# Embodied Construction Grammar

Benjamin Bergen
Nancy Chang

## 1. Introduction

### 1.1 A cognitive science approach to grammar

Theories of grammar are defined by the questions they ask. In the analytical linguistic tradition, this question is often something like:

> What is the most parsimonious formal account that predicts all grammatical (or acceptable) utterances in a language, and does not predict ungrammatical (or unacceptable) ones?

Construction-based approaches to grammar are driven by the insight that meaning plays a crucial role in answering this question: the meanings of both words and grammatical structures can affect grammaticality, and they should therefore be incorporated into grammatical representations. To wit, the various approaches presented in this volume take this tack in predicting the grammaticality of utterances of all sorts, from the creative to the conventional.

But focusing on utterance grammaticality removes utterances from their physical and social contexts of use. The result is an approach to language that resembles traditional anatomy, in which utterances are dissected, categorized, and explained in terms of their structural and distributional properties. Insights from such an approach can be both useful and informative in accounting for what utterances are or potentially could be. They do not, however, explain human language *use*—that is, the mechanisms and representations that human beings use to learn, produce, and understand utterances.

If we are interested in these latter issues, issues more properly situated within cognitive science than analytical linguistics, then we find our investigations of grammar (and language more generally) guided by a different sort of question:

> What cognitive and neural mechanisms do people engage while using human language?

Embodied Construction Grammar (ECG) is an approach to language defined by this question, one that of course is not unique in adopting this orientation (see, e.g. Bybee, 2006). Work in this paradigm follows the conventional scientific cycle that has proved fruitful in the study of other aspects of human cognition. Namely, observations about the acquisition, production, or comprehension of language (which we'll bundle together as "use" in this chapter) spur hypotheses about mechanisms that might account for the observations. Then theorists develop models, which are formally expressed so as to communicate claims explicitly to other researchers. They are often also computationally implemented, which allows assumptions to be validated, and often, new phenomena to be predicted. The models' predictions are tested against new observations, and the models are changed or rejected when observations contravene them.

ECG shares with other construction-based approaches an interest in how constructions of varying types contribute meaning and function to utterances. But its emphasis is not just on *what constructions look like* but on *how they are used*. That is, constructions are incorporated into models of language use; rather than just descriptive objects, as they are in most flavors of construction grammar, these form-meaning pairings are components of the hypothesized set of mechanisms engaged by language users. Any construction proposed to participate in the production or comprehension of a given utterance should therefore have observable consequences in that language usage event; there should be no formally vacuous constructions. In this sense, each constructional form-meaning pair represents a hypothesis to be validated through observations of behavior in natural and experimental settings. To facilitate the building of models that can validate such hypotheses, constructions in ECG are expressed in a formal notation that has a straightforward computational implementation. In short, ECG takes the insight that people use grammar meaningfully and functionally, and uses it to build an empirically driven, computationally implemented, predictive theory of language use.

## 1.2 The computational level

ECG asks what cognitive and neural mechanisms people engage when using language. The answers one arrives at will depend on the intended level of analysis. It might be possible in principle to answer the question at a relatively low level—entirely in terms of the biology, chemistry, and physics of the individual humans involved in language events. This sort of reduction is certainly an attractive ultimate outcome; cognitive science aims to explain a host of cognitive functions in terms of their material (including biological) underpinnings. But for language, we don't yet have the means to find answers at this level of detail; single-cell recording from living humans is only rarely possible, and animal models provide only limited evidence about specifically human language. Moreover, it's not obvious that such a low-level explanation would be especially informative to the analyst; knowing how billions of cells individually behave might not shed much light on the higher-order processes in which those cells participate, their causes, effects, properties, and so on. In fact, a full reduction to biology might not be any more predictive of new observations than a high-level, functional computational model of the mechanisms involved. For these reasons, ECG is articulated at a higher, functional level. We develop computational models, in which mechanisms are described in terms of structures and processes: what people appear to know and how they appear to learn and use that knowledge. At the same time, these models are proposed with an eye toward seeking connections to their underlying biological substrate; toward developing a linking theory between computational and biological levels of description (Feldman, 2006; Marr, 1982).

The computational level of description used in ECG is considerably more detailed, however, than the kinds of description found in most construction-based accounts and other work in cognitive linguistics. These accounts, though insightful, tend to stop short of the level of formal precision needed for building detailed computational models of how constructions are represented, used and learned. The ECG formalism (as introduced in Bergen & Chang, 2005) is intended as a means of bridging the gap between these levels, using a set of notations for the kinds of high-level relations and constraints expressed (typically discursively) in the literature. The formalism itself is thus not a linguistic theory *per se*; rather, it is a theory of what conceptual distinctions are necessary and sufficient to account for language phenomena.[1]

---

[1] Some additional notations have been introduced since our original formulation; these allow inherited constraints to be overridden by subcases, explicit marking of optional constituents, and

The exposition in this chapter will make use of formal notation and will mention computation quite a bit, but it is not intended as a thorough introduction to either the ECG formalism or its associated computational models. In fact, we will present relatively simple analyses and avoid most technical details; these are available elsewhere (e.g., Bergen & Chang, 2005; Bryant, 2004, 2008; Chang, 2008; Dodge, 2010; Feldman, 2006; Feldman & Narayanan, 2004; Mok, 2008; Mok & Bryant, 2006; Mok, Bryant, & Feldman, 2004). Formal notation will be used only at the service of highlighting or explicating properties of theoretical components of ECG. Likewise, we will describe the computational models expressing these theoretical ideas, but say relatively little about their concrete implementations (the systems that realize the proposed structures and processes). This is because it's the theory itself and the properties of computational models designed to express that theory that are of broadest relevance to cognitive scientists of language.

### 1.3 Mental simulation

Since ECG aims to account for the mechanisms of language use, the structures and processes we propose should naturally be constrained by evidence about how people actually produce and comprehend language. The past decade has seen an explosion of behavioral and brain imaging research on language comprehension, focusing on meaning. One foundational finding is that people understand utterances by performing mental simulations: they engage perceptual, motor and other brain systems to create internal experiences similar to those they would have when experiencing the described scenes. To accommodate this finding, ECG includes mental simulation as a component of language use, and represents words and other constructions in terms of how they drive the specific perceptual and motor experiences that the language user mentally simulates. Section 2 reviews mental simulation results, introduces the notion of language as an *interface* to simulation, and briefly illustrates how form, meaning, and constructional knowledge are represented in the ECG formalism. Section 3 describes grammatical constructions and the kinds of contributions they can make to mental simulation. In Section 4 we focus more directly on language processing, in particular on a model of language comprehension developed within the ECG framework that is compatible with both the idea of mental simulation and relevant evidence from the sentence processing literature.

## 2. Mental simulation and language

### 2.1 Meaning is grounded in mental simulation

Over the past century, the gradual spread of the study of language into the behavioral and physical sciences has been driven predominantly by investigations of linguistic form—i.e., observable characteristics of linguistic behavior. There is a practical reason for this. It is relatively straightforward to observe word order or measure the amplitude of a stop consonant's burst. By contrast, meaning is invisible, internal, and subjective. Thus, while speakers' intuitions can provide some clues to the nature of semantic and pragmatic structure, the study of how people access and construct meaning during online language use has lagged far behind the study

---

the incorporation of distributional information essential for the probabilistic version of the ECG analyzer described below. The basic structural components remain stable, however.

of linguistic form. The past decade, however, has seen substantial progress toward understanding how meaning works, especially in language comprehension. This work has attacked the problem from several angles, based on indirect measures of cognition (such as reaction times, eye movements, changes in fMRI signal, etc.) taken while people are processing language under tightly constrained experimental conditions. Crucially, a variety of designs and measures have been converging on the increasingly stable conclusion that when people process words or sentences, they activate "mental simulations" of the described scenes.

