6. Language, Spatial Cognition, and Vision

Annette Herskovits

6.1. Introduction: Speaking About the Spatial World

One essential function of language is to refer to objects and situations in the world. This process is mediated by nonlinguistic mental representations, most prominently by perceptual representations in different modalities. Human minds have the ability to establish systematic relationships between linguistic forms and perceptually based knowledge. This *grounding* of linguistic symbols in perceptual representations (Harnad, 1990), though often overlooked in linguistics and artificial intelligence, is essential to understanding linguistic abilities and linguistic structure. And a good way to examine it is to investigate our ability to talk about space; the spatial world seems amenable to precise and objective description – unlike, say, the world of smells and feelings – and much is known about visual and spatial perception.

As this chapter shows, close semantic analysis of spatial expressions opens a window onto the interplay of faculties that supports our ability to describe the spatial world. There are several kinds of spatial expressions, but spatial relation terms are key elements in a large proportion of them, as they are the principal means available to speakers for the description of location and path. So, in this chapter, I examine the meaning and use of spatial prepositions in English against the background of current knowledge on vision and spatial cognition and infer the consequences of this exploration for the processes linking language and spatial knowledge.

In speaking about the spatial world, we tap several sources of knowledge: linguistic knowledge; conceptual knowledge; world knowledge; perceptual knowledge (particularly the output of visual perception); and

O. Stock (ed.), *Spatial and Temporal Reasoning*

nonperceptual spatial representations. I offer brief characterizations of these before sketching the connections between semantic facts and spatial cognition brought out in this study.

6.1.1. LINGUISTIC AND CONCEPTUAL KNOWLEDGE

The conceptual system consists in the set of concepts and principles of conceptual combination by which we categorize our experience. The semantic system of a language cannot be identical with the conceptual system because we think and reason with concepts that have no ordinary linguistic expression – mathematical ones, for instance. But are semantic structures simply a subset of conceptual structures – that is, the linguistically expressible subset? Some linguists assume this is the case (Jackendoff, 1983; Langacker, 1991); others (Bierwisch, 1988; Pinker, 1989) propose separate conceptual and semantic systems, assigning distinct semantic and conceptual structures to a sentence.

The linguistic analyses of Pinker differ significantly from those of Bierwisch, but both are moved to distinguish two levels of structure by the assumption that nonlinguistic conceptual thought is universal. Given that semantic distinctions can be language-specific, "equating linguistic semantic representations with the conceptual categories underlying nonlinguistic thought is tantamount to a very strong and implausible Whorfian claim" (Pinker, 1989, p. 357). Yet the evidence so far for a language-independent, universal conceptual system is rather weak. Bowerman (1996), Langacker (1991), and Levinson (1993) describe differences between languages that would be hard to reconcile with a common underlying conceptual base.

This study assumes only one level of meaningful structure, which I call indifferently semantic or conceptual. Collapsing semantic and conceptual structure compels us to assume that conceptual systems are significantly shaped by the language one speaks, a position congruent with the views of lexical knowledge advocated in this chapter.

6.1.2. COMMONSENSE KNOWLEDGE OF THE WORLD

Commonsense knowledge of the world includes general knowledge about location, shape, gravity, time, motion, and causality; about the properties of liquids, solids, gases, and various particular substances; about the many kinds of objects with which we interact (people, animals, plants, buildings, lakes, rivers, and so on); and about the self and its range of behaviors. This commonsense world is the world "as it is," the reality in the background of our linguistic productions; it is the default reference world, the one we typically talk about; it is "objective," not in any scientific sense but according to our naive ontology.

Work in artificial intelligence typically assumes that a representation of commonsense knowledge is the key to semantics, that the meaning of a lexical item corresponds to some structural configuration in the knowledge representation. The semantics of a language is thus definable as a mapping from lexical items to such configurations. But, in fact, world knowledge can take us only so far toward understanding semantics. Languages manifest subjective ways of structuring the objective, commonsense world, and much of this chapter is concerned with understanding that structuring in the domain of space.

6.1.3. SPATIAL COGNITION

Spatial cognition – the collection of mental structures and processes that support our spatial behavior – involves a complex and various set of abilities:

- *Visual abilities*: Vision is a primary source of spatial information.
- *Other perceptual abilities*: The aural, kinesthetic, haptic, and olfactory systems all deliver spatial information.
- *Motor abilities*: Walking, reaching, interacting with, and using various objects (avoiding an obstacle, sitting in a chair, putting a cap on a bottle, and so on) depend on coordinating perceptual and general knowledge to direct motion.
- *Navigational abilities*: When moving through medium- to large-scale spaces, we use *cognitive maps* as guides (Kuipers, 1983). Unlike real maps, these representations are fragmented – some regions are rich in detail, others only roughly characterized – and distorted. They may include visual memories, but encode essentially the location of landmarks and routes connecting the landmarks.
- *Mental imagery*: We can evoke in our minds scenes observed from a specific viewpoint, imaginary or remembered scenes. Kosslyn (1980) and Finke (1989) claim these mental images are functionally similar to perceptual representations.
- *Spatial mental models*: As we attempt to comprehend linguistic spatial descriptions, we create models of the situations described (Johnson-Laird, 1983; Tversky, 1993b). These share properties with mental images (they are in part analog) but are apparently three-dimensional rather than dependent on a particular vantage point.
- *Spatial memory*: This complex ability involves several distinguishable mechanisms (Wadell and Rogoff, 1987; Stiles-Davis et al., 1988). It

intersects with mental imagery and cognitive maps.

- *Spatial reasoning and problem solving*: To plan a trip or design an appropriate configuration of furniture in a room, one must engage in spatial problem solving. This joins logical reasoning with the ability to create and manipulate various mental images and models.

6.1.4. FROM SPATIAL COGNITION TO LANGUAGE

Spatial cognition, then, includes many abilities and distinguishable representations. Which representations can language access? One possible answer is that it has access to a veridical (but partial) three-dimensional map of the environment – a map constructed from perceptual representations and world knowledge, yet consistent with cognitive maps and with the representations guiding motion. Such a map would explain our ability to coordinate different spatial operations – for instance, when we use both a cognitive map and sights along the route to find our way. Moreover, some such view of the *real world* – whether explicitly represented or not – is always in the background of our utterances. Otherwise there would be neither truth nor reference, and hence no communication.

But we refer using linguistic categories (of objects, spatial relations, actions, and so on). A sentence is not a *copy* of a scene; it is a statement that the scene belongs to a certain category, which every part of the sentence helps specify. The categories available in languages are only partially determined by the world; they can be defined in many different ways as shown by variation among languages. Thus, this (presumably essentially universal) view of the *real* physical world does not determine what concepts are used to describe a scene. And, as we will see, the categorizing of scenes by spatial relations does in fact show that language accesses representations distinct from this hypothetical veridical map.

The question of how scenes are categorized by spatial relations is tied to the question of *schematization.* A phrase like *the tree lying across the road* reduces the rich information contained in the referent scene to the mention of two object categories (tree and road), and the assertion that the axis of the tree leads from one edge of the road to the other. The relation expressed is based on simple geometric descriptions of two complex entities – the tree as an axis and the road as a ribbon – and their detailed arrangement reduced to a very simple geometric relation. This is schematization.

It seems sensible to assume that schematization is based on general spatial cognitive processes and that the geometric representations of objects in spatial expressions are clues to the spatial representations accessed by language. It is therefore essential to examine schematization closely; to see exactly which schematizations are used in which contexts. For example, it is often assumed that schematization means that the actual arguments of a preposition are always simple geometric representations of the objects: points, lines, planes, or blobs. A close look shows that such reduction happens sometimes but not always. Where reduction does occur, it reveals a great deal indeed about the representations accessed by language. For example, it provides evidence that spatial expressions often invoke cognitive maps, in which landmarks and moving objects are seen as points. But at times fully detailed close-up representations of scenes, including precise representations of objects' shape, are needed to decide whether a preposition is applicable. In fact, for many prepositions, there is evidence of two categories of use: one in which the objects are close to each other and requiring full representations of objects shape and placement, and one in which the objects are far apart, when a representation analogous to a cognitive map is used.

Objects are represented as points in almost all descriptions of motion (Section 6.3.3). Other representations accessed by language include a two-dimensional representation homologous to the visual array, and two-dimensional ground maps for small-scale environments (those for large-scale environments are simply cognitive maps). In some cases, Gestalt processes of perceptual organization provide the applicable schematic representations of objects. One may not want to consider each of these as distinct, but it is important to be precise about which object representations are used in spatial expressions and about their origin.

Schematization also sheds light on a hypothesis proposed by Landau and Jackendoff (1993), which posits a division of spatial language and ascribes it to a division in the processing of spatial information in the brain. The hypothesis, based on a mistaken view of schematization, is shown to be invalid (see 6.5.2).

The representations accessed by language are part and parcel of nonlinguistic spatial cognition, but the spatial relations expressed need not be explicit elements of these representations. In the last part of the chapter (6.5.3), I argue that perceiving and expressing a spatial relation requires, at least at times, visual processing beyond a simple act of attention. The related objects must be configured together, and their configuration assigned to a prepositional category. These operations might not take place without a linguistic goal; thus, there are language-induced percepts.

The balance of this chapter is divided into four main parts. Section 6.2 presents the English system of spatial prepositions and general patterns in the way they are used to specify spatial characteristics; Section 6.3 examines schematization; Section 6.4 concerns the striking fluidity of prepositional meaning and questions of polysemy and prototypicality; and Section 6.5 looks at the interface between language and spatial cognition.

6.2. Usages of the Prepositions

Prepositions denote spatial relations, which are principally used to predicate constraints on such attributes as location, trajectory, orientation, direction, and disposition. Most prepositions are polysemous (Herskovits, 1986; Brugman, 1988), but one finds patterns in the polysemy that allow one to classify them. Thus, the *location prepositions* (Table 6.1) have only location senses; the prepositions can be used in motion sentences, but these do not require defining additional (motion) senses. The *primarily motion* prepositions, on the other hand, have distinct motion senses and stationary senses, some of which are systematically related to the motion senses. For convenience, I term these prepositions the *motion prepositions.* The prepositions labeled *misfits* follow patterns of their own.

6.2.1 BASIC USAGES OF THE LOCATION PREPOSITIONS

The prepositions in the first column fit in a syntactic frame:

NP is preposition NP

as in

The ball is ...
...on the table / against the wall / among the stones / at the store / behind the wall / ...

In these expressions, the location of the first object (the ball) is constrained with respect to the second by the spatial relation denoted by the preposition. Following Talmy (1983), I will call the located object the *Figure* and the reference object – the object of the preposition – the *Ground*:

The ball	is on	the table
Figure		*Ground*
Located object		*Reference object*

It is often assumed that the [Preposition + Ground] phrase defines a region of space, and the meaning of a locative sentence is an assertion that the Figure is located in that region. This is often the case, but not always:

The milk is in the cup.
The cat is under the bed.

These statements do mean that the milk is located in the interior of the cup, and the cat in the region *under the bed.* But prepositions implying contiguity – such as *on*, *against*, and *on top of* – cannot be defined in terms of inclusion in a region, however the region is defined: contiguity cannot be reduced to inclusion in the surface of the object, or in the region above the object, or in the region around it. An object *on the desk* is not in the

Table 6.1. English spatial prepositions

Location	**Primarily Motion**	**Misfits**
at/on/in	across	over
upon	along	about
against	alongside	throughout
inside/outside	around	after/before
within/without	away from	ahead of
near/ (far from)	toward	for
next	up/down	
beside	to/from	
by	into/ (out of)	
between	onto/off	
beyond	past	
opposite	through	
amidst	via	
among		
above/below		
under		
beneath		
underneath		
on top/bottom of		
on the top/bottom of		
behind		
in front/back of		
at the front/back/right/left of		
on/to the left/right of		
left/right of		
north/east/west/south of		
to the east/north/south/west of		

surface of the desk; inclusion in the regions above or surrounding the desk does not entail contiguity with the desk.

