Spatial Schematicity of Prepositions in Neural Grammar

Benjamin K. Bergen and Nancy C. Chang
University of California at Berkeley and International Computer Science Institute
{bbergen|nchang}@icsi.berkeley.edu

1. INTRODUCTION

The semantic content of prepositions and other spatial relations terms has been the object of a great deal of cognitive linguistic investigation over the past two decades (Brugman 1981, Lindner 1981, Casad and Langacker 1985, Lakoff 1987, i.a.). The main foci of these investigations have been the precise senses of spatial relations terms and how those senses are related. An exciting result from this line of research has been the hypothesis that multiple, related senses of words have distinct but related image-schematic content. Image schemas as described by Johnson (1987) and Lakoff (1987) are abstractions over relatively simple perceptual and motor structures that recur in everyday bodily experience. At present, very little has been said about how to represent image-schematic content, or how to model the combination of different image-schematic representations to produce a unified understanding of a given construct, e.g. a sentence.

This paper proposes some formal mechanisms required for a concrete model of image schemas. We outline a computationally implementable formalism for linguistic units, including prepositions, nominals, verbs, and larger constructions, and show how image schemas fit into such a formalism. This formalism has been developed as part of the Neural Theory of Language Project (Feldman et al. 1996, Chang and Bergen To appear), which, among other things, endeavors to model language understanding as a process through which linguistic input is converted into a mental simulation of a described scene or other intended meaning. We briefly describe this process, showing how a set of linguistic constructions with semantic correlates is assembled into a unified situation to be simulated. As an argument for the necessity of simulation for understanding, we demonstrate that for particular combinations of image schemas, incoherencies and idiosyncracies can be detected only by putting to use the embodied motor- and image-schematic knowledge underlying image schemas, not just their abstract schematizations.

2. MODELING IMAGE SCHEMAS

Crucial details of how a sentence is construed, including the scope of the action and the granularity at which the landmark is viewed, can hinge on the choice of preposition:

(1a) Harry walked to the cafe.
(1b) Harry walked into the cafe.

(1a) and (1b) elicit different interpretations with respect to the walking action asserted. In (1a), the entire cafe serves as the end point, or goal, of the trajectory, while it is some more specific point in the interior of the cafe that plays this role in (1b). Inferences about the action's starting point also exhibit a subtle difference in focus, where (1a) suggests a source location away from the cafe and (1b) merely requires that the source be some point exterior to the cafe.

The observed pattern of inferences can be explained in terms of the different image-schematic contributions of the two prepositions. Whereas *into* (at least in its central sense) evokes both the CONTAINER and the SOURCE-PATH-GOAL (SPG) image schemas, *to* evokes only the SPG schema (Johnson 1987, Lakoff 1987). These distinct image-schematic structures combine with the semantic content of the rest of the sentence is quite different ways. Stated more precisely, they involve different **bindings** between the abstract entities that participate in each image schema (or image-schematic **roles**) and the the remaining sentential elements. Some bindings appear in both (1a) and (1b): the subject (*Harry*) must be bound to the Trajector required by the SPG schema, and, by the same token, the prepositional object (*the cafe*) must be bound to the Goal of the SPG trajectory. But in (1b), *the cafe* must additionally be bound to the Container of the CONTAINER schema, which, as noted above, affects the precise specification of the SPG Goal. In the remainder of this section, we demonstrate a means for representing image schemas so as to facilitate the specification of lexical items that involve image-schematic content. We also show how these representations fit together to produce a structure that can be used to simulate the meaning of an utterance.

A representation for image schemas

Our representation assumes that the linguistically relevant properties of a given image schema can be identified and named; these properties are precisely the image-schematic roles mentioned above. Although image schemas have often been likewise characterized in terms of such roles (also called components or elements), it is crucial to note that these roles are abstractions over individual perceptual experiences, and that a full representation of image schemas must at some level involve representations based on the perceptual system (cf. Regier 1996). That is, although these roles can be represented in symbolic terms, this symbolic representation serves only to parameterize, and not to replace, the perceptual properties of the schema in question.

For example, in the case of the CONTAINER schema, the content is principally visual: enclosure and boundedness are primary visual properties that are frequently experienced together. The full image schema also includes haptic content, since the body is construed as a container, and since humans physically manipulate containers. Nevertheless, certain aspects of the CONTAINER schema that have important linguistic consequences can be distinguished as the image-schema's abstract roles; these are the Interior, Exterior, Boundary, Container, Contents, and Portal. A simple attribute-value matrix representation of these image-schematic roles is shown in Figure 1a, where the variables following each role denote possible **fillers** (or instantiations). The SPG schema can be represented in the same manner, with roles Source, Path, Goal, and Trajector, as shown in Figure 1b.