Mental simulation is the internal engagement of modality-specific brain systems to create or recreate non-present experiences (Barsalou, 1999). Mental simulation can be visual, in which case parts of the brain's visual system become active in ways similar to how they would react if external stimuli were actually present (Kosslyn et al., 2001). Mental simulation can also be motor; people activate parts of their brains dedicated to motor control as an internal model of physical action, usually without actually engaging their skeletal muscles. People automatically (and mostly unconsciously) engage motor simulation during a variety of cognitive behaviors, such as intentional mental imagery (Ehrsson, Geyer, & Naito, 2003), recall (Nyberg, et al., 2001), and other tasks.

Recent work on language comprehension points to a role for mental simulation: when people process language about perceptible things and events, they activate perceptual representations of what the mentioned items would look like (Stanfield & Zwaan, 2001), sound like (Winter & Bergen, submitted), and so on. These perceptual simulations encode implied details not explicitly mentioned in the language. For instance, people represent mentioned objects with the most likely shapes (Zwaan, Stanfield, & Yaxley, 2002), orientations (Stanfield & Zwaan, 2001), or colors (Connell, 2007) they would have, depending on the context. These simulations may adopt a particular perspective, or viewpoint, from which the comprehenders simulate perceiving the event, and grammatical person (e.g. *you* vs. *he*) affects which perspective is adopted (Brunyé et al., 2009). Similarly to language about perceivable things, when people process language about actions, they engage motor representations of those actions (Glenberg & Kashak, 2002). These motor representations are quite specific, down to the body part used to perform the action (Bergen et al., 2010; Pulvermueller et al., 2005), or the most likely hand shape (Bergen & Wheeler, 2005; Masson, Bub, & Warren, 2008). Roughly, then, understanding seems to involve activating internal representations of perceivable things or performable actions, even without the things being there or the actions being performed. (By comparison with the growing body of work on mental simulation in comprehension, there has been relatively little work on mental simulation in language production, though Hostetter & Alibali (2008) and Sato (2010) provide evidence from gesture and reaction time studies, respectively, that it plays a role there, too.)

Experimental evidence leaves little doubt that people perform mental simulations during comprehension. It remains less clear precisely what role such simulations play in language use, and further investigation is needed to ascertain to what degree and under what conditions simulation is necessary for particular aspects of language understanding. In principle, however, mental simulation could offer elegant explanations for a variety of phenomena associated with meaning. In particular, we hypothesize that comprehenders may be able to understand language about non-present entities and events because the processing of such language triggers brain states similar to those that result from experiencing those entities and events (a claim compatible with arguments by Barsalou, 1999, Glenberg & Robertson, 2000, Pulverlmueller et al., 2005, i.a). Further, during simulation, these brain states may activate various related processes, in much the same way as the actual perception or the performance of real actions, thus allowing comprehenders to enrich their understanding by drawing detailed, relevant inferences that are grounded in sensorimotor experience and sensitive to contextual conditions. And finally, attention to perceptual and motor simulation can effect the comprehender's subjective feeling of understanding the topic being discussed. In sum, simulation has the potential to explain how we ground meaning in the brain's perceptual and motor systems.

ECG is a intended to be consistent with the evidence above. It is a theory in which the potential functions of simulation can be proposed and modeled, such that detailed predictions can be generated for testing. More specifically, it provides the means for exploring the hypothesis that linguistic constructions serve as an interface to mental simulation. The remainder of this section outlines minimal requirements for such an interface and briefly describes how ECG satisfies them, beginning with words and their meaning representations.

## 2.2 An interface to mental simulation

The idea that language understanding involves mental simulation calls for a radical rethinking of what words and other linguistic constructions mean, as well as how this meaning should be represented. Traditionally, linguists use quick-and-dirty approximations of meaning, often dodging the issue of what meaning is by merely labeling word meanings (for instance, the meaning of the word *cat* might be represented as CAT or as a collection of ungrounded features, like [+FELINE, +DOMESTIC]). But work on mental simulation suggests that what language users know is how to take a linguistic form (such as a word or grammatical structure) and activate perceptual or motor representations corresponding to the denoted entities and events, and vice versa. That is, linguistic units drive mental simulations: instead of fizzling out in static, formal, arbitrary representations (like CAT), meaningful words and constructions serve as interfaces that activate and constrain modality-rich mental simulation.

Stated more concretely, in ECG, words and other constructions serve as pathways connecting detailed, modality-specific knowledge about their forms with detailed, modality-specific knowledge about their meaning. Critically, however, these pathways are mediated in both domains by categorization. The range of possible phonetic and graphical realizations of a word varies dramatically, not just across but even within modalities (consider the difference between cursive *cat* and block print stencil *cat*). Yet we categorize these different tokens, which vary continuously within visual and auditory space, as instances of the same category—instances of a single word form (Dahaene, 2009). The possible denotations of *cat* vary similarly; even restricting ourselves to the domestic cat (*felis catus*), the perceptual and motor components of relevant simulations vary widely not only in visual features (color, size, speed, and adorableness) but also in haptic, auditory, perhaps olfactory, and even gustatory features. Just as the formal realization of the word form *cat* varies across uses, so does the realization in simulation of the word's meaning. And critically, for the most part, the two do not co-vary: it is not the case that *cat* written in cursive denotes a particular type a cat (say, a calico), while stencil block letters denote a tabby. Rather, the multifarious formal realizations of a word activate a mentally represented category of experiences with things in the world.

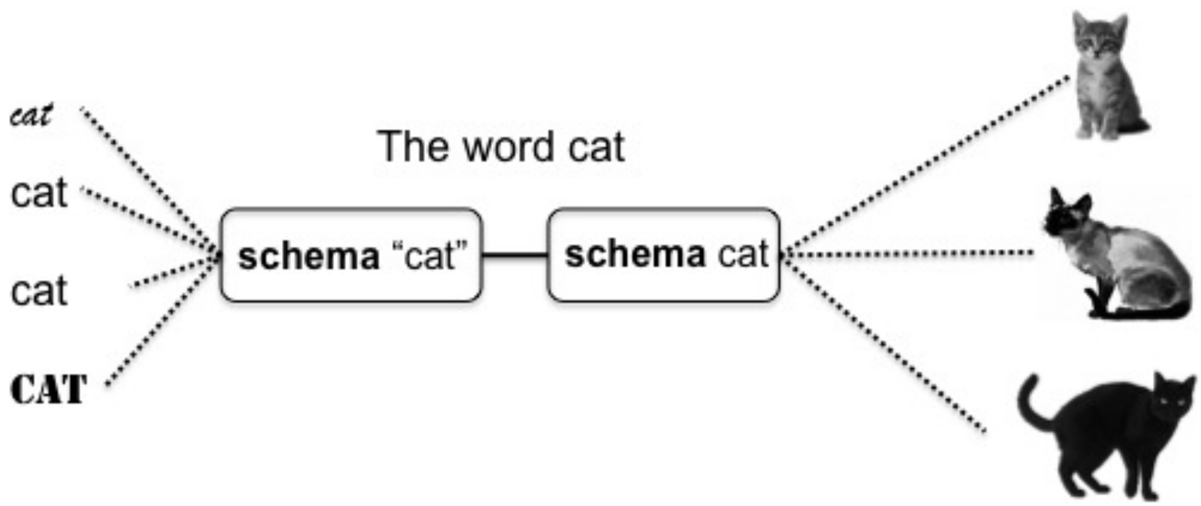


Figure 1. Words as bipolar constructions linking schematic representations in the domains of form and meaning.

Following the cognitive linguistic literature, we call categorical mental representations that generalize over instances (like the form or meaning generalizations relevant to *cat*) "schemas" (Johnson, 1987; Lakoff & Johnson, 1987). And following construction-based approaches to grammar, we call the entire structure that links such schemas a "construction". As suggested by the image in Figure 1, constructions are "bipolar": they serve to connect schemas in the two domains of "form" (phonological, orthographic, gestural, etc.) and "meaning" (conceptual, ontological, situational, etc.). As we will see below, the ECG formalism provides notations for defining these terms that are designed to be consistent with the literature.

