A computational model of comprehension-based construction acquisition

Nancy Chang
UC Berkeley and International Computer Science Institute

1. Introduction

Many models of the acquisition of early multi-word constructions assume significant interaction among multiple cognitive structures and processes, including the learner's conceptual world, existing linguistic inventory, and processes of language use (e.g., Slobin 1985, Tomasello 1992, 2003, Clark 1993, 2003). Thus far, however, the available theoretical frameworks for describing the relevant structures and processes—including how children comprehend and produce utterances, how conceptual knowledge and constructions may be represented in biologically and developmentally grounded ways, and how situational and discourse context interact with processes of language use—have not been precisely enough characterized to illuminate their possible interactions. This paper describes a computational model of early constructional acquisition that is consistent with linguistic and developmental evidence (Chang, in prep.) and focuses on how language comprehension may drive acquisition. The model offers technically precise formulations of constructions, concepts and background context, as well as a simplified model of language comprehension (Bergen & Chang, in press, Bryant 2003); these in turn permit concrete definitions of notions like frequency, similarity, and simplicity, which can then be used to investigate the precise nature of the link between usage and learning. We survey the main structures and processes of the model, including a construction-based grammar formalism that serves as the target of learning, a representation of the input data as pairings of utterances and communicative contexts, and a set of usage-driven heuristics for proposing and evaluating new constructions. Together these supply a framework within which the acquisition of early constructions can be both qualitatively described and quantitatively assessed, and in which questions about the complex interactions involved can be more sharply formulated and thus more satisfyingly answered.

2. Meaningful language acquisition

The transition from single words to complex grammar ranks among the most contentious matters in the cognitive sciences. Different framings of the problem have made clashing assumptions about issues including what kinds of innate or pre-linguistic knowledge children may be endowed with, how or whether domain-general knowledge and processes are involved, and even the core issue of what kind of information constitutes linguistic knowledge. In particular, Chomsky's (1957) argument from the poverty of the stimulus established a framework in which knowledge of language is equated with knowledge of abstract symbolic patterns, and the input data for language learning is conceptualized as strings of symbols, devoid of conceptual or semantic content. These assumptions, in combination with the observed lack of explicit negative evidence, led him and others to conclude that learning is constrained by a relatively limited set of innately specified parameters that determine the space of grammatical possibilities.

This work explores an alternate framework that makes more cognitively plausible assumptions about the innate structures guiding language learning. Such a viewpoint is consistent with the growing body of theoretical proposals and empirical findings that argue for a radically different conception of both the nature of the human capacity of language and its developmental course. Among the key ideas that depart from the Chomskyan paradigm are:

- **Construction grammar.** Linguistic knowledge is characterized by the pairing of aspects of *meaning*, including both semantic and pragmatic factors, with aspects of *form*. Such pairings, or *constructions*, are recognized within construction-based approaches to grammar as the basic unit of language at all levels of abstraction (Goldberg 1995, Kay & Fillmore 1999, Langacker 1987, Croft 2001).

- **Meaning as conceptualization.** Linguistic meaning is the result of human *conceptualization*, rather then an objective description of the world, and as such is grounded in cognitive and neural structures (Lakoff 1987; Talmy 2000; Langacker 1987). The basic units of meaning associated with early language are likewise both grounded in basic scenes of human experience and subject to crosslinguistic grammaticalization patterns (Slobin 1985; Choi & Bowerman 1991).
- **Context-dependence.** Language has a *communicative function*; utterances are rooted in specific speaker-hearer discourse contexts and must be interpreted with respect to the speaker's underlying intentions. Pragmatic differences in context are not somehow segregated from linguistic competence; rather, differences in form-function pairings over time serve as direct evidence of the evolution of the child's grammatical system (Bates 1976, Clark 1993, 2003, Budwig 1995).
- **Usage-based acquisition.** Statistical properties of language *use* drive acquisition and generalization. The earliest constructions appear to be motivated by lexically specific exemplars (Tomasello 1992) rather than based on rigid parameters, suggesting that children are sensitive to distributional correlations between form and meaning (Lieven *et al*. 1997, Maratsos & Chalkley 1980). The course of acquisition thus reflects not just structural properties of linguistic constructions but also their usefulness as part of the processes of comprehension and production (Clark 2003).

