Grounded Learning of Grammatical Constructions

Nancy C. Chang* **and Tiago V. Maia**†

International Computer Science Institute
1947 Center Street, Suite 600, Berkeley, CA 94704
{nchang, maia}@icsi.berkeley.edu

Abstract

We describe a model of grammar learning in which all linguistic units are grounded in rich conceptual representations, and larger grammatical constructions involve relational mappings between form and meaning that are built up from smaller (e.g., lexical) constructions. The algorithm we describe for acquiring these grammatical constructions consists of three separate but interacting processes: an *analysis* procedure that uses the current set of constructions to identify mappings between an utterance and its accompanying situation; a *hypothesis* procedure that creates new constructions to account for remaining correlations between these two domains; and *reorganization* processes that generalize existing constructions on the basis of similarity and co-occurrence. The algorithm is thus grounded not only in the twin poles of form and meaning but also, more importantly, in the relational mappings between the two.

Introduction

Language is often branded as a prime example of the human capacity for abstract symbol manipulation, with formal approaches to syntax in the forefront. Treating units of language (typically words or parts of speech) as autonomous and disembodied – that is, disregarding any conceptual basis or meaning with which they may be associated – has yielded notable success in speech recognition and parsing technology. As yet, however, this approach has made comparatively little headway toward effective natural language understanding and learning systems.

This disparity is readily explained within the cognitive linguistic tradition, which views linguistic knowledge at all levels as consisting of mappings between the domains of form and meaning (Langacker 1991), where form typically refers to the speech or text stream and meaning refers to a rich conceptual repertoire. Language use is therefore bipolar, and linguistic representations are inherently grounded by way of the meaning pole.

Meaning-free formal approaches to language retain an important advantage, however, in that they are amenable to corpus-based techniques. While not grounded in conceptual structure, the sophisticated methods computational linguists have developed for discovering statistical regularities are nevertheless grounded in real linguistic data. In the terms adopted here, they have focused on grounding the form pole.

In this paper, we describe an approach to language learning in which linguistic representations are grounded both in the conceptual world of the learning agent and in the statistical properties of the data to which it is exposed. Although models along these lines have previously been proposed for learning individual words, such as terms for spatial relations, objects and actions (Regier 1996; Roy & Pentland 1998; Bailey 1997; Siskind 1997), we concentrate here on the acquisition of larger grammatical constructions that can likewise be viewed as mappings between form and meaning.

We take as both inspiration and constraint the kinds of constructions observed in studies of child language acquisition: such studies provide valuable clues from the only natural exemplar of the learning process in question, as well as a standard by which we restrict our domain of inquiry. Although we believe that the acquisition of the more complex constructions found in adult language follows a similar process, the key aspects of the learning process are easier to unravel and illustrate at the stage of children's earliest word combinations. In this paper, we use acquisition of some simple English constructions as a case study, though our claims are meant to apply crosslinguistically.

We first describe a number of assumptions we believe apply to language learning and use in general, and then discuss some of the complexities introduced with grammatical constructions; these larger constructions usually involve multiple entities in both form (e.g., multiple words and/or phonological units) and meaning (multiple participants in a scene). We then present an algorithm for learning such constructions, using data from child-language studies as illustration. Finally, we discuss some of the broader implications of the model for language learning and use.

Prerequisites: Making Sense of the World

Our model of grammar learning makes several crucial assumptions that acknowledge the significant prior knowledge the language learner brings to the task. Above all, acquiring grammatical constructions depends on largely the same skills needed for acquiring ontological distinctions: the ability to form representations capturing regularities in the input, and to use these representations to make sense of subsequent stimuli. Learning ontological and linguistic knowledge is uniform in this respect; the former involves regularities within the conceptual domain, while the latter involves regularities across the domains of form and meaning.

Prelinguistic Representations

Much of the difficulty in formulating language acquisition in computational terms is representational: infants inhabit a dynamic world with many continuous percepts, and how they represent and process the fluid sensations that make up

* University of California, Berkeley
† State University of New York, Buffalo

their experiences remains poorly understood. Well before the first recognizable words appear, however, a substantial repertoire of concepts corresponding to people, objects, settings and actions will have emerged from the chaos as the beginnings of a stable ontology.

