CHAPTER 21

# Language and action: creating sensible combinations of ideas

Arthur M. Glenberg

COULD two topics be less related than language and action? Both historical (e.g. Descartes, 1992) and contemporary (e.g. Fodor, 1975) philosophers have argued that language is separate from perception and action, that it is a higher faculty, or that it is what separates human from animal. The data, however, present an overwhelming case in favor of an intimate relation between language and action. Much of the data and theory derive from considerations of embodied cognition, and so this chapter begins with a brief overview of that notion. Then, the relation between language and action is considered from the perspectives of neuroscience, cognitive development, and behavioral research. The chapter will conclude with a theoretical rationale for the relation: the mechanism of action planning is the mechanism that allows us to sensibly combine meanings across words and sentences.

## 21.1 Embodied cognition

The embodied cognition framework has been applied to variety of areas including memory (e.g. Glenberg, 1997), concepts (Barsalou, 1999), language (Glenberg and Kaschak, 2004; Lakoff, 1987; Zwaan, 2004), social psychology (Barsalou et al., 2003), and development (Thelen and Smith, 1994). A common theme that runs through these endeavors is that facts of the body (e.g. morphology, perceptual processes) play a prominent role in cognition. Glenberg et al. (forthcoming) characterize embodied approaches to language:

> Linguistic symbols are embodied to the extent that a) the meaning of the symbol ... depends on activity in systems also used for perception, action, and emotion, and b) reasoning about meaning, including combinatorial processes of sentence understanding, requires use of those systems.

At first blush, the notion that language calls upon processes of perception, action, and so on would seem to flounder on the apparent fact that language deals with abstractions such as the meanings of a words, grammatical categories, and generalized syntactic processes (see Vigliocco and Vinson, Chapter 12 this volume). In contrast, Barsalou (1999) discusses conceptual systems built from perceptual symbols. Perceptual symbols are collections of neural activity based on attended aspects of experience, such as shape, sound, and so on. Repeated experience results in different aspects being attended and hence greater knowledge. Importantly, the neural activities which compose perceptual symbols are maintained in the neural systems originally used in perceiving and acting. Thus, for example, the shape of an airplane is represented as a trace of neural activity in the visual system, whereas the sound of the airplane is represented in the auditory system. Perceptual symbols can enter into simulations of various situations, such as an airplane landing or taking off. These simulations are a type of imagery, although there is no requirement that the imagery be consciously experienced, particularly when dealing with familiar situations which may have resulted in simulations becoming automated. Furthermore, Barsalou discusses how a conceptual system based on perceptual symbols and simulators has the power to account for inferences, abstract ideas such as negation, and compositionality. Thus, perceptual symbols constitute a fully functional conceptual system.

The idea that conceptual processes are embodied is supported by several types of data. First, work in the neurosciences demonstrates that language comprehension results in activation in many areas of the brain that bear a systematic relation to the content of the language. For example, language about visual motion produces activation in those areas of the brain (MT/V5) strongly associated with the perception of visual motion (Tettamanti et al., 2005). If language were purely symbolic and divorced from bodily mechanisms such as perception, then this systematicity would have to be treated as an amazing coincidence. Second, behavioral work has demonstrated interactions between language comprehension and tasks that are demonstrably related to perception, action, and emotion. Finally, work in language development shows strong links between bodily mechanisms and language acquisition. We turn to these sources of evidence now.

## 21.2 The neuroscience of language and action

Neuroscience has provided multiple demonstrations of the overlap between areas of the brain contributing to language and those contributing to action (particularly useful reviews are provided by Pulvermuller, 2005, and Rizzolatti and Craighero, Chapter 47 this volume). What happens in our brains when we hear words such as *pick*, *kick*, and *lick*? Of course, there is activation in the left temporal lobe (e.g. Wernicke's area) traditionally associated with language perception. Somewhat surprisingly, there is also activation in the prefrontal area (e.g. Broca's region) normally thought of as contributing to speech production (e.g. Fadiga et al., 2002). The big surprise, however, concerns activation in areas of the brain associated with motor activity. For example, Hauk et al. (2004) used functional magnetic resonance imagery (fMRI) to record brain activity while people listened to verbs. When they were listening to verbs referring to leg actions, regions of the motor cortex that control the leg were particularly active, when they were listening to verbs referring to hand actions, regions of the motor cortex that control the hand were particularly active, and so on. Similarly, Tettamanti et al. (2005) tracked areas of activation while people listened to sentences using verbs requiring mouth actions (e.g. *I eat an apple*), hand actions (e.g. *I grasp the knife*), and leg actions (e.g. *I kick the ball*). As predicted by the embodiment position, these sentences selectively activated areas of the brain associated with mouth, hand, and leg actions, respectively.