Each of these strands is independently well motivated by the evidence, and together they constitute a coherent and multifaceted challenge to formalist approaches to language acquisition, sharing the hypothesis that *language acquisition is driven by meaningful language use in context.* This overarching framework holds great promise of accounting for crosslinguistic developmental phenomena while accommodating convergent evidence from multiple disciplines. Delivering on that promise, however, requires a clear description of the structures and processes involved and the complex interactions among them. Specifically, a full account will include:

1. A **language formalism** for representing the target of learning, i.e., constructions including both form and meaning, where the latter is taken to be conceptualized meaning (in the sense above) and to encompass context and pragmatic function.
2. Models of the **processes** associated with language use, including comprehension and production, as well as general cognitive processes that may interact with these, such as categorization, generalization, and reinforcement.
3. A **learning framework** explaining how the linguistic system changes in response to experience — i.e., how the structures of (1) adapt over time based on the processes of (2).

Such an account would facilitate moving beyond identifying the various factors in play during acquisition to experimenting with how these factors may interact under different conditions to produce different patterns of learning, both across and within languages. To date, however, there have been few attempts to formalize these ideas in a computational setting. A few computational models of lexical acquisition are consistent with the meaning-oriented principles outlined above (e.g., Regier 1996, Bailey 1997). These approaches have shed light on how lexical mappings can be acquired from pairings of word forms with embodied meanings, and how the formation of semantic and conceptual categories might be driven by linguistic input and thus sensitive to crosslinguistic variation. Although these are among the major challenges faced by the language-learning child, they do not address the distinct challenges that arise in learning constructions with more internal structure, either word-internally (in the form of morphological structure) or in phrasal and clausal constructions. The model to be described focuses on this latter problem, and further restricts attention to how comprehension (and not production) may drive the acquisition of these structured multi-unit constructions.

Figure 1 provides a high-level view of the proposed model of construction learning. The structures (shown in ovals) correspond to *conceptual knowledge* (including events, actions, objects, people, etc., both individual and categorical), *linguistic knowledge* (form-meaning constructions, both simple and structured); and the current *utterance and situation* (the specific instances of linguistic form and conceptual available in context). These structures support the various processes of language acquisition and use (shown as wide arrows) in the model. Comprehension is modeled as an *analysis* process that produces an interpretation of an utterance based on the current linguistic and conceptual knowledge. Typically, the child's linguistic knowledge is incomplete, so this interpretation is *partial* (often compared to the more complete understanding of the situation available based on pragmatic sophistication). The other processes depicted in the figure capture the two main kinds of construction learning to be modeled: the *hypothesis* of new constructions based on a partial interpretation, and the *reorganization* of existing constructions to reflect structural and usage-based regularities.

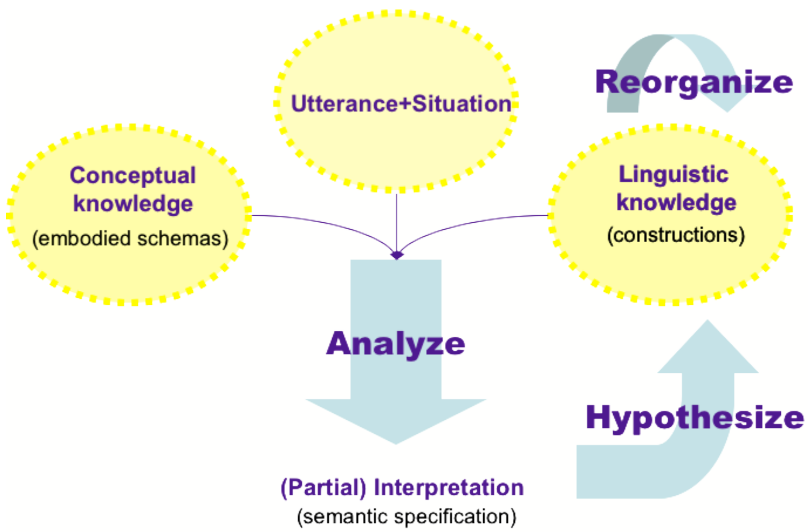

Figure 1. Learning by partial understanding: A schematic diagram of how the language understanding process draws on both conceptual and linguistic knowledge to produce an interpretation of a specific utterance-situation pair. Partial interpretations prompt learning processes that update linguistic knowledge.

