limits. We summarize the main regularities observed.

The observed distribution of nouns in the corpus is consistent with Zipf’s law — the numerical frequency of words decays roughly as an inverse power of their rank frequency. Moreover, the most frequent subjects and objects in the corpus are pronouns, as shown in Figures 1 and 2.

Our quantitative analysis of co-occurrence relationships is based on the log likelihood ratio as described by Dunning (1993) and recommended by Manning & Schütze (1999). Suppose we have observed N clauses with m subject-(or object-) types and n verb-types. Let

$$S = s_1, s_2, \dots, s_m \text{ and } V = v_1, v_2, \dots, v_n$$

represent the subjects and the verbs respectively. Furthermore, let

$$K_S = k_{s_1}, k_{s_2}, \dots, k_{s_m} \text{ and } K_V = k_{v_1}, k_{v_2}, \dots, k_{v_n}$$

represent the observed frequencies of the subjects and verbs respectively and

$$K_{SV} = k_{s_1v_1}, k_{s_1v_2}, \dots, k_{s_1v_n}, k_{s_2v_1}, \dots, k_{s_mv_n}$$

² Cohen’s kappa coefficient (Cohen, 1960), which adjusts for chance agreement, is only applicable for two raters. We know of no chance-corrected multiple rater agreement measures that are widely used in the language acquisition literature. However, given the number of variables, the number of levels of each variable (3 speakers, 3 addressees, 7 frames, and 6 syntactic relations), and the number of coders (5) in the present study, the probability of chance agreement is low.

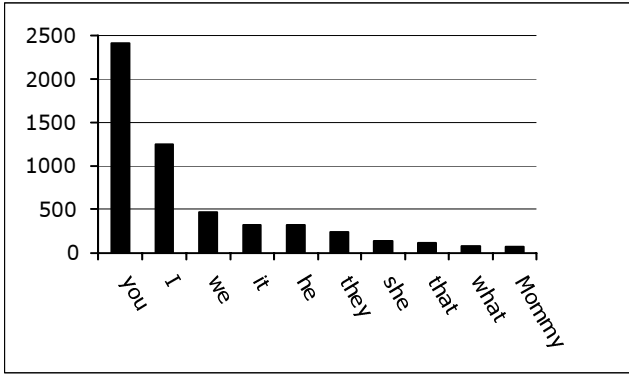


Figure 1: The 10 most frequent subjects in PCDS by their number of occurrences.

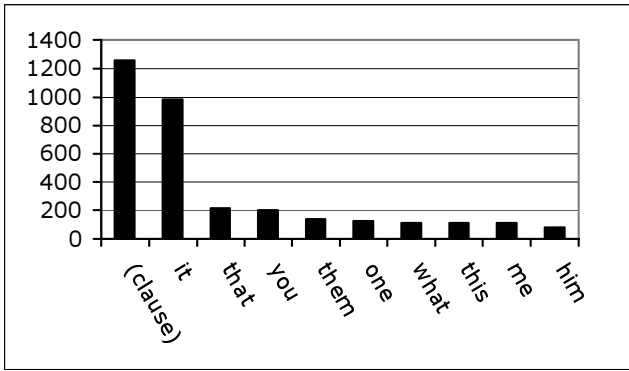


Figure 2: The 10 most frequent objects in PCDS by their number of occurrences.

represent the observed frequencies of subject-verb co-occurrences. Then the test statistic is the quantity:

$$-2 \log \lambda = 2[\log L(p_1, k_{s_i v_j}, k_{v_j}) + \log L(p_2, k_{s_i} - k_{s_i v_j}, N - k_{v_j}) - \log L(p_0, k_{s_i v_j}, k_{v_j}) - \log L(p_0, k_{s_i} - k_{s_i v_j}, N - k_{v_j})]$$

Where the components are defined as follows:

$$L(p, k, n) = p^k (1 - p)^{n-k} = k \log p + (n - k) \log(1 - p)$$

$$p_0 = p(s_i) = \frac{k_{s_i}}{N}$$

$$p_1 = p(s_i | v_j) = \frac{k_{s_i v_j}}{k_{v_j}}$$

$$p_2 = p(s_i | \neg v_j) = \frac{k_{s_i} - k_{s_i v_j}}{N - k_{v_j}}$$

The test statistic $-2 \log \lambda$ is χ^2 distributed with 1 degree of freedom. Intuitively, it represents how much more likely it is that s_i and v_j go together than should be expected purely

by chance. It has also been demonstrated that this statistic identifies natural collocations in text corpora.