The acquisition of these concepts from natural or naturalistic input has been addressed by models in probabilistic, connectionist, clustering and logical frameworks.[1] For our current spotlight on the acquisition of grammatical structures, we require only that conceptual representations exhibit the kinds of category and similarity effects known to be pervasive in human cognition (Lakoff 1987). That is, concepts should cluster into categories with radial structure, preferably with graded category membership. Representations should also facilitate the identification of similar concepts and provide some basis for generalization.

An important additional requirement comes from the assumption that many early concepts involve multiple entities interacting within the context of some unified event (Tomasello 1992). Prelinguistic children are competent event participants who have accumulated structured knowledge about the roles involved in different events and the kinds of entities likely to fill them. Again, although the formation of such concepts is not our current focus, we assume that biologically grounded processes such as those described in the companion paper give rise to frame-based representations that capture the crucial relational structure of many sensorimotor actions and events. Similar processes should apply to conceptual knowledge of all kinds: while sensorimotor concepts presumably dominate the earliest stages of concept formation, aspects of the surrounding social and cultural context are also firmly in place prelinguistically; frame-based social knowledge seems to underlie children's early grasp of basic interaction (e.g., turn-taking), as well as more cultural frames associated with, for example, meals, play and clothing.

Frames may be expressed using simple role-filler notation, e.g., throw [Thrower = *Human*. Throwee = *Object*.], for an action-specific throw frame with a *Human* thrower acting on an *Object* throwee. It will also be convenient to represent frames in terms of individual role bindings: Throw.thrower:*Human* and Throw.throwee:*Object*. Note that although these representations highlight relational structure and obscure lower-level features of the underlying concepts, both aspects of prelinguistic knowledge will be crucial to our approach to language learning.

Linguistic Units as Cross-Domain Correlations

Linguistic units are based on the same kinds of correlations as those described above, where one domain happens to be the acoustic signal in the learner's environment. The other domain ranges over the entire conceptual ontology, from simple representations of people and objects to much more complex actions and interactions whose physical referents may be more transient and difficult to identify.

Not surprisingly, early words behave much like other categories, in both form and meaning poles.[2] Lexical items are initially tightly coupled with the specific events, contexts and even purposes with which they have co-occurred. They are also subject to polysemy effects, since the same form may be encountered in multiple distinct (though possibly related) contexts, which may be diverse enough to resist a single generalization. The word *up*, for example, may initially have several distinct uses: as a request to be picked up; as a comment on an upward movement; as a remark about a highly placed item; etc. (Bloom 1973).

As this example makes clear, children in the one-word stage have already acquired a number of primitive speech acts – remarkably enough, single words (often in conjunction with intonational cues) seem sufficient for expressing toddlers' requests, comments, refusals and even queries. It is important, however, to distinguish such pragmatic sophistication from the form-meaning associations we address in this paper; despite their initial context-bound nature, most word meanings eventually become generalized toward representations that are neutral with respect to speech act.

Though less is known about children's early language comprehension abilities, general pragmatic skills play a similarly dominant role early on, with simple lexical associations supplementing the complex reasoning children have developed to behave successfully in social situations. As the language learner amasses an increasing collection of relatively stable form-meaning pairs, however, the linguistic cues become better correlated with the environment and accordingly more informative.

To a large extent, we can characterize adult language in a similar vein: mature users of language also rely on relatively impoverished form cues that, in combination with extensive pragmatic reasoning, evoke vastly more complex conceptual structures. The larger theoretical framework in which the current work is situated takes a simulation-based approach to language understanding (Narayanan 1997; Bergen, Chang & Paskin, to appear). Under such an approach, linguistic constructions serve as a bridge between utterances and active, embodied representations (see also submission by Feldman to this symposium). In other words, understanding an utterance involves using constructional mappings between form and meaning to specify parameters for a simulation, and the inverse mapping holds for language production.

These claims are relatively uninformative in the case of single-word constructions, but place strong requirements on how larger constructions are represented. We now discuss some of these representational complexities, before addressing how such constructions are learned.