Schema	**CONTAINER**
Interior	i
Exterior	e
Boundary	b
Container	c
Contents	n
Portal	o

(a)

Schema	**SPG**
Source	s
Path	p
Goal	g
Trajector	t

(b)

Figure 1. Representations for the (a) CONTAINER and (b) SPG image schemas, showing image-schematic roles and fillers (here as uninstantiated variables) in an attribute-value matrix.

The lexical representations we discuss in the next section exemplify how linguistic units can make reference to these simple image-schematic representations and their roles, binding the roles both internally and externally in different ways.

Lexical representations

As mentioned earlier, the English prepositions *to* and *into* as used in (1a) and (1b) differ in which image schemas they evoke and how the image-schematic roles are bound. Figure 2 below shows lexical representations for each of these words that include this information, along with structures containing more familiar form and meaning information. In both cases, the line beginning with the word "*with*" indicates that the concept is **dependent** in the sense of Langacker (1991); both a trajector and landmark are necessary for the relevant senses of each word and must eventually be bound to other elements in the sentence. The next line, with "*uses*", indicates which image schemas are used in the meaning pole of the word. All of these mentioned entities (Trajector, Landmark) and image schemas (SPG, CONTAINER) appear with variables that allow them to be referred to elsewhere, often in binding constraints. In Figure 1a, for example, the SPG schema is referred to as s, and its roles (referred to as s.Source, s.Goal and s.Trajector) are bound to the Trajector and Landmark (or, in the case of lm.Away, an appropriate related location).

to
with **Trajector** tr, **Landmark** lm
uses **SPG** s

Form : Phon = $[t^h u^w]$
Orth = "to"

Meaning : s.Source = lm.Away
s.Goal = lm
s.Trajector = tr

(a)

into
with **Trajector** tr, **Landmark** lm
uses **SPG** s, **CONTAINER** c

Form : Phon = $[Int^h u^w]$
Orth = "into"

Meaning : s.Source = c.Exterior
s.Goal = c.Interior
s.Trajector = tr
c.Container = lm

(b)

Figure 2. Representations for (a) *to* and (b) *into.*

In using not just the SPG schema but also the CONTAINER schema, (1b) involves image-schematic role bindings not necessary in (1a). Specifically, *into* requires that the Exterior of the CONTAINER schema be bound to the Source of the SPG schema and that the Interior of the CONTAINER schema be identified with the Goal of the SPG schema. These bindings are shown in the lexical representation for *into* in Figure 2b. As in 2a, the Trajector and Landmark must be provided elsewhere in the utterance or context. We will show how these other semantic entities are bound through a higher-level clausal construction in the next section.

Other linguistic units can also provide image-schematic specifications. Nouns bound to the Goal of a SPG schema must be construed as locations, and only a subset of English nouns can be so characterized. As in Cognitive Grammar (Langacker 1991), we posit a semantic type hierarchy allowing multiple inheritance, in which a *cafe*, for example, is a physical location, as well as a possible container. This type of information is specified in the line beginning "*denotes*" in the construction for *cafe* (Figure 3a). (Other complex aspects of the meaning of *cafe* have been omitted and are shown here as simply CAFE.) Verbs also typically involve the coordination of image-schematic information with features of a particular action or event. In the case of *walked*, for example, we represent the complex motor pattern for walking as simply WALK and bind this value to the Schema role of the CONTROLLER, a structure that summarizes the motor control parameters for an action (Narayanan 1997). This structure is linked to the FORCE-DYNAMIC (or FORCE) image schema in such a way that the entity performing the action is the Energy-source of the FORCE schema, which is in turn bound to the Trajector of the SPG. In other words, *walked* enforces bindings between some entity's motor control, force expenditure and movement, a pattern typical for verbs of self-propelled motion. As shown in Figure 3b, the *walked* construction includes additional information about the temporal grounding of the event with respect to speech time.

Noun cafe
denotes **Location**, **Container**

Form: Phon = [kæfej]
Orth = "cafe"

Meaning: CAFE

(a)

Verb walked
with **Trajector** tr
uses **SPG** s, **FORCE** f

Form: Phon = [wakt]
Orth = " walked"
before(tr, m)

Meaning: Controller.Schema = WALK
f.Energy – source = tr
s.Trajector = tr
Ground.Time = past

(b)

Figure 3. Representations for (a) *cafe* and (b) *walked.*

Clausal constructions and the simulation specification

The SPG's Trajector can, for *into*, be bound to either the clause's subject or its object, as seen in (2a) versus (2b) and (2c):

(2a) Harry marched into the cafe.
(2b) Harry sent the youngsters into the cafe.
(2c) Harry marched the youngsters into the cafe.

