Why do Speakers Mix Perspectives?*

BARBARA TVERSKY[1], PAUL LEE[1] and SCOTT MAINWARING[2]
1Department of Psychology, Stanford University, U.S.A.
2Interval Research Corporation

Abstract. Although considerations of discourse coherence and cognitive processing suggest that communicators should adopt consistent perspectives when describing spatial scenes, in many cases they switch perspectives. Ongoing research examining cognitive costs indicates that these are small and exacted in establishing a mental model of a scene but not in retrieving information from a well-known scene. A perspective entails a point of view, a referent object, and terms of reference. These may change within a perspective, exacting cognitive costs, so that the costs of switching perspective may not be greater than the costs of maintaining the same perspective. Another project investigating perspective choice for self and other demonstrates effects of salience of referent object and ease of terms of reference. Perspective is mixed not just in verbal communications but also in pictorial ones, suggesting that at times, switching perspective is more effective than maintaining a consistent one.

Key Words. perspective, referent object, frame of reference, route, survey, deictic, intrinsic, spatial cognition, spatial mental model

Perspective

When people talk about space, or anything for that matter, they take a perspective on it. There are many different perspectives that speakers can take, but it has been generally assumed that in a particular discourse, speakers adopt one perspective consistently. There is an a priori reason for consistency of perspective: a consistent perspective provides coherence for a description, a framework for the discourse. There is also a practical reason for maintaining the same perspective: perspective switching can have cognitive costs, at least for the listener or reader. When one sentence uses "come," and the next, "go," readers take more time to comprehend the sentences than when the viewpoint of the narrative is consistent (Black, Turner, & Bower, 1979)

Perspective is one of those extraordinarily useful words that, as a consequence, has many senses. Some senses are strictly spatial, but others are broader. The concern here is with spatial perspective. Even within the domain of space, *perspective* has been used narrowly and broadly.

Perspective, narrowly viewed. Narrowly, perspective refers to the reference systems that are used to describe the location of a target object in space. At a minimum, a reference system includes a referent object or frame and terms of reference; it may also include a second referent object and a viewpoint. A compelling way to distinguish reference systems is by the type of referent object: for the deictic system, the speaker; for the intrinsic system,

* We are grateful for the insightful comments of two anonymous reviewers.

an object with inherent front, back, top, bottom, left, and right; for the extrinsic system, an external reference frame, determined by the environment (e. g., Levelt, 1984). Researchers investigating way-finding have also distinguished egocentric frameworks based on the wayfinder and allocentric frameworks based on an environment.

In language, reference systems distinguished by speakers, objects, or environments have posed conceptual difficulties. Because speakers, like some objects, have inherent sides, deictic thus designated is formally indistinguishable from intrinsic. However, a speaker can serve as the only referent object or a speaker can serve as a referent object in conjunction with another referent object. In the first case, a speaker may say "the fountain is in front of me," a usage parallel to the intrinsic use, "the fountain is in front of City Hall." In the second case, a speaker may say, 'the fountain is left of City Hall." Here, there is a ternary relation; the fountain is left of City Hall from the viewpoint of the speaker. Levinson (1996) called the latter case "relative." It subsumes some deictic uses, but leaves the simple deictic uses in the same category as intrinsic. Levinson's analysis yields three separable systems, relative, intrinsic, and absolute (extrinsic), but without the elegant alignment of reference system to type of referent object. It has had the additional virtue of conferring a peace treaty in the deixis wars.

Perspective, broadly viewed. More broadly, perspective has been used for the reference systems adopted in extended spatial descriptions, for example, those used to describe a room or an apartment or a town. Three basic perspectives have been distinguished. In a *gaze* tour, typically adopted in describing a room from an entrance, the speaker takes a fixed point of view level with but external to an environment, and describes objects relative to other objects from the point of view of the speaker, using right, left, front, and back (Ehrich & Koster, 1983). In a route tour, often invoked to describe an environment that is too large to be seen from a single viewpoint, such as an apartment, the environment is described from the changing viewpoint of a traveler in the environment, typically called "you." Objects are described relative to the traveler in terms of "your" left, right, front, and back (Perrig & Kintsch, 1985; Taylor & Tversky, 1992a, b, 1996). Yet another way to describe an environment too large to be seen from a single embedded viewpoint is with a survey perspective. In a survey perspective, the speaker takes a fixed viewpoint above the environment and describes objects relative to other objects using environmental directions, north, south, east, and west (e. g., Perrig & Kintsch, 1985; Taylor & Tversky, 1992a, b, 1996).