[1] Typically, input data corresponding to sensorimotor input is described using a set of continuous and/or discrete features, and standard machine learning techniques are used to acquire categories based on supervised or unsupervised training. See (Bailey 1997; Roy & Pentland 1998), as well as references in the companion paper (Maia & Chang 2001).

[2] Some researchers have modeled the acquisition of word form categories, using data from the acoustic domain (Brent 1999). This issue, and that of how children learn to correlate articulatory gestures with auditory signals, will not be discussed here; discovering these regularities requires the same ability to detect intra- and inter-domain regularities assumed in the previous section.

Grammatical Constructions

We base our representations of grammatical knowledge on ideas from Construction Grammar (Goldberg 1995) and Cognitive Grammar (Langacker 1991). In these approaches, larger phrasal and clausal units are, like lexical constructions, pairings of form and meaning. A key observation in the Construction Grammar tradition is that the meaning of a sentence may not be strictly predictable from the meaning of its parts; the syntactic pattern itself may also contribute a particular conceptual framing. For example, the CAUSED-MOTION construction underlying *Pat sneezed the napkin off the table* imposes a causative reading on the typically non-causative verb *sneeze*, and the need for an agentive recipient in the DITRANSITIVE construction renders *Harry kicked the door the ball* somewhat anomalous.

On this account, syntactic patterns are inextricably linked with meaning, and grammaticality judgments are rightly influenced by semantic and pragmatic factors. That is, the interpretation and acceptability of a particular utterance depends not only on well-formedness conditions but also on the structure of the language user's conceptual ontology and on the situational and linguistic context.

The main representational complexity introduced with these multi-word constructions is the possibility of structure in the form pole. That is, although single words can evoke complex frames with multiple participant roles (e.g., *bye-bye*, *baseball*), the actual mapping between the form and meaning pole is necessarily straightforward. With multiple form units available, however, additional structures arise, both within the form pole itself and, more significantly, in the relation between the form and meaning poles.

In addition to the sound or stress patterns of individual words, the form pole includes intonational contours, morphological inflections and word order. As with single words, the meaning pole encompasses the much larger set of frame-based conceptual knowledge. The constructional mapping between the two domains typically consists of a set of form relations (such as word order) corresponding to a set of meaning relations (in the form of frame bindings). Figure 1 gives an iconic representation of some of the possible constructions involved in an analysis of *I throw the ball*.

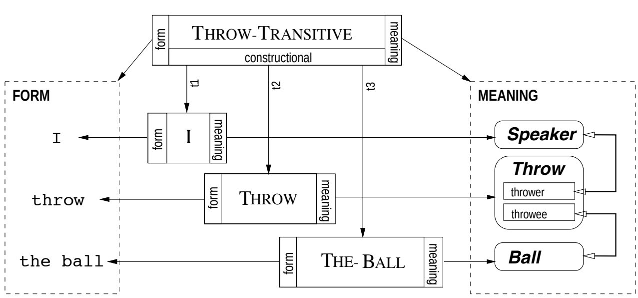

Figure 1: A constructional analysis of the sentence, *I throw the ball*, showing elements of form on the left, elements of meaning on the right and constructions linking the two domains in the center. We assume a verb-specific THROW-TRANSITIVE construction, as well as some lexical constructions that serve as its constituents; see text for details.

The lexical constructions for I, THROW and THE-BALL[3] all have simple poles of both form and meaning. But besides the individual words and concepts involved in the utterance, we have a number of word order relationships (not explicitly represented in the diagram) that can be detected in the form domain, and bindings between the roles associated with Throw and other semantic entities (as denoted by the double-headed arrows within the meaning domain). Finally, the larger clausal construction (in this case, a verb-specific one) has constituent constructions, each of which is filled by a different lexical construction.[4]

A more formal representation of the THROW-TRANSITIVE construction is given in Figure 2. Further details about the formalism employed here can be found in Bergen, Chang & Paskin (to appear). For our current purposes, it is sufficient to note that this representation captures the constituent constructions, as well as constraints on its formal, semantic and constructional elements. Each constituent has an alias used locally to refer to it, and subscripts f and m are used to denote the constituent's form and meaning poles, respectively. A designation constraint specifies a meaning type for the overall construction.