As expected, the inanimate pronoun “it” was more likely as the object of verbs of physical motion than as the object of psychological attitude verbs, whereas complement clauses were more likely to occur with psychological attitude verbs than with verbs of physical motion. As shown in Table 1, “it” tended to occur with physical motion verbs far more often than would be predicted by chance, and clauses occurred with most physical motion verbs, if at all, only about as much as would be predicted by chance. The verb “put” is an exception to this general rule, occurring with a clause more often than would be predicted by chance. As shown in Table 2, complement clauses tended to occur with psychological attitude verbs more often than would be predicted by chance, whereas “it” only occurred more often than would be predicted by chance with two of five psychological attitude verbs. The exceptions were *want* (uses such as “Oh, I want it now”) and *know* (“No, that’s wrong and you know it”). In any case, as shown in Figure 3, it is somewhat more likely that a physical motion verb will occur with “it” than with a complement clause, and substantially more likely that a psychological attitude verb will occur with a complement clause than with “it”.

Table 1: The log likelihood ratio for uses of object “it” or a clause with physical motion verbs. * indicates $p < 0.01$; — indicates no co-occurrences.

	“it”	(clause)
put	102.79*	70.70*
turn	72.58*	—
throw	39.55*	6.14
hold	32.17*	—
push	24.87*	3.02

Table 2: The log likelihood ratio for uses of object “it” or a clause with psychological attitude verbs. * indicates $p < 0.01$; — indicates no co-occurrences.

	“it”	(clause)
think	—	399.13*
want	12.00*	283.28*
know	69.53*	134.44*
remember	—	37.22*
mean	0.91	15.81*

Table 3: The log likelihood ratio for uses of subject “I” or “you” with epistemic verbs. * indicates $p < 0.01$; — indicates no co-occurrences.

	“I”	“You”
think	605.01*	24.7*
know	200.05*	108.17*
guess	60.00*	—

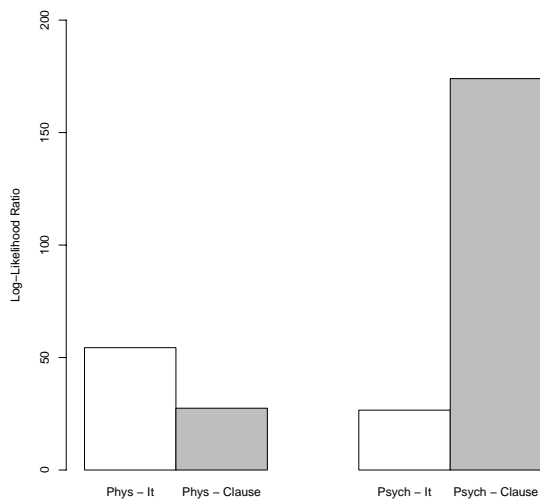


Figure 3: Mean log-likelihood ratio of the object pronoun “it” (white bars) versus a clause (gray bars) given a physical motion verb (left) or a psychological attitude verb (right).

We also found that “I” is more likely to be the subject of epistemic verbs, whereas “you” is more likely to be the subject of deontic verbs. As shown in Table 3, “I” occurred with epistemic verbs far more often than would be predicted by chance. The subject “you” also occurred more often with “think” and “know” than would be predicted by chance, but with a much lower likelihood.

As shown in Table 4, “you” tended to occur with deontic verbs far more often than chance would predict. The subject “I” was no more likely than chance would predict to appear with the verbs “like” and “need” and was only slightly more likely than chance to occur with the verb “want”. In any case, as demonstrated in Figure 4, it is substantially more likely that the subject “I” will appear with an epistemic verb than it is that the subject “you” will appear with an epistemic verb. It is also somewhat more likely that “you” will appear with a deontic verb than “I” will appear with a deontic verb.

