Spatial perspective choice in ASL

Karen Emmorey and Barbara Tversky
The Salk Institute for Biological Studies / Stanford University

Two studies investigated the ramifications of encoding spatial locations via signing space for perspective choice in American Sign Language. Deaf signers ("speakers") described the location of one of two identical objects either to a present addressee or to a remote addressee via a video monitor. Unlike what has been found for English speakers, ASL signers did not adopt their addressee's spatial perspective when describing locations in a jointly viewed present environment; rather, they produced spatial descriptions utilizing shared space in which classifier and deictic signs were articulated at locations in signing space that schematically mapped to both the speaker's and addressee's view of object locations within the (imagined) environment. When the speaker and addressee were not jointly viewing the environment, speakers either adopted their addressee's perspective via referential shift (i.e., locations in signing space were described as if the speaker were the addressee) or speakers expressed locations from their own perspective by describing locations from their view of a map of the environment and the addressee's position within that environment. The results highlight crucial distinctions between the nature of perspective choice in signed languages in which signing space is used to convey spatial information and spoken languages in which spatial information is conveyed by lexical spatial terms. English speakers predominantly reduce their addressee's cognitive load by adopting their addressee's perspective, whereas in ASL shared space can be used (there is no true addressee or speaker perspective) and in other contexts, reversing speaker perspective is common in ASL and does not increase the addressee's cognitive load.

Keywords: spatial perspective, shared space, classifiers

1. Introduction

When English speakers talk about locations within a scene that they are both viewing, they are often faced with a coordination problem with respect to reference frames. That is, speakers must choose a particular spatial perspective from which to describe the locations of objects. What is on the left from one person's point of view might be
in front of or on the right from another person's point of view. For example, speakers can describe object locations from their own viewpoint, e.g. "Pick the one that's farthest from me" or "It's the one on my right," or they can adopt the point of view of their addressee, saying "Pick the one closest to you," or "It's the one on your left." Speakers may also adopt a more neutral perspective by describing object locations with respect to other objects or landmarks, e.g. "Pick the one near the water cooler" or "It's the one next to the door." In the experiments reported here, we investigate the nature of spatial perspective choice in American Sign Language (ASL).

Signed languages generally convey spatial information using "classifier" constructions in which spatial relations are expressed by where the hands are placed in signing space or with respect to the body (e.g. Supalla 1982; Engberg-Pedersen 1993). Classifier predicates are complex forms in which the handshape is a morpheme that encodes information about object type, and the movement and position of the hand may specify the movement and location of a referent (see papers in Emmorey in press a, for an in-depth discussion of classifier constructions in various signed languages). For these constructions, there is a schematic correspondence between the location of the hands in signing space and the position of physical objects in the world (e.g. Emmorey & Herzig in press). When describing spatial scenes, the identity of each object is usually indicated by a lexical sign (e.g. TABLE, T.V, CHAIR). The location of objects, their orientation, and their spatial relation vis-a-vis one another is indicated by where the appropriate classifier predicates are articulated. Where English uses prepositions and directional terms to express spatial relations, ASL uses the visual layout displayed by classifier signs positioned in signing space. Figure 1 provides an illustrative example of a room description in ASL.

In the example in Figure 1, neither the addressee nor the signer is currently observing the room that is being described. In this situation, spatial descriptions are almost always produced from the signer's (the "speaker's") perspective. "Speaker" will be used here to refer to the person who is signing, in order to parallel the speaker/addressee contrast for English. In Figure 1, the speaker describes a scene where a table is on his left as he enters a room. He uses the sign glossed as ENTER at the beginning of the discourse which signals that the scene should be understood from his perspective. The speaker indicates that the table is to the left by producing the classifier sign appropriate for tables at a spatial location on his left. Because the addressee is usually facing the speaker, the spatial location for table is actually positioned on the addressee's right. There is a mismatch between the location of the table in the room being described (the table is on the left as seen from the entrance) and what the addressee actually observes in signing space (the classifier sign for table is produced to the

---

1. The status of handshape as a classifier in these constructions has been recently called into question (see papers in Emmorey in press a).
situation depicted, the speaker (the signer) and addressee are facing each other, and between them are two boxes. Suppose the box on the speaker's right is the one that he wants to identify. If the speaker uses signing space (rather than just pointing to the actual box), he would indicate the desired box by placing the appropriate classifier sign on the right side of signing space. Note that in this situation, no mental transformation is required by the addressee. The speaker's signing space is simply "mapped" onto the jointly observed physical space — the right side of the speaker's signing space maps directly to the actual box on the left side of the addressee (see Figure 2a). However, if the speaker were to adopt the addressee's spatial perspective, producing the classifier sign on his left, the location in signing space would conflict with the location of the target box observed by the addressee (see Figure 2b). We predict that in such situations, signers will use shared space, rather than adopt their addressee's spatial viewpoint.

It is not impossible to adopt the addressee's viewpoint when physical space is jointly observed by both interlocutors. For example, the speaker could describe an action of the addressee. In this case, the speaker would indicate a referential shift through a break in eye gaze, and within the referential shift, the signer could sign LIFT-BOX using a handling classifier construction articulated toward the left of signing space. The signing space in this case would reflect the addressee's view of the environment (i.e., the box is to the addressee's left).

In general, however, for situations in which the speaker and addressee are both observing and discussing a jointly viewed physical environment, there is no true speaker versus addressee point of view in signed descriptions of that environment. The signing space is shared in the sense that it maps to the observed space and to both the speaker's and addressee's perspective of the physical space. Furthermore, the speaker's description of the box would be the same regardless of where the addressee was standing (e.g. placing the addressee to the signer's left in Figure 2, would not alter the speaker's description or the nature of the mapping from signed space to physical space). Thus, in this situation, an ASL signer (unlike an English speaker) does not need to take into account where his addressee is located and can simply describe what he sees. This difference between languages derives from the fact that signers use the actual space in front of them to represent observed physical space.

To investigate the nature of perspective choice in ASL, we adapted a task developed by Mainwaring, Tversky & Schiano (1996) to elicit spatial descriptions. In this experimental task (adapted from Schober 1993), subjects are invited to join a "Secret Operations Agency" in which the subject must communicate with another secret agent (Agent Z) about the locations of various objects (e.g. a bomb, microfilm). The subject is given maps of various simple scenes with a description of each situation (see Figure 3). English speakers in the Mainwaring et al. (1996) study were told that for security reasons they must communicate via an "Encoder Pad" which allows Agent Z to send the subject a simple question that the subject can answer with a short written message. For the example situation shown in Figure 3, an English speaker might write "It's on my right," "It's on your left," or "It's near the water cooler." The results from Mainwaring et al. (1996) indicate that English speakers generally preferred to adopt their addressee's perspective. Overall, 81% of spatial descriptions adopted the addressee's viewpoint, and the percentage rose to 92% when no landmark was present in the environment. In contrast, we predict that ASL signers will not adopt their addressee's spatial perspective but will use shared space.

For English speakers, the presence of a relevant landmark in the environment altered the tendency to adopt the addressee's perspective (in these situations, 23% of the descriptions were perspective-neutral and used the landmark to locate the target object) (Mainwaring et al. 1996). We investigate whether the presence of a landmark also affects the nature of spatial descriptions for ASL signers. Another factor that affects perspective choice for English speakers is whether they are conversing with actual conversational partners or with imaginary addressees (Schober 1993). Speakers are less likely to adopt their addressee's perspective when their addressee is actually present, rather than imaginary: 2% versus 13% of spatial descriptions were from the speaker's perspective for present vs. imaginary addressees, respectively (Schober 1993). Our experiments also investigate the effects of real vs. imaginary addressees for spatial perspective.
Agent M is being blackmailed. The incriminating documents are hidden in a filing cabinet in the blackmailer’s office. You and Agent Z have just broken into the office. On your map, the squares are the filing cabinets, the circle is the water cooler, and the arrow points to the filing cabinet with the documents. [English text was translated into American Sign Language]

![Sign Language Image]

Agent Z signs: “Where are the documents?”

Figure 3. An example scenario from Experiment 1.

perspective choice in ASL. Experiment 1 parallels the Mainwaring et al. (1996) study, except that the addressee is actually present in the (imagined) environment and asks the speaker for information regarding target locations (e.g. “Where are the documents?”; Figure 3). However, in Experiment 2, the speaker communicates with an imaginary addressee via a videophone. We predict that, unlike English speakers, the presence of an addressee does not increase the use of speaker perspective. Rather, when the addressee is present, we predict that ASL signers are much more likely to use shared space.

Finally, it is important to point out that we are comparing communication in written English to signed communication with an actual addressee in Experiment 1.

Our previous research with English speakers producing spoken descriptions of environments revealed that speakers often use gesture to indicate perspective (Emmmorey, Tversky & Taylor 2000). In their speech, subjects adopted either a route perspective (addresses were taken on a tour of the environment) or a survey perspective (the environment was described from one, unchanging, viewpoint). Speakers’ gestures generally mirrored the spatial perspective of the spoken description; for example, 3-D gesture space was used for route descriptions as if moving through an environment, and 2-D planar gesture space was used for survey descriptions, as if illustrating the environment on a blackboard. In addition, there were no differences in perspective choice for the spoken descriptions from Emmmorey, Tversky, and Taylor (2000) and the written descriptions of the same environments from Taylor and Tversky (1996). These results suggest that the ability to gesture does not change the nature of perspective choice for English speakers, at least when comparing route and survey perspectives. Thus, it is not unreasonable to compare the results from Mainwaring et al. (1996) with those of the present study, even though we are comparing a primary language mode (ASL) with a derived language mode (written English).

2. Experiment 1: Co-present situation

Experiment 1 was designed to test our predictions concerning perspective choice by ASL signers and to investigate the nature of spatial descriptions in ASL. As with the Mainwaring et al. (1996) study, we asked subjects to imagine themselves in various environments (e.g. an office, a zoo, a casino). However, unlike the English speakers, Deaf ASL signers communicated with an actual addressee who took on the role of Agent Z and moved to different locations within the room. For each environment and scenario, subjects were told to imagine that they were in a particular environment with Agent Z, as depicted on a set of maps.

2.1 Method

2.1.1 Subjects

Twenty Deaf subjects participated in the experiment. 15 subjects had Deaf families and were native signers of ASL, and 5 subjects were near-native signers, exposed to ASL at a mean age of 6 years. 14 subjects were deaf from birth, and 6 became deaf before age 3. Subjects were tested either at Gallaudet University in Washington, D.C., at the Salk Institute in San Diego, or at California State University, Northridge.
2.1.2 Materials and procedure

Subjects were first shown an instructional videotape in which a native ASL signer (dressed in suit and tie) welcomed the subject into the Secret Organization and identified himself as its head. He explained that the subject's mission was to help Agent Z accomplish various dangerous but important tasks in a variety of environments. Subjects were told that there would be two parts to their job. For the first part, the subject would meet Agent Z in the same location, and the person in the room with them now was Agent Z (the Deaf research assistant in the experiment room). It was not feasible to use "encoder pads" as in the English study because we were interested in signed, rather than written responses. Furthermore, pilot research indicated that signs overwhelmingly preferred to sign to a live addressee, rather than to imagine an addressee as present in the environment. In the second part (Experiment 2 below), subjects were told they would work with another secret agent and that this mission would be explained later. Subjects were told to use ASL as their "secret language" and that Agent Z would only be able to ask a simple question, and subjects must provide a short answer. Agent Z would not be able to ask for clarification so their answer must be clear and concise.

Subjects were given a set of maps, and for each map the Agency head described their mission and the environment in ASL on videotape. Subjects were told that "you" indicated where they were standing and that "Z" indicated where Agent Z was standing (see Figure 3, but note that for ASL subjects, maps were only labeled with the name of the location, e.g. "Blackmailer's office," the English description was not printed on the map). The map always matched the real world scene, i.e. you was always at the bottom of the map representing the subject's view (subjects stayed in one position), and the Deaf research assistant moved to the appropriate locations indicated on the map for Agent Z (i.e. to the side or directly across from the subject). After each scene description, Agent Z would ask the subject for information (e.g. "Where's the microfilm?"). All subjects' responses were recorded on videotape for later analysis.

Each map description included a brief key to the diagram (e.g. "the squares are the wine vats, and the arrow points to the poisoned one"). For all maps, two containers were shown as identical filled squares with an arrow pointing to the target container. If a scenario included a landmark, it was shown as an unfilled circle.

