Chapter 15
Space and Language

Mary A. Peterson, Lynn Nadel, Paul Bloom, and Merrill F. Garrett

15.1 Introduction

Functioning effectively in space is essential to survival, and sophisticated spatial
cognitive systems are evident in a wide range of species. In humans, the emergence of
language adds another level of complexity to the organization of spatial cognition.
We use language for many purposes, not least of which is the conveying of informa-
tion about where important things are located (food, safety, enemies) and how to get
to and from these places (for discussion of these evolutionary issues, see O'Keefe and
Nadel 1978, Pinker and Bloom 1990). Given the fundamental nature and importance
of spatial cognition, it is of considerable interest to determine the ways in which it
connects to language. The hope that study of such connections might shed light on
both the spatial cognitive faculty and on the language faculty has generated consider-
able interest in the domain of “language and space.”

We are interested in how people talk about space and what they can and do choose
to say about it. By exploring the boundaries of these cognitive domains, we hope to
uncover their structure and to elucidate the ways in which they can relate to one
another. By considering the role of development and culture in shaping the language-
space interaction, we hope to discover the extent to which fundamental aspects of
spatial cognition are given a priori and the extent to which spatial cognition can be
altered by experience. And by analyzing the ways in which neural systems organize
spatial and linguistic knowledge, we hope to shed light on how these two capacities
relate to one another.

In the present chapter we analyze what we take to be the consensually accepted
framework within which the relations between language and space have been consid-
ered. Based on this framework, we critically discuss some influential proposals as to
the precise nature of this relationship. Finally, we return to the set of issues and
questions with which we began, and reach some tentative conclusions.
### 15.2 Framework

The framework we adopt here for how we talk about space is based on the proposal by Jackendoff (1983, 1987), who took Fodor’s (1975) “language of thought” hypothesis as a starting assumption. Fodor argued convincingly that one cannot learn a language unless one already has an original language to structure the learning process; he referred to this original language as the “language of thought.” The language of thought includes the building blocks from which our concepts are constructed.

Extending Fodor’s analysis, Jackendoff has argued for something along the lines of the situation represented in figure 15.1. There exist language representations (LRs), spatial representations (SRs), and conceptual representations (CRs). LRs include all aspects of language structure, including the lexicon and the grammar; SRs include all aspects of spatial structure as it is represented in the brain; and CRs are primitives that form the components of meaning, both linguistic (CR₄) and spatial (CR₅) meaning.

![Diagram showing the relationship between language and space](image)

**Figure 15.1**
A schematic depiction of Jackendoff’s analysis of the relationship between language representations (Language Rs), spatial representations (Spatial Rs), and conceptual representations (CRs) of both language and space.

### 15.3 Space and Spatial Information

What do we mean by “space”? It is clear that space can contain objects and events, but it need not. Empty space, and unoccupied places, exist. O’Keefe and Nadel (1978, 86) note that “the notions of place and space are logical and conceptual primitives which cannot be reduced to, or defined in terms of, other entities. Specifically, place and spaces are not, in our view, defined in terms of objects or the relations between objects.” On the other hand, objects and events both occupy spatial locations and have intrinsic spatial properties. Objects are partly defined by the spatial relations among their parts, and events are partly defined by the spatial relations among the various entities (e.g., objects, people, actions) that compose them. Are the logical distinctions among objects, events, and spaces maintained in language, and are they paralleled by dissociable neural representations? In order to determine which aspects of spatial representation are transparent to language, it would be useful to know something about the various ways in which space is represented in the mind/brain. Is there a single multimodal or amodal representation of space per se? Or are there a number of independent modules for spatial representation? Can the study of these representations help us understand space, language, and the interface between the two, and can it shed light on the nature of the “spatial” primitives in CR?
15.3.1 Evidence from Studies of the Brain

Available neurobiological evidence suggests that a relatively large number of distinct representations or “maps” of space and spatial information exist. For example, various investigators have discussed maps of motor space, auditory space, visual space, and haptic space; maps of body space, near space, far space; maps of egocentric space and allocentric space; and, maps of categorical space and coordinate space. Quite separate representations for the spatial features of objects and their locations have also been investigated. Neuroscientists have linked a variety of brain structures and systems to one or another of these spatial representations, providing converging evidence for some degree of independence of many of these forms of spatial information.

Neurobiological evidence also suggests the existence of multimodal spatial representations. It is known, for example, that the mammalian superior colliculus integrates sensory spatial maps and motor maps, such that nearby neurons are activated either by sensory inputs from, or motor outputs directed to, a particular part of egocentric space (cf. Gaither and Stein 1979; Meredith and Stein 1982). Indeed, this integrative system is quite primitive phylogenetically; comparable integration of sensory spatial maps occurs in snakes, where visual and thermal “spatial” maps are brought into register in the superior colliculus (Newman and Hartline 1981).

There is also evidence demonstrating the presence of multimodal sensory spatial maps in various areas of the mammalian neocortex, including especially parts of the parietal cortex (Pohl 1973; Kolb et al. 1994). All of these multimodal maps seem to share the critical feature of representing space egocentrically, that is, with reference to the organism, or some part of the organism (e.g., hand, eye, head, torso). The “maps” in these multimodal brain association areas, as well as those in unimodal regions such as the visual and somatosensory cortex, are all laid out in topographic fashion, such that neighboring regions of neural space represent neighboring regions of the ego-centered world.