Table 4: The log likelihood ratio for uses of subject “I” or “you” with deontic verbs. * indicates $p < 0.01$.

	“I”	“You”
want	6.72*	116.97*
like	0.03	74.24*
need	2.69	15.26*

We conclude by noting that there are many other significant co-occurrences in the corpus, some of which involve triadic correlations between specific verbs, specific nouns, and pronouns. For example the objects “book” and “story” are more likely to appear with the verb “read” than would be predicted by chance (LLR=131.51, 128.39). Both the object “book” and the object “this” are likely to appear with the phrasal verb “look at” (LLR=67.28, 88.01).

Similarly, not only is “it” likely to appear as the object of “turn” (as discussed above), but so is “page” (LLR=81.89). Likewise for “play,” which makes not only the objects “ball”, “blocks”, “game”, and “house” more likely, but also the objects “this” and “it”. These are potentially important on several fronts. The child may learn an association between pronouns such as “this” and “it” and inanimate objects, like books and pages. The pronouns “this” and “it” may then be used to help the child understand the meanings of other verbs that take inanimate objects as their objects. Conversely, the verb “tell” selects strongly for the pronouns “us” and “me” as well as for “Mommy” and “Daddy”. Hence, the child may learn that verbs taking “us” and “we” as objects have to do with communicating with or directing attention toward other people.

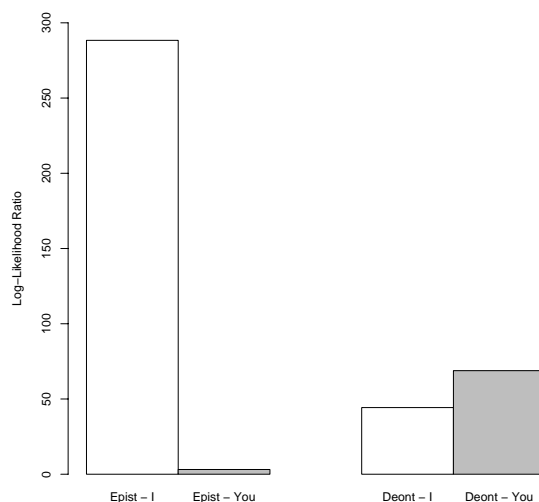


Figure 4: Mean log-likelihood ratio of the subject pronoun “I” (white bars) versus “you” (gray bars) given an epistemic verb (left) or a deontic verb (right).

Discussion

Although pronouns are “light” in their meaning, their referents determinable only from context, they may nonetheless be potent forces on early lexical learning by identifying some kinds of verb meanings as more likely than others. The results of Experiment 1 show that there are statistical regularities in the co-occurrences of pronouns and verbs that the child could use to discriminate verb meanings. Verbs that describe physical motion or transfer are likely to be followed by “it,” whereas verbs attributing psychological state are likely to be followed by a relatively complex complement clause. Verbs having to do with thinking or knowing are likely to occur with subject “I,” whereas verbs having to do with wanting or needing are likely to occur with subject “you.” This regularity most likely reflects the ecology of parents and children—parents “know” and children “want”—but it could nonetheless be useful in distinguishing these two classes of meanings. The results thus far show that there are potentially usable regularities in the statistical relations between pronouns and verbs.

Experiment 2

To demonstrate that the regularities in pronoun-verb co-occurrences in parental speech to children can actually be exploited by a statistical learner, we trained a connectionist network to auto-associate subject-verb-object “sentences” from the input, then tested it on individual verbs and pronouns. We predict that the network should be able to learn the statistical regularities demonstrated in Experiment 1, specifically: (1) physical transfer verbs are likely to have “it” as an object, whereas psychological verbs are likely to take a complement clause, and (2) epistemic verbs are likely to have “I” as a subject, whereas deontic verbs are likely to have “you” as a subject.