The fMRI findings can be interpreted in several ways. The data could indicate that understanding these verbs requires activity in motor areas of the brain. Or the results might simply reflect a habit of envisioning action after hearing action verbs. The balance of the research points strongly to the first interpretation. For example, Pulvermuller et al. (2003) demonstrated that activity in the motor areas occurs very soon (e.g. 20 msec) after the word produces peak activation in areas of the brain traditionally associated with language. The speed of activation would appear to rule out a conscious or optional process. Pulvermüller et al. (2005) activated motor areas using transcranial magnetic stimulation (TMS). When left-hemisphere leg motor areas were activated, people were fast at identifying leg-related words, and when left-hemisphere arm motor areas were activated, people were fast at identifying arm-related words. Finally, Buccino et al. (2005) reported related findings for whole sentences. That is, when people listened to sentences describing leg (or arm) movements while TMS was applied to leg (or arm) motor centers, there was differential modulation of electrical activity recorded in the legs (or arms). Thus, there are strong connections between language and action that can be found in the brain and that extend out to the periphery of the body. The time-course of the effect is too quick, and the causal pathways demonstrated by TMS too convincing, to believe that the link between language and action is optional.

Another finding from the neuroscience literature helps to cement the relation between language and action. Rizzolatti and Arbib (1998) review data on the mirror neuron system and the relation between that system and language. Mirror neurons were first discovered in an area of monkey prefrontal cortex (F5). The mirror neurons in this area respond when the animal takes a particular action, such as ripping a sheet of paper. The same neuron will respond when the animal observes another monkey or a human performing the same action. In fact, the neuron will also respond when the animal simply hears paper being ripped (Kohler et al., 2002). The mirror neuron system has been associated with the ability to recognize the intent of a conspecific's actions (e.g. you understand the intent of another person's gesture because the mirror neurons firing are those that would fire when you take action with the same intent), and hence the system is thought to be particularly important for empathy and social interaction (Gallese et al., 2004).

The story becomes more interesting for the connection between language and action because area F5 in the monkey brain is a homolog of Broca's area (which is involved in speech production) in the human brain. Once this correspondence was noted, research demonstrated that the human Broca's area also contains neurons with mirror-like properties (Fadiga et al., 1995), and that parts of Broca's area control not just the speech articulators but also the hand (Binkofski et al., 1999). From these various correspondences, Rizzolatti and Arbib (1998) suggest that oral language developed from the ability to recognize the communicative intent of actions and gestures. That is, Broca's area evolved into a language area because of its prior usefulness in gestural communication.

The tremendous overlap between neural structures contributing to language and hand/arm movement may help to explain the prevalence of hand gesture in language (McNeill, 1992). Gestures while speaking are nearly universal. Even congenitally blind people speaking to blind listeners gesture (Iverson and Golden-Meadow, 2001), so gesture is unlikely to be something learnt or consciously planned for the benefit of the listener. Nonetheless, gestures do enhance communication (e.g. Valenzeno et al., 2003; Kelly et al., 1999). With two assumptions, it becomes clear why gestures are so prevalent while speaking. The first assumption is that much of meaning is based on action, i.e. what a sentence means to a listener consists of how that sentence describes, suggests, or modulates actions. Data supporting this assumption is reviewed in section 21.4. The second assumption is that meaning can only be determined by utilizing the motor system (section 21.5.2). Thus, sentences convey meanings which have action at their core, and many of our humanly important actions involve the hand (e.g. giving, eating, threatening, or appeasing). Broca's area controls both the speech articulators (e.g. lips, tongue) and the hand. Thus, there can be near-simultaneous activation of the speech apparatus and the hand in the service of the same message.

## 21.3 **Developmental support for the language and action connection**

Several recent reports in the language acquisition literature provide confirmation of the relation between language and action. One particularly interesting set of studies was reported by Smith (2005), who documented the causal relations amongst action, naming, and categorization in children between 24 and 35 months old. In one experiment, children were given a graspable toy shaped roughly like an asymmetric barbell—i.e. there were bumps on both ends, but one bump was larger. The children were told that the object was a *wug*. Half the children were taught to hold the wug using both hands, with one hand on each bump, and to play with it by rotating the wrists. Note that this activity treats the wug in a symmetrical manner. The other children were taught to hold the wug by using one hand to grasp the smaller bump and to wave the wug about. This activity treats the wug in an asymmetrical manner. Following these activities, children were shown other objects that were either more symmetrical or less symmetrical than the original, and were asked to determine which were wugs. Children who acted on the original wug in a symmetrical manner were more likely to classify symmetrical variations as wugs than were the children who acted on the original in an asymmetrical manner, and the reverse was found for the asymmetrical variations. In other words, how the children interacted with the original wug helped to determine what other objects would be called the same name.