Crucially, however, the approach pursued here further requires an active, grounded notion of meaning capable of supporting mental simulation. For this we turn to "simulation semantics" (Bailey 1997; Feldman 2006; Naranayan, 1998), an approach to semantic representation developed with precisely this aim in mind. Broadly, the same dynamic control systems that people use to perform actions or perceive events are hypothesized to also be used to simulate actions and events. Though a full exposition goes beyond the scope of this chapter, Figure 2 gives a taste of how actions (in this case, the action of jumping) might be represented using the "executing schema" (or "x-schema") formalism (Narayanan, 1998). The x-schema shown captures the flow of activation through a motor control system responsible for the performance or simulation of jumping. First, assuming all enabling conditions are met (e.g., the jumper is positioned on a stable surface and has sufficient energy to expend), the jumper prepares by coiling its legs. Then comes the explosive phase in which the legs (and body) are rapidly extended and the jumper is propelled upwards from the ground, resulting in a ballistic period (corresponding to the node marked "done") in which the airborne jumper is subject to the force of gravity. Like the cat category above, this jumping action schema is intended to support considerable variation: the jumper might be a human, a cat, or some other animate agent; the action may vary in force expended, launch height attained, direction of motion, or bodily orientation; and it might precede a host of other actions (landing in the same place, landing on a higher or lower surface, grabbing a trapeze, flying through the air, falling, etc.).

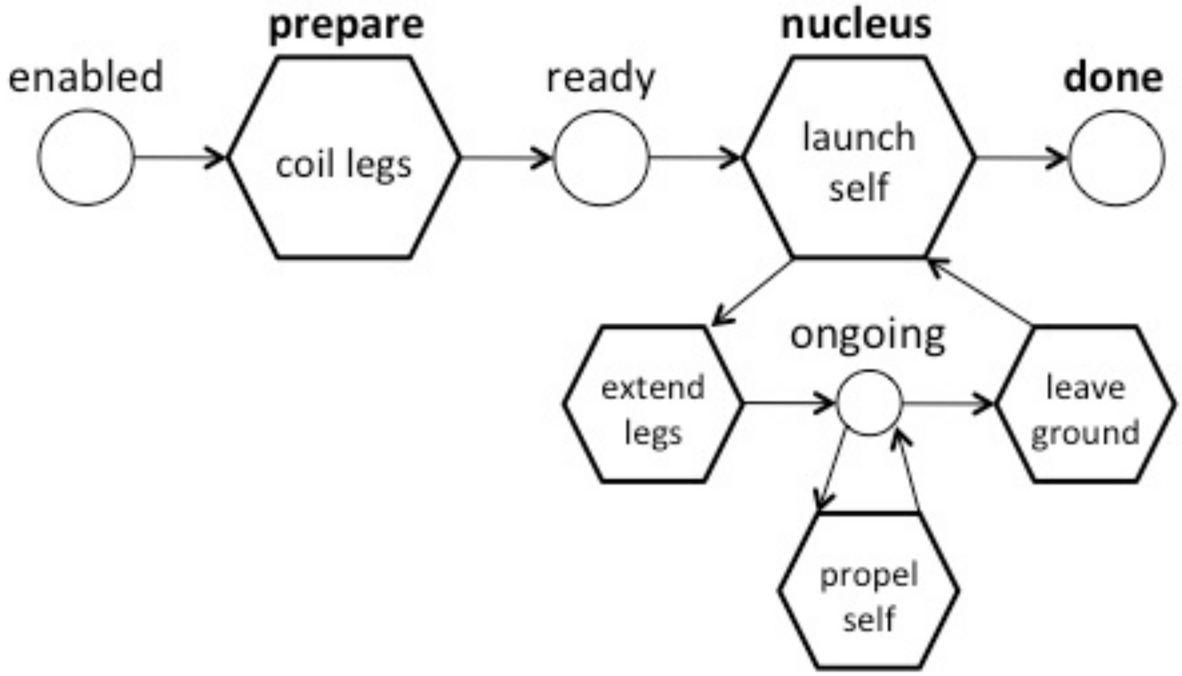

Figure 2. An x-schema for jumping, showing different stages of the motor action.

Language doesn't usually need to reach into sensorimotor knowledge at this level of detail, of course. In fact, the simulation-based framework hypothesizes that specifically linguistic knowledge need concern only a limited subset of the rich underlying action schema, enough to describe (or "parameterize") the relevant simulation. For example, a linguistic unit like the verb *jump* might evoke this motor action as a gestalt, perhaps allowing a few features (e.g., identity of jumper, rate, direction, goal) by other linguistic units. In other words, linguistic meaning provides a limited interface to mental simulation, which in turn activates dynamic, embodied schemas that combine with each other and the current context. The result of simulative inference constitutes the full understanding of an utterance in context.

## 2.3 Form, meaning, and constructions

Now that we've drawn the broad outlines of how ECG uses constructions, schemas, and how these interface with mental simulation, we turn briefly to how ECG represents this linguistic knowledge. In comprehension, schemas serve as the means by which a recognized word activates perceptual or motor components of a simulation. The simplest schemas do little more than this, serving essentially as pointers to experientially grounded gestalts, in either the form or meaning domain. A simple conceptual schema illustrating how the meaning category for *cat* is represented in the ECG formalism appears below. It includes a name of the schema (Cat), along with a notation indicating an inheritance relationship: part of what you know about the category Cat is that it is a subtype of a superordinate category Animal. The Cat schema does not explicitly list various other salient features that are by default associated with the category (e.g., that it is an animate physical entity), since these are inherited from the Animal schema. While representations like this one may appear similar to classical stand-ins for meaning (like CAT), it is crucial to recall that they are meant to serve as pointers to detailed, modality-specific structures that are activated during language use.

```
schema Cat
  subcase of Animal
```

Form schemas are represented in a similarly simple way, where again, schemas serve as pointers to more complex and variable phonological, articulatory and auditory representations. The Cat-Form schema below states that it is a subcase of a more general Word-Form schema (the category of forms of words in the language), and that among its "roles" (or features) is its orthography, which is the string "cat".

```
schema Cat-Form
  subcase of Word-Form
  roles
    orth ← "cat"
```

These form and meaning schemas are linked by a construction, Cat-Cxn, whose form is defined to be an instance of the Cat-Form schema above, and whose meaning is defined as an instance of the Cat schema, also above. (This construction corresponds to the line connecting the form and meaning schemas in Figure 1.) The Cat-Cxn is also a subcase of a class of similar constructions, which for current purposes we'll call Noun. The Noun construction, not shown, pairs an unspecified Word-Form on the form side with some entity on the meaning side.[2]

```
construction Cat-Cxn
  subcase of Noun
  form : Cat-Form
  meaning : Cat
```

Other schemas and constructions are more complex. Words like prepositions, verbs, or adjectives, whose meanings are necessarily computed in combination with the meanings of other words, involve relations among multiple structures, some of whose internal components must be supplied during language use by the linguistic or physical context. The word *jumped*, for example, invokes the action of jumping, which inherently includes some animate entity doing the jumping. As noted earlier, the x-schema associated with this action (Figure 2) involves rich details about various stages and subactions, but the conceptual Jump schema might be more simply defined in terms of its key parameters (or roles). Below we represent the Jump schema as a subcase of Action (not shown), which is a conceptual category of events in which some animate agent acts intentionally. It is defined here as having one participant, specified as the jumper under the **roles** block, which is bound to (or "identified with") the agentive actor (inherited from the Action schema), as indicated by the double-headed arrow.

```
schema Jump
  subcase of Action
  roles
    jumper ↔ actor
```

[2] This version of the Cat-Cxn is a massive simplification of both the form and content that would be hypothesized in a fuller analysis, but it captures the general flavor of the formalization.

