

Annette Herskovits  
Linguistics Department, Stanford University

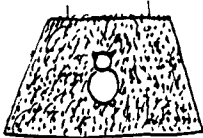
At first glance, the spatial uses of prepositions seem to constitute a good semantic domain for a computational approach. One expects such uses will refer more or less strictly to a closed, explicit, and precise chunk of world knowledge. Such an attitude is expressed in the following statement:

"Given descriptions of the shape of two objects, given their location (for example, by means or coordinates in some system of reference), and, in some cases, the location of an observer, one can select an appropriate preposition."

This paper shows the fallacy of this claim. It addresses the problem of interpreting and generating "locative predications" (expressions made up of two noun-phrases governed by a preposition used spatially). It identifies and describes a number of object characteristics beyond shape (section 1) and contextual factors (section 2) which bear on these processes. Drawing on these descriptions, the third section proposes core meanings for two categories of prepositions, and describes some of the transformations these core meanings are subject to in context. The last section outlines the main directions of inquiry suggested by the examples and observations in the paper.

### 1. OBJECT CHARACTERISTICS

Throughout the paper, I use the term "object", meaning, strictly speaking, the object together with some lexical label. In effect, the choice of preposition depends on the lexical category associated with the object by the noun-phrase used to refer to it. And such a category is not uniquely defined. There are different levels in the categorization hierarchy (e.g. "end table", "table", "piece of furniture"), but also different perspectives on the object. Consider the picture below.



That patch of grass could be referred to alternately as a *front-yard*, a *lawn*, *grass*, a *patch of grass*, etc. (to assume that these phrases refer to the same object, one must see the *grass* as a metonymic substitute for *this patch of grass*, and the *front-yard* as some "area" rather than a "slice" including air above and ground under; neither view is unreasonable). The permissible prepositions, and their interpretation, vary with each referring phrase: compare *in/on the grass*, *in/on the patch of grass*, *in/(on) the front-yard*, *on/(in) the lawn* [Fillmore 1971]. With this warning, I will go on speaking of "object characteristics", "object identity", etc.

Some of the object characteristics used in production and interpretation can be computed from the shape of the objects -- the axes of symmetry (needed for *across the road* and *along the road*), the "top surface" (*on the table*), the "outline" (*the bird in the tree*), etc. (for a description of some of these characteristics, and of their role in

comprehension, see [Bogges 1979]).

Other characteristics are not deducible from shape. These include:

#### 1.1. ALTERNATE GEOMETRIC DESCRIPTIONS

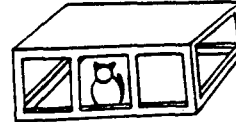
Objects identical in shape may be "conceived" differently, for instance as surface or as enclosure. This may be a choice available to the speaker to emphasize certain aspects (*in/on the rug*), or it may be determined for the category of the reference object (*on the football field*).

In *under the water*, the water stands for the upper free surface of the water; in *in the water*, it is conceived as a volume. A whole category of objects follows this rule: see (*under/in*) *the (snow/lake/ocean/sand)*. Such objects tend to be viewed only as volumes with "underneath": *underneath the lake* is generally interpreted as meaning "under the lower surface of that body of water".

In *the crack in the bowl*, the crack is in the volume defined by the normal surface of the bowl in its uncracked state. In *the milk in the bowl*, the milk is in the volume enclosed by the bowl and limited upward by a plane through the rim.

#### 1.2. FUNCTION

One says *in the dish* and *on the tray* though these objects may be essentially identical in shape. One will not ordinarily say *the cat is in the table*, but *under the table*, even with the cage-like table below.



*John is at X* often means that John is using X as one normally uses it (*John is at his desk*). If normal use implies being *on* or *in* X, then *at* is not used (*John is in or on the bed*, but not *at*). And to the right of the chair is defined by reference to a typical user of the chair.

