DISCOURSE PROCESSES, *28*(1), 61–80
Copyright © 1999, Lawrence Erlbaum Associates, Inc.

COMMENTARIES

Language Comprehension: Archival Memory or Preparation for Situated Action?

Lawrence W. Barsalou
Department of Psychology
Emory University

Current paradigms study language comprehension as if archival memory were its primary function. Participants only receive linguistic material and are later tested on memory for its contents. In contrast, the 2 target articles in this issue—Glenberg and Robertson, and Roth—examine comprehension as if preparing for situated action were its primary function. Besides receiving linguistic materials as input, participants study objects, actions, and interactions among agents. Rather than simply being tested on memory for linguistic materials, participants also produce actions and enter into group interactions. Although these researchers focus their attention on specific genres—the comprehension of verbal instructions and the comprehension of scientific theories—their methods and findings have wider implications. In particular, the primary function of comprehension is not to archive information but is instead to prepare agents for situated action. Arguments from the evolution of cognition and language are brought to bear on this thesis, and perceptual simulation is proposed as a mechanism well suited for supporting situated comprehension. Finally, it is conjectured that studying comprehension in the context of situated action is likely to produce significant scientific progress.

> Sense fades into reference.
> —Roth (p. 30)

If an outsider reviewed research on language comprehension, what conclusions might he or she reach? After reviewing this literature myself for a text on cognitive

Correspondence and requests for reprints should be sent to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, GA 30322. E-mail: barsalou@emory.edu

Copyright ©2000. All Rights Reserved.

psychology (Barsalou, 1992, chap. 8 and 9), I concluded that comprehension is essentially *archival memory,* describing it as follows:

1. Words enter the cognitive system through phonemic and graphemic processing.
2. Word representations are translated into amodal syntactic structures and amodal semantic propositions.
3. Syntactic structures and semantic propositions are combined to form an integrated representation of the text, phrase by phrase, sentence by sentence.
4. Inferences may be added minimally or richly from amodal background knowledge.
5. A mental model or situation model of the text may be constructed.
6. The structures in Points 3, 4, and 5 are used to produce memories of the text and inferences about it, as measured by recall, recognition, and reading time.

By focusing on the storage and retrieval of texts, the dominant paradigms for studying language comprehension imply that it is essentially an archival process. Information is extracted from texts and stored away to reproduce information later as needed. Not only do the dominant paradigms suggest this, so do the dominant theories, which specify particular mechanisms for implementing archival functions.

The paradigms studied by Glenberg and Robertson and by Roth suggest a much different way of thinking about comprehension. Rather than primarily performing an archival function, language comprehension prepares agents for situated action. Several aspects of these paradigms highlight this substantially different emphasis. First, rather than simply assuming that the sole input to comprehension is words, these researchers assume that the input is much more complex, including the perception of physical objects, physical events, the body, and other agents. Second, the information stored from this input is not merely a description of what has transpired but also potential perceptions, actions, and interactions that may be relevant in the future. Third, rather than simply assuming that the outputs of comprehension are verbal memories, binary responses, and reaction times, these researchers assume that the outputs include actions on objects, interactions with other agents, information seeking, and bodily reenactments, such as manual and facial gestures.

Together, these added complexities in input, storage, and output suggest a fundamentally new way of thinking about comprehension. Rather than primarily establishing an archival memory, language comprehension prepares agents for situated action.

SPECIFIC FINDINGS AND GENERAL CLAIMS

If we were to revise our basic view of language comprehension, what form might a new view take? Glenberg and Robertson as well as Roth are careful to limit the scope of their claims. Glenberg and Robertson focus on the comprehension

Copyright ©2000. All Rights Reserved.

of verbal instructions that serve the later execution of a physical task (i.e., using a compass). When comprehenders index (i.e., associate) verbal instructions with the perception of physical objects and events, they become optimally prepared for subsequent task performance. For example, associating the expression "compass arrow" with the perception of a compass arrow links the expression with corresponding perceptual memories. When "compass arrow" is heard later, the associated perceptual memories provide a precise meaning and help establish reference to a physical referent, if present. Indexing is related to *reference* in philosophy and linguistics, and various theorists have noted its absence in accounts of cognition (e.g., Barsalou et al., 1993; Johnson-Laird, Herrmann, & Chaffin, 1984).

Roth focuses on how high school science students comprehend scientific theories. Rather than only using language to discover and represent scientific understandings, students also use physical objects and events in collaboration with other agents. As a result, the process of comprehension is distributed, not just residing within a single individual.

Neither Glenberg and Robertson nor Roth make general claims about language comprehension. Glenberg and Robertson are careful to stress the importance of indexing only with respect to the comprehension of verbal instructions, not to all genres of comprehension. Roth is careful to stress the importance of distributed representation only in the novice comprehension of a few scientific theories, not in the comprehension of all scientific theories, not in the cognition of expert scientists, and certainly not in all comprehension genres.

Although these researchers focus on specific genres, their methods and observations suggest wider implications. Following Glenberg and Robertson, we might consider the possibility that indexing is of widespread importance in comprehension. On encountering language, associating its elements with perceived referents in the accompanying situations is often central to the basic goals that comprehenders seek to achieve. Rather than simply establishing an internal representation of the text for archival purposes, the cognitive system attempts to link these internal representations with the physical world either immediately or later. Following Roth, we might consider the possibility that language is just one dimension of a complicated system that discovers and represents the conceptualization of a domain. Rather than restricting input, storage, and output to linguistic elements, the comprehension system distributes these representations and processes across the physical environment, the body, events, and other agents. In this spirit, the remainder of this commentary develops the general view that comprehension is essentially preparation for situated action.