The Jump schema appears in the meaning pole of the Jumped construction defined below, paired with a form specified by the Jumped-Form schema (not shown, but similar to the Cat-Form schema, mutatis mutandis). The Jumped construction is also defined as a kind of verb; the Verb construction (not shown) pairs an unspecified Word-Form with a perceptual or motor schema. Verbs may also specify constraints on the time of the reported event (relative to speech time), as shown in the meaning pole of the construction below, where the time role is inherited from the Verb construction. In the Jumped construction, for example, the time of the jumping event is constrained to be in the past, i.e., it took place prior to speech time.

**construction** Jumped
 **subcase of** Verb
 **form:** Jumped-Form
 **meaning:** Jump
  time ← past

These examples illustrate the general approach to defining ECG schemas and lexical constructions: together, the formal notations above provide ways of specifying which form and meaning schemas are linked, as well as which embodied schemas should be simulated with what parameters. It is fair to ask, however, how exactly one decides what schemas and constructions to hypothesize—what participant roles might appear, how many different schemas there are, how related concepts might be organized, etc. After all, the space of possible concepts ranges over a variety of modalities and domains (motor, perceptual, affective, social), which involve different specialized neural and cognitive structures. But just as the dynamics of the complex motor action of jumping can be summarized, or schematized, in terms of several parameters relevant to the verb *jump* and other related constructions, so we assume that the space of possible concepts can be schematized. Indeed, many schemas may be relevant for multiple words (variants of the Jump schema might be used, for example, for *jump*, *leap*, and *hop*). At the same time, many words may evoke multiple schemas. For instance, in the context of playing checkers, jumping your opponent's pieces doesn't involve anything like the same motor control as the type of jumping a cat does; moreover, the relevant schema invariably includes a participant that undergoes the jumping, with implications for differences in how this sense of the verb may be realized (for example, it is transitive). To make matters worse, some schemas may not correspond directly to any particular construction; we may have schematic knowledge about proprioceptive, affective, gustatory, social, and other categories of experience that we never learn to put into words.

In fact, speakers of different languages converge on different organizations of conceptual space into schemas and constructions, and even different speakers of the same language may vary in the particular representations they acquire. Previous work in simulation semantics has shown how machine learning techniques can be used to learn mappings between action schemas and word senses (see, e.g., Bailey, 1997), modeling how the same verb may have multiple related (embodied) senses (e.g., pushing a block on a table or a button on a keyboard differ distinctly in the direction of movement and default hand posture). Ideally, such learning models would provide a dynamic, experientially driven basis for defining schemas and constructions.

More generally, however, schemas and constructions may also be defined by the analyst to account for a particular linguistic phenomenon or set of data. In such cases, our procedure is conservative: we hypothesize a schema for each sense of each morpheme in a language (where enumerating senses is admittedly an enterprise fraught with uncertainty), except where another hypothesized schema can already account for the processing of that morpheme. In addition, we hypothesize higher-level schemas where linguistic patterning suggests them; such schemas overlap with what have been called "image schemas" in the literature, such as Containment, Source-Path-Goal, and Trajector-Landmark (Johnson, 1987). These schemas are often relatively simple, and when notated in the ECG formalism along with their associated roles they bear a strong (and deliberate) resemblance to informal descriptions appearing in the literature. Like other ECG schemas, they are posited to serve as interfaces to the richer embodied structures supporting mental simulation. They are also, however, more general in their applicability; they recur across a wide range of event types and are lexicalized and grammaticized crosslinguistically.

## 3. Grammatical constructions and the specification of simulation

Consistent with a constructional view, ECG treats all linguistic units, including lexical items like those illustrated above, as well as larger phrasal constructions and other traditional grammatical notions, as variants of the same kind of thing—mappings between form and meaning. What distinguishes many grammatical constructions from simple words, however, is that they exhibit complex internal structure with multiple constituents, each of which instantiates a constructional form-meaning mapping. From a simulation-based perspective, this means that they are particularly well-suited for parameterizing complex mental simulations involving multiple, variable conceptual schemas.

To illustrate some features of grammatical constructions in ECG, we give a simplified definition for one common construction, the DirectedMotion construction, below. This is an argument structure construction (as noted in the subcase line) licensing sentences like *The horse raced past the barn* or *The cat jumped down,* in which a mover (designated by some Referring-Expression) moves along some path (designated by a Path-Expression) by means of some action (designated by the verb). The DirectedMotion construction specifies these three constituents, and in the notation shown below, each is given a local alias (r, v, or p) by which the construction may refer to them. Each of these constituents is also restricted to be of a particular construction type. (Though not shown here, Referring-Expressions are roughly analogous to noun phrases, while Path-Expressions express a path, and include particles and prepositional phrases.) Similar constructions are defined in, e.g., Bergen & Chang (2005) and Chang (2008).

**construction** DirectedMotion
  **subcase of** ArgumentStructure
  **constituents**
    r: Referring-Expression
    v: Verb
    p: Path-Expression
  **form**
    $r_f$ **before** $v_f$
    $v_f$ **before** $p_f$
  **meaning:** Motion
    mover $\leftrightarrow r_m$
    action $\leftrightarrow v_m$

```
    path ↔ p_m
```

```
schema Motion
  roles
    mover : Animate
    action: Action
    path : SPG
  constraints
    mover ↔ action.actor
```

The form of this construction, specified in the "**form**" block of the construction, is more complex than the forms of most lexical items: it states that the form of the referring expression precedes the form of the verb, and that the form of the verb precedes that of the path expression; this is meant to hold whether the construction is used in writing or speech, or whether in producing or comprehending an utterance. In its meaning, the DirectedMotion construction above denotes a motion, specified by the Motion schema (also shown). In the DirectedMotion construction, the roles of the Motion schema are explicitly bound together in the "**meaning**" block to the meaning poles of the Referring-Expression, Verb, and Path-Expression constituents, respectively. That is, the meaning of the Referring-Expression (e.g., a structure involving the Cat schema in the case of *the cat*) will be bound to the mover of the Motion schema. (The constraint in the Motion schema further binds this to the actor of its associated action schema.)

Of course, the DirectedMotion construction defined here has been simplified for expositional purpose, and it elides many issues that would arise in defining a larger-scale grammar. For such grammars, one may wish to handle not only *The cat jumped* but also sentences like *Into which bucket did the cat jump* and *Which bucket did the cat jump into,* which violate the ordering constraints above, and involve a different predication type. Two potential solutions have been adopted in the ECG literature. The first is to proliferate argument structure (and other) constructions into families, such that the construction shown above (which might be renamed DeclarativeDirectedMotion) is a subcase of a more schematic DirectedMotion construction, which itself does not specify ordering of its constituents. Other subcases of this general construction identify the specific constituent orderings and predications types required for interrogative and other uses of DirectedMotion. The second solution is to factor out word order and predication type entirely from argument structure constructions, and allow it to be provided by other sentence-level constructions, like a Subject-Auxilliary-Inversion Construction, and so on. These two solutions (families of argument structure constructions versus factoring out of word order) are explored in more detail in Dodge (2010).

Much of the work in ECG has focused on grammatical constructions like these, by specifying their forms and meanings, inheritance relations, and how they with interact with one another (see, for instance, Bryant, 2008). While such analyses vary in precisely how they partition the space of constructions and schemas, and what level of specificity they aim for, they make a common set of predictions about how grammatical constructions can help specify mental simulation and thereby affect meaning processing, falling into three broad classes.