6.2.2. BASIC USAGES OF THE MOTION PREPOSITIONS

Every motion preposition fits in a syntactic frame:

NP [activity verb][1] Preposition NP

as with

The ball rolled ...

...across the room / along the street / toward the boy / away from the curb / ...

Here the Figure is the moving object; the Ground is still the referent of the object of the preposition; the preposition constrains the trajectory, or path of the Figure.

The ball	rolled across	the street
Figure		*Ground*
Moving object		*Reference object*

To illustrate how these prepositions constrain the Figure's path:

1. One sense of *across* implies motion from one side to the opposite side of a ribbonal Ground (such as a road). The directed lines in Figure 6.1a show possible paths satisfying the description *across the road.*
2. *Along*, in one main sense, requires a linear Ground (Figure 6.1b); the path may be in, on, or alongside (parallel to and outside) the Ground. Since the Ground need not to be straight, we must extend the notion of parallelism to lines of any shape.
3. *Around*, in one main sense, requires that the Ground be a region bounded by a closed line; the trajectory can be any closed line either circumnavigating the Ground or within it at a short distance from the boundary (Figure 6.1c).
4. *Toward* applies to any path leading in the direction of the Ground (Figure 6.1d), though not necessarily reaching it. So in *Walk toward the tower for half a mile, then turn right!*, the turn may be any distance from the tower. The path need not be straight; it may be the customary, or the only possible, route to the Ground. Thus, interpretations of *toward* and *away from* must take into account the geography of the earth.
5. *Through* describes any trajectory whatsoever within a three-dimensional Ground.

In conclusion, a motion preposition defines a *field of directed lines* with respect to the Ground. A motion sentence using such a preposition asserts that the path of the Figure coincides with one of the directed lines of the field defined by preposition and Ground.

The [Preposition + Ground] phrase does not refer to a path; it is a *predicative* phrase, which can be used to express a constraint on a path referred to by other means.[2] So in

Jim walked across the street.

Jim walked implies the existence of a particular path. *Across the street* predicates that this path satisfies a constraint: it is one of the continuous infinity of paths defined by *across the street.* The "extension" of *across the street* is not one path, but a field of paths. A prepositional phrase alone does not specify "one" path.

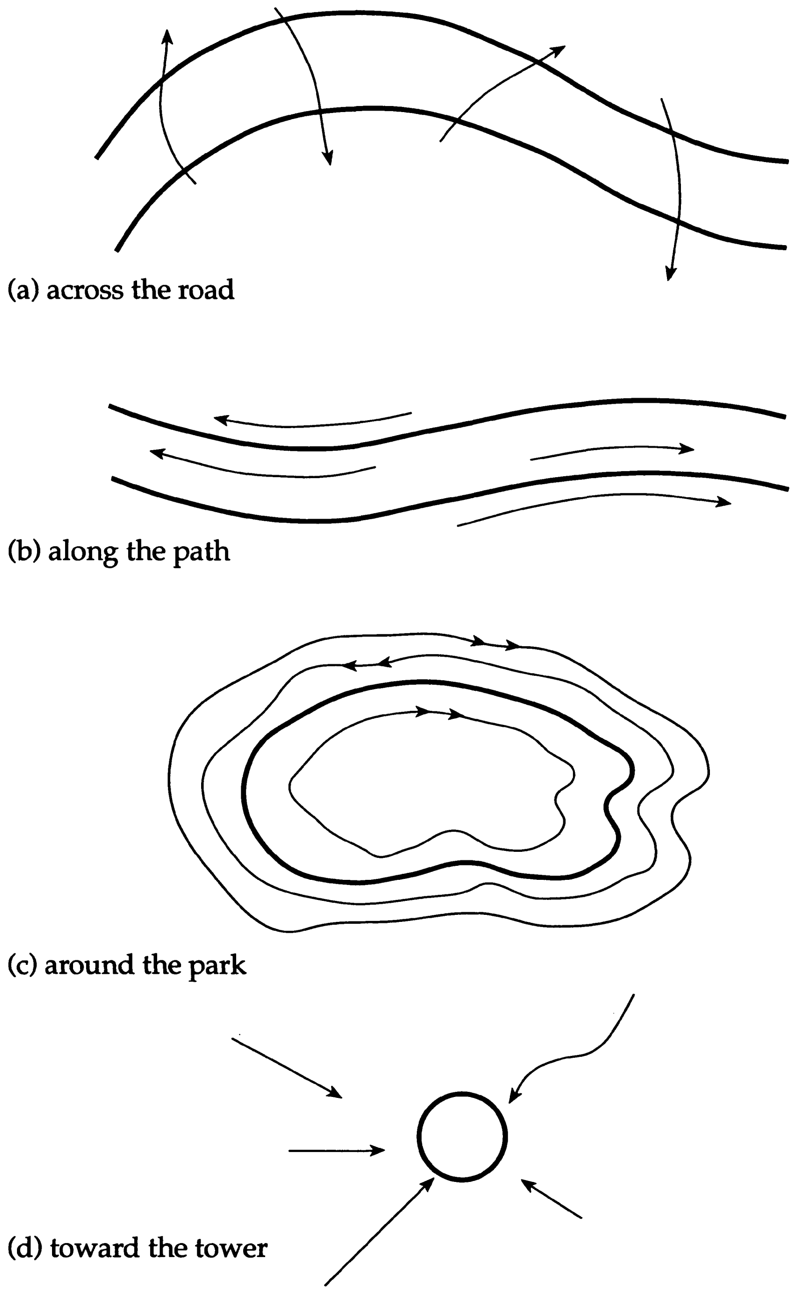

Figure 1. Fields of directed lines associated with various motion prepositions

Although it is often said that motion prepositions specify the orientation of a path, this is not generally the case, for any reasonable and consistent definition of orientation. Since a trajectory can be a directed line of any shape, no single value of orientation can be called "its" orientation. Orientation can be defined in accord with usual intuitions for straight lines: all straight parallels have the same orientation; thus parallel straight trajectories running in opposite directions have the same orientation – mathematically, orientation is an equivalence class of parallel lines. But if a trajectory twists and curves, its orientation changes at every point.

6.2.3. MOTION USES OF THE LOCATION PREPOSITIONS

Location prepositions are used in motion sentences in three ways:

1. To constrain the location of the entire trajectory:

   The butler was walking ...
   ... among the guests.
   ... on the floor.
   ... under the trees.
   ... near the park.

   This class of motion uses has an interesting property, which differentiates it from the basic uses of the motion prepositions; at every point of its trajectory, the figure satisfies the relation denoted by preposition and Ground. So the walking butler remains always *among the guests*, *on the floor*, and so on. But a man walking to the store is not always *to the store*; a chicken running across a road is not always *across the road*.
2. To constrain the endpoint of the trajectory:

   The cat ran ...
   ... under the bed.
   ... outside the room.
   ... between the two chairs.
   ... behind the curtain.

   These sentences can also mean that the entire path is *under the bed*, *outside the room*, and so on, though that interpretation is pragmatically somewhat less natural.
3. To constrain the location of a point or segment internal to the path:

   Jack walked by the house.
   The cart passed in back of the house.
   The geese flew over the house.

   This interpretation may be triggered by the verb (pass) or by the preposition (*by* strongly favors this interpretation). Or it may be suggested by pragmatic factors (the geese probably did not fly in a minimal arc over the house).

I think it more appropriate to assume here different types of use rather than three additional senses. The different interpretations involve locating the event referred to by the clause, or a part of that event, using the same sense of the preposition as is used to locate objects; it seems therefore more natural to attribute the differences to different syntactical meanings associated with the clause structure – as is easily done in a construction grammar approach (Goldberg, 1995) – than to different senses of the preposition.

6.2.4. STATIONARY SENSES OF THE MOTION PREPOSITIONS

Though most of the motion prepositions have several motion senses, one usually stands out as highly salient; and two stationary senses systematically related to that motion sense – the *Figure disposition* and *vantage point* senses – are frequently associated with the preposition.

6.2.4.1. Figure disposition senses

They are illustrated by

The snake lay...
...across the road.
...along the wall.
...around the tree.

The Figure must be a linear object coaxial with one of the directed lines of the field defined by preposition and Ground object. These are sometimes called *Figure orientation* uses, but *disposition* seems a better term; just as a motion preposition does not define "the" orientation of a trajectory, it does not define the orientation of objects whose shape is not fixed by their category, such as snakes and pathways.

There are restrictions on the Figure disposition uses of the motion prepositions. Consider

The snake lay ...
... ♯ up/down the tree trunk.[3]
... * up to the edge of the ditch.
... * from the rock to the tree.
... ♯ past the stone.[4]

The two first sentences cannot mean that the snake's body is aligned vertically with the tree trunk. The last cannot mean that the snake's body is like a path running past a stone. It is not clear why these are not acceptable. Acceptability depends on subtle factors, as shown by the following sentences, which are analogous to the ones above:

He has bangs down to his eyebrows.
His arm is scarred from the elbow to the shoulder.[5]

All the motion prepositions can be used in instances of *virtual motion* such as

The road runs ...
... up to the top of the hill.
... past the stone.
... from the rock to the stream.

It is as if these expressions included an instruction to mentally scan along the road.[6]

6.2.4.2. Vantage point senses

These involve asserting that the Figure is located at the end of a (virtual or real) path constrained as specified by preposition and Ground.

The car is ...
... across the street.
... across the corner.
... up the hill.

In every such use, there is an implicit vantage point where the virtual path starts. So *across the street* is across from some contextually defined vantage-point, usually associated with the speaker.

Here also we find restrictions, but they are clearly motivated:

The car is ...
... * to the church.
... * onto the hill.
... * from the church.

Saying **The car is to the church* to mean *The car is at the end of a path leading to the church* is needlessly indirect: *The car is at the church* says the same thing. The unacceptability of the other two examples can be explained in a similar way.

Different senses, rather than different usages of the same sense, are involved here. The primary sense of the preposition (the salient motion sense) cannot be used in describing the meaning of sentences with the corresponding stationary senses, unless one brings in a "hypothetical" path; but one would not want such a notion to be introduced as part of the meaning associated with clause structure.

6.2.5. THE MISFITS

Over is one preposition that has a highly salient location sense but an equally salient motion sense:

The lamp is over the table.
He walked over the hill.

Over in the second example does not mean the same as in the first; the path is not only above (on) the hill: it leads from a point at the base of the hill to a point opposite by way of the summit. *Over* has the Figure disposition and vantage point senses derived from that motion sense:

The rope goes over the wall.
The post-office is over the hill.

About and *throughout* also have equally salient motion and location senses:

There were newspapers about/throughout the room.
He walked about/throughout the apartment.

and do not fit in the frame [NP is Preposition NP]. They require an existential or a verb implying a *distributed* configuration – that is, a collection of objects scattered through space with an approximately uniform distribution:

*Newspapers were about/throughout the room.
Newspapers were strewn about/throughout the room.

As for *ahead*,

Jack walked ahead of Mary.

does not constrain the trajectory of Jack with respect to Mary; instead its semantic contribution includes the presupposition that Jack and Mary were simultaneously walking on the same path in the same direction, and the assertion that Jack preceded Mary on that path. *Before* and *after* have similar spatial meanings. *Before*, but not *after* and *ahead*, has also a simple locative sense:

Before me was my long-lost friend.

For can specify spatial extent:

Jane walked for six miles.

For expository purposes, this section was organized in terms of patterns found in the polysemy of English prepositions. But these patterns manifest only tendencies, not necessity, logical or other, since the misfits are exceptions.

6.2.6. FIGURE AND GROUND ASSIGNMENTS

The Figure is the object whose location, disposition, or path is at issue (Talmy, 1983); that is to say, it is moving, or *conceptually movable* (which means the same as saying that location or disposition are at issue). The Ground is an object conceived of – perhaps provisorily – as stationary: its location is generally assumed known to the addressee, so that characterizing the location, disposition, or path of the Figure with respect to it will provide adequate localization information.

But what does *knowing the location of an object* mean? Most typically, we assume that the earth is stationary, and knowing an object's location involves knowledge of a place fixed with respect to the earth, as with the living room in

Jack's glasses are in the living room.