The model is intended to capture the intuition that the child is always situated within a rich situational context, and can exploit relatively well-developed domain-general sensori-motor, cognitive and social skills to make sense of her environment, even when she has only a tenuous grasp on language. But over time, specific sound patterns recur reliably with specific entities and events, and these correlations become useful for guiding and constraining the child's interpretations of utterances containing these patterns even in the absence of direct contextual support. We assume that many such associative mappings have been learned by the time the child reaches our stage of interest, when the first multi-word utterances are learned (around 18-24 months), and we thus focus on the formation of more complex constructions that combine these simpler associations. Throughout all these stages, the child encounters new data that drives the learning and reorganization of constructional mappings; these mappings in turn facilitate better constructional analysis of further data. In sum, the structures learned are motivated not by abstract generalizations over symbolic forms alone but instead tied directly to their attested utility in accounting for the child's experiences.[1] The next section elaborates on the representations and processes used in the model.

3. Overview of the model

This section summarizes the components of the model, including the formalism used for representing conceptual and linguistic knowledge (Section 3.1), the input data representing the utterance and situation (Section 3.2), and the learning processes used to update the set of constructions based on a statistical learning framework (Section 3.3).

3.1 A simple grammar formalism

Conceptual and linguistic knowledge is represented using a simplified version of *Embodied Construction Grammar* (ECG), a computationally explicit unification-based formalism motivated by cognitive and constructional approaches to language (Bergen & Chang, in press). The ECG formalism is part of a broader effort to explore the hypothesis that language understanding exploits many of the same neural structures used for action, perception, imagination, memory, and other cognitive processes. Under this model, linguistic structures provide parameters for simulations drawing on these embodied structures. Recent neurobiological and behavioral findings also support the notion that perceptual and motor systems are activated during language understanding (see Bergen *et al.* 2004 and references cited there). The ECG notation allows constructions to specify how surface cues in an utterance can be mapped to such embodied representations through an *analysis* process. The resulting structure (or *semantic specification)* provides the parameters for an active *simulation* of these embodied representations with respect to the current context, producing new inferences. For current purposes, we focus on the representation of complex constructions with internal structure; see Bergen and Chang (in press) for more discussion of simulation-based language understanding and a detailed introduction to the formalism, and Bryant (2003) for a technical description of the language analysis process.

Figure 2 shows some simple example structures defined using the ECG notation, including the lexical PUSH construction, its associated Push-Action conceptual representation (or **schema**), and two more complex constructions. The notation captures the idea that the word *push* has a **form** component, simplified here as the orthographic form "push" (though it could be replaced or enriched with other form information, such as phonological or phonetic strings, intonational information and perhaps gestural representations), and a **meaning** component related to the pushing concept as captured by the Push-Action schema. This action is associated with participant **roles** (a pusher and pushee) that are part of the larger event or scene associated with the action. The type constraints on these roles (marked with a colon) reflect children's knowledge of what kinds of entities typically take part in different events (Nelson 1996; Tomasello 1992). The Push-Action's action role is filled by a dynamic motor-perceptual representation called an *x-schema*, or *executing schema* (Bailey 1997; Narayanan 1997). Though not shown here, the relevant Push x-schema incorporates notions including force application, directionality, and hand posture. It is this association that grounds the Push-Action schema and the PUSH lexical construction in terms of its underlying motor-perceptual representations.

[1] The model makes many simplifications, corresponding to various arrows that are missing or simplified in the diagram. We omit discussion of how conceptual knowledge arises from experience; how it is guided and constrained in part by language acquisition; and how the analysis process may update the context (particularly the goals and intentions of one's interlocutors). The lexical acquisition models mentioned earlier address some of these issues; these are theoretically compatible with the current model and will be integrated in future work.

schema Push-Action
  roles
    pusher : Human
    pushee : Physical-Object
    action : Push-xschema

construction PUSH
  form : "push"
  meaning : Push-Action

construction PUSH-BALL
  constituents
    p : PUSH
    b : BALL
  form
    p_f before b_f
  meaning
    p_m.pushee ↔ b_m

construction X-PUSH-Y
  constituents
    x : REF-EXPR
    p : PUSH
    y : REF-EXPR
  form
    x_f before p_f
    p_f before y_f
  meaning
    p_m.pusher ↔ x_m
    p_m.pushee ↔ y_m

Figure 2. Examples of the ECG formalism defining a conceptual schema for a pushing action, a lexical construction for *push*, and two complex constructions.