In addition to these data demonstrating the existence of various unimodal and multimodal ego-centered spatial maps, there is evidence for a superordinate allocentric amodal or multimodal spatial representation that subsumes or somehow integrates the spatial representations (SRs) provided by each of the various spatial maps. It is now well established that the vertebrate hippocampus subserves a spatial mapping function that is both multimodal and allocentric; that is, external space is represented independent of the momentary position of the organism, in terms of the relations between objects and the places they occupy in what appears to be an objective, absolute framework (O’Keefe and Nadel 1978; and see O’Keefe, chapter 7, this volume). This system contains information about place, distance, and direction. The “place” cells first identified by O’Keefe (e.g., O’Keefe and Dostrovsky 1971; O’Keefe 1976; O’Keefe and Nadel 1978) are active when an animal is in a given location in space, as defined by the relationship between that spatial location and other places in the environment. O’Keefe and Nadel (1978) have postulated that information about distance is provided to the hippocampus via the septal region, and a pattern of activation termed theta, driven by inputs from brain stem movement systems. More recently, Taube and others (Taube 1992; Taube, Muller, and Ranck 1990a,b) have described a population of “head direction” neurons in the dorsal subiculum and thalamus of the rat; these cells are active when an animal faces in a particular direction in the environment, wherever its specific location. Existing data show that the place cells and the head direction cells are tightly linked together (Knierim et al. 1993). The representation of allocentric space created in the hippocampus uses multimodal information from cortical systems including the parietal and temporal regions, as well as inputs conveying egocentric information about directions and distances.

Interestingly, this allocentric spatial representation is not neurally topographic in the way the egocentric representations are. As far as we can tell, neighboring regions of hippocampal neural space do not represent neighboring regions of the external world. While this hippocampal system has been postulated to provide the basis for certain spatial primitives (e.g., places), it does not appear to be necessary for a wide range of nonallocentric spatial representations, such as those subserved by the superior colliculus and neocortical regions already noted.

The spatial maps in the superior colliculus, parietal cortex, and hippocampus are thought to represent space without regard for the exact nature of the objects occupying any part of the represented space. The spatial (and other) properties of objects appear to be captured in separate neural systems. Thus considerable neurobiological evidence suggests the existence of two streams of visual processing about objects: a ventral pathway, incorporating regions of the temporal lobe that is concerned with what an object is, and a dorsal pathway, incorporating regions of the parietal cortex that is concerned with where an object is located with respect to the organism (e.g., Ungerleider and Mishkin 1982). Neurpsychological investigations of brain-damaged individuals have been taken as support for this “what” versus “where” distinction in the object representation system, as in the well-known cases of “blindsight,” where subjects express a lack of awareness for the presence and nature of an object, all the while demonstrating by their behavior that they know “where” the object is located (e.g., Wexler 1986). However, recent evidence indicates that the ventral and dorsal processing streams are not nearly as isolated as originally conjectured (Van Essen, Anderson, and Fellman 1992). Nor is the neurpsychological evidence for independent streams completely convincing; some neuropsychologists now argue for notions like degrees of modularity (for review, see Shallice 1988). Whatever the status of these visual processing streams, they both provide inputs to
the hippocampal system, presumably contributing to its ability to construct the allocentric representation discussed above.

The evidence from studies of the brain thus suggests that (1) there are a variety of spatial maps in the brain, which makes it unlikely that there is a single amodal spatial representation that gives rise to the entire set of spatial primitives; (2) at least some neural representations of space do not include detailed representations of objects, reflecting the logical distinction between environmental space and the spatial aspects of objects discussed above; and (3) there is some, but not total, separation within the systems representing objects between those representing what an object is, and those representing where it is located.

15.3.2 Evidence from Studies of Perception
Behavioral evidence is consistent with the idea of a variety of spatial maps. For example, consider the elegant study conducted by Loomis et al. (1992), who showed observers two targets located at different distances from the observer in an open field. The observers performed two tasks with respect to these targets. In one task, observers used a matching response to report about the perceived distance between the two targets at different distances (i.e., they adjusted the apparent horizontal distance between two objects located at a standard distance to match the apparent depth interval between the two test targets). In the second task, observers viewed the display from the same stationary vantage point used for the first task and, closing their eyes, walked first to one distal target and then to the other (the targets were removed once the observers closed their eyes and began walking). Thus the distance between two distal objects that had been visually apprehended from the same vantage point as used in the first task was motorically expressed in the second task, once the observers had walked to the location of the first object.

Loomis et al. (1992) found a dissociation between these two different estimates of the distance between the two objects, with the walking responses suggesting a more veridical perception of distance than the matching responses. The latter reflected the operation of organizing factors; the error in the perceived distance between the two distal objects appeared to increase systematically as a function of the distance from the observer’s vantage point to the objects (this effect may be an instance of Gogel’s 1977 equidistance tendency). On the other hand, no such increase as a function of distance from the stationary vantage point was observed when the interobject distance was assessed via walking responses. The experiment by Loomis et al. demonstrates that although certain distance representations are veridical, in that they can support accurate navigating to the two targets in turn, other representations of the distance between the two objects are systematically distorted by the operation of perceptual organizing factors.