First, a grammatical construction may align (or "identify" or "bind") different aspects of the meanings of its constituents, bringing together their various contributions to mental simulation. For instance, an argument structure construction like the DirectedMotion construction designates one of its constituents (the referring expressing) as the mover engaged in the motion denoted by the verb along a path denoted by the path expression. There's nothing new or specific to ECG in this claim; the only new contribution ECG makes is to predict that manipulations of such constituent-aligning constructions will affect the mental simulations comprehenders generate in ways predictable from constructional meaning. For instance, switching the subject and prepositional object (*The fence jumped over the cat*) should result in a (potentially unlikely) mental simulation in which the mover and landmark are switched. To our knowledge, this claim has not been tested experimentally, perhaps because it is largely shared with other accounts, though there's some evidence (Glenberg & Kaschak, 2002) that switching subject and object affects simulated direction of motion.

Second, a grammatical construction may contribute content directly to mental simulation—evoking categories of experience (schemas) that supply part of the comprehender's modal representation of the described entities and events. This corresponds more closely to the type of contribution of grammatical constructions unique to the construction grammar literature. For instance, the DirectedMotion construction, described above, might lead comprehenders to simulate events as though they involved an animate mover moving along a path, even when the verb doesn't explicitly specify motion (as in *The cat meowed down the street*). Other examples include the Double-Object or Ditransitive construction, which has been argued to activate a Transfer-of-Posession schema, as contrasted with the Prepositional Dative, which may activate a Caused-Motion schema (Goldberg, 1995). The ECG view proposes that these argument structure constructions activate the relevant schemas, bind them with the meaning contributions of the constructional constituents, and then together drive mental simulation of the unified scene. There has been a bit of experimental work addressing construction-specific hypotheses of this sort, comparing the transitive and ditransitive (Kaschak & Glenberg, 2000), which shows comprehenders accessing different meanings for sentences using the different constructions. But this experimental work is in its infancy.

Third, grammatical constructions may simply modulate second-order properties of simulation, like perspective or locus of attentional focus. For instance, active sentences might lead the comprehender to simulate an event from the perceptual or motor perspective of the agent, while passive sentence might lead to simulation from the perspective of the patient. To take another example, aspectual constructions (like the English progressive versus perfect) might modulate what part of a scene comprehenders simulate in more detail. Such predictions appear to be specific to simulation-based approaches to language use, like ECG, and have been borne out experimentally (e.g. Bergen & Wheeler, 2010; Choe & Bergen, In prep; Madden & Zwaan, 2003).

In sum, grammatical constructions in ECG are hypothesized to contribute to mental simulation by binding together the conceptual schemas evoked by their constituents as appropriate, by contributing conceptual schemas directly to be mentally simulated, and by imposing second-order constraints on how simulation is to be enacted. A theory of how people use grammar in the ECG framework includes both an account of how grammar constrains simulation in these broadest terms, as well as computationally implemented models fleshing out empirically testable claims about how specific constructions affect mental simulation in one of (or combinations of) these three ways. The next section provides further detail on these two aims, with specific reference to language processing.

## 4. Language processing

How does ECG support processes of language use? In this section we present the processing side of ECG: a model of language comprehension that uses ECG constructions, and uses mental simulation on the back end. (See also Bergen & Chang, 2005; Bryant 2004, 2008; Chang, 2008; and Mok, 2008.) Our models of language comprehension are continually evolving in response to new computational techniques and new findings about human language comprehension, as well as in response to our efforts to address more language phenomena. The core ideas, however, are relatively stable. In the sections below, we first discuss how a specific sentence might be understood and what general requirements such a process might impose; we then describe a computational model of language comprehension that satisfies those requirements in a cognitively plausible way.[3]

## 4.1 Comprehension as constraint satisfaction

The foundational "simulation hypothesis" of our language comprehension model is that linguistic constructions of all kinds parameterize mental simulations: they evoke the experiential schemas (corresponding to events, actions, objects, etc.) involved in a particular utterance and specify how these are combined. The resulting gestalt is simulated to yield the inferences that effect useful comprehension of the utterance. The notational mechanisms described in Section 3 provide a means of specifying which form and conceptual categories are involved, and what relationships in the form and meaning domains must be satisfied. Consider the sentence below:

(1) The cat jumped down.

Restricting ourselves for simplicity to the sentence in written form (we'll assume for present purposes that the input is just an orthographic string, pre-segmented into words), a number of surface cues are already available: the individual word forms, the order in which they appear, and a punctuation mark. Understanding this sentence involves using these surface cues to activate various pieces of linguistic knowledge—constructions for the words *the*, *cat*, *jumped*, and *down*, as well as larger grammatical constructions, like DirectedMotion. These in turn evoke various conceptual schemas, including some we have already seen (Cat and Jump) and others like a Down schema capturing the spatial direction evoked by the word *down*. As before, these schemas are not mere symbols, but rather are generalizations associated with experientially grounded attributes and functions: the notion of Down evokes visual, motor, and proprioceptive knowledge about orientation and motion of some trajector towards the earth.

During sentence processing, lexical constructions activate these concepts, which are combined by more complex grammatical constructions, as described above. For instance, the comprehender may recognize the string *The cat jumped down* as an instance of the clause-level DirectedMotion construction, defined in Section 3. DirectedMotion includes a PathExpression constituent in the postverbal position, and if the schema associated with the Down construction satisfies the type constraints that DirectedMotion imposes on its path, then it can bind to the schema denoted by Down to the meaning of its PathExpression constituent. The comprehender also binds the Action event in the meaning of DirectedMotion to the Jump schema, and the

[3] Work in language processing in the ECG framework has focused on language comprehension, since there has been more experimental and computational work on comprehension than on production, resulting in more tightly constrained models.

Mover role to the Jumper of the Jump schema and also with the Cat schema. Moreover, the jumping event is constrained to take place prior to speech time (as indicated by the past-tense verb *jumped*). And the use of *the* picks out a specific cat of which this jumping event is asserted. The hearer may use a more general determined noun phrase construction (not shown here, but see Bergen & Chang, 2005) underlying phrases like *the cat* and *a cat*, which specifies the word order of the determiner and noun, and combines their constraints each imposes on the identity of the referent. In the example, we assume the determiner *the* restricts the referent to a particular member of the relevant category (i.e., that specified by the noun *cat*) that is identifiable (or given) to the hearer in the current context.

The resulting configuration of schemas is precisely the parameterization we have claimed is used to drive a mental simulation; we call this configuration a "semantic specification", or "semspec". The semspec dictates what scenario to simulate and how, and the comprehender uses perceptual and motor systems to perform this simulation. A semspec corresponding to a possible interpretation of the example sentence using some of the constructions defined earlier is shown in Figure 3. The forms in the input utterance are shown in the box on the left; some constructions instantiated in the analysis are shown in the central area; and their associated semspec (i.e., collection of bound meaning schemas) is shown in the box on the right. Another way to think about the figure is as depicting the linear structure of the input sentence string on the left; the constituent structure of the constructions involved in the middle; and the situational structure of the identity and role-filler relations among the meaning schemas evoked by the sentence on the right. (Note that we've left out the constructional and meaning details associated with "the" for the purposes of this exposition.)