The narrow and broad perspectives distinguished here map to one another. Each of the extended description perspectives corresponds to a narrow perspective: gaze to relative, route to intrinsic, and survey to absolute. Note also that each corresponds to a natural way of experiencing an environment: gaze from a fixed external viewpoint, route from the changing viewpoint of a traveler, and survey from a fixed viewpoint above (Taylor & Tversky, 1996).

Inconsistency in Use of Perspective. Despite the a priori and practical rationales for consistency of perspective, when asked to produce descriptions of a wide variety of environments, people mix perspectives in extended descriptions of environments about half the time (Taylor & Tversky, 1992a; 1996). In addition, in studies of comprehension, when extended descriptions of environments from either route or survey perspectives are well-learned, people respond to (inference, not verbatim) statements from either the read (route or survey) or the new (survey or route) perspective equally quickly and accurately (Taylor & Tversky, 1992b). These findings suggest that if there are cognitive costs of switching perspective, they disappear when environments are well-learned and presumably abstracted into a perspective-free representation. But what about on-line comprehension? Might there be costs to perspective-switching then? The studies of Black, Turner, and Bower (1979) suggest that there should be comprehension costs to switching perspective. In ongoing research, Lee and Tversky have found costs for switching perspective in on-line reading times of descriptions of small environments (Lee & Tversky, in preparation).

Studies of Comprehension of Perspective Switches

Comprehension. Reading Times are Expedited by Consistency of Perspective. In Lee and Tversky's experiments, participants read a short introduction to an environment followed by three sentences describing an environment from either a route or survey perspective. Each environment had a route and a survey instantiation for different participants. Here is an example of a survey description. First, the introduction, which opened both survey and route instantiations: "Bonanmpak is one of the lesser known Mayan temples, nestled in the heart of the jungle. The entrance to Bonanmpak leads to Eagle-Snake Path, which is a small path that is still partially covered by jungle foliage". The survey instantiation proceeded: "Eagle-Snake Path runs north-south, intersecting a main pathway called Road of the Dead, which runs east-west. West of Eagle-Snake Path on the north side of Road of the Dead is the Palace of the Kings. West of Eagle-Snake Path on the south side of Road of the Dead is the Sun Temple." After the three survey sentences, half the participants read a fourth sentence in the same perspective: "On the south side of the Road of the Dead, west of the Sun Temple is the Moon Temple." The other half of the participants read a sentence from a route perspective: "From the Road of the Dead when you face the Sun Temple, on your right is the Moon Temple."

During construction of a spatial mental model from text, switching perspective slowed reading times. Reading times per syllable for the consistent survey target sentence averaged 471 ms but averaged 585 ms for the inconsistent route target sentence. Now for an example of a route description, beginning with the introduction: "The Saturday morning fruit market has a large variety of summer fruits. Once you enter the market, there is a large main aisle with large crates of fruits on each side." Then the three route sentences; the first contains a cardinal direction to make the survey sentence comprehensible. "If you go south on the

main aisle, you will see a side aisle that crosses it. Turn left on the side aisle and you your left, you will see a crate full of apricots. On your left, on the side aisle, in front of the apricots, you will see a crate full of grapes." The consistent route target was: "From the side aisle when you face the grapes, behind you are the plums." The inconsistent survey target was: "On the south side of the side aisle, south of the grapes are the plums." As for switching from survey to route, switching from route to survey exacted a cost in reading times. Average time per syllable for the consistent route target was 328 ms, but was 560 ms for the inconsistent survey target. The route target, when consistent, was read considerably faster than the survey target when consistent. This is unlikely to be due to something inherent in the route target as the route advantage holds only for consistent targets. The most probable explanation for the advantage of the route target for consistent descriptions is that in the route description, the sentence prior to the target sentence indirectly oriented the reader toward one of the landmarks so that that part of the information in the route target was redundant to those who had read the route description. The main effect of interest, the cost of switching perspective was, on average, more than 150 msec per syllable in reading time.