Figure 15.2
An illustration of the Duncker induced-motion display in which a small dot is enclosed in a large rectangle. A stationary display is shown in (A). A moving display is shown in (B), where the actual motion of the rectangle is indicated by a solid arrow pointing to the right, and the perceived motion of the dot is indicated by a dashed arrow pointing to the left.

Similarly, Bridgeman, Kirch, and Sperling (1981) demonstrated a striking dissociation between the location of a small target relative to an enclosing frame and its location relative to the observer. These investigators showed their observers a Duncker display, containing a small target enclosed within a larger rectangular frame, like that shown in figure 15.2. In displays such as this one, an “induced motion” illusion occurs when the frame is displaced abruptly in one direction, say, to the right. In a number of situations, abrupt displacements have been shown to mimic motion signals to the visual system, and consequently, to result in apparent motion. The unique characteristic of the Duncker induced-motion illusion is that observers perceive motion (or displacement) of the stationary dot, rather than of the displaced frame. For example, in the display shown in figure 15.2, rather than perceiving the displaced rectangular frame as moving to the right, observers perceive the stationary target as moving to the left, that is, in the direction opposite to the direction in which the frame was displaced.

The induced illusory motion of the small target is very compelling visually, and it can be canceled (i.e., the small target can be made to appear stationary when the frame is displaced to the right) by the addition of a real displacement of the small target in the same direction as the frame displacement (in this example, to the right). Thus, as with many other kinds of illusory motion (e.g., Gogel 1982; Peterson and Shyi 1988), induced motion and real motion add perceptually and may be indistinguishable.
Bridgeman, Kirch, and Sperling gathered two kinds of responses about the location of the small target in the Duncker display. One response was a cancellation response, as described. By this measure, and by self-report about what they saw, the observers in their experiment indicated perceiving the (actually stationary) target as having moved from its original location. The magnitude of the change in location inferred from the cancellation responses was about half the distance through which the frame had been displaced. On another block of trials, Bridgeman, Kirch, and Sperling asked the same observers to point, with an unseen hand, to the final perceived location of the target after viewing the induced-motion display, which disappeared from sight before they made the pointing response. Surprisingly, at least with respect to the cancellation responses given by the same subjects, the magnitude of the illusion measured by the pointing response was negligible. Under these conditions, observers pointed much closer to the actual location of the target than to the perceived location of the target (as inferred from the cancellation responses).

Thus the experiment revealed a distinction between the spatial representations mediating the cancellation response (and presumably, visually perceived location) and the spatial representations mediating the pointing response (and perhaps, motoric responses in general). It is clear that the visually perceived location reflects visual organizing factors—in this case, an organization that depends on the enclosing relationship between the frame and the target—whereas the representation of location accessed by the motor response seems relatively free of such effects. (We will return to this point below.) Although it may not be clear how best to characterize this distinction (see Loomis et al. 1992 for a lucid discussion), the results of Bridgeman, Kirch, and Sperling and those of Loomis et al. strongly imply that the maps mediating visually perceived spatial relationships differ from those mediating movement-expressed distances, locations, and/or directions.

In addition to the behavioral evidence for differential encoding of egocentric versus allocentric spaces, and for locations versus directions, there is evidence that spatial experience can reflect the combination of inputs from different modalities. For example, Lackner and his colleagues have shown that auditory, visual, and kinesthetic inputs regarding an observer's orientation in space are combined to yield a perceived space that does not correspond to the space signaled by any one input (for review, see Lackner 1985). The behavioral evidence is therefore consistent with the idea of multimodal spatial representations, as well as with the idea of various unimodal spatial maps.

What do perceptual studies indicate about the independence of object and spatial representations? To a certain extent, the independence of these two systems has simply been assumed (see Marr 1982 and Wallach 1949 for explicit statements of this assumption). Consistent with this assumption, object recognition does appear to exhibit location invariance (see Biederman and Cooper 1991 for recent evidence), and accurate distance perception is clearly possible for novel objects.

On the other hand, behavioral evidence has occasionally suggested that these two systems may influence one another. For example, Carlson and Tassone (1971) found that the perceived egocentric distance to objects in naturalistic settings is influenced by the familiarity of the objects. The initial experiments could not rule out a number of alternative explanations based on response tendencies or differences in the complexity of familiar and unfamiliar objects, but subsequent work excluded these possibilities (Predebon 1990). Similarly, object recognition may not be completely independent of location: recognition accuracy for individual objects located within contextually appropriate scenes is reduced when the objects are presented in inappropriate locations (e.g., a fire hydrant in a street scene is less likely to be recognized when it is located inappropriately on top of a mailbox than when it is located appropriately at street level; Biederman 1981).