## 21.4 **Behavioral support for the connection between language and action**

Behavioral work with adults has also produced strong evidence for the interaction of action and language across multiple levels of language processing: grammatical, basic word meaning, and the meanings of sentences about concrete and abstract situations, as well as the interpretation of extended dialog (Noice and Noice, forthcoming; Ochs et al. 1996; Roth, 1999). An important concept that occurs throughout this work is that of an "affordance." Gibson (1979) coined the term to refer to possible interactions between biological and physical systems. Thus, a chair affords sitting for organisms with the right sort of body morphology (e.g. humans, but not elephants). Some types of chair also afford hiding under for toddlers, but not for adults, whereas some chairs afford throwing in anger by adults, but not by toddlers because they are not strong enough. In brief, what an object affords depends on the body of the organism interacting with the object. Gibson proposed that the function of perceptual systems is to detect affordances that direct action.

Chambers et al. (2004) used the concept of affordances to show the relation between action and grammatical parsing. In their experiments, participants were faced with real situations, such as that illustrated in Figure 21.1, and they heard instructions regarding how to move the objects. One instruction was *Put the whistle on the folder in the box*. Note that there are two whistles in the situation. Thus, clear communication must use some sort of language (or gesture) to discriminate between the two whistles, and that is exactly the function of the phrase *on the folder*. That is, which whistle should be moved? The whistle that is *on the folder*. In another condition, people held a hook and used it to move objects. In this condition, even though both whistles were physically present, there was only one whistle that afforded moving, namely the one with the lanyard that can be grabbed with the hook. Thus, when holding the hook, there is really only one whistle to consider, and the phrase *on the folder* is redundant at best (indicating which whistle to move, but we already know that it is the one with the lanyard) or easily misinterpreted at worst. That is, when holding the hook and listening to the sentence *Put the whistle on the folder…*, people could reasonably interpret *on the folder* as a phrase indicating where the (afforded) whistle might be moved. In fact, when holding the hook, people tended to look at the empty folder (as if they were preparing to move the whistle there) much more frequently than when they were not holding the hook. The point is that the affordances of the situation determined how people parsed the phrase *on the folder*. With the "no-hook" affordances, the phrase was parsed as a reduced relative clause indicating which whistle to move; with the "hook" affordances, the phrase was parsed as a prepositional phrase describing where the whistle was to be moved.

Borghi et al. (2004) also used the concept of affordances to demonstrate that the basic meaning of a word taps action. In one of their experiments, participants read a sentence that mentioned an object, e.g. *There is a car in front of you*. Then, the participant pressed the middle button of a vertically oriented three-button panel. Pressing this button revealed a target word such as *roof*, *wheel*, or *road*. At that point, the participant was to determine if the target was a part of the object named in the sentence (yes for *roof* and *wheel*) or not (for *road*). Note that interacting with the roof of a car normally requires action directed upwards, whereas interacting with the wheel of a car normally requires action directed downwards. Suppose that these different affordances are part of the basic meanings of *car*, *roof*, and *wheel*. That is, suppose that just thinking of the meaning of a word such as *roof* prepares one to act upwards. In this case, participants required to move upwards to the top response button to indicate "yes," should respond faster to a target such as *roof* than participants required to move downwards to indicate "yes." In contrast, for a target word such as *wheel*, those participants required to move downwards should respond faster. This is exactly what Borghi et al. (2004) found. Apparently, when we think about

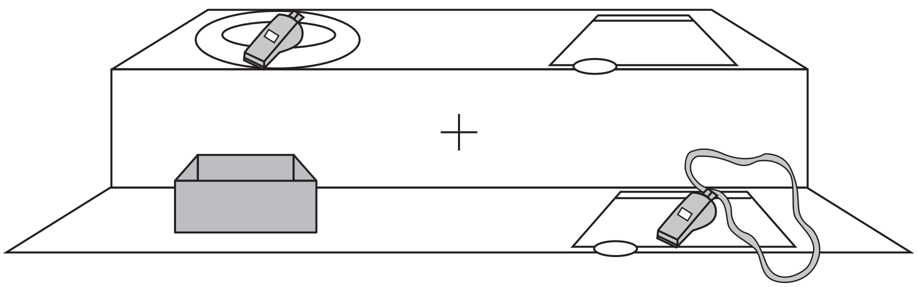

**Figure 21.1** Illustration of one environment used in Chambers et al. (2004). Reprinted from *Journal of Memory and Language*, vol. 30: C. G. Chambers, M. K. Tanenhaus, and J. S. Magnuson, 'Actions and affordances in syntactic ambiguity resolution', pp. 687–96 (copyright © 2004), with permission from Elsevier.

the meaning of a word, at least part of that meaning is in terms of how to act on the object named by the word.