Subjects were presented with 16 maps. Eight maps corresponded to the "baseline" condition of Mainwaring et al. (1996) which only depicted Agent Z and the two containers (no landmark or compass direction), and eight contained a landmark which was skewed toward either the target container or the non-target container. For half of the maps, Agent Z was facing the subject, and for the other half, Agent Z was either to the left or right of the subject. Finally, the container squares were either aligned with the direction that the subject was facing (i.e. one in front of the other from the subject's view, as shown in Figure 4) or perpendicular to the subject (i.e. side-by-side from the subject's view, as shown in Figure 3).

2.2 Results and discussion

2.2.1 Response coding

Each ASL description was coded by a native ASL signer as using shared space, as taking the addressee's or the speaker's perspective, or as using the landmark to locate the target object. Figure 5 below provides an illustration of each response type for a single scene. A description was coded as using shared space if the subject located the target object by articulating classifier signs or deictic signs within signing space, such that the specified locations were isomorphic with the imagined environment in front of the signer. Figure 4a and 4b provide examples of ASL descriptions using shared space (actual subject responses are illustrated in all of the figures, unless otherwise noted). Subjects either articulated appropriate classifier signs at specific locations in signing space (Figures 4a and 6), or they articulated deictic pointing signs or other deictic signs in signing space or toward the imagined environment (Figures 4b and 5a).

Responses were coded as from the addressee's viewpoint if the target location was lexically described with respect to the addressee's position (e.g. near you; see Figure 5b). A response would also have been coded as from the addressee's perspective if locations in the speaker's signing space represented the addressee's view of the environment, e.g. locating the target object on the left side of signing space, if the object was on the addressee's left; however, no subject produced such a description, which would have been analogous to the starred description in Figure 2b.

Responses were coded as from the speaker's viewpoint if the target location was lexically described with respect to the speaker (e.g. near me, see Figure 5c for another example). The use of shared space also reflects the speaker's perspective of the environment and could be thought of as a description from the speaker's point of view. However, the use of shared space also reflects the addressee's view of the imagined scene, as illustrated in Figure 2. Again, we suggest that when shared space is used, there is no true addressee vs. speaker perspective.

Finally, when a landmark was present, signers produced some neutral non-spatial descriptions, such as the one shown in Figure 5d, but more often signers described the location of a target with respect to the landmark using shared space. Figure 6 provides an example in which the signer describes the location of two slot machines and a roulette wheel in a casino. She indicates the location of target slot machine with respect to the landmark (a roulette wheel). The locations in signing space representing the slot machines and the roulette wheel correspond to the locations indicated on the map (and thus to the locations in the imagined environment) and with the addressee's view of the imagined environment. Thus, although the target location is described with respect to a landmark, the signer expresses this relationship using shared space.
"You see two cars located here. It’s this one."

Figure 4a. Example of an ASL description using shared space and classifier signs. The subject described the location of the target car as depicted in the scenario shown on the map.

"It’s not the one located here. It’s the one located here."

Figure 4b. Example of an ASL description using shared space and deictic signs. The subject described the location of the target car as depicted in the scenario in Figure 4a.

2.2.2 Analysis of perspective choice
Re-descriptions were excluded from the analysis — the first response was considered primary. For example, if a subject used shared space to indicate the location of a target object, and then added "on your right," this would be counted as a shared space description, and the re-description with the addressee perspective would be excluded. The percent of all responses from each perspective is shown in Table 1. Figure 5 provides

"Now you are looking around a library. Earlier Agent X placed a secret code book in a drawer of a card catalogue cabinet. Agent Z now needs to retrieve it. The squares are the card catalogue cabinets and the arrow points to the one with the secret code book. The circle is the reference table."

a) Shared Space

b) Addressee’s perspective

"It’s in the card catalogue on that side, there."

c) Speaker’s perspective

"It's near you."

d) Landmark as reference point

"Two card catalog filing cabinets are located here. It's the one on the left."

"The secrets are not near the reference table. They are in the other table."

Figure 5. Example response types from Experiment 1. The signers are modeling responses from four subjects who described the location of the secret code book in the library, as depicted on the map at the top of the figure.
examples of each type of description for a single map. There was a total of 160 descriptions in the landmark condition, and 158 descriptions in the no landmark condition.\footnote{Percent of descriptions that used shared space in combination with referencing a landmark element (see Figure 6).}

Table 1. Percent of descriptions from each point of view for the co-present condition (Experiment I)

<table>
<thead>
<tr>
<th></th>
<th>No landmark</th>
<th>Landmark present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared Space</td>
<td>84.2</td>
<td>79.4\footnote{Percent of descriptions that referenced a landmark and did not use shared space (see Figure 5d).}</td>
</tr>
<tr>
<td>Addressee’s Point of View</td>
<td>7.6</td>
<td>8.8</td>
</tr>
<tr>
<td>Speaker’s Point of View</td>
<td>8.2</td>
<td>5.0</td>
</tr>
<tr>
<td>Landmark</td>
<td>N/A</td>
<td>6.8\footnote{Percent of descriptions that referenced a landmark and did not use shared space (see Figure 5d).}</td>
</tr>
</tbody>
</table>

As can be seen from Table 1, ASL signers overwhelmingly preferred to express locative information using shared space when the addressee was present in the environment. When subjects adopted the addressee’s perspective, they produced descriptions that did not make use of locations within signing space. As noted, no subject adopted the spatial perspective of the addressee using referential shift to “take on the role” of Agent Z. In such a description, the locations in signing space would represent the locations in the environment as observed by Agent Z. For example, for the situation depicted in Figure 3, no signer indicated the location of the incriminating documents by articulating a classifier or deictic sign on the left of signing space.

When shared space was not used to describe target locations, descriptions were fairly evenly divided between those that adopted the addressee’s viewpoint and those that adopted the speaker’s viewpoint. In contrast, the English descriptions in the Mainwaring et al. (1996) study were most often from the addressee’s perspective. In the baseline (no landmark) condition, 92% of the descriptions adopted the addressee’s viewpoint (only 8% adopted the speaker’s viewpoint, and in the (skewed) landmark condition, 70% of the descriptions adopted the addressee’s viewpoint, 8% adopted the speaker’s viewpoint, and 23% used the landmark to indicate the target location. However, for ASL signers, choice of addressee vs. speaker viewpoint was affected by whether the target location was nearest the addressee (as in Figure 5). For these scenarios, 12% of the descriptions were from the addressee’s perspective, and only 1% of descriptions adopted the speaker’s perspective. When the target location was nearest the speaker, descriptions were evenly divided between addressee (8.8%) and speaker (8.8%) viewpoints.

\footnote{One subject’s description utilized an absolute reference frame and was excluded (this subject indicated that the target was the “North barn” assuming North was at the top of the map), and a second subject misunderstood the first map and did not produce a description for that environment.}

\[ YOU \text{SEE SLOTS TWO MACHINE} \] (not shown)

\[ LOCATED-HERE \]
\[ R(-O-U-L-E-T-T-E) \]
\[ WHEEL-LOCATED-HERE \]

\[ NEAR \]
\[ WHEEL-HERE \]
\[ RIGHT-THERE \]

“You see two slot machines located here and there. A roulette wheel is located here. It’s near the roulette wheel. Right there.”

“You are now looking around a casino. A mischievous person has placed a gas canister in a slot machine. If someone hits the jackpot, the canister will explode releasing the gas. Agent Z has arrived and is an expert at removing and defusing bombs. The black squares are the slot machines. The arrow indicates where the canister has been placed. The circle indicates a roulette wheel.”

Figure 6. Example of a description using shared space with the landmark as the reference object.

In sum, most ASL signers produced spatial descriptions in which classifier signs and deictic signs were articulated at locations in signing space that schematically mapped to both the speaker’s and addressee’s view of object locations within the imagined environment. When shared space was not used to express spatial information, ASL signers exhibited a preference for descriptions from the addressee’s perspective, but only when the target location was nearest the addressee. Otherwise, neither the speaker’s nor the addressee’s perspective was preferred. This pattern of performance contrasts with English speakers who overwhelmingly prefer to produce
the addressee (Agent Y) must perform a mental transformation of locations observed in the speaker’s signing space to correspond to locations observed in the field (e.g., the target object is in front of Agent Y, rather than to the left as observed on the videoscreen). A mental transformation is also required when the speaker adopts Agent Y’s perspective via referential shift (Figure 7a). The addressee (Agent Y) is facing the speaker via the videophone screen. Thus, a location observed as farthest away from the addressee on the video screen (closest to the speaker) must be understood as closest to the addressee (Agent Y) in the actual environment.

Figure 7. Illustration of three possible spatial perspectives that a speaker could adopt to describe the greenhouse scenario depicted at the top of the figure.

A third possible way of describing the environment to Agent Y would be to assume shared space. That is, speakers know that Agent Y is facing the video monitor, and they imagine that Agent Y is facing them in the room. In this case, speakers must mentally transform the location of Agent Y on the map so that Agent Y is now facing them, as illustrated in Figure 7c. In this case, the addressee does not need to perform any transformation of the speaker’s signing space. The locations observed in signing space on the videoscreen map directly to the locations observed in the environment, e.g., the target location observed as closest to the addressee (Agent Y) on the video screen is also closest to the addressee in the actual environment (for the scenario in Figure 7).

In sum, for the situation in which the speaker and addressee are not in the same environment, ASL signers have several options for describing the location of target objects. A speaker can express the spatial viewpoint of the addressee by using referential shift, or the speaker can express his or her view of the map, describing the location of the addressee with respect to the target object. Both of these options require some type of mental transformation on the part of the addressee who is facing the speaker on the videoscreen. The final option is to use shared space by imagining the addressee as...
3.1 Method

3.1.1 Subjects
The same subjects from Experiment 1 participated in Experiment 2.

3.1.2 Materials and procedure
Subjects were told that they were now working with Agent Y, an expert spy who works out in the field, and that Agent Y has a small TV monitor that receives information from the camera in front of them. As in Experiment 1, subjects were presented with 16 maps. Eight maps depicted only Agent Y and the two containers (no landmark), and eight contained a landmark which was skewed toward either the target object or the non-target object. For half of the maps, Agent Y was at the top of the map "above" the objects (e.g. Figure 9), and for the other half, Agent Y was either to the left or right of the objects (from the speaker's view of the map). Finally, the container squares were either aligned horizontally (as in Figure 7) or vertically (as in Figure 8).

3.2 Results and discussion

3.2.1 Response coding
Each ASL description was coded by a native ASL signer as adopting the addressee's point of view, the speaker's point of view (of the map), or using shared space (via the video connection). Descriptions were coded as from the addressee's viewpoint if the target location was described lexically from the addressee's point of view (e.g. "near you") or if the locations in signing space represented Agent Y's view of the environment. Descriptions were coded as from the speaker's viewpoint if the locations in signing space represented the subject's view of the map. No subject lexically described a target location from their viewpoint as speaker (e.g. "near me") because the subject was not in the environment. Examples of spatial descriptions from the addressee's perspective and from the speaker's view of the map are given in Figure 8.

Descriptions were coded as using shared space if the subject indicated that he or she imagined the addressee (Agent Y) as facing them, either by physically rotating the map so that Agent Y was located at the top (effectively facing the subject) or by describing the location of the target object as if Agent Y were facing the subject (as illustrated in Figure 7c). In addition, shared space could be indicated by articulating opposite the speaker (via the videophone). If this option is chosen, the addressee need not perform any type of mental transformation, but the speaker must mentally transform the locations observed on the map such that the locations correspond with a shared space situation (see Figure 7c). Experiment 2 examines which spatial perspective is chosen most often by ASL signers.
signs outward from the speaker, as if creating a signing space between Agent Y and the speaker. An example is shown in Figure 9.

Descriptions were coded as "landmark only" if the subject produced a perspective neutral description of the location of the target object with respect to the landmark (e.g., near water fountain). If the description referenced the landmark, but signing space clearly reflected the subject's view of the map, then the description was coded as from the speaker's viewpoint. Similarly, if the description referenced the landmark, but signing space clearly reflected the addressee's perspective, then the description was coded as from the addressee's viewpoint.

Finally, a few subjects produced a small number of descriptions that expressed a type of "mixed" perspective that we had not anticipated. For these descriptions, the locations in signing space reflected the speaker's view of the map, but the identity of the target location was specified lexically from the addressee's point of view. In the example shown in Figure 10, the subject indicates that the target location is on the addressee's left by signing on your left, but the sign left is articulated in signing space with respect to the subject's view of the target location on the map.

3.2.2 Analysis of perspective choice
The percent of descriptions from each perspective is shown in Table 2. There was a total of 160 descriptions in the No Landmark condition and 159 descriptions in the Landmark condition (one subject produced a description that was unclear and uncodable). The vast majority of descriptions (over 90% in both conditions) expressed either the addressee's perspective or the speaker's view of the map, and these descriptions were roughly split between the two perspective types. Combining the conditions, there were slightly more descriptions from the addressee's perspective (48.6%) than from the speaker's view of the map (43.5%). Nine subjects consistently adopted the addressee's perspective (over 80% of their descriptions were from this perspective); eight subjects consistently described the target locations with respect to their view of the map, and three subjects alternated between different perspective types.

