9

Explanations in Gesture, Diagram, and Word

BARBARA TVERSKY, JULIE HEISER, PAUL LEE, and MARIE-PAULE DANIEL

9.1 Beginning: Characterizing Explanations

People are constantly explaining things to one another. Parents explain to children how to build a tower of blocks or why they need to go to bed. Friends explain to each other how to find their homes or why they were late. Teachers explain to students how nerve conduction happens or why the World War I began. In person, explanations, in common with most face-to-face communication, are typically multimodal. Not simply talk, explanations include gestures, props, and often diagrams (e.g. Engle, 1998). Each mode has relative advantages and disadvantages, and they work in concert, complementing and supplementing one another (e.g. Clark, 1996; Goldin-Meadow, 2003). Frequently, explanations are restricted to a single mode. Giving directions over the phone limits them to words. Providing instructions to operate a camera or assemble a piece of furniture to international consumers limits them to diagrams. Although gesture is often used alone for brief interchanges, like signalling to a partner a desire to leave a party, it is less likely to be solely used in complex explanations. Those who have got lost in a country whose language they don't know learn the power of gesture alone. Limiting communicators to a single mode requires completeness of that mode. This reveals the structure of explanations and the parallel ways each mode expresses that structure.

Not all explanatory tasks readily lend themselves to words, diagrams, and gestures. Fortunately, two of the most common kinds of explanatory tasks do: navigation and construction. One of each was selected in order to investigate the structure and semantics of explanations. The navigation task was to communicate how to get from one place to another. The construction task was to communicate how to put together a piece of furniture, a TV cart. The data analysed here were gathered from several experiments differing somewhat in methods. In all cases, participants first learned or already knew the specifics of the task and then provided instructions so that another person could perform the task. In some experiments, participants were asked to provide verbal instructions,

in others, diagrammatic instructions, and in others, explanations using gestures with props, either with or without words. In still other experiments, participants could use various combinations of modes, such as diagrams and words or gestures and words.

The two tasks, constructing an object and learning a route, are representative of the kinds of tasks people learn in ordinary as well as technical situations. Both tasks require explaining actions in space. They both stipulate the arrangement of parts in a particular spatial-temporal configuration. As such, they are similar to tasks such as making pot-au-feu or operating a voting machine or understanding how the heart works or performing surgery or maintaining a power plant. At the same time that they are representative of tasks requiring knowledge of the spatial-temporal organization of parts, they are readily learned by novices in a laboratory session. Route learning and assembly tasks have been studied extensively from different perspectives (e.g. Allen, 2000; Denis, Pazzaglia, Cornoldi, and Bertolo, 1999; Novick and Morse, 2000; Shi and Tenbrink, this volume; Striegnitz, Tepper, Lovett, and Cassell, this volume; Tversky, Agrawala, Heiser, Lee, Hanrahan, Phan, Stolte, and Daniel, 2007). Because these tasks entail explanations of the spatial configuration of parts in a temporal sequence, they are likely to elicit numerous gestures (e.g. McNeill, 1992; Krauss, Dushay, Chen, and Rauscher, 1995; Wagner, Nusbaum, and Goldin-Meadow, 2004).

Like narratives, explanations have a discourse structure consisting of a beginning, middle, and end. For explanations, the beginning is an introduction, the middle, a step-by-step set of procedures (some with qualifications and embellishments), and the end, an indication that the task has been completed. This discourse structure has been observed in spontaneous verbal instructions (e.g. Denis, 1997) and in diagrams as well (Tversky et al., 2007). The semantics crucial for explanations includes referring expressions for objects (or object parts) and for actions. Objects are typically static whereas actions are dynamic. Of interest here is how the discourse structure and the semantics are expressed visually, in diagrams or gesture.

Of the three modes of communication, words are perhaps the most common. Although words are purely symbolic, there are a great many of them, allowing nuanced expression of concepts, actions, and relations. Words can do a number of things that are difficult to do in pictorial or visual media; for example, words can be used to qualify, to negate, and to hypothesize. Diagrams are also familiar modes of communication, most certainly maps for navigation and diagrams for construction. Diagrams have a number of advantages over words. Words are purely symbolic, whereas diagrams bear visual similarity to what they are meant to communicate, both objects and actions. Diagrams can depict objects and object parts, their structural relations, and even their manner of assembly to convey construction. Diagrams can indicate landmarks by names or icons and can schematize the turns that constitute a route by a turning line to convey a route.