Glenberg and Kaschak (2002) used a similar methodology to determine the contribution of action to the interpretation of whole sentences. The task was to judge if a sentence was sensible (e.g. *Courtney handed you the notebook* or *You handed Courtney the notebook*) or nonsense (e.g. *You drank the house to Joe*). The sensible judgement was made by moving to a button requiring movement away from the body (in one condition) or toward the body (in the other condition). As with the *Courtney* sentences, half of the sensible sentences described action toward the reader and half away. If sentence understanding requires a determination of direction using action systems, then readers literally moving a hand toward the body to make the "sensible" judgement should respond faster to sentences describing action toward the body than to sentences describing action away. The opposite should be found for those readers required to respond "sensibly" by a moving a hand away from the body. This interaction was found, thus demonstrating a contribution of action to sentence comprehension.

The data described so far involve language about concrete objects and activities. But language can also be used to describe abstract feelings, events, transitions, and so on. At first glance, it would appear that action could not possibly contribute to understanding language of this sort; but the data indicate otherwise. One such illustration comes from the Glenberg and Kaschak (2002). In addition to sentences describing the transfer of concrete objects, some described transfer of information from one person to another, e.g. *Liz told you the story* or *You told Liz the story*. The same interaction of transfer direction and literal response direction was found, implying a contribution of mechanisms of action to language understanding of at least some abstract situations.

More impressive are the data from Matlock (2004). Her participants first read descriptions of terrains, for example of a desert described as smooth and flat or as rocky and hilly. Matlock then timed the participants' reading of target sentences such as *A fence runs through it.* These sentences describe "fictive" motion (Talmy, 1996), i.e. nothing mentioned in the sentence is literally moving. Nonetheless, Matlock observed that people took substantially longer to read target sentences describing fictive motion through complex terrains than through simple terrains. That is, people seemed to be simulating movement through the terrain as they cognitively followed the fence. See Vigliocco and Vinson (Chapter 12 this volume) for further discussion of the representation of abstract knowledge.

## 21.5 Language as an opportunistic system

The data from a variety of literatures are convincing: contrary to first impressions, there are strong connections between language and action. Why? The Indexical Hypothesis (Glenberg and Robertson, 1999; 2000; Kaschak and Glenberg, 2000) begins to answer this question. According to the Indexical Hypothesis, language understanding makes use of three processes (none of which need be consciously experienced). The first process is using words and phrases to index (i.e. map to) objects and events in the world or their perceptual symbols. That is, upon hearing a sentence such as *Courtney handed you the pizza*, one indexes *the pizza* to a real pizza in the environment or to a perceptual symbol of a pizza. Second, one derives affordances from the indexed objects. Note that affordances cannot be derived directly from words because words do not have affordances in the traditional sense of supporting bodily interaction; only objects do. Any object, however, may have a tremendously large number of affordances, so which are considered? Derivation of the affordances is controlled in part by syntax. According to construction grammar (e.g. Goldberg, 1995), many sentence forms carry with them information about goals. For example, double-object sentences, such as the one about Courtney (the two objects are *you* and *the pizza*), carry the goal of transfer. Therefore, in deriving affordances for objects mentioned in a double-object sentence, the focus is on those affordances that might be relevant to transfer, in contrast, say, to eating.

The third process specified by the Indexical Hypothesis is combining, or meshing, affordances as specified by syntax. As just noted, the double-object syntax indicates that the pizza is transferred to *you* from Courtney by means of handing. The third process determines how the affordances can be integrated to satisfy the goals (e.g. transfer) and specifications (to you from Courtney) provided by syntax. Only when all three processes—indexing, derivation of appropriate affordances, and meshing—are successful is the sentence understood. For example, if you happen to have a dog named Courtney and you indexed *Courtney* to your dog, you would have difficulty understanding the sentence because dogs do not have the body morphology that allows the affordance of handing. As another example, the sentence *You drank the house to Joe* is difficult to understand because people do not have the body morphology for drinking to transfer a solid object from one person to another. In a world inhabited by creatures with long, large, and flexible digestive systems, however, the sentence might be perfectly acceptable.