Table 2. Percent of descriptions from each point of view for the remote condition (Experiment 2)

<table>
<thead>
<tr>
<th>Description Type</th>
<th>No Landmark</th>
<th>Landmark Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addresser's viewpoint</td>
<td>47.5</td>
<td>49.7</td>
</tr>
<tr>
<td>Speaker's view of map</td>
<td>45.6</td>
<td>41.5</td>
</tr>
<tr>
<td>Shared space (via video)</td>
<td>3.3</td>
<td>3.8</td>
</tr>
<tr>
<td>Landmark only</td>
<td>N/A</td>
<td>4.4</td>
</tr>
<tr>
<td>Mixed view</td>
<td>3.8</td>
<td>0.6</td>
</tr>
</tbody>
</table>

In sum, when the speaker and addressee were not jointly viewing the environment, speakers either adopted their addressee's perspective via referential shift (i.e., locations in signing space were described as if the speaker were Agent Y) or speakers expressed locations from their own perspective by describing locations on the map and Agent Y's position in the environment. Two-thirds of the participants consistently adopted a single perspective for most scenarios (either addressee or speaker perspective), and one-third of the participants varied perspective choice across scenarios. Unlike English speakers, we found no general bias toward adopting the addressee's perspective. Participants rarely chose to use shared space to describe locations. The use of shared space occurred when speakers imagined the addressee as facing them (either in the imagined environment or via the videophone). Perspective neutral descriptions referencing the landmark were also relatively rare. In general, when landmarks were mentioned, signing space clearly reflected either the speaker's view of the map or the addressee's view of the environment. Finally, for a few location descriptions (4%) the perspective indicated by signing space and the perspective indicated by lexical forms were not the same (Mixed View descriptions; see Figure 10).
deictic signs are positioned in signing space. Although lexical encoding via spatial terms and phrases is also possible, it is not the preferred means of expressing spatial information. When interlocutors are present in an environment, there is a relationship between signing space and the environment (whether imagined or real). Locations in signing space associated with referent objects must map to the location of objects in the environment described and to the interlocutors' views of that environment. These locations reflect a shared view of the environment and do not strictly reflect either the speaker's or the addressee's perspective. Lexical encoding (in either ASL or English) tends to force the adoption of a particular viewpoint (e.g. 'my left' vs. 'your right'), unless a neutral perspective is available (e.g. 'near the fountain'). The use of signing space neutralizes the need to adopt a particular perspective when describing jointly viewed (or imagined) environments.

In a separate study, Emmorey (in press b) examined how signing space was used when ASL interlocutors were not discussing a present environment. ASL signers (speakers) memorized the locations of landmarks in a town (from a map) and described these locations to an addressee who was later asked to draw a map of the town. All but one of the eleven speakers in this study described the environment from their own perspective, i.e. from their view of the map. When addressees asked a question or re-described the locations of landmarks (e.g. for clarification), they "reversed" the locations in the speaker's signing space. For example, speakers described Maple Street as looping to the left (observed as to the right by the addressee facing the speaker), and when addressees subsequently described Maple Street, they also indicated that Maple Street looped to the left in their signing space. Both speakers and addressees generally described the map of the town as if they were each looking at the map. Thus, locations observed on the addressee's right (i.e. locations on the speaker's left) were described as on the left by the addressee.

In addition, addressees in the study by Emmorey (in press b) produced re-descriptions and questions that utilized a form of shared space. In these instances, the addressee and speaker did not maintain separate signing spaces. In one type of shared space, an addressee referred directly to referent locations by directing a pronoun or classifier sign toward the appropriate location in the speaker's signing space. Such shared space is common for non-spatial discourse when an addressee points toward a location in the speaker's space in order to refer to the referent associated with that location. In another type of example from Emmorey (in press b), the signing space of the speaker and addressee physically overlapped. For both of these example types, signing space is shared because both interlocutors use the same locations in signing space to refer to locations and/or referents associated with those locations.

In sum, the notion of shared space derives primarily from how the addressee interprets the speaker's signing space. The addressee shares the speaker's signing space either because it maps to the addressee's view of present objects (as in the present
study) or because the addressee uses the same locations within the signer's space to refer to non-present objects (as in the Emmorey in press b study). Because addressees can share the speaker's signing space, speakers rarely produce spatial descriptions that are from the addressee's spatial viewpoint, i.e. in which locations in signing space match the addressee's view of an environment, rather than the speaker's view.

Experiment 2 investigated the unusual situation in which a speaker described locations in an environment to an addressee who was in that environment via a type of videophone. In our scenario, the speakers imagined themselves in a control room with maps of various environments, while the addressee was imagined as present in these environments, communicating with the speaker via a monitor and video camera. This situation highlighted the unusual difference between spoken and signed languages. For environments in which speakers know only the location of the target object and the position of the addressee (Agent Y), English speakers can only naturally describe the location of the target object from the addressee's perspective. However, as we have discussed, ASL signers have at least three perspective choices: they can describe locations from the addressee's viewpoint, from their own view of the map, or they can use shared space via the videophone. The results indicated that perspective choice was roughly split between addressee and speaker viewpoints.

Adoption of the speaker’s viewpoint lessens the cognitive communication load for the speaker who does not need to perform a spatial mental transformation because locations on the map are directly mapped to locations in signing space, but the cognitive burden is greater for the addressee who must perform a mental transformation to map locations in the speaker's signing space to locations in the observed environment (see Figure 7). The cognitive load for the speaker is greater when the addressee's perspective is adopted, but the addressee must still perform a mental transformation (because the addressee is facing the speaker), but this "reversal" transformation occurs frequently during sign conversations and does not appear to be difficult for addressees (see Emmorey et al. 1998). Given the even split in perspective choice, it suggests that ASL speakers know that the reversal transformation required of their addressee is not difficult. In contrast, English speakers appear to know that reversing left and right (e.g. "my right is your left") is difficult for their addressee, and they avoid producing descriptions that require this transformation (Tversky 1996).

In conclusion, these studies highlight some of the consequences of using signing space to represent space for choice of viewpoint in spatial descriptions. The fact that signers can view both the environment and the speaker's signing space allows for speaker and addressee points of view to be "combined" via shared space. Because spatial notions are encoded by lexical items and phrases in English, these speakers must adopt a particular viewpoint (most often the addressee's viewpoint), and the location of the addressee will affect which spatial terms are chosen. Experiment 2 revealed that the use of signing space permits additional points of view in situations where English allows only one. An unusual aspect of Experiment 2 was that the addressee was not allowed to respond to the speaker (being imaginary). It is possible that choice of perspective might change if the speaker and addressee were communicating using a true videophone. In this case, we might find that speakers and addressees make use of "virtual" shared space (e.g. referring to locations in signing space observed on the videophone). Further research may reveal other factors that influence the nature of perspective choice in signed languages and how speakers and addressees negotiate the use of signing space with respect to these spatial viewpoints.

Acknowledgments

This research was supported by grants from the National Institutes of Health (NICHHD RO1-13349) and the National Science Foundation (BRS-9809002) awarded to Karen Emmorey, and by grants from the Office of Naval Research (N00014-PP-1-0659 and N000140110717) awarded to Barbara Tversky. We thank Scott Mainwaring for the use of his stimulus materials, and Brenda Falgier and Melissa Herzig for help with data coding and with testing Deaf subjects. We also thank Sam Hawk, Helsa Borininstein, and Kelly Gunderson for help with the ASL illustrations.

References


Corresponding author's address
Karen Emmorey
Laboratory for Cognitive Neuroscience
The Salk Institute for Biological Studies
10010 North Torrey Pines Rd.
La Jolla, CA 92037
emmorey@salk.edu
Fax: (858) 452–7052
Phone: (858) 453–4100

Sign Language & Linguistics

Subscription rates
Subscription rates for Vol. 5 (2002) (2 issues; ca. 300 pp.) Postage included; airmailed to all countries.

Institutions/Companies/Libraries EUR 135.00
Excl. VAT EC customers should provide their VAT-ID number plus country code and will be invoiced

John Benjamins Publishing Company offers special, reduced, subscription rates for private subscribers. For information on the rates and the conditions, please contact our subscription department.

For general information, contact the publisher

John Benjamins Publishing Company
P.O. Box 36224, 1020 ME Amsterdam, The Netherlands
E-mail: customerservices@benjamins.nl
http://www.benjamins.com
Using space to describe space: Perspective in speech, sign, and gesture

KAREN EMMOREY¹, BARBARA TVERSKY² and HOLLY A. TAYLOR³
¹Laboratory for Cognitive Neuroscience, The Salk Institute for Biological Studies, 10010 North Torrey Pines Rd., La Jolla, CA 92037, U.S.A. (Fax: (858) 452-7052, E-mail: emmorey@salk.edu); ²Stanford University; ³Tufts University

Received 11 October 2000; accepted 30 May 2001

Abstract. Describing the location of a landmark in a scene typically requires taking a perspective. Descriptions of scenes with several landmarks use either a route perspective, where the viewpoint is within the scene or a survey perspective, where the viewpoint is outside, or a mixture of both. Parallel to this, American Sign Language (ASL) uses two spatial formats, which the described space is conceived of as in front of the speaker, or the described space is conceived of as from outside, usually Enlish or ASL described one of two memorized survey perspectives than English speakers, perspective choice. In ASL, descriptions from whereas descriptions from a route perspective accompanying route descriptions used the full set of accompanying survey descriptions used diagrammatic space. Thus, the two modes of from without, are expressed naturally in speech.

... question, or one like it, is among the first in any travel or directions language. Answering the question, describing the location of a landmark, typically requires taking a perspective (e.g., “The train station is on Bahnhoff Strasse, north of the bus terminal” or “You go down Bahnhoff Strasse, and you’ll see the bus terminal on your left and then the train station”). How I tell you where the train station is depends on how I think about the environment or how I think you’re thinking about the environment. In describing environments with many landmarks, such as a convention center or a town, speakers of spoken languages such as English or Dutch generally take either a route or a survey perspective, or a mixture of
the two (Taylor and Tversky 1996). In a route perspective, the viewpoint is within the scene, and the addressee is taken on a mental tour of the environment, so the viewpoint changes with the addressee's position. Landmarks are described relative to the current imagined position of the addressee in terms of "left," "right," "front," and "back." For example, "After you turn right on Michigan Avenue from Randolph, you will see the Chicago Symphony on your right and the Art Institute on your left." In a survey perspective, the imagined viewpoint is stationary and outside and above the environment. Landmarks are described relative to each other in terms of "north," "south," "east," and "west" (Perrig and Kintsch 1985; Taylor and Tversky 1992a, b, 1996). For example, "The Museum of Natural History is west of Central Park and the Metropolitan Museum is on the east side of the Park, a little north of the Natural History Museum." Occasionally, speakers may adopt a gaze perspective, which can be viewed as an amalgam of route and survey perspectives. In a gaze perspective, speakers take a point of view just at the edge of a scene, for example, an entrance to a room; and they describe the locations of objects to other objects relative to that point of view, e.g., "the lamp is left of the reading chair" (Ehrich and Koster 1983). Gaze perspective descriptions are relatively rare for large-scale environments and thus won't be discussed here.

**Route and survey descriptions**

Route and survey perspectives differ with respect to 1) point of view (moving within the scene vs. fixed above the scene), 2) reference object (the addressee vs. another object/landmark), and 3) reference terms (right-left-front-back vs. north-south-east-west). Route and survey perspectives also correspond to two natural ways of experiencing an environment. A route perspective corresponds to experiencing an environment from within, by navigating it, and a survey perspective corresponds to viewing an environment from a single outside point at a height, such as a tree or a hill. In a survey description, the viewer is clearly outside the environment, looking on it almost as an object with parts; whereas in a route description, the viewer is immersed within the environment. The correspondence of the description perspective to natural ways of experiencing environments may account for the consistency of perspective that appears in many spatial descriptions (Taylor and Tversky 1992a).

For spoken languages, which perspective is adopted seems to depend at least in part on features of the environment. In research eliciting descriptions of a number of environments learned from maps or by exploration and varying on many features, participants tended to use relatively more survey perspectives when the environment had features on several size scales and when there were several routes through the environment. Participants used relatively more route perspectives when environmental features were on a single size scale and when there was only a single natural route through the environment (Taylor and Tversky 1996). In about half the cases, participants mixed perspectives; frequently, perspective switching occurred where portions of the environments matched the ideal conditions for one or the other perspective.