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 121 of 131 August 26, 2008 18:13

Gestures, like diagrams, are a visual mode of communication. They serve a number of roles in discourse as well as in thinking, that is, for those using gestures as well as for those observing them (e.g. Goldin-Meadow, 2003; Kessell and Tversky, 2005; Krauss, 1998; McNeil, 1992; Sowa and Wachsmuth, this volume). Several taxonomies of gestures have been developed (e.g. Ekman and Friesen, 1969; Kendon, 1988; McNeil, 1992; Rimé and Schiaratura, 1991). Some gestures serve to structure the discourse, for example, using the hands to separate arguments, often accompanied by 'on the one hand' and 'on the other hand'. Others, termed emblems, have conventional meanings like words, for example, waving goodbye or the sign for *OK*. Here the focus is on two other types of gestures, those central to explanations, specifically, the deictic and iconic gestures that carry semantic meaning and gestures that form the discourse structure characteristic of explanations. On the whole, *deictic* gestures point to or indicate things in the environment; *iconic* gestures resemble what they are meant to convey. Some gestures function by interacting with things in the world, by referring to and acting on things in the environment. In the present cases, the things are maps for route directions and object parts for assembly instructions. Gestures, then, are inherently both embodied and situated.

Partly because they are both situated and embodied, gestures can support or even convey explanations in a rich set of ways. Some gestures, like *deictics*, can refer to particular aspects of a situation rather than others, directing and focusing attention on the critical aspects of a situation. Unlike most words, gestures, specifically *iconic* gestures, may bear physical similarities to the things they refer to, for example, illustrating a T intersection by making a T with the hands; that physical resemblance may make gestures easier to interpret and easier to remember. Note that the T may be conveyed in gesture in different ways, by the whole hands, by two fingers, by writing a T in the air. At the level of semantics, the particular hand and finger positions can vary considerably and still convey the same meaning. Because gestures occur in a spatial medium, and because people think and talk about many abstract concepts in spatial metaphors, gestures can also bear *metaphoric* relations to the things they represent. For example, when telling a friend that she and another friend had had little recent contact, one person said, 'We've grown apart', while separating her two hands. The metaphor of spatial distance representing psychological distance was expressed symbolically by the words, but concretely by the hands. The actual distance between the hands can be used to convey the degree of separation, an aspect of communication not conveyed in the words. Gestures share both these qualities, iconicity and metaphoricity, with diagrams. Part of the effectiveness of diagrams derives from physical and metaphoric similarities to the things they represent (e.g. Tversky, 1996, 2001). When gestures are redundant with speech, they provide a second way of encoding information, in addition to words. For memory, two codes are better than one (e.g. Paivio, 1986). What's more, for concepts that can be depicted, a pictorial code is superior to a verbal one, presumably because of the resemblance of pictures to the things they

represent. By analogy, iconic gestures should have an advantage over arbitrary words.

Gestures can do more than depict, they can also enact. When describing how he swerved his car to avoid a potential collision, one speaker swayed his body deeply to show the swerve. Notice that it was the car that swerved, not the body, so the swerving was metaphoric, though both passengers and drivers are known to swerve when concerned about collisions. Swerving and other demonstrations of action, such as showing another's strutting or demonstrating a knot, *embody* the knowledge they are meant to convey, coding it motorically as well as pictorially. Motor codes, like pictorial codes, are also known to augment memory (e.g. Engelkamp, 1998). Finally, gestures can situate knowledge in the relevant context, by using actual or virtual props. When describing how to make a pendulum from a rope hanging from the ceiling, one speaker first used his hand to indicate the position of the rope, then to indicate the weight to be tied on the end of the rope, then to show the tying, then to indicate grasping the rope and setting it in motion, and, finally, to indicate the swinging of the rope (Kessell and Tversky, 2005). It is remarkable how many different meanings the hand was used to convey, smoothly and effortlessly. In situating the knowledge, gestures can create a mental model of the objects and the space as well as of the actions (e.g. Emmorey, Tversky, and Taylor, 2000). Because they are spatial and (in the case of gestures) temporal, diagrams and gestures can convey models, structural, behavioural, or causal more directly than words.

