On a possible role for pronouns in the acquisition of verbs

Aarre Laakso and Linda Smith
Department of Psychology
1101 E. 10th Street
Bloomington, IN 47408
{alaakso,smith4}@indiana.edu

Abstract

Given the restrictions on the subjects and objects that any given verb may take, it seems likely that children might learn verbs partly by exploiting statistical regularities in co-occurrences between verbs and noun phrases. Pronouns are the most common NPs in the speech that children hear. We demonstrate that pronouns systematically partition several important classes of verbs, and that a simple statistical learner can exploit these regularities to narrow the range of possible verbs that are consistent with an incomplete utterance. Taken together, these results suggest that children *might* use regularities in pronoun/verb co-occurrences to help learn verbs, though whether this is *actually* so remains a topic for further research.

1 Introduction

Pronouns stand for central elements of adult conceptual schemes—as Quine pointed out, pronouns "are the basic media of reference" (Quine, 1980, p. 13). In fact, most syntactic subjects in spontaneous spoken adult discourse 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 (e.g., Valian, 1991). Parents use the inanimate pronoun *it* far more frequently as the subject of an intransitive sentence than of an transitive one (Cameron-Faulkner et al., 2003, p. 860). 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, we would note, that the use of the inanimate pronoun might be a cue for the child as to whether the verb is transitive or intransitive. Similarly, Lieven and Pine (Lieven et al., 1997; Pine and Lieven, 1993) have suggested that pronouns may form the fixed element in lexically-specific frames acquired by early language learners—so-to-speak "pronoun islands" something like Tomasello's (1992) "verb islands."

Many researchers have suggested that word-word relations in general, and syntactic frames specifically, are particularly important for learning verbs (e.g., Gleitman, 1990; Gleitman and Gillette, 1995). What has not been studied, to our knowledge, is how *pronouns* specifically may help children learn verbs by virtue of systematic co-occurrences. We have begun to address this issue in two steps. First, we measured the statistical regularities among the uses of pronouns and verbs in a large corpus of parent and child speech. We found strong regularities in the use of pronouns with several broad classes of verbs. Second, using the corpus data, we trained a connectionist network to guess which verb belongs in a sentence given only the subject and object, demonstrating that it is possible in principle for a statistical learner to use the regularities in parental speech to deduce information about an unknown verb.

2 Experiment 1

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

2.1 Method

Parental utterances from the CHILDES database (MacWhinney, 2000) were coded for syntactic categories, then subjected to cluster analysis. The mean age of target children represented in the transcripts that were coded for this experiment was 3;0 (SD≈1;2).

2.1.1 Materials

The following corpora were used: Bates, Bliss, Bloom 1970, Brown, Clark, Cornell, Demetras Working, Gleason, Hall, Higginson, Kuczaj, MacWhinney, Morisset, New England, Post, Sachs, Suppes, Tardiff, Valian, Van Houten, Van Kleeck and Warren-Leubecker. Coding was performed using a custom web application that randomly selected transcripts, assigned them to coders as they became available, 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 the use of the system.

2.1.2 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 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 clause, up to 3 subjects, auxiliaries, verbs, direct objects, indirect objects and obliques. Because many utterances were multi-clausal, the unit of analysis for assessing pronoun-verb co-occurrences was the clause rather than the utterance.

The syntactic frames were: no verb, question, passive, copula, intransitive, transitive and ditransitive. These were considered to be mutually exclusive, i.e., each clause was tagged as belonging to one and only one frame, according to which of the following frames it matched first: (1) The *no verb* frame included clauses – such as “Yes” or “OK” – with no main verb. (2) 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?” (3) The *passive* frame included clauses in the passive voice, such as “John was hit by the ball.” (4) The *copula* frame included clauses with the copula as the main verb, such as “John is angry.” (5) The *intransitive* frame included clauses with no direct object, such as “John ran.” The *transitive* frame included clauses with a direct object but no indirect object, such as “John hit the ball.” (6) The *ditransitive* frame included clauses with an indirect object, such as “John gave Mary a kiss.”