Given the power of other approaches to meaning (see Vigliocco and Vinson, Chapter 12 this volume), it is difficult to incisively demonstrate the operation of affordances and mesh. Kaschak and Glenberg (2000) attempted to do so by using innovative language. Because innovations are made up and interpreted on the fly, it is unlikely that the words have long-term semantic representations; instead, meaning must be derived from the situation. Kaschak and Glenberg (2000), asked participants to understand sentences containing innovative denominal verbs. Denominal verbs are made from nouns, such as *to bicycle*. Innovative denominal verbs such as *porched* in *The newsboy porched the paper* (Clark and Clark, 1979) appear to be understood with little effort. Kaschak and Glenberg investigated the hypothesis that people will only understand a sentence with an innovative denominal verb when the object named by the verb has the affordances needed to accomplish the goal specified by the syntactic construction. For example, consider this scenario with two alternative endings:

> A man with a broken leg was eating a hard-boiled egg on a park bench when a soccer ball rolled up from a game being played by school girls. He crutched the girls the ball/He egg-shelled the girls the ball.

Note that neither *to crutch* nor *to egg-shell* is a verb in standard English. Thus, on many accounts, both versions should be seen as nonsense. Nonetheless, the first alternative ending is comprehensible whereas the second is not. Kaschak and Glenberg's (2000) interpretation makes use of the Indexical Hypothesis. The double-object syntax carries the goal of transfer. Here, the soccer ball is transferred from the man to the girls. In the case of *to crutch*, the mechanism of transfer is a crutch, and crutches have the right affordances to transfer a soccer ball. Hence, the affordances can be meshed, as directed by syntax, into a smooth and coherent set of actions that accomplish the goal of transfer, and the sentence is understood. In contrast, egg-shells do not afford transfer of soccer balls. Consequently, comprehension fails for the egg-shell sentence.

### 21.5.1 Why language and action are related: an evolutionary argument

Whereas the Indexical Hypothesis helps us to understand how action systems are used in language comprehension, it does not provide a reason for the intimate connection between language and action. Did it have to be this way? There are two ways to answer this question in the affirmative: yes because of evolutionary pressure; and yes because of function. The evolutionary idea was briefly noted before. Rizzolatti and Arbib (1998) begin with the observation that mirror neurons are a type of communicative device in that they allow signals, i.e. actions produced by one animal, to be comprehended by another. Furthermore, mirror neurons are found in what was to evolve into Broca's area. Finally, there is some evidence (see Craighero and Rizzolatti, this volume) that hand shapes used to grasp large objects invoke an opening of the mouth consistent with vowels used in words designating large objects, and hand shapes used to grasp small objects invoke an opening of the mouth consistent with vowels used in words designating tiny objects. Thus, Craighero and Rizzolatti speculate that oral sounds began to accompany gestures automatically by virtue of the overlap between hand and mouth control in Broca's area. Then, the system of oral communication advanced with evolutionary pressure to communicate quickly, accurately, and under circumstances where gesture is not useful, such as in the dark.

### 21.5.2 Why language and action are related: a functional argument

The functional answer to the question of why language and action are related is based on three conjectures. First, language is a system that depends on combinations of parts (e.g. words). Second, language is opportunistic, i.e. it will use whatever brain mechanisms are available to accomplish the task. Third, action planning is the brain's most well-developed system for producing effective, goal-directed combinations of parts, namely combinations of actions to accomplish a goal.

The first conjecture does not need much defense. Language makes use of a relatively small number of words and rules to create a very large number of meanings and communications. The key is to combine these parts in the right ways.

To what extent is language opportunistic? This chapter has already reviewed data demonstrating that language calls upon action systems. The case can also be made for language calling on perception and emotion systems. For example, Kaschak et al. (2005) had participants listen to and judge the sensibility of sentences conveying motion in specific directions such as toward an observer (e.g. *The car approached you*) or away (e.g. *The squirrel scurried away*). Simultaneously, the participant looked at a spiral spinning so as to convey visual motion toward or away. Kaschak et al. found that sensibility judgements were slowed by the visual stimulus conveying motion in the same direction as implied by the sentence. Apparently, the same visual processing system required for perceiving the spiral motion was also required for understanding the sentence. Because that system was engaged, sentence understanding was slowed (see Tettamanti et al., 2005 for neural imaging data demonstrating activation of visual processing areas in understanding language about actions).

