A lexically based proposal of argument structure meaning∗

Adele E. Goldberg^, Devin Casenhiser^ and Nitya Sethuraman#
University of Illinois^ and Indiana University#

Introduction

Children conservatively produce syntactic patterns on a verb-by-verb basis (Akhtar and Tomasello, 1997; Baker, 1979; Bates and MacWhinney, 1987; Bowerman, 1982; Braine, 1976; Brooks and Tomasello, 1999; Gropen et al., 1989; Ingram and Thompson, 1996; Lieven et al., 1997; MacWhinney, 1982; Olguin and Tomasello, 1993; Schlesinger, 1982; Tomasello, 1992). Ultimately, however, generalizations over specific verbs are made, forming speakers’ knowledge of argument structure patterns (Akhtar, 1999; Bowerman, 1982; Brooks and Tomasello, 1999).

We address the question of how it is possible for children to go from specific knowledge of individual verb usage to knowledge of more general linking patterns, using general inductive strategies. We do not address the question of exactly when generalizations emerge in this study; that is, no specific time line is suggested (see Fisher, 2002; Tomasello, 2000 for a range of views on this issue).

For some time, linguists have observed that within a given language, there exist certain formal patterns that correlate strongly with the meaning of the utterance in which they appear. Such correlations between form and meaning have been variously described as linking rules projected from the main verb’s specifications (e.g., Bresnan and Kanerva, 1989; Davis, 1996; Dowty, 1991; Grimshaw, 1990; Jackendoff, 1983) as lexical templates overlain on specific verbs (Hovav and Levin, 1998), or as phrasal form and meaning correspondences (constructions) that exist independently of particular verbs (Goldberg, 1995; Jackendoff, to appear).

One way to account for the association of meanings with particular forms is to claim that the association is innate (Gleitman, 1994; Grimshaw, 1990; Pinker, 1984; Pinker, 1989; Pinker, 1994). This claim generally rests on the idea that the input is not rich enough for the relevant generalizations to be learned; this is the well-known “poverty-of-the-stimulus” argument (Chomsky, 1980; Chomsky, 1988; Pinker, 1994). This view of learning a grammar can be likened to customizing a software package: everything is there, and the learner simply selects the parameters that are appropriate for his environment (Jackendoff, to appear, Chapter 7). Many have criticized this approach for its biological implausability (Bates and Goodman, 1998; Deacon, 1997; Elman et al., 1996; Sampson, 1997). Moreover, there have been virtually no successful proposals for
what any specific aspect of the parameters might look like (Ackerman and Webelhuth, 1988; Culicover, 1999; Jackendoff, to appear; Newmeyer, 1998).

An alternate way to account for the association of meanings with particular forms is to argue that the association is learned. This paper joins the growing body of literature that detracts from the poverty-of-the-stimulus argument by presenting evidence that the nature and properties of at least certain patterns in language are learnable on the basis of general categorization strategies (see also e.g., Bybee and Slobin, 1982; Bybee and Moder, 1983; Jackendoff, Chapter 7 draft; Lakoff, 1987; Taylor, 1995). In this paper we argue that language input provides more than adequate means by which learners can induce the association of meaning with certain argument structure patterns insofar as well-established categorization principles apply straightforwardly to this domain. Throughout this paper, we adopt constructional terminology, but the ideas we present are not exclusive to a constructionist account. Those who favor one of the other terminologies mentioned above need only construe this account as a proposal for how children can learn linking rules or learn the semantics associated with various lexical templates on the basis of the input. What is crucial is the uncontroversial notion that there do in fact exist correlations between formal linguistic patterns and meaning.

In Table 1, we provide a partial list of such form and meaning correspondences (lexical templates, combination of linking rules, constructions) along with the labels we use as mnemonics throughout the paper to refer to them.