construction THROW-TRANSITIVE
constituents:
construct t1 **of meaning type** *Human*
construct t2 **of type** THROW
construct t3 **of meaning type** *Object*
formal constraints:
$t1_f$ *before* $t2_f$
$t2_f$ *before* $t3_f$
semantic constraints:
$t2_m$.thrower $\longleftrightarrow$ $t1_m$
$t2_m$.throwee $\longleftrightarrow$ $t3_m$
designates $t2_m$

Figure 2: Formal representation of the THROW-TRANSITIVE construction, with separate blocks listing constituent constructions, formal constraints (e.g., word order) and semantic constraints (role bindings).

Although this brief discussion necessarily fails to do justice to Construction Grammar and related work, we hope that it nevertheless manages to convey the essential representational demands on the structures to be learned.

Learning Constructions

Given the nature of constructions – mappings linking relations in form and meaning – it is clear that the simple classification techniques used in single-word learning will not

[3] The definite determiner *the* explicitly depends on a representation of the situational and discourse context that supports reference resolution. For simplicity, we will ignore the internal structure of "the ball" and treat it as an unstructured unit.

[4] This example, like the rest of those in the paper, is based on utterances from the CHILDES corpus (MacWhinney 1991) of child-language interaction. Note that the earliest constructions children learn seem to be verb-specific, but at later stages of development this utterance might be seen as an instance of a more general transitive construction (Tomasello 1992).

suffice. In this section we present an algorithm for acquiring grammatical constructions on the basis of examples, focusing on the earliest multi-word constructions. In accord with our discussion of conceptual prerequisites, a training example is taken to consist of an utterance paired with a representation of a situation, where the former is a sequence of familiar and novel forms, and the latter a set of conceptual entities and role bindings representing the corresponding scene.

Our approach rests on three separate but interacting procedures, shown in Figures 3, 4 and 5. In the broadest terms, we assume that the learning agent expects correlations between what is heard and what is perceived. Some of these correlations have already been encoded and thus accounted for by previously learned constructions; the tendency to try to account for the remaining ones leads to the formation of new constructions. In other words, what is learned depends directly on what remains to be explained.

The identification of the mappings between an utterance and a situation that are predicted by known constructions can be seen as a precursor to language comprehension, in which the same mappings actively evoke meanings not present in the situation. Both require the learner to have an **analysis** procedure that determines which constructions are potentially relevant, given the utterance, and, by checking their constraints in context, finds the best-fitting subset of those.

Once the predictable mappings have been explained away, the learner must have a **hypothesis** procedure for determining which new mappings may best account for new data. The mappings we target here are, as described in the previous section, relational. It is crucial to note that a relational mapping must hold across arguments that are themselves *constructionally correlated*. That is, mappings between arguments must be in place before higher-order mappings can be acquired. Thus the primary candidates for relational mappings will be relations over elements whose form-meaning mapping has already been established. This requirement may also be viewed as narrowing the search space to those relations that are deemed *relevant* to the current situation, as indicated by their connection to already recognized forms and their mapped meanings.

But structure hypothesis is not the only way constructions can arise. The same kinds of generalization that we assume for conceptual and single-word learning can also apply to constructions. Generalizations driven in a bottom-up fashion by similar or co-occurring constructions lead to the **reorganization** of the set of known constructions (or *constructicon*). We extend previous work using Bayesian model merging as the basis for both types of generalization (Stolcke 1994) to handle relational structures.

Details of each procedure are best illustrated by example. Consider the utterance U_1 = "you throw a ball" spoken to a child throwing a ball. The situation S consists of entities S_e and relations S_r; the latter includes role bindings between pairs of entities, as well as attributes of individual entities. In this case, S_e includes the child, the thrown ball and the throwing action, as well as potentially many other entities, such as other objects in the immediate context or the parent making the statement: S_e = {Self,Ball,Block,Throw,Mother,…}. Relational bindings include those encoded by the Throw frame, as well as other properties and relations: S_r = {Throw.thrower:Self, Throw.throwee:Ball, Ball.Color:Yellow, …}.