The remaining PUSH-BALL and X-PUSH-Y constructions represent the primary target of learning for the model. In both cases, the notation indicates that each construction has some internal **constituents**, each of which is itself a form-meaning construction. The constituents have local variable names and are constrained to instantiate a specific construction type (again using the colon notation). The p constituent, for example, is an instance of the PUSH construction described above; the others, though not shown, instantiate the REF-EXPR (or referring expression, a general construction similar to the more traditional NP) and BALL constructions. The key representational complexity is that the form and meaning components (also called *poles)* for complex constructions typically involve *relations* among the form and meaning poles of their constituents. These poles are referred to using a subscripted f (for form) or m (for meaning) on the relevant constituent name. The only form relation shown here is word order (the before relation), though other form relations are in principle allowed. Likewise, the only meaning relation shown here is *identification*, or unification, between two meaning entities (denoted with a double-headed arrow). In particular, both complex constructions identify roles of one constituent's meaning pole with the meaning pole of another constituent. Each of the example constructions thus pairs word order constraints over its constituents' form poles with identification constraints over its constituents' meaning poles (thus specifying the role fillers of the Push-Action schema). The two constructions differ in the number of fillers specified and in level of specificity. Both are instances of the lexically specific "verb island" constructions described by Tomasello (1992, 2003). But while the PUSH-BALL construction also specifies that its *pushee* constituent must be *ball*, the X-PUSH-Y construction encompasses a greater range of referring expressions (e.g., *my sock, the doll, block*) for its two non-verb constituents.[2]

In short, the construction notation provides a detailed way of expressing the notion that utterances of the form *push ball* and *X push Y* involve pushing events, where the word order tells us who is pushing and what is being pushed. These examples represent a particularly simple subset of English; among the phenomena they simplify or omit are complex noun phrases with determiners, quantifiers, and modification; complex predications involving auxiliary or modal verbs; and issues of agreement and tense. Nevertheless, constructions like these capture many of the earliest multi-word constructions in English that are the focus of the current model and represent an appropriate first step toward more inclusive future models.

[2] These referring expressions are still, however, constrained to refer to physical objects, as required by the Push-Action schema's type constraint on the pushee.

This formal representation has a number of advantages for computational modeling (as well as linguistic description). On a practical level, they are suitable for use in the language analysis process mentioned earlier, which can take sentences like "You can push the ball" and automatically produce interpretations, even for sentences including unfamiliar words not yet in the grammar. (For example, the analyzer can partially interpret the sentence even without constructions for the determiner *the* and modal *can*; see Section 3.2 for the interpretation of this example). Also, the formalism provides a straightforward interface for capturing the relationship between linguistic items and detailed world knowledge (for example, the degree of force or direction associated with a pushing action), so that even sentences not directly specifying motor-perceptual action parameters can be understood as nevertheless setting those parameters (perhaps to default or context-dependent values). This allows language to tap into embodied representations and encyclopedic world knowledge while relieving it from the burden of accounting for every nuance or inference relevant to action in the world.

For the learning model, the key advantage of having an explicit construction description is that it affords clear criteria for choosing one construction (or set of constructions) over another during learning. In particular, the processes described in Section 3.3 rely on notions that can be objectively measured in terms of the formal description of each construction, including *similarity*, based on the degree of shared structure between two constructions; and *simplicity*, based on the length of the construction's description. Constructions can also be assessed for non-structural properties, such as the *frequency* with which they are useful for analyzing data encountered.

3.2 Input data

The model assumes that the basis for learning new constructions is a sequence of utterances paired with the communicative contexts in which they appear; we will call each of these pairs an *input token*. The utterance is described as a string of known and novel word-forms (for simplicity, these are represented using only orthographic forms and assumed to be pre-segmented at a prior stage), along with a feature representing its intonational contour (neutral, falling, rising); the communicative context is described in terms of the salient scene and discourse features available in the situation. Figure 3 shows an example input token:

Utterance:		**Scene**:	
form:	"You can push the ball"	Push-Action	
intonation:	neutral	pusher :	Naomi
		pushee :	Ball-1
Discourse:		action :	Push-XSchema
speaker:	Mother		
addressee:	Naomi	**Participants**: Naomi, Mother, Ball-1	
speech-act:	imperative		
joint-attention:	Ball-1		

Figure 3. A sample input token representing a single utterance and its communicative context, including features of the discourse context as well as the salient participants and events.