Similar arguments about the central role of situated action in cognition can be found in A. Clark (1997), Johnson (1987), Newton (1996), and Greeno (1998). Arguments about the importance of situated action in language specifically can be found in H. H. Clark (1992), Gibbs (1994), Lakoff (1987), and MacWhinney (1999). Research on other species clearly indicates that communication serves

Copyright ©2000. All Rights Reserved.

situated action (e.g., Gouzoules & Gouzoules, 1995; Gouzoules, Gouzoules, & Ashley, 1995; Smith, 1977). Glenberg (1997) addressed the implications of situated action for memory, and Barsalou (1991) and Ross (1996) addressed the implications of situated action for categorization. In general, the arguments offered in this commentary for language comprehension extend to most other areas of cognition.

FUNCTIONS OF COMPREHENSION

The move to generalize indexing and situated action broadly across comprehension might elicit negative reactions in some readers. First, much comprehension occurs outside the situations described. Rather than indexing a speaker's current situation, language often describes absent situations as well as unfamiliar situations. Second, much text is not about actions but is instead about states of affairs, both physical and psychological. Rather than instructing someone how to do something, language often describes situations in which no actions are to be taken. How can we resolve these reactions with the claim that the primary function of comprehension is to prepare agents for situated action?

Modern Biases

It is essential to acknowledge the biases that we as highly trained scientists bring to bear on these issues. Indeed, it may well be in our own best interests to do everything we can to suppress these biases, assuming that our goal is to understand language in everyday life. Three biases seem particularly important to acknowledge: First, we often focus on language as it is used in formal education. Traditionally (but not optimally!), teachers stand up in front of students and convey information that students are to archive for later reproduction on examinations. The texts that accompany such instruction often function similarly. Second, as academics, we pride ourselves on our ability to archive and reproduce facts that we acquire through language. Indeed, it is our responsibility to become archival experts about our respective areas of study. Third, perhaps the easiest way to study language is as an archival process. Compare the relative ease of performing traditional studies on comprehension with the difficulty of performing the studies in Glenberg and Robertson and in Roth.

Via availability, these three biases may have considerable impact on how we conceptualize language comprehension (Tversky & Kahneman, 1973). To the extent that comprehension in formal education, scientific knowledge, and current research paradigms are highly available in our experience, our conceptualization of comprehension will emphasize archival functions. Conversely, viewing the primary function of comprehension as preparing agents for situated action will seem misguided.

Copyright ©2000. All Rights Reserved.

The Evolution of Language

Thinking about comprehension from the perspective of evolution suggests a very different conceptualization. When we replace formal education and science with the coordination of hunting, gathering, simple industry, and social organization, comprehension no longer seems to be an archival process. Instead, it seems much more oriented toward preparing agents for situated action. Consider recent proposals of how human cognition evolved (e.g., Bickerton, 1990; Corballis, 1991, 1998; Donald, 1991, 1993; Mithen, 1996; Tomasello, Kruger, & Ratner, 1993). Although these proposals differ in important ways, they generally assume that increasing social coordination was one of the primary pressures that shaped human cognition. By coordinating the actions of multiple agents, humans increased their fitness. The larger a coordinated group, and the more specialized the roles its members play, the greater its control over environmental resources. Good or bad, the power of sophisticated social coordination is obvious in humans' ability to control and utilize the environment.

These evolutionary theorists have further argued that increasingly sophisticated social coordination resulted, at least in part, from increasingly sophisticated language. Although these accounts again differ in important ways, they agree that evolving linguistic abilities were critical. Highly developed language helps achieve social coordination in at least three ways. First, language can be used to establish shared beliefs about the environment. If an agent views a particular location, he or she can describe it to others who have not seen it, specifying the geography, people, objects, and events present. Second, language can be used for describing actions to perform on the environment. If an agent discovers a procedure that is effective at acquiring environmental resources, he or she can describe it to others, specifying the enabling conditions, the action sequence, troubleshooting, optimization, and so forth. Third, language can be used to specify roles in groups, thereby implementing a division of labor. If a group develops a complex procedure that requires its members to play different roles, its members can describe these roles to novices and discuss ways to optimize them with each other.

To make these functions of language more concrete, imagine the sorts of coordinated activities in which early humans were probably engaged, such as hunting, gathering, simple industry, and social organization. As evolutionary theorists have noted, language could have easily evolved to support such activities. For example, language can be used to describe where prey might be found, procedures for hunting them, and different roles that individual hunters might play to achieve success. Hunts for large land and sea mammals clearly benefit from such coordination. Gathering, simple industry, and social organization similarly benefit from being able to describe the world, procedures, and roles.

When thinking about language in these contexts, situated action appears much more important than archival memory. Language was probably not used to archive information, as it is in formal education and science; instead, it was used to control action in specific situations. Most important, because language evolved

Copyright ©2000. All Rights Reserved.

over a period of human culture in which activities like hunting, gathering, simple industry, and social organization became increasingly important, it would make sense that the foundational properties of human language today reflect those evolutionary pressures then. Formal education and science have occurred much too recently to have had such impact.