In the following sections we describe what the learner might do upon encountering this example, given an existing set of constructions $\mathcal{C}$ that has lexical entries for BALL,THROW,BLOCK,YOU,SHE, etc., as well as a two-word THROW-BALL construction associating the `before(throw,ball)` word-order constraint with the binding of Ball to the throwee role of the Throw frame.

Language analysis

Given this information, the analysis algorithm in Figure 3 first extracts the set F_{known} = {`you`,`throw`,`ball`}, which serves to cue constructions that have any of these units in the form pole. In this case, C_{cued} = {YOU,THROW,BALL,THROW-BALL}. Next, the constraints specified by these constructions must be matched against the input utterance and situation. The form constraints for all the lexical constructions are trivially satisfied, and in this case each also happens to map to a meaning element present in S.[5] Checking the form and meaning constraints of the THROW-BALL construction is only slightly less trivial: all relations of interest are directly available in the input utterance and situation.[6]

Analyze utterance. Given an utterance U in situation S and current set of constructions $\mathcal{C}$, produce the best-fitting analysis A:

1. Extract the set F_{known} of familiar form units from U, and use them to cue the set C_{cued} of constructions.
2. Find the best-fitting subset C_A of C_{cued} for utterance U in situation S. Let F_A be the set of form units and relations in U used in C_A, and M_A be the set of meaning elements and bindings in S accounted for by C_A. Then $A = < C_A, F_A, M_A >$. A has an associated cost $Cost_A$ providing a quantitative measure of how well A accounts for U in S.
3. Reward constructions in C_A; penalize cued but unused constructions, i.e., those in $C_{cued} \setminus C_A$.

Figure 3: Construction analysis.

In the eventual best-fitting analysis A, the constructions used are C_A ={YOU,THROW,BALL,THROW-BALL}, which cover the forms and form relations in F_A = {`you`,`throw`,`ball`,`before(throw,ball)`} and map the meanings and meaning relations in M_A = {Self,Throw,Ball,Throw.throwee:Ball}. (Remaining unused in this analysis is the form a.)

[5] We assume the YOU construction is a context-dependent construction that in this situation maps to the child (Self).

[6] Many complications arise in adult language – category constraints on roles may apply only weakly, or may be overridden by the use of metaphor or context. At the stage of interest here, however, we assume that all constraints are simple and few enough that exhaustive search should suffice, so we omit the details about how cueing constructions, checking constraints and finding the best-fitting analysis proceed.

Construction hypothesis based on explaining away

We proceed with our example by applying the procedure shown in Figure 4 to hypothesize a new construction. All form relations and meaning bindings, respectively, that are *relevant* to the form and meaning entities involved in the analysis are extracted as, respectively, F_{rel} = {`before(you,throw)`, `before(throw,ball)`, `before(you,ball)`} and M_{rel} = {Throw.thrower:Self, Throw.throwee:Ball}; the *remainder* of these not used in the analysis are F_{rem} = {`before(you,throw)`, `before(you,ball)`} and M_{rem} = {Throw.thrower:Self}. The potential construction C_{pot} derived by replacing terms with constructional references is made up of form pole {`before`(YOU_f,THROW_f),`before`(YOU_f,BALL_f)} and meaning pole {THROW_m.thrower:YOU_m}. The final construction C_{U_1} is obtained by retaining only those relations in C_{pot} that hold over correlated arguments:

({`before`(YOU_f,THROW_f)}, {THROW_m.thrower:YOU_m})

> **Hypothesize construction**. Given an analysis A of utterance U in situation S, hypothesize a new construction C_U including correlated but unused form and meaning relations:
>
> 1. Find the set F_{rel} of form relations in U that hold between the familiar forms F_{known}, and the set M_{rel} of meaning relations in S that hold between the mapped meaning elements in M_A.
> 2. Find the set $F_{rem} = F_{rel} \setminus F_A$ of relevant form relations that remain unused in A, and the set $M_{rem} = M_{rel} \setminus M_A$ of relevant meaning relations that remain unmapped in A. Create a potential construction C_{pot} = (F_{rem},M_{rem}), replacing terms with references to constructions in C_A where possible.
> 3. Create a new construction C_U consisting of pairs of form-meaning relations from C_{pot} whose arguments are constructionally related.
> 4. Reanalyze utterance using $\mathcal{C} \cup \{C_U\}$, producing a new analysis A' with cost $Cost_{A'}$. Incorporate C_U into $\mathcal{C}$ if $Cost_A - Cost_{A'} \geq MinImprovement$; else put C_U in pool of potential constructions.
> 5. If U contains any unknown form units, add the utterance-situation pair (U, S) to the pool of unexplained data.