<table>
<thead>
<tr>
<th>Form/Example</th>
<th>Meaning</th>
<th>Construction Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subj V Obl\text{path/loc}</td>
<td>X moves Y\text{path/loc}</td>
<td>Intr. motion (VL)</td>
</tr>
<tr>
<td>\text{e.g., The fly buzzed into the room.}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Subj V Obj Obl\text{path/loc}</td>
<td>X causes Y to move Z\text{path/loc}</td>
<td>Caused-Motion (VOL)</td>
</tr>
<tr>
<td>\text{e.g., Pat sneezed the foam off the cappuccino.}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Subj V Obj Obj2</td>
<td>X causes Y to receive Z</td>
<td>Double Object (VOO)</td>
</tr>
<tr>
<td>\text{e.g., She faxed him a letter.}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Examples of Correlations Between Form and Meaning

We examined three constructions in children's early utterances and found a strong tendency for there to be one single verb occurring with very high frequency in these constructions in comparison to other verbs used. We conducted a second study to demonstrate that the existence of a high frequency token facilitates the acquisition of constructional meaning.
2. Children’s Early Use of Verbs in the Bates Corpus

Database

In order to examine more closely children’s early uses of particular constructions, we investigated a corpus of children’s early speech. The main language corpus used in this study is the Bates corpus (Bates et al., 1988) on the Child Language Data Exchange System (CHILDES) database (MacWhinney, 1995). This corpus contains transcripts from the Bates/Bretherton Colorado longitudinal sample of 27 middle-class children, 13 boys and 14 girls at age 20 and 28 months. There are transcripts for 15 minutes, equally divided into three types of mother-infant interaction: free play, reading of the book *Miffy in the Snow*, and snack time.

Data collection and coding

The speech of all 27 children at 28 months and the speech of 15 mothers to their children aged 28 months were extracted from the Bates et al. transcripts. Utterances were hand-coded independently by the two coders and then combined into one full list. Disagreements were resolved through discussion with a third person. Agreement between the two independent coders was 96% initially and 100% after discussion (n = 840) for classifying children’s utterances by construction type.

Three construction types are focused on in this paper:

<table>
<thead>
<tr>
<th>Label</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>VL:</td>
<td>(Subj) V Obl&lt;sub&gt;path/loc&lt;/sub&gt;</td>
</tr>
<tr>
<td>VOL:</td>
<td>(Subj) V Obj Obl&lt;sub&gt;path/loc&lt;/sub&gt;</td>
</tr>
<tr>
<td>VOO:</td>
<td>(Subj) V Obj Obj&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Classifications were based primarily on form: categorizing an utterance as an instance of the VOO pattern required a verb and two NPs. Categorizing an utterance as an instance of the VL pattern required that there be a verb and some type of locative: a preposition phrase indicating location, a particle indicating location (e.g., *down*, *in*), a locative (*there*, *here*), or some combination.

Results

In analyzing the children’s speech we found a strong tendency for there to be one single verb occurring with very high frequency in comparison to other verbs used in each of the constructions analyzed. For example, for the VL construction, (Subj) V Obl<sub>loc</sub> (e.g., I went to the store) the verb *go* accounted for a full 54% (121/224) of the instances of that construction in the children’s speech. Other
verbs occur in the construction, but with much less frequency. Go and the next six most frequent verbs are provided below:

\[
\begin{array}{ll}
go & 54\% (121/224) \\
get & 6\% \\
fall, come & 5\% \\
look, live, sit & 4\% each
\end{array}
\]

A similar pattern emerged in the case of the VOL construction, (Subj) V Obj Obj\textsubscript{path/loc}, exemplified by the sentence, I put the book on the table. The verb put accounts for a full 31\% (16/51) of the instances of the construction in the children’s speech. Again, other verbs appear in the construction, but with markedly less frequency:

\[
\begin{array}{ll}
put & 31\% (16/51) \\
get & 16\% \\
take & 10\% \\
do, pick & 6\%
\end{array}
\]

The ditransitive or VOO construction, (Subj) V Obj Obj\textsubscript{2} (e.g., Pat gave Chris a book), is represented below:

Only 6 instances of the VOO construction were found in the children’s speech in the Bates et al. corpus. These included 2 instances of give, 2 of make, and 2 of bring. This paucity of data is due to the fact that the VOO construction is only just beginning to be learned at 28 months. We therefore examined data collected by by Gropen et al. (1989) from longitudinal data in the Brown (1973) corpus of Adam, Eve, and Sarah together with MacWhinney’s (1995) data on Ross and Mark. Adam was recorded in 55 two hour samples taken every 2-4 weeks between age 2;3 and 4;10. Eve was recorded in 20 two hour samples taken every 2-3 weeks between age 1;6 and 2;3. Sarah’s speech was recorded in 139 one hour samples taken at 2-19 day intervals between the ages of 2;3 and 5;1. Ross and Mark were both recorded by their father Brian MacWhinney in 62 samples of varying sizes at
varying intervals, Ross between the ages of 2;6 and 8;0, and Mark between the ages of 0;7 and 5;6.