9.2 **Explaining How to Put Something Together**

9.2.1 Production of assembly explanations

In the experiments eliciting explanations of assembly, participants first assembled a TV cart on their own, using the photograph on the carton as a guide (e.g. Heiser, Phan, Agrawala, Tversky, and Hanrahan, 2004). The TV cart was 17″ x 25″ × 21″ in size, and consisted of two sideboards, an upper shelf, a lower shelf, a support board, pegs for attaching the support board, screws, screwdriver, and wheels. After assembling the TV cart, participants in the various studies were asked to design instructions explaining how to assemble the TV cart. Some groups produced diagrams and words, some only diagrams; still other groups made explanatory videos as they reassembled, free to use both speech and gesture, while other participants were asked to make explanatory videos as they reassembled, but told that the videos didn't have a sound track, and might be used by people who don't speak English. Here, the nature of the verbal, diagrammatic, and gestural explanations is analysed. The focus is on gestures used in explanations, as that work is new.

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 123 of 131 August 26, 2008 18:13

9.2.2 Gestures

Narrative Structure. Explanations, like other forms of discourse, have a narrative structure, notably, a beginning, middle, and end. Think of recipes: they begin with the name of the dish to be cooked, and then the list of ingredients, typically in order of use. They continue with the list of procedures in the order to be followed. The end of a recipe is the outcome, typically the number of people it will serve. Explanations of TV cart assembly and route directions also frequently had a discourse structure. That structure was often carried by gestures, especially in the gesture alone group. A variety of gestures served narrative roles, some used by some participants, others by others. Most participants began with some sort of an introduction. Many waved to the camera by way of introduction. Others began as a recipe begins, by presenting the parts to be used to construct the whole object. The narratives continued with a step-by-step explanation. Frequently, the steps themselves had an internal structure, a beginning, middle, and end. Finally, the explanations often used gestures to signal the end of the narrative, that the task had been accomplished. Some of the devices used to carry the narrative and convey the explanation are described in the following sections, beginning with presenting parts. Parts were often presented as part of the opening of the entire explanation and also as the opening of each step.

Introduction: Presenting Parts. Presenting parts typically took one of two forms, exhibiting and pointing. Both of these involve *deictic* gestures whose primary function is to point to or indicate something in the environment. Typically, the large parts were exhibited by holding them up to viewers, and the small parts were pointed to. Parts were presented not only as an introduction to the entire task but also as an introduction to each step.

Conveying Steps. After presenting the parts, the explanatory narrative typically continued by providing each assembly step in turn. Discourse structures are often hierarchical, and these were no exception. At the more general level, the discourse structure of procedural explanations has a beginning, a step-by-step specification, and an ending. But each of these discourse parts may have subparts, especially the steps that constitute the middle. Each step had a beginning, middle, and end. Those who could speak used language to convey narrative structure. For those discouraged from speaking, a variety of gestures served to convey narrative structure. Gestures were used to mark the beginning and end of each step, and to explain each step by demonstrating it. The beginnings of steps were often marked by holding up a finger to indicate the step number, and the endings by a gesture such as 'OK' or 'thumbs up', or flattening the two hands palms down and moving them away from the body and outwards as if to push away to indicate 'done'. Perhaps because of the relative ease of marking steps in language, using words like 'next', 'after that', or 'now', speakers marked steps explicitly more often than those restricted to gesture. Gestures were also used at the beginnings of steps to present the parts that would be used for that step.