All nouns were coded in their singular forms, whether they were singular or plural (e.g., “boys” was coded as “boy”), and all verbs were coded in their infinitive forms, whatever tense they were in (e.g., “ran” was coded as “run”).

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 each other coder) was 86.1%. 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), the probability of chance agreement is very low. Although there are some substantive errors (usually with complex embedded clauses or other unusual constructions), many of the discrepancies are simple spelling mistakes or failures to trim words to their roots.

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, including 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.

2.2 Results

The most frequent nouns in the corpus—both subjects and objects—are pronouns, as shown in Figures 1 and 2. The objects divided the most common verbs into three main classes: verbs that take the pronoun *it* and concrete nouns as objects, verbs that take complement clauses, and verbs that take specific concrete nouns as objects. The subjects divided the most common verbs into four main classes: verbs whose subject is almost always *I*, verbs whose subject is almost always *you*, verbs that take *I* or *you* almost equally as subject, and other verbs. The verbs divided the most common object nouns into a number of classes, including objects of telling and looking verbs, objects of having and wanting verbs, and objects of putting and getting verbs. The verbs also divided the most common subject nouns into a number of classes, including subjects of having and wanting verbs, and subjects of thinking and knowing verbs.

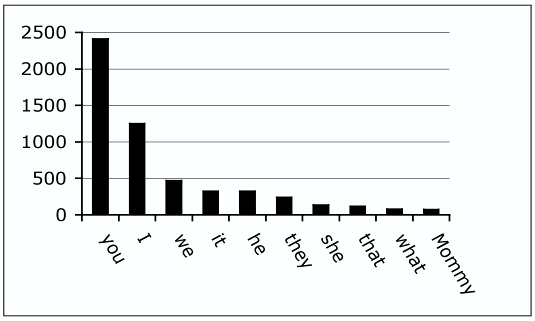

Figure 1: The 10 most frequent subjects in PCDS by their number of occurrences

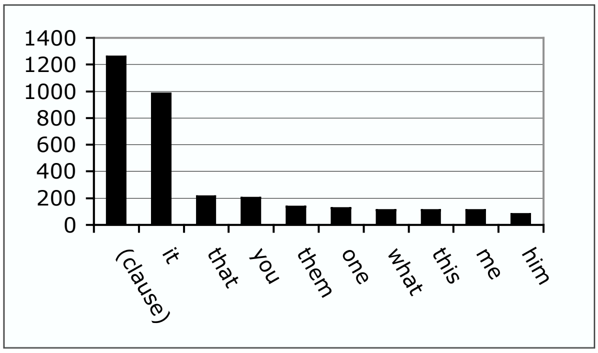

Figure 2: The 10 most frequent objects in PCDS by their number of occurrences.

2.2.1 Verbs that take *it* as an object

The verbs that take *it* as their most common object include verbs of motion and transfer, as shown in Table 1.

2.2.2 Verbs that take complement clauses

Most verbs that did not take *it* as their most common object instead took complement clauses. These are primarily psychological verbs, as shown in Table 2.

2.2.3 Verbs that take concrete nouns as objects

Most remaining verbs in the corpus took unique sets of objects. For example, the most common object used with *read* was *book*, followed by *it* and *story*; the most common object used with *play* was *game*, followed by *it*, *block*, and *house*.

2.2.4 Verbs that take *I* as a subject

Verbs whose most common subject is *I* include *bet* (23 out of 23 uses with a subject, or 100%), *guess* (21/22, 95.4%), *think* (212/263, 80.6%), and *see* (95/207, 45.9%). Parents were not discussing their gambling habits with their children – *bet* was being used to indicate the epistemic status of a subsequent clause, as were the other verbs.