Method

Data The network training data consisted of the subject, verb, and object of all coded utterances that contained the 50 most common subjects, verbs and objects. There were 5,835 such utterances. The inputs used a localist coding wherein there was exactly one input unit out of 50 activated for each subject, and likewise for each verb and each object. Absent and omitted arguments were counted among the 50, so, for example, the utterance “John runs” would have 3 units activated even though it only has 2 words—the third unit being the “no object” unit. With 50 units each for subject, verb and object, there were a total of 150 input units to the network. Active input units had a value of 1, and inactive input units had a value of 0.

Network Architecture The network consisted of a two-layer 150-8-150 unit autoassociator with a logistic activation function at the hidden layer and three separate softmax activation functions (one each for the subject, verb and object) at the output layer—see Figure 5. Using the softmax activation function, which ensures that all the outputs in the bank sum to 1, together with the cross-entropy error measure allows us to interpret the network outputs as probabilities (Bishop, 1995). The network was trained by the resilient backpropagation algorithm (Riedmiller & Braun, 1993) to map its inputs back onto its outputs. It is well known that this sort of network performs nonlinear dimensionality reduction at its hidden layers, extracting statistical regularities from the input data.

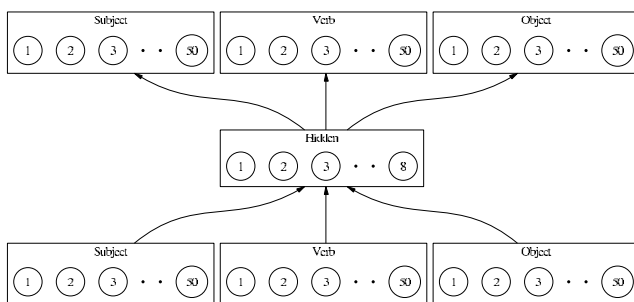


Figure 5: Network architecture

Training The data was randomly assigned to two groups: 90% was used for training the network, while 10% was reserved for validation. Starting from different random initial weights, 10 networks were trained until the cross-entropy on the validation set reached a minimum for each of them. Training stopped after approximately 150 epochs of training, on average. At that point, the networks were achieving about 81% accuracy on correctly identifying subjects, verbs and objects from the training set.

Testing After training, the networks were tested with incomplete inputs corresponding to isolated verbs and pronouns. For example, to see what a network had learned about *it* as a subject, it was tested with a single input unit activated—the one corresponding to *it* as subject. The other inputs were set to 0. Output unit activations were recorded and averaged over all 10 networks.

Results

To test the hypothesis that the network learns that psychological attitude verbs are more likely than physical motion verbs to take a clause as an object, we tested the networks with the frames “I ___ (clause)” and “You ___ (clause)” using psychological and physical verbs. The psychological verbs were “think,” “want,” “know,” and “remember.” The verb “mean” was not among the top 50 verbs used in the corpus and therefore was not used in the network training experiments. The physical verbs were “put,” “turn,” “throw” and “hold.” The verb “push” was not among the top 50 verbs used in the corpus and therefore was not used in the network training experiments. As shown in Figure 6, the networks activated the psychological verbs more strongly at the output ($\bar{M} = 0.047$, $SD = 0.152$) than the physical verbs ($\bar{M} = 0.002$, $SD = 0.014$). This difference was significant, $t(80) = -2.62$, $p = 0.01$. Results are similar for the converse (physical verbs are significantly more activated when the object is “it”) and for the epistemic / deontic distinction (epistemic verbs are significantly more activated when the subject is “I,” whereas deontic verbs are significantly more activated when the subject is “you”).

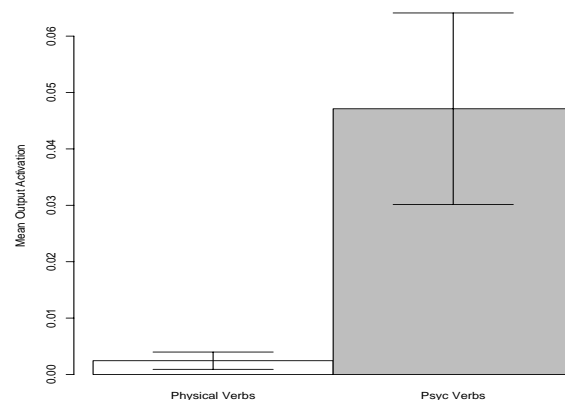


Figure 6: Network output activations for physical verbs versus psychological attitude verbs for the frames “I ___ (clause)” and “You ___ (clause)”.