Models of Structure and Action. Another way that steps were introduced was by previewing the structural change or the action needed to accomplish the structural change. In gesture, this was accomplished by a gestural model. A little-noted feature of gestures is that sequences of related gestures can be used to convey a mental model (Enfield, 2003; Engle, 1998; Emmorey, Tversky, and Taylor, 2000). For example, in describing environments, participants used as many as 15 related gestures in a sequence to locate landmarks in space, maintaining a spatial display 'drawn' in the air, with consistent size and position (Emmorey et al., 2000). In order to explain the assembly of the TV cart, participants often made models using gestures, frequently prior to an assembly step. For TV cart assembly, the models often used parts of the TV cart as props. The models were of two types, structure or action. For structure models, explainers used their hands, sometimes with the parts, to show the desired structural relations among the parts, for example, using the palm flattened horizontally to indicate the top of the cart, and both hands flattened vertically to indicate the sideboards. For action models, explainers used their hands, sometimes with the parts, to show the actions required to achieve the desired structure, for example, holding real or imaginary parts and moving them in place. Both structure and action models used a combination of *deictic* gestures that point to things in the environment and *iconic* gestures that bear resemblance to what they are communicating. The demonstrations of action, for example, were iconic gestures, as were sequences of gestures showing the shape of the object.

Ending. Participants often took care to indicate when the task was done, even though the completed cart was a self-evident ending. One way they did this was by a gesture of presentation. In contrast to the step-ending gesture, the palms of the two hands faced upwards as the hands unfolded before the completed cart. Other endings included an 'OK' sign or thumbs up. The sequence of gestures, then, can tell a story, a particular kind of story, namely an explanation. In this case, gestures were used to indicate a beginning, often including a greeting, commonly with a presentation of the parts to be assembled. Gestures marked assembly steps, and introduced them with a demonstration of the action to be performed. They marked ends of steps. Gestures can also end the story, often by presenting the completed object.

It should be remembered that not all of the participants used gestures to indicate all the components of the narrative. Gesture narratives were more frequent among explainers restricted to gestures. For them, the gestures carried the entire message. Those using only gestures acted as if they needed to create a coherent and integrated set of gestures in order to convey the assembly task. So, for example, if they used a gesture to introduce the task, they also tended to use a gesture to end the task, often gestures that were related, such as waving 'hello' at beginning and waving 'goodbye' at the end. Just as in speech words are related and integrated, so for gestures in those not allowed to speak. For those who could speak, the language

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 125 of 131 August 26, 2008 18:13

could carry most if not all the message, so the gestures were more optional, and did not have to be integrated into a coherent set.

Action gestures are also actions. As such, they *embody* the information that is crucial for task performance. They also *demonstrate* the information that is crucial to performance. In addition, the action gestures *situate* the information necessary for task performance: they were formed with respect to virtual or actual parts and performed with respect to a virtual or actual object. These properties are likely to render gestures as especially important to communication and comprehension, especially for information literally or metaphorically about spatial relations and about action. Gestures can directly map spatial relations to be conveyed, such as the structure of an object; they can also directly map actions to be comprehended, such as moving or rotating parts. Diagrams also have a special status in communication and comprehension because they too can map spatial relations in the information to be conveyed to spatial relations on paper. Diagrams also have techniques for conveying actions, notably arrows (e.g. Tversky, Zacks, Lee, and Heiser, 2000; Tversky, Heiser, Lozano, MacKenzie, and Morrison, 2007). However, the techniques for conveying action using gesture are more direct, just as the techniques for conveying structure are more direct in diagrams. Now we turn to examine briefly how explanations are accomplished in another visual medium, diagrams.

9.2.3 Diagrams

The diagrams explaining construction produced by users also had a narrative structure, especially those drawn by people high in spatial ability (Tversky et al., 2007). Like a recipe, the beginnings of explanatory diagrams were often a list of ingredients, namely the parts to be used in construction of the TV cart. As for the explanations in gesture, diagrammatic explanations typically had a middle consisting of the sequence of steps needed to assemble the object. The better diagrams showed each step in the perspective needed for assembly, and embellished the depictions of the object parts with arrows and guidelines indicating how the parts should be moved into position, that is, the better diagrams showed the actions needed to put the parts together, not just the structure of the parts. Each new step was a new part to be attached. Finally, diagrammatic explanations often used a deliberate ending. In many cases, participants drew lines surrounding a sketch of the completed TV cart, like rays around a sun.