2.2.5 Verbs that take *you* as a subject

Verbs whose most common subject is *you* include *like* (86 out of its 134 total uses with a subject, or 64.2%), *want* (192/270, 71.1%), and *need* (33/65, 50.8%). These verbs are being used to indicate the deontic status of a subsequent clause, including disposition or inclination, volition, and compulsion.

2.2.6 Verbs that take you or I as a subject

Verbs that take *I* and *you* more or less equally as subject include *mean* (15 out of 32 uses, or 46.9%, with *I* and 12 of 32 uses, or 37.5%, with *you*), *know* (*I*: 159/360, 44.2%; *you*: 189/360, 52.5%), and *remember* (*I*: 9/23, 39.1%; *you*: 12/23, 52.2%).

Verb	Total	it (#)	it (%)
turn	56	33	58.9
throw	36	20	55.5
push	25	13	52.0
hold	42	19	45.2
break	36	16	44.4
leave	27	12	44.4
open	36	15	41.7
do	256	105	41.0
wear	25	10	40.0
take off	24	9	37.5
put	276	93	33.7
get	348	74	21.3
take	106	22	20.8
put on	42	8	19.0
buy	50	9	18.0
give	85	14	16.5
have	340	26	7.6

Table 1: Verbs most commonly used with object *it.*

Verb	Total	<clause> (#)	<clause> (%)
think	187	179	95.7
remember	31	23	74.2
let	78	57	73.1
know	207	141	68.1
ask	29	17	58.6
go	55	32	58.2
want	317	183	57.7
mean	25	14	56.0
tell	115	45	39.1
try	51	18	35.3
say	175	53	30.3
look	48	14	29.2
need	64	18	28.1
see	266	73	27.4
like	123	32	26.0
show	36	9	25.0
make	155	23	14.8

Table 2: Verbs most commonly used with complement clauses.

Verb	Total	I (#)	I (%)	you (#)	you (%)
bet	23	23	100	0	0
guess	22	21	95.4	0	0
think	263	212	80.6	38	14.4
see	207	95	45.9	50	24.1
mean	32	15	46.9	12	37.5
know	360	159	44.2	189	52.5
remember	23	9	39.1	12	52.2
like	134	20	14.9	86	64.2
want	270	34	12.6	192	71.1
need	65	5	7.7	33	50.8

Table 3: Some verbs commonly used with subject *I* or *you*.

2.2.7 Objects of *tell* and *look at*

The objects *me*, *us*, *Daddy* and *Mommy* formed a cluster in verb space, appearing frequently with the verbs *tell* and *look at*.

2.2.8 Objects of *put* and *get*

The objects *one*, *stuff*, *box*, and *toy* occurred most frequently with *get*, and frequently with *put*. The objects *them*, *him*, *her*, *bed*, and *mouth* occurred most frequently with *put* and, in some cases, also frequently with *get*.

2.2.9 Objects of *have* and *want*

The objects *cookie*, *some*, *money*, *coffee*, *milk*, and *juice* formed a cluster in verb space, appearing frequently with verbs such as *have* and *want*, as well as, in some cases, *give*, *take*, *pour*, *drink*, and *eat*.

2.2.10 Subjects of *think* and *know*

The subject *I* appeared most frequently with the verbs *think* and *know*.

2.3 Discussion

Although pronouns are semantically "light," their particular referents determinable only from context, they may nonetheless be potent forces on early lexical learning by statistically pointing to some classes of verbs as being more likely than others. The results of Experiment 1 clearly show that there are statistical regularities in the co-occurrences of pronouns and verbs that the child could use to discriminate classes of verbs. Specifically, when followed by *it*, the verb is likely to describe physical motion, transfer, or possession. When followed a relatively complex complement clause, by contrast, the verb is likely to attribute a psychological state. Finer distinctions may also be made with other objects, including proper names and nouns. Verbs followed by *me*, *us*, *Daddy*, and *Mommy* are likely to have to do with telling or looking. Verbs followed by *one*, *stuff*, *them*, *him*, or *her* are likely to have to do with getting or putting. Verbs followed by certain concrete objects such as *cookie*, *milk*, or *juice* are likely to have to do with having or wanting. Fine distinctions may also be made according to subject. If the subject is *I*, the verb is likely to have to do with thinking or knowing, whereas if the subject is *you*, *she*, *we*, *he*, or *they*, the verb is likely to have to do with having or wanting. 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 verbs.