This information represents only a subset of the possible information available to the child, corresponding to the most relevant information in the given communicative context. We assume that children can infer the appropriate communicative intent and make sense of their environments, even in the absence of much linguistic knowledge, by drawing on social and pragmatic knowledge. Well before the two-word stage, children have learned to use gaze, pointing, and other cues to establish joint attention with their parents and from these infer their communicative intentions (Tomasello 1995). Lexical acquisition is sped along in part by children's ability to pragmatically infer lexical distinctions (Clark 2003, Bloom 2000). Moreover, by the time they enter the two-word stage, children have developed a wealth of structured knowledge about the participant roles involved in different events and the kinds of entities likely to fill them (Nelson 1996). All of this experience allows children to robustly interpret utterances beyond their productive abilities and respond appropriately to multi-word comments and queries from their parents even when they are producing only one word at a time (Bloom 1973).

One practical problem that arises is the lack of child language corpora appropriate for our task. Most transcripts in the CHILDES database include enough commentary for a human reader to disambiguate the situation, but they have not been systematically annotated with the semantic and pragmatic information assumed to be available to the learner. Several CHILDES corpora, however, have been annotated for semantic and pragmatic features available in the scene as part of a crosslinguistic study of motion utterances by Dan Slobin and his students (p.c.). Experiments on the learning model are based on a subset of this data, taken from the Sachs (1983) Naomi corpus, where the requisite input token features have been determined based on these additional annotations. (The data also includes Spanish, Italian, and French corpora to be used for future studies. In addition, comparable German and Mandarin Chinese data are also being similarly annotated to support crosslinguistic learning experiments.)

Besides the input data itself, the model assumes an initial set of ECG schemas for people, objects, locations, and actions familiar to children by the time they enter the two-word stage, as well as an initial set of ECG lexical constructions corresponding to these. It also takes the language analysis process mentioned above as a given. This process provides the means for mapping an utterance to its interpretation according to the current state of linguistic knowledge. Note that the same utterance may thus yield different interpretations at different stages of learning. For example, when the model (or the child) has only lexical constructions, the utterance "You can push the ball" from the input token in Figure 3 may allow only a partial interpretation. Specifically, assuming the model has not yet learned *can* or *the*, but does have some possible constructions mapping the forms *you* (for simplicity, mapping directly to the child Naomi), *push* (shown in Figure 2) and *ball* (mapping to a particular ball here called Ball-1), the analysis process would interpret the sentence as involving each of those independent meanings but not capturing their participant relations (as in Figure 4a). Further inference (drawing on the pragmatic skills mentioned above) would allow the likely role-filler relations to be deduced, based primarily on animacy constraints, but the language itself would not lead to the correct analysis. In contrast, at a later stage of learning when complex constructions like those in Figure 2 have been learned, a more coherent interpretation (as in Figure 4b) is available from the analysis.

To summarize: the information assumed available to children entering the two-word stage includes conceptual knowledge (ECG schemas), lexical items (ECG lexical constructions), a language analysis process that (partially) interprets utterances according to the current constructions, and pragmatic abilities that can supplement the analysis process. Incoming data is characterized as a set of (positive) examples pairing an utterance with its communicative context.

(a)

```
Push-Action
    pusher :
    pushee :
    action : Push-XSchema
Naomi
Ball-1
```

(b)

```
Push-Action
    pusher :   Naomi
    pushee :   Ball-1
    action :   Push-XSchema
```

Figure 4. Two possible (partial) interpretations of "You can push the ball" drawing on grammars at different learning stages. The structure in (a) shows a push action whose participants are not known, while the structure in (b) includes some of the appropriate role-fillers.

3.3 Learning processes

How does the model draw on information described so far to learn new constructions? The proposed learning processes are driven by comprehension: every incoming utterance presents the child with an opportunity to compare her pragmatically inferred view of the situation with her linguistically driven interpretation. The difference between these can prompt the formation of new constructions or the reorganization of existing constructions. We summarize the different operations possible; see Chang and Gurevich (2004) and Chang (in prep.) for details.