But consider

Jack's glasses are on his nose.

The location of Jack's nose is known relative to his body; in this instance, Jack's body is conceived of as stationary, though it can move about. The truth conditions are independent of the location of Jack himself.

In

Jack's glasses are in the next car of the train.

the train, moving or not, is conceived of as stationary, and knowing the location of the car involves only knowledge of a place fixed relative to the train. One could say Jack's glasses are *in the same place as yesterday*, meaning either that they are on his nose (wherever he may be), or in the next train car (even though the train traveled to California). Clearly, what it means to know the location of an object is rather tricky.

Reference objects tend to be – but need not be – larger and less mobile than Figure objects. This is a matter of communication, not spatial cognition; addressees are more likely to know the location of large, fixed objects. But we can say either

The newspaper stand is near Trafalgar Square.

Trafalgar Square is near the newspaper stand.

although in some spatial tasks, Trafalgar Square, but not the newspaper stand, would be a "reference point" (Sadalla et al., 1980).[7] Only when the Ground is considerably smaller and more mobile than the Figure is a sentence odd:

? The house is near the bicycle.

?* The bottle is under the cap.

This may suggest rigid conventional restrictions, but is more likely a product of the workings of a usage-based system – a system affected by practice and by the strength of memory traces, as is true of language (Bybee, 1985). The process of checking size and mobility, practiced innumerable times, is not disabled in these cases, though its original purpose may not apply.

6.3. Schematization

Schematization is characterized in Talmy (1983, p. 225) as

> a process that involves the systematic selection of certain aspects of a referent scene to represent the whole, disregarding the remaining aspects

and in Herskovits (1986, p. 2) as follows:

> there is a fundamental or canonical view of the world, which in everyday life is taken as the world as it is. But language does not directly reflect that view. Idealizations, approximations, conceptualizations, mediate between this canonical view and language.

Systematic selection, idealization, approximation, and conceptualization are all facets of schematization, a process that reduces a physical scene, with all its richness of detail, to a sparse and sketchy semantic content. This reduction is often said to involve applying some abstract spatial relation to simple geometric objects. So

> The village is on the road to London.

would imply contiguity between a point (*the village*) and a line (*the road*).

Schematization has been discussed mostly in linguistics and psycholinguistics; no artificial intelligence work provides explicit computational accounts of it. There are two sets of questions here:

1. Which schematic representations of the objects are used in which contexts? How are objects and trajectories related to their schematic representation? Is it true that objects in the context of a preposition are always represented as either points, lines, planes, or blobs?
2. Is schematization related to language-independent spatial cognitive processes? What precise computational processes underlie linguistic schematization?

6.3.1. ABSTRACTION, GEOMETRIC IDEALIZATION, AND SELECTION

Schematization involves three distinguishable processes: abstraction, idealization, and selection. Abstraction, of course, is an essential characteristic of *all* linguistic meaning. Every linguistic category abstracts from the distinguishing characteristics of its individual members. In saying

> Joe is running.

we abstract away from particular distinguishing characteristics of Joe's running – speed, style, location, goal, and so on. Similarly, in saying

> There is a tree lying across the road.

we abstract away from such characteristics as

– the position of the tree along the road,
– the angle between tree and road axes,
– the positions of the ends of the tree with respect to the the road's edges,
– the width of the road, and

– whether the road is in an horizontal plane or inclines steeply.

The facet of schematization particular to spatial language is *geometric idealization.* We *idealize* features of the real scene so they match simple geometric objects: points, lines, and so on. Idealization goes beyond abstraction: the real geometric features do not exactly match the geometric categories in which we fit them. Thus, the top surface of the road in the above example, though it may be bumpy and of varying width, is (arguably) idealized to a ribbon.

Selection involves using a part or aspect of an object to represent the whole object, as with

the cat under the table

where the top of the table stands for the whole table. Including selections stretches the ordinary meaning of schematization – yet selections do fit Talmy's definition, and they commonly produce the reduced object geometry relevant to spatial expressions.

6.3.2. TREATING OBJECTS AS POINTS: A FALLACIOUS ARGUMENT

Talmy (1983, p. 234) writes that, typically, the prepositions "treat the focal object [the Figure] as a point or related simple form." This is a frequently expressed intuition, but it is not clear what *treating an object as a point* might mean. One justification often given for this view is that if a preposition puts no constraint on the geometry of one of the objects related, then that object is treated as a point. As most prepositions do not restrict Figure shape, it follows that the Figure must generally be treated as a point. Let us examine each step of this argument.

Some stationary senses of the prepositions do in fact put constraints on Figure shape. Figure disposition senses require a linear Figure:

The snake lay across/along the trail.

For one sense of *over*, the Figure must be a surfacy object:

The tablecloth lay over the table.

For one sense of *throughout*, the Figure must be a composite aggregate:

There were blackbirds throughout the tree.

But, other than these few instances, the stationary senses of the prepositions put no constraint on Figure shape. So is the Figure then treated as a point?

There are clearly cases where the Figure is *not* treated as a point. The Figure can be infinite or unbounded:

He contemplated the firmament above him.
The land beyond the river is fertile.

The firmament extends infinitely upward; the land stops somewhere, but this outer boundary is not part of the conceptualization – it is outside the scope of the mental eye. It would be absurd to claim that infinite or unbounded objects are seen or treated as points; only bounded objects can be idealized as points. But even a bounded Figure is not necessarily seen as a point: idealization to a point appears irrelevant to

The orange juice is in the bottle.
The sheets are on the bed.
The Atlantic is between Europe and America.

If the Figure is often not treated as a point, perhaps the problem is with the assumption that when a preposition places no constraints on the shape of one of the objects related, then that object is idealized to a point. And indeed, closer examination shows the assumption to be invalid. The error may stem from the logical misstep: *if a prepositional predicate applies to objects of any shape, then its truth in particular cases can be assessed without referring to the object's shape.* But that is clearly false.

Consider the preposition *in.* It does not restrict the Figure shape in any way; an object of any shape or even any dimensionality will do. The selection restrictions for the Ground are equally loose; any Ground shape will do, except a point (nothing can be "in" a point). But it certainly does not follow that Figure and Ground are treated as points in uses of *in.* Ullman (1985) describes various algorithms for deciding whether a point is in a closed curve. These, as one would expect, require full knowledge of the shape of the curve; hence the Ground (the curve) is not seen as a point. If the Figure is *extended,* we must also know the position of every one of its points to decide whether it is *in* the curve. Therefore we treat neither Figure nor Ground as points.

Perhaps objects are seen as "blobs" rather than points. Blobs may be what Talmy meant by "related simple form." A blob must be how we apprehend an extended object (in two or three dimensions) whose precise shape and extent are not known. Its representation could consist of the position of a center point of the object, together with the assumption that the object extends outward from this center to an indeterminate boundary; the boundary could be additionally known to lie outside a given area surrounding the point. But without the precise extent of the Ground, we cannot in general[8] decide whether an object is in it; in fact, we need also to know the exact region of space occupied by the Figure. So blobs will not do; neither will lines or planes.

The belief that *all* prepositions treat Figure and Ground as points, lines, planes, or blobs is unfounded; some do, and some do not; some do in some uses and not in others (6.3.5). At times, we need to know the precise regions occupied by Figure and Ground to decide whether a preposition applies.

Figure and Ground are then seen (as far as shape is concerned) just as they are.

6.3.3. FIGURE AS A POINT: THE MOTION PREPOSITIONS

While not all prepositions idealize Figures to points, there is some reliable evidence of such a process. The basic meanings of the motion prepositions are all cast in terms of linear paths. Such paths can be traced out by a point, or by a linear, deformable object sliding along its own axis – for instance (ignoring their cross-section), a snake or a train. As the latter objects are not all that common, we must somehow make use of predicates defined in terms of motion of a point to talk about motion of any extended object.

The path of a rigid object undergoing translation has the shape of a generalized cylinder, a kind of snake, but with possible overlaps and not necessarily with a circular cross-section. If the object rotates on itself, changing its orientation as it translates, the volume described is typically too complicated to visualize precisely. Kinematics shows we can analyze this movement as a succession of infinitesimal motions, each combining a translation and a rotation around some axis, but for the most part we lack the ability to represent subcategories of paths involving different such combinations. Fine conceptual distinctions among the possible paths of a *nonrigid* object are of course even harder to make. The restriction is not in our awareness of the object's changing appearance as we watch it move; it is in our ability to analyze the motion, albeit unconsciously and nonverbally – that is, to note separable aspects of it. This ability is needed to form different subcategories of motion and to assign a particular instance to a subcategory.[9]

Because of this limitation, we typically idealize the motion of a three-dimensional object to that of a center point – probably its *centroid* or center of gravity assuming uniform density – ignoring rotations around the centroid. Note, for instance, that there exist no two prepositions that contrast only in that one entails a pure translation and the other translation accompanied by rotation. A few verbs do describe trajectories involving different combinations of these motions – for example, *roll, flip, tumble, turn, twirl, revolve* – but the list is not a long one, which supports the notion that our ability to conceptually distinguish motions along these dimensions is limited.

Consider

The child danced around the Maypole.

Although we would assume that the child rotates on herself, no such notion is expressed by the preposition. The child would still be *dancing around the Maypole* if the pole were on a stage and she continued facing the audience while dancing around it.

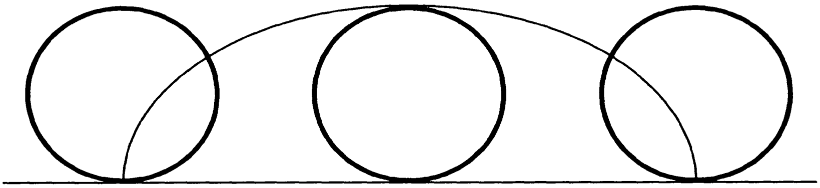

Figure 2. Trajectory of a point on the circumference of a rolling log

In some expressions, the preposition does not describe the trajectory of the centroid:

He rolled the log over.

The earth turns around its axis.

use prepositions specifying the motion of a point in a different way. Brugman (1988) provides a nice explanation for the first sentence. As the log rolls along the ground a distance equal to half its circumference (Figure 6.2), the point of the log originally in contact with the ground comes to the top; as the log rolls on another half-circumference, the point goes back to the bottom. Its trajectory is a curve called a *cycloid.* It is as if this point passed *over* the object. In other words, *over* applies to that point only, though we talk about motion of the entire log.

With the earth turning around its axis, each point of the sphere, except the points on the axis, describes a circle around the axis. *Around* constrains the trajectory of every such point.

Consider also

The sea rushed onto the sand.

The water rose up to the roof.

In the first sentence, each wave (a part of a superficial layer of the sea) follows a path constrained by *onto the sand*; in the second, every part of water follows a path leading *up to the roof.* The prepositional phrases then describe the motions of each part (however defined) of the Figure object. Note that this *is* a schematization: the actual motions of the sea and of the rising water are far from so orderly. But it is not idealization of sea and water to a point.

In conclusion,

1. The basic meanings of the motion prepositions are all cast in terms of motion of a point;

2. Most often, such predicates are adapted to the motion of a three-dimensional object by abstracting the rotation components of the motion and reducing the motion to that of the object's centroid. This is effectively idealization to a point, and it follows in part from limitations in spatial cognition.
3. There are a few exceptions to this pattern: idiomatic ones (*roll over*); and cases where the preposition applies to all or most points (or parts) of a two- or three-dimensional Figure.

So far, we have considered the use of motion prepositions from the point of view of perception, but prepositions are not uniquely, or even primarily, for recording perceptual experience; they can be used to describe action – specifically, possible navigation paths.

6.3.4. MOTION PLANNING

The linguistic representation of objects' paths as lines is fundamental to motion planning. In planning to walk across a room, or to move one's hand to reach an object, an essential step is the mental construction of a *line of travel* that leads to the goal while avoiding obstacles. It is certainly a line: the plan would not usually include detailed descriptions of how to move feet and arms – this will happen automatically. Such a line probably still acts as a high-level guide during the execution of the move, while at a lower level, execution relies on modules that coordinate vision and motion and probably operate to a large degree autonomously (Goodale, 1988).

