Available online at www.sciencedirect.com
SCIENCE @ DIRECT®
Brain and Language 89 (2004) 385–392
Brain and Language
www.elsevier.com/locate/b&l

Embodied meaning in a neural theory of language

Jerome Feldman* and Srinivas Narayanan

International Computer Science Institute, University of California, Berkeley, CA 94720, USA

Accepted 26 August 2003

1. Introduction

Imagine that you are dining with a group of people and one of them asks you to pass the salt. Ignoring for now any social or emotional issues, what does it mean for you to carry out the request? You must locate the salt container, determine which grip to use, reach out and grasp the container, and move it to a location near the requester. If the salt is too far for you to easily reach, you should convey the request to someone who can pass it to you, possibly following a further request. Now think about your favorite theory of language and what it could describe of this scenario.

In this paper, we outline an explicitly neural theory of language (NTL) that attempts to explain how many brain functions (including emotion and social cognition) work together to understand and learn language. The focus will be on the required representations and computations, although there will be some discussion of results on specific brain structures. In this approach, one does not expect to find brain areas specialized only for language or to find language processing confined to only a few areas.

Our first sentence asked you to imagine yourself at some dinner, being asked to pass the salt. To understand this sentence you need to know what it means to pass the salt. The NTL assumption is that people understand narratives by subconsciously imaging (or simulating) the situation being described. When asked to grasp, we enact it. When hearing or reading about grasping, we simulate grasping, being grasped, or watching someone grasp.

We will start with the meaning of words for simple actions like those involved in passing the salt. Consider the word “grasp”. Everyone will agree that the meaning of the word grasp involves the motor action of grasping in some way. The NTL approach to language suggests that the complex synergy that supports grasping *is* the core semantics of the word. We choose this particular example because there is a great deal known about the intricate distributed neural circuitry involved in grasping by monkeys and humans. We will briefly review some of the key findings and then demonstrate their relevance for the NTL theory of meaning.

The action of *grasping* has both a motor component (what you do in grasping) and various perceptual components (what it looks like for someone to grasp and what a graspable object looks like). There are other modalities involved as well, such as the somato-sensory component (what it feels like to be grasp something or to be grasped yourself). Both the meaning of a word and its defining behavior are context dependent - you grasp differently for different objects and purposes.

The theory also entails that the meaning of a noun (e.g. cup) depends on its possible uses or *affordances* (Gibson, 1979). There is both linguistic evidence (from classifier languages) and imaging data (Tettamanti et al., 2002) supporting this idea that the meaning of a noun depends on the uses of the underlying thing.

Mirror neurons in monkeys and their homologs in people (Buccino et al., 2001) suggest an overlapping substrate for the execution of actions and for the perception of the same action. This is a plausible neural basis for the fact that an action word, like *grasp*, denotes grasping, being grasped, or observing grasping.

More generally, there is increasing evidence for the multi-modal neural substrate for actions and action words. Rizzolatti and coworkers, over the last 20 years have shown that frontal area F4 contains neurons that integrate motor, visual, and somato-sensory modalities for the purpose of controlling actions in space and perceiving peri-personal space (the area of space reachable by body parts) (Fogassi, Gallese, Fadiga, & Rizzolatti, 1996a; Fogassi et al., 1996b; Fogassi et al., 1992; Gentilucci et al., 1988; Gentilucci, Scandolara, Pigarev,

---

* Corresponding author. Fax: 510-643-1543.
E-mail address: jfeldman@cs.berkeley.edu (J. Feldman).

doi:10.1016/S0093-934X(03)00355-9

& Rizzolatti, 1983; Rizzolatti & Gentilucci, 1988; Rizzolatti, Fogassi, & Gallese, 2000; Rizzolatti and Gallese, 2003; Rizzolatti, Matelli, & Pavesi, 1983; Rizzolatti, Fadiga, Fogassi, & Gallese, 1997; Rizzolatti, Scandolara, Gentilucci, & Camarda, 1981a; Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981b). Similar results about multisensory integration in area F4 were independently obtained by Graziano, Gross and their coworkers (Graziano & Gross, 1995; Graziano, Hu, & Gross, 1997a, 1997b; Graziano, Yap, & Gross, 1994). More recently, Graziano, Reiss, and Gross (1999) showed that F4 neurons integrate not only visual but also auditory information about the location of objects within peripersonal space.