The results thus far show that there are potentially usable regularities in the statistical relations between pronouns and verbs. However, they do not show that these regularities can be used to cue the associated words.

3 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 an autoassociator on the corpus data, then tested it on incomplete utterances to see how well it would "fill in the blanks" when given only a pronoun, or only a verb. An autoassociator is a connectionist network that is trained to take each input pattern and reproduce it at the output. In the process, it compresses the pattern through a small set of hidden units in the middle, forcing the network to find the statistical regularities among the elements in the input data. The network is trained by backpropagation, which iteratively reduces the discrepancies between the network's actual outputs and the target outputs (the same as the inputs for an autoassociator).

In our case, the inputs (and thus the outputs) are subject-verb-object "sentences." Once the network has learned the regularities inherent in a corpus of complete SVO sentences, testing it on incomplete sentences (e.g., "I ___ him") allows us to see what it has gleaned about the relationship between the given parts (subject "I" and object "him" in our example) and the missing parts (the verb in our example).

3.1 Method

3.1.1 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 one and only 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.

3.1.2 Network Architecture

The network consisted of a two-layer 150-8-150 unit autoassociator with a logistic activation function at the hidden layer and a three separate softmax activation functions (one each for the subject, verb and object) at the output layer—see

Figure 3. 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 and Braun, 1993) to map its inputs back onto its outputs. We chose to use eight units in the hidden layer on the basis of some pilot experiments that varied the number of hidden units. Networks with fewer hidden units either did not learn the problem sufficiently well or took a long time to converge, whereas networks with more than about 8 hidden units learned quickly but tended to overfit the data.

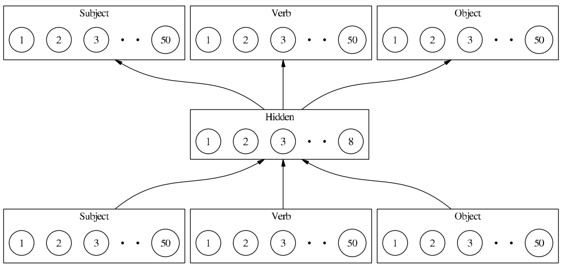

Figure 3: Network architecture

3.1.3 Training

The data was randomly assigned to two groups: 90% of the data was used for training the network, while 10% was reserved for validating the network's performance. Starting from different random initial weights, five 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. Near perfect accuracy on the training set could have been achieved by further training, with some loss of generalization, but we wanted to avoid overfitting.

3.1.4 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 input units were set to 0. Activations at the output units were recorded. The results presented below report average activations over all five networks.

3.2 Results

The networks learn many of the co-occurrence regularities observed in the data. For example, when tested on the object *it* (see Figure 4 on page 7 below), the most activated verbs are *get*, *hold*, *take* and *have*, which are among the most common verbs associated with *it* in the input (see Table 1). Similarly, *tell*, *make* and *say* are the most activated verbs when networks are tested with the *clause* unit activated in the object position (figure not shown), and they are also among the verbs most commonly associated with a *clause* in the input (see Table 2).