6.3.5. NAVIGATION AND COGNITIVE MAPS

As the question of how to go from here to there is a central concern of human beings, the descriptions of paths in large-scale spaces accounts for a major part of our uses of prepositions. Navigation in large-scale spaces is guided by cognitive maps whose major components are landmarks and routes represented respectively as points and lines. Moreover, in the context of a cognitive map, a moving Figure is conceptualized as a point and its trajectory as a line. Prepositions are frequently used to describe locations, pathways, and trajectories within a cognitive map:

Penny is at the market
Penny walked to the market.
This street goes to the market.

There is much linguistic evidence that a main sense of *at* is "coincidence of a movable point object with a point place in a cognitive map". For instance, *at* cannot be used to describe location in small-scale spaces:[10]

* The ashtray is at the bottle.

and

Jack is at the supermarket.

is typically infelicitous if the speaker herself is in the supermarket. A space that surrounds you is not naturally seen as a point; representing a fixed object as a point requires seeing it from a distance. In

I am at the supermarket.

the speaker may be "in" the supermarket, but the sentence evokes a context where he is on the phone and taking the point of view of an addressee located at a distant place.

Consider also

?* She is at the garden.

She is at the community garden.

The first is odd, as the vantage point evoked is from the house adjoining the garden; but the second is fine – the vantage point can be at a place distant from the community garden. The overall geometric context evoked is a cognitive map with landmarks and Figures represented as points.

Finally, one can show that, although the use of *at* requires only close proximity in "real" space, in the geometric idealization of a cognitive map, located object and reference object are collapsed to a point and close proximity becomes coincidence. So if

Jane is at the store.

she may be on the sidewalk next to the store. But in

One focus of the ellipse is at the intersection of the two lines.

the focus of the ellipse is necessarily coincident with the intersection point. Intuitions are very clear; *near* is not enough – the two points must coincide. This supports the meaning of *at* proposed above. One might object that intuitions in the domain of abstract geometry are irrelevant to use of the prepositions with real objects. But *by* means "in close proximity," and it does not shift its meaning when used in the domain of geometry:

Jack was sitting by his sister.

This point is by that one.

So *at* and *by* must differ in meaning. If we assume *at* means coincidence, we have no need to call upon an arbitrary shift in meaning to explain uses in the domain of geometry: with real objects as arguments, coincidence applies in a cognitive map, in which the locations of moving object and landmark are collapsed when they are "very close" to each other.

6.3.6. IDEALIZATION TO A POINT AND DISTANCE BETWEEN FIGURE AND GROUND

The *projective* prepositions (*to the right, above, behind*, and so on) exhibit a dichotomy: when Figure and Ground are far apart, they are viewed as points; when close, their shape and precise relative placement matter. This should come as no surprise. So in Figures 3a and 3b we can decide that *F is to the right of G* by approximating *F* and *G* to points – their shape and orientation are irrelevant. But if Figure and Ground are close together (Figures 6.3c through 6.3e), shape and exact relative placement matter: the rectangle in 6.3c is *to the right of G*, but the rectangle in 6.3d and the cross in 6.3e are not, though the location of the centroid of F is the same in all three cases.

So objects perceivable from a single vantage point (this defines "small-scale" as opposed to "large-scale" environments[11]), but "far" apart, are integrated in a spatial representation quite similar to a cognitive map – they are represented as points.

6.3.7. TREATING AN OBJECT AS A LINE: AXES

Axes play an important role in spatial cognition. Several models of object recognition proposed in computer vision are based on shape representations structured by axes (Binford, 1971; Marr, 1982). Axes also play a role in spatial language. In all Figure disposition uses of the motion prepositions, one can legitimately say that a preposition treats an object as if it were a line, that is, reduced to its major axis. The Figure is then elongated, and the only object knowledge required to check the preposition's applicability is the position and shape of the axis. But note that shape is not fully abstracted. The elongation axis may not be straight, and applicability of the preposition depends on its shape:

The hose lay ...
...across the road.
...around the flower bed.
...along the trail.

Moreover, except with *along*, the Ground cannot be reduced to a line. With *across the road*, we cannot judge whether the hose leads from side to side if the road is reduced to a line, and we must know the exact region covered by the flower bed to decide whether the hose goes around it. Only with *along the trail* do the median axes of trail and hose suffice to decide whether the sentence is true. So again, it is not the case that prepositions always treat objects as if they were points, lines, or blobs.

The kind of axis selected to represent an object supports Marr's 3D-model (1982). Usually, it is the *model axis* for the entire shape.[12] So in

Jack lay across the bed

the model axis of Jack's body is orthogonal to the bed's main axis. But the relevant axis may also be the "principal reference axis" – that is, an axis of the frame of reference used to specify the location and orientation

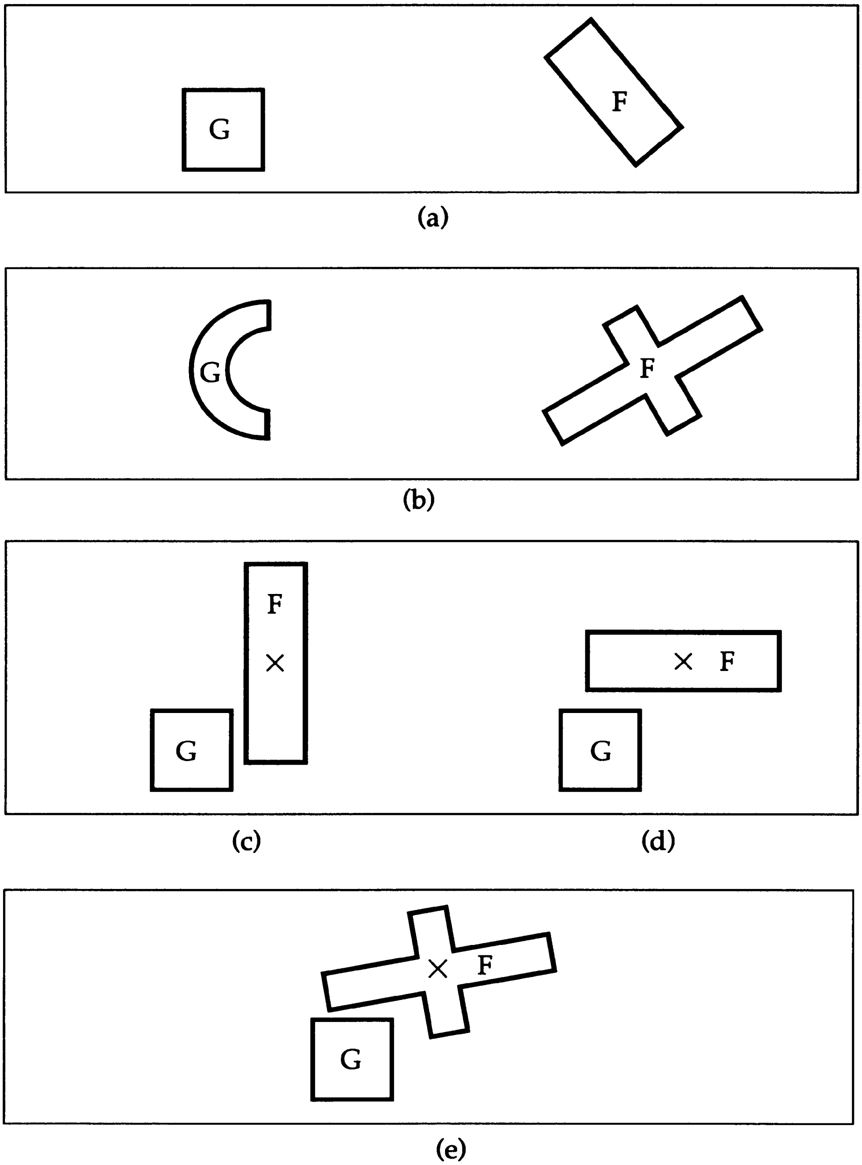

Figure 3. Effect of the distance between *F* and *G* on the truth of *F is to the right of G*

of the whole object's immediate parts. Model and principal axis need not be identical; Marr suggests that the principal reference axis for the human shape is the torso's axis, as the torso is connected to the greatest number of other parts. This turns out to be the Figure axis relevant to

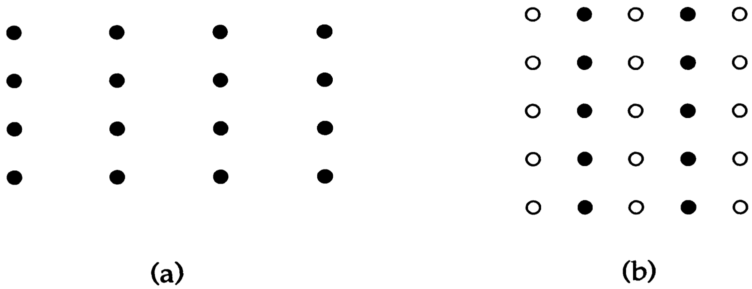

Figure 4. Grouping: effects of proximity and similarity

The driver leaned across the front-seat toward the passenger window.

Here, the axis of the driver's torso projects on the front seat orthogonally to the seat's median plane.

Recent theories of object recognition emphasize viewpoint-dependent knowledge rather than object-based componential representations. Two-dimensional views (especially canonical views) may play an important role in object recognition (Palmer et al., 1981; Tarr and Pinker, 1989), but with rare exceptions (see 6.3.8), what matters to language are parts and axes, not two-dimensional views.

6.3.8. PERCEPTUAL GROUPING, LAYOUT PLANE, AND PLANE OF VIEW

6.3.8.1. Perceptual grouping.

The visual system has a powerful tendency to group objects in the visual field on the basis of proximity, alignment, and similarity (in size, orientation, color, and so on). So in Figure 6.4a, we see vertical lines – the dots being grouped by proximity; in 6.4b grouping is by color similarity.

Some spatial expressions clearly manifest linear grouping:

the lights of the boat along the horizon
the stepping stones across the river
the houses around the lake

Grouping can also involve two- or three-dimensional regions:

the nest in the tree
the brown sugar in the strawberries
The butler made his way through the guests.

Here the Figure is within the global contour of the branches of the tree, the strawberries, the guests – it is in the volume inside the surface bounding the group of objects or parts constituting the Ground. To make sense of *in* or *through*, we must assume that the Ground is viewed as a volume.

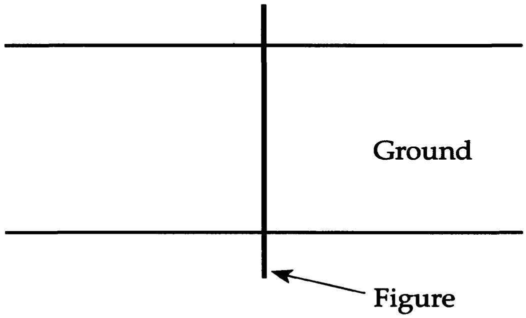

Figure 5. Diagrammatic representation of Talmy's schema for "across"

6.3.8.2. Layout plane

Shepard and Hurwitz (1985, p. 166) say that "the point of contact with the ground provides the best indicator of an object's location." As creatures bound to the ground, we accord great importance to the ground level arrangement of things and spaces. Thus it is natural to represent the location of an object at some height in terms of its projection on the layout plane together with its height. Language yields evidence of this: frequently, the relations expressed are meant to apply in that plane rather than in three-dimensional space.

Here is an example involving *across*. Figure 6.5 is a diagrammatic representation of Talmy's definition of *across*, a "schema" consisting of a ribbonal Ground and a linear Figure.

Now consider

A curtain hung across the room.

The main discrepancies between this example and the schema are that a curtain is a vertical surface – not a line – and a room is a three-dimensional space – not an area, ribbonal or not. But we get a line and an area – and much closer to the schema – if we project the room-curtain configuration onto the floor. Since Figure and Ground are pretty much restricted respectively to lines and areas in uses of *across* (Section 6.4), it is very likely that projection on the layout plane explains the acceptability of this example.