In summary, over two decades of work in neuroscience to suggests that cortical premotor areas contain neurons that respond to multimodal (visual, somatosensory, and auditory) sensory stimuli. On the other hand, areas that were conventionally considered to process only sensory information like posterior parietal areas, have been found to play a major role in motor control. It appears that premotor and parietal areas are neurally integrated not only to control action, but also to serve the function of constructing an integrated representation of (a) actions together with (b) objects acted on and (c) locations toward which actions are directed. We hypothesize that this complex serves as the neural substrate of the meaning of action words. Pulvermueller (2001) reviews a wide range of evidence for perception/action circuits as the neural basis of word meaning. If we accept this complex of neural circuits and behaviors as the core meaning of grasping, it remains to show how a word like "grasp" gets associated with the embodied concept.

Obviously enough, the perception of a word form, whether through text, speech or sign, is itself an elaborate neural computation. Assuming that the modeling of perception is treated separately, we still need a theory of how the form of an utterance is associated with its meaning. NTL adopts the simple and elegant approach from Construction Grammar (CG) to the form-meaning linking problem. In CG, every linguistic element, from a morpheme to a complete discourse, is considered as a <*form, meaning*> pair. The meaning of a complex utterance is determined by constructional composition of the meanings of its parts. For now we focus on a computer model of how children learn to pair form and meaning for verbs like "grasp". We will discuss the general implications of Embodied Construction Grammar (ECG) (Bergen & Chang, 2002) in the final section.

2. Learning the meaning of simple action verbs

David Bailey (Bailey, 1997) built a computer program that modeled child word learning in the challenging domain of actions. To limit the complexity of the task, Bailey restricted consideration to actions that could be carried out by one hand with objects on a table. The program is modeling a scenario in which a child is doing an action and hearing her parent's (one word) label. The child's (and program's) main task is solving the correlation problem – what features of the situation and of my actions is my parent talking about. The basic idea is that the child must learn to associate her perception of a sound with an action (embodied as described above) that she is currently carrying out.

The problem is that languages differ widely in the way that they label actions. Even for actions involving only one hand, there are quite a lot of verbs in English including:

> seize, snatch, grab, grasp, pick up, hold, grip, clutch, put, place, lay, drop, release, pull, push, shove, yank, slide, flick, tug, nudge, lift, raise, lower, lob, toss, fling, tap, rap, slap, press, poke, punch, rub, shake, pry, turn, flip, rotate, spin, twirl, squeeze, pinch, twist, bounce, stroke, wave, caress, stack, salute, and many, many more...

And that's only English. Other languages make distinctions that English does not. Moreover, each language has its own unique collection of linguistic gaps that reflect differences in the concepts explicitly named. Here a few examples:

- In Tamil, *thallu* and *ilu* correspond to English *push* and *pull*, except that they connote a sudden action as opposed to a smooth continuous force. The continuous reading can be obtained by adding a directional suffix, but there is no way to indicate smooth pushing or pulling in an arbitrary direction.
- In Farsi, *zadan refers* to a large number of object manipulations involving quick motions. The prototypical *zadan* is a hitting action, though it can also mean to snatch (*ghaap zadan*) or to strum a guitar or play any other musical instrument.
- In Cantonese, *meet* covers both pinching and tearing. It connotes forceful manipulation using the two finger posture, but is also acceptable for tearing larger items when two full grasps are used. Cantonese has no distinct word equivalent to *drop*; there is a word meaning *release*, but it applies whether or not the object is supported.
- In Spanish, there are two separate words for different senses of the English verb *push.* The word *pulsar* corresponds to pushing a button and *presionar* covers most of the other uses.

Bailey was faced with the problem of building a program that needed to capture the conceptual differences across languages in order to learn word meanings. Building in too many assumptions would preclude learning some languages and leaving everything unspecified gives the program no chance at all of learning. It will come as no surprise that Bailey's solution on what structure to build into the system was to base it on the body and on neural control networks. The idea is that all people share the same neural circuitry and thus the same semantic potential. Various languages, as we saw above, explicitly label different aspects of a motor or other activity.