#### 1.3. TYPICAL PHYSICAL OR GEOMETRICAL CONTEXT

When using *in* with areas, it is not sufficient that the reference object be two-dimensional; that object must be part of a surface divided into cells. One does not draw a line *in a blackboard*; but *in the margin* is acceptable, because the margin is a subdivision of a page. In the same fashion, geographical areas (*England, the county*, etc.) are sections of a divided surface. Some objects are exclusively conceptualized as parts of a "cell structure" and cannot then follow *at* (*at his room, at England*). Other objects can be conceptualized both as elements of a cell structure (*in the village*), or as one of a set of separate places (*at the village*). Or consider *yard*: when it is a part of the grounds of a house, one is restricted to *in*. But of somebody working in a junkyard, one could say *he is at the yard*, reflecting a view of the yard as one of a set of separate locations.

If a door is in its typical context, i.e. part of a

wall, then interpretation of *to the right of the door* must be based on the door's own axes. Otherwise (in a hardware store for example) an observer's line of sight may override the door's cross-axis.

#### 1. 4. RELATIVE MOBILITY

The mobility of the reference object relative to the located object influences the order of the nominals around the preposition: the more mobile object normally precedes the preposition. One will not say *the cognac bottle is the one in a cap*, but *the one with a cap on it*. Following Talmy [1978a], I will call the located object the "Figure", and the reference object the "Ground", when discussing the order of the nominals.

Human beings tend to play havoc with the relative mobility rule, either because they are the preferred topic (*the man in a blue coat*), or -- as center of the universe -- preferred reference object (*the Empire State building is in front of me*).

\*\*\*

Typicality plays an important role in determining an appropriate locative predication (and no doubt other types of expressions). The choice of expression tends to depend not on particular (non prototypical) attributes of the objects considered, but rather on typical attributes of the category to which they are assigned by the linguistic expression. If typical conditions do not obtain, they tend to be ignored, unless one has some special reason to bring attention to the atypical conditions. If for instance the cap of a bottle were glued to the wall, one would still say *the bottle with a cap on it*. Even if a tray has very high sides, one will say *on the tray*. Consider also the table pictured above. Imagine the space under it progressively more solidly enclosed; there is a point at which one might be struck by this and say *in the table*. But this point is rather far along; even with a table with a solid shelf at floor level, people consistently describe objects on that shelf as *under the table*.

## 2. CONTEXTUAL FACTORS

The choice of an appropriate locative predication also depends on various aspects of the context. Some of these contextual characteristics are discussed in this section as if they were neatly separable; in fact, all are interdependent in complex ways, and these interrelations must become clear before we can design models of comprehension and production.

### 2. 1. CONTEXT DEPENDENT PARAMETERS.

These include the location of an observer, for the deictic uses of some prepositions (*in front of the tree*), and an implicit (fuzzy) distance threshold for the prepositions indicating proximity [Denofsky 1976].

In *the gas-station is at the freeway*, an implicit cross-path is assumed. To say that "freeway" occurs as a metonymic substitute for "at the intersection of a cross-road with the freeway" is not very useful, since no general rule of metonymy will predict this one (as natural as such a substitution may sound to English speakers, it is not acceptable in French: see *le poste d'essence est à la route*).

### 2. 2. FIGURE/GROUND ASSIGNMENT

The assignment of the roles of Figure and Ground depends primarily on which of the two objects' location is at issue. The object whose location is at issue precedes the preposition: compare *the house near the church*

and *the church near the house*. But the assignment must also respect the relative mobility rule. *The house near the church* is reversible because both house and church are equally immobile; but *the bicycle near the church* is not. When one wishes to locate a less mobile object with respect to a more mobile one, there are a number of periphrastic devices -- one being the use of "with" as in the earlier example (*the bottle with a cap on it*); "with", not being basically locative, is not subject to this relative mobility rule. See also *the house is near where the bicycle is* (but *\*the house is near the bicycle* [Talmy 1978a]); this turns *bicycle* into an immovable entity, namely a *place*.