Clearly, language does perform archival functions. In early human cultures, language played a central role in cultural transmission. Just like today, adults imparted their knowledge to children through language, thereby implementing a sort of archival memory. Yet, such transmission was probably not at all like formal education, taking instead the form of apprenticeship in contexts of situated action. Either while performing a specific activity or while discussing it offline, adults described relevant situations to children, the procedures for performing the activity, and various roles to be played in it. Rather than the goal being to reproduce knowledge verbally, the goal was to perform situated action competently. Only in very recent history have adults established formal schools that convey abstract bodies of knowledge for verbal reproduction outside contexts of action. Notably, many recent theorists have questioned the viability of this approach (e.g., Brown, Campione, Webber, & McGilly, 1992; Resnick & Resnick, 1992), arguing instead that education should be oriented toward situated action!

The information stored on a particular occasion may not always have obvious applications. Instead, its applications may only become relevant on later occasions. For example, an adult describing the location of a cave to a child might not have any immediate use for the cave in mind. On a later occasion, however, the child might need a safe haven in a severe storm, remember that the cave is nearby, and seek shelter in it. In such circumstances, stored information eventually serves situated action—it is not simply a fact to be reproduced at a later time.

The more one thinks about language in its evolutionary context, the more available indexing and distributed representation become in one's conceptualization of comprehension. If the goal of communication is largely to coordinate situated action, then indexing the meanings of words becomes critical, as do actions on objects and interactions with agents. Comprehension starts looking a lot more like the paradigms in Glenberg and Robertson and in Roth than like the paradigms that dominate current research.

The Immediacy of Indexing

Clearly, language does not always describe situations that are immediate (present); it also describes situations that are displaced (absent). In evolutionary terms, this is probably nothing new. Early humans probably spent considerable time talking about displaced situations, just like us. Indeed, the ability to discuss displaced situations may have been a primary pressure on language evolution (Donald, 1991, 1993). How are we to resolve this with the view that indexing and situated action are central to comprehension?

Copyright ©2000. All Rights Reserved.

To lay the groundwork for a potential solution, consider some possible ways in which language can index situations. First, in *immediate indexing,* conversationalists simultaneously view the physical situation under discussion. For example, several hunters might discuss a herd of antelope while observing them from afar. Second, in *displaced indexing on actual experience,* conversationalists discuss an absent situation that they viewed earlier. Thus, the hunters might be sitting around a fire later that evening discussing the antelope herd they had seen earlier in the day. Third, in *displaced indexing on similar experience,* listeners have a situation described to them that is similar but not identical to one they have experienced previously. Thus, another hunting party who has seen antelope herds before but not the one seen by the first hunting party might have it described to them. Fourth, in *displaced indexing on componential experience,* listeners have a situation described to them that they have not experienced as a whole but whose components have been experienced in previous situations. For example, another hunting party that has never seen an antelope herd might have antelope described to them in terms of the shapes, colors, and body parts of other animals that are familiar. If this other hunting party has previously encountered zebras, brown things, and animals with horns, then the antelope could be described as "like zebras but brown and with horns."

So far, all these examples of displaced indexing refer to past events. However, displaced indexing can also refer to future events. Thus, the hunters might discuss a future event involving the same herd that they had all seen earlier (displaced indexing on similar experience). Similarly, another hunting party that has never hunted antelope might hear about the hunt planned for the next day, described with components from other kinds of hunts that are familiar (displaced indexing on componential experience).

These various forms of indexing can combine to produce more complex forms. Imagine that the hunters are watching an antelope herd as they discuss it, perhaps trying to identify the dominant male and the females without young. During this part of the conversation, most of the indexing is immediate, relating linguistic expressions to perceived entities. While still viewing the herd, however, the hunters may begin to develop a specific plan for hunting it. At this point, they have switched from immediate to displaced indexing on similar experience because the planned event has not occurred but is similar to events that they have experienced. Finally, as the hunters formulate their plan, they may refer to previous episodes of hunting antelope as they try to converge on an optimal procedure. At this point, they have switched to displaced indexing on actual experience, assuming that they have all experienced the past hunts under discussion. In this manner, indexing can become highly complicated, not just over the course of a conversation but in a single utterance.

Although indexing takes many complex forms in conversations, the focus typically remains grounded in particular objects, events, and agents in the world. Indeed, the whole conversation revolves around preparing for situated action.

Copyright ©2000. All Rights Reserved.

Even when indexing refers to something not present, it nevertheless refers to something in the past or future that is essential to performing situated action. The purpose of communication is not simply to archive information about the world and the events that occur in it, the purpose is to coordinate the actions of multiple agents toward achieving goals in specific situations. Thus, the fact that language is often about displaced situations does not pose a problem for indexical and situated views of language. To the contrary, the whole point of such language is typically to optimize situated action later as the envisioned situations become immediate.

Of course, a critical question is whether language still plays the same role in modern cultures. Excluding academia and other professions in which archiving information is central, what role does language typically play? Clearly, rigorous scientific analysis of this question is essential for a definitive answer. Over the course of writing this commentary, however, I have been informally observing the conversations around me. So, for example, I have recently heard discussions of good places to shop for fresh produce, good restaurants with short waits, where best to park on campus, how best to treat children with sinus infections, the best route for getting to work, neighborhoods where houses are still affordable, how best to ensure that house remodelers finish on time, sports to watch during the NBA strike, and so forth. Admittedly, I am probably biased in my sampling, but nevertheless, I would not be at all surprised to find that this collection of conversational topics is quite representative of what many people in our culture discuss much of the time.