But there seems to be a complexity barrier. How could the meaning of an action word be the activity of a vast distributed network of neurons? The key to solving this in the model and, we believe also in the brain, is parameterization. A motor action such as grasping involves many coordinated neural firings, muscle contractions, etc. but we have no awareness of these details. What we can be aware of (and talk about) are certain parameters of the action – force, direction, effector, posture, repetition, etc. The crucial hypothesis is that languages only label the action properties of which we can be aware. That is, there is a fixed set of embodied features that determine the semantic space for any set of concepts, such as motor actions.

Fig. 1 presents an overview of Bailey's model for learning words that describe one-hand actions. The first thing to notice is that there is an intermediate set of features, shown as a large rectangle in the middle of the figure. As discussed above, what we can consciously know about our own actions can be described by a relatively small number of features. People do not have direct access to the elaborate neural networks that coordinate our actions. This parameterization of action is one key to the success of the program. The features chosen as available for language are consistent with known neural parameterizations, but there was no attempt to produce circuit level models of this.

A second critical feature of the model is the schematic representation of actions, called executing schemas (x-schemas), shown at the bottom of Fig. 1. In addition to parameters like force, actions are characterized by control features. For example, some actions are repetitive, some conditional, etc. Depicted in Fig. 1, is a generic control diagram showing an action that is comprised of a set of concurrent actions, both of which have to complete before the next action occurs. This kind of abstract action schema is common in the motor control literature and has also been used effectively in various computational models. The crucial point is that control of action can also be parameterized and thus made available to language learning.

Also notice in Fig. 1 that the arrows are bi-directional. The system not only learns to label actions with words, but will also carry out requests expressed using the words that it has learned. The upward arrows on the left describe the labeling pathway–features are extracted from executing schemas (bottom right arrow) and then these features are used decide which verb is the most appropriate label for the action.

The two arrows on the right capture an ability not present in any system based on supervised weight change learning. This pathway is used by the system to carry out actions that are requested using a verb that has been learned. For example, suppose that the program has learned (as it does) that the word *shove* involves using the slide executing schema with high force and short duration. This information on which schema and parameters define the word *shove* would be stored as part of its definition. When asked to *shove* something, the system would activate the definition and select the appropriate schema and parameters from the large collection available in the middle of the figure. These can then be used to activate the appropriate schema (lower right arrow), here slide; this then leads to the simulated android *Jack* carrying out the requested shoving.

Bailey trained and tested his program extensively in English and more sparsely in several other languages. In the main experiment, he presented the system with 165 labeled examples of actions corresponding to 15 English verbs and 18 word senses. Using learning techniques that are neurally plausible (Shastri, 2002), the program was able to deduce the correct number of words and word senses. The system was then tested by asking it to label 37 novel actions according to the definitions it had learned. The performance was quite good; 80% of the scenes were given exactly the right label. Moreover, all of the errors involved overlapping concepts; for example *move* for *push* or *jerk* for *lift*.

With no further training, the program was then tested for its ability to carry out the actions specified by the words it had learned. The results were quite similar; around 80% of the actions were optimal and the rest were near misses, which of course children also make. So, there is at least a prima facie case that embodied verb meanings can be learned by biologically plausible models.

3. Extension to abstract words and to narrative

There are many words, even for children, that do not directly label immediate experiences. In addition, fairly early in life children start to use the same word in both concrete and metaphorical senses. On common example is when a parent says (with a pained tone) "Now, see what you've done". The child knows that she is supposed to learn (see) as well as perceive her transgression. In the NTL, abstract and metaphorical words derive their meanings from concrete words.

There are several metaphorical uses of the word "grasp" e.g.,

Grasp an idea;
Grasp an opportunity;
Grasp at straws.