Figure 4: Construction hypothesis.

At this point, the utility of C_{U_1} can be evaluated by reanalyzing the utterance to ensure a minimum reduction of the cost of the analysis. As noted in Step 4 of Figure 4, a construction not meeting this criterion is held back from immediate incorporation into $\mathcal{C}$. It is possible, however, that further examples will render it useful, so it is maintained as a candidate construction. Similarly, Step 5 is concerned with maintaining a pool of examples that involve unexplained units of form, such as the unfamiliar article a in this example. Further examples involving similar units may together lead to the correct generalization, through the reorganization process to which we now turn.

Constructicon reorganization

So far we have described an analysis-hypothesis approach to learning constructions on the basis of a single example and a set of existing constructions. A separate process that takes place in parallel is the data-driven, bottom-up reorganization of the constructicon on the basis of similarities among and co-occurrences of multiple constructions. Figure 5 gives a high-level description of this process; we refrain from delving into too much detail here, since these processes are closely related to those described for other generalization problems (Stolcke 1994; Bailey 1997).

> **Reorganize constructicon**. Incorporate a new construction C_n into an existing set of constructions $\mathcal{C}$, reorganizing $\mathcal{C}$ to consolidate similar and co-occurring constructions if necessary:
>
> 1. Find potential construction pairs to consolidate.
>   - **Merge** constructions involving correlated relational mappings over one or more pairs of similar constituents, basing similarity judgments and type generalizations on the conceptual ontology.
>   - **Compose** frequently co-occurring constructions with compatible constraints.
> 2. Evaluate constructions; choose the subset maximizing the posterior probability of $\mathcal{C}$ on seen data.

Figure 5: Constructicon reorganization.

Continuing our example, let us assume that the utterance U_2 = "she's throwing a frisbee" is later encountered in conjunction with an appropriate scene, with similar results: in this case, both the unfamiliar inflections and the article are ignored; the meanings are mapped; and constraints with appropriate correlations are found, resulting in the hypothesis of the construction C_{U_2}:

({`before`(SHE_f,THROW_f)}, {THROW_m.thrower:SHE_m})

C_{U_1} and C_{U_2} bear some obvious similarities: both constructions involve the same form relations and meaning bindings, which hold of the same constituent construction THROW. Moreover, the other constituent is filled in the two cases by SHE and YOU. As emphasized in our discussion of conceptual representations, a key requirement is that the meaning poles of these two constructions reflect their high degree of similarity.[7] The overall similarity between the two constructions can lead to a merge of the constructional constituents, resulting in the merged construction:

({`before`($\mathbf{h}_f$,THROW_f)},{THROW_m.thrower:$\mathbf{h}_m$})

where $\mathbf{h}$ is a variable over a construction constrained to have a *Human* meaning pole (where *Human* is a generalization over the two merged constituents). A similar process, given appropriate data, could produce the generalized mapping:

({`before`(THROW_f,$\mathbf{o}_f$)},{THROW_m.throwee:$\mathbf{o}_m$})

[7] The precise manner by which this is indicated is not at issue. For instance, a type hierarchy could measure the distance between the two concepts, while a feature-based representation might look for common featural descriptions.

where $\mathbf{o}$ is constrained to have an Object meaning pole.[8]

Besides merging based on similarity, constructions may also be composed based on co-occurrence. For example, the generalized *Human*-THROW and THROW-*Object* constructions just described are likely to occur in many analyses in which they share the THROW constituent. Since they have compatible constraints in both form and meaning (in the latter case even based on the same conceptual Throw frame), repeated co-occurrence eventually leads to the formation of a larger construction that includes all three constituents:

$$(\{\texttt{before}(\mathbf{h}_f,\text{THROW}_f),\texttt{before}(\text{THROW}_f,\mathbf{o}_f)\},\\ \{\text{THROW}_m.\text{thrower:}\mathbf{h}_m,\text{THROW}_m.\text{throwee:}\mathbf{o}_m\})$$