What is clear about such conversations is that they typically serve situated action. In most cases, the conversationalists do not simply provide information for archival purposes. Although listeners store information, its purpose is usually to support situated action either in the immediate situation or in a later one—not simply to be reproduced later. Thus, a tremendous amount of energy is often spent on indexing. Conversationalists attempt to index parking lots, good restaurants, symptoms of sinus infections, indications of problematic remodelers, and so forth. In each case, considerable effort goes into describing referents in the world or into actually pointing them out, so that they can later be recognized accurately and easily. Similarly, a tremendous amount of energy is spent on describing procedures as well as on describing interactions with other agents. Conversationalists discuss procedures for treating sinus infections, for ensuring that remodelers finish on time, for finding affordable houses, and so forth.

What about reading newspapers and other sources of nonfiction? What about escapism in fiction? How do these activities prepare people for situated action? First of all, some people probably do focus on archiving information some of the time. Nevertheless, this appears to be relatively unusual relative to preparing for situated action. Speaking from personal experience as both a speaker and a listener, people whose social strategy is oriented toward archiving information are not the most successful conversationalists. Further witness the decline in

Copyright ©2000. All Rights Reserved.

reading relative to other media. One possibility is that television has replaced print media because the visual information available in television makes linguistic indexing, and therefore comprehension, much easier. Finally, when people do archive information, it often appears to serve situated action. Thus, voters are admonished to follow politics and read history so that they can act as informed citizens in voting and in expressing their views.

Escapism and fiction might seem least relevant to situated action, but consider the possibility that these modes allow people to experience situated action that they normally would not experience, either because it is impossible or too dangerous. It is certainly not the case that the readers of fiction are simply archiving the contents of novels. To the contrary, they are experiencing the situated action and events that novels are about. Indeed, a mark of good fiction is how experientially real it seems. Furthermore, from experiencing the situations described in novels, readers may learn moral lessons, acquire wisdom, and come to exercise better judgment. Arguably, the oral tradition that perpetuated fiction across generations before the advent of printing similarly served this purpose (cf. Rubin, 1995).

Although the specific contexts of situated action might appear to have evolved considerably from those of earlier humans, the role of comprehension still appears much the same. Rather than simply serving to archive information, comprehension serves to prepare agents for situated action, or to at least create the experience of situated action.

PERCEPTUAL SIMULATIONS OF SITUATED ACTION

Up to this point, the focus has primarily been on the purposes of comprehension and on the contexts in which it is used. In this section, the focus will be on the cognitive mechanisms that underlie comprehension. The argument will be that mechanisms of perceptual simulation are central because they are well suited for perceiving, monitoring, and executing situated action. Because perceptual simulations are represented in essentially the same format as perception and action, they readily support these activities during goal-directed behavior. The argument will be further that, even in so-called abstract situations, perceptual simulation is well suited for supporting situated action.

This discussion adopts the framework of perceptual symbol systems (Barsalou, 1999). In this framework, perceptual simulation works as follows. During perception of the external world, proprioceptive events, and introspective states, selective attention focuses attention on components of experience (e.g., shapes, colors, actions, sounds, smells, movements, emotions). As a result, associative areas in the brain capture patterns of activation in sensory-motor systems and in systems that represent introspection (e.g., emotions, cognitive operations). Later, these associative areas partially reactivate these perceptual representations in the

Copyright ©2000. All Rights Reserved.

absence of perceptual input, thereby simulating the experience of what an external or internal event was like. Barsalou (1999) illustrated how these simulation mechanisms implement the type–token distinction, categorical inference, the productive construction of novel simulations, and the representation of propositions. Barsalou also illustrated how these simulation mechanisms could underlie the knowledge that supports basic cognitive processes, including perception, categorization, memory, language, and thought.

Perceptual Simulations in Immediate Indexing

To see how perceptual simulation supports comprehension during situated action, first consider immediate indexing. Imagine that several people are discussing an immediate situation. For example, as several hunters observe an antelope herd, one might assert that the herd contains a dominant male and several females without young. How might the other hunters establish the referents of this assertion? Most simply, if the other hunters previously perceived male, female, and young antelope, and if they stored perceptual representations of them, they can use these representations to index the speaker's assertions. For example, by running a perceptual simulation of a previously experienced male, and then comparing this simulation to individuals in the herd, listeners may be able to establish joint reference with the speaker. It is difficult to think of a form of representation that would be better suited for this function than perceptual simulations. Because the simulations are in essentially the same representational format as the perceptions, the two can be readily compared and integrated (see Glenberg's, 1997, construct of *mesh*).

Perceptual Simulations in Displaced Indexing

Next, consider the mechanisms that might support displaced indexing on similar experience. Imagine that the hunters watching the antelope start predicting how various individuals will behave when hunted. For example, the hunters might predict that the herd will initially head away from the hunters but then circle back behind them, anticipating that the hunters will continue forward into the distance. What might be the best form of representing such inferences? Again, perceptual simulation seems well suited to the task. First of all, on previous occasions, the hunters presumably saw antelope adopt this escape tactic. Thus, a natural way for the hunters to represent this inference would be to reactivate those earlier perceptual experiences. Second, such simulation would optimally serve later situated action based on these inferences. When the hunt actually begins, if the hunters represent this prediction with simulations of it, they will quickly be able to recognize its presence in actual events. If the herd does indeed circle back, the hunters' perception of this event should match their earlier envisionments of it, such that it can be recognized and the appropriate actions taken.