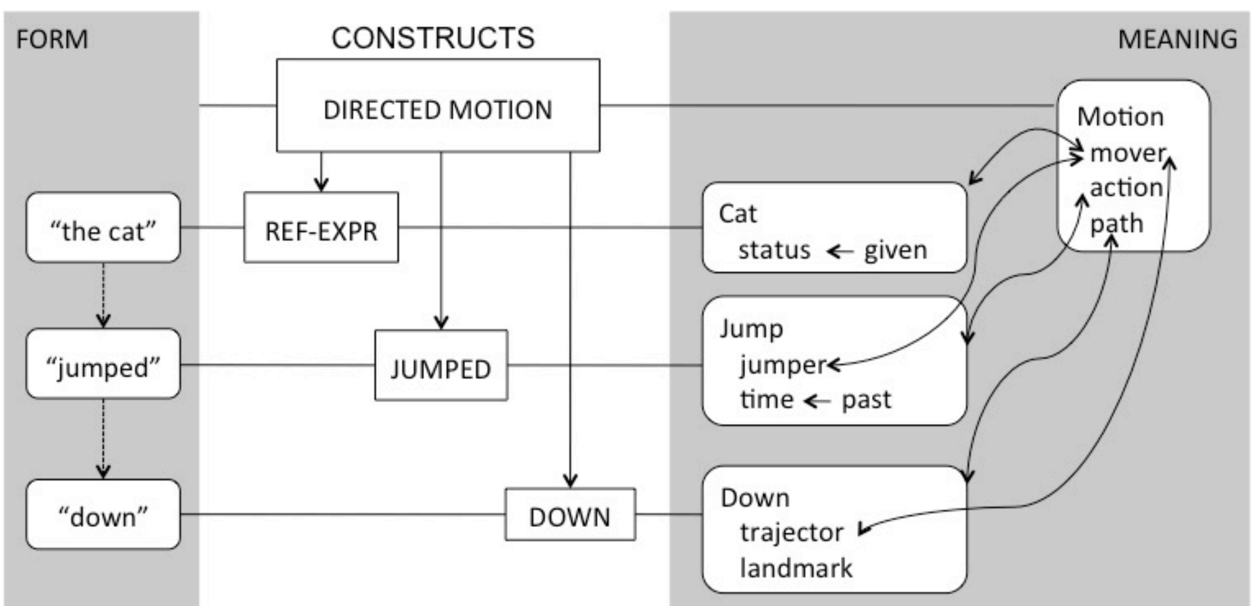


Figure 3. A (simplified) constructional analysis of the example sentence, with *constructs* (instances of constructions), shown in the center. They link their associated form in the input sentence on the left to their meaning in the semantic specification on the right.

The semspec captures much of what might be considered the traditional meaning of the sentence, such as basic thematic relations (who is doing what to whom), spatial and temporal relations, and discourse constraints about the information status of the various referents (e.g. the cat's status as given). But the semspec alone doesn't provide *everything* we need to know to perform the relevant simulation and account for the many other inferences a hearer might make. In fact, many specific inferences may depend on the particular jumper and path, which may be based, among other things, on the current context, as well as on general knowledge about the relevant categories. Specific participants identified (or instantiated) in context lead to simulations that can produce much finer-grained inferences about the jumper's position before, during and after the event; the amount of energy expended and the involvement of feline femurs and paws; the default height of jumping, and so on. That is, knowing how to simulate involves more than merely what categories the language evokes; it involves resolving references made using those categories to specific instances of them.

We therefore distinguish the semspec above—which captures the general bindings and constraints imposed by the relevant constructions—from a fully "resolved semspec" that further constrains these based on the hearer's knowledge and the current situational and discourse context. Only once reference has been resolved does the comprehender have enough information to engage in a contextually appropriate simulation; the hearer can then use the results of simulation to update her beliefs about the situation, or prepare to respond appropriately to the utterance. For our example, the resolved semspec corresponding to the semspec above might instantiate the general Cat category with a specific cat salient in the physical or linguistic context, or (in the absence of any such instance) with an instance selected arbitrarily or due to prototypicality.

It is worth noting that individuals are overwhelmingly likely to bring different constructional knowledge to bear on language use. The different constructions that each language user has learned over the course of his or her unique exposure to language may lead to the construction of subtly or substantially different (unresolved) semspecs when processing the same utterance. To the extent that these different semspecs lead to different resolved semspecs and different simulation-based inferences, communication will fail. But because an unresolved semspec serves as only an intermediate step in comprehension, it is possible, on an ECG view, for dyads to meet with some communicative success even with non-identical linguistic knowledge, again, to the extent that the various processes the comprehender engages produce inferences compatible with the intentions of the speaker.

## 4.2 A computational model of language comprehension

The account of comprehension described above is summarized in Figure 4, which distinguishes three main processes involved in understanding an utterance in context:

- constructional "analysis"**:** the identification of which constructions are instantiated by a given utterance and how they are related, along with the assembly of an associated "semantic specification" (or "semspec") identifying what meaning schemas are evoked and how they are related
- contextual "resolution": the mapping of objects and events in the semspec to the current communicative context, producing a "resolved semspec "
- embodied "simulation": the invocation of dynamic embodied structures in the (resolved) semspec to yield contextually appropriate inferences

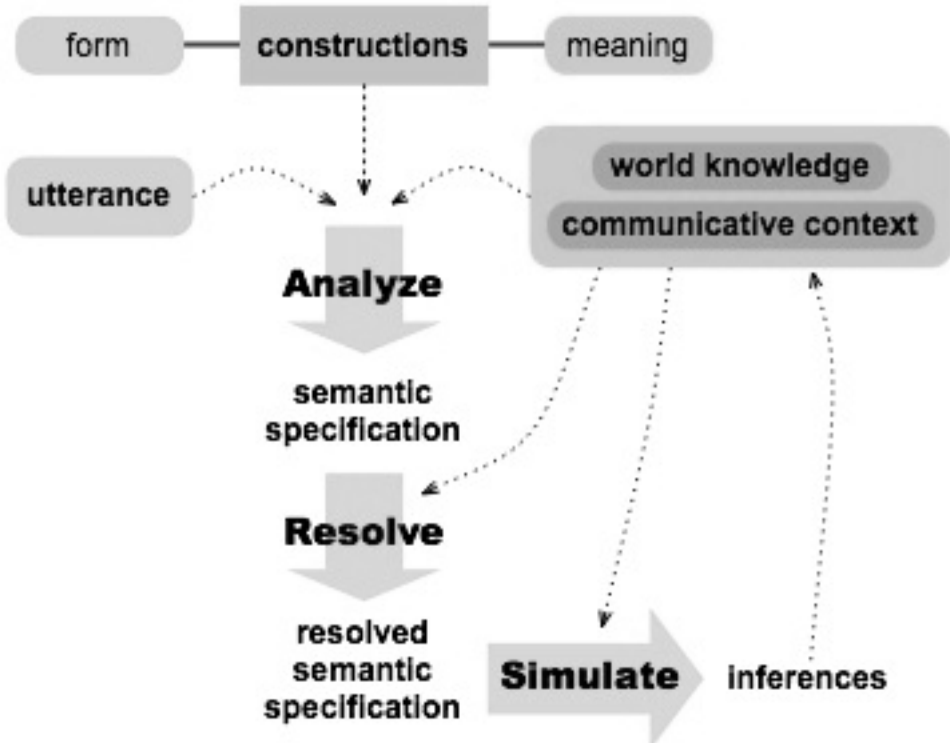


Figure 4. Overview of structures and processes involved in simulation-based language comprehension.