9.2.4 Words

Just as for explanations relying on gestures and explanations relying on diagrams, explanations relying or using words also had a narrative structure (Daniel, Tversky, and Heiser, 2006). Most participants introduced the task, many as in

the gestures and diagrams, by a listing of the parts. Others began by providing a structural mental model how the parts fit together to make a whole. Still others gave general advice. Some participants used no special beginning, but rather started right in with the step-by-step instructions. The middle, as for gestures and diagrams, was a hierarchical set of step-by-step instructions that specified the actions and subactions to be taken on each object part, that is, actions on objects or object parts. The instructions also often included qualifications, for example, perceptual details that helped to identify the relevant part or action details that specified the manner of action. The steps were often explicitly marked, as in 'first', 'next', and 'finally'. Endings varied, sometimes simply saying the task is now finished, sometimes adding a suggestion about how it can now be used or suggesting that the user can be proud of finishing.

9.2.5 Assembly instructions and mental model of assembly

The three modes of communicating share a number of features that suggest that they also share the same underlying discourse structure and the same underlying mental model of assembly. Spontaneously produced explanations, like stories, typically had a narrative structure; they were not merely a list of a steps but an integrated list sandwiched between a beginning and an end, an introduction to the task and an indication that the task was completed. The heart of the instructions was the set of actions on objects, whether expressed gesturally, diagrammatically, or verbally. The actions on objects were the goals and subgoals that constituted the task rather than motions. That is, rather than 'slide the shelf horizontally', participants said or showed the goal of the sliding. Typically, each step involved a new object part. Thus, the mental model of assembly is a hierarchical set of actions and subactions on objects or object parts. These insights have come from people's production of procedures for carrying out an organized set of actions with a beginning, a middle, and an end, which is an accomplishment or achievement. The processes that complement production are perception and comprehension of organized sets of actions, having beginnings, middles, and accomplishments, termed *events*. Research on event perception and cognition have revealed the same mental model as the present research on production of assembly instructions, that is, events are perceived and conceived to be a hierarchical set of actions on objects accomplishing goals and subgoals (e.g. Tversky, Zacks, and Martin, 2008; Zacks, Tversky, and Iyer, 2001).

9.3 **Explaining a Route**

Wayfinding is a skill that precedes humankind, indeed, remarkably so. Think of migrating butterflies, ants, fish. And *Lassie Come Home*. What humans have added is ways to communicate routes to others, including describing them, sketching

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 127 of 131 August 26, 2008 18:13

them, or gesturing them. Anyone who has got lost in a country with an unknown language has had to rely on the gestures of foreigners to find their way. The structure of routes, whether described, gestured, or depicted, consists of a sequence of actions, typically turns, on paths at landmarks (Denis, 1997; Tversky and Lee, 1998, 1999). How are these components, and others, of routes communicated in each mode?

9.3.1 Gestures

Participants were given four maps (London, Paris, Palo Alto, San Francisco Bay Area), each with a start point and an endpoint of a highlighted route. They were asked to devise a route from the start point to the endpoint, and to explain that route to a camera while making the map visible to the camera so that someone else viewing the video could find their way. Some of the participants were free to use gesture and speech and others were told that the sound track was off because viewers might not understand English.

In explaining how to assemble the TV cart, communicators used deictic gestures to indicate parts and iconic gestures to demonstrate actions. For explaining routes, communicators used both deictic and iconic gestures. Deictic gestures were frequently used to point to places; iconic gestures were often used to demonstrate action, notably turns. Here, communicators demonstrated action using their hands rather than their feet. Since communicators had maps, they could gesture on the map by pointing to landmarks where turns occur and tracing the paths. Gestures like tracing the path are iconic with respect to the map but metaphoric with respect to the action.

Narrative Structure. As for assembly explanations, the gestures used in route explanations, especially by those restricted to gesture, often had a narrative structure. Those restricted to gesture tended to produce an integrated, related set of gestures. For a *beginning*, many communicators greeted the recipients by waving; others pointed vigorously at the start point. Some pointed first to the viewer, then to the start point, a way of orienting the viewer in the map, analogous to 'you are there' indicators. The *middle* was typically a step-by-step explanation of the route, consisting primarily of deictics to indicate locations, landmarks, or turning points and iconics to indicate form of paths or turns.