Copyright ©2000. All Rights Reserved.

Perceptual simulations support not only the recognition of entities that are currently present but also the recognition of entities and events inferred to become present later.

Perceptual Simulation of Actions and Interactions

Similarly, consider how perceptual simulation supports displaced indexing on actions and interactions. Once the hunters have used perceptual simulation to identify individuals and to infer their expected behavior, they may start planning actions to hunt their prey as well as interactions with other hunters during the hunt. For example, imagine that one hunter's job is to circle around to the other side of the herd, thereby being in a position to encounter the antelope as they attempt their initial escape. If the hunters have performed or perceived this action before, they can use perceptual memories of it to simulate this part of the plan, thereby establishing a fairly precise idea of what is likely to happen. Furthermore, the hunters can discuss various parameters of the action that may need to be adjusted for this particular hunt. For example, because it is essential that the antelope not see, hear, or smell the hunter circling around them, the hunter must take a path that does not come too close to the herd or place him in a position that is upwind from it. On the other hand, the hunter does not want to take a path that is unnecessarily distant or the antelope may have moved on by the time he is in position.

By jointly simulating this action and by varying the parameters of the simulation, the hunters not only design the action optimally, they also coordinate their timing. Thus, simulations of interactions become relevant as well. Once the hunter who circles around is in place, his next planned action might be to mimic an animal call that signals completion of the action but that does not scare the antelope away. By jointly simulating these cooperative interactions with each other, the hunters can further design and coordinate their plan. In this manner, perceptual simulation supports preparation for action and interaction. It also supports monitoring of the plan during its execution. By running a simulation of how the plan should proceed, the hunters can determine if something has gone wrong so that adjustments can be made. Because the simulated plan is in the same representational format as the perceived event, the simulation is well suited to serving as a standard by which to evaluate the event.

Perceptual Simulations of Unfamiliar Situations

Finally, consider how perceptual simulation supports displaced indexing on componential experience. Again, imagine that a hunter is describing a herd of antelope to hunters who have never seen one. What might be a good strategy to optimize situated action for these hunters? The obvious answer is that the speaker should capitalize on what these hunters know to help them anticipate what to expect on

Copyright ©2000. All Rights Reserved.

encountering antelope. Perhaps these hunters have experienced similar animals, such as zebras. If so, the speaker can use language to initially establish simulations of zebras in the minds of his listeners. He can then use language to modify these simulations, using productive mechanisms for constructing complex hierarchical simulations (Barsalou, 1999). For example, the speaker could say that, rather than being black and white, the antelope are brown. Presumably these hunters have experienced brown, such that displaced indexing on similar experience can provide reasonably accurate information about color. Similarly, the speaker could say that antelope, unlike zebras, have horns and that they are somewhat smaller than zebras. Again, because the listeners presumably have knowledge about horns and relative size, they can adjust their simulations of antelope accordingly. The result is the simulation of an animal that they have perhaps never seen but that may be close enough to antelope to recognize them.

Once the speaker conveys some sense of what antelope look like, he could similarly convey a plan for hunting them. Again, he might begin by having his listeners simulate a plan they know and then having them adjust it accordingly by describing components that must be changed. Presumably, these adjustments succeed if listeners have experienced referents of these components before and can use them to revise their simulations appropriately. Most important, once the listeners have this plan in place, it may be accurate enough to execute a hunt that works. For all the reasons cited earlier, these hunters may experience some success in recognizing relevant objects, executing effective actions, and coordinating their interactions.

Expertise

Who would be more likely to carry out a plan for hunting antelope successfully, hunters who have hunted them before or hunters who have just heard about them for the first time? The answer is obvious. Hunters who have experienced the plan before can index its sensory-motor components with considerable precision, thereby producing simulations that guide action effectively. As these experts plan a hunt, linguistic descriptions of the plan's components index past sensory-motor experience directly from the relevant domain. As a result, the plan representation closely resembles what the hunters are likely to experience during the actual hunt, thereby facilitating their ability to monitor the situation. These hunters are much like Glenberg and Robertson's participants, who were able to integrate perceptual experiences of a compass with verbal instructions for using them.

In contrast, the inexperienced hunters, on hearing a description of a plan for hunting antelope, construct plan simulations that approximate the critical situation. Because the components of these simulations have been borrowed from other domains of situated action, they may not fit the target situation precisely. As a result, these hunters may fail to recognize key entities and events, or they may false alarm to incorrect entities and events. Furthermore, the description that

Copyright ©2000. All Rights Reserved.

they heard from the experienced hunter probably omitted key components that he forgot to mention.

Again, one might ask whether the sorts of examples provided here really apply to modern life. Consider some of the examples given before, such as handling a sinus infection in a child, finding an affordable house, and ensuring that remodelers proceed on schedule. In all of these situations, relevant physical entities must be recognized, inferences must be drawn about them, actions must be taken, and interactions must be managed. Thus, all of the same basic activities that characterize situated action in hunting antelope similarly characterize modern domains of situated action. Furthermore, perceptual simulation appears to provide a satisfactory means of supporting action in these situations as well. Consider handling a sinus infection in a child. If a child's symptoms match past perceptions of sinus infections (as opposed to colds or allergies), drawing the conclusion that the child has a sinus infection is warranted or, at the least, that it should be entertained. Should an intervention be attempted, establishing a plan can be readily envisioned through perceptual simulation. For example, one could imagine the procedure for administering antibiotics. The relevant interactions can be envisioned as well, such as imagining how to induce the child's cooperation. As this example illustrates, perceptual simulation provides just as much leverage on situated action today as it would appear to have provided long ago.