6.3.8.3. Plane of view

Relations are sometimes meant to hold in the two-dimensional plane of view:

The moon is to the right of the tree.

does not locate the moon in the vicinity of the (ascribed) right axis of the tree; instead, the image of the moon in the plane of view is to the right of the image of the tree. Language has access to the two-dimensional representation of a scene in the plane of view, though in the vast majority of cases, what matters are relations in three-dimensional space.

6.3.9. OBJECT GEOMETRY SELECTION

We have seen that schematization is a finely modulated process, dependent on many cognitive and linguistic variables. In effect, the mappings between real objects and their schematic representations can be accounted for by a set of functions, several of which we have considered in some detail. The argument of these functions is the full region occupied by the object, and their value is the region defining the object's schematic representation. Table 6.2 lists some "schematization functions" or "object geometry selection set of possible mappings.[13] So with

Jane walked across the streaming crowd.

three functions apply in succession: first, individual entities are turned into a volume bound by the "contour" of the crowd; second, this volume is projected on the layout plane to an area; third, directionality – the common direction of motion of the crowd's individual members – is assigned to that area. The final value satisfies the selection restrictions of the Ground for one sense of *across* (directionality intrinsic to an area, see 6.4.3), and we can then compute the Figure's path: it is in the layout plane, orthogonal to the area's directionality.

When use of a preposition depends on full definitions of the shape and placement of Figure (or Ground), then the schematization function for that object is the *identity* (this does sound oxymoronic, hence the rather unwieldy coinage: "object geometry selection function").

The linguistic examples in this chapter show that many factors may play a role in determining the applicable schematizations: the preposition, any sentence element, geometric and functional properties of the objects, contextual factors, and so on. From a linguistic point of view, the application of the functions can be a matter of semantics, of pragmatics, or of both. To fit the objects to the geometric types that a preposition selects for, it may be necessary to employ a pragmatic process of *coercion*.[14] For instance, one sense of *across* selects for a line and an area, which can be obtained by projecting the objects on the layout plane – a transformation best described as a pragmatic process, since it may apply with any preposition. In *The ball rolled across the street*, however, describing the path of the ball by focusing on its centroid is only a matter of the meaning of *across*.

Table 6.2. Schematization or object geometry selection functions

1. Idealizations to a
   - point
   - line
   - surface
   - plane
   - ribbon
2. Gestalt processes:
   - linear grouping (yields a two- or three-dimensional linear object)
   - two- and three-dimensional grouping (yields an area or volume)
   - completed enclosure
   - normalized shape
3. Selections of axes and directions:
   - model axis
   - principal reference axis
   - associated frame of reference
   - direction of motion
   - direction of texture
   - direction of maximum slope of surface
4. Projections:
   - projection on layout plane
   - projection on plane of view
5. Part selections (triggered by the high salience of the part):
   - three-dimensional part
   - oriented free top surface
   - base

6.3.10. SCHEMATIZATION AND SPATIAL COGNITION

Relations between objects and their schematic representation can be accounted for by a set of functions, but how are these functions related to spatial cognitive processes? The list in Table 6.2 is heterogeneous in this respect: some functions are specified by spatial cognitive processes (Gestalt processes); others by geometric processes (idealizations and projections); still others by a description of the function's value (all the selections). There is overlap, for example, axis selections are types of idealizations to a line. This heterogeneity appears because we have looked at schematization from different angles. But for every case considered in some depth, we found clear evidence that the functions are grounded in (language-independent) spatial cognitive processes and representations. Sometimes, several spatial representations give rise to a particular mapping; for instance, "idealization to a line" calls on componential axis-based representations for ordinary three-dimensional objects but on cognitive maps for pathways and rivers.

The functions not described in this chapter are similarly grounded in nonlinguistic spatial cognition. For instance, "associated frame of reference," which refers to the assignment of right, front, back, and so on, axes to an object, is a function that invokes frames of reference motivated by nonlinguistic spatial cognition (Herskovits, 1986); the part selections involve spotlighting a (functionally or interactionally) salient part of an object.

The possible geometric representations of an object are thus strongly constrained by spatial cognition. The list of Table 6.2, arrived at by the consideration of a large number of examples, is probably close to complete, and presumably the same functions apply cross-linguistically. But the simplifications of object geometry in locative expressions are not as drastic as commonly assumed: objects are not always represented as points, lines, planes or blobs – at least for any precise (computational) understanding of this statement; many prepositional uses depend on full representations of the shape and placement of one or both of the objects related.

This deepened understanding of schematization has important implications for the interface between language and spatial cognition. Before we consider this question, it will be worthwhile to discuss the fluidity of prepositional meaning.

6.4. The Fluidity of Prepositional Meaning

Prepositional meaning is very fluid. Introspection is not a good guide to this semantic polymorphism: typically, a couple of senses come quickly to mind when one is asked the meaning of a preposition, but actual texts yield a wealth of examples that do not fit the senses accessed in this "zero" context. This section uses the example of one preposition, *across*, to examine whether polysemy and/or prototypicality are useful in accounting for this fluidity and for the salience of particular usage types.

6.4.1. ACROSS: TALMY'S SCHEMA

Talmy (1983) defined *across* as follows:

(F = the Figure object; G = the Ground object)

a. *F* is linear (and generally bounded at both ends).
b. *G* is ribbonal (a two-edged plane).
c. The axis of *F* is (and the axis of *G* is typically, but not necessarily) horizontal.
d. The axes of *F* and *G* are roughly perpendicular.
e. *F* is parallel to the plane of G.
f. *F* is adjacent to (not in) the plane of G.
g. *F*'s length is at least as great as *G*'s width.
h. *F* touches both of *G*'s edges (without this stipulation, the conditions so far would also fit this configuration: | +)
i. Any extension of *F* beyond *G*'s edges is not enormously greater on one side than on the other, nor than the width of *G* itself.

(Figure 6.5 contains a diagrammatic representation of this definition; being a view from above, it does not represent the fact that the Figure is just above the Ground.)

The object geometry selection functions represent one dimension of the flexibility of prepositional meaning. We can view Figure and Ground in different ways beyond their intrinsic geometry. So we can use *across* with nonlinear Figures and nonribbonal Grounds, provided applicable geometry selection functions allow us to obtain a line and a ribbon from them. But in many cases no applicable geometry selection functions will yield a line and a ribbon arranged as in Talmy's schema:

1. The duck swam across the pond.
2. with his hair combed across the top of his skull
3. The snow was drifting across the land.
4. The man started swimming across the current.
5. One could see children skiing across the slope.
6. Ripples spread across the pond.
7. He searched across the city for an apartment.
8. There were clothes strewn across the floor.

A pond (1) need not be ribbonal. It is hard to relate hair on a skull (2) to the schema (and what accounts for the clear intuition that the hair is combed from side to side?). There are no two edges to "the land" in 3; it is unbounded. In 4, the man need not swim from one edge of a body of water to its opposite edge; the sentence only entails a direction orthogonal to the current. A slope (5) is not horizontal nor is it bounded by two opposite edges (and why is motion approximately along contour lines – compare *along the slope*?). In 6, the path of the ripples is not a line, it is two-dimensional; the advancing "front" could be a straight line, or a circle if the ripples expand in all directions from a center. The most salient interpretation of 7 involves a path distributed over the city – the meaning is the same as *all over*. In 8, the Figure consists of points distributed over a surface, it is a *multiplex*.

So *across* has a great diversity of uses not fitting Talmy's schema. I will show that one can account for almost all the examples I have encountered (in the Brown corpus, Kucera and Francis, 1967, and various readings) using ten distinct senses of *across* obtained by cross-classifying five configurations (or schemas) with the three-way distinction between motion, Figure disposition, and vantage point (Figure 6.6). Five of the fifteen combinations are impossible because of incompatibilities; for instance, there cannot be a vantage point sense with a distributed path.

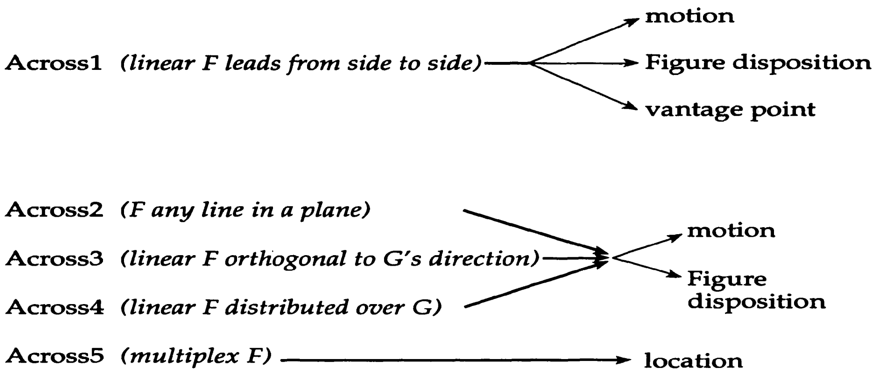

Figure 6. The senses of *across*

6.4.2. ACROSS1

Across1 is a generalization of Talmy's schema. Talmy (1983, p. 235) defines a ribbon as "a plane bounded along two parallel edges," the complementary boundaries are assumed nonexistent, out of sight, or irrelevant. A road or a river provides a good match for such a geometric object, but *across* can be used with an area of any shape as Ground.[15] Talmy clearly intends his definition to be applicable to any closed-contour objects, but he has not spelled out how his schema is to match such cases and how loose the match can be. As there are difficulties involved here, I will first describe the precise geometry of configurations to which Across1 applies, and then consider whether they can be characterized as rough instantiations of the schema. I will call "ribbons" only those things that are unmistakable ribbons – such as pathways and rivers.

For simplicity, I define only the motion sense of Across1:

1. The Ground is a ribbon or any area bounded by a closed line.
2. The Ground is a plane at any orientation.
   See

   A fly was walking ...

   ... across the window.

   ... across the ceiling.

3. The Path leads from one side to the opposite side of the Ground, starting and ending near the Ground's boundary (within, on, or beyond it).

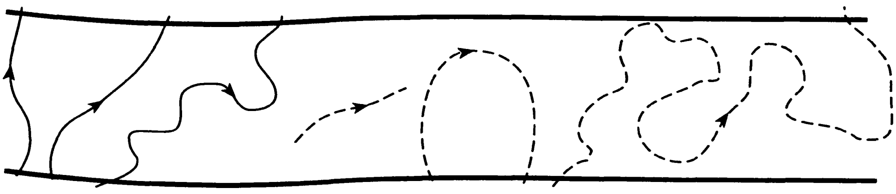

Figure 7. Acceptable and unacceptable paths *across* a ribbon

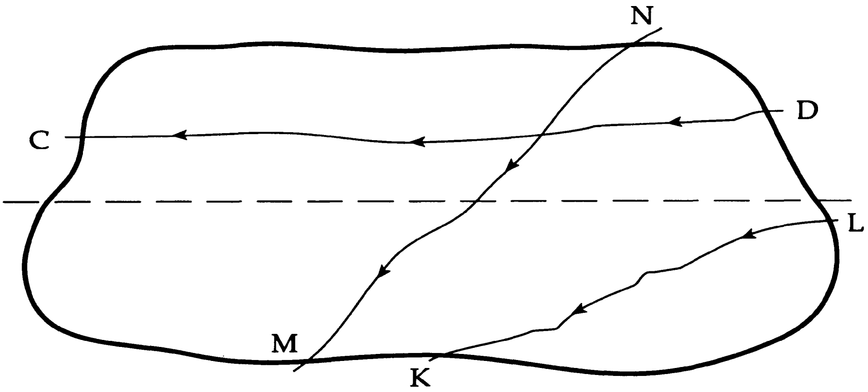

Figure 8. Acceptable paths *across* an elongated shape

Since in *He walked across the room*, the path need not reach the walls of the room, conditions *g* and *h* of Talmy's schema do not hold.