However, the network does not merely learn the relative frequencies of pronouns with verbs. For example, the verbs most activated by the subject *you* are *have* and *get* (see Figure 5 on page 8 below), neither of which appears in Table 3. The reason for this, we believe, is that the subject *you* is strongly associated with the object *it* (note the strong activation of *it* in the right column of Figure 5), and the object *it*, as mentioned in the previous paragraph, is strongly associated with the verbs *have* and *get*. The difference may be observed most clearly when the network is prompted simultaneously with *you* as the subject and *clause* as the object (see Figure 6 on page 8 below). In that case, the verb *want* is strongly preferred and, though *get* still takes second place, *tell* and *know* rank third and fourth, respectively—consistent with the results in Table 1. This demonstrates that the network model is sensitive to high-order correlations among words in the input, not merely the first-order correlations between pronoun and verb occurrences.

These results do not depend on using an autoassociation network, and we do not claim that children in fact use an autoassociation architecture to learn language. Any statistical learner that is able to discover higher-order correlations will produce results similar to the ones shown here. An autoassociator was chosen only as a simple means of demonstrating in principle that a statistical learner can extract the statistical regularities from the data.

4 Conclusion

We have shown that there are statistical regularities in co-occurrences between pronouns and verbs in the speech that children hear from their parents. We have also shown that a simple statistical learner can learn these regularities, including subtle higher-order regularities that are not obvious in a casual glance at the input data, and use them to predict the verb in an incomplete sentence. How might this help children learn verbs? In the first place, hearing a verb framed by pronouns may help the child isolate the verb itself—having simple, short consistent, and high-frequency slot fillers could make it that much easier to segment the relevant word in frames like "He ___ it." Second, the information provided by the particular pronouns that are used in a given utterance might help the child isolate the relevant event or action from the blooming, buzzing confusion around it—in English, pronouns can indicate animacy, gender and number, and their order can indicate temporal or causal direction or sequence (e.g., "You ___ it" versus "It ___ you"). Finally, 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 unknown verb is similar to 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. The pronoun/verb co-occurrences in the input may thus help the child narrow down the class to which an unknown verb belongs, allowing the learner to focus on further refining her grasp of the verb through subsequent exposures.

Whether children are actually sensitive to these regularities remains an open question. To the extent that children have actually picked up on the regularities, two predictions should follow. The first is that children's utterances should exhibit roughly the same co-occurrence patterns as we found in their parents' speech to them. Therefore, the next step in our research is to determine whether children are using pronouns and verbs together with roughly the same frequencies that they hear in their parents' speech. This is the subject of research in progress using the coded corpus data from Experiment 1. Because our hypothesis concerns broad-class verb acquisition, we are focusing on children younger than the age of 3, by which time most children can produce the most common verbs (Dale and Fenson, 1996).

The second prediction that follows from the hypothesis that children might be sensitive to the regularities demonstrated in this paper is that children's comprehension of ordinary verbs should be better when they are used in frames that are consistent with the regularities in the input than when they are used in frames that are inconsistent with those regularities. Assessing whether this is true requires an experiment testing children's comprehension of real but relatively infrequent verbs in two conditions: a "consistent" condition (in which the verb is used with nouns or pronouns that are consistent with the regularities in the corpus) and an "inconsistent" condition (in which the verb is used with nouns or pronouns that are inconsistent with the regularities in the corpus). This experiment is in the planning stages.

Even if children are sensitive to the regularities, this knowledge might not help them learn new verbs. That is, whether these regularities actually play a role in language acquisition also remains an open question. To the extent that they do, a third prediction follows: children should be better able to generalize comprehension of *novel* verbs when they are presented in frames consistent with these regularities. We are designing an experiment to test this hypothesis.

The argument that the frequency of pronouns and their co-occurrences with verb classes play a role in the acquisition of verbs could be strengthened by showing that it is true in many languages. The present study considered only English, which is a relatively noun-heavy language in which argument ellipsis is rare. Some other languages, by contrast, tend to emphasize verbs and frequently drop nominal arguments. We are especially keen to find out what sorts of cues children might be using to identify verb classes in such languages. Hence, work is underway to collect comparable data from Japanese and Tamil, verb-heavy languages with frequent argument dropping and case-marked pronouns reflecting various degrees of social status.