Simulation Richness and Inference

Although perceptual simulations support displaced indexing, their richness and accuracy may vary with the degree of displacement. As the situation to be simulated becomes increasingly displaced, it may become increasingly sketchy and inaccurate. Thus, simulations related to immediate situations should be the richest in accurate detail, followed by simulations of recently perceived situations, then simulations of situations perceived long ago, and finally simulations of unfamiliar situations, both past and future. As simulations become sketchier and less accurate, they should become decreasingly effective in supporting situated action. Agents should find it increasingly difficult to understand discourses about the situations and to formulate successful plans for acting in them.

The varying richness of perceptual simulations similarly explains minimal versus rich inference during comprehension (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992). Minimal inference occurs when comprehenders put minimal effort into constructing simulations that represent discourse. They simulate entities and events for salient words in the discourse but not much else. As a result, the simulation for the discourse is somewhat fragmented, containing a collection of somewhat coherent but disjoint simulations of discourse parts. When simulating an entity minimally requires an inference, one is made. For example, on hearing, "The surgeon scowled at the clock when she realized it was late," many comprehenders are garden pathed by *she* (Carreiras, Garnham, Oakhill, &

Copyright ©2000. All Rights Reserved.

Cain, 1996). If comprehenders must simulate a particular individual to minimally represent "surgeon," they must make a commitment to gender, given it is unusual and perhaps difficult to simulate a genderless person. Thus, the perceptual simulation view explains minimal inferences as those necessary to constructing minimal simulations.

In contrast, richer inferences result when comprehenders are motivated to construct more detailed simulations. Graesser et al. (1994) reviewed a wide variety of factors that induce comprehenders to produce richer inferences, including comprehension goals, high standards for coherence, and explanation seeking. For example, when comprehenders have the goal of determining which character in a mystery novel is the murderer, they make all sorts of inferences about the unstated motives and actions of characters. From the perceptual perspective, simulations of mental states and physical actions represent these inferences and become integrated into the larger simulation that represents the novel's meaning.

High standards for coherence can similarly be viewed as motivated comprehenders attempting to construct simulations whose components are well integrated rather than fragmented. If the simulation of a sentence's meaning cannot be integrated with the larger simulation that represents the text, comprehenders attempt to construct additional simulations that link them. Explanation seeking could similarly reflect the construction of linking simulations that integrate fragmented simulations of text components. As these examples illustrate, inferential richness can be viewed as how much effort comprehenders put into constructing detailed and coherent simulations of discourse meaning. Increasing evidence suggests that perceptual simulation is indeed central to comprehension (e.g., Black, Turner, & Bower, 1979; Bransford & Johnson, 1973; Gernsbacher, Varner, & Faust, 1990; Gibbs, 1994; Glenberg, Meyer, & Lindem, 1987; Glenberg & Robertson, in press; Intraub & Hoffman, 1992; MacWhinney, 1999; Morrow, Greenspan, & Bower, 1987; Potter & Faulconer, 1975; Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986; Potter, Valian, & Faulconer, 1977; Rinck, Hahnel, Bower, & Glowalla, 1997; Von Eckardt & Potter, 1985; Wilson, Rinck, McNamara, Bower, & Morrow, 1993).

ALTERNATIVES AND EXCEPTIONS

Amodal Eliminativism

Should one become convinced that situated action is more central to understanding language comprehension than is archival memory, one might nevertheless remain unconvinced that perceptual simulation is the best way to think about it. Why not simply adapt current amodal theories to explain comprehension in situated action? Why not assume that recognition, inference, action, and interaction are all represented by amodal propositions of the sort that we have all come to know

Copyright ©2000. All Rights Reserved.

and love (e.g., Barsalou, 1992)? Why not replace perceptual simulations with descriptions of them in amodal representation languages?

If we replace perceptual simulations with systems of amodal descriptions, we run into a long list of problems associated with this way of doing things (Barsalou, 1999). For example, what direct evidence supports the presence of amodal symbols in the human cognitive system? There is little if any direct evidence for such representations. Instead, researchers have adopted them primarily because they implement successful theories that explain important empirical findings. Certainly, this is no small achievement, and it suggests that amodal symbols might be implicated in cognition. Given the unlimited expressive power of amodal languages, however, together with their unfalsifiability, perhaps it is not so surprising that they can be used to implement successful theories. Because they can represent and explain anything, they can explain the comprehension of situated action. Again, though, it is essential to have direct evidence.

Many other problems also face amodal symbol systems. It is not clear how the amodal symbols purported to underlie cognition relate to perceptual states and to the world (Harnad, 1987, 1990; Searle, 1980). It is also not clear how amodal symbols emerged over the course of evolution or how (or where) they are implemented in neural systems. Furthermore, after decades of attempting to implement spatial and temporal constructs in amodal symbols, many knowledge engineers have become convinced that this is the wrong way to represent these constructs (e.g., A. Clark, 1997; Glasgow, 1993; McDermott, 1987; Winograd & Flores, 1987). For all of these reasons, adapting amodal views to explain the comprehension of situated action may not be a viable move.