Although verbs behave fairly consistently with respect to the binding between the Trajector role of the SPG and the grammatical subject or object, many verbs allow both possibilities. The choice of binding thus depends on the interaction between the verb and the larger clausal construction in which it appears (Goldberg 1995). The exact form of these constructions and how they fit together with verbs is beyond the scope of the present paper (but cf. Chang and Bergen To appear). All that is presently relevant is that clausal constructions may also have image-schematic content that must cohere with that of its constituent parts. For example, the sentences in (1) are both instances of the directed-motion construction, which is shown in Figure 4.

directed-motion(**Mover** m, **Motion** v, **Direction** d)
uses **SPG** s
denotes **MotionEvent** e

Form: order(m,v,d)
subject - agree(m,v)

Meaning: e = v
e.SPG = s = d
e.SPG.Trajector = m

Figure 4. Directed-motion construction.

For current purposes it is sufficient to observe that this construction also uses the SPG schema, and its meaning constraints specify how its constituents (listed within parentheses on the first line) fit together: much of the semantic content of the motion event denoted by the clause (referred to in the construction as e) derives from that of its Motion constituent (as indicated by the reference variable v, and bound in our example to *walked*); the SPG of the overall clause is bound to the SPG of the Direction constituent d (*in/into the cafe*); and finally, the Trajector of this SPG is bound to the Mover constituent m (*Harry*).

Constructions and lexical items represented in this way – that is, using form and meaning structures that make reference to common image-schematic components – are the basis for a model of sentence comprehension in which meaning arises from **simulation**; experientially and bodily grounded structures that are useful for performing actions can also be used to understand utterances about those actions (Bailey 1997). The analysis and binding process that matches a set of constructions and lexical items to a sentence culminates in the production of a **simulation specification**, which is a feature structure description of the scene or set of scenes to be simulated or imagined. The simulation specifications for the sentences in (1) are shown in Figure 5. These include both image-schematic descriptions (which resemble the image schema representations shown in Figure 1 but are instantiated appropriately) and other necessary simulation structures, such as the GROUND structure that locates the event (here only temporally) and the CONTROLLER structure that provides crucial motor control information.

SIMULATION SPECIFICATION

SPG	Source	CAFE.Away
	Goal	CAFE
	Trajector	HARRY
FORCE	[Energy-source	HARRY]
GROUND	[Time	past]
CONTROLLER	[Schema	WALK]

(a)

SIMULATION SPECIFICATION

SPG	Source	CAFE.Outside
	Goal	CAFE.Inside
	Trajector	HARRY
CONTAINER	Interior	CAFE.Inside
	Exterior	CAFE.Outside
	Container	CAFE
FORCE	[Energy-source	HARRY]
GROUND	[Time	past]
CONTROLLER	[Schema	WALK]

(b)

Figure 5. Simulation specifications for (a) *Harry walked into the cafe* and (b) *Harry walked into the cafe*. (Uninstantiated image-schematic roles not relevant for these simulations have been omitted.)

It should be clear that the simulation specification includes exactly the schematic content of the different elements of the sentence, bound appropriately. As noted earlier, the two representations differ with respect to which image schemas are involved – as reflected by the additional CONTAINER schema in Figure 5b – and in the precise bindings of aspects of the cafe to the SPG schema. Like the image schema representations, the simulation specifications can be viewed as a summary of the much more complex structures that are active when an event is simulated or imagined. Activating these structures – that is, "running" the simulation – can thus provide the much richer basis for inference necessary for accounting for many linguistic phenomena.

3. SIMULATION-BASED INFERENCE

When image-schematic content is combined in simulation, detailed inferences about meaning emerge that may not have been provided in the simulation specification. In this section, we show how several kinds of inference become greatly simplified through the use of simulation.

Two differences between the sentences in (3) are (i) the final location of the Trajector with respect to the Container and (ii) the parts of the *house* that correspond to the Portal of the CONTAINER schema evoked by *into*.

(3a) The preacher drifted into the house.
(3b) The smoke drifted into the house.