The case can also be made that language is opportunistic in regard to emotion. Glenberg et al. (2005) report an effect of emotion on the ease of language comprehension. They induced emotions using the Strack et al. (1988) procedure. Participants were asked to hold a pen using only their teeth (inducing a smile and a pleasant emotional state) or lips (inducing a frown or pout and a negative emotional state). Concurrently, participants read sentences describing pleasant situations (e.g. *Your lover chases you playfully around the bedroom*) or unpleasant situations. The pleasant sentences were read more quickly when the participant was smiling than when frowning, and sentences describing negative situations were read more quickly when frowning than when smiling. (The question of when embodied states facilitate language comprehension, as in Glenberg et al., 2005, and when they interfere, as in Kaschak et al., 2005, remains to be answered.) Thus, understanding language about emotional situations may call upon neural systems controlling emotion.

Is language an opportunistic system? Although the case for language being opportunistic is just starting to be made, in fact, wherever psychologists and neuroscientists have looked, they have found neural systems that evolved for one purpose (e.g. action, perception, emotion) being used for language comprehension.

The third conjecture is that the action planning system is the brain's best-developed system for producing effective, goal-directed combinations of parts. That is, even simple transitive actions (e.g. picking up a coffee cup and moving it to the lips) requires combinations of multiple muscles (e.g. in the hand and arm) and muscle systems (e.g. those muscles controlling eye movements and posture). Consider further that the goals of similar actions can be tremendously variable (a coffee cup can be picked up for drinking, for throwing, for admiring, for trapping an insect, etc.). Apparently, actions cannot be based solely on fixed routines; instead, the action components need to be flexibly combined. Thus, action planning systems are ready made for the sort of flexible combinations demanded for language comprehension (for a related idea, see Steedman, 2002). Furthermore, the action planning system has learned to avoid the combination of actions that would be physically impossible. Thus, if language were to use this system, there would be a mechanism for determining (at least in part) when sentences do not make sense, or when one is misunderstanding: when the parts cannot be successfully combined to satisfy the goal.

Kaschak and Glenberg's (2000) work with innovative denominal verbs is one demonstration of flexibility in combining novel affordances (actions) in the service of language. Another example is provided by Glenberg and Robertson (2000). They presented participants with sentences describing novel situations which ended in one of three ways, and the participants judged the sentences as sensible or nonsense. For example:

> *Bill needed to paint the top of his barn wall so he stood on his …*
>
> afforded and associated: *ladder*
> afforded: *tractor*
> non-afforded: hammer

The objects named in the "afforded and associated condition" afforded accomplishing the goal and were the usual object used to do so. In the "afforded" condition, the object had the right affordances to accomplish the goal, but an unusual use of the object was required. In the "non-afforded" condition, the object could not be used to accomplish the goal. Participants easily judged the non-afforded condition as nonsensical and the others as sensible, but how? Many of the procedures standardly invoked for determining sensible combinations do not work here. For example, note that both tractors and hammers are tools, both are common words, neither are animate, etc. That is, the afforded and non-afforded objects are equivalent on many of the standard linguistic criteria used to determine if a word combines with its context. Also, because most people have never attempted to paint the top of a barn wall, differential familiarity with the situation cannot contribute to the judgement. Third, it is unlikely that people go through a complex logical reasoning process to determine that a tractor can be used but not a hammer, because participants read and judged the "afforded" sentences in about the same time as they read and judged the "afforded and associated" sentences. Finally, note that there is nothing intrinsically impossible about standing on a hammer (e.g. *The braggart stood on his hammer to show off his sense of balance*); however, standing on a hammer does not afford lifting the body to accomplish the goal of painting the top of a wall. Thus, many of the linguistic processes that we might invoke do not seem to discriminate between, on the one hand, sensible sentences (the "afforded and associated" and the "afforded") and, on the other, nonsense sentences (the "non-afforded"). Instead, the judgement seems to be made by using action planning systems to determine if the components of the sentence can be combined to describe coherent action.

A possible mechanism for this sort of action planning is the forward model. Forward models have been proposed (e.g. Wolpert and Kawato, 1998) as a solution to the feedback problem in motor control. The problem arises because the time required to sense feedback from actions often exceeds the time needed to make the movement. Thus, the sensory feedback arrives too late to help correct quick movements. Grush (2004) extended the formal work of Wolpert and Kawato by proposing an articulated, organism/environment forward model. The model is articulated in the sense that it takes into account the joints, masses, strengths, etc. of the real body. The model is an organism/environment model in that it takes into account "some number of objects and surfaces, their rough sizes and shapes, their dynamical properties (especially movements), and their egocentric locations" (Grush, 2004: 391). When the model is fed the same motor commands as are issued to the real body, it can be run fast enough to provide quick feedback both to correct action and to aid in interpretation of the actual sensory feedback (Wilson and Knoblich, 2005). Grush discusses how such a model can be used for motor control, to aid perception, and as a source of motor and visual imagery.