As seen in Table 2, *give* is the most frequent verb to appear in this construction for the majority of children. The numbers for *give* do not include uses that Gropen et al. considered idiomatic. In the case of one child, Mark, there is no statistical difference between *give* and *tell*; the data set is too small for this particular child to draw any clear inferences. The overall frequency of *give* in the double object construction in this child’s speech outnumbers *tell* if idiomatic uses were included.

<table>
<thead>
<tr>
<th>Child</th>
<th>Most frequent verb in VO</th>
<th>Percentage of Tokens</th>
<th>Total # of Verb Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adam</td>
<td><em>give</em></td>
<td>53% (59/112)</td>
<td>13</td>
</tr>
<tr>
<td>Eve</td>
<td><em>give</em></td>
<td>36% (4/11)</td>
<td>5</td>
</tr>
<tr>
<td>Sarah</td>
<td><em>give</em></td>
<td>43% (29/67)</td>
<td>12</td>
</tr>
<tr>
<td>Ross</td>
<td><em>give</em></td>
<td>43% (69/160)</td>
<td>13</td>
</tr>
<tr>
<td>Mark</td>
<td><em>tell</em> <em>give</em></td>
<td>32% (11/34) 29% (10/34)</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2. The most frequent verb in ditransitive (VOO) construction based on Gropen et al. (1989)

To summarize the data so far, one verb accounts for the lion’s share of tokens of each of the constructions analyzed. Zipf long ago noted that highly frequent words account for most linguistic tokens (Zipf, 1935). Although he did not claim that there should be a single most highly frequent word for each clause pattern, nor did Zipf’s work prepare us for the fact that a single verb accounts for such a lion’s share of the tokens, his observation suggests that we may find a similar pattern in constructions other than argument structure constructions.

3. General Purpose Verbs and Constructional Meaning

As represented in Table 3, the meanings of the most frequent verbs used in particular argument structure constructions bear a striking resemblance to the meanings independently posited for those argument structure constructions (Goldberg, 1995).

- *put*  X causes Y to move Z  VOL: Caused Motion
- *go*   X moves Y         VL: Intransitive Motion
- *do*   X acts on Y       VO: Transitive
Table 3. Main verbs and the constructional meanings they correspond to

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning Description</th>
<th>VOR</th>
<th>VOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>make</td>
<td>X causes Y to become Z</td>
<td>VOR: Resultative</td>
<td>VOO: Ditransitive</td>
</tr>
<tr>
<td>give</td>
<td>X causes Y to receive Z</td>
<td>VOO: Ditransitive</td>
<td>VOO: Ditransitive</td>
</tr>
</tbody>
</table>

“General purpose” verbs, including *put, go, do, make* are among the first and most frequent verbs in many languages (Clark, 1978; Clark, 1996). Clark cites data from Bowerman (1973) for Finnish, Grégoire (1937) for French, Sanchés (1978) for Japanese, and Park (1977) for Korean; Ninio (1999), discussed below, provides similar data from Hebrew.

4. Experimental evidence

The corpus data demonstrate that a single verb accounts for the lion’s share of tokens for all of the constructions considered here. However, it does not demonstrate that the existence of a high frequency token facilitates the acquisition of constructional meaning. In order to show a causal relationship, we performed an experiment designed to test just this.

81 University of Illinois undergraduates took part in the experiment as an extra-credit course assignment. All subjects were native speakers of English and ranged from 18 to 24 years of age.

Subjects were randomly and equally divided into three conditions: the *no-training* condition, the balanced condition and the high token frequency condition. Subjects in both conditions other than the no-training condition saw a short film that consisted of 16 film clips in which puppets acted out scenes. Subjects wore earphones and listened to a tape recording that described each of the scenes as they unfolded.