Metacomments. Explanations often include a variety of metacomments in addition to the step-by-step information. Remember that such comments were common in the verbal instructions for assembling the TV cart. Gestures were used in these. For example, on some occasions, gesturers explicitly chunked the information, first showing a few turns one by one, and then reviewing a group of them. Gesturers also made off-map gestures to clarify components of the route. They sometimes included iconic gestures specifying landmarks, such as a series of gestures that conveyed a hotel, first by outlining a building, then by pantomiming its function, for sleeping. They also used off-map gestures to

clarify a route, switching from the survey or overview perspective of the map to the route or embedded perspective of the traveller (cf. Taylor and Tversky, 1992).

Endings. Finally, gesturers ended their narratives most typically with an 'OK' or thumbs up or presentation gesture, both palms out and upwards as if presenting the explanation to the viewer.

9.3.2 Diagrams

A number of years ago, university students were approached outside a dormitory just before dinner hour and asked if they knew the way to a local restaurant (Tversky and Lee, 1998, 1999). If they did, they were asked to either sketch a map or to write directions to the restaurant. Both sketches and directions were analysed. A single diagram was used, and the beginning, the start point, the middle, or the set of turns at landmarks, and the ending, the restaurant, were marked. Notably, although the sketch maps could have been analogue, reflecting actual distances and angles, they were instead schematic; that is, they did not show exact metric relations, either distance or angle. They left out streets and landmarks not directly involved in the route, retaining primarily the streets and landmarks of the route. Turns, irrespective of actual angle, were shown as right angles. Short distances with complex actions were relatively enlarged whereas longer distances with no change of action were relatively reduced. Thus, the diagrams were distorted in ways that made the route more readily apparent.

9.3.3 Words

The language of route maps parallelled the sketch maps. Denis and his colleagues (Denis, 1997; Denis, Pazzaglia, Cornoldi, and Bertolo, 1997) have analysed the language of routes, and our results replicate those. Those directions were given in response to a question about how to get to a destination, so the start points and endpoints, the beginnings and ends, were already established and did not need to be communicated. Nevertheless, route directions often had an explicit *beginning*, orienting the traveller. The *middle* consisted of iterations of locating a landmark and specifying an action. Landmarks are typically referred to by names, names of buildings or of street intersections. Actions, typically turns, are conveyed by terms such as 'turn right', 'take a right', or 'make a right'. Other actions, typically progressions, were referred to as 'go down' for straight paths or 'follow around' for curved paths. Note that the descriptions of actions, parallelling the sketch maps, did not specify the exact angle of turn or the exact distance of procession. As for the verbal assembly instructions and the gestural route directions, the verbal route directions often included redundancies and metainformation, such as using more than one perspective, route, and survey, local summaries of

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 129 of 131 August 26, 2008 18:13

set of turns, specifications of landmarks and warnings of pitfalls. Finally, route directions often explicitly marked an *ending*, such as 'there you are' or 'now you've arrived'.

9.3.4 Route directions and mental models of routes

As for assembly, the data from the three modes of communication of routes suggest a common underlying mental model for routes: a sequence of turns or actions at landmarks where distance and angle are schematized, that is, not specified. Correspondingly, communications of routes, whether by gesture, diagram, or word have beginnings that orient a recipient, middles that provide a step-by-step set of actions at landmarks, and ending that indicate arrival.

9.4 **End: Modes of Explanation**

Explanations are a common form of discourse. Like stories and expository prose, the present kind of discourse, explanations are structured, with beginnings, middles, and ends. Yet stories differ from explanations in other ways. Whereas stories typically have a narrative voice, explanations typically do not. Stories are usually about life events of people, explanations about possible events of systems, objects, or people. As such, stories are typically imbued with emotion whereas explanations are not. Stories tend to be episodic, explanations semantic, in Tulving's (1972) sense of the terms. Here we have examined two paradigmatic kinds of explanation, how to put something together and how to get from here to there.