Grush also speculates on how the forward model (or "emulator," as he calls it) can be used in language comprehension. For example, consider how it can be combined with the Indexical Hypothesis to analyze understanding of two sentences, *You kicked Joe the telephone/skyscraper*. Upon hearing *kicked*, the motor system related to the leg becomes active (Hauk et al., 2004; Tettamanti et al., 2005). Upon hearing *the* it is clear that a second noun phrase will be encountered. Thus, the sentence is likely to be a double-object sentence implying transfer of a second object named by the noun phrase (*the telephone* or *the skyscraper*) to the first (*Joe*). Once a perceptual symbol for the second object is inserted into the model (through the indexing process), there is enough information to begin an emulation, namely using a kicking action to transfer the object to Joe. The articulated model will reveal that it is possible (albeit quite unusual) to transfer the telephone by kicking, but that it is not possible to transfer the skyscraper this way (at least not if indexing inserted a realistically sized skyscraper into the model).

Clearly, this is a very speculative account. It has the advantage, however, of being highly testable. For example, the model predicts that many aspects of motor control (e.g. the common coding principle; Hommel et al., 2001) should apply to language understanding. Also, forward models need to be updated or recalibrated to take into account the current sensorimotor environment. When using a tool, for example, literal movement in one direction can cause an effect in another direction. Would such a recalibration affect language comprehension? Or, if one fatigues an effector system through repeated use, and presumably the forward model is updated to reflect this fatigue, will that fatigue affect language comprehension about that effector (e.g. about kicking a telephone)? Whereas the scientific literature currently offers no answers to these question, it is easy to envision how they might be answered.

## 21.6 Conclusions

Advances in scientific understanding have come at the cost of demoting humans from a unique position in the universe. We no longer believe that the sun revolves around the earth, or that humans are unrelated to other animals. A corollary of the belief that humans are unrelated to other animals is that the mechanisms of language constitute a higher faculty unrelated to our bodily nature. That corollary is also succumbing to the weight of the data. For example, language seems to be related to mirror systems found in other primates. Furthermore, language uses neural systems highly evolved in many animal species.

Clearly, much work needs to be done to secure the ideas (a) that language is opportunistic in the sense described above and (b) that the mechanisms of action control provide the basis for combining ideas conveyed by language. For example, as currently formulated, the Grush emulator account is far from being able to handle the majority of what is known about syntax. Nonetheless, because clear and novel predictions can be derived from embodied accounts, it is certain that they will be tested and lead to a clearer understanding of the close relationship between language and action.

## Acknowledgements

This work was supported by NSF grant BCS-0315434 to Arthur Glenberg. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

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

Barsalou, L. W., Niedenthal, P. M., Barbey, A., and Ruppert, J. (2003) Social embodiment. In B. Ross (ed.), *The Psychology of Learning and Motivation*, vol. 43, pp. 43–92. Academic Press, San Diego, Calif.

Binkofski, F., Buccino, G., Stephan, K. M., Rizzolatti, G., Seitz, R. J., and Freund, H.-J. (1999). A parieto-premotor network for object manipulation: evidence from neuroimaging. *Experimental Brain Research*, 128: 21–31.

Borghi, A. M., Glenberg, A. M., and Kaschak, M. P. (2004) Putting words in perspective. *Memory and Cognition*, 32: 863–73.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005) Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research*, 24: 355–63.

Chambers, C. G., Tanenhaus, M. K., and Magnuson, J. S. (2004) Actions and affordances in syntactic ambiguity resolution. *Journal of Memory and Language*, 30: 687–96.

Clark, E. V., and Clark, H. H. (1979) When nouns surface as verbs. *Language*, 55: 767–811.

de Vega, M. (forthcoming) Levels of embodied meaning: from pointing to counterfactuals. In A. M. Glenberg, M. de Vega, and A. C. Graesser (eds.), *The Garachico Workshop on Symbols and Embodiment*. Oxford University Press, Oxford.

Descartes, R. (1992 [1641]) *Meditations on first philosoph : in which the existence of God and the distinction of the human soul from the body are demonstrated*. Translated by G. Heffernan. University of Notre Dame Press, Notre Dame, Ind.

Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience*, 15: 399–402.

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

Fodor, J. A. (1975) *The Language of Thought*. Harvard University Press, Cambridge, Mass.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004) A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8: 396–403.

Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*. Houghton Mifflin, New York.