The mobility rule is in fact a consequence of the principle that the object whose location is at issue should precede the preposition. The Ground is typically bigger and less mobile than the Figure, since those objects whose location is most commonly at issue are those which move around, and a good reference object is one whose location can be inferred from its name, and thus had better be the same over some time.

What is at issue in turn depends on the speaker's purpose in constructing the locative predication, and how it fits into his/her overall discourse plan.

### 2. 3. VARYING VIEWPOINT ON THE OBJECT

Mainly this involves the contrast between a close-up and a remote view of the objects. Most often, this is not a matter of actual distance, but a way of viewing an object for a given purpose: one may choose to ignore one or more dimensions, or internal characteristics of the object. For example, a road may be seen as a strip (*a truck on the road*), or a line (*a village on the road to London*). Normally *behind the house* will be based on the house's own axes. But when looking from some distance, one may use one's line of sight as axis.

Another aspect of viewpoint, is the bounded/unbounded distinction. Compare *walking through versus across the water* [Talmy 1978b]: in the former the boundaries of the body of water are ignored, but in the latter, the extension of the body of water from one end to another is involved.

### 2. 4. RELEVANCE

Given the pictures below, one will say *the bread under the bowl*, but *the bulb in the socket*.



The socket is still functioning as a socket when facing down, the bowl not as a bowl. If function is the relevant aspect, it is of no interest to distinguish between situations where bulb and socket are as above, and their upside down versions. With the bowl, this distinction matters.

Similarly, the pear in B is *in the bowl*. It is not normally useful to distinguish between situation A and B.



For the two examples just described, one could contrive contexts in which the distinctions normally ignored would be important. And certainly adequate models of

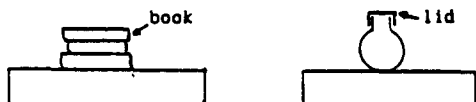
language should account for this possibility.

A locative expression may describe the general intention of a person over some time, rather than his precise location at the time of speaking. I could say *Lynn is at the store* even if I knew Lynn might still be on her way. But this may not be appropriate (e.g. if I know the addressee is at the store).

Relevance is important for Gricean inferences. For instance, from *Jon is near his desk*, one can generally infer *Jon is not at his desk*. If however I asked a friend on the phone "are you near your desk? could you look up the address...", an appropriate answer is "yes", even if my friend is at his/her desk. In this context proximity is the relevant aspect, and "near" becomes appropriate.

## 2.5. SALIENCE

The book below left is *on the table*, the lid (right) is not, because the intervening relation between the lid and the jar is salient. Such salience is not primarily a matter of the size of some intervening object.



One generally says that X is *in the field* and *in the bowl*, whenever field or bowl contain X. One may however say *the dust on the bowl*, and *the fertilizer on the field*. An adhering thin lamina brings attention to contact rather than inclusion.

## 2.6. HIGHLIGHTING SOME BACKGROUND ELEMENT

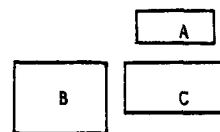
The choice between expressions is often a matter of bringing attention to some background element rather than signalling differences of fact. Thus *to the right*, as contrasted with *on the right*, tends to highlight the distance between the two objects, and to evoke travel away from the reference object; the contrast cannot always be described in terms of objective differences in the situations (it sometimes is: thus if a third object of the same kind is between the two considered, only *to the right* is appropriate). And *on the right side of the building* as contrasted with *on the right of the building* brings attention to the wall. Consider also *Bogota is at/on the equator*; "at" will be preferred if one wishes to signal the presence of some transverse line (e.g. a travel trajectory).

## 2.7. INDETERMINACY

Most spatial relations are true given a certain tolerance. The tolerance has a lower limit defined by the nature of the objects; its effective value depends on one's purpose, and the precision of one's knowledge. Thus, the angular precision with which *directly to the right* is defined varies with silverware on the table, chess pieces on a board, or houses on a block.

