

From Molecule to Metaphor: A Neural Theory of Language

Jerome A. Feldman

(University of California, Berkeley)

Cambridge, MA: The MIT Press (A Bradford book), 2006, xx+357 pp; hardbound, ISBN 0-262-06253-4, \$36.00

Reviewed by

Stefan Frank

Radboud University Nijmegen

Over the last decade or so, it has become increasingly clear to many cognitive scientists that research into human language (and cognition in general, for that matter) has largely neglected how language and thought are embedded in the body and the world. As argued by, for instance, Clark (1997), cognition is fundamentally **embodied**, that is, it can only be studied in relation to human action, perception, thought, and experience. As Feldman puts it: “Human language and thought are crucially shaped by the properties of our bodies and the structure of our physical and social environment. Language and thought are not best studied as formal mathematics and logic, but as adaptations that enable creatures like us to thrive in a wide range of situations” (p. 7). Although it may seem paradoxical to try formalizing this view in a computational theory of language comprehension, this is exactly what *From Molecule to Metaphor* does. Starting from the assumption that human thought is neural computation, Feldman develops a computational theory that takes the embodied nature of language into account: the *neural theory of language*.

The book comprises 27 short chapters, distributed over nine parts. Part I presents the basic ideas behind embodied language and cognition and explains how the embodiment of language is apparent in the brain: The neural circuits involved in a particular experience or action are, for a large part, the same circuits involved in processing *language* about this experience or action.

Part II discusses neural computation, starting from the molecules that take part in information processing by neurons. This detailed exposition is followed by a description of neuronal networks in the human body, in particular in the brain.

The description of the neural theory of language begins in Part III, where it is explained how localist neural networks, often used as psycholinguistic models, can represent the meaning of concepts. This is done by introducing **triangle nodes** into the network. Each triangle node connects the nodes representing a concept, a role, and a filler—for example, “pea,” “has-color,” and “green.” Such networks are trained by a process called **recruitment learning**, which is described only very informally. This is certainly an interesting idea for combining propositional and connectionist models, but it does leave the reader with a number of questions. For instance, how is the concept distinguished from the filler when they can be interchanged, as in “cats, feed-on, mice” versus “mice, feed-on, cats.” And on a more philosophical note: Where does this leave embodiment? The idea that there exists a node representing the concept “pea,” neurally distinct from its properties and from experiences with peas, seems to introduce abstract and arbitrary symbols. These are quite alien to embodied theories of cognition, which generally assume modal and analogical **perceptual symbols** (Barsalou 1999) or even no symbols at all (Brooks 1991).

In Part IV, Feldman explains how systems can be described at different levels of abstraction. A description at the computational level involves role-filler (or feature-value) structures and rules for manipulating these. The reader is regularly reminded that this is just a higher-level description of a connectionist, neural structure. Nevertheless, it remains unclear how exactly such a computational system can be implemented in a neural network, especially as the models' complexity increases in the later chapters of the book.

Part IV also introduces the notion of **conceptual schemas**. These are cognitive structures that emerge from the organization of our bodies and of the world, and are therefore universal across cultures. Examples are basic concepts such as "grasp" and "support." Although the conceptual schemas themselves are universal, the mapping between words and schemas differs across languages. For instance, there is no one-to-one translation between the spatial prepositions of different languages. The acquisition of words denoting spatial relations in different languages is simulated by Regier's (1996) connectionist model, which is the first specific model discussed in the book.

Part V discusses the acquisition of words for actions. Knowledge of particular actions is encoded as **execution schemas** for controlling motor behavior. People are not aware of complete actions, but only of particular parameters that can be set in the execution schema—for instance, the action's speed, direction, and duration. Learning the meaning of a verb comes down to associating the word with the corresponding parameters of the execution schema. These ideas are implemented in Bailey's (1997) model of verb learning, presented at the end of this part of the book.

Part VI explains how abstract expressions are understood through metaphor. When a metaphor is used, words in the abstract domain are understood by using words and knowledge from more-concrete domains. **Primary metaphors** are grounded directly in perception and action. For instance, the expression *a warm greeting* makes use of primary metaphor by talking about affection in terms of temperature. **Complex metaphors** are conceptual combinations of primary metaphors. One of the most important complex metaphors is used for describing the structure of events, and combines mappings such as "causes are physical forces," "states are locations," and "changes are movements."