- **Hypothesis (relational mapping)**: New constructions can be hypothesized to capture information not covered in the linguistic analysis but available in context (e.g. the missing role-filler bindings of Figure 4a). These missing semantic relations can be associated with unused form relations (such as word order), resulting in complex (but lexically specific) constructions that incorporate previously known constructions as their constituents, as in the PUSH-BALL construction in Figure 2.
- **Reorganization**: Existing constructions can be reorganized on the basis of shared structure (using similarity as the criteria, as mentioned earlier). Three reorganization operations are possible: (1) Constructions with similar structure but different type constraints may be *merged* into a single construction through generalization (e.g., a PUSH-BALL and PUSH-BLOCK construction can be merged to create a new PUSH-Y construction that does not specify the pushed object). (2) Constructions with overlapping content may be *split* into smaller constructions based on non-overlapping portions (e.g., if SIT-DOWN and PUT-DOWN are learned as gestalts (each mapping to an entire corresponding scene), they may be reanalyzed as sharing a DOWN constituent and split to produce SIT and PUT constructions). (3) Constructions with overlapping but otherwise complementary content may be *joined* into one larger composite construction (e.g., the X-PUSH and PUSH-BALL constructions, which have similar meanings but assert fillers of different roles, may be joined into a larger X-PUSH-BALL construction).
- **Reinforcement**: A different kind of update to the set of constructions is not structural but usage-based: each construction is associated with a *frequency* that is incremented whenever it is successfully used to comprehend an utterance.

All of these operations are usage-based, motivated by the particular sequence of input tokens encountered. But multiple operations will often be possible in a given circumstance. The learning model also needs a way to evaluate potential operations, to determine which (if any) should be applied. The criteria for this evaluation is motivated by the need to recognize statistically common patterns while generalizing to previously unseen examples. This balance is achieved using a statistical framework based on a minimal description length principle (Rissanen 1978) that favors compactness in describing both the grammar and the statistical properties of the data. Every grammar (or potential grammar) can be evaluated in terms of its *size*—a sum over all its constructions of the number of constituents and constraints—and its *data complexity*—a measure of how well it captures the data (based on how many constructions are involved in the analysis, how frequently those constructions arise, and how many role-bindings are successfully identified). These competing factors embody a key tradeoff between generalization power (facilitated by smaller grammars with more abstract constructions) and predictive power (facilitated by larger grammars with more specific and frequent constructions and hence simpler analyses of the data). The extreme cases serve to illustrate this point: a grammar consisting of a single rule allowing any combination of words will fit all new data, though it may not be very informative about the specific examples likely to be encountered (or their semantic relationships); on the other hand, a grammar consisting of one (maximally specific) rule for each example seen captures the data perfectly but does not generalize well to new situations. The evaluation criteria provide a heuristic for guiding the grammar toward an optimal balance between these extremes.

The model has been tested in an experiment illustrating its ability to acquire simple English motion constructions, based on a subset of the input data described in Section 3.2. As described in Chang and Gurevich (2004), model performance was evaluated by measuring its comprehension (in terms of the proportion of situational bindings correctly identified using the current grammar) at several intervals throughout the learning. Results of this initial study showed that (1) the model improved its comprehension performance, climbing steadily from an initial state of producing mostly partial interpretations (as in Figure 4a) to a later stage of producing more complete interpretations (as in Figure 4b); the model acquired complex constructions, starting with fully lexically specific constructions and gradually including more abstract verb-island-like constructions (similar to those in Figure 2) that exhibited some ability to generalize to previously unseen data; (3) verbs differed in the rate of generalization and number of constructions learned in a data-driven manner. In current experiments we are further testing the model on a larger English corpus as well as comparable crosslinguistic data.

4. Discussion

The computational model we have described is intended to validate and formalize an approach to language acquisition that gives a primary role to meaning and usage in context. The model makes minimal assumptions about specifically linguistic biases and exploits semantically and pragmatically rich structures and processes that capture the wealth of experience brought to the task by children entering the two-word stage. This informal discussion omits many technical details, focusing on how the model's main structures and processes (the target grammar formalism, input data representation and learning framework) facilitate a meaning-oriented, usage-based and cognitively motivated approach. The model's ability to learn multi-unit constructions from data representing child-directed utterances provides at least a suggestive demonstration of how ideas in the literature can be precisely and consistently defined in a framework that approximates some characteristics of child language learning, including: incrementally improving comprehension over time, ability to generalize to new data, no explicit negative evidence, and the lexically specific nature of the earliest constructions.