The descriptions involved novel verbs in a novel construction, “SOV-o” e.g., *the king the ball moo-po-ed*. The construction was novel formally in that it had SOV word order as opposed to standard English SVO order. Additionally, the construction was morphologically marked in that all nonsense verbs had the nonsense suffix *-o* attached.

In each film clip, a scene of appearance was depicted: e.g., a dot appeared on the queen’s nose, a ball rolled onto the stage from off screen, a clown dropped into a chair from out of sight, a rabbit appeared in a hat, a dragon wriggled out from a hole in a cloth etc. Thus the construction could be said to designate a scene of appearance. Since there is no construction in English devoted to this meaning, the semantics of the construction was novel as well.

In the balanced condition, subjects heard 5 different novel verbs, each with a relatively low token frequency (4-4-4-2-2). In the high token frequency condition, subjects again heard the same 5 novel verbs, but this time one had especially high token frequency (8-2-2-2-2). Each scene was repeated exactly twice in each condition. To be consistent with the finding that the most frequent verb had a very general meaning, the high token frequency exemplar was designed to encode the meaning of “appearance” in a very general way, without
designating a particular manner. It was viewed 8 times (in 4 distinct scenes). The same exemplar (paired with a subset of the same scenes) appeared in the balanced condition 4 times (in two distinct scenes). The balanced condition included 2 additional scenes that involved a specific manner of appearance. Subjects were instructed only that they would see some films and that they were to pay close attention because there would be a quiz at the end. In the no-training condition, subjects went directly to the test phase without watching either video.

The test was designed to determine whether subjects had learned anything about the construction’s semantics by watching the scenes and listening to the novel constructional descriptions of each scene. In a forced choice comprehension task subjects saw two scenes simultaneously on the computer screen and were asked to choose which scene best matched the description they heard. Over earphones, subjects heard the same novel construction with new novel verbs that had not been used in the training conditions. They had to choose between a new scene of appearance and some other scene that involved a related action. For example, in one trial, subjects saw a flower grow from out of the ground in one clip and the same flower simply getting taller, but continuously in view in the other. Given that the description involved the novel construction and this construction was intended to designate appearance, the correct choice would be to identify the scene of appearance.

Both film clips replayed automatically until subjects pointed to the scene that matched the description. Clips designating scenes of appearance were randomly assigned to the left or right of the computer screen. Seven trials with new novel verbs in the novel construction were recorded for accuracy. These trials were randomly interspersed with additional distractor items that used new novel verbs but in a simple English transitive construction.

Results are given in Table 6.

<table>
<thead>
<tr>
<th>#correct responses (out of 7)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High token frequency condition</td>
<td>5.1</td>
</tr>
<tr>
<td>Balanced condition</td>
<td>3.8</td>
</tr>
<tr>
<td>Control (no training) condition</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Table 6: Experimental results varying token frequency in the learning of a novel construction with novel verbs

As expected, subjects in the control no-training condition did no better than chance at choosing the correct scene. The balanced condition showed a statistically significant improvement over the control condition, indicating that they had learned something about the construction’s semantics from the
training film involving 5 relatively low token frequency verbs. As predicted by our hypothesis, the high-token frequency condition showed a statistically significant improvement in accuracy over the balanced condition.\textsuperscript{i} Analysis of variance confirms a significant effect for type of training $F(2, 78)=23.29$; $p < .001$. Post Hoc comparisions using Bonferroni's correction indicate significant differences between all groups at or below the $p < .01$ level.

Thus with less than 3 minutes of training, subjects demonstrated an ability to learn constructional meaning: they were able to extend the semantics of the construction to new novel verbs and new scenes in this forced choice comprehension task. Moreover, the results demonstrate that high token frequency of a single general exemplar does indeed facilitate the acquisition of constructional meaning.