Immediate Verification Times are Expedited by Consistency of Perspective. The target sentences were followed by four statements from both perspectives for true-false verification. These assured that participants did in fact construct accurate mental models of the environments. They also provided additional evidence for costs of perspective switching, and for diminishing costs. As for reading times, verification times were slower to different-perspective statements than to same perspective statements; however, for verification times, the effect of switching perspective decreased to 50 ms per syllable averaged over the four statements. A decreased effect of perspective switching is to be exected as by the second or third true/false statement, all readers will have experienced both perspectives. There are at least two non-conflicting possibilities to explain the decrease. First, through repeated questions, readers could get used to understanding the alternative terms of reference. This is unlikely to be a complete explanation as participants experience both perspectives over trials. More likely is that as they answer questions, readers' mental models of the situation become better integrated and more perspective-free, as in the previous experiments of Taylor and Tversky (1992b).

These results were replicated in a second study. In that study, on half the trials prior to the target sentence, participants read a description of one of the landmarks. For the Mayan Temple, that description was: "Eagle-Snake Path has a slight S-shaped curve that somewhat resembles a snake." For the Fruit Market, it was: "The main aisle has many exotic fruits that are grown from tropical regions." The effect of reading a description of a landmark was also to reduce the effect of switching perspective on reading the target sentences. The decrease in cost of perspective switching could be for one or several reasons. The landmark description itself entailed a perspective switch, zooming in and focusing on a particular landmark. The landmark description sentence also momentarily directed attention away

from establishing the layout of the environment, perhaps allowing for integration of the representation of the environment into the more perspective-free form that Taylor and Tversky (1992b) found.

Components of Perspective Switching

It is now clear that during construction of spatial mental models from descriptions, switching perspective from route to survey or survey to route exacts a cost. To understand what that cost may be we need to analyze what perspective switching entails. Because a perspective consists of a referent object or frame, a viewpoint, and terms of reference, each of these changes when perspective is changed. As noted earlier, for a route description, the referent object is typically "you," but for a survey description, the referent object is another object in the scene. Thus, for a route description, the same referent object serves all targets, but for a survey description, the referent object keeps changing. For route descriptions, the viewpoint is embedded in the environment, and keeps changing as the traveler moves about. In contrast, for survey descriptions, the viewpoint is external and above, and fixed, so that the viewpoint stays the same for an entire survey description. Thus, route descriptions have an advantage in that the object doesn't change but survey descriptions have an advantage in that the viewpoint doesn't change. Finally, switching perspective requires switching terms of reference, from left-right-front-back for route descriptions to north-south-east-west for survey descriptions.

Switching perspectives, then, entails changes in referent object, viewpoint, and terms of reference. Presumably, each change has a cognitive cost. Yet, even for discourse with a consistent perspective, successive utterances entail mental transformations with cognitive costs of their own. Let us consider for a moment the sentence to sentence changes that occur within a perspective. In both route and survey descriptions, the target object changes from sentence to sentence as the environment is constructed by the discourse. In addition, for route descriptions, the viewpoint keeps changing, and for survey descriptions, the object keeps changing. This may be the reason why, when constructing spatial descriptions, people switch perspective around half the time (Taylor & Tversky, 1996). In some cases, the costs of switching perspective may not be greater than the cost of staying with the same perspective. Some spatial situations may be more readily described from one perspective rather than another. Indeed, some spatial arrays seem to encourage route perspectives, and others, survey perspectives. Having a single prominent pathway through an environment and landmarks at the same size scale, for example, encourages route descriptions (Taylor & Tversky, 1996).

Because referent objects and viewpoint change systematically within consistent perspectives, survey and route respectively, the most likely cost in switching between perspectives is switching terms of reference, which ordinarily do not change within either perspective. Although we know of no data comparing costs of changing referent objects,

viewpoints, and terms of reference (we are currently collecting such data ourselves), we do know that constructing or computing terms of reference is time-consuming, especially for symmetric direction terms like "left" and "right " (e. g.. Bryant & Tversky, 1991; Farrell, 1979; Franklin & Tversky, 1990; Maki & Braine, 1985; Maki, Grandy, & Hauge, 1979; Sholl & Egeth, 1981). So difficult is computing direction terms that, as we know from data to be described below (Mainwaring, Tversky, Ogishi, & Schiano, 2000), people avoid using them when other options are available.