The high-level structure of the model, though consistent with findings in the literature, should most properly be considered a starting point for refinement and experimentation. The learning processes and quantitative learning framework proposed offer particular scope for improvement. For instance, the hypothesis process leads to constructions associating form and meaning relations over previously known constructions, thus capturing some aspects of Slobin's proposed Operating Principles (1985), especially those related to mapping and frequency, and Clark's (1993) transparency of meaning principle. Similarly, the criteria for evaluating potential grammars can be seen as extending Clark's (1993) proposed principle favoring simplicity of form to include semantic description, as appropriate for the constructional domain; the tradeoff between grammar size and data complexity also produces overall behavior consistent with usage-based learning (Tomasello 2003, Lieven *et al.* 1997). But many other heuristics could be incorporated to model crosslinguistic patterns of language acquisition in more detail, and especially to investigate competing accounts of the acquisition of more abstract grammatical patterns generalizing over item-based constructions. Future work can also relax the simplifying assumptions mentioned in Section 2 to allow an explicit interaction between construction learning and concept acquisition, and to incorporate production-based learning processes. The current model thus offers not only a particular set of assumptions and claims about a restricted subset of the language acquisition problem but also the means of stating such assumptions and claims explicitly enough to support more detailed investigations.

References

Bailey, D. 1997. When push comes to shove: A computational model of the role of motor control in the acquisition of action verbs. PhD thesis, University of California at Berkeley.

Bates, E. 1976. *Language and context: The acquisition of pragmatics*. New York: Academic Press.

Bergen, B.K. & Chang, N. In press. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Ostman (editor), *Construction grammar(s): Cognitive and cross-language dimensions*. John Benjamins.

Bergen, B.K., Chang, N. & Narayan, S. 2004. Simulated action in an Embodied Construction Grammar. *Proceedings of the 26th Cognitive Science Society conference*. 108-113.

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

Bloom, P. 2000. *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bryant, J. 2003. Constructional analysis. M.S. report, University of California at Berkeley.

Budwig, N. 1995. *A developmental-functionalist approach to child language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chang, N. In preparation. Constructing grammar: A computational model of the acquisition of early constructions. Ph.D. dissertation, University of California at Berkeley.

Chang, N. & Gurevich, O. 2004. Context-driven grammar learning. *Proceedings of the 26th Cognitive Science Society conference*. 204-209.

Choi, S. & Bowerman, M. 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41: 81-121.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton & Co.

Clark, E. V. 1993. *The lexicon in acquisition*. Cambridge, UK: Cambridge University Press.

Clark, E. V. 2003. *First language acquisition*. Cambridge, UK: Cambridge University Press.

Croft, W. 2001. *Radical construction grammar*. Oxford: Oxford University Press.

Goldberg, A. 1995. *Constructions*. Chicago: University of Chicago Press.

Kay, P. & Fillmore, C. J. 1999. Grammatical constructions and linguistic generalizations: The What's X doing Y? construction. *Language* 75/1: 1-33.

Lakoff, G. 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Langacker, R. 1987. *Foundations of cognitive grammar: Theoretical perspectives 1*. Stanford: Stanford University Press.

Lieven, E., Pine, J. & Baldwin, G. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24, 187-219.

Maratsos, M. & Chalkley, M.A. 1980. The internal language of children's syntax. In Nelson, K.E. (ed.), *Children's Language*, 2: 127-214. New York: Gardner Press.

Narayanan, S. 1997. KARMA: Knowledge-based Action Representations for Metaphor and Aspect. Unpublished PhD dissertation, University of California at Berkeley.

Nelson, K. 1996. *Language in cognitive development: The emergence of the mediated mind*. Cambridge, UK: Cambridge University Press.

Regier, T. 1996. *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Rissanen, J. 1978. Modeling by shortest data description. *Automatica* 14, 465-471.

Sachs, J. 1983. Talking about the there and then: the emergence of displaced reference in parent-child discourse. In Nelson, K.E. (ed.), *Children's language*, 4: 1-28. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Slobin, D.I. 1985. Crosslinguistic evidence for the language-making capacity. In Dan Isaac Slobin (ed.), Theoretical Issues. *The crosslinguistic study of language acquisition*, vol. 2. Mahwah, NJ: Lawrence Erlbaum Associates.

Talmy, L. 2000. *Toward a cognitive semantics*. Cambridge, MA: MIT Press.

Tomasello, M. 1992. *First verbs*. Cambridge, UK: Cambridge University Press.

Tomasello, M. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.