It is important to note that the linear depiction of the three processes is a simplification, and in processing a given utterance may take place at least partly in parallel. That is, partial utterances may be analyzed, resolved and simulated even before the entire utterance is recognized, and results of these processes may feed back into the ongoing analysis.
The figure also identifies several kinds of information comprehenders bring to bear in understanding an utterance:

- Linguistic knowledge: what they know about their language, i.e., a set of lexical and grammatical constructions, each a structured association between knowledge about form (such as phonological schemas) and knowledge about meaning (such as conceptual schemas)
- The utterance itself: what they perceive, i.e., the written/spoken/gestured signal
- The communicative context: both the physical situation and the preceding discourse they've produced and apprehended
- World knowledge: what they know about the entities and eventualities of the world

Essentially, the analysis process is the construction-based analogue to traditional syntactic parsing, except that all structures, including the input set of constructions (or "constructicon") and the output semspec, include meaning along with form. Indeed, this analogy may be evident in Figure 3: the structures in the center column capture constructional constituency relations and thus correspond to a syntactic parse tree. (We call this collection of structures an "analysis graph" or "construct graph".) Contextual resolution slightly broadens the traditional notion of reference resolution to apply to any structures in the semspec, grounding meaning in a specific context. Both of these first two processes can be seen as constraint-based searches: a comprehender perceiving an utterance must find the subset of her constructicon that best explains what she's perceiving—that is, it should satisfy all relevant constructional constraints, account for the forms of the utterance and yield a coherent semspec—and then find the particular mapping (or "resolution") of the semspec to the current context that makes the most sense (broadly speaking, in which the structures in the semspec are most easily identifiable). By contrast, the simulation process is not one of search but involves dynamic activation of structures and potentially unbounded inferential reasoning based on semantically and pragmatically rich sources of information; the resulting inferences corresponds to the broadest and deepest sense of the "meaning" that the comprehender constructs for the utterance.[4]

## 4.3 Solving the comprehension puzzle: basic requirements and psycholinguistic constraints

In the discussion so far, we have outlined the ECG view of the comprehension process. But both the analysis and resolution processes are in fact a good deal more complex than we've suggested. Analysis, for instance, applies constraints to determine which particular constructions and schemas are relevant to a given utterance, and how they are related. Here we might draw an (imperfect) analogy to solving a jigsaw puzzle (or one region of a puzzle), where constructions are like individual puzzle pieces that fit together in specific ways, according to the (relatively hard) constraints imposed by their specific shapes, as well as the (often softer) constraints imposed by features of their respective images. These constraints apply both locally (between neighboring pieces) and to progressively larger sets of pieces. From this perspective, Figure 3 shows a solution to the puzzle that satisfies the various applicable constraints, but it doesn't show precisely *how* a comprehender might arrive at that solution.

There are several minimal capabilities required for a model of language comprehension:

- Given an input utterance and a set of constructions, the model must select a subset of constructions that can cover the input.
- Given such a subset of constructions, it must compose these constructions into a coherent construct graph.
- Given such construct graph, the model must verify that (1) the utterance forms are accounted for; and (2) the associated semspec makes sense in context.

A model equipped with the above minimal requirements could, we claim, perform the basic task of activating a collection of constructions to be used in building a construct graph that covers the input string and makes sense. It could also, incidentally, be used to determine whether the input grammar can account for a given input utterance.

But this is only a simplified version of the story. When human language users apprehend utterances, they exploit a number of heuristics for prioritizing search and pruning irrelevant constructions from consideration. These heuristics work both bottom-up and top-down. For one,

[4] This view of comprehension is consistent with that presented in Bergen & Chang (2005), though it draws a conceptual distinction between the processes of contextual resolution and constructional analysis. It is important to note, however, that the linear depiction of the three processes is a simplification, and in processing a given utterance may take place at least partly in parallel. That is, partial utterances may be analyzed, resolved and simulated even before the entire utterance is recognized, and results of these processes may feed back into the ongoing analysis.

words whose concrete forms are present in the utterance play the most direct role in triggering constructional selection. At the same time, conceptual and distributional expectations cause some constructions to be primed or prioritized for earlier consideration. And these processes interact. For example, the recognition in the input of instances of one or more constructions that can act as constituents of a construction may facilitate the recognition of any remaining constituents and of the whole construction.

Human language processors also have to deal with substantial inherent uncertainty, ambiguity and noise present during all stages of language comprehension. For any apprehended utterance, there may be multiple possible constructional analyses and multiple ways of resolving referents to the context. There may also be errors in the perceived input utterance or communicative context. Moreover, some inferences may be only probabilistically licensed. And in many cases (notably in modeling child language learning), the input may simply not be covered by the current grammar. In general, all the capabilities identified in this section are more realistically seen as producing outputs that are not categorical but stochastic; interpretations can be judged as relatively more or less complete, coherent and consistent with the context. A more realistic model of comprehension thus requires the means of combining partial and uncertain information and evaluating candidate structures at every stage to choose those that best contribute to effective and efficient comprehension (e.g., by maximizing utterance interpretability in context, or minimizing constructional and contextual ambiguity).

Last but far from least, studies of human language processing provide strong evidence that the comprehension process is incremental: it begins as soon as the hearer begins to perceive an utterance, and even before a given utterance (or clause, or phrase) is completed, results of intermediate processing can affect the processing of the entire utterance (Altmann & Steedman, 1988).

## 4.4 A cognitively plausible model of constructional analysis

We now briefly summarize an implementation of constructional analysis within the ECG framework that satisfies the basic constraints of comprehension while addressing some complications of cognitively plausible language processing. The construction analyzer described and implemented by Bryant (2008) uses unification as the basic mechanism for composing constructions and verifying that their constraints are consistent, where both constructions and schemas are represented as typed feature structures with unification constraints as specified by the ECG formalism. But the search for the best analysis also exploits many heuristics to improve efficiency, limit search and approximate the uncertain, incremental, robust nature of human language processing, including:

- Incremental interpretation: the analyzer allows incremental left-to-right interpretation of the utterance. To do this, it employs left-corner parsing techniques (Manning & Carpenter, 1997) to keep track of the current set of competing analyses and update their scores, where partially matched subportions of complex constructions provide top-down expectations about which constructions may next be encountered.
- Best-fit interpretation: the analyzer defines a quantitative heuristic for combining information from disparate domains and ranking candidate interpretations (and thus guide parsing decisions). The technical implementation is a Bayesian probabilistic model that integrates any available information affecting the likelihood of the analysis, drawn where possible from corpus data. Such information includes, for example, lexical and

constructional frequencies; the likelihood that one construction has another as a constituent; and the likelihood that a given schema has a particular kind of filler in a given role.

- Partial interpretation: the analyzer produces partial analyses even when the input utterance is not covered by the grammar, or when parts of the utterance have been omitted. An extension to the analyzer permits analyses with omitted constituents (as often encountered in, for example, Mandarin) by integrating the score of an interpretation with the results of the contextual resolution process.

In sum, the analyzer is consistent with the constructional view, drawing on all available information at every step to ensure that syntactic, semantic and constructional constraints are satisfied. Crucially, the early incorporation of semantic, pragmatic and statistical constraints can dramatically reduce the search space that may result from purely syntactic approaches.

The constructional analyzer has been applied to a variety of linguistic phenomena, including modeling families of related argument structure constructions (Dodge, 2010), early Mandarin constructions (Mok, 2008) and Hebrew morphological constructions (Schneider, 2010). Besides serving as a platform for linguistic analysis, it has also been applied as a psycholinguistic model of reading time data (Bryant, 2008), and versions of the analyzer have been integrated in models of child language acquisition (Chang, 2008; Mok, 2008). Ongoing research has integrated ECG representations of mental spaces and metaphor into the constructional analysis process (Feldman & Gilardi, In prep.), similar to earlier proposals (Bergen & Chang, 2005; Chang et al., 2002; Mok et al., 2004).

## 5. Conclusion and future directions

The discovery that language comprehension uses mental simulation has dramatic ramifications for theories of language knowledge and use. In this chapter, we have laid out a construction-based approach to grammar consistent with this finding. Embodied Construction Grammar provides both formal notations and computational models designed to support the exploration and validation (or falsification) of such simulation-based approaches, while also accounting for the kinds of phenomena addressed by constructional approaches more generally. We have focused in this chapter on giving a broad overview of how constructions, both lexical and grammatical, can contribute to and parameterize embodied simulations based on perceptual, motor and other cognitively and neurally grounded structures. The structures and notations used in the ECG formalism are more than just a way of expressing linguistic analyses; they are designed to support processes of language use, as illustrated by the model of language comprehension described above.