Understanding a metaphorical use of a word like grasp is obviously different from physically grasping something. As another example, consider processing a literal use of a word, like when you read (or hear) a

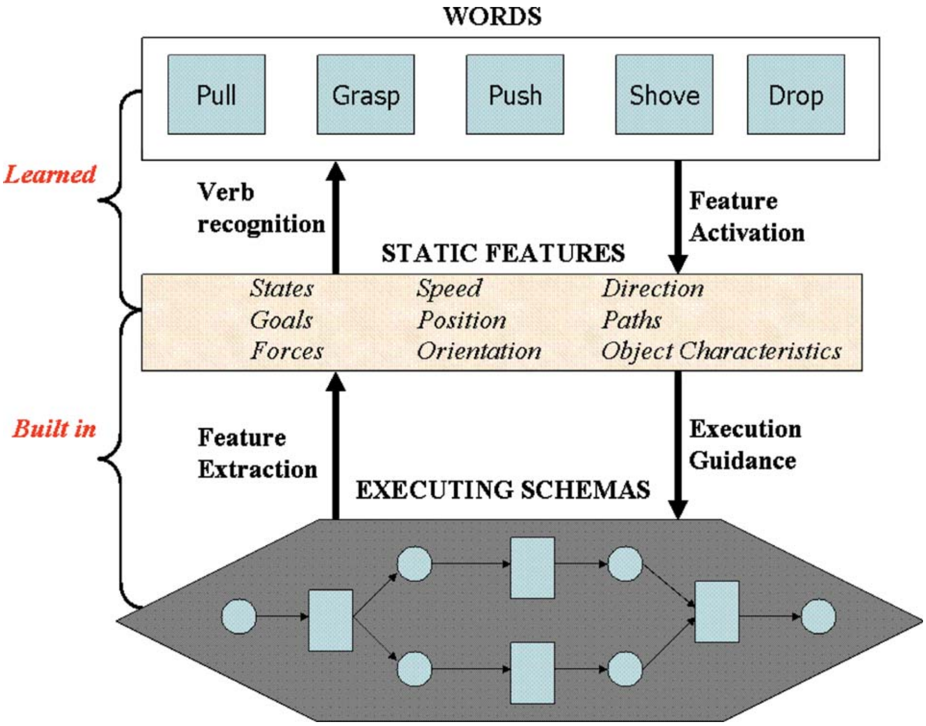

Fig. 1. Overview of Bailey's Model for Learning Action Verbs.