We know that frequency and order of acquisition play key roles in category formation in that training on prototypical instances, frequently and/or early, facilitates category learning (Bruner et al., 1956; Kruschke, 1996; Maddox, 1995; Nosofsky, 1988). This generalization together with the experimental evidence discussed in the previous section suggests that the very frequent and early use of one verb in a pattern facilitates the learning of the semantics of that pattern. The corpus findings demonstrate that exactly this sort of tailor-made input is available to language learners. We suggest, for example, after using many sentences with *put* in the VOL construction as in (1), children come to associate the meaning of *put* with the construction even when the verb is not present as in (2):

(1) She put a finger on that.
(2) He done boots on. (STE, 28 months, Bates et al., 1988)

The result is that the meaning of roughly “X causes Y to move Z\textsubscript{loc}” comes to be associated with the Subj V Obj ObI\textsubscript{path/loc} formal pattern.

We need to emphasize that we are not claiming that general purpose verbs are necessarily the very first verbs uttered. Neither our cross-sectional corpus data nor our experimental data address this question. Moreover, longitudinal studies have yielded differing answers (see Ninio 1999 for a claim that such verbs should be the very first verbs uttered; but see Campbell and Tomasello (Campbell and Tomasello, to appear (2002?)) for evidence that they are not always the very first verbs; also see Tomasello and Stahl (to appear) for arguments that extreme care must be taken to avoid confusing high frequency with early acquisition when intermittent sampling techniques are used.).

It also must be noted that it is not necessary for there to be a single verb with frequency far greater than other verbs for successful learning to take place. The
correlation between form and meaning can be learned by noting their association across several distinct verbs, each with relatively low frequency. This is in fact evident from our data insofar as subjects in the balanced condition outperformed those in the control condition: they clearly did learn something from witnessing several verbs in the construction, each with relatively low frequency.

Strikingly asymmetric frequencies are found in the corpus data, however, with one verb having far greater frequency than other verbs for each of the constructions discussed here. We have argued that the especially high frequency makes the learning of the correlation that much easier: the meaning associated with the construction requires minimal abstraction from the semantics of the verb as used in the construction. The highly frequent verb serves as a readily available prototype with which other verbs may be associated.

5. Conclusion
This work suggests an account of how children form the argument structure generalizations they do. We argue that the input is structured in such a way as to make the generalization from verb islands to argument structure constructions straightforward. One particular verb accounts for the lion’s share of tokens of each argument frame considered in an extensive corpus study on the Bates et al. (1988) corpus, in both mothers’ and 28 month old children’s speech. The dominance of a single verb in the construction facilitates the association of the meaning of the verb in the construction with the construction itself, allowing learners to get a “fix” on the construction’s meaning. Previous research in general categorization as well as new experimental results on construction-learning presented here support this idea. In this way, grammatical constructions may arise developmentally as generalizations over lexical items in particular patterns.

The implications of this work are potentially far reaching insofar as tokens of constructions are typically centered around one or a few specific words, or around a semantic prototype, even when they potentially occur with a much broader range of words or meanings (Diessel and Tomasello, 2001; Goldberg, 1996; Hunston and Francis, 1999).

The present proposal for how the semantics associated with constructions is learnable from the input directly undermines the “paucity of the stimulus” argument as it is aimed at this particular problem. Before we decide that language-specific properties must be innate, it is worth investigating how they might be learned, given general cognitive mechanisms such as categorization, together with a closer look at the input children receive.


CAMPBELL, AIMEE L. and TOMASELLO, MICHAEL. to appear (2002?). The Acquisition of English Dative Constructions.
CASENHISER, DEVIN and GOLDBERG, ADELE E. submitted. Fast Mapping of a Phrasal Form and Meaning.


—. Chapter 7 draft. The lexicon: Storage vs online construction.
—. to appear.


* This paper is a condensed version of (Goldberg et al., to appear). The authors are grateful for helpful discussions with Cindy Fisher, Dedre Gentner, Gregory Murphy, and Mike Tomasello. We thank Aarre Laakso and Dan Jackson for writing programs to help us analyze the corpus data, and Giulia Bencini, Bill Croft, Dan Jackson, Aarre Laakso, Yasuhiro Shirai for helpful comments. We are especially grateful to Kathryn Gelder for running the subjects in the Moopo experiment. This research was funded in part by the first author’s NSF Grant SBR-9873450.

† This experiment raises a number of questions that we have begun to follow up on. For example, a similar pattern of results is found with Kindergarteners, even when all three conditions watch
exactly the same film, the control condition involving watching the film without sound (Casenhiser and Goldberg, submitted).