Conclusions

We have shown that there are statistical regularities in co-occurrences between pronouns and verbs in the speech that children hear from their parents, including regularities that distinguish between psychological and non-psychological verbs, as well as between deontic and epistemic psychological verbs. We have also shown that a simple statistical learner can exploit these regularities, not to learn the meanings of verbs per se (the network obviously does not know the meanings of the verbs), but to learn the formal associations between tokens of verbs and pronouns. How might this help the child learn the meanings of verbs? In the first place, hearing the verb framed by pronouns may help the child isolate the relevant event or action from the blooming, buzzing confusion around it; the pronouns can indicate animacy, gender, number and direction of causality. This would allow the child to focus on the relevant things. Second, if we suppose that the child has already learned one verb and its pattern of correlations with pronouns, and then hears another verb being used with the same or a similar pattern of correlations, the child may hypothesize that the meaning of the unknown verb is similar to the meaning of the known verb. For example, a child who understood “want” but not “need” might observe that “you” is usually the subject of both and conclude that “want,” like “need,” has to do with his desires and not, for example, a physical motion or someone else’s state of mind.

Now that we have shown that the regularities exist and are learnable in principle, the next step is to show that children actually pick up on these regularities. We are planning a series of experiments with children from 24-36 months that will involve priming the children with movies showing actions corresponding to nonsense verbs in the context of various pronoun frames (“He zorps it,” “It zorps,” “They zorp this to her” and so on) and testing whether this influences their interpretations of these verbs, for example by having them select which of a pair of novel movies shows “zorping.” This will provide a strong test of the hypothesis that children actually use pronoun distributional statistics to pick up on the meanings of novel verbs.

Future modeling experiments will attempt to capture not only the statistical relationships among verb-pronoun token co-occurrences but also their relationships with shared meanings, by associating words with featural representations of their meanings. We are also working on a mechanism for manipulating the statistical properties of the relevant conditional distributions to be used for generating network inputs in future simulations—the distribution of syntactic frames, the distribution of verbs given a syntactic frame, and the distribution of nouns (including pronouns) in each argument position given a verb. We expect that such a model could be used to test the utility of statistical associations between pronouns with verbs for a theoretical learner—simulations run with varying degrees of correlations could demonstrate not only whether but also just how much correlation is useful in principle.

Finally, work is underway to collect crosslinguistic data from Japanese and Tamil, verb-heavy languages with

frequent argument dropping and case-marked pronouns referring to various levels of social status. We are especially keen to find out what sorts of cues children might be using in languages where pronouns are both rarer and “heavier” than they are in English.

Acknowledgments

This research was supported by NIMH grant number ROI MH 60200. Thanks to our coders for their hard work, to members of the Cognitive Development Laboratory at IU for useful discussions, and to three anonymous reviewers for thoughtful comments on an earlier draft.

References

- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.
- Cameron-Faulkner, T., Lieven, E. V. M., & Tomasello, M. (2003). A construction-based analysis of child directed speech. *Cognitive Science*, 27, 843-873.
- Chafe, W. L. (1994). *Discourse, Consciousness and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*, 37(6), 739-748.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), 61-74.
- Gleitman, L. R. (1990). The structural sources of word meaning. *Language Acquisition*, 1(1), 3-55.
- Gleitman, L. R., & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 413-427). Cambridge, MA: Blackwell.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed. Vol. 2: The Database). Mahwah, NJ: Lawrence Erlbaum Associates.
- Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.
- Riedmiller, M., & Braun, H. (1993). *A direct adaptive method for faster backpropagation learning: The Rprop algorithm*. Paper presented at the IEEE International Conference on Neural Networks 1993 (ICNN 93), San Francisco, CA.
- Tomasello, M. (2003). *Constructing a Language*. Cambridge, MA: Harvard University Press.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40(1-2), 21-81.