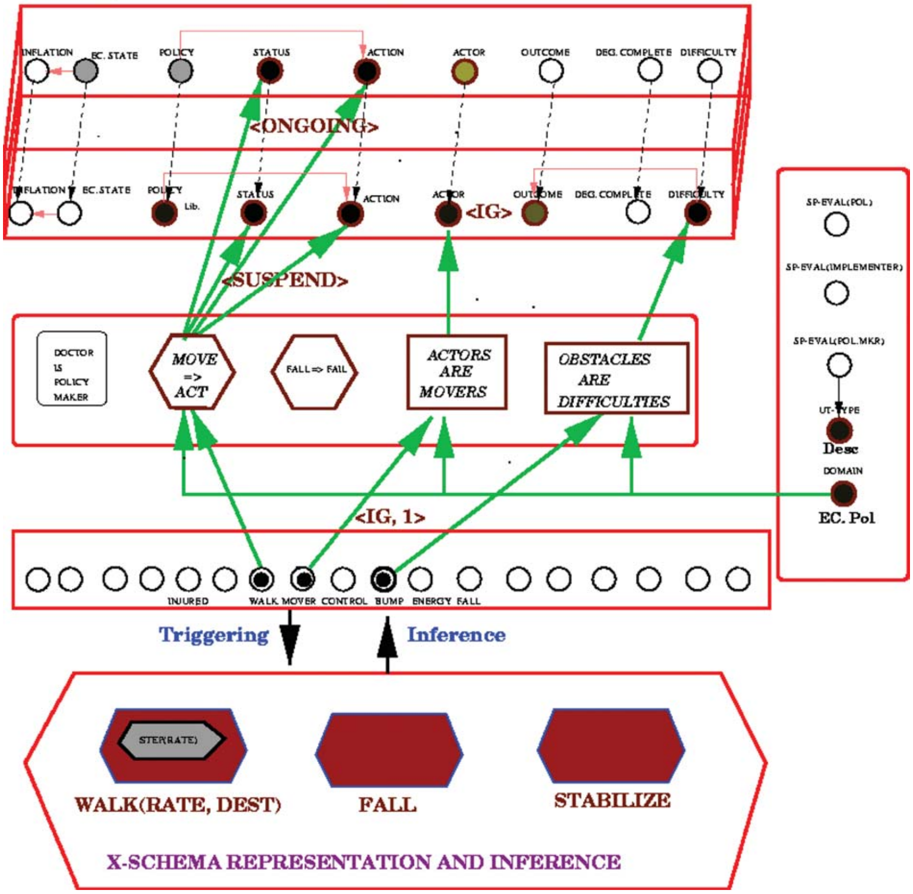

Fig. 2. Architecture of the KARMA metaphor understanding model.

narrative about grasping, such as the introductory paragraph of this paper. Our introduction asked you to imagine yourself at some dinner, being asked to pass the salt. To understand our introduction, or the preceding sentence about it, you need to understand what it means to pass the salt. The NTL assumption is that people understand narratives by subconsciously imaging (or simulating) the situation being described. More generally, NTL suggests that all understanding involves simulating or enacting the appropriate embodied experience. When asked to grasp, we enact it. When hearing or reading about grasping, we simulate grasping or watching someone grasp.

Narayanan (1997, 1999) has built a biologically plausible model of how such metaphorical uses can be understood by mapping to their underlying embodied meaning. The central idea behind Narayanan's model is that the reader interpreting a phrase that corresponds to a motion term is in fact performing a mental simulation of the entailed event in the current context. The basic idea is simple. We assume that people can execute x-schemas with respect to structures that are not linked to the body, the here and the now. In this case, x-schema actions are not carried out directly (as in passing the salt), but instead trigger simulations of what they would do in the imagined situation. This ability to simulate or imagine situations is a core component of human intelligence and is central to our model of language. The system models the physical world as other x-schemas that have input/output links to the x-schema representing the planned action.

In the computational implementation, the spatial motion domain (*source domain*) is encoded as connected x-schemas. Our model of the source domain is a dynamic system based on inter-x-schema *activation*, *inhibition* and *interruption*. In the simulation framework, whenever an executing x-schema makes a control transition, it potentially modifies state, leading to asynchronous and parallel triggering or inhibition of other x-schemas. The notion of system state as a graph marking is inherently distributed over the network, so the working memory of an x-schema-based inference system is distributed over the entire set of x-schemas and source domain feature structures. Of course, this is intended to model the massively parallel computation of the brain.

Fig. 2 gives an overview of KARMA, which uses projections of source domain simulations outlined earlier to interpret narrative fragments about international economics. In the model, *effects* of x-schema execution not only propagate to dependent source domain features but may also be projected by metaphor maps to features of non-physical actions in a domain like economics.

The metaphor understanding system (KARMA) uses conventionalized metaphor maps from the source domains of embodied action and health and well being to project international economics as the target domain. Knowledge of international economics was modeled using a Temporal Probabilistic Network (TBN) (Jensen, 1996). The metaphor maps from the source domains of embodied actions (Lakoff and Johnson 1980, 99) project the x-schema based inferences onto the target domain of abstract actions and economic policies, actors and outcomes.

The KARMA system has been tested on narratives from the abstract domain of international economics. The implemented model has about 100 linked x-schemas, and about 50 metaphor maps from the domains of *health* and *spatial motion*. These were developed using a database of 30 2–3 phrase fragments from newspaper stories all of which have been successfully interpreted by the program. Results of testing the system have shown that a surprising variety of fairly subtle and informative inferences arise from the interaction of the metaphoric projection of embodied terms with the topic-specific target domain structure (Narayanan, 1999). Among the inferences made were those related to *goals* (their accomplishment, modification, concordance, or thwarting), *resources* (consumption, production, depletion, and level), *aspect* (temporal structure of events) *frame-based* inferences, *perspectival* inferences, and inferences about *communicative intent*.