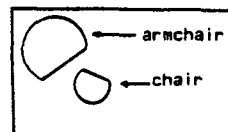
## 2.8. CONTRAST

"Polar concepts", i.e. terms like *to the right*, may behave like implicit comparatives. In some sense, *to the right* is better realized the closer the located object is to the "right axis". Thus, if I said *put the chair to the right of the desk*, I would expect you to put it more or less on the right axis of the desk. And, in the figure below, *A is to the right of B* only in the absence of C. The location of A must be contrasted with that of similar objects in the relevant part of space.



(One could however say here: *A is to the right and behind B*, or *A is diagonally to the right of B*. This suggests that even in the presence of C, *A is to the right of B* is true, but "uncooperative" [Grice 1974]. However, it is "uncooperative" precisely because of some intrinsic property of the concept *to the right* -- i.e. because "closer to the axis" is in some sense a better way to realize *to the right*. Even if one grants some usefulness to the semantic/pragmatic distinction, it does not neatly apply here.)

A similar use of contrast can be seen with *the chair is in the corner* in the figure below. It is not appropriate unless the armchair be removed.



The concept of a "corner" has built in that *in the corner* becomes less appropriate as one gets further from the vertex itself.

## 2.9. OTHERS

Many uses of the prepositions cannot be explained in terms of any of the above factors. One then needs a description of the context of use at a rather specific level. Consider for example the contexts in which one will say *Suzy is at the playground* versus *in the playground*. *In* would be (i) preferred if the speaker can see Suzy, (ii) required if the addressee expects Suzy to be just outside the playground, (iii) required if the speaker her/himself is in the playground (an analogous contrast exists between *at the beach* and *on the beach*). These conditions "suggest" a close-up view, and that the speaker's knowledge is precise; by contrast, "at" suggests a remote view, and imprecise knowledge. But "to suggest" is not to imply: one cannot infer these conditions of use from the ideas of a remote versus a close-up view.

## 3. CORE MEANINGS

With most of the examples given, the explanation suggested for the choice of a preposition assumes the existence of a "core meaning". This core meaning is basically a geometrical relationship between geometrical entities. Thus, in a given context, "geometrical descriptions" (say a point, line, surface, volume, lamina, etc.) are mapped onto the subject and object of the preposition. Strictly speaking, the core meanings are -- at best -- true only of these geometric descriptions. In fact, they may not even hold for any such geometric description -- see the *pear in a bowl* example above, assuming the natural core meaning for "in", i.e. "inclusion". Yet, the core meaning is then present as "prototype".

Here are informal definitions of the core meanings for two categories of prepositions, designated as "topological" (*at, on, in*), and "projective" (*to the right, behind, etc.*).

**Topological prepositions:**

- in*: partial inclusion of a geometrical construct in a volume, an area, or a line.
- on*: contiguity, adjacency of a geometrical construct with a surface, or a line.
- at*: coincidence of a point with a point in space.

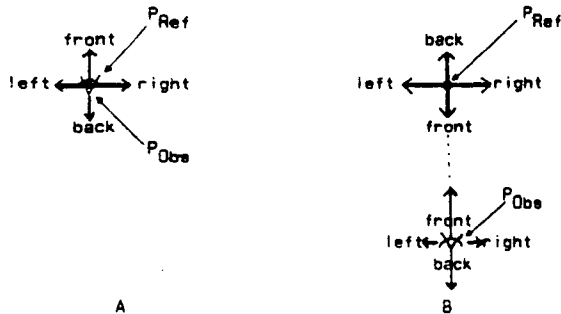
In actual context, inclusion, contiguity, and coincidence need not be true. Thus the book on top of a pile of books on the table can be said *on the table*, and *Mary is at the gate* when very close to it. But the relations represent the "ideal" around which particular instances gravitate. Thus *at* implies the closest reasonable relationship between two objects, and coincidence where sensible (*the center of the circle is at the intersection of the axes*). Of course, the core meanings are not sufficient to determine the conditions of use of a given preposition: one must also know precisely which deviations from the ideal are permitted. One principal process mediating between core meanings and actual conditions of use is the mapping of objects onto points, surfaces, and volumes.