The last condition – that the Path lead from one side to the opposite side of the Ground – requires some elaboration. Take the case of a "true" ribbon (Figure 6.7; continuous lines represent acceptable paths; dotted lines unacceptable ones). An acceptable path must indeed lead from one side to the opposite side of the ribbon, but it need not be straight nor orthogonal to the ribbon's axis. However, wild zigzagging disqualifies the path as an instance of *across* for some speakers.

Consider a nonribbonal Ground (Figure 6.8). The most natural way to divide a shape into two opposite sides is to cut it along its major axis, here an elongation axis. And indeed, MN is an excellent path across; but so are KL and CD, although they do not intersect the major axis.

In fact, the function that discriminates between instances and non-instances of *across* appears to depend on two factors: (1) how good opposites the endpoints of the path are and (2) the straightness of its average heading. A path becomes more acceptable as its endpoints become better opposites and its average heading is closer to a straight line. If the end-points are very good opposites, the path can go off-track; if not, the

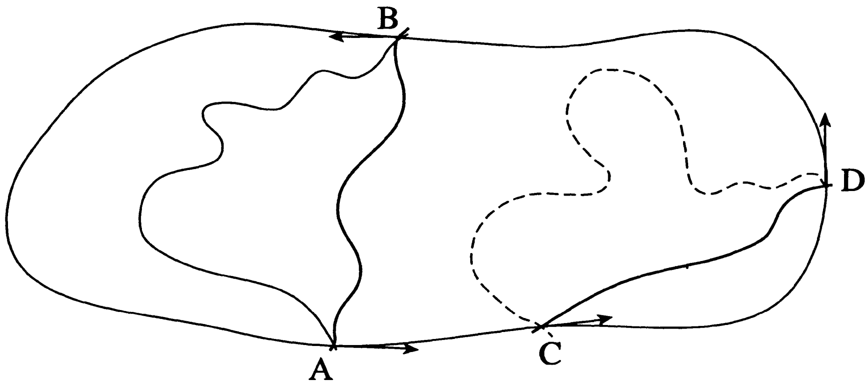

Figure 9. Acceptable and unacceptable paths *across* an elongated shape

path must go straighter to the goal. So A and B (Figure 6.9) are very good opposites and the circuitous path from A to B is acceptable. C and D are not such good opposites, so an "across" path joining them must go more or less straight: the curving C to D path, though quite similar to the acceptable A to B path, is not acceptable. Finally, A and C are very bad opposites; no "across" path can join them.

What is a measure of how "good opposites" two points on a boundary are? Assuming a smoothed out (no small irregularities) and smoothly curving (no deep concavities or singularities) contour, the angle between the tangents at the points provides such a measure: if the angle of the tangents is 180 degrees (when all tangents are oriented counterclockwise, say), the points are perfect opposites (Figure 6.9); as the angle decreases, the points become worse opposites until 0 degrees, where the points are not opposites at all. Subjects asked to rate different paths across an ellipse (Herskovits et al., 1996) confirmed the validity of this measure.

Note that the measure defining the goodness of across paths is a continuous function of the two factors mentioned (the straightness of the path heading and the opposition of the path's endpoints). To decide whether or not to use *across*, we need a threshold, and there will naturally be cases where no clear-cut decision is possible.

This measure works well for all shapes with a smoothly varying curvature and no deep concavities. If there are singularities in the curvature (vertices), judgments may be "bistable": for instance, a path joining adjacent vertices in a rectangle might look good if viewed as linking opposite sides, but unacceptable if viewed as linking adjacent sides. For shapes with deep concavities and several lobes, judgments become uncertain and unstable. The difficulty is analogous to that found when trying to assign width and length to such shapes. Clearly, parallelepipedic and rectangular shapes play a central role in our conceptual system. Systems of dimensional adjectives (*high, deep, wide, long*) are based on assuming objects not too dissimilar to these.

In any case, division of the Ground shape into opposite sides, very flexibly interpreted, is the principal idea underlying Across1. This flexibility, however, disappears if geometric or interactional properties produce a salient division of the Ground; then, the Path or Figure must lead from one side of the dividing axis to the other and be loosely orthogonal to that axis, as the following illustrations show:

- *Elongation:* If there is great disparity between the Ground's width and length, then Figure or Path must be orthogonal to the Ground's long axis:

  She walked across the vegetable row
  The road leads across the ridge.

- *Canonical use:* If the Ground is primarily used for travel along a given direction, then the Path must be orthogonal to it:

  Shė walked across the pier.

  Similarly, there is a canonical way to sit on a saddle and lie on a bed: *across* in

  to sit across the saddle
  He was lying across the bed.

  is orthogonal to that canonical direction.

- *Symmetry:*

  She laid the stick across her lap.
  with his hair combed across the top of his skull

  The stick is orthogonal to the symmetry plane of the lap, and the hair combed orthogonally to the symmetry plane of the head.

- *Verticality:* A vertical plane could be generated by the sweep of an horizontal or a vertical line, but we think of it as made up of verticals; we conceptualize it as having a vertical intrinsic orientation. So across a vertical plane is (loosely) horizontal:

  He drew a line across the blackboard.

Does the "goodness" measure of *across* paths measure goodness of fit of Talmy's schema to the Path-Ground configuration? We will return to this question in 6.4.6.

6.4.3. THE OTHER SENSES OF "ACROSS"

6.4.3.1. Across2

1. The Ground is an unbounded plane surface.
2. The Ground has no intrinsic directionality.
3. The Path includes salient segments of straight translation.

Across2 is illustrated by

They walked across the sand for hours.
We followed a track across the grass.

where the limits of the sandy area and of the grass are irrelevant. The third condition is meant to explain the following contrasts:

An ant wandered across the ground.
She was pacing across the floor.
* The camel walked in a circle across the sand.
* He drew a circle across the sand.

One can wander and pace across a surface but not draw a circle or walk in a circle across. The trajectory need not be a single straight line but a circular trajectory is not acceptable. Across2 may be used with a clearly bounded object when those boundaries are deemed irrelevant; the object is then viewed simply as a surface. So in

She pushed the ashtray across the desk.

the path of the ashtray need not lead from one edge to another – the top of the desk is viewed as a surface. This alternation between a two-dimensional bounded entity and an unbounded surface appears in a surprising example found in the Brown corpus (Kucera and Francis, 1967):

We did 80 miles an hour across a hard dirt road to a cluster of shacks.

Although the ribbonal geometry of a road appears almost inescapable, it is not impossible to foreground the road as surface; Across2, then, applies instead of Across1. This switch is facilitated by the modifiers *hard* and *dirt*, which bring the road surface into focus.

6.4.3.2. Across3

1. The Ground is an unbounded plane surface.
2. The Ground has an intrinsic directionality.
3. The Path is loosely orthogonal to that directionality.

The Ground may be seen as bounded by edges under certain conditions, yet there is no implication that the Path runs from edge to edge. Intrinsic directionality can arise through

– *Motion*:

cutting across the streaming crowd
swimming across the flow

– *Texture*:

cutting across the grain

– *Inclination*:

skiing across the slope

The intrinsic directionality of a slope is defined by its lines of maximal incline, so motion across a slope is orthogonal to these – along contour lines.

6.4.3.3. Across4

1. The Ground is a bounded plane area.
2. The Path is distributed over the Ground.

One interpretation of

For a whole year, I traveled across India.

involves a path distributed over India.

6.4.3.5. Across5

1. The Ground is a bounded plane area.
2. The Figure is a set of "points" distributed over the Ground.

Across5 applies in

Across the United States, people are listening to the President.
the sprinkle of freckles across her face

6.4.4. RELATIONS BETWEEN SENSES

We can discern relations among the ten senses of across. The senses can be arranged in a network of the kind used by Brugman (1988) and Lindner (1981) to describe prepositional polysemy. Part of the network is represented in Figure 6.10; with the exception of Across4, only motion senses are included.

A link in the network represents close semantic similarity; the two nodes differ minimally in meaning. For instance, Across1 and Across2 entail a punctual Figure moving on a plane; but while in Across1, the path must run between opposite edges of a bounded Ground, with Across2, with the Ground being simply a plane surface, there is no constraint on the position of the path's end-points. An interesting way to think of the link between the two senses is to imagine viewing a case of Across1, and zooming in until the boundaries of the Ground disappear from view; the new configuration will match Across2.

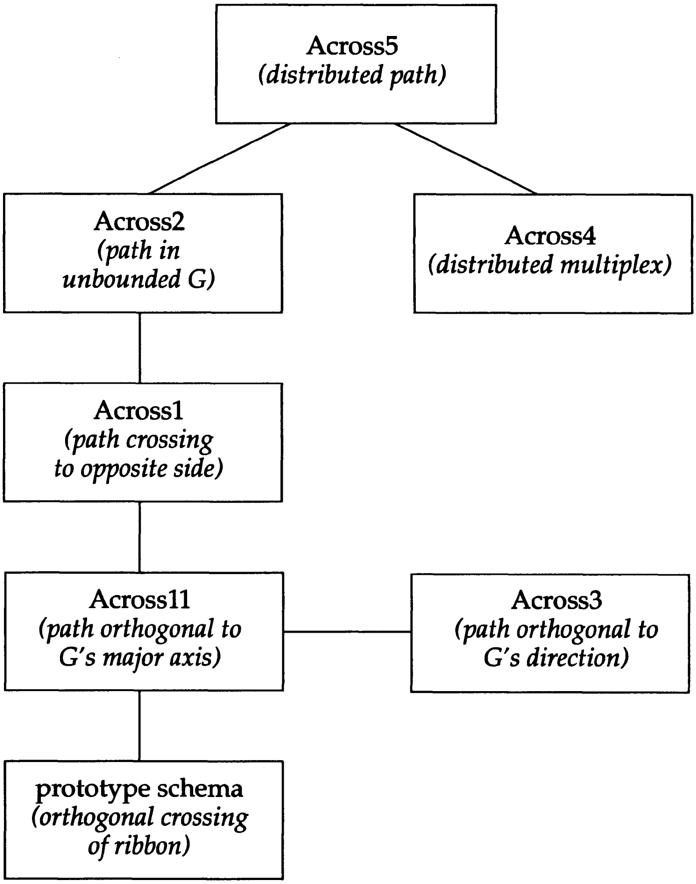

Figure 10. Polysemy network for *across*

Across11 is a specialization of Across1, in which the Path is orthogonal to the major axis of the Ground; it provides a bridge between Across1 and Across3 on one hand, and Across1 and Talmy's schema on the other. Thus Across1 and the schema are separated by two arcs, each representing close similarity; and so are Across1 and Across3. Across5 has one entailment added to Across2 – the path must cover the Ground. Finally, Across4 and Across5 share the feature of distribution over an area. The category is made up of senses related by similarity, but in network fashion, so several "similarity steps" may separate two senses.

The polysemy of *across* parallels that of *over*:

The power line stretches over/across the yard.
He walked carefully across/over the ice.
He traveled across/(all over) India.

His clothes were strewn across/over the floor.

The fact that two prepositions have the same four senses is strong evidence that some natural conceptual plasticity has led to cocategorizing these four geometric patterns. The relations described are plausible explanations for the tendency of these notions to cluster under the same word.

Family resemblance thus makes the different senses cohere. But it does not follow that the entire extension of the category "across" is predictable from a prototype. This is family resemblance but conventionalized. These links need not play a role in comprehension and production; few speakers may even ever be aware of them. Instead, their mental lexicon includes the conventional knowledge needed to use and comprehend all these uses, namely associations of *across* with the various interpretations.

Yet the relations between senses must have been active for some speakers at some time, otherwise there would be no way to explain how these distinct uses came to be expressed by the same word. In other words, I am suggesting that the proper level of understanding for these links is diachronic and statistical: extensions in language change may frequently follow along such links; and a pair of senses tied by a single link will often be expressed by the same word across many languages. Both assumptions, of course, would need to be tested.