Expression of spatial perspective is especially interesting to study in a signed language because of a unique resource afforded by the visual-manual modality: the use of space to represent space. In fact, American Sign Language (ASL), the primary language of the Deaf in most of North America, uses two different spatial formats to express locations and spatial relations between objects (Emmorey and Falgier 1999). *Signing space* is the term used for the three-dimensional space in front of the signer, extending from the waist to the forehead, where signs can be articulated. In the diagrammatic spatial format, signing space represents a 2-D or 3-D map-like model of the environment and takes a fixed bird's-eye view from above the environment. Signing space is either a low horizontal plane (tabletop) or a vertical plane (blackboard). In the viewer spatial format, signing space reflects the individual's current 3-D view of the environment, so the viewpoint can change as the turn is described.2

The two spatial formats proposed for ASL bear resemblance to the two perspectives that appear in spoken language descriptions of environments. The essential quality separating route and survey perspectives and use of viewer and diagrammatic space seems to be whether the viewer is conceived of as immersed in an environment or whether the viewer is conceived of as outside an environment, looking at it as an object. These two modes of conceptual relationship between a viewpoint and an environment have been evident in previous work on perspective (Bryant et al. 1992; Bryant and Tversky 1999; Franklin et al. 1992). In these experiments, internal or immersed and external perspectives could be induced in several ways, by the nature of the description of viewpoint, by instructions, by conveying the environment through diagrams or models. In each case, the immersed or internal viewpoint led to different consequences compared to the external viewpoint, specifically with respect to the time required to respond to direction probes (e.g., above, behind, etc.).

**Comparison of English, ASL, and gesture**

Given the similarities of the two ASL spatial formats to the two frequent spoken perspectives, and given their correspondence to two basic ways of experiencing the world, two questions arise. First, does the use of diagrammatic and viewer spatial formats in ASL coincide with the choice of survey
and route perspective descriptions by signers? And second, do the same environmental features that elicit survey or route perspective descriptions for English speakers, elicit the same perspective choices for ASL signers?

The comparison of spoken and signed languages also invites the study of gesture, where, as for signed language, space can also be used to describe space, though not in any conventionalized way. In fact, spatial descriptions elicit relatively large numbers of gestures (Iverson and Goldin-Meadow 1998). Of course, in spoken languages, gestures are not the primary means of communication. Nor are gestures universally used, to the frustration of those who study them. Three aspects of gesture are of special interest to the study of spatial descriptions. First, describing environments is likely to elicit a number of iconic gestures, gestures that schematically depict the situation they are intended to convey (McNeill 1992). Even more than that, describing environments may elicit gestures that construct models. By model, we mean three or more successive gestures that are used to convey structural features of the environment. In research in which an expert explained the workings of a lock to a novice, experts used gestures to construct models of the components and dynamics of the locks (Engle 1998). The third and perhaps most interesting aspect of gestures investigated here is gesture space (i.e., the space in front of a speaker where gestures are made). Studying gestures produced for spatial descriptions allows us to examine the spatial format of gestures. For example, do the spatial features attributed to diagrammatic and viewer spatial formats in ASL apply to gesture space? What is the relation between the perspective indicated by spoken descriptions and the nature of iconic gestures that accompany these descriptions? This study examines the correspondence between the spoken perspective and the gesture space, and the possible relations of gesture space to spatial format in signing.

To investigate how space is used to describe space in sign and in gestures accompanying speech, both ASL signers and English speakers studied one of two maps of two environments, a town or a convention center (see Figure 1). They then described the environment from memory. In a previous study using written English, Taylor and Tversky (1992b, 1996) found that the town elicited primarily survey perspective descriptions, while the convention center elicited primarily route perspective descriptions.

**Method**

**Participants**

Forty Deaf ASL signers participated in the study (26 females; 14 males). Twenty-seven signers had Deaf families and learned ASL from birth. Ten signers acquired ASL between age two and seven, and three acquired ASL...
after age fourteen. Thirty-eight signers were deaf at birth or became deaf before 1 year of age, and two signers became deaf at or before age three. Signers participated in the study either at Gallaudet University, The Salk Institute, California State University at Northridge, or at Deaf Community Services in San Diego. All participants were college educated. Forty hearing English speakers also participated in the study (21 females, 19 males); thirty-nine hearing participants were students at Stanford University, and one was a student at the University of California, San Diego.

Materials and procedure

Half of the participants were given the map of the town (Figure 1a), and half were given the map of the convention center (Figure 1b). They were asked to study the map until they had memorized it. Participants were told to describe the environment so that if someone unfamiliar with the area were shown the videotape of their description, they would know what the environment (town/convention center) looked like and where all the landmarks were. The instructions were given either in ASL by a Deaf native signer or in English by a native English speaker.

Results and discussion

Each description was judged as adopting a route perspective, a survey perspective, or a mixed perspective. For the ASL descriptions, two Deaf native signers were asked to decide if the description felt more like a "tour", a bird’s eye view description, or a mixture of both. They were shown examples of written English route and survey descriptions from Taylor and Tversky (1996) to familiarize them with how English descriptions were coded. The ASL coders agreed on 88% of judgments for ASL descriptions. When disagreements occurred, the signers discussed the description and came to an agreement. For the English descriptions, two of the authors (HT and BT) independently rated each description and agreed on all of the judgements. The results are shown in Table 1.

The perspectives adopted by ASL signers differed from those of English speaking participants. ASL signers were significantly more likely to adopt a survey perspective when describing the convention center, whereas English participants preferred a route perspective ($X^2 = 8.98, p < 0.05$). For the town, English and ASL participants did not differ significantly in perspective choice.

Why do ASL signers use more survey descriptions than English speakers for the convention center? One possibility is that signers prefer survey perspectives in general, perhaps because signing space can be used so effectively to represent a map. That is, participants can locate landmarks on a horizontal plane in signing space in a manner that is isomorphic to the locations of landmarks on a map (in fact, this is how signing space is utilized for survey descriptions). However, ASL signers appear to choose either route or mixed descriptions when describing environments that they have actually experienced themselves. In a pilot study, ASL signers described either their houses ($N = 8$) or the locations of the dormitories on the Gallaudet Campus ($N = 5$). Only one person produced a description with a survey perspective. Thus, the difference between English speakers and ASL signers does not appear to be due to an unqualified preference for the use of a survey perspective.

The nature of the linguistic system may affect which spatial perspective is chosen. The fact that signers studied a map may have influenced how they structured signing space within their descriptions, which in turn may have affected their perspective choice. That is, a mental representation of the map may be more easily expressed using a horizontal plane in signing space, and this type of spatial format may be more compatible with a survey perspective, as expressed in ASL. There is no such natural correspondence in English. However, the apparent preference for survey perspectives in ASL requires further substantiation.

It is possible that the differences between ASL and English regarding perspective choice are due to differences in linguistic judgment criteria used by the ASL and English judges. To determine whether ASL and English participants used similar language, we examined the use of motion verbs and spatial terms. Although ASL signers tended to rely on classifier constructions (see below) and the topographic use of signing space in their environment descriptions, signers did produce some lexical spatial terms. We compare the use of these terms with their English counterparts.

Table 1. Perspective choice by ASL signers and English speakers

<table>
<thead>
<tr>
<th></th>
<th>Route</th>
<th>Mixed</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convention Center - English ($N = 20$)</td>
<td>11</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>ASL ($N = 20$)</td>
<td>3</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Town</td>
<td>English ($N = 20$)</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>ASL ($N = 20$)</td>
<td>7</td>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>
result replicates Emmorey’s (1996) finding of shorter description times for ASL signers when describing the spatial layout of rooms in a doll house. Emmorey (1996) hypothesized that the use of signing space to represent physical space and the use of the hands to represent objects may result in the relatively rapid and efficient expression of spatial relations in ASL. For nonspatial language, the proposition rate is equal for ASL signers and English speakers, but individual ASL signs take about twice as long to articulate as English words (Bellugi and Fischer 1972).

**Lexical comparison of ASL and English**

Finally, the data in Table 2 indicate that the lexical encoding of spatial perspective within a discourse is similar for both ASL and English. However, for ASL signers, lexical spatial terms are not the primary mechanism for expressing spatial perspective, as attested by the rarity of these terms. For English speakers, directional terms such *left/right* and *East/West* are frequent, but they are hard to produce and comprehend (e.g., Franklin and Tversky 1990; Scholl and Egath 1981). For example, some speakers specify a landmark as *near* another in order to avoid the use of directional terms (Tversky 1996). The use of signing space in ASL generally obviates the need for directional terms because signers have a more direct way of encoding directional information, e.g., by articulating signs to the left or to the right in signing space itself. We next examine those aspects of spatial language that are unique to signed languages.

**Perspective choice and spatial format in ASL**

When describing environments in ASL, the identity of each landmark is generally indicated by a lexical sign (e.g., BULLETIN-BOARD, SCHOOL, STORE), and the location of the landmarks, their orientation, and their spatial relation with respect to one another are indicated by where the appropriate accompanying “classifier” sign is articulated in the space in front of the signer (see Emmorey, in press). For example, the B handshape (a flat hand) is the handshape used to refer to rectangular, flat, surface-prominent objects, such as a bulletin board. Signers can also use pointing signs to indicate the locations of landmarks in signing space. Where English uses a linear string of prepositions and adjunct phrases to express spatial relations, ASL uses the visual layout displayed by various signs positioned in signing space. This visual layout can utilize either a diagrammatic or a viewer spatial format. Table 3 summarizes the properties associated with these two spatial format types (Emmorey and Falgier 1999).

<table>
<thead>
<tr>
<th>Table 3. Properties associated with two spatial formats in ASL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagrammatic space</strong></td>
</tr>
<tr>
<td>• Signing space represents a map-like model of the environment</td>
</tr>
<tr>
<td>• Space can have either a 2-D “map” format or a 3-D “model” format</td>
</tr>
<tr>
<td>• The vantage point does not change (generally a bird’s eye view)</td>
</tr>
<tr>
<td>• Relatively low horizontal signing space or a vertical plane</td>
</tr>
</tbody>
</table>

To determine the relation between the use of spatial formats and perspective choice, for each ASL description, a native signer with linguistic training determined whether the location of each landmark was described using viewer space or diagrammatic space, according to the criteria listed in Table 3. When a survey perspective was adopted, 91% of landmarks were located using a diagrammatic spatial format. When a route perspective was adopted, 88% of landmarks were located using viewer space.

Figure 2 provides an example of the use of diagrammatic space, illustrating the pointing signs produced by one participant to indicate the locations of the outer rooms of the convention center. The locations within signing space map isomorphically to the locations of the rooms on the convention center map (Figure 1b). This particular signer is unusual because she did not rotate the map. That is, most signers (80%) “shifted” the map so that the entrance was located at the chest and the bulletin board extended outward on the left of signing space. This pattern may reflect a convention for spatial descriptions of buildings (and rooms) in ASL: position a multi-entrance at the front of the body. In fact, most sketch maps are drawn with the starting point at the bottom page, ignoring cardinal directions, as if the user could shrink and walk out on the map (Tversky 1981; Tversky 2000).

Viewer space is used when signers conceptualize the environment as present, describing the scene as if they were viewing the landmarks and other surrounding elements. Signers know that their addressees cannot see the environment, and therefore the description is not the same as if both discourse participants were simultaneously observing the environment. When both signers can view the environment, they use a type of *shared space* (Emmorey, in press).
Figure 2. Illustration of the diagrammatic spatial format in ASL. The figure shows pointing signs used in a survey description of the outer rooms of the convention center. The intervening lexical signs are not shown, and the lexical and pointing signs for the inner rooms are also omitted. The rooms associated with the pointing signs are given in parentheses.

Figure 3. Illustration of the viewer spatial format in ASL. The figure shows pointing and other signs used in a route description of the outer rooms of the convention center. The intervening lexical signs for the rooms are not shown, and the lexical and pointing signs for the inner rooms are also omitted. The rooms associated with the pointing signs are given in parentheses.

marks were described using viewer space for route descriptions, but when a survey perspective was adopted, landmarks were most often described using diagrammatic space.

Comparing gesture accompanying speech and ASL signs

We next examine the similarities and differences between the spatial gestures used by speakers (i.e., those gestures that convey spatial information, rather than beat gestures or abstract metaphorical gestures) and classifier constructions produced by ASL signers when describing the same environments. As with any study of spontaneous gestures, some individuals produced many gestures and others close to none. For this reason, analyses of gestures are not like analyses of other responses that are generated by all participants. Two coders determined whether gestures were iconic or non-iconic and whether a string of gestures created a model. A model was defined as three or more gestures that referred consistently to the same spatial structure, building it.
up in segments. By consistently, we mean gestures referring to connecting entities that use the same viewpoint and size scale. These criteria are similar to those of Engle (1998). The coders agreed on 70% of the judgements and came to agreement on the others.