I am not saying that the core meanings presented here are the only possible ones. Only when core meanings are incorporated in a global explanatory system will it be possible to make rigorous arguments for alternate choices. Those proposed here represent a good starting point.

**Projective prepositions:**

Each of these prepositions involves -- through fact, supposition, or metaphor -- a "point of observation". A point of observation consists of two vectors, one indicating the intrinsic vertical of the observer (it will not be the gravitational vertical if the observer is lying down, or not in the gravitational field), and the other orthogonal to the first along the line of sight. These two vectors completely specify four coplanar orthogonal half-line axes associated with the point of observation: the "front", "right", "back", and "left" axes, in clockwise order.

In the core meaning definition of these prepositions, reference and located objects are points. Given a point of observation, one can specify axes associated with the reference object -- the "base axes" (right, left, front, and back) by reference to which *to the right, behind, etc.*, will be defined. These axes originate at the reference object. If the point of observation ( $P_{Obs}$ ) coincides with the reference object ( $P_{Ref}$ ) (figure A below), the base axes are identical to those of the point of observation. If the point of observation is away from the reference object (figure B), the base axes are a mirror image of those of the point of observation -- the mirror plane being the bisector of the segment joining point of observation to reference object.



There are thus two possible orders of the base axes, as shown in A and B.

I will define the core meaning of each projective preposition as follows: given a punctual reference object ( $P_{Ref}$ ), punctual located object ( $P_{Loc}$ ), and a point of observation, base axes can be constructed according to the procedure outlined above:  $P_{Loc}$  is *to the left* of  $P_{Ref}$  iff it is located on the left base axis. Analogous definitions for the other prepositions are easy to formulate.

A few examples will help understand how these core meanings are manifest in the actual uses of the prepositions.

In *in front of a rolling stone*, the point of observation is "virtual" -- i.e. it is an hypothetical location and direction for viewing: the location is coincident with the stone, and the direction is the direction of movement. One must of course assume -- as with the objects in the examples that follow -- that the stone is assimilated to a point.

In *to the right of the chair*, the base axes may be specified as those intrinsic to the chair -- i.e. by reference to a typical user. The point of observation is then coincident with the reference object -- namely the chair. Coincidence may be the case when the base axes are not intrinsic to the reference object: for instance *on the right of the stool* may be defined with respect to somebody sitting on it, given a round stool with no intrinsic front axis.

Reference object and point of observation are separated in *the moon behind the cloud*. They can also be separated when the base axes are intrinsic to the reference object: with *on the right side of the closet*, the point of observation is defined by a typical user facing the closet.

Again, one might define the core meanings differently. In particular one could define the core meaning of "to the right" say, as implying location in the whole right-hand half-space instead of on the axis. The choice adopted here reflects the fact noted in earlier examples that the "ideal" realization of *to the right* is with the located object on the right base axis.

Processes other than the mapping of objects onto points may mediate between core meanings and actual conditions of use. The reference object may rotate: where is *on the right side of the painting* when the painting is tilted? *The tree to the right of the road* actually means "at some point of the road" -- think of a curving road), and *the city behind the barbed wire fence* assumes "integration" along the length of the fence (note one cannot say *the city to the right of the fence* to the same effect -- that is referring in this way to the whole city. The line of sight is a favored axis, as compared to right, left, and back axes).

**4. SOME CONCLUSIONS**

Here are the main problems and directions of inquiry suggested by the examples in this paper.