The detailed relation between the SPG's Trajector and the CONTAINER's Interior, while not necessary in the selection between parses of the sentences, is part of the inferential content of their simulations. In (3a), the preacher's final location is within the Interior of the house, while in (3b), the smoke may have permeated the entire Interior of the house. Such knowledge of the physical properties of smoke versus preachers accounts for the strangeness of (4a) but not (4b):

(4a) ?The preacher drifted into the house and filled it.
(4b) The smoke drifted into the house and filled it.

Similarly, because of our detailed knowledge about how people interact with houses, we can guess that sentence (3a) involves a door: *drift* implies ease and slow pace, and the portal of a house that best provides these properties for people is a door. On the other hand, we know that smoke can travel just as easily through windows as through doors, so we are likely to imagine multiple and/or varied portals in (3b). This detail is not relevant in selecting the correct sense of *into,* but it is clearly necessary for further inferencing: while (5a) seems like an odd piece of reasoning, (5b) is perfectly felicitous:

(5a) ?The preacher drifted into the house because the window had been left open.
(5b) The smoke drifted into the house because the window had been left open.

Note that the word *drift* might be considered ambiguous between a more physical floating sense and a more abstract aimless-attitude sense, where the difficulty of simulating preachers with the former and smoke with the latter helps in the selection of the appropriate sense. This kind of disambiguation through simulation may have widespread applicability. For example, the sentences in (6) involve two senses of *into*: (6a) uses the central sense used in the previous examples, while (6b) seems more likely to evoke a difference sense of *into,* one involving CONTACT with an obstacle. In both cases, the construal of the prepositional object as a container is crucial to its binding with the Container role of (central) *into*'s CONTAINER image schema. While a *laboratory,* as a canonical location, is quite easily construed as such, a *wall* can only be so construed given the right context, leading to another (more likely) reading with the alternate sense.

(6a) The scientist walked into the laboratory.
(6b) The scientist walked into the wall.

4. CONCLUSION

We have briefly discussed structures and processes that we have argued are necessary for modeling language understanding. In this model, the interpretation of a sentence depends on how the image-schematic content of prepositions interacts with that of nominals, verbs and larger constructions. All of these image-schematic structures must in turn be coordinated with motor control and other aspects of world knowledge to produce a structure that can drive a mental simulation. We have additionally argued that richly detailed inferences crucial for explaining a variety of linguistic phenomena can only arise as a result of such a simulation.

In this paper, we have focused on inference and disambiguation with respect to only literal meanings, but the model we have described can easily be extended to metaphorical meanings as well. Decisions about which sense of a polysemous word is appropriate must sometimes be mediated by metaphors (Lakoff and Johnson 1980). In a simulation-based model, an implausible literal simulation may give rise to the activation of a related metaphor, especially if the metaphor's target domain is also activated by contextual elements. Metaphorical mappings can then license a more plausible literal simulation in the source domain and allow the resulting fine-grained inferences to be given appropriate metaphorical interpretations (Narayanan 1997). The addition of such mappings to our image-schematic representations would enable the model described to ground the simulation of both literal and metaphorical language.

REFERENCES

Bailey, David. 1997. *When Push Comes to Shove : A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs*. Unpublished Ph.D. Dissertation, Computer Science Division, UC Berkeley.

Brugman, Claudia. 1981. *Story of Over*. M.A. Thesis. UC Berkeley.

Casad, Eugene and Ronald Langacker. 1985. `Inside' and `Outside' in Cora Grammar. *International Journal of American Linguistics* 51.3:247-81.

Chang, Nancy and Benjamin K. Bergen. To appear. Constructions as structured dependencies and constraints in a simulation-based framework. ICSI Technical Report.

Feldman, J, G. Lakoff, D. Bailey, S. Narayanan, T. Regier, A. Stolcke (1996). Lzero: The First Five Years. *Artificial Intelligence Review*, v10 103-129, April 1996.

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

Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
Lakoff, George. 1993. The contemporary theory of metaphor, In Andrew Ortony (Ed.) *Metaphor and Thought*, 2nd Edition. Cambridge: Cambridge University Press.
Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
Langacker, Ronald. 1991. *Concept, Image, and Symbol*. New York: Mouton de Gruyter.
Lindner, Susan. 1981. *A lexico-semantic analysis of English verb-particle constructions with UP and OUT*. Unpublished Ph.D. Dissertation, UCSD.
Narayanan, Srini. 1997. *KARMA: Knowledge-Based Active Representations For Metaphor and Aspect.* Ph.D. Dissertation, Computer Science Division, University of California, Berkeley, 1997.
Regier, Terry. 1996. *The Human Semantic Potential*. Cambridge: MIT Press.