Recent evidence indicates that figure-ground organization, which entails the perception of the relative distance between two adjacent regions in the visual field, is influenced by the familiarity (or recognizability) of the regions (Peterson 1994; Peterson and Gibson 1993, 1994). These findings have led Peterson and her colleagues to propose that a rapid object recognition process (a "what" process) operates before the determination of depth segregation (a classic "where" process) and that the former exerts an influence on the latter in combination with more traditional depth cues, such as binocular disparity (see Peterson 1994; Peterson and Gibson 1994). Similarly, Shiffrar and Freyd (1998; Shiffrar 1994) have shown that perceived direction of motion through space ("where") is constrained by the types of movements that are possible given the nature of the objects in motion ("what").

In addition to these effects of object identity on perceived spatial organization, there is evidence that another type of spatial information is fundamental for object identity. Object identification fails when the parts of the object are spatially rearranged (see, for example, Biederman 1987; Hummel and Biederman 1992; and Peterson, Harvey, and Weidenbacher 1991), and is delayed when a picture depicts an object misoriented with respect to its typical upright orientation (Gibson and Peterson 1994; Jolicoeur 1988; Tarr and Pinker 1989).

In sum, and consistent with conclusions drawn from neurobiological analysis, the study of perception shows that (1) a variety of independent modules for spatial representation exist; and (2) some representations deal with objects, some with spaces, and some with the interaction between the two. What can be said about how, if at all, language hooks up with each of these proposed modules?
Talking about Space and Spatial Relations

Taking the existence of independent spatial modules—some dealing with spaces, some dealing with objects, and some dealing with the interactions between the two—as a starting assumption, one can pose the following question. Does language express the information available in all, or only some, of these spatial modules? And are particular parts of language used to express specific forms of spatial information?

Landau and Jackendoff (1993; Jackendoff, chapter 1, this volume; Landau, chapter 8, this volume) have recently taken a modular position on the question of how language relates to representations of objects versus spatial relations between objects. Their position rests on a linguistic analysis that emphasizes the differences between the manner in which languages code for spatial relations and for objects. For example, in English, objects are described by nouns, which are open-class linguistic elements, whereas spatial relationships are described by prepositions, which are closed-class linguistic elements. The prepositions of English—and members of the corresponding grammatical class in other languages—may be specialized for speaking about space, as opposed to speaking about objects, in that they may be applied with few constraints to objects of different sizes and different sorts (e.g., see Talmy 1983). In addition, spatial relationships between objects tend to be coarsely coded by languages, at least in comparison to object identities. That is, spatial prepositions such as near and far would not support accurate motor behavior of the type studied by Bridgeman, Kirch, and Sperling (1981) and by Loomis et al. (1992). Landau and Jackendoff (1993) also stress the fact that the number of spatial prepositions in English (around 75) is quite small relative to the number of object names (30,000 or so, according to a count by Biederman 1987).

Landau and Jackendoff (1993) took these differences between prepositions and nouns as evidence that prepositions and nouns mapped onto different sorts of spatial relationships. In particular, their proposal suggests that closed-class linguistic spatial terms might map to a subset of CRs that are about the spatial relations between objects, or between an observer and objects (i.e., relations that represent the locations of objects without regard for the specific properties of the objects occupying those locations), whereas nouns might enjoy a privileged mapping to a subset of CRs that are specialized for object representation (e.g., the 3-D object models of Marr 1982 or Biederman 1987). Landau and Jackendoff pointed out that the linguistic distinctions in the meanings of nouns and prepositions fit nicely with neurobiological and computational evidence indicating that “what” and “where” are represented independently (e.g., Ungerleider and Mishkin 1982; Rueckl, Cave, and Kosslyn 1989).

By incorporating modern research and theory about “what” and “where” systems, Landau and Jackendoff’s (1993) approach usefully builds on Jackendoff’s (1983) insight that we can learn about spatial conceptual representations by studying how we talk about space. However, in our view, research programs that attempt simply to identify subdivisions of language with neural spatial systems will not be able to fully elucidate the nature of spatial conceptual representations (see also Bierwisch, Chapter 2, this volume). This follows from the fact that words express abstract conceptual notions that do not appear to be captured in any one-to-one fashion by sensory, perceptual, or neural representations. The literature on word learning illustrates this point.

In much of the discussion of the relationship between language and space, particularly by developmental psychologists, it is assumed that nouns are equivalent to object names and that objects are equivalent to entities generalized on the basis of shape. But neither of these equivalences exists. For adults, only a minority of nouns refer to material objects; most nouns are like day, family, joke, factors, information, and so on. Nouns that do not refer to objects are also used, with appropriate syntax and meaning, by two- and three-year-old children, and there is considerable experimental evidence showing that children have no special problems learning such words (see Bloom 1994, in press). Even infants appear to possess CRs that do not correspond to objects: infants six months of age are capable of mentally representing, and counting, discrete sounds (Starkey, Spelke, and Gelman 1990) and individual actions, such as the jump of a puppet (Wynn 1995).

What about the claim that “just for nouns that do name objects,” a property such as shape, which can be derived from sensory/perceptual inputs, is criterial? Even this is too strong. Children learn superordinates, like animal and furniture, that refer to categories that share no common shape, relationship terms like doctor and sister, and functional terms like clock and switch (see Soja, Carey, and Spelke 1992 for further discussion). In fact, even for those objects that do have characteristic shapes, children know that shape is not criterial. If you alter a porcupine so that it has the shape of a cactus, three- and four-year-old children insist it remains a porcupine; they view ontological boundaries (e.g., you cannot transform an animal into a plant) as more significant than shape (Keil 1989).