Mixed Amodal and Perceptual Models

Perhaps amodal knowledge underlies our general meanings for words, with perceptual experiences of their referents becoming associated with them to provide indexing. This view faces two problems. First, all of the problems just mentioned for amodal symbols apply to the amodal side of mixed models. What direct empirical evidence do we have that such symbols exist in the brain? How are these symbols related to their perceptual counterparts? How did they evolve, and where do they reside in the brain? How do they represent the general meanings of words in the domains of space and time?

Second, what purpose do amodal symbol systems serve in such a system? Why include them in the first place? Why not just use perceptual simulations to keep things parsimonious? The obvious answer is that amodal symbols implement important functions that perceptual simulations cannot. This may be true, but the central argument of Barsalou (1999) is that perceptual symbol systems can implement all of the critical cognitive functions that amodal symbol systems have implemented traditionally. If so, then why did evolution add a redundant layer

Copyright ©2000. All Rights Reserved.

of amodal symbols? Of course this is possible, but again direct empirical evidence for such symbols is necessary to justify this conclusion.

Abstract Concepts

Another concern might be that some concepts do not seem associated with situated action, in particular, abstract concepts. However, Barsalou (1999) conjectured that all abstract concepts are situated and that situations are central to defining them. As Barsalou suggested, abstract concepts are not really abstract, they are simply complex and temporally extended. Whereas more concrete concepts index well-specified objects, actions, and properties in situations, abstract concepts index complex configurations of information distributed over multiple modalities and over time. On this view, abstract concepts are just as central to situated action as are concrete concepts; they simply refer to different, more complex aspects of situations.

Consider the abstract concept of *truth.* As Barsalou (1999) noted, truth is a polysemous concept. One core sense, however, refers to a situation in which an agent constructs an internal simulation and then compares it to a perceived situation. If the simulation fits the perceived situation, the simulation is a true construal of it. Thus, if a hunter states that the antelope herd he saw earlier in the day contains an albino, a listener who has not seen the herd would construct a simulation of this individual, assuming that he has seen an albino on another occasion or can construct one through productive simulation. The next day, on seeing the herd himself, the listener could determine whether it is true that the herd does indeed contain an albino. Presumably, he would do this by running the simulation he had constructed earlier and then comparing it to members of the herd, with the speaker's original statement being true if a matching individual were found.

The concept of truth is almost certainly a key component of most plans for performing situated action. Generally, all such plans contain steps like the following:

> If it is true that entity *X* is present, carry out action *Y*.

For example, carrying out the plan to hunt antelope is contingent on it being true that antelope are present. Similarly, as one hunter circles around the herd, the remainder of the plan is contingent on it being true that the hunter has completed this preliminary move successfully. As these examples illustrate, truth is not a concept that lies outside situated action. To the contrary, truth is central to it. Situated action could not occur in the absence of this concept. However, truth is not a simple circumscribed entity that can be easily pointed to in a situation. Instead, it is an aspect of situated action that is distributed across perceptual

Copyright ©2000. All Rights Reserved.

simulations, perceptions of physical situations, and introspective comparisons between them. It is a complex temporally extended property of situated action.

CONCLUSIONS

Some readers are probably wondering how any reasonable member of the community could arrive at such a naive, confused, and misguided view of language comprehension. Based on the assumptions, paradigms, and theories that dominate current thinking, such wonderment certainly seems reasonable. Conversely, however, a small but rapidly increasing number of researchers are coming to behold the dominant view with wonderment. How could we have ever held it?

Consider for the moment that the situated action view of comprehension is correct. If so, how could our original view have been so distorted yet so difficult to relinquish? Why did we not see the central importance of situated action much sooner? One conjecture is that our original view of comprehension, like our original view of cognition, has been shaped substantially by 20th-century developments in logic, statistics, and computer science (Barsalou, 1999). We have become so entranced with information processing and computational devices that they have come to totally permeate our thinking. Thus, in the study of language, we have come to believe that comprehension is essentially an archival process and not much else. Because of our foundational assumptions and frameworks, we do not see other deeply important dimensions of comprehension and cognition. In addition, our close associations to formal education and scientific archiving make an archival view of comprehension seem even more compelling.

What other foundational assumptions might we adopt that would motivate a shift in our view of comprehension? What other framework could cause us to perceive comprehension as preparation for situated action? Biology. Comprehension and cognition are at least as much biological phenomena as they are computational ones. Once we adopt the biological perspectives of evolution, ethology, and neuroscience, cognition and comprehension take on important new dimensions essential to understanding them. When we think about how cognition and communication evolved, how they manifest themselves in other species, and how neural systems and bodies implement them, we begin to conceptualize them much differently. There is no doubt that memory and information processing are central to comprehension. However, these computational phenomena are the outcomes of evolutionary history, species-specific pressures, neural mechanisms, and bodies. As we have seen, taking such factors into account can motivate new perspectives on familiar phenomena, such as the view that comprehension is grounded in perceptual simulations that prepare agents for situated action.

Should one find inspiration in this view, the articles by Glenberg and Robertson and by Roth point the way to go. They illustrate the type of research that will eventually allow us to understand comprehension as preparation for situated

Copyright ©2000. All Rights Reserved.

action. In particular, this research indicates that we should present participants with more than just words in our experiments. We should simultaneously present the objects, events, and agents that these words are about. Analogously, we should not just ask participants to reproduce the language they heard earlier, we should ask them to perform situated activity related to this language, including actions on objects and interactions with other agents. Should the resources of this research community be applied to such paradigms, I predict that we will observe dramatic growth in our understanding of language comprehension.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation Grant SBR–9796200 to Lawrence W. Barsalou. I am grateful to Arthur Graesser not only for the opportunity to write this commentary but for an inspiring initial exposure to cognitive psychology and cognitive science as his research assistant. I am also grateful to Arthur Glenberg, Harold Gouzoules, Arthur Graesser, Laura Namy, Robert McCauley, David Robertson, and Brian Ross for helpful comments on an earlier draft of this article.