Participants produced frequent non-iconic gestures, e.g., a gesture in which both palms are turned up, as if to present something to the listener (the "conduit" metaphor in which the discourse itself is presented as an object to the listener, see McNeill, 1992) or a gesture with the circular motion that accompanied a search for words. However, our comparative analysis will only concern the iconic gestures produced by speakers.

Comparing gestures and signs. First, there are many critical differences between gestures that accompany speech and sign language (see Emmorey, 1999, for a review; see also McNeill, 1992). Some of the most important distinctions are listed below:

- Signs have a sublexical (phonological) structure, not found in gestures. That is, signs, but not gestures, are governed by phonological rules such as assimilation and have both a segmental and a syllabic structure.

- Signs participate in a hierarchical constraint-based system of generative lexical processes, whereas gestures are generally holistic without internal structure.

- Signs are combined into sentences subject to language specific and universal constraints on grammatical form; gestures occur in combination only in special circumstances, and successive gestures do not form a larger hierarchical structure with internal ordering constraints.

- Unlike signs, gestures are timed to coincide with a structural element (the prosodic stress peak) of a separate co-expressive system (i.e., speech).

- Gestures, unlike signs, are relatively idiosyncratic with no agreed upon standard of form (excluding emblems like "thumbs up").

Certainly, these distinctions hold for the comparison of spatial gestures and ASL classifier constructions, which are used to specify spatial information. Nevertheless, iconic gestures produced by speakers during their spatial descriptions exhibited some similarities with ASL classifier constructions. Twenty-six participants used either their fingers (17) or hands (9) to trace paths or shapes. An example of a gesture that traces a path is shown below in Figure 6, gesture (b); and an example of a gesture that indicates the extent of an object (the bulletin board) is shown in Figure 7 below, gesture (c). Parallel to these gesture types, ASL has a type of classifier construction in which the hands trace the extent of an object or trace a path of motion (Supalla, 1986; Valli and Lucas, 1995). Examples of such constructions can be found in the description of the extent of the Bulletin Board in Figure 2 and in the description of the C-D and Computer rooms in Figure 3.

Iconic gestures were similar to ASL classifier constructions in that the hand could represent a landmark or an object. Areas, such as rooms, were frequently indicated by either forming the thumb and index finger into a "U" (8 participants) or by forming both hands into an open rectangle, fingers together, apart from the thumbs (8 participants) or by moving open hands in parallel (6 participants). This "rectangular" hand configuration is similar to the first syllable of the ASL sign ROOM, but the U-type handshape could never be used to represent a room in ASL. The gesture handshapes and some of their functions (e.g., to represent an object or to trace a path) were similar to classifier constructions in ASL, but of course for ASL, linguistic constraints determined which handshape could be chosen to represent a particular object.

Although the majority of gestures produced by speakers were singleton gestures, twenty-five participants produced at least one string of three or more gestures that constituted a model. Overall, 91 models were produced, and these were roughly distributed across route, survey, and mixed perspective types. Some of the models involved a long sequence of gestures; for example, one participant describing the convention center, conveyed the locations of bulletin board, office, rest rooms, cafeteria, CD displays, stereo displays, computer displays, and the entrance in a continuous coherent sequence of 15 gestures.

One similarity between gesture models and signed descriptions was the use of "anchor" gestures or signs when describing the location of multiple landmarks. Figure 4 provides an example from an ASL description of the town. After tracing the outline of Maple Street, the signer leaves his index finger in space, which serves as an anchor point representing the end of Maple street. He then locates the park, school and store in signing space while maintaining the point to the north end of Maple street. The signer appears to be deliberately holding this pointing sign because he does not release the point when making the two-handed sign SCHOOL (the citation form of this sign is made with both hands in a 5 hand configuration).

Similarly, many of the speakers (44%) produced gesture strings in which one hand appears to represent an anchor landmark (and its location), and the other hand indicates the location of other landmarks. In the example shown in Figure 5, the speaker’s left hand produces the anchor gesture representing the location of the school in the town and is held throughout the speech. The speaker represents the location of the school on the map with his left hand, which is held in space while he describes the location of the store and the town hall. The gestures that accompany these descriptions are produced at the appropriate locations in gesture space with respect to an imagined map with the school positioned on the left (see Figure 1a).
Thus, the gestures of speakers resembled some aspects of ASL classifier constructions in that the hand was used to represent a landmark, and speakers co-ordinated their hands in space during their descriptions to represent the relative locations of one landmark to another, creating a type of model of the environment in gesture space. Gestures differed from signs with respect to consistency (e.g., not all speakers gestured or produced gesture strings) and linguistic constraints (e.g., choice of handshape).

Comparing gesture space and signing space. Both signers and speakers use sign/gesture space to represent locations; yet, there are some interesting differences. ASL signers never used the vertical plane in signing space when describing the convention center. This constraint may be due to the fact that the vertical plane is limited to two dimensions in ASL, and the use of the sign ROOM to describe the convention center invokes a three dimensional space — therefore, signers used only the horizontal plane which can be three dimensional (representing a model, rather than a map, in diagrammatic space). English speakers were not constrained by such linguistic requirements, and 40% of the speakers produced gestures using a vertical plane parallel to the body when describing the convention center. The vertical plane seemed to be adopted to make the model created by gesture visible to listeners, and an example is provided in Figure 6. The spatial layout is in the vertical plane (e.g., the speaker's gestures indicate that the Cafeteria is above the restrooms in space).

Figure 4. Illustration of the use of anchor signs within the vertical plane in signing space. English translation: "(The street) is like this. Here's a park; here's a school; and here's a store."

(a) left of the park is uh is a [store] or is no is a [school] [and north of the park]
(b) (c) (d)
(e) (f) (g)

ok so [west of the park] is the school [north of the park] is the store, [east of the park]

(d) (e) (f)

(g)

across from mountain road is uh [the town hall]

(g)

Figure 5. Illustration of the gestures accompanying a description of the town by speaker 31. The extent of the gesture is shown by enclosing the concurrent speech in square brackets. The lower case letter underneath each section of bracketed text corresponds to the accompanying gesture shown above the text. This example illustrates the use of an anchor gesture.
One striking similarity between gesturing and signing is that speakers also appeared to structure gesture space using a type of diagrammatic vs. viewer spatial format. To determine whether speakers consistently matched their use of gesture space to the verbal perspective, two new coders who were naive to the study first analyzed all gestures. They coded gestures without the soundtrack, so that there were no cues to the linguistically encoded perspective. Each hand gesture was marked as tabletop, blackboard, viewer-like, or other. Tabletop and blackboard gestures were characterized by the use of a 2-D horizontal or vertical plane, which corresponds roughly to the use of diagrammatic space in ASL (see Figures 5 and 6). The viewer-like gestures were characterized as gestures that appeared as though the speaker were in the environment indicating directions from their viewpoint (see Figure 7). Viewer-like gestures generally moved away from the speaker and used the full 3-D space (rather than a 2-D plane) which corresponds roughly to the use of viewer space in ASL. Gestures that were coded as “other” could
not be unambiguously coded into one of the other three categories. Coders agreed on 80.5% of judgements, and disagreements were eliminated from the analysis. These gestures were then compared with their accompanying spoken description to examine the correspondence between the use of gesture space and perspective choice. To avoid any ambiguities, descriptions that used mixed perspectives were excluded from the analysis. Also excluded were participants who did not gesture. This left 10 route descriptions and 6 survey descriptions.

Gesture space and description perspective were highly related. For survey descriptions, 50.6% of the gestures used diagrammatic space and 23.3% used viewer space. In contrast, for route descriptions, 54.7% of the gestures used viewer space and 17.1% of the gestures used diagrammatic space (interaction between description perspective and gesture space, $F(1,14) = 7.07, p < 0.02$).

The examples in Figures 5 and 6 illustrate the use of diagrammatic space for gestures in which the gesture space represents a map-like model of the environment being described. The examples shown in Figure 7 illustrate the use of viewer space for gesture in which the speakers appear to imagine the environment as present in front of them rather than as a map, and their gestures reflect this conceptualization. In both examples, the speakers produce gestures that seem to refer to a three-dimensional imagined environment rather than to a two-dimensional map.

These variations in gestures and gesture space are not unlike McNeill's (1992) analysis of character vs. observer viewpoint or "voice" in gesture:

The use of space also differs for these voices [character vs. observer] . . . With the character voice the space envelops the narrator - it is a space for the enactment of the character, and includes the locus of the speaker at its center. With an observer's voice, in contrast, the narrative space is localized in front of the narrator - as if it were an imaginary stage or screen - and in this space the narrator moves the relatively undifferentiated figures (blobs). (p. 190)

Thus "character viewpoint" may reflect a description of the space in which the speaker imagines herself as in the environment, and "observer viewpoint" characterizes gesture space that reflects a view of the environment as a map or diagram.

**Correspondence between gesture space and signing space.** Combining the gesture space results with the ASL space results yields a provocative correspondence. In both spontaneous gesture and signing, the space forward from the body, the space of viewing and action, is more likely to accompany a route perspective, where the traveler is described as facing, usually moving forward, and viewing the environment from within. In these cases, the spatial relations of the landmarks are described relative to the viewer. Similarly, for both spontaneous gesture and ASL, the depictive space typically of the table-top, but also of the blackboard, is more likely to accompany a survey perspective, where the viewpoint is stationary and from above, and the spatial relations between landmarks are described relative to the reference frame of the external environment rather than the viewer. The correspondence suggests that the use of gesture and signing space may arise naturally from the two most common ways in which humans experience their environment.

**General discussion**

In describing environments with several landmarks, users of spoken languages frequently adopt one of two perspectives, either within the environment surrounded by it, or outside the environment observing it in entirety. For a route description, the viewpoint is embedded within an environment, changing direction within it. Landmarks are described relative to the changing point of view in terms of right, left, front, and back. For a survey perspective, the viewpoint is external to the environment and above it. Landmarks are described relative to each other in terms of north, south, east, and west. In a previous corpus, speakers mixed perspectives about half the time, usually without signaling (Taylor and Tversky 1992a).

A special feature of signed languages is that they can use space itself to convey spatial information. Indeed, in ASL, there are two dominant modes of describing objects in space. When viewer space is adopted, it is as if the speaker were traveling through the space, indicating turns but always facing forward, and locating the landmarks relative to the body. When diagrammatic space is adopted, it is as if the speaker were above the environment seeing it as a whole, and locating landmarks relative to each other. These two modes bear striking similarities to the two spoken perspectives, route to viewer space and survey to diagrammatic space. Does the special feature of using space to convey space affect perspective?

Research using a broad spectrum of environments (including indoor, outdoor, large and small, public and private) has shown that English speakers tended to use relatively more survey perspectives when there were several paths through the environment and when landmarks were on several size scales (Taylor and Tversky 1992a, 1996). Among the environments, that of a map of a town elicited the highest proportion of survey descriptions and the map of a convention center elicited the highest proportion of route descriptions. These findings were replicated here for English speakers, but not for ASL signers. Both speakers of English and of ASL preferred survey perspectives for describing the town. However, for the convention center, English speakers preferred a route perspective whereas ASL signers...
preferred a survey perspective. This indicates a bias in ASL relative to English toward a survey perspective. We propose two possible and non-contradictory explanations for the bias toward survey perspectives. One reason might be an overall preference for the use of diagrammatic space to convey complex spatial information. Since the diagrammatic spatial format is most consistent with a survey perspective, ASL signers are more likely than English speakers to adopt a survey perspective when describing environments. Another reason may be that the environments were learned from maps. The spatial transformation from a map to diagrammatic space may be so direct as to be compelling in signing. Whatever the explanation for the bias toward survey perspective in ASL, this research makes clear that features of the language as well as features of the environment affect linguistic perspective choice.

The comparison of iconic gestures and ASL signs revealed some intriguing similarities and differences. Similar to ASL classifier constructions, several speakers used their hands to represent landmark objects, and the spatial relation between landmarks was depicted by where their hands were positioned in space. For both ASL and gesture, one hand can serve as an anchor landmark, and other landmarks can be located in space with respect to this landmark. However, signing, unlike gesturing, is subject to linguistic constraints on handshape choice for representing objects, on the choice of a vertical or horizontal plane in signing space, and on the height of signing within a description.