Production Experiments. Selecting a Perspective. Constructing a sentence locating an object in space entails selecting a target, a referent object, and terms of reference. It also entails selecting a viewpoint. For route and survey descriptions, the viewpoint is typically, though not always, confounded with the terms of reference. In practice, some extended descriptions that in every other way appear to be route descriptions nevertheless slip into cardinal directions on occasion (Taylor & Tversky, 1996). For now and for simplicity, let us consider only selection of target, referent object, and terms of reference. In many cases, the target is selected for the participant, for example, by a query from a lost traveler. In extended spatial descriptions, the participant chooses or is given an environment, but may choose the order of mentioning target objects in that environment. Taylor and Tversky (1992a) argued that selecting the sequence of targets comes prior to selecting a perspective. In their experiments, participants ordered targets in the same way for the same environment, irrespective of the perspective adopted. The order of mentioning targets reflected the spatial organization of the environment, a process that preceded perspective selection. Eliminating selection of target leaves us with selection of a referent object and selection of terms of reference as the two basic components of perspective selection. Let us turn to some studies of production to see how this is done.

Two Principles for Selecting Referent Objects and Terms of Reference.

Before reviewing evidence, let us propose two principles, one for referent objects and one for reference terms. For referent objects, we propose the Principle of Salience: Select a referent object that is salient to communication partners. For terms of reference, we propose the Principle of Ease: Select terms of reference that are relatively easy to produce and comprehend.

Principle of Salience of Referent Object. Like "perspective," "salience" is a useful word with many senses. One sense is perceptual, foremost, relative size. In fact, larger objects are typically selected to serve as referent objects for smaller targets (e. g., Shanon, 1984). But size is not the only factor. Salience is context-bound, and conceptual issues enter the selection. The referent object must be known to the communication partners. In fact, communication partners quickly discover their common ground with respect to environments, and select referent objects appropriate to their mutual knowledge (e. g.,

Isaacs & Clark, 1987; Shanon, 1983). The multiple considerations for selecting good referent objects are quite similar to those for selecting landmarks (e. g., Couclelis, Golledge, Gale, & Tobler, 1989). Salience, then, has both perceptual and conceptual determinants.

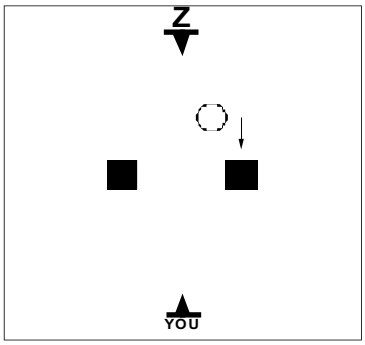
Principle of Ease of Reference Terms. As earlier discuss indicates, computing spatial directions can be difficult. Nevertheless, some direction terms, in particular the asymmetric ones, are easier to compute than others. East-west is more difficult than north-south (Sholl & Egeth, 1985), perhaps because north-south are absolute directions, anchored at the poles, whereas east-west are relative terms. For the direction terms that are projections from the intrinsic sides of objects, including people, right and left take longer to produce and understand and produce more errors than front, back, above or head, and below or feet (e. g., Franklin & Tversky, 1990). Front/back and head/feet project from easily discernible sides of an object. The relative difficulty of left/right compared to the other axes holds for computations from one's own body, from the body of another, from an object, from a description, from a diagram, and from a model (Bryant, Lanca, & Tversky, 1995; Bryant & Tversky, 1995, 1999; Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1992). The relative difficulty of front/back and head/feet depends on the posture of the observer. For an upright observer, head/feet is fastest, seemingly because of the alignment of the body to gravity, the only asymmetric natural axis of the world. For a reclining observer who turns from side to side, gravity no longer correlates with a natural axis of the body. In that case, front/back is faster than head/feet, and, of course, left/right.