5 Acknowledgements

This research was supported by NIMH grant number ROI MH 60200. Additional thanks go to our coders, to members of the Cognitive Development Laboratory at IU for useful discussions of these results, and to several anonymous reviewers for helpful comments.

References

Christopher M. Bishop. 1995. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Thea Cameron-Faulkner, Elena V. M. Lieven, and Michael Tomasello. 2003. A construction-based analysis of child directed speech. *Cognitive Science* 27:843-873.

Wallace L. Chafe. 1994. *Discourse, Consciousness and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

P. S. Dale, and L. Fenson. 1996. Lexical development norms for young children.

Behavioral Research Methods, Instruments & Computers 28:125-127.

Lila R. Gleitman. 1990. The structural sources of word meaning. *Language Acquisition* 1:3-55.

Lila R. Gleitman, and Jane Gillette. 1995. The role of syntax in verb learning. In *The Handbook of Child Language*, eds. Paul Fletcher and Brian MacWhinney, 413-427. Cambridge, MA: Blackwell.

Elena V. M. Lieven, Julian M. Pine, and Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24:187-219.

Brian MacWhinney. 2000. *The CHILDES Project: Tools for Analyzing Talk*.vol. 2: The Database. Mahwah, NJ: Lawrence Erlbaum Associates.

Julian M. Pine, and Elena V. M. Lieven. 1993. Reanalysing rote-learned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language* 20:551-571.

Willard Van Orman Quine. 1980. On what there is. In *From a Logical Point of View*, ed. Willard Van Orman Quine. Cambridge, MA: Harvard University Press.

Martin Riedmiller, and H. Braun. 1993. A direct adaptive method for faster backpropagation learning: The Rprop algorithm. Paper presented at *IEEE International Conference on Neural Networks 1993 (ICNN 93)*, San Francisco, CA.

Michael Tomasello. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

Virginia Valian. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40:21-81.