Finally, the spatial perspectives observed for spoken language map naturally to the use of spatial formats in signed language: moving within an environment maps to viewer space and observing an environment from without maps to diagrammatic space. The gestures of English speakers complete this elegant correspondence between conceptual perspective and use of space. When the linguistic description was from within the environment (a route perspective), speakers produced gestures within a 2-D space as if they were in the environment. When the description was from an external viewpoint (a survey perspective), speakers produced gestures along a 2-D horizontal or vertical plane as if they were drawing on a chalkboard or tabletop. Gestures, unlike signed language, do not have a conventional format; they are optional and not used at all by some speakers. Thus, the correlation of gesture space (and spatial formats in ASL) with linguistic perspective suggests that there is a natural correspondence between how speakers and signers conceptualize the space around their bodies as they communicate with sign, speech, and gesture. This correspondence further suggests that the use of viewer and diagrammatic formats in ASL may derive from the same natural correspondence, another example of using space to describe space.

Acknowledgements

This research was supported in part by NSF grants SBR-9510963 and SBR-9809002 and NIH grant HD13249 awarded to Karen Emmorey and by a Varenius seed grant awarded to Barbara Tversky and Holly A. Taylor. We thank Steve McCullough, Brenda Falgier, and Kevin Clark for their help in analyzing the ASL descriptions; Bonita Ewan, Brenda Falgier, and Steve McCullough for their help collecting the ASL data; Kulin Tantod for his help collecting the spoken English data; and Eurie Chung, Rachel Groner, David Megathlin, Shelley Wyncocooj, No Vo, and Jesse Zolla for their help analyzing the gesture data. Portions of the ASL data were described previously in Emmorey and Falgier (1999).

Notes

1 Following convention, lower case deaf refers to audiological status, while upper case Deaf is used when sign language use and/or membership in the Deaf community is at issue.
2 Diagrammatic space is very similar to Liddell's notion of token space (Liddell 1985) and to Schick's model space (Schick 1990). Viewer space is parallel to Liddell's surrogate space and Schick's real world space.
3 The ASL and English verbs and directional terms were counted by the first author, KE, a linguist and fluent signer.
4 Underlining indicates an initialized sign.
5 The hearing participants' faces are masked to protect their privacy. The Deaf participants shown in Figures 2-4 gave permission for their images to be published.

References


Recent titles in
NEW DIRECTIONS IN COMPUTERS AND COMPOSITION STUDIES
Gail E. Hawisher and Cynthia L. Selfe
series editors

Gail Hawisher, Cynthia L. Selfe, Charles Moran, and Paul LeBlanc

Opening Spaces: Writing Technologies and Critical Research Practices
Patricia Sullivan and James Porter

A Rhetoric of Electronic Communities
Tharon W. Howard

Nostalgic Angels: Rearticulating Hypertext Writing
Johnalan Johnson-Ellola

The Computer and the Page: The Theory, History and Pedagogy of Publishing, Technology and the Classroom
James R. Kalmbach

Transitions: Teaching Writing in Computer-Supported and Traditional Classrooms
Mike Painquists, Kate Kiefer, James Hartvigsen, and Barbara Goodlew

Rhetorical Ethics and Internetworked Writing
James Porter

Feminist Cyberscapes: Mapping Gendered Academic Spaces
Kristine Blair and Pamela Takayoshi, editors

Electronic Writing Centers: Computing in the Field of Composition
David Coogan

Digital Fictions: Storytelling in a Material World
Sarah Stone

Working with Words and Images

Naps in an Old Dance

by

LEN

Graphics Communicate

onomatopeia with his

Ables Publishing

Westport, Connecticut

London
century and a half ago, a young English woman sent the letter in the

Chapter Four

Some Ways That

Communicate

Barbara Tversky

GRAPHICS COMMUNICATE
Children Invent Arithmetic

one group is based on a different kind of graphic, one on maps, one on drawings, and
produce different kinds of graphic elements. Each of our projects is

Some Ways That Graphics Communicate

LINE, CURVES, CROSSES, AND BLOBS: CONVEYING ROUTES

Lines appearing in one-to-one correspondence to the number of things, then, have been invented by children and by many cultures to represent numbers of things. Lines have been used to represent other concepts as well—for example, segments of roads. A number of years ago, we stopped passers-by near a student dormitory and asked them if they knew how to get to a popular fast-food place off campus (Tversky and Lee 1998, 1999). If they said they did, we asked them either to sketch a map or to write down directions to the restaurant. Figure 4.3 contains two of the route directions, and figure 4.4 contains two of the maps given by informants. Note that route maps differ from regional maps in that they only include the information needed to get from the start to the destination, meaning that much detail, even of the roads followed and structures encountered, is omitted.

Following an analysis developed by Michel Denis (1997), we coded both depictions and descriptions into their elements. Significantly, al-
Some Ways That Graphs Communicate

1. Emphasizes patterns and trends in the data. Graphs make it easier to see patterns and trends that might not be obvious in a table of numbers. For example, a line graph can show how a variable changes over time, while a bar chart can compare quantities at different points in time. Graphs can also highlight outliers or unusual data points.

2. Improves comprehension. Graphs are more effective than text at conveying complex information. They can help readers understand relationships between variables and trends over time. Graphs can also simplify complex data by reducing the amount of information that needs to be processed at once.

3. Facilitates comparisons. Graphs allow readers to compare different data sets or variables side by side. This can be particularly useful when comparing data from different sources or time periods. Graphs can also help readers identify similarities and differences between data sets.

4. Enhances retention. Graphs are easier to remember than raw data. This is because graphs provide a visual representation of information, which can be more memorable than numbers or text. Graphs can also help readers retain information for longer periods of time.

5. Encourages critical thinking. Graphs require readers to think critically about the data they are viewing. They must interpret the information presented and draw conclusions based on that interpretation. This process can help readers develop their analytical skills and improve their ability to think critically about information.

In summary, graphs are a powerful tool for communicating complex information in a clear and effective way. They can help readers understand patterns, trends, and relationships in data, and encourage critical thinking and retention of information.
The ways in which lines and containers are interpreted and used to communicate abstract information, then, is cognitively compelling. Lines connect and bars separate. These perceptual units carry conceptual meaning, affecting how depictions are interpreted and how descriptions are visualized.

**ARROWS: CONVEYING ORDER**

Arrows are another perceptual unit that appears in depictions, both concrete and abstract. In route maps, they are used to indicate direction of movement. In fact, about half the participants in the study of Tversky and Lee (1998) used them for that purpose. In diagrams of systems, arrows appear to indicate direction of power or control or causality. As for maps of environments, diagrams of systems by themselves are structural and neutral with respect to direction. Are arrows a cognitively natural perceptual unit comparable to lines, curves, and bars? A case can be made that they are. One place that arrow-like figures appear in nature is in river junctions or, on a smaller scale, in water runoff. In both cases, the arrow-like form created from the juncture indicates the direction of flow of the water. Another place arrows commonly occur is as tools created by humans. Here, too, the arrowhead point indicates the direction of movement—in this case, of the projectile. The transition from movement in space to movement in time to movement in causality seems to be a natural one, certainly one reflected widely in the ways people talk.

To ascertain whether arrows serve to indicate causal direction in diagrams, in ongoing research Julie Heiser and Barbara Tversky (Tversky et al. 2000) presented students with one of three diagrams: of a bicycle pump, a car brake, or a pulley system. Arrows were added to half the diagrams. Students were simply asked to interpret the diagrams. When there was no arrow, students’ descriptions were more structural. For example, for the pulley system without arrows, one student wrote: “A three-pulley system with a load/weight.” In contrast, when arrows were present, the descriptions were more causal or functional. For example, for the bicycle pump with arrows, one student wrote: “Pushing down on the handle pushes the piston down on the inlet valve which compresses the air in the pump causing it to rush through the hose.” Interestingly, most of the time, the structure was implicit in the functional descriptions. This suggests that it may be easier to infer structure from function than function from structure, a possibility we are now testing. We are also investigating the mirror-image situation, asking students to produce diagrams from either structural or functional descriptions. The expectation is that arrows are more likely to be included in the diagrams produced from functional rather than from structural descriptions.

Some Ways That Graphics Communicate

**UNDERLYING CONCEPTUAL STRUCTURE: DEPICTIVE AND DESCRIPTIVE UNITS**

These three studies point to powerful correspondences between depictions and descriptions of the same conceptual material. The correspondences in turn suggest that depictions and descriptions are similarly schematized, because both are driven by the same conceptual analysis of the domain. For example, in order to construct an external representation of instructions to get from A to B, the route is first schematized to an ordered list of actions around landmarks and route segments. Either a depiction or a description can then be constructed from the schematization. For route directions, then, the correspondence between depictions and descriptions is at the level of words or phrases and the structured linking of them.

For graphs, the correspondence between depictions and descriptions is at a level more general than the level of words. For interpreting or producing bars and lines, the common underlying conception is of the relationship between the variables, as a trend or discrete comparison. For diagrams, the correspondence between depictions and descriptions is yet more general, at the level of conceiving of the entire system as a structural one or a causal one.

Despite these variations in the generality of the correspondences between depictions and descriptions, in all cases, perceptual units—lines, curves, containers, arrows—map onto conceptual structures. The underlying meaning suggests a natural way of interpreting graphics as well as a natural way of constructing graphics from an interpretation.

**Graphic Space**

Like the graphic elements themselves, the space between graphic elements has also been used to convey meaning in cognitively compelling ways. An example so obvious that it typically goes unnoticed is the space between words. There was a time when language was written as strings of letters without breaks between words. Grouping the letters that belong to one word separately from those that belong to another by a spatial device, an empty space between the letters, makes it easier to distinguish the words. Written language has other examples of spatial devices that convey meaning naturally. For example, ideas are separated by paragraphs, which are signaled by indentation and/or skipping a line, and outlines make successive use of indentation to signal subordination.

Space can be used meaningfully at several levels, depending on the degree of spatial information that is intended to be conveyed. The weakest level is the nominal, or categorical, level, where things are merely separated into groups by a common feature or features, like the letters that
belong to different words (see Stevens [1946] for a discussion of scale types). Stronger constraints come when order is indicated spatially, as in indentation for paragraphing or successive indentation in outlines. Order can be indicated in other ways. A straightforward way is to order things in a list in the way they are ordered on some other variable: children by age, scientific discoveries by dates, countries by GNP, groceries by route through the store. Partial orders are commonly represented by hierarchical trees, where one direction, usually horizontal, is meaningful, and the other is not. Still more information is represented when space is used intervally. In many X-Y graphs, not only is the order of elements in space meaningful, but also the distances between the elements. Interval representations allow inferences such as that the age time between scientific discoveries and commercial applications is getting shorter and shorter. Finally, when a zero point located in space is meaningful, then ratios of spatial distances between elements are also meaningful. Ratio representations allow inferences such as that the distance between Chicago and San Francisco is more than twice the distance between Chicago and New York. With the notable exception of pie charts, which are appropriate for ratio relations but not for interval ones, uses of space to represent interval and ratio relations are often the same, differing in the interpretation.

Let us now turn to some examples of graphic inventions that use space in these ways. The examples come primarily from cross-cultural studies on children of varying ages. The historical inventions that will be discussed echo those of the children.

**CHILDREN INVENT WRITING**

In a series of studies, Lillianna Tolchinsky Landsmann and Iris Levin asked preliterate children from several different cultures, all with alphabetic scripts, to take dictation (Tolchinsky Landsmann and Levin 1986, 1987). Of course, there was no expectation that they would write real words. Rather, the goal was to characterize the graphic symbol systems the children would invent. Naturally, their inventions were not pure, as the children had been exposed to writing even if they did not know how to decipher the code, or even what the code was. Early on, children wrote down one mark for each word, much like the ideographic scripts that preceded the alphabet in which each character corresponded to a word. Older children often used several characters for each word, much like the scripts they would eventually learn. The words children invented often resembled the concepts they were representing. For example, larger concepts got larger words, and when a choice of colors was given, the choice corresponded to the color of the thing represented. In some cases, written words were longer for concepts that took longer to say, again indicating that the children were absorbing something of the nature of the alphabetic scripts they would acquire. In all cases, spaces separated words. Children began writing from the top of the page and always placed words on a line.

These features characterize all written languages as well, although sometimes the lines are horizontal, sometimes vertical, sometimes beginning at the left, sometimes at the right, but always from the top.

The devices that children inventing writing produce, then, strongly resemble the devices invented across cultures for the same purposes. Of course, children’s exposure to the writing systems of the surrounding culture may have biased at least some of their inventions, but nevertheless, the ubiquity of the inventions suggests that they are cognitively compelling.