Studies of Perspective Selection

In many face-to-face communication situations, the communication partners themselves are salient both perceptually and cognitively. Adapting a task developed by Shober (1993), Mainwaring, Tversky, Ogishi, and Schiano (2000) presented participants with maps of simple spatial scenes consisting of two identical objects arrayed on the projected sagittal or frontal planes of speaker and addressee. Speaker and addressee either faced each other or were offset at 90 degrees. In addition, some arrays had landmarks and/or indications of the cardinal directions. The participants, the speakers, were told that they were secret agents, Agent U, working with their partners, the addressees, Agent Z. The task of the participant-agents varied. In the Tell Other task, they were told that one of the objects was critical, containing a bomb or a hiding a treasure for example, and that they were to signal their partners which object was critical by writing a brief description on their secret communicator. In the Ask Other task, participants were told that their partners knew which object was critical, and that they were to construct a question that could be answered by yes or no that would inform them, the speakers, which object was critical. In the Tell Self task, participants were told that they needed to write a brief message for themselves to be used by themselves at a later time to identify the critical object. On each trial, the communicator saw

a sketch of the positions of Agent U, Agent Z and the objects. See Figure 1 for an example. In the Tell Other and Tell Self cases, the critical object was identified.

A gold bar has been buried in one of the sandboxes at a playground. You've been told which one, but Agent Z has the metal detector and so can best dig up the gold. The squares are sandboxes, the arrow points to the gold, and the circle is a jungle gym.



Agent Z looks at you and signals: Where is the gold?

You signal back to her:

Figure 1: Example of a "Tell Other" scenario from Mainwaring, Tversky, Ohgishi, and Schiano, 2000.

We constructed this situation and these tasks to explore what perspectives speakers would take on their own, their addressees', or perhaps neither, when landmarks or cardinal directions were offered. In fact, the tasks had a large effect on whose perspective was taken. One key factor was the relative cognitive load of speaker and addressee, not politeness. If politeness were the predominant factor in perspective choice, speakers would take addressees' perspectives equally in both the Tell Other and Ask Other tasks, and they didn't. Relative cognitive load does vary between the two tasks. In the Tell Other task, Agent U has the critical information and is the one who produces an utterance; Agent Z has to understand the utterance and use it to determine the critical object. Assuming that comprehending an utterance is more difficult than producing one, and that acquiring information is more difficult than knowing it ahead of time, the addressee's cognitive burden is greater for the Tell Other task. In the Ask Other Task, the Agent Z has the critical information. The speaker has to construct a question and also to understand the answer. Thus, in the Ask Other Task, the cognitive burdens of speaker and addressee are more balanced. Finally, in the Tell Self task, the speaker is the addressee, so there is no one else to accommodate. By this reasoning, speakers should take their own perspective least often in the Tell Other task, more in the Ask Other task, and most in the Tell Self task. The data, from both American and Japanese participants, confirmed this pattern. For the scenarios

that included only the positions of Agent U, Agent Z, and the two objects, speakers took addressees' perspectives more than 90% of the time in the Tell Other task, more than 6% of the time in the Ask Other task, and around 40% of the time in the Tell Self task.

The strong tendency to take addressee, Agent Z's, perspective in the Tell Other task and weaker tendency to do so in the Ask Other task can be accounted for by the relative cognitive burden of speaker and addressee. But why should the speaker take Agent Z's perspective at all in the Tell Self task, where the addressee is the self, Agent U, and not Agent Z? The most compelling explanation is that speakers are using the other person in the scene as a landmark. In fact, the tendency to use Agent Z's perspective was greater when the alternative was to use speaker's right and left than when the alternative was to use speaker's front. This suggests an explanation for using the communication participants as landmarks. It allows using simpler terms of reference. This in turn implies that selection of a referent object and selection of terms of reference are not independent. They cannot be viewed as sequential steps in constructing spatial utterances, but rather as constraints.

The finding that speakers used Agent Z's perspective in constructing messages for themselves calls to question the meaning of using Agent Z's perspective when Agent Z is the addressee. Is it taking someone else's perspective, or using someone else as a landmark? The same question holds for using speaker's perspective. The use of personal perspective in many cases may be using people's positions as landmarks. This does not discount that speakers are taking consideration of the relative cognitive burdens. In those scenarios where the more difficult terms of reference, right and left, cannot be avoided, speakers are more likely to take Agent Z's perspective in the Tell Other task than in the Ask Other task.