6.4.5. IS TALMY'S SCHEMA A PROTOTYPE?

Talmy's schema can be considered a prototype for the category in the sense that it is saliently associated with the form *across* and most frequently accessed in a neutral context. So when subjects were asked to draw a diagrammatic representation of *across* and list five sentences with *across*, half the diagrammatic representations looked like the schema and 39 percent of the sentences involved a ribbon as Ground (Herskovits et al., 1996). But it is clear that membership in the category "across" is not in terms of similarity with Talmy's schema. The similarity between Across5 and the schema is minimal, yet *clothes strewn across the room* is a good example of *across*.

Is the schema a prototype for Across1 in a stronger sense – that is, is the measure of goodness described above (6.4.2) a measure of how similar instances of Across1 are to the schema? Deviations from some schema characteristics (horizontality of the Ground, Path/Figure slightly above the Ground or extending beyond the Ground's edges) do not lead to "worse" instances of Across1:

He drew a line across the blackboard.
the ditch across the road
the curtain across the room

are as good as

the pedestrian overpass across the freeway.

The lexical entry speakers use to judge these examples must be a generalization of Talmy's schema: it is abstracted from the orientation of the Ground plane; from the Path/Figure's precise position within close distance of that plane; and from the precise position of the endpoints of the Path/Figure within close distance of the Ground's boundary. It is not impossible that a process of matching this generalized schema to the situation at hand underlies the measure described, but we would need a better understanding of similarity to decide that point. As the shape of the path can deviate greatly from a straight line, and the edges of the Ground can be far from parallel, the relevant similarity is probably based on function (defined in terms of reaching the other side of a region) rather than perception.

Barsalou (1985) studied variation in typicality among category members. He showed that such "graded structure" depends on (1) the member's similarity to the central tendency of the category, (2) its similarity to the ideal of the category, (3) its frequency of instantiation. The prototype of the category is the member with the highest typicality rating but it need play no role in membership decisions.

The central tendency of Across1 is similar to Talmy's schema, only more abstract, along the three dimensions mentioned above. The ideal of the category would have the symmetry of the schema, the parallelism of the Ground's edges, and also a path extending beyond the Ground's edges, since this "really" takes you across the Ground. And *across* is probably most frequently instantiated by cases of motion on horizontal ground. Thus the schema scores highest on all three counts – no wonder it is the "best" example of Across1. But membership judgments are unlikely to be based on the evaluation of similarity to this "prototype."

6.4.6. CONCLUSIONS

We were able to classify all usages of *across* (excluding some semi-idiomatic forms) by means of ten "related" senses; the relations bring the senses together in a family resemblance network. But what of the mental lexicon itself? What form does the lexical knowledge supporting the comprehension and production of phrases with *across* take? This semantic analysis does not allow precise answers. It is probable that there are many more "entries" than these ten senses, that multiple, more specific "subsenses" have entries of their own, even though they are redundant with the more abstract senses that subsume them (Langacker, 1991). There is even evidence that we use different entries to judge pictures and sentences instantiating the same sense (Herskovits et al., 1996). The precise form of lexical knowledge remains an enduring puzzle, as it is embedded in a complex system of interactions between context-dependent access properties, syntactic and semantic constructional processes, and memorized form-meaning units at various levels of specificity.

6.5. The Interface Between Language and Spatial Cognition

What can we deduce from this detailed study of schematization and category structure concerning (1) the spatial representations accessed by language, and (2) the computations in the interface between language and spatial cognition?

6.5.1. THE SPATIAL REPRESENTATIONS ACCESSED BY LANGUAGE

We found evidence that language may access the following spatial representations:

1. Cognitive maps (large-scale environments),
2. Representations of observable scenes with objects as points (at least in some regions),
3. Representations of motion with objects as points,
4. Representations of observed scenes with objects shapes fully specified,
5. Representations of layout in the ground plane in small-scale environments,
6. Gestalt groupings of objects,
7. Axis-based componential representations of objects, and
8. Representation of plane of view.

These are probably not all distinct. So the first three – in which objects are represented as points – may be integrated into a single seamless representation, but close-up representations of observed scenes (4) with detailed representations of shape ought to be distinguished from it. Layout maps (5), Gestalt groupings (6), and axis-based componential representations of objects (7) could be characterized as aspects of representations of observed scenes (4), but they do provide alternate geometric representations of the objects. Finally, language has access to a two-dimensional representation of the plane of view (8),[16] a representation clearly distinct from the others.

Some authors (such as Bryant, 1993; Landau and Jackendoff, 1993) hypothesize that a single system of spatial representation underlies all spatial activities, as well as language. But this assumption is not confirmed by a careful analysis of schematization. It is true the various representations mentioned are tied together as facets of the "real world"; they are like snapshots taken of the three-dimensional veridical map we believe we sample. We do take a point in a cognitive map and a view of the corresponding landmark to be manifestations of the same object – and, under all circumstances, the object is what we mean to refer to. Yet language calls upon different geometric representations of objects, which are sometimes components of distinct nonlinguistic representations of space.

6.5.2. LANDAU AND JACKENDOFF'S HYPOTHESIS

Landau and Jackendoff (1993) argue for a dichotomy in the expressive power of language: we use detailed geometric properties of the objects when naming them (with nouns); but coarse representations – as points, lines (axial structure), and blobs – when locating them (with the help of a preposition). They suggest that this reflects a division in the way the brain represents spatial information. Assuming object identification and object localization are performed by separate neural subsystems – the "what" and "where" systems (Ungerleider and Mishkin, 1982) – they attribute the contrast between nouns and prepositions to a parallel modular division of the language system: the preposition system accesses only the encoding produced by the "where" system; the noun system only the encoding produced by the "what" system. To quote them directly (Landau and Jackendoff, 1993, p. 257):

> One kind of object description gets through the interface between spatial representations and language for naming (the "whats"), and another kind of object descriptions does so for locating (the "wheres").

The difference between nouns and prepositions would then follow: the "what" system represents fine distinctions of shape; the "where" system represents objects only as place-markers of roughly specified shapes. (The markers are cross-indexed with object representations in the "what" system, so we can know which "what" is at a given "where.")

We have seen that their basic premise is invalid: objects referred to in the context of a preposition are not necessarily represented as points, lines, or blobs; one may need to know their shape and precise placement to decide upon the applicability of a preposition. Landau and Jackendoff confuse the selection restrictions of the preposition with the knowledge of object shape needed for their appropriate use.

As for the interface between language and spatial representations, they skirt the following central question: does it compute abstract spatial relations on coordinate representations produced by the "where" system, or do the encodings produced by the "where" system include all abstract spatial relations necessary to linguistic expression? Either way, the hypothesis cannot hold: If the "where" system computes all relational primitives necessary to linguistic expression, the interface need only focus attention on the relevant combination of primitives and associate it with the right morpheme. Sometimes (as with some uses of *in*), the "where" system must operate with full knowledge of object shape; then, reduction of objects to points or blobs is not needed at any point in the processes leading to the selection of the preposition. If the interface computes the relations, then it must sometimes have access to objects'shapes.

6.5.3. THE COMPUTATION OF LINGUISTIC SPATIAL RELATIONS

We certainly perceive the relations we express. We see that a cup is *on* a table, a chair *to the right of* a door, someone sits *across* a desk. This implies that the relations have a visual (nonsymbolic) representation. But are such abstract relations part of the visual representations constructed in the absence of linguistic goals, or are they (at least sometimes) computed on-line, upon some "command" of the linguistic system? Visual processing may provide different information depending on attention. For instance, a brief look at a scene only yields a sense of its gist – the setting, some awareness of its characteristic objects. Attentional processing of limited regions will lead to more specific shape perception and object recognition. Some elements of information are made available only through sustained attention or scrutiny (Julesz, 1980). Finally, practice improves subjects' abilities to discriminate and note certain characteristics of visual stimuli.

In other words, it is plausible that the perception of certain spatial relations requires special visual processing. It will be useful at this point to look at the relevant work in psychology and computer vision.

6.5.3.1. Abstract spatial relations in vision.

Abstract spatial relations could play a role in vision as part of the representation of objects or as part of the representation of location. I consider first object representation.

In componential theories of object recognition (Marr and Nishihara, 1978; Biederman, 1987), objects are represented as assemblies of parts. The parts' boundaries are determined by geometric features of the objects' surface (Hoffman and Richards, 1985a). The parts themselves are represented by shape primitives, such as generalized cones (Marr and Nishihara, 1978), or "geons" (Biederman, 1987). There is very little specific about the encoding of relations between parts, other than the suggestion that abstract spatial relations would solve the problem nicely, since the resulting representations for articulated objects would be stable (Marr and Nishihara, 1978; Kosslyn et al., 1989). For instance, in any position, the arm is connected to the shoulder – *connected* is an invariant under normal motion. However, fully abstract relations will not in general suffice: for instance, we know that an arm can move only within a certain solid angle, and we can immediately recognize an arm at an odd angle. So *connected* by itself will not do; some measure of angle must be added.

Biederman (1987) proposes a set of spatial relations to represent the connections between geons:

- For any pair of connected geons, whether one is greater than, equal to, or smaller than the other;
- Whether a geon is above, below, or at the same height as another;
- Whether two geons are joined end-to-end, end-to-side and centered, or end-to-side and uncentered;
- Whether a geon is joined to another's "long" or "short" surface.

This is a very short catalogue and unlikely to account for all needed distinctions between shapes, but the relevant point here is that, except for the second, these relations do not appear fundamental to language. They are not lexicalized in English or even easy to express (see the fourth), and it seems unlikely that any language lexicalizes them – although English lexicalizes a great many other relations, as do most other languages.

Ullman (1985, p. 98) suggests that "objects are often defined visually by abstract shape properties and spatial relations among components." He illustrates this with a representation of a face in which the internal features have been shuffled; the resulting picture is not recognizable as a face. Yet, it is unclear that a set of abstract spatial relations of the kind used in language is what brings features together into a face. It is the simultaneous satisfaction of many spatial relations – at least some of which must be in terms of metric parameters – that makes a face into a face, and this simultaneity might be better captured by a pattern-matching process than some articulated (proposition-like) representation. In short, there is no strong support for the assumption (expressed in Kosslyn et al., 1989 and Hayward and Tarr, 1995) that linguistic spatial relations originate in those used in object representations.

Abstract spatial relations may also play a role in location representation. But researchers (such as Shepard and Hurwitz, 1985; Pinker, 1985; Sedgewick and Levy, 1986; Hinton and Parsons, 1988) have usually assumed that location is represented by means of metric coordinates (or some equivalent) within three types of reference frames:

1. Egocentric, based on the top, bottom, front, back, right, and left of the perceiver;
2. Allocentric, anchored on prominent landmarks in the environment (for example, the walls of a room);
3. Object-based.

Coordinate representations yield the precision evident in visual representations. Moreover, any abstract relationship can be computed from coordinate representations.

So there is little substantive evidence that visual representations include explicit representations of the spatial relations we express – except when we are specifically prompted to compute such relations. Yet the question is certainly not settled. It looks intuitively plausible, for instance, that simple spatial relations between proximate objects, such as contiguity, inclusion, or "immediately above," would be explicitly encoded in visual representations independently of higher-level cognitive activities.

6.5.3.2. Perceiving for speaking.

There are definitely cases where some "extra" visual processing – beyond a simple act of attention – is needed to extract a spatial relation from visual representations. I will give five arguments to support the proposition that visual processing beyond that accompanying the attentional, but not goal-directed, apprehension of a scene is needed to compute and perceive at least some linguistic spatial relations (see Logan, 1994, for supporting psychological evidence). The arguments are based on careful examination of the processing needed to compute a given relation, and on introspective evidence that such processing takes application, effort, and time.

Assume a scene of moderate scope (such as a desk top supporting some objects); in looking at it at leisure but with no particular goal in mind, we construct a visual representation.

1. It is implausible that all the relational primitives needed to ascertain any lexically expressible relation would be encoded in that representation. There would have to be one or more primitives for every pair of objects – not only proximate objects but distant ones. It is highly unlikely that we have explicit representations of every expressible relation between distant objects. If asked to describe the relationship between one book among several between book ends at one corner of the desk and a cup at a diagonal corner I may produce *The book is to the left of the cup.* But that requires "configuring" the book and cup together in an act of attention[17] that ignores the objects in between, and then evaluating the degree of fit of the configuration with *to the left.* There is no reason to perform these operations absent a linguistic goal – and clear introspective evidence that it requires a special effort.