REFERENCES

Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 27, pp. 1–64). New York: Academic.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Barsalou, L. W., Yeh, W., Luka, B. J., Olseth, K. L., Mix, K. S., & Wu, L. (1993). Concepts and meaning. In K. Beals, G. Cooke, D. Kathman, K. E. McCullough, S. Kita, & D. Testen (Eds.), *Chicago Linguistics Society 29: Papers from the parasession on conceptual representations* (pp. 23–61). Chicago: University of Chicago, Chicago Linguistics Society.

Bickerton, D. (1990). *Language and species.* Chicago: University of Chicago Press.

Black, J. B., Turner, T. J., & Bower, G. H. (1979). Point of view in narrative comprehension, memory, and production. *Journal of Verbal Learning and Verbal Behavior, 18,* 187–198.

Bransford, J. D., & Johnson, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383–438). New York: Academic.

Brown, A. L., Campione, J. C., Webber, L. S., & McGilly, K. (1992). Interactive learning environments: A new look at assessment and instruction. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 121–211). Boston: Kluwer Academic.

Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *Quarterly Journal of Experimental Psychology. A Human Experimental Psychology, 49A,* 639–663.

Clark, A. (1997). *Being there: Putting brain, body, and world together again.* Cambridge, MA: MIT Press.

Copyright ©2000. All Rights Reserved.

Clark, H. H. (1992). *Arenas of language use.* Chicago: University of Chicago Press.

Corballis, M. C. (1991). *The lopsided ape: Evolution of the generative mind.* New York: Cambridge University Press.

Corballis, M. C. (1998). Evolution of the human mind. In M. Sabourin, M. Robert, & F. Craik (Eds.), *Advances in psychological science: Vol. 2. Biological and cognitive aspects* (pp. 31–62). Hove, England: Psychology Press/Lawrence Erlbaum Associates, Inc.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition.* Cambridge, MA: Harvard University Press.

Donald, M. (1993). Precis of "Origins of the modern mind: Three stages in the evolution of culture and cognition." *Behavioral and Brain Sciences, 16,* 739–791.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 430–445.

Gibbs, R. W., Jr. (1994). *The poetics of mind: Figurative thought, language, and understanding.* New York: Cambridge University Press.

Glasgow, J. I. (1993). The imagery debate revisited: A computational perspective. *Computational Intelligence, 9,* 309–333.

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

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26,* 69–83.

Glenberg, A. M., & Robertson, D. A. (in press). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language.*

Gouzoules, H., & Gouzoules, S. (1995). Recruitment screams of pigtail monkeys (*Macca nemestrina*): Ontogenetic perspectives. *Behavior, 132,* 431–450.

Gouzoules, H., Gouzoules, S., & Ashley, J. (1995). Representational signaling in non-human primate vocal communication. In E. Zimmermann, J. D. Newman, & U. Jurgens (Eds.), *Current topics in primate vocal communication* (pp. 235–252). New York: Plenum.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53,* 5–26.

Harnad, S. (1987). Category induction and representation. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 535–565). New York: Cambridge University Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42,* 335–346.

Intraub, H., & Hoffman, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. *American Journal of Psychology, 105,* 101–114.

Johnson, M. (1987). *The body in the mind: The bodily basis of reason and imagination.* Chicago: University of Chicago Press.

Johnson-Laird, P. N., Herrmann, D., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychological Bulletin, 96,* 292–315.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

MacWhinney, B. (Ed.). (1999). *The emergence of language.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McDermott, D. (1987). A critique of pure reason. *Computational Intelligence, 3,* 151–160.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.

Mithen, S. (1996). *The prehistory of the mind: The cognitive origins of art, religion and science.* London: Thames & Hudson.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26,* 165–187.

Newton, N. (1996). *Foundations of understanding.* Philadelphia: Benjamins.

Copyright ©2000. All Rights Reserved.

Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature, 253,* 437–438.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115,* 281–294.

Potter, M. C., Valian, V. V., & Faulconer, B. A. (1977). Representation of a sentence and its pragmatic implications: Verbal, imagistic, or abstract? *Journal of Verbal Learning and Verbal Behavior, 16,* 1–12.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37–75). Boston: Kluwer Academic.

Rinck, M., Hahnel, A., Bower, G. H., & Glowalla, U. (1997). The metrics of spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 622–637.

Ross, B. H. (1996). Category learning as problem solving. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 35, pp. 165–192). San Diego, CA: Academic.

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes.* New York: Oxford University Press.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3,* 417–424.

Smith, W. J. (1977). *The behavior of communicating: An ethological approach.* Cambridge, MA: Harvard University Press.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–552.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.

Von Eckardt, B., & Potter, M. C. (1985). Clauses and the semantic representation of words. *Memory & Cognition, 13,* 371–376.

Wilson, S. G., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1993). Mental models and narrative comprehension: Some qualifications. *Journal of Memory and Language, 32,* 141–154.

Winograd, T., & Flores, F. (1987). *Understanding computers and cognition: A new foundation for design.* Reading, MA: Addison-Wesley.

Copyright ©2000. All Rights Reserved.