Rather than assuming that nouns map directly onto a “what” system that encodes objects in terms of shape, an alternative outlined in Bloom (1994, in press) assumes that nouns map onto CRs that are nonspatial (and thus can include notions like joke and day). This is not to deny that shape is important for learning object names, as demonstrated by Landau and her colleagues (see Landau, chapter 8, this volume). In some cases, considerations of shape are relevant for determining category membership, implying an interface between CRs and the shape of an object. For instance, there is evidence that children and adults have an essentialist notion of natural kind concepts, so that the CR for porcupine is, roughly, “everything that has the same internal ‘stuff’ as previously encountered porcupines” (e.g., Putnam 1975; Keil 1989).
But since internal stuff—essence—is unobservable, we normally use an observable property that is highly correlated with essence—shape—to determine whether something is a porcupine. Shape also correlates with membership in certain functional kinds; given the purpose for which chairs are designed, they are likely to have a certain configuration (i.e., they are likely to have shapes that afford sitting).

As noted above, however, there are many words and semantic categories for which no such correlation with shape exists. Jokes and days have no shapes, doctors and fundamentalists do have typical shapes but not ones that distinguish them from lawyers and agnostics, and although categories like animals and furniture refer to entities with shapes, the entities forming the category all have different shapes. These considerations show that although there is a relationship between the category of nouns and the notion of object shape, it is not direct. Rather, it is mediated through a more abstract conceptual system of CRs. As a result, the link between language and the shapes of objects is, at least for open-class categories such as nouns, nowhere near as direct as many researchers assume it to be.

Similarly, it is clear that spatial terms cannot be derived simply from an interface between language and a set of sensory/perceptual maps. Consider what is actually conveyed by the spatial representation used to describe the relationship between the butterfly and the jar in figure 15.3A, captured by the following sentence:

The butterfly is in the jar.

In this description, the relationship described by the spatial preposition in cannot be reduced to one of mere surroundedness in the visual display. The butterfly in figure 15.3B is not in the tabletop, although the contours of the tabletop surround it, as the contours of the jar surround the butterfly in figure 15.3A; whereas the butterfly in figure 15.3C is correctly described as in the canyon, although the contours of the canyon do not surround the butterfly. Clearly, the meaning of the spatial preposition in cannot be defined by appealing to attributes of a sensory spatial representation only. Instead, one must appeal to some abstract relationship, such as a capacity for containment that jars and canyons share, but tabletops do not. The abstract notion of containment may be one of the conceptual representations linking space and language, and it may be by virtue of this CR that canyons and jars can be categorized similarly by language, but the notion of containment simply cannot be accounted for by complexes of sensor information (for discussion, see Bowerman, chapter 10, this volume, and Mandler, chapter 9, this volume).

Even though CRs will not map to complexes of sensory information, can similarities in the way linguistic terms and spatial maps are characterized help us identify an underlying isomorphism between a linguistic category such as prepositions and a particular spatial representation? For example, might the "categorizing" role played by the term in in the situations illustrated in figures 15.3A and 15.3C imply that spatial terms have a privileged mapping to "categorical" as opposed to "coordinate" spatial representations within the "where" system, such as those postulated by Kosslyn and his colleagues (Kosslyn et al. 1989; Kosslyn et al. 1992)? Such a solution would appeal to the surface similarity implied by the use of the term categorical in these two cases, but the similarity may not go beyond the surface. Categorical maps treat a set of spatial locations as equivalent; that is, these maps have low resolution. Yet terms such as in could as easily reflect high-resolution as low-resolution spatial representations.

Those situations in which spatial prepositional usage can depend upon the entities being related may be more revealing about the nature of conceptual representations than those situations that fit within a "what" versus "where" dichotomy. There are a number of examples in addition to the one illustrated in figure 15.3, showing that
nonspatial factors govern the use of spatial terms in English. For example, the considerable variability in applying the terms *front* and *back* to churches implicates functional factors as relevant to axial descriptions, given that the front of a church can be defined by functional factors (e.g., the direction people attending a service face) as well as by structural factors coded in an object representation. (See Vandeloise 1991 for a discussion of functional meanings of French prepositions). Current neuropsychological evidence, supported by computational evidence, indicates that functional representations may be critically important attributes of object meaning (see Shallice, chapter 14, this volume). Thus the nonspatial semantics of an object may govern prepositional usage. Other nonspatial semantic factors such as salience are relevant when we say whether something is “near” or “far”; these semantic factors are evident in some spatial memories as well (for review, see McNamara 1991).

A last example that language does not divide up in ways that map directly onto particular neural spatial representations can be found in linguistic directional terms. Directions are coded and lexicalized by various languages in deictic (egocentric) terms; in intrinsic (object-centered) terms; or in absolute (cardinal) terms. Does this variability imply that language interfaces directly with spatial representations of all these types, at least with regard to direction? Would this count as a case of direct linkage between particular linguistic elements and specific spatial representations?