One cannot fully explain the use of a locative predication in a given situation in terms of the core meanings together with some inferences from general principles involving object knowledge, salience, relevance, the precision desired, etc. There will always remain uses involving some degree of arbitrariness (most uses are "motivated", if not "compositional" [Fillmore 1979] -- i.e. the morphemes composing the appropriate expression are normally selected from "reasonable" candidates). Even where such principles are at play, they may operate not at the comprehension/production level, but rather at the

phylogenetic level. To sort principled aspects of use from arbitrary ones, and to understand exactly where such principles operate, one must of course first establish their existence and nature.

In terms of knowledge of the physical world, I believe one important step toward adequate representations is to put the experiencer back into the picture. That is, it is not enough -- or even always necessary -- to know where what objects are; one must also consider how much fits into one field of view, how things "appear" as opposed to how they "are", how this changes with viewing distance, visibility, obstruction, etc.

General principles of salience should be studied: how some object parts or relations are selected as most important -- either in all imaginable contexts, or in some contexts. Salience underlies many instances of metonymy (in *at the front of the theatre*, "theatre" actually refers to the part occupied by the audience).

Many questions revolve around the issue of "relevance" -- of "what matters, to whom, in what circumstances" rather than the traditional concern with what is true. All existing artificial intelligence programs have ignored this problem by using a limited vocabulary in a limited domain, so that the question of selecting relevant utterances never arises.

Relevance is linked to the speaker's purpose, as many of the contextual factors described in this paper -- indeterminacy, Gricean inferences, highlighting of background elements, determination of the Figure/Ground relationship, etc. The set of "expressible" goals is constrained by the "potential" of the language, i.e. by a semantic system with finitely many options. One can only want to say what can be said, and said in a reasonable amount of time. Clearly, "planning" for natural language processing is a very important problem.

Purpose however, will not explain everything one says. Simple associative mechanisms must sometimes be responsible for what one says. For instance, some background element may be highlighted -- provided some linguistic means to do so exists -- only because some passive associative link has brought it to attention.

Once general principles are better understood, it is an open question whether they are used by speakers, or whether their explanatory power is at the phylogenetic level (and will thus be only implicit in the structure of the knowledge representation). For instance, although there is a general principle that the located object should be more mobile than the reference object, production may not proceed by inferences from this general principle together with scenarios involving the two objects. The linguistic expression (or pattern for expressions) may be attached to some representation of a "situation type" involving the two objects (or two superordinate objects). And although "at" generally implies the closest reasonable relationship between two objects, such a definition may never be used by a speaker -- or used only in the creation or understanding of novel types of expressions, metaphors, witticisms, etc.

What speakers do, and what computer models of comprehension and production processes should be made to do, are two different things: the latter depends on the constructor's goals, which should be subjected to some scrutiny.

A computational treatment of the use of prepositions will require much greater sophistication than naive representation theory would lead us to expect.

#### REFERENCES

Bogges, Lois C. 1979. Computational interpretation of English spatial prepositions. PHD thesis, University of Illinois, Urbana.

Denofsky, Murray E. 1976. How near is near: a near specialist. M.M.T. Artificial Intelligence Lab. memo no. 344, Cambridge, Mass.

Fillmore, Charles J. 1971. Santa Cruz lectures on deixis, presented at the University of California, Santa Cruz. Mimeographed, Indiana University Linguistics Club, Bloomington.

\_\_\_\_\_. 1979. Innocence: a second idealization for linguistics. *Proceedings, Fifth Annual Meeting, Berkeley Linguistics Society*, 63-76. University of California, Berkeley.

Grice, H. Paul. 1974. Logic and conversation. *Syntax and Semantics: speech acts*, vol. 3, ed. by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.

Talmy, Leonard. 1978a. Figure and ground in complex sentences. *Universals of Human Language*, vol. 4, ed. by Joseph H. Greenberg et al, 625-649. Stanford, Cal: Stanford Univ. Press.

\_\_\_\_\_. 1978b. The Relation of grammar to cognition -- a synopsis. *Theoretical Issues in Natural Language Processing-2*, 14-24. Coordinated Science Lab., University of Illinois, Urbana-Champaign.