Both generalization operations we describe are, like the hypothesis procedure, merely means of finding potential constructions. Due to space considerations, we do not discuss the many complexities that arise in evaluating these constructions using Bayesian criteria. Briefly, a prior based on minimum description length favors merged and composed constructions that compactly encode previously seen data; this measure combats the inevitable drop in likelihood associated with these more general constructions. The learning algorithm chooses the set of constructions that maximizes the posterior probability of the constructicon given the data.

Discussion

The model we have proposed for the acquisition of grammatical constructions makes some strong claims about the relationship between comprehension, production and learning. We take these three processes to be tightly linked: new constructions are hypothesized specifically to make up for correlations not covered by currently known constructions, and productions are based largely on the most entrenched subset of previously acquired constructions.

The model is compatible to the extent possible with evidence from child language acquisition. The principles guiding construction hypothesis, in particular those for mapping relevant form and meaning relations, have counterparts in some of Slobin's (1985) Operating Principles for mapping. Construction reorganization allows more general constructions to result from the merging of lexically specfic constructions like those described by (Tomasello 1992).

More broadly, since the algorithm produces constructions based on any utterance-situation pair and existing constructicon represented as described above, it can apply equally well for more advanced stages of language development, when the learner has more sophisticated meaning representations and more complex constructions. The potential continuity between early language acquisition and lifelong constructional reorganization offers hope for the modeling of adaptive language understanding systems, human and otherwise.

Acknowledgments

We are grateful for the input and influence of Jerry Feldman and the rest of the NTL group at ICSI; opinions and errors remain ours alone. This work was supported in part by an IBM Research Fellowship granted to the first author, and a Praxis XXI Fellowship from the Portuguese "Fundação para a Ciência e a Tecnologia" to the second author.

[8]Although not further discussed here, examples with unexplained forms (such as the a in U_1 and U_2) may also undergo merging, leading to the emergence of common meanings.

References

Bailey, D. R. 1997. *When Push Comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs*. Ph.D. Dissertation, University of California at Berkeley.

Bergen, B. K.; Chang, N. C.; and Paskin, M. A. To appear. Simulation-based language understanding in Embodied Construction Grammar. In Östman, J.-O., ed., *Construction Grammar(s): Cognitive and Cross-language dimensions*. John Benjamins.

Bloom, L. 1973. *One word at a time: the use of single word utterances before syntax*. The Hague: Mouton & Co.

Brent, M. R. 1999. An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning* 34(1-3).

Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things*. University of Chicago Press.

Langacker, R. W. 1991. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Cognitive Linguistics Research. Berlin and New York: Mouton de Gruyter.

MacWhinney, B. 1991. *The CHILDES project: Tools for analyzing talk*. Hillsdale, NJ: Erlbaum.

Maia, T. V., and Chang, N. C. 2001. Grounding the acquisition of grammar in sensorimotor representations. *Proc. 2001 AAAI Spring Symposium on Learning Grounded Representations*.

Narayanan, S. 1997. *Knowledge-based Action Representations for Metaphor and Aspect (KARMA)*. Ph.D. Dissertation, University of California at Berkeley.

Regier, T. 1996. *The Human Semantic Potential*. MIT Press.

Roy, D., and Pentland, A. 1998. Learning audio-visually grounded words from natural input. In *Proc. AAAI workshop on Grounding Word Meaning*.

Siskind, J. M. 1997. *A computational study of cross-situational techniques for learning word-to-meaning mappings*. Chapter 2. MIT Press.

Slobin, D. I. 1985. Crosslinguistic evidence for the language-making capacity. Volume 2 of *The Crosslinguistic Study of Language Acquisition*. Chapter 15. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Stolcke, A. 1994. *Bayesian Learning of Probabilistic Language Models*. Ph.D. Dissertation, University of California, Berkeley.

Tomasello, M. 1992. *First verbs: A case study of early grammatical development*. Cambridge, UK: Cambridge University Press.