Consider the fact that the terms *right*/*left* can be used to denote both the speaker’s *right*/*left* (e.g., egocentric use), or the right and left of some other object or person. Yet spaces are apprehended egocentrically by an exploring animal, either by the act of moving through them or by the behavior of visually scanning them. In either case, the inputs are initially coded in terms of egocentric relations between the observer and entities such as places and objects in the environment. Spatial relations of distance and direction, for instance, and even of the arrangement of parts within an entity, are “computed” by the organism from its various egocentric inputs. Human factors work has demonstrated the primacy or importance of egocentric coding as well. For example, if one has to discriminate which of two adjacent objects on a display screen is brighter (larger, more familiar), the best response mapping is one in which a choice of the object on the right is indicated by a right-hand key press, and a choice of the object on the left by a left-hand key press. There is much evidence from developmental work showing that egocentric knowledge about spatial location precedes allocentric knowledge (e.g., Mangan and Nadel 1992; Wilcox, Rosser, and Nadel 1994). In addition, the neuropsychological deficit of neglect points to the primacy of egocentric *right*/*left* coding. Individuals who sustain damage to the right parietal lobe often “neglect” left hemispace. For example, when there are objects in their right hemispace, neglect patients ignore objects in their left hemispace (Heilman 1979; Volpe, LeDoux, and Gazzaniga 1979). Likewise, when neglect patients are asked to imagine a scene, well known to them before their brain damage, they are unable to imagine those objects that lie to their left from their imagined vantage point, and the objects omitted change as the imagined vantage point changes (Bisiach and Luzzatti 1978). Thus right/left egocentric relationships appear to be coded early in processing and appear to be critically important in spatial understanding.

Notwithstanding the evidence attesting to the importance of egocentrically based spatial information, many normal individuals have severe difficulties in mapping linguistic terms onto egocentric relationships. Part of the problem in using the terms *right*/*left* may arise because egocentric spaces depend on the direction the individual is facing; that is, the regions of space that lie to the right and left are interchanged by a 180° rotation. For a speaker and a listener who face each other, the space to the right of the speaker lies to the left of the listener. Even for speakers who evidence no overt difficulty in using the terms *right* and *left*, considerable effort is required to translate the frame of the speaker into that of the listener (see Tversky, chapter 12, this volume).

Those same individuals who have difficulties mapping linguistic terms onto egocentric relationships have no trouble in reaching to the right or left to catch an object falling off a table. This dissociation suggests that right/left terms do not map directly to motoric egocentric neural representations. Consistent with the possibility that the perspective taking evident in language use does not simply reflect the use of a perspectivefixed spatial map, Levelt (chapter 3, this volume) demonstrates that the spatial representations accessed for speaking about arrays of dots affording either an egocentric or an intrinsic description are not already coded for the egocentric or intrinsic directions speakers choose to express. In our view, the underlying spatial representations may instead be allocentric spatial representations such as those found in the hippocampus (see O’Keefe and Nadel 1978).

The discussion of linguistic evidence has thus far focused on objects and spatial relations, including distances and directions. We have yet to touch upon a critical aspect of space and that is place. How does language treat places? In contrast to the spatial relations of distance and direction, places are described by open-class elements rather than by closed-class elements. Some place names are count nouns, like *center*, *basement*, and *border*, and these can be extended to novel instances in much the same way as names for kinds of entities like ball, house, and country. Others are proper names, like Paris, Times Square, or the Equator, which behave much like the individual names Bill, Joan, and the Salvation Army, and are certainly as informative. We know of no count of the number of places that can be named in English to stack up against the number of spatial prepositions counted by Landau and Jackendoff (1993; n = 75) or against the number of object names estimated by Biederman (1987;
n = 30,000), but certainly the number of places that can be named is several orders of magnitude larger than 75.

This fact suggests to us a rather different way of imagining the relation between aspects of language such as open- and closed-class elements and aspects of spatial (and object) representations. Systems concerned with both space per se and with objects contain information about entities and about relations between entities. The linguistic evidence could be taken to suggest that nouns are used in the case of entities (be they places, objects, or other things—see below), and that prepositions are used in the case of relations (be they about places, objects, or other entities). In this view, prepositions are not limited to describing the spatial relationships between objects. Even putting aside more abstract usages of prepositions (as in "John went from rich to poor"; see, for example, Jackendoff 1990), sentences such as the following are perfectly acceptable:

The mist hovered over the sea.
John put the poison into the soup.
A swarm of bees flew into the forest.
There was an explosion next to my house.
Boston is near New York.
He swept the space in front of the fireplace.

In these examples, there is no problem whatsoever using prepositions to describe spatial relationships between and among substances, collections of objects, events, locations, and even empty space itself.

In sum, studies of language provide no reason to go beyond the basic framework spelled out at the outset, in particular, to propose privileged one-to-one mappings between parts of language and particular spatial representation systems. While it may be the case that the information carried by some spatial maps can be talked about, and the information carried by others cannot, we believe this is not a result of connections between specific types of language elements and specific spatial maps. This does not mean that we reject the idea that neurobiological and perceptual/cognitive research can shed light on the nature of the spatial conceptual representations. Rather, we suspect that investigations of the ways in which CRs interact with the maps identified by neurobiological and behavioral research will be fruitful in elucidating the nature of spatial conceptual representations. This leads us to the following modest proposal.