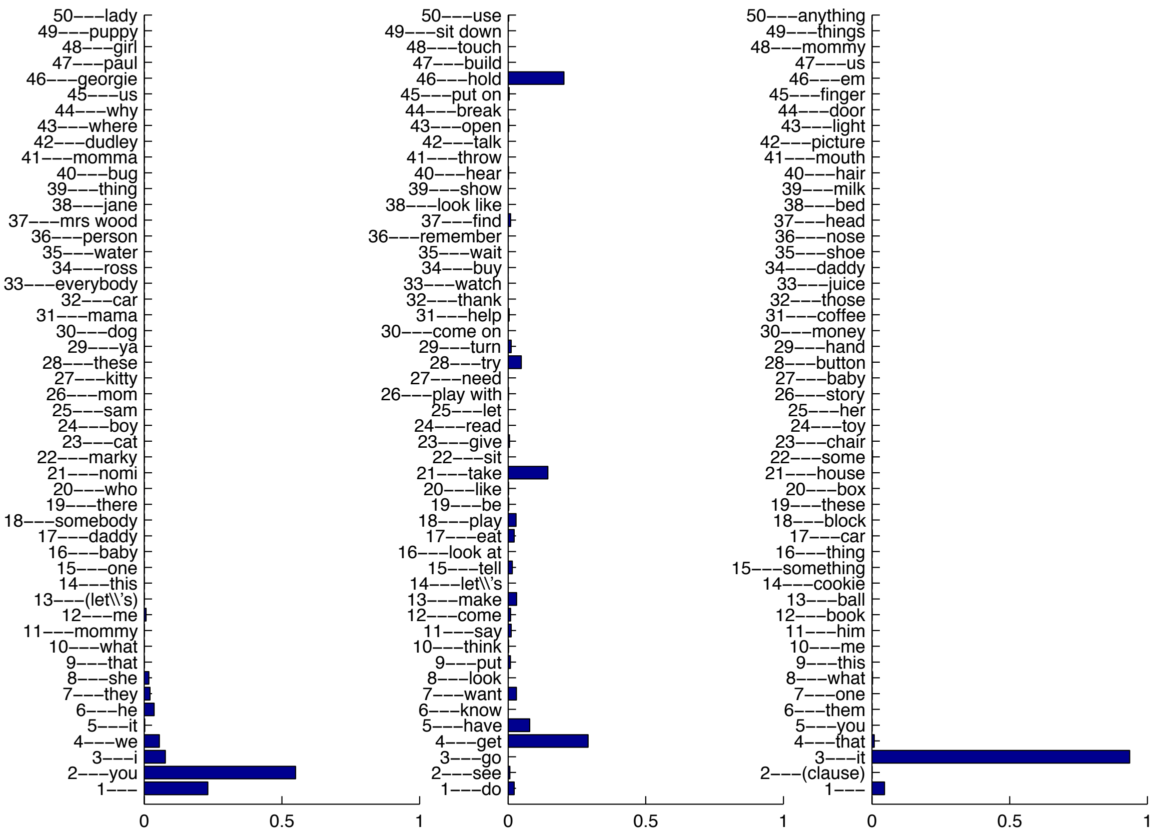

Figure 4: Average network output response to the object *it*. Subjects are shown in the left column, verbs in the middle, and objects on the right. Within each syntactic category, output units are ordered according to the frequency of the corresponding words in the input (lower numbers are higher frequency). The width of each bar reflects the average activation of the corresponding unit in our networks.

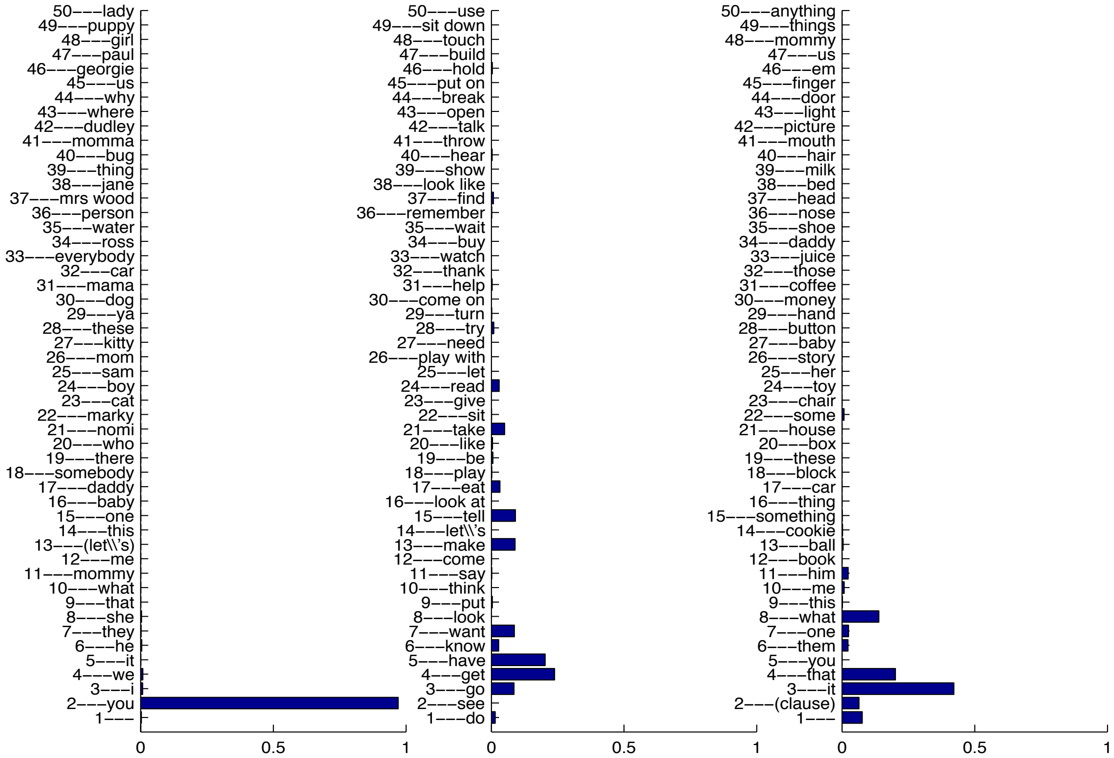

Figure 5: Average network output response to the subject *you*. Same conventions as previous figure.