An example of the use of embodied metaphors to communicate important discourse information comes from sentences like the following, which were both successfully interpreted by the KARMA system.

> In 1991, the Indian government *deregulated* the business sector.
>
> In 1991, the Indian government *loosened its stranglehold* on business.

While both headlines talk about privatization and liberalization policies of the Indian government, clearly the use of the embodied predication, *loosen stranglehold*, communicates important information about speaker opinion and communicative intent. Furthermore, the counterfactual implication that if the government had not intervened, the business sector was unlikely to improve and likely to *choke to extinction* is a standard source domain entailment of simulating *strangle* and is immediately available and mapped to the target domain as the *dire consequences of continued government intervention*. Narayanan (1997, 1999), Chang, Gildea, and Narayanan (1998) and Chang, Narayanan, and Petruck, 2002 report on the different types of inferences produced by the KARMA system.

An important and novel aspect of the KARMA representation is that the same system is able to respond to either direct sensory-motor input or to other ways of setting the agent state (such as linguistic devices). This allows for a single mechanism to be used for high-level control and reactive planning as well as for inference through imaginative simulation in language understanding

There is considerable biological support for the idea of simulation semantics. The monkey studies discussed above have been interpreted as suggesting that the various action loops enable the animal to simulate and thus predict consequences of possible actions. A number of recent studies have suggested strongly that homologous circuits in human brains support similar abilities.

First, the *Action-Location Neurons*. Recent brain imaging experiments probed a circuit in humans homologous to F4-VIP in monkeys. Neurons in this circuit were activated when subjects heard or saw objects moved in their peri-personal space (Bremmer et al., 2001). The significance of this is that the area activated during such perception is in the premotor area, the area that would most likely control movements aimed at objects in peri-personal space.

Second, the *Mirror Neurons*. Several studies using different experimental methodologies and techniques have demonstrated also in humans the existence of a similar mirror system, matching action observation and execution (see Buccino et al., 2001; Cochin, Barthelemy, Lejeune, Roux, & Martineau, 1998; Decety et al., 1997; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Hari et al., 1998; Iacoboni et al., 1999; Rizzolatti et al., 1996). In particular, brain imaging experiments in humans have shown that, during action observation, there is a strong activation of premotor and parietal areas, which very likely are the human homolog of the monkey areas in which mirror neurons were found (Buccino et al., 2001; Decety & Grézes, 1999; Decety et al., 1997; Grafton et al., 1996; Iacoboni et al., 1999; Rizzolatti et al., 1996).

Finally, the *Canonical Neurons*. In several recent brain imaging experiments, subjects were asked to (a) observe, (b) name silently, and (c) imagine using various man-made objects (e.g., hammers, screwdrivers, and so on). In all these cases, there was activation of the ventral premotor cortex, that is, the brain region activated when using those same tools to perform actions (Chao & Martin, 2000; Grafton et al., 1996; Martin, Ungerleider, & Haxby, 2000; Perani et al., 1999).

These experiments argue both for the general idea of simulation and also for our particular notion of the embodied meaning of action words. They also provide robust biological evidence that planning, recognition and imagination share a common representational substrate. Our model of actions and their use in narrative predicts that the same structures are being exploited in understanding language about both concrete and abstract events.

Experiments are under way to test this idea, extending the results of (Tettamanti et al., 2002) on premotor activity during reading. The existing experiments show that reading sentences about concrete motor actions (e.g. grasping) leads to activity in somotopically appropriate areas of pre-motor cortex. Subjects will be asked to read sentences involving metaphorical uses of action words and fMRI imaging will ascertain if there is activity in brain areas corresponding to the underlying physical action.

While the discussion so far has focused on action words and phrases, NTL postulates that all linguistic constructions attain meaning through embodiment, not just individual lexical items. This will be discussed in the next section.

4. Beyond words–embodied construction grammar and the NTL

So far the paper has considered only embodied meaning for individual words. The extension to larger (and smaller) linguistic units follows naturally from the NTL adaptation of Construction Grammar (Bergen & Chang, 2002), called ECG. Recall from Section 1 that in ECG, every linguistic unit is a <*form, meaning*> pair, where the meaning part is assumed to link directly to appropriate brain circuits.