Some, but not all, of the spatial maps identified by neurobiological and behavioral research impose a structure that goes beyond, and in consequence alters, our interpre-

tation of the information available in the input alone. For example, the hippocampus appears to impose a Euclidean framework onto non-Euclidean inputs (O'Keefe and Nadel 1978, who see in this process the instantiation of a Kantian a priori notion of absolute space). Other examples are revealed by the organizing factors that structure some behavioral representations—factors like the equidistance tendency (Gogol, 1977; Gogol and Teitz 1977) and the constraints due to gravity identified by Shepard (for summary, see Shepard 1994). We propose that in "distorting" the sensory inputs, these spatial maps may impose an order and a structure that our spatial conceptual representations require. If this were the case, studies of language use and other spatial behaviors that revealed the operation of these organizing factors might lead to some understanding of the CRs themselves.

Before engaging in this sort of analysis it is important to first look at the ways in which the mappings between language, behavior and space vary across cultures. If there are internally imposed structures that reflect primitive spatial CRs, one would expect to see these structures preserved across cultural and linguistic boundaries. This follows from the assumption, central to our guiding framework, that the CRs are part of a universal "language of thought" that makes understanding of the world possible. If, on the other hand, spatial frameworks and perception itself can be shown to vary across culture, their utility as stable indicators of the nature of spatial CRs is questionable.

15.4 Effects of Experience

Speakers of Tzeltal code spatial relations with respect to absolute directions; they simply do not use egocentric terms to speak about space or objects (see Levinson, chapter 4, this volume). In this respect they differ from speakers of Dutch and English, for example. The critical feature of absolute directions is that they remain invariant as vantage point changes. In Tzeltal, the absolute directions that are used originate in a feature of the environment—uphill/downhill—are applied even when the geographical feature is out of sight. A tremendous amount of effort is required to keep track of the absolute directions; nevertheless, these directions seem to be well preserved in the memories of events and scenes experienced by the speakers of Tzeltal (Brown and Levinson 1993; Levinson, chapter 4, this volume). This certainly raises the question that led us to consider the effects of experience in the first place. Are there differences in the CRs between speakers of Tzeltal and speakers of English, Dutch, or other languages that lexicalize egocentric relations rather than cardinal directions?

This possibility is difficult to address, but Levinson and his collaborators have shown that speakers of Tzeltal and speakers of Dutch behave differently in old/new
perceptual recognition tasks, problem-solving tasks, and memory tasks (Levinson, chapter 4, this volume). Furthermore, gestures employed by speakers recounting remembered scenes and events are different. The gestures employed by speakers of Tzeltal indicate absolute directions, and the gestures employed by speakers of Dutch and English indicate relative directions (Haviland 1993). Does this mean that the language one speaks, or the culture in which one lives, can change the nature of the underlying CRs? Or does it support the less radical claim that the culture in which one lives, and the language one speaks, affects the availability of different CRs because of differential degrees of practice utilizing them? And in either case, do such findings imply that the conceptual representations at the interface between language and thought are themselves different?

It seems clear that different languages and/or cultures can utilize different cognitive skills to different degrees. For example, Emmorey and her colleagues (see Emmorey, chapter 5, this volume) have shown that sign language may engage mental rotation skills, and consequently, may improve these skills due to practice. Of course, differential prowess at mental rotation does not imply that the CRs are different. Nor is this implied by differences in performance on memory tasks and problem-solving tasks, such as those discussed by Levinson (chapter 4, this volume), although the existence of such differences provides evidence relevant to theories of perception, memory, and problem solving. For example, Levinson’s finding (chapter 4, this volume) that differential encoding of absolute versus egocentric directions by speakers of Tzeltal and of Dutch, respectively, is evident in problem-solving tasks is consistent with psychological evidence that world knowledge (which differs from speaker to speaker) influences problem-solving behavior (Murphy and Medin 1985).

What about Levinson’s claim that differential encoding of absolute directions versus egocentric directions is also evident in performance on old/new perceptual recognition memory tests? If the perceptual representations accessed, for example, by speakers of two languages were different, by virtue of the differential attention each had paid to particular aspects of the situation at encoding, that would be consistent with recent evidence that knowledge influences perception more than traditionally assumed (for summaries, see Peterson 1994; and Shiffrar 1994). Note that the effects of knowledge on perceptual organization may be highly constrained in that the relevant structural, semantic, or functional representations mediating such knowledge must be accessed within the normal time course of perceptual processing (see Carpenter and Grossberg 1987; Gibson and Peterson 1994; and Peterson, Harvey, and Weidenbacher 1991). Thus, if Levinson’s findings are shown to reflect differences in perceptual representations per se, they might justify a search for correlates of absolute directions in perceptual input.