**CHILDREN INVENT GRAPHS**

How do children use space to convey abstract concepts? We investigated this by providing children with square pieces of paper and stickers, and asking them to arrange the stickers in space to represent a number of relative concepts: time, quantity, and preference (Tversky, Kugelmass, and Winter 1981). Children were first acquainted with the task by representing spatial relations—specifically, the arrangement of small dolls on a line. Then, to elicit representations of time, the experimenter sat next to the child and asked the child to think about the times of day when the child ate breakfast, lunch, and dinner. The experimenter then put a sticker down in the middle of a blank, square piece of paper to represent the time for eating breakfast, and asked the child to put down another sticker showing the time for eating lunch and another sticker for the time for eating dinner. Other time questions followed. One question about quantity asked the child to think about the amount of candy in a handful, the amount of candy in a bagful, and the amount of candy collected at Halloween. A preference question asked the child to think about a food the child loved, a food the child didn’t like, and a food somewhere in between. The participants included Hebrew-speaking Israelis, Arabic-speaking Israelis, and Americans. These cultures are of interest partly because of the directions of their writing systems. English is written from left to right, whereas Hebrew and Arabic are written right to left. The right to left tendencies are much stronger in Arabic than in Hebrew for several reasons: Arabic script is connected and each character is formed from right to left, whereas Hebrew has no script and most letters are formed from left to right. In Hebrew, the arithmetic system follows the Western left-to-right order, but in Arabic, arithmetic is taught from right to left until the middle-school years when the Western conventions are adopted. Participants ranged in age from preschool to college.

One question of interest was how much information from the conceptual relations the children preserved in the spatial mappings? Would this vary with age or culture? In fact, the amount of information preserved in the mappings increased with age, but did not vary across culture. Some of the youngest children only preserved nominal or categorical relations in their mappings; that is, they regarded time for breakfast, lunch, or dinner or preference for television shows as different, but not as on an ordered
continuum. These children placed their stickers haphazardly on the page. Such children were unusual. Most of even the youngest children ordered the stickers on a line, mappings that preserved ordinal information. Some of the relations children were asked to map had clearly unequal distances between the items; for example, time to wake up, time to go to school, and time to go to bed. Despite the clear differences, the older children’s mappings preserved interval information. To test the limits of this, a new group of children was given special procedures designed to call attention to the different intervals and to elicit interval mappings. For the most part, these failed. By twelve years, however, children began to map intervals spontaneously.

Another question asked of the data concerned the directionality of the increases. Would they vary with language or with concept? In fact, directionality varied both with language and with concept. For preference and quantity, increases were mapped approximately equally across cultures from right to left, from left to right, and from bottom to top. The only direction to indicate increases that was avoided was top to bottom. These practices reflect what seem to be biases about horizontal and vertical space. Horizontal space is neutral—the right and left halves of the body are relatively symmetric, especially in comparison to the top and bottom or front and back halves, which are clearly different. What is on the right and what is on the left in the space around the body are for the most part arbitrary, an accident of one’s current point of view. What is up and what is down in the world, by contrast, are no accident. What is up defies gravity, exhibits strength. People grow stronger as they grow taller. Larger piles, of goods or money, are higher. So “up” is associated with more, better, stronger. This apparently natural association between space and meaning is reflected in language and gesture as well. We say that someone is at the top of the world or the top of the heap and we give a thumbs up or a high five.

Spatial displays of temporal relations, however, were different. All groups mapped temporal order from down to up. They also used the horizontal axis, but the specific direction depended on the direction of writing in the language. Many English speakers plotted increases in time from left to right, whereas Arabic speakers tended to plot temporal increases from right to left. As noted, in Hebrew, writing is less lateralized than in Arabic; similarly, the graphic mappings of Hebrew speakers were also more evenly distributed from right to left and from left to right. Some of the Arabic speakers, and only the Arabic speakers, mapped later times to lower. This, of course, corresponds to the way calendars and date books are organized: the beginning, the earliest time at the top.

Spatial relations, then, can be used spontaneously to convey abstract, nonspatial relations by children and adults in different cultures. Space is readily used to convey categorical and ordinal relations; it may also be used to convey interval relations. There is some consistency in the spatial direction used to convey increases. First, increases are nearly always conveyed vertically or horizontally, not diagonally or circularly. The particular horizontal direction is neutral except for the case of temporal increases, wherein horizontal direction tends to follow writing order. As for the vertical axis, increases in quantity and preference (but notably not in time) overwhelmingly correspond to upwards direction, in correspondence with language and gesture, where up indicates more, better, and stronger. Moreover, these correspondences seem grounded in the world, where, in general, more things make higher piles and stronger things are taller.

Depicting Abstractions

Graphics have been produced by different cultures throughout history for different ends. They portray reality and myth; they record history and present proclamations; they convey models of things and systems. To convey these various meanings, they use characteristics of elements as well as the spatial relations among elements. Many of these depictive devices have been invented and reinvented across cultures and ages for similar meanings. As such, they are readily interpreted, even if novel; thus, they appear to have a degree of cognitive naturalness. Elements may resemble the elements to be mapped, or they may represent them figuratively. As the preceding studies have shown, a small group of geometric forms may be used to convey abstract meanings directly. For example, in sketch maps, straight-line segments indicate straight roads, curved-line segments indicate curved roads, crosses indicate intersections, and blobs suggest environmental structures. Note that these depictive units, like language, are schematic or categorical; none normally captures exact metric relations. In graphs, lines link and bars separate, so that line graphs are interpreted and produced for trends, whereas bar graphs are interpreted and produced for discrete relations. In diagrams, arrows indicate temporal sequence from which causal sequence is readily inferred. The presence of arrows in diagrams encourages causal—over and above structural—interpretation.

Spatial relations among elements are also readily produced and interpreted using the basic underlying metaphor that proximity in space reflects proximity in an abstract dimension. Graphic space may preserve abstract spaces at several levels of information.

At the nominal or categorical level, items are separated into groups that share a common feature or features, but there is no relationship implied between groups. At the ordinal level, items have differing values on the same underlying feature, yet the distance between them is not intended to be meaningful. Distance between items is meaningful at the interval level (and ratios of distances at the ratio level, where zero is meaningful).

General abstract meanings may also be expressed in depictions. Equiv-
alence, for example, can be expressed by grouping items that are equivalent and spatially separating them from items that are not. Equivalence can also be indicated pictorially, by various frames such as boxes, bars, and parentheses and by similar appearance, as in fonts, sizes, colors. Connections between items can be depicted by lines of various sorts. Order among items may be indicated both spatially and pictorially as in indentation of paragraphs or in outlines as well as the order in which a set of items is listed—for example, children in order of age, groceries in order of a path through the supermarket. Pictorially, orders, especially partial orders, can be represented as trees. Similarly, degrees of relationship such as similarity, salience, or strength can be suggested, for example, spatially by degree of proximity, or pictorially by degree of appearance, color, or size. Proportion can be indicated by spatial proportion. Direction is conveniently conveyed by arrows, whether the direction is spatial, temporal, causal, or other.

Change can be easily expressed by actual change, as in animation, though animations can present both perceptual and cognitive difficulties (see Tversky, Morrison, and Betancourt [forthcoming]). Other effective ways to indicate change are by artfully selected successive stills, as in comics (McCloud 1984), instructional materials (Zacks and Tversky forthcoming), or flowcharts with arrows.

Conclusion

Words are certainly the prototypic medium of communication. They can be concrete or abstract, succinct or expansive. They can be audible or viewable and are portable. But words bear only symbolic relations to the concepts they represent—therein lies their limits and their power, the power of abstraction. By contrast, depictions use elements and the spatial relations among them to convey concrete and abstract meanings quite directly. Using space and the elements in it to convey meaning also capitalizes on the impressive capacity people have to process and store spatial and visual information—therein lies the powers of depictions. They can grab and keep attention by being attractive or humorous or frightening. They can demonstrate knowledge directly rather than indirectly, as in maps and models. They can promote inferences based on spatial and visual reasoning. They can serve as external representations of thought, alleviating the mental-processing load. As external representations, they are open to a community of users who can inspect, reinspect, and revise them. There is more to depictions than meets the eye.

References

Navigating by Mind and by Body

Barbara Tversky

Stanford University
Department of Psychology
420 Jordan Hall
Stanford, Ca 94305
(bt@psych.stanford.edu)

Abstract. Within psychology, at least two research communities study spatial cognition. One community studies systematic errors in spatial memory and judgement, accounting for them as a consequence of and clue to normal perceptual and cognitive processing. The other community studies navigation in real space, isolating the contributions of various sensory cues and sensori-motor systems to successful navigation. The former group emphasizes error, the latter, selective mechanisms, environmental or evolutionary, that produce fine-tuned correct responses.

How can these approaches be reconciled and integrated? First, by showing why errors are impervious to selective pressures. The schematization that leads to errors is a natural consequence of normal perceptual and cognitive processes; it is inherent to the construction of mental spaces and to using them to make judgments in limited capacity working memory. Selection can act on particular ways of errors, yet it is not clear that selection can act on the general ways that produce them. Next, in the wild, there are a variety of
c errors, for example, over-shooting in dead reckoning across species.

Finally, closer examination of navigation in the wild shows
too, environments may provide corrections, specifically, landmarks.

general cognitive mechanisms generate general solutions. The errors inevitably

produced may be reduced by local specific sensori-motor couplings as well as
local environmental cues. Navigation, and other behaviors as well, are a
consequence of both.

1 Two Research Communities in Psychology

Yes, the title evokes the mind-body problem. However one regards the venerable monumental mind-body problem in philosophy, there is a contemporary minor mind-body problem in the psychological research on spatial cognition. While the major

---

1 I am grateful to Christian Freksa for helpful comments and encouragement and to two anonymous reviewers for critiques of an earlier version of this manuscript. Preparatory of the manuscript was supported by Office of Naval Research, Grants Number NONO14-PP-1-O649 and N000140110717 to Stanford University.
problem is how to integrate the mind and the body, an additional minor problem in spatial cognition is how to integrate the approaches—and the researchers—on the mind and on the body. The community studying spatial judgments and that studying wayfinding rarely interact. Or have rarely interacted. These conferences of minds may be a meeting point and a turning point.

The two communities, the mind community, and the body community, differ in their agendas and differ in the tools to carry them out. The mind community studies spatial judgments: what is the direction between San Diego and Reno? How far is Manchester from Glasgow? Manchester from Liverpool? The Eiffel Tower to Jaque’s house? How do I get to San Marco? The questions are cleverly chosen. They are designed to yield errors. The design works because the errors are a consequence of the way spatial information is represented and used. In fact, one goal of this approach is to reveal those cognitive representations and mechanisms, many of which appear not only in spatial judgments, but in other domains as well (e.g., Tversky, 1993; 2000a; 2000b).

In contrast, the body community studies the cues, visual, auditory, kinesthetic, vestibular, that people and animals use to arrive at their destinations. The research reduces the sensory input and diminishes the environmental richness in order to isolate the role of a particular cue or system in guiding the organism. In many cases, the goal is to reveal the elegant fine-tuning of a particular cue or sets of cues or sensory-motor systems to specific aspects of environments (see, for examples, Gallistel, 1990 and papers in the volume edited by Golledge, 1999, especially the paper by Berthoz, Amorim, Glassauer, Grasso, Takei, and Viaud-Delmon and Loomis, Klatzky, Golledge, and Philbrick).

To caricature the approaches, the emphasis of the mind community is to reveal the systems generating error and the emphasis of the body community is to reveal the systems generating precision.

No wonder the community of mind and the community of body pass each other by like the proverbial ships in the night. They differ in the tasks they give, in the responses they collect, in the processes they propose to account for the responses to the tasks. And, perhaps most significantly, they differ philosophically, in their fundamental attitudes toward human nature. For the mind group, being human is fundamentally about limitations, limitations in representations and in processing, in capacity and in computation. Those limitations can be revealed in errors. The errors provide clues to normal operations. For the body group, being human is fundamentally about evolution and learning, about selection and adaptation, pressures toward perfection. Again, these are caricatures of the positions, hence not attributed to any of the fine reasonable people in the fields, but caricatures that are close enough to the truth to warrant further discussion. And perhaps, approachment, even integration, of the approaches.

Neither evolution nor adaptation are doubted. Both communities believe that organisms have evolved in and continue to live in environments, and that the environments have selected successful behaviors across the millennia through evolution and across the lifespan through learning. So the real puzzle is not why some spatial behaviors are exquisitely precise and fine-tuned, but rather why systematic errors persist. Before that question can be addressed, a review of some of the documented errors is in order. Then these errors must be accounted for by an analysis of the general mechanisms that produce and maintain them.