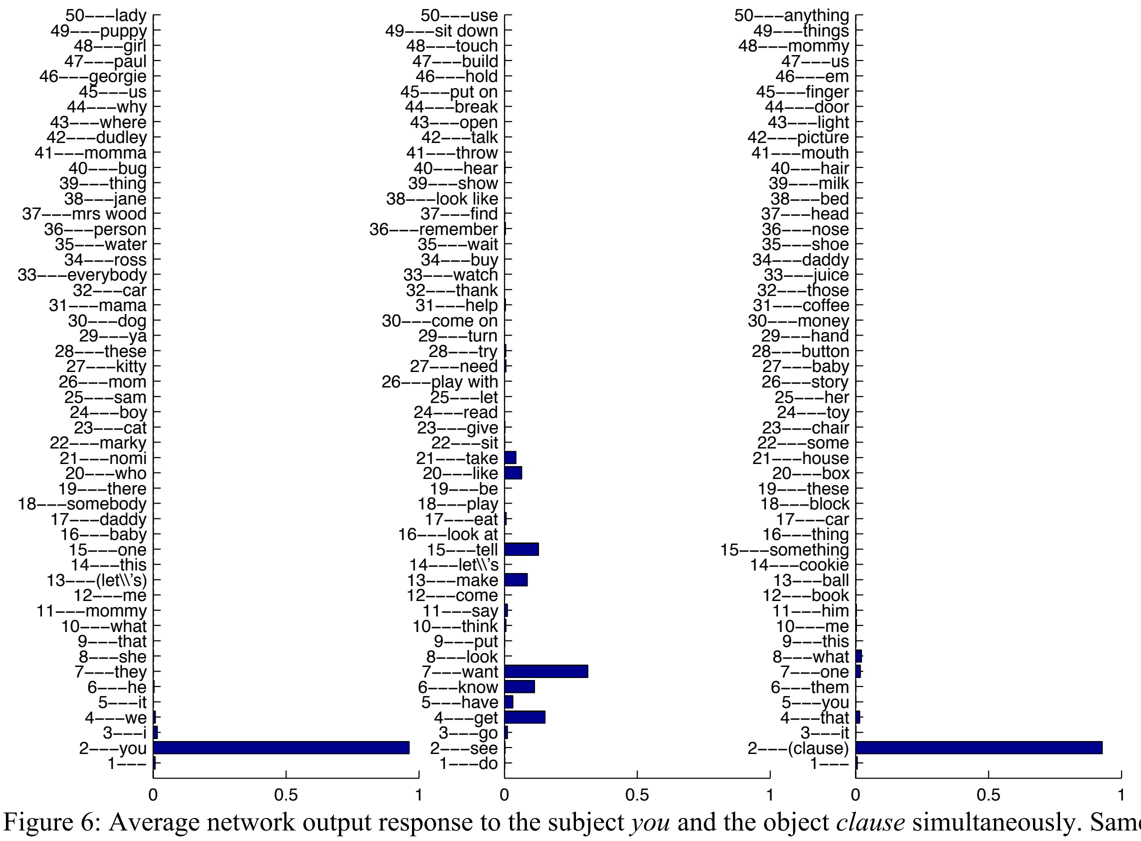

Figure 6: Average network output response to the subject *you* and the object *clause* simultaneously. Same conventions as Figures 4 and 5.