Alternatively, the tasks employed by Levinson may reflect differences in semantic representations between speakers of different languages and/or members of different cultures. It has always been supposed that different languages or different cultures might combine primitive CRs differently so that certain meanings are more or less salient to speakers of a given language (Bowerman 1989; Slobin 1995). Neuro-psychological findings discussed by Shallice (chapter 14, this volume), suggest that qualitatively different semantic representations may be accessed in the course of identifying artifacts and living things (see also Farah and McClelland 1991). Similarly, semantic representations of remembered scenes and events could vary in their emphasis on absolute or egocentric directions, depending upon one’s culture and experience. Thus, while Levinson’s results might mean that semantic representations are different for speakers of different languages (see also Bowerman, chapter 10, this volume), they would not entail that the primitive CRs themselves, from which the semantic representations are constructed, have been changed by language (or culture).

It is important not to overemphasize the differences between speakers of different languages: it is clear that spatial cognition is not necessarily constrained by the language that one knows. For example, speakers of languages that do not habitually employ absolute direction terms can do so when these terms are suited to the task (see Tversky, chapter 12, this volume). In addition, speakers of Tzeltal may use egocentric (or deictic) relations, especially when these are not overshadowed by absolute direction relations (Brown and Levinson 1993). More generally, it follows as a matter of logic that some understanding of spatial relationships must be available prior to the acquisition of spatial language; otherwise, it would be impossible for spatial language to be acquired in the first place. While this leaves open the possibility that exposure to different languages can engange certain aspects of spatial cognition to a greater extent than others, it does not support the strongest Whorfian hypothesis that one’s manner of thinking about space is entirely determined by the language one learns. (For more general discussion of this point, see Fodor 1975.)

The preceding discussion and much of the evidence in this volume, implies that the exact mappings between CR_A and CR_B are plastic in that they can look rather different in different languages. Need this have implications for the structure of CR_A and CR_B themselves? The short answer is no. One need not assume that how one talks about one’s spatial concepts necessarily influences those concepts in some fundamental way. What we can talk about is certainly less than what we know, and what we know consciously, whether we can state it precisely or not, is certainly less than what we know in toto. Cultures may influence how we choose to refer to spatial attributes, and even which spatial attributes we choose to refer to, but there is little support for the view that they, or the languages they use, fundamentally alter our spatial understanding of the world. Under the assumption that experience does not fundamentally


Space and Language


alter perceptual and cognitive processes, the study of intrinsic organizing factors should offer a window upon the underlying conceptual representations. By comparing how we use language to refer to space and spatial relations with how we behave in space, we can gain insight into some of these “distortions.”

15.5 Conclusions

At the outset we posed four central questions in the study of space and language: (1) Which aspects of space can we talk about? (2) Which aspects of language reflect particular spatial attributes? (3) Are spatial CRs changed by experience? (4) What light can the study of space and language shed on the nature of conceptual representations?

There are aspects of spatial knowledge we cannot naturally talk about (for example, absolute distance between two objects or between an observer and an object), and aspects of spatial knowledge we can talk about (for example, spatial relations), but we cannot at present provide a satisfying distinction between these two classes of spatial knowledge. Although the distinctions can be described by terms like precise and coarse spatial representations, we do not believe that those terms accurately express the CRs that might underlie the distinction. The suggestion by Landau and Jackendoff (1993) that nouns and prepositions describe objects and spatial relations, respectively, is an important start to the project of understanding how language maps to space, but we suspect that a broader view, namely, that nouns describe entities (including, but not limited to, places and objects) and prepositions describe relations, is likely to be closer to the truth. There is evidence that different cultures refer to space in different ways, but there is no reason to suppose this involves a change in the underlying conceptual representations, as long as the distinction between CRs and semantic representations is kept clear.

We pointed out the importance of a careful analysis of the intrinsic “organizing factors” that interact with environmental information to structure our knowledge of the spatial world. These organizing factors act as a kind of “syntax” in accord with which inputs to spatial systems are ordered, and in so doing they contribute meaning to the spatial representations themselves. This is perhaps clearest in the allocentric map observed in the hippocampus, but it is also observable in other cases. It is our view that careful study of the way language reflects these organizations or “distortions” should help illuminate the CRs. By itself, however, such a study will not accomplish the entire task. The relationship between spatial language and other aspects of cognitive processing, such as our intuitive understanding of motion (Talmy, chapter 6, this volume), the on-line recognition of spatial relationships (Logan and

Space and Language

Sadler, chapter 13, this volume), and our deductive inferences about these relationships (Johnson-Laird, chapter 11, this volume), must also be carefully unpacked if we are to derive maximum benefit from the study of language and space. Progress in these areas should improve our understanding of the relations between space and language, which in turn could illuminate the nature of conceptual representation.

Acknowledgments

We thank the McDonnell-Pew Cognitive Neuroscience Program for fostering our collaboration on this chapter, which was written while Mary Peterson and Lynn Nadel were on sabbatical from, and supported by, the University of Arizona.

Notes

1. Levinson’s claims rest on the assumption that performance on old/new recognition memory tests predominantly reflects differences in perceptual organization or processing. See Hochberg and Peterson (1987); Peterson and Hochberg (1983) for criticism of this assumption.

2. Both Fodor (1975, 85–86) and Jackendoff (1983, 17) allow that, by combining the primitive CRs differently, different languages may render a given idea more or less salient to the speakers of those languages.

References