Glenberg, A. M. (1997) What memory is for. *Behavioral and Brain Sciences*, 20: 1–55.

Glenberg, A. M., de Vega, M., and Graesser, A. C. (forthcoming) Framing the debate. In A. M. Glenberg, M. de Vega, and A. C. Graesser (eds.), *The Garachico Workshop on Symbols and Embodiment.* Oxford University Press, Oxford.

Glenberg, A. M., Havas, D. A., Becker, R., and Rinck, M. (2005) Grounding language in bodily states: the case for emotion. In R. Zwaan and D. Pecher (eds.), *The Grounding of Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, pp. 115–28. Cambridge University Press, Cambridge.

Glenberg, A. M., and Kaschak, M. P. (2002) Grounding language in action. *Psychonomic Bulletin and Review*, 9: 558–565.

Glenberg, A. M., and Robertson, D. A. (1999) Indexical understanding of instructions. *Discourse Processes*, 28: 1–26.

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

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

Grush, R. (2004) The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences,* 27: 377–442.

Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41: 301–307.

Hommel, B., Muesseler, J., Aschersleben, G., and Prinz, W. (2001) The theory of event coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences*, 24: 849–878.

Iverson, J. M., and Goldin-Meadow, S. (2001) The resilience of gesture in talk: gesture in blind speakers and listeners. *Developmental Science*, 4: 416–422.

Kaschak, M. P., and Glenberg, A. M. (2000) Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, 43: 508–529.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A., and Zwaan, R. A. (2005) Perception of motion affects language processing. *Cognition*, 94: B79–B89.

Kelly, S. D., Barr, D. J., Church, R. B., and Lynch, K. (1999) Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, 40: 577–592.

Kohler, E., Keysers, C., Umiltá, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297: 846–8.

Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, Chicago.

Matlock, T. (2004) Fictive motion as cognitive simulation. *Memory and Cognition*, 32: 1389–1400.

McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought.* University of Chicago Press, Chicago.

Noice, T., and Noice, H. (forthcoming) What studies of actors and acting can tell us about memory and cognitive functioning. *Current Directions in Psychological Science*.

Ochs, E., Gonzales, P., and Jacoby, S. (1996) "When I come down I'm in the domain state": grammar and graphic representation in the interpretive activity of physicists. In E. A. Schegloff and S. A. Thompson (eds.), *Interaction and Grammar*, pp. 328–69. Cambridge University Press, Cambridge.

Pulvermüller, F. (2005) Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6: 576–582.

Pülvermuller, F., Hauk, O., Nikulin, V. V., and Ilmoniemi, R. J. (2005) Functional links between motor and language systems. *European Journal of Neuroscience*, 21: 793–797.

Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R. J. (2003) Spatio-temporal patterns of neural language processing: an MEG study using Minimum-Norm Current Estimates. *Neuroimage*, 20: 1020–1025.

Roth, W.-M. (1999) Discourse and agency in school science laboratories. *Discourse Processes,* 28: 27–60.

Rizzolatti, G., and Arbib, M. A. (1998) Language within our grasp. *Trends in Neuroscience*, 21: 188–194.

Smith, L. B. (2005) Action alters shape categories. *Cognitive Science*, 29: 665–

Steedman, M. (2002) Plans, affordances, and combinatory grammar. *Linguistics and Philosophy*, 25: 723–753.

Strack, F., Martin, L. L., and Stepper, S. (1988) Inhibiting and facilitating condition of facial expressions: a non-obtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology,* 54: 768–777.

Talmy, L. (1996) Fictive motion in language and "ception." In P. Bloom, M. A. Peterson, L. Nadel, and M. F. Garrett (eds.), *Language and Space*, pp. 211–76. MIT Press, Cambridge, Mass.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P. Fazio, F., Rizzolatti, G., Cappa, S. F., and Perani, D. (2005) Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17: 273–281.

Thelen, E., and Smith, L. B. (1994*) A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, Cambridge, Mass.

Valenzeno, L., Alibali, M. W., and Klatzky, R. (2003) Teachers' gestures facilitate students' learning: a lesson in symmetry. *Contemporary Educational Pscyhology,* 28: 187–204.

Wilson, M., and Knoblich, G. (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131: 460–473.

Wolpert, D. M., and Kawato, M. (1998) Multiple paired forward and inverse models for motor control. *Neural Networks*, 11: 1317–1329.

Zwaan (2004) The immersed experiencer: toward an embodied theory of language comprehension. In B. Ross (ed.), *The Psychology of Learning and Motivation*, vol. 44: pp. 35–62. Academic Press, San Diego, Calif.