These studies revealed a related bias in the terms of reference preferred in constructing spatial descriptions. Participants overwhelming preferred to use "near" than to use "far." The near bias accounts for some of the use of Agent Z's perspective in the Tell Self task. When the critical object was lined up between Agent U and Agent Z but closer to Z, speakers preferred "near Z" to "near me." The near bias is in spite of the fact that, of the pair of polar adjectives, it is "far" that is the neutral, unmarked term and so should be preferred (cf. Bierwisch, 1967; Clark, 1973). The near bias is better explained as a spatial consideration as well as a linguistic one. "Near" defines a relatively small spatial area at any direction from the landmark or target (cf. Morrow & Clark, 1988). "Far," on the other hand, refers to a relatively larger spatial area, potentially anywhere that is not near, in any direction from the target or landmark. "Near," then, is a more specific term, so that when both "near" and "far" are possibilities, by Gricean principles, "near" should be, and in fact, is, preferred to "far."

The earlier work on extended descriptions also supports the idea that the spatial layout and communication task affect perspective globally and locally. Taylor and Tversky (1996) found that environments with multiple routes and landmarks on multiple size scales encouraged relatively more survey perspectives and environments with single routes and

landmarks at a single size scale encouraged relatively more route descriptions. About half the descriptions mixed perspectives, frequently systematically. For example, the part of the Town that had landmarks on several size scales was typically described using a survey perspective, but the town square, with a salient route and same-sized landmarks was typically described using a route perspective. For these cases, switching perspective was not only common, but predictable, based on the configuration of the environment.

Deconstructing Perspective

By now we have a partial answer to the question we posed, why do speakers mix perspectives? First we have shown that there is indeed a cognitive cost to switching perspective, but the cost is small and transient. Switching perspectives can entail changes of target object, terms of reference, and viewpoint. Each of these changes requires different mental transformations, and the relative costs of the transformation are not yet fully known. In addition, however, there are cognitive costs to the mental transformations entailed in understanding successive utterances within a perspective. Within a route perspective, for example, the viewpoint changes, and within a survey perspective, the referent object changes. For both, the target object changes for each utterance. The major difficulty with perspective switching, like the major difficulty in constructing and understanding spatial descriptions, seems to be changes in terms of reference. Some terms of reference, like "left" and "right" or "east" and "west" are particularly difficult. The cognitive costs of switching perspective may be balanced by the cognitive costs of retaining the same perspective. In constructing spatial descriptions for self or other, at a minimum, speakers need to select a referent object and terms of reference. Ceteris paribus, speakers select referent objects that are salient to communication partners and terms of reference that are relatively easy to produce and comprehend. Switching perspective may enable the more salient referent object and easier reference terms. Moreover, the elements of spatial descriptions discussed here, selection of referent object and terms of reference, are not the only considerations relevant to selection of perspective. Putting all this together, switching perspective may be more effective in communication than not switching perspective.

But relative cognitive costs are the only reason for switching perspectives. Perhaps a more compelling reason is that when people perceive and represent environments, they seem to do so from more multiple perspectives simultaneously. This is suggested by evidence from receptive fields for tactile and visual stimuli (e. g., Graziano & Gross, 1994) by perceptual tasks (e. g., Bridgeman, 1999; Robertson & Rafal, in press; Sholl, 1999; Tipper & Behrmann, 1996) and by language tasks (e. g., Carlson-Radvansky, 1999; Carlson-Radvansky & Irwin, 1993; Werner, 1999).

Let us now broaden our perspective from science to art. Consider, for example, perspective in painting. Consistent, one-point perspective is a recent Western invention, an invention that has largely been dropped in 20th century art. Children's art, so-called

primitive art, and high art all over the world, from ancient Egypt to the masters of the Far East to Cezanne, Picasso, and Hockney, freely and beautifully mix perspectives. And none of it is jarring. On the contrary. Literature, too, throughout the ages has mixed perspective, spatial, temporal, personal. Different perspectives provide different views and highlight different aspects of a scene or experience, enriching appreciation of it.

Given the overwhelming evidence for mixed perspectives in perception and behavior, it might be time to turn the original question upsidedown; instead of asking why people mix perspectives, perhaps we should be asking why people use perspective consistently. An even more radical proposal is to deconstruct the notion of reference frame and propose instead that neither perceivers nor speakers adopt reference frames as whole, but rather as fragments.