Understanding a sentence, so it is claimed in Part VI, is not akin to logical inference, but comes down to mentally simulating the event or situation described by the sentence (after metaphorically mapping to a more-concrete domain, if necessary). In support of this view, the book gives evidence from brain-imaging experiments but, oddly, ignores compelling behavioral findings by, for instance, Glenberg and Kaschak (2003), Stanfield and Zwaan (2001), and Zwaan, Stanfield, and Yaxley (2002).

In Part VII, this idea of mental simulation is put into computational form by presenting Narayanan's (1997) model of story comprehension. In this model, background knowledge is implemented in temporal belief (Bayes) networks. These perform probabilistic inference by means of a process of parallel multiple constraint satisfaction, which is quite plausible from a neural perspective. Again, metaphorical or abstract language is understood by mapping to a more-concrete domain. As an example, it is shown how the model would process an excerpt from a newspaper article about economics.

It isn't until Part VIII that grammar begins to play a role, reflecting that the neural theory of language relies less on the syntactic structure of sentences than do more traditional theories of language comprehension. According to the proposed theory of grammar (based on Bergen and Chang's [2005] Embodied Construction Grammar), the basic element of linguistic knowledge is the **construction**: a pairing of linguistic form and meaning. By applying linguistic and conceptual knowledge, an utterance (in the context of a situation) is transformed into a network of conceptual schemas that

specifies the meaning of the sentence. “The job of grammar is to specify which semantic schemas are being evoked, how they are parameterized, and how they are linked together in the semantic specification. The complete embodied meaning comes from enacting or imagining the content . . . and always depends on the context in addition to the utterance” (p. 288). This is clarified by a complete and detailed analysis of the simple sentence *Harry strolled into Berkeley*.

The book’s final part presents a model (Chang 2006) that simulates how constructions are learned by associating an utterance (a linguistic form) to its meaning as inferred from the situational context.

To conclude, *From Molecule to Metaphor* is a good introduction to embodied theories of language and computational modeling in cognitive linguistics. It discusses a wide range of issues, many of which were not included in this review. The view that language should not be described as a mathematical or logical system but is foremost a part of human behavior and a function of the brain could well be new and thought-provoking to readers of this journal. However, the same readers are also likely to desire a high level of computational detail which the book does not provide, as it does not present any formal specification of the discussed models.

References

- Bailey, D. 1997. *When Push Comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs*. Ph.D. thesis, University of California, Berkeley.
- Barsalou, L. W. 1999. Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–660.
- Bergen, B. and N. Chang. 2005. Embodied construction grammar in simulation-based language understanding. In J.-O. Östman and M. Fried, editors, *Construction Grammar(s): Cognitive and Cross-Language Dimensions*. John Benjamins.
- Brooks, R. A. 1991. Intelligence without representation. *Artificial Intelligence*, 47:139–159.
- Chang, N. 2006. *Construction Grammar: A Computational Model of the Acquisition of Early Constructions*. Ph.D. thesis, University of California, Berkeley.
- Clark, A. 1997. *Being There*. The MIT Press, Cambridge, MA.
- Glenberg, A. M. and M. P. Kaschak. 2003. The body’s contribution to language. In B. H. Ross, editor, *The Psychology of Learning and Motivation*, volume 43. Elsevier Science, pages 93–126.
- Narayanan, S. 1997. KARMA: *Knowledge-Based Action Representation for Metaphor and Aspect*. Ph.D. thesis, University of California, Berkeley.
- Regier, T. 1996. *The Human Semantic Potential*. MIT Press, Cambridge, MA.
- Stanfield, R. A. and R. A. Zwaan. 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12(2):153–156.
- Zwaan, R. A., R. A. Stanfield, and R. H. Yaxley. 2002. Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2):168–171.

Stefan Frank is a postdoctoral researcher at the Radboud University, Nijmegen. He develops psycholinguistic models of sentence and text comprehension, using recurrent neural networks and other dynamical systems. His address is: Nijmegen Institute for Cognition and Information, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands; e-mail: S.Frank@nici.ru.nl; URL: www.nici.ru.nl/~stefanf.