There remain, however, a host of details that this overview has elided, in the hope that readers inclined to dig deeper will do so. As forewarned, we have omitted detailed specification of the ECG formalism, or any algorithmic or implementational detail of the computational models associated with language comprehension. A major strand of related research has focused on theories and computational models of child language learning (Chang, 2008; Chang & Gurevich, 2004; Chang & Mok, 2006; Mok, 2008); these have shown how a construction-based grammar formalism, a model of (partial) language comprehension like that described above, and usage-based learning strategies can be integrated into a cognitively and developmentally plausible model of how children learn their earliest lexical and grammatical constructions. Also, the x-schema formalism and simulation-based semantics was originally motivated by

Narayanan's (1997) model of linguistic aspect and conceptual metaphor (Lakoff & Johnson, 1980); more recent research has been extending the ECG formalism and implementations to accommodate such cross-domain mappings.

All of the above provide encouraging evidence of Embodied Construction Grammar's potential for supporting theories of language structure, learning, and use. Many areas need more attention, of course; most notably, a model of language production would provide a complementary and crucial component of language use. But we hope the basic outlook laid out by investigations to date have demonstrated the benefits (and challenges) of taking seriously the enterprise of understanding grammar as part of cognitive science. This commitment binds us to investigate much more than just the possible nature of representation of linguistic structures: we must also ask how they are used, and how they interact with other cognitive and neural structures. Only by asking—and answering—the right questions can the study of grammar contribute to a broader cognitive science of language.

## Acknowledgments

The authors would like to express their gratitude to researchers in the Neural Theory of Language project at UC Berkeley/International Computer Science Institute for their substantial contributions to this work, as well as to an anonymous reviewer for insightful comments, and to colleagues at the Sony Computer Science Laboratory Paris for helpful discussion.

## References

Altmann, G. T. M., and Steedman, M. J. (1988). 'Interaction with context during human sentence processing'. *Cognition*, 30(3), 191-238.

Bailey, David. (1997). *A Computational Model of Embodiment in the Acquisition of Action Verbs*. Doctoral dissertation, Computer Science Division, EECS Department, University of California, Berkeley.

Barsalou, L. (1999) 'Perceptual symbol systems'. Behavioral and Brain Sciences 22, 577–660.

Bergen, Benjamin, Ting-Ting Chan Lau, Shweta Narayan, Diana Stojanovic, and Kathryn Wheeler. (2010). 'Body part representations in verbal semantics'. *Memory and Cognition* 38(7):969-981.

Bergen, B. & Wheeler, K. (2005). 'Sentence Understanding Engages Motor Processes'. In Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society.

Bergen, B. & Wheeler, K. (2010). 'Grammatical Aspect and Mental Simulation'. Brain and Language.

Brunye, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). 'When you and I share perspectives: Pronouns modulate perspective-taking during narrative comprehension'. Psychological Science, 20, 27-32.

Bryant, J. (2004). 'Scalable Construction Based Parsing and Semantic Analysis'. In proceedings of the workshop on Scalable Natural Language Understanding.

Bryant, J. (2008). Best-Fit Constructional Analysis. Ph.D. Dissertation. Computer Science Department, University of California, Berkeley.

Bybee, Joan. (2006). 'From usage to grammar: the mind's response to repetition'. *Language* 82: 711-733.

Chang, N. (2008). *Constructing grammar: A computational model of the emergence of early constructions.* Computer Science Division, University of California at Berkeley dissertation.

Chang, N., Feldman, J., Porzel, R. & Sanders, K. (2002). 'Scaling cognitive linguistics: Formalisms for language understanding'. In Proc. 1st International Workshop on Scalable Natural Language Understanding, Heidelberg, Germany.

Chang, N. and Gurevich, O. (2004). 'Context-driven construction learning'. Proceedings of the 26th Annual Meeting of the Cognitive Science Society. Chicago, IL.

Chang, N. and Mok, E. (2006). 'A Structured Context Model for Grammar Learning'. *Proceedings of the 2006 International Joint Conference on Neural Networks.* Vancouver, BC.

Choe, J & Bergen, B. (In prep) 'Motor simulation is suppressed by the passive voice'.

Connell L. (2007). 'Representing object colour in language comprehension'. *Cognition*, 102, 476-485.

Dahaene, Stanislaus. (2009). *Reading in the Brain*. New York: Penguin.

Dodge, E. (2010). *Constructional and Conceptual Composition*. Linguistics Department, University of California at Berkeley dissertation.

Ehrsson, H.H., Geyer, S., and Naito, E. (2003). 'Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part specific motor representations'. J. Neurophysiol. 90: 3304-3316.

Feldman, J. (2006). *From Molecule to Metaphor*. Cambridge, MA: The MIT Press.

Feldman, J. & S. Narayanan. (2004). 'Embodied Meaning in a Neural Theory of Language'. Brain and Language 89 (2004), 385-392.

Glenberg, A. M., & Kaschak, M. P. (2002). 'Grounding language in action'. Psychonomic Bulletin & Review.

Glenberg, A. & Robertson, D. (2000). 'Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning'. *Journal of Memory & Language* 43(3), 379-401.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hostetter, A. B. & Alibali, M. W. (2008). 'Visible embodiment: Gestures as simulated action'. Psychonomic Bulletin and Review.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). 'Neural foundations of imagery'. *Nature Reviews Neuroscience, 2*, 635-642.

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

Lakoff, G. and M. Johnson. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Madden, C.J. & Zwaan, R.A. (2003). 'How does verb aspect constrain event representations?' *Memory & Cognition*, 31, 663-672.

Manning, C. & Carpenter, B. (1997). 'Probabilistic parsing using left-corner language models'. In *Proceedings of the 5th International Workshop on Parsing Technology.*

Marr, D. (1982). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman and Company.

Masson, M. E. J., Bub, D. N., & Warren, C. M. (2008). 'Kicking calculators: Contribution of embodied representations to sentence comprehension'. Journal of Memory and Language, 59, 256-265.

Mok, E. (2008). *Contextual Bootstrapping for Grammar Learning.* Ph.D. Dissertation. Computer Science Department, University of California, Berkeley.

Mok, E. & Bryant, J. (2006). 'A Best-Fit Approach to Productive Omission of Arguments'. In Proceedings of the Berkeley Linguistics Society.

Mok, E., Bryant, J., & Feldman, J. (2004). 'Scaling Understanding up to Mental Spaces'. In Proceedings of the Workshop on Scalable Natural Language Understanding.

Narayanan, S. (1997.) *Knowledge-based Action Representations for Metaphor and Aspect (KARMA)*. Computer Science Division, University of California at Berkeley dissertation.

Nyberg, L., Petersson, K.-M., Nilsson, L.-G., Sandblom, J., Åberg, C., & Ingvar, M. (2001). 'Reactivation of motor brain areas during explicit memory for actions'. *NeuroImage*, 14, 521-528.

Pulvermüller, F., Hauk, O., Nikulin, V. & Ilmoniemi, R.J. (2005). 'Functional links between motor and language systems'. *European Journal of Neuroscience*, 21 (3), 793-797.

Sato, M. (2010). *Message in the "Body": Effects of Simulation in Sentence Production*. Ph.D. Dissertation. Linguistics Department, University of Hawaii, Manoa.

Schneider, N. (2010). 'Computational cognitive morphosemantics: Modeling morphological compositionality in Hebrew verbs with Embodied Construction Grammar'. In proceedings of the Berkeley Linguistics Society.

Stanfield, R.A. & Zwaan, R.A. (2001). 'The effect of implied orientation derived from verbal context on picture recognition'. *Psych Science*, 12, 153-156.

Winter, B., & Bergen, B. (Submitted). 'Language comprehenders represent object distance both visually and auditorily: Evidence for the immersed experiencer view'.

Yaxley, R.H. & Zwaan, R.A. (2007). 'Simulating visibility during language comprehension'. *Cognition*, 150, 229-236.

Zwaan, R.A., Stanfield, R.A., Yaxley, R.H. (2002). 'Language comprehenders routinely represent the shapes of objects?' *Psychological Science*, 13, 168-171.