References

Bierwisch, M. (1967). Some semantic universals of German adjectivals. *Foundations of Language, 3,* 1-36.

Bridgeman, B. (1999). Motor response to cognitively mislocalized targets. Manuscript submitted for publication.

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

Bryant, D. J. & Tversky, B. (1992). Assessing spatial frameworks with object and direction probes. *Bulletin of the Psychonomic Society, 30,* 29-32.

Bryant, D. J. & Tversky, B. (1999). Mental representations of spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 137-156.

Bryant, D. J., B Tversky & Franklin N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language, 31,* 74-98.

Bryant, D. J., Lanca, M., & Tversky, B. (1995). Spatial concepts and perception of physical and diagrammed scenes. *Perceptual and Motor Skills, 81,* 531-546.

Carlson-Radvansky, L. (1999). Selecting a reference frame. Manuscript submitted for publication.

Carlson-Radvansky, L. & Irwin, D. E. (1993). Frames of reference in vision and langugage: Where is above? *Cognition, 46,* 223-244.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* Pp. 27-63. N. Y.: Academic Press

Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology, 7,* 99-122.

Farrell, W. S. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 42-51.

Franklin, N. & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119,* 63-76.

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models acquired from text. *Memory and Cognition, 20,* 507-518.

Graziano, M. S. A. & Gross, C. G. (1994). Mapping space with neurons. *Current Directions in Psychological Science, 3,* 164-167.

Ehrich, V. & Koster, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes, 6,* 169-195.

Isaacs, E. A. & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of General Psychology, 116,* 26-37.

Lee, P. U., & Tversky, B. (In preparation). Switching perspectives has a cognitive cost.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind, & J. J. Koenderink (Eds.), *Limits on perception*. Pp. 323-358. Utrecht, The Netherlands: VNU Science Press.

Levinson, S. (1996). Frames of reference and Molyneux's question: Cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, and M. F. Garrett, *Space and Language*. Pp. 109-169. Cambridge: MIT Press.

Mainwaring, S. D., Tversky, B., Ohgishi, M., & Schiano.D. J. (2000). Descriptions of simple spatial scenes in English and Japanese. Manuscript submitted for publication.

Maki, R. H. & Braine, L. G. (1985). The role of verbal labels in the judgment of orientation and location. *Perception, 14*, 67-80.

Maki, R. H., Grandy, C. A. & Hauge, G. (1979). Why is telling right from left more difficult than telling above from below? *Journal of Experimental Psychology: Human Perception and Performance, 5,* 52-67

Morrow, D. G. & Clark, H. H. (1988). Interpreting words in spatial descriptions. *Language and Cognitive Processes, 3,* 275-291. Shanon, B. (1983). Answers to where- questions. *Discourse Processes, 6,* 319-352.

Perrig, W. & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24,* 503-518.

Shanon, B. (1983). Answers to where-questions. *Discourse Processes, 6,* 319-352.

Shanon, B. (1984). Room descriptions. *Discourse Processes, 7,* 225-255.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition, 47,* 1-24.

Sholl, J. (1999) Types of self-reference systems used for retrieving knowledge of large-scale spatial relations. Manuscript submitted for publication.

Sholl, J. & Egeth, H. E. (1981). Right-left confusion in the adult: A verbal labeling effect. *Memory and Cognition, 9,* 339-350.

Taylor, H. A. & Tversky, B. (1992a). Descriptions and depictions of environments. *Memory and Cognition, 20,* 483-496.

Taylor, H. A. & Tversky, B. (1992b). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language, 31,* 261-292.

Taylor, H. A. & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory & Language, 35,* 371-391.

Tipper, S. P. & Behrmann, M. (1996) Object-centered not scene-based visual neglect. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1261-1278.

Robertson, L. C. & Rafal, R. (In press). Disorders of visual attention. In M. Gazzaniga (Editor). *The cognitive neurosciences*. Second Edition. Cambridge, MA: MIT Press.

Werner, S. (1999). Reference systems for spatial knowledge at environmental scale. Manuscript submitted for publication.