Let's reconsider the original example about being asked to pass the salt. One common way to make this request is to say:

> Could you pass the salt?

The surface form of this sentence is a yes/no question, but no one treats it that way. This is one common example of what is called a speech act [Searle, 1969], an utterance that is an action – here requesting the salt. From the ECG perspective, there is a general construction:

> Could you X?

That under certain conditions has the semantics of a request to X. The conditions include several features of the action X, the setting, the social relationships involved, etc. But within the general ECG framework, this is just another example of a contextually conditioned <*form, meaning*> pair fundamentally the same as meaning of a word like grasp.

The full description of how an ECG system would analyze and enact an utterance is beyond the scope of this paper (Bergen & Chang, 2002), but an example should help convey the flavor of the model.

The basic operation in ECG analysis is constructional composition, which we will now illustrate. Consider the phrase "on the table". Following the principles of embodied semantics, the ECG meaning of (this sense of) "on" is an instance of the image schema for *support*. The parser places a support schema in the semantic specification (SemSpec) with two open roles: one for the supported item and one for the supporting item. The semantics of (one sense of) "table" includes the fact that it is a probable supporting item so the parser unifies the correct senses of on and table, yielding a composed SemSpec element for subsequent analysis. A discussion of how the system determines which table is involved is can be found in Feldman (2002).

As the analysis continues, each construction that is recognized gives rise to additional elements and connections in the semantic specification, which is the basis for enactment. In the case of being asked to pass the salt, the result would be a general SemSpec for a personal request linked to an X-schema for passing the salt, assuming the hearer decided to comply with the request. The salt-passing schema would itself be the composition of sub-schemas for locating the salt, reaching for it, grasping it, and placing it near the requester. The enactment of any such schema can not be totally specified by language and depends also on the situation. But it does seem that ECG can explain what is conveyed by language and how it links to the underlying embodied semantics.

The ECG paradigm is also being extended to gestures accompanying language as well as to the intonation and emotional tone of the utterance. You can imagine many ways that someone might ask for the salt that would be radically different in emotion. There are planned ECG constructions that link from prosodic features e.g. a strident voice to meaning, e.g., lack of respect. Much work remains to be done, but the basic ideas on embodied word learning, active simulation, and metaphorical interpretation appear to from the basis for a biologically plausible model of language acquisition and use.

As should be obvious by now, the theory predicts that any utterance will simultaneously match multiple constructions and thus activate multiple brain circuits. The synthesis of embodied meaning and its subsequent enactment is essentially the same task faced in visual or other perception of a complex ongoing situation. One should not expect language to be any more (or less) localized than other perception and action.

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. Dissertation, Computer Science Division, University of California, Berkeley.

Bergen, B. K., Chang, N. C. (2002). Embodied construction grammar in simulation-based language understanding. Technical report 02-004, International Computer Science Institute.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Chang, N. C., Gildea, D., & Narayanan, S. (1998). A dynamic model of aspectual composition. In Proceedings of the twentieth Cognitive Science Society Conference (CogSci 1998), Madison, WI.

Chang, N. C., Narayanan, S., & Petruck, M.R.L. (2002). Putting frames in perspective. In Proceedings of the nineteenth international conference on Computational Linguistics (COLING 2002), Taipei.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12*, 478–484.

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalography and Clinical Neurophysiology, 107*, 287–295.

Decety, J., & Grézes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences, 3*, 172–178.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain, 120*, 1763–1777.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608–2611.

Feldman, J. A. (2002). The meaning of reference in embodied construction grammar. Technical Report, International Computer Science Institute, ICSI TR 02-11, Berkeley, CA.

Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., Matelli, M., Pedotti, A., & Rizzolatti, G. (1992). Space coding by premotor cortex. *Experimental Brain Research, 89*, 686–690.

Fogassi, L., Gallese, V., Fadiga, L., & Rizzolatti, G. (1996a). Space coding in inferior premotor cortex (area F4): Facts and speculations. In F. Laquaniti & P. Viviani (Eds.), *Neural basis of motor behavior. NATO ASI Series* (pp. 99–120). Dordrecht: Kluwer Academic Publishers.