We have compared and contrasted spontaneous explanations that use one or more communication mode, gesture, diagram, or word. The explanations were of actions in space, putting something together or navigating a route. Each had objects, landmarks or parts, and actions on or at them. The exact form of the beginnings, middles, and ends varied, depending on the affordances of the communication mode adopted. For each kind of explanation and each mode, the beginning presumed an initial state of ignorance and a goal to learn the task. Correspondingly, beginnings for the assembly task sometimes greeted the recipient of the communication, sometimes overviewed the task, and sometimes presented a menu of parts. For the route task, beginnings often greeted the recipient or oriented the recipient in the environment. The middle consisted of a set of steps or procedures, often explicitly marked, that were actions on objects for assembly or actions at landmarks for routes, that is, actions with respect to a spatial configuration. Finally, explanations ended with some indication that the task was completed.

All modes, gesture, word, and diagram, appear to serve thought as well as communication. People have been known to talk to themselves, to think aloud,

especially in the shower. Designers, mathematicians, scientists are often at a loss without the proverbial cocktail napkin to sketch on. Indeed, it has been proposed that designers use sketches to hold conversations with themselves (Schon, 1983; also Goel, 1995; Goldschmidt, 1991). In fact, designers, especially experienced ones, get new ideas from inspecting their own sketches; they 'see' new and unintended relations, patterns, functions that emerge from the sketches (Suwa, Tversky, Gero, and Purcell, 2001). Interestingly, when architects are asked to design while blindfolded, they gesture profusely, as if the gesturing replaces the sketching they normally do (Z. Bilda, personal communication, 2005). Diagrams are thought to be effective for communicating a broad range of ideas for a number of reasons (Tversky, 2001). Diagrams use elements that may bear a physical or metaphorical relation to what they represent, for example, icons in software or airports. They use spatial relations to convey relations that are directly spatial, such as maps, or metaphorically spatial, such as degree of productivity or attractiveness. Communications that are iconic or metaphoric are more direct than purely symbolic communications. Indeed, diagrams (and gestures) can often be understood in situations where the local language is not. Diagrams encourage completeness of thought; unlike words or gestures, they do not readily tolerate ellipsis (e.g. Tversky and Lee, 1999).

Gestures, like diagrams, are effective in part because their relationship to meaning is more direct, less mediated. In addition, and in contrast to words and diagrams, gestures can embody the knowledge they are meant to convey. This is particularly true for action information as gestures are frequently mini-actions, such as pushing and pulling and putting. There is evidence that embodiment itself is privileged for learning action information, even when retrieval is verbal (Engelkamp, 1998). Finally, gestures situate knowledge in the world in which it will be used. They do this by pointing to the objects and the places that they hold or will hold with respect to other objects. In the present cases, situating knowledge was facilitated by the props provided, the object parts or the paper map. However, in cases without props, gestures are often used to establish virtual or proxy props, and other gestures act on them. For example, in describing environments without props, people frequently created a map in the air, adding landmark after landmark to the appropriate place in the imagined map (Emmorey et al., 2000). Situating knowledge embeds it in a rich structure, part of which may already be known, that provides a scaffold for new information. As for many cognitive phenomena, then, the advantage of gesture for both communicators and recipients has more than one cause. We have proposed four—that (deictic) gestures draw attention selectively to the critical aspects of the message, that (iconic and metaphoric) gestures bear a literal or figurative likeness to what they convey, that gestures embody knowledge, and that they situate knowledge in the world—but there may be others.

Gestures have advantages, but they are not, per se, a complete language, though they can evolve into one (see Goldin-Meadow, 2003). Words, of course,

978–0–19–955420–1 09-Coventry-c09 OUP339-Coventry (Typeset by SPi, Delhi) 131 of 131 August 26, 2008 18:13

are, though it must be noted that spontaneous spoken words typically lack the completeness that written text or prepared talks have. For some tasks, especially explanations of routes and assembly, diagrams are complete and sufficient. Yet, despite completeness, combining these modes seems to create the most effective explanations as each has something special to contribute. People have a great many stories to tell, expositions to relate, and explanations to convey; fortunately, they have a rich set of means of expression to do so.