2 Systematic Distortions of Distance and Direction

2.1 Errors of Distance

First, what errors do we mean? Errors of distance estimates, for one. They are affected by irrelevant factors, such as hierarchical organization. Elements, like cities or buildings, within the same group are perceived as closer than those in different groups. The groups might be states or countries. The groups need not be geographic; they can be functional or conceptual. Distances between a pair of academic buildings or a pair of commercial buildings in Ann Arbor are perceived as shorter relative to distances between an academic and a commercial building (Hirtle and Jonides, 1981). Arabs perceive distances between pairs of Arab settlements to be smaller than distances between an Arab and a Jewish settlement; similarly, Jews perceive distances between Jewish settlements to be shorter than distances between an Arab and a Jewish settlement (Portugali, 1993). Grouping is reflected in reaction times to make distance estimates as well; people are faster to verify distances between geographic entities such as states or countries than within the same entity (e.g., Maki, 1981; Wilton, 1979). Another factor distorting distance estimates is the amount of information along the route. Distance judgments for routes are judged longer when the route has many turns (e.g., Sadalla and Magel, 1980) or landmarks (e.g., Thorndike, 1981) or intersections (e.g., Sadalla and Staplin, 1980). Similarly, the presence of barriers also increases distance estimates (e.g., Newcombe and Liben, 1982). Most remarkably, distance judgements are not necessarily symmetric. Distances to a landmark are judged shorter than distances from a landmark to an ordinary building (Sadalla, Burroughs, and Staplin, 1980; McNamara and Djwadik, 1997). Similar errors occur for prototypes in similarity judgments: people judge atypical magenta to be more similar to prototypic red than red to magenta (Rosch, 1975). Landmarks seem to define neighborhoods and prototypes categories whereas ordinary buildings and atypical examples do not. Ordinary buildings in the vicinity of a landmark may be included in the neighborhood the landmark defines.

2.2 Errors of Direction

Systematic errors occur for judgments of direction as well. Hierarchical organization is again a factor. For example, the overall direction between pairs of states appears to be used to judge the direction between pairs of cities contained in the states. The example so famous that it has become a Trivial Pursuit question is the direction between San Diego and Reno. Students in San Diego erroneously indicated that San Diego is west of Reno (Stevens and Coupe, 1978). That is, the overall direction of the
states is used to infer the directions between cities within those states. But errors of direction occur within groups as well, for example, informants incorrectly report that Berkeley is east of Stanford (Tversky, 1981). This error seems to be due to mentally rotating the general direction of the surrounding geographic entity, in this case, the south Bay Area to the overall direction of the frame of reference, in this case, north-south. In actuality, the south Bay Area runs nearly diagonally with respect to the overall frame of reference, that is, northwest to southeast. Geographic entities create their own set of axes, typically around an elongated axis or an axis of near symmetry. The axes induced by the region may differ from the axes of its external reference frame. Other familiar cases include South America, Long Island, Japan, and Italy. In this error of rotation, the natural axes of the region and those of the reference frame are mentally brought into greater correspondence. Directions also get straightened in memory. For example, asked to sketch maps of their city, Parisians drew the Seine as a curve, but straighter than it actually is (Milgram and Jodelet, 1976). Even experienced taxi drivers straighten the routes they ply each day in the maps they sketch (Chase and Chi, 1981).

2.3 Other Errors

These are not the only systematic errors of spatial memory and judgment that have been documented; there are others, notably, errors of quantity, shape, and size, as well as errors due to perspective (e.g., Tversky, 1992; Poulton, 1989). Analogous biases are found in other kinds of judgements: for example, people exaggerate the differences between their own groups, social or political, and other groups, just as they exaggerate the distances between elements in different geographic entities relative to elements in the same geographic entity. The errors are not random or due solely to ignorance; rather they appear to be a consequence of ordinary perceptual and cognitive processes.

3 Why Do Errors Exist?

3.1 Schematization Forms Mental Representations

A number of perceptual and cognitive processes are involved in establishing mental representations of scenes or depictions, such as maps. Isolating figures from grounds is one of them; figures may be buildings or roads, cities or countries, depending on what is represented. Figures are then related to one another and to a frame of reference from a particular perspective (e.g., Tversky, 1981; 1992; 2000a). Natural as they are, essential as they are, these perceptual organizing principles are guaranteed to produce error. They simplify, approximate, omit, and otherwise schematize the geographic information. Schematization thereby produces error.

How does this happen? Consider these examples. Relating figures to one another draws them closer in alignment in memory than they actually are. Evidence comes from a task where students were asked to select the correct map of the Americas from a pair of maps in which one was correct and the other had been altered so that South America was more aligned with North America. A majority of students selected the more aligned map as the correct one (Tversky, 1981). The same error was obtained for maps of the world, where a majority preferred an incorrect map in which the U.S. and Europe were more aligned. Alignment occurred for estimates of directions between cities, for artificial maps, and for blobs. Relating a figure to a reference frame yields the rotation errors described in the section on errors of direction. Like alignment, rotation occurs for directions between cities, for artificial maps, and for blobs.

3.2 Schematization Allows Integration

Many environments that we know, navigate, and answer questions about are too large to be perceived from a single point. Acquiring them requires integrating different views as the environment is explored. Even perceiving an environment from a single point requires integration of information, from separate eye fixations, for example. How can the different views be integrated? The obvious solution is through common elements and a common reference frame. And these, elements and reference frames, are exactly the schematizing factors used in scene perception. To make matters more complex, knowledge about environments comes not just from exploration, but from maps and descriptions as well, so the integration often occurs across modalities. Again, the way to link different modalities is the same as integrating different views, through common elements and frames of reference.

3.3 Schematization Reduces Working Memory Load

A third reason for schematization is that the judgments are performed in working memory, which is limited in capacity (e.g., Baddeley, 1990). Providing the direction or distance or route between A and B entails retrieving the relevant information from memory. This is unlikely to be in the form of a prestored, coherent memory representation, what has been traditionally regarded as a cognitive map. More likely it entails retrieving scattered information and organizing it. Moreover, whatever is stored in memory has already been schematized. All this, and the judgment as well, is accomplished in working memory. Like mental multiplication, this is burdensome. Anything that reduces load is useful, and schematization does just that. This is similar to reducing bandwidth by compression, but in the case of constructing representations in working memory, the compression is accomplished by schematization, by selecting the features and relations that best capture the information.
3.4 Spatial Judgments Are Typically Decontextualized

Unlike navigation by the body, navigation in the mind is without support of context. This is in sharp contrast to the spatial behaviors that are precise, accurate, and finely-tuned, such as catching balls, playing the violin, wending one's way through a crowd, finding the library or the subway station. Context provides support in several ways. First it provides constraints. It excludes many behaviors and encourages others. The structure of a violin constrains where the hands, fingers, chin can be placed and how they can be moved. The structure of the environment constrains where one can turn, where one can enter and exit. The world does not allow many behaviors that the mind does. Second, natural contexts are typically rich in cues to memory and performance. For memory, contexts, like menus on computer screens, turn recall tasks into recognition tasks. A navigator doesn't need to remember exactly where the highway exit or subway entrance is as the environment will mark them. The presence of context means that an overall plan can leave out detail such as exact location, direction, and distance. In fact, route directions and sketch maps leave out that level of detail, yet have led to successful navigation across cultures and across time (e.g., Tversky and Lee, 1998, 1999). For performance, context facilitates the specific actions that need to be taken. In the case of playing the violin, this includes time and motion, the changing positions of the fingers of each hand. In the case of wayfinding, this also includes time and motion of various parts of the body, legs in walking, arms, hands, and feet in driving.

4 Why Do Errors Persist?

4.1 Rarely Repeated

Context and contextual cues provide one reason why spatial behaviors by the body may be highly accurate and spatial behaviors by the mind biased. Context constrains behaviors and cue behaviors. Contexts are also the settings for practice. As any violin player or city dweller knows, the precise accurate spatial behaviors become so by extensive practice. The efforts of beginners at either are full of false starts, error, and confusion. Practice, and even more so, practice in a rich context supporting the behavior, is the exception, not the rule, for navigation by the mind, for judgements from memory. Indeed, for the judgments that we are called upon to make numerous times, we do eventually learn to respond correctly. I now know that Rome is north of Philadelphia and that Berkeley is west of Stanford.

4.2 Learning Is Specific, Not General

But knowing the correct answer to a particular case corrects only that case, it does not correct the general perceptual and cognitive mechanisms that produce schematizations that produce the errors. Knowing that Rome is north of Philadelphia doesn't tell me whether Rome is north of New York City or Boston. Knowing that Rome is north of Philadelphia doesn't inform me about the direction from Boston to Rio either. Learning is local and specific, not general and abstract. Immediately after hearing an entire lecture on systematic errors in spatial judgments, a classroom of students made exactly the same errors.

The mechanisms that produce the errors are multi-purpose mechanisms, useful for a wide range of behaviors. As noted, the mechanisms that produce errors derive from the mechanisms used to perceive and comprehend scenes, the world around us. The schematizations they produce seem essential to integrating information and to manipulating information in working memory. In other words, the mechanisms that produce error are effective and functional in a multitude of ways.

4.3 Correctives In Context

Another reason why errors persist is that they may never be confronted. Unless I am a participant in some abstruse study, I may never be asked the direction between Rome and Philadelphia, from Berkeley to Stanford. Even if I am asked, I may not be informed of my error, so I have no opportunity to correct it. And if I am driving to Berkeley, my misconception causes me no problem; I have to follow the highways. Similarly, if I think a particular intersection is a right-angle turn when in fact it is much sharper, or if I think a road is straighter than it is, the road will correct my errors, so I can maintain my misconception in peace. In addition, these errors are independent of each other and not integrated into a coherent and complete cognitive map, so there is always the possibility that errors will conflict and cancel (e.g., Baird, 1979; Baird, Merrill, and Tannenbaum, 1979). Finally, in real contexts, the extra cues not available to working memory become available, both cues from the environment, like landmarks and signs, and also cues from the body, kinesesthetic, timing, and other information that may facilitate accuracy and overcome error. In short, schematic knowledge, flawed as it is, is often adequate for successful navigation.

5 Systematic Errors in the Wild

Now the caricature of the communities that has been presented needs refinement. Despite millennia of selection by evolution and days of selection by learning, navigation in the wild is replete with systematic errors. One studied example is path integration. Path integration means updating one's position and orientation while navigating according to the changes in heading and distances traveled, the information about one's recent movements in space (Golledge, 1999, p. 122). A blindfolded navigator traverses a path, turns, continues for a while, and then heads back to the start point. How accurate is the turn to home? Ants are pretty good, so are bees, hamsters, and even people. But all make systematic errors. Bees and hamsters overshoot (Etienne, Maurer, Georgakopoulos, and Griffin, 1999). People overshoot small distances and small turns and undershoot large ones (Loomis, Klatzky,
Golledge, and Philbeck, 1999), a widespread error of judgment (Poulton, 1989). But the situation that induced the errors isn’t completely wild; critical cues in the environment have been removed by blindfolding or some other means. In the wild, environments are replete with cues, notably, landmarks, that may serve to correct errors.

6 Implications

How do people arrive at their destinations? One way would be to have a low-level, finely-detailed sequence of actions. But this would only work for well-learned routes in unchanging environments; it wouldn’t work for new routes or vaguely known routes or routes that encounter difficulties, such as detours. For those, having a global plan as well as local actions seem useful. These are global and local in at least three senses. Plans are global in the sense of encompassing a larger environment than actions, which are local. Plans are also global in the sense of being general and schematic, of being incompletely specified, in contrast to actions, which are specific and specified. Plans are global in the sense of being amodal, in contrast to actions, which are precise movements of particular parts of the body in response to specific stimuli. A route map is a global plan for finding a particular destination, much as a musical score is a global plan for playing a particular piece on the violin. Neither specifies the exact motions, actions to be taken.

Several approaches to robot navigation have recommended the incorporation of both global and local levels of knowledge (e.g., Chown, Kaplan, and Kortenkamp, 1995; Kuipers, 1978, 1982; Kuipers and Levitt, 1988). The current analysis suggests that global and local levels differ qualitatively. The global level is an abstract schematic plan, whereas the local level is specific sensori-motor action couplings. Integrating the two is not trivial.

The gap between the mind navigators and the body navigators no longer seems so large. True, the focus of the mind researchers is on judgments and the challenge is to account for error and while the focus of the body researchers is on behavior and the challenge is to account for success. Yet, both find successes as well as systematic errors. And in the wild, the correctives to the errors are similar, local cues from the environment.

Systematic errors persist because the systems that produce them are general: they are useful for other tasks and they are too remote to be affected by realization of local, specific errors. Spatial judgment and navigation are not the only domains in which humans make systematic errors. Other accounts have been made for other examples (e.g., Tversky and Kahneman, 1983). It makes one think twice about debates about the rationality of behavior. How can we understand what it means to be rational if under one analysis, behavior seems replete with intractable error but under another analysis, the mechanisms producing the error seem reasonable and adaptive.

References