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1996b). Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of Neurophysiology, 76*, 141–157.

Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements. *Experimental Brain Research, 71*, 475–490.

Gentilucci, M., Scandolara, C., Pigarev, I. N., & Rizzolatti, G. (1983). Visual responses in the postarcuate cortex (area 6) of the monkey that are independent of eye position. *Experimental Brain Research, 50*, 464–468.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by PET: 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103–111.

Graziano, M. S. A., & Gross, C. G. (1995). The representation of extrapersonal space: a possible role for bimodal visual-tactile neurons. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1021–1034). Cambridge, MA: MIT Press.

Graziano, M. S. A., Hu, X., & Gross, C. G. (1997a). Visuo-spatial properties of ventral premotor cortex. *Journal of Neurophysiology, 77*, 2268–2292.

Graziano, M. S. A., Hu, X., & Gross, C. G. (1997b). Coding the locations of objects in the dark. *Science, 277*, 239–241.

Graziano, M. S. A., Reiss, L. A. J., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature, 397*, 428–430.

Graziano, M. S. A., Yap, G. S., & Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science, 266*, 1054–1057.

Hari, R., Forss, N., Avikainen, S., Kirveskari, S., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Science, USA, 95*, 15061–15065.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526–2528.

Jensen, F. (1996). *An introduction to Bayesian networks*. Berlin: Springer Verlag.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: The sensory/motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (2nd ed., pp. 1023–1036). MIT Press, Cambridge Ma: A Bradford Book.

Narayanan, S. (1997). KARMA: Knowledge-based active representations for metaphor and aspect. Ph.D. Dissertation, Computer Science Division, University of California, Berkeley.

Narayanan, S. (1999). Moving right along: A computational model of metaphoric reasoning about events. In Proceedings of the National Conference on Artificial Intelligence AAAI-99. Orlando, FL.

Perani, D., Schnur, T., Tettamanti, M., Gorno-Tempini, M., Cappa, S. F., & Fazio, F. (1999). Word and picture matching: A PET study of semantic category effects. *Neuropsychologia, 37*, 293–306.

Pulvermueller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences, 5*(12).

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*, 190–191.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (2nd ed., pp. 539–552). Cambridge, MA: MIT Press.

Rizzolatti, G., & Gentilucci, M. (1988). Motor and visuomotor functions of the premotor cortex. In P. Rakic & W. Singer (Eds.), *Neurobiology of neocortex* (pp. 269–284). Chichester: Wiley.

Rizzolatti, G., Matelli, M., & Pavesi, G. (1983). Deficits in attention and movement following the removal of postarcuate (area 6) and prearcuate (area 8) cortex in macaque monkeys. *Brain, 106*, 655–673.

Rizzolatti, G., Scandolara, C., Gentilucci, M., & Camarda, R. (1981a). Response properties and behavioral modulation of 'mouth' neurons of the postarcuate cortex (area 6) in macaque monkeys. *Brain Research, 255*, 421–424.

Rizzolatti, G., Scandolara, C., Matelli, M., & Gentilucci, M. (1981b). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behaviour and Brain Research, 2*, 147–163.

Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Shastri, L. (2002). Episodic memory and cortico-hippocampal interactions. *Trends in Cognitive Sciences, 6*, 162–168.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Perani, D., Cappa, S. F., Fazio, F., & Rizzolatti, G. (2002). Sentences describing actions activate visuomotor execution and observation systems, MS.

Further reading

Bailey, D., Chang, N., Feldman, J., & Narayanan, S. (1998). Extending embodied lexical development. In Proceedings of the twentieth annual meeting of the Cognitive Science Society COGSCI-98, Madison.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., & Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. *Brain, 124*, 571–586.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 202–251). Cambridge: Cambridge University Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books, ISBN 0-465-05673-3.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press, ISBN: 0226468011.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology, 11*, 194–201.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience Reviews, 2*, 661–670.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology, 12*, 149–154.

Small, S., & Howard, C. N. (2003). On the neurobiological investigation of language understanding in context, *Brain and Language* (this issue).