2. The evaluation of some spatial relations requires configuring *virtual structures* with visible ones. These processes are sometimes clearly language dependent.

Assume a speaker looking obliquely at the TV in Figure 6.11a who says *The basket is to the right of the TV.* Ascertaining the truth of the relation required her to establish which side of the TV is its canonical front; imagine the symmetry axis leading from the center of the TV toward the

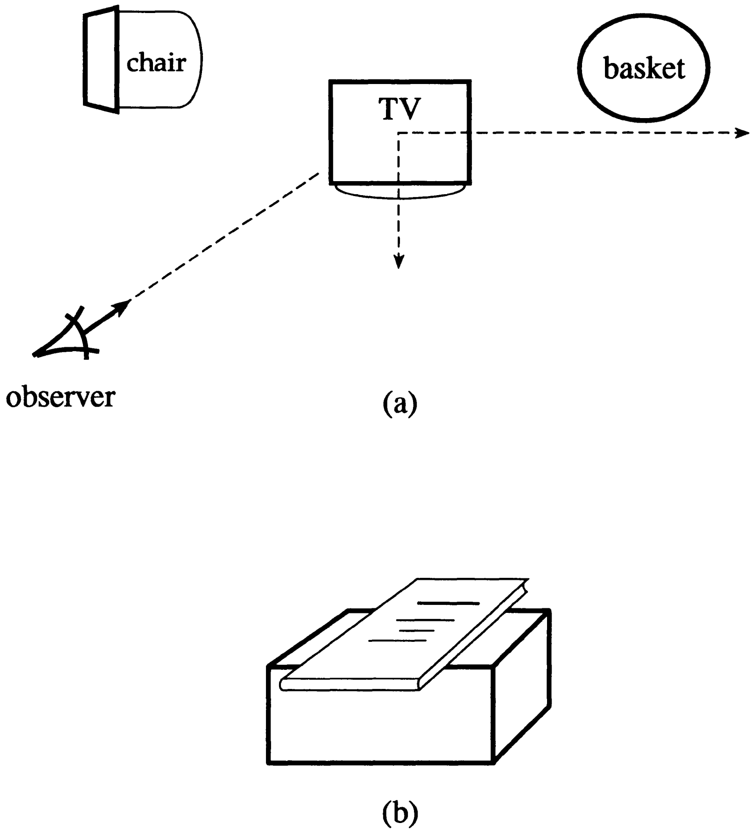

Figure 11. Perceiving spatial relations

front; imagine the line orthogonal to that axis running through the center; choose the appropriate half of this orthogonal; and ascertain the presence of the basket within a cone centered on that half-axis (Gapp, 1995). While egocentric right and left are relatively easy to access,[18] the right and left directions intrinsic to an object rotate with it; keeping track of the divisions of surrounding space induced by these directions is unlikely to be done at all times by automatic vision processes.

3. In describing a configuration of two objects, we can often express either of two converse relationships, provided the objects' size and mobility are not disproportionate. It is unlikely that these converse relationships are both encoded in visual representations. The perception of non-symmetric spatial relations must be anchored on one of the objects. One might perceive two

objects as *near each other*, but it is hard to imagine how one would perceive simultaneously that *The chair is to the left of the TV* and *The TV is in front of the chair* (Figure 6.11a).

4. Categorizing a configuration by a spatial relation may require approximate fitting and yield uncertain results; this is a high-level operation, not the kind performed by undirected, unintentional vision processes.

Several lexically expressible relations can generally categorize a given configuration of two objects. So in Figure 11b, one could say any of the following:

The book is ...

...on the box
...on top of the box.
...on the top of the box.
...over the box.
...across the box.

The box is under the book.

The choice between these is a matter of what is most salient and/or relevant in the context. For the first three expressions, the only relational primitives that need be accessed are contiguity and support – relations so basic they may well be encoded in "minimal" visual representations. But consider *over* and *across*: *over* highlights the fact that the book *almost* covers the box; *across* the fact that the long (top-bottom) axis of the book is *almost* orthogonal to the long axis of the box. Prepositional categories, like many linguistic categories, are fuzzy. Why perform such an act of categorization, an act involving approximate fitting, in the absence of a high-level cognitive goal? The complex computations underlying some uses of *across* (6.4.2) certainly support this point.

5. The basic purpose of the motion prepositions is to describe navigation paths. Their use in describing static configurations in visual scenes is derivative. It is improbable that we would "see" relations like *across* or *along* in a visual scene, were it not that language makes them available, prompting us to look for the patterns that define them (between and within objects; see *the carving across the handle of the knife*).

To compute a spatial relation between two objects, one must:

1. Configure them together – that is, select them for attention in a way that makes it possible to apprehend the applicable spatial relations; and
2. Categorize the configuration.

Each of these steps may require visual processing beyond that involved in observing the scene in undirected fashion.

Thus, we can look at a scene, even focus on two particular objects, and not see the spatial relations between them that we later express. This means there are *language-induced percepts*. Slobin (1996) gives evidence of "thinking for speaking," a special kind of thinking used on line in the act of speaking, which is evanescent and does not affect the way we think outside the act of speaking.[19] There must also be "perceiving for speaking." I do not doubt that speakers of Dyirbal and of Thai see essentially the same thing when they simply look at a scene; the basic processes of vision are certainly universal. But using language involves a kind of "visual cognition" beyond these basic processes. Here, perceptual and conceptual operations cease to be clearly distinguishable (Talmy, 1996), and language prompts the visual system to constructive operations that are not a necessary part of undirected perception.

Ullman (1985) addresses the question of the computation of abstract spatial relations from coordinate representations, and proposes a set of operators that can be assembled to construct "visual routines." I am not concerned here with the details of his architecture, but with the design constraints he posits: visual routines would not apply throughout the image but operate on selected locations of particular interest when triggered by high-level (possibly linguistic) goals. They would constitute a task-specific vision system. This system would allow the computation of an open-ended variety of abstract predicates – thus enabling perceptual learning. Ullman bases these design constraints on the requirements of object recognition, but they are clearly useful for the extraction and perception of the abstract relations used in language[20] and spatial problem-solving.

6.6. Conclusions

We have examined the semantics of English prepositions, the schematization processes involved in their use, and the forms of nonlinguistic spatial cognition underlying these processes. This exploration has taken us a step closer to understanding the interface between language and spatial cognition, showing that the flexible choice of object representation employed in descriptions of location is tied to access to various nonlinguistic spatial representations; and that the expression of at least some spatial relations requires processes of configuration and categorization beyond those performed by automatic visual perception.

Rather than a closed system with limited access to nonlinguistic representations, language seems to be flexibly connected to a variety of spatial representations. This does not mean that spatial cognition puts no restriction on the spatial relations languages can express: we brought out one such restriction relating to the perception and representation of motion. Also, given the importance of cognitive maps in spatial cognition, it is likely that every system of spatial relation terms will have a significant subset used primarily with objects represented as points. But the conceptual system is not restricted to these. The visual system has the ability to compute an extended variety of linguistic relations, requiring a flexible computational process controlled by linguistic goals.

This study of schematization and of the structure of prepositional categories also provides a first sketch of the complex way spatial linguistic categories are grounded in perception. The problem of object recognition has driven much of the research in computer vision in the last thirty years; other, less transparent connections between vision and conception have been neglected – those having to do with relations, actions, and events. The basic categories that the visual system must compute to identify these are not as clearly apparent as Rosch's (1977) "basic-level" object categories. They need to be brought out by fine-grained semantic analysis and psycholinguistic studies. This chapter sets the stage for such developments.

Notes

[1] Activities are one of the "aspectual classes"; the others are states, achievements, and accomplishments (Vendler, 1967). An activity extends over time (contrary to achievements, such as *reach*, which are punctual events), but does not specify a completion point (in contrast with accomplishments, such as *cross* in *Jo crossed the road*).

[2] In Jackendoff's conceptual structure (Jackendoff, 1983; 1990), phrases such as *from the bridge* are said to refer to trajectories. The motion prepositions are Path-functions that map a reference object onto a trajectory. It is actually impossible to define such functions; given a preposition, there is no way to assign a unique trajectory to every given reference object.

[3] Examples preceded by ♯ are acceptable, but do not have the intended intepretation. So *The snake lay up the tree trunk* is acceptable but with the whole snake located toward the top of the tree trunk.

[4] The sentence is acceptable with the entire snake located past the stone. These examples are adapted from Talmy (1983).

[5] The difference in acceptability may be due to Figure and Ground being body parts, which frequently are treated differently from other kinds of objects in spatial sentences (Herskovits, 1986).

[6] Again, Talmy (1983) uses very similar examples. See also Talmy (1996) on fictive paths.

[7] Sadalla et al. (1980) found that some locations in cognitive maps are *anchors*; other places are seen in relation to them. One of the facts associated with the role of anchor is that subjective distance is asymmetric; subjects judge the distance from A to B longer than the distance from B to A, if A is an anchor and B is not.

[8] *Inness* is certainly often inferred rather than directly assessed; the location of every point of the Figure need not always be checked. So ascertaining that an object is *in a room* often only requires making sure it is visible.

[9] Rock (1972, p. 671) defines perception "to mean what is 'noted,' 'described,' attended to, or apprehended about a figure, albeit unsconsciously and nonverbally." There can be awareness without perception. Experiments show that perception, so defined, is necessary for memory formation. It must also be a condition for the formation of conceptual categories dividing the range of shapes (or motions) considered.

[10] One sense of *at* entails a canonical interaction between Figure and Ground:

Jane is at the desk.

This sense can be extended to chairs but not to objects playing no role in the canonical interaction:

The chair/*vase is at the desk

[11] "A large-scale environment is one whose structure is revealed by integrating local observations over time" (Kuipers, 1983, p. 347).

[12] Marr's 3-D model (1982) is hierarchical: the whole shape is divided into its "immediate" parts, which are in turn divided into parts, and so on down. The location of parts is represented only with respect to the entity immediately above in the hierarchy, which allows for stability in the representation of articulated objects: the representation of a finger will be with respect to a frame of reference attached to the hand, not to the whole body. The entire shape and each part have a model axis, which gives coarse information about length and orientation.

[13] The list here is a revised and augmented version of a similar list in Herskovits (1986) where examples of application of the functions not considered in this chapter can be found. See also Section 6.5 for additional illustrations.

[14] Hays (1987) uses the term *coercion*, from programming language theory, to indicate the "forced" matching of the argument(s) of a linguistic predicate to its selection restrictions. It is always associated with metonymy (Herskovits, 1986), since the actual arguments of the predicate are geometric constructs distinct from the primary referents of the complement noun phrases.

[15] I will, for conciseness, talk about Figure and Ground in what follows, when actually meaning "coerced Figure" and "coerced Ground" – that is, the values of the applicable geometry selection functions (the actual arguments of the relation *across*).

[16] Levinson (1994) makes a similar point, using examples from Tzeltal.

[17] Niyogi (1995) proposes a model of the computation of spatial relations in which the location of the focus of attention itself serves as input to the "daemons" carrying out the computation; configuring, then, would involve moving the focus from one object to the other.

[18] Egocentric relations are stable under projective transformations; so, from a given vantage point, right and left in three-dimensional space always correspond to right and left in the plane of view. We can easily judge whether two objects are to the right of another; they both appear on the same (right) side of it. By contrast, two objects to the intrinsic right of a TV could project in the plane of view right and left of the TV, given that the Figures are not required to be exactly on the axes.

[19] But Levinson's study of Guugu Yimidhirr (1993) shows that not all spatial thought supporting language use is without consequences for spatial thought outside language.

[20] Chapman (1991) and Niyogi (1995) use visual routines in artificial intelligence models of linguistic abilities; they argue against vision models that involve processing an entire retinal image, and assume that relational knowledge is computed only when needed by higher-level cognitive processes.