

Various Views on Spatial Prepositions

Gudula Retz-Schmidt

In this article, principles involving the intrinsic, deictic, and extrinsic use of spatial prepositions are examined from linguistic, psychological, and AI approaches. First, I define some important terms. Second, those prepositions which permit intrinsic, deictic, and extrinsic use are specified. Third, I examine how the frame of reference is determined for all three cases. Fourth, I look at ambiguities in the use of prepositions and how they can be resolved. Finally, I introduce the natural language dialog system CITYTOUR, which can cope with the intrinsic, deictic, and extrinsic use of spatial prepositions, and compare it with the approaches dealt with in the previous sections as well as to some other AI systems.

This article is a modified and extended version of a paper presented at the AAAI Workshop on Spatial Reasoning and Multisensor Fusion in October 1987 in St. Charles, Illinois.

When describing spatial relations in natural language, we often use spatial prepositions such as in, behind, or above. Spatial prepositions such as these, which are used in order to describe the location of one object in relation to another, are called relational prepositions (Clark 1973). Some of these prepositions, such as in, at, and near, only refer to topological relations between the objects. Others, such as in front of, behind, left of, right of, beside, above, and below, also convey information about the direction in which one object is located with respect to the other. Following Herskovits (1986), I call these projective prepositions. They are also called directional prepositions (Richtungspräpositionen) (Wunderlich and Herweg 1988).

Projective prepositions can be used in different ways. For example, if we say, "The ball is in front of the car," it can mean that we want to locate the ball in relation to the car from the point of view of the speaker, with respect to the orientation of the car itself, or with respect to the actual direction of the motion of the car. In the first case, I speak of deictic use; in the second case, of intrinsic use; and in the third case, of extrinsic use (Wunderlich 1985).

This article presents a survey of various approaches to the issues concerning the deictic, intrinsic, and extrinsic use of projective prepositions. These approaches belong to different disciplines, such as linguistics, psychology, and AI. In addition, the system CITYTOUR is introduced and compared to some other AI systems that deal with spatial prepositions.

In the following sections, I call the object that is to be located the primary object and the object in relation to

which the primary object is located the reference object. In order to use projective prepositions, the speaker has to establish a reference frame (that is, an orientation) that determines the direction in which the primary object is located in relation to the reference object.

The reference frame can be established in different ways. One way is to use the intrinsic orientation of the reference object. In this case, the regions that are above, below, in front of, behind, to the left of, and to the right of the reference objects are the regions which are adjacent to the top, bottom, front, back, left side, and right side, respectively. If the intrinsic orientation of the reference object is used to establish the reference frame, I am referring to the intrinsic use of the corresponding prepositions. Thus, in intrinsic use, two arguments are needed for a locative description: the primary object and the reference object (Herrmann et al. 1986).

If the reference object doesn't have an intrinsic orientation, or its intrinsic orientation isn't used for establishing the frame of reference, factors of the situational context determine the reference frame (Wunderlich 1986) and, thus, impose an extrinsic orientation on the reference object. In this case, I am referring to the extrinsic use of prepositions. Relevant contextual factors are, for example, the accessibility of the reference object, its motion, other objects in its vicinity, or the gravitation of the earth (see How is the Reference Frame Determined in Extrinsic Use?). Thus, in extrinsic use, three arguments are needed for a locative description: the primary object, the reference object, and a contextual factor (Hays 1987; Herrmann et al. 1986; and Wunder-

	Primary Object	Reference Object	Point of View
Bennett 75, p.83 Bühler 82, p. 13ff Bürkle et al. 86, p. 4		anchor	reference point origo perspective, point of view, reference point
Hays 87, p. 6 Herskovits 86, p. 156f. Jackendoff 83 Jackendoff 87, p. 202	located object theme figural object, theme figure	reference object reference object reference object landmark reference object ground	point of observation
Kautz 85, p. 2-33 Klein 78, p. 20 Klein 82, p. 162 Klein 83, p. 291 Lakeoff 87, p. 419 Langacker 86 Levelt 86, p. 188 Nirenburg, Raskin 87, p. 366 Schulze 87 Sennholz 85 Talmy 83	trajector figure Verweisobjekt figure figure primary object figure	landmark ground Bezugsobjekt ground ground secondary object ground, reference object, primary reference object	origo, reference point origo, reference point origo secondary reference object
Ullmer-Ehrich 82, p. 228 Vandeloise 84, p. 3f., 187f.	trajector	landmark	origo, reference point point of reference (consciously taken on) or (virtual) viewpoint (unconsciously taken on by the speaker)

Table 1. Other Terms for Primary Object, Reference Object, and Point of View.

lich and Herweg 1988).

An important contextual factor that can impose an orientation on the reference object is the point of view from which the reference object is viewed (either in actual fact or mentally through an act of imagination). In particular, the speaker's location can serve as the point of view. The hearer's location might also serve as the point of view but is used less frequently (Bürkle, Nirmaier, and Herrmann 1986) (see *When Are Prepositions Used Intrinsically, Deictically, or Extrinsically?*). In both cases, I am referring to deictic use. Thus, in deictic use, the third argument is the point of view (Hays 1987; Wunderlich and Herweg 1988). Deictic use can be explicitly marked in natural language descriptions, as in the sentence "From here, the bank is behind the department store." A deictic variant of intrinsic use also exists, for example, "The ball is to the right of me" (Herrmann et al. 1986).

In theory, locations other than those of the speaker or hearer can serve as the point of view, too, as in

the sentence "The travel agency is to the left of the post office, as seen from the town hall." This kind of extrinsic use is classified as deixis at phantasma in Bühler's (1982) terminology and as displacement deixis (*Versetzungsdeixis*) in Sennholz's (1985) terminology because the point of view doesn't coincide with the location of the speaker (or hearer). The particular sentence pattern in this example is quite rare in natural language. Modifications of it, however, can occur, for example, in route descriptions, as in "In order to get to the travel agency, you have to go to the town hall. From there, you walk to the post office. Then, it's on your left."

Different terms for primary object, reference object, and point of view can be found within the relevant literature. A selection is listed in table 1. Additional terms for point of view, or origo are given in Sennholz (1985).

Ehrich, in an analogy to Reichenbach's (1947) three-part system for the interpretation of tenses, distinguishes between the speaker's place, the denotation space, and the reference space

(Ehrich 1982). Her denotation space comprises our primary object and reference object, and our point of view corresponds to her reference space, which can coincide with her speaker's place (Ehrich 1982).

Different terms also exist for intrinsic use and deictic use. Levelt (1982, 1986) and Bürkle, Nirmaier, and Herrmann (1986) use the terms *intrinsic system* and *deictic system*. Saile (1984) speaks of objective interpretation and subjective interpretation. Rauh (1984) uses the terms *nonegocentric use* and *egocentric use*. Klein (1983) speaks of *intrinsic* and *unmarked origo*.

Talmy's (1983) terminology and classification are slightly different. He distinguishes between characterizing location by one and more than one reference object. In the case of only one reference object, we have *intrinsic use*. Talmy's category of localizations with two reference objects includes, among others, our *extrinsic* and, in particular, *deictic uses*.

A similar way of classifying different uses of projective prepositions is that of Herrmann et al. (1986). They distinguish between two-point and three-point localizations. Examples of two-point localizations are "The ball is to the right of me" and "The ball is behind you" (which, in my terminology, are examples of *intrinsic*, as well as *deictic*, use because the reference object, as well as the point of view, are the speaker's or hearer's location), and "The ball is in front of the car" (which I regard as *intrinsic*). Examples of three-point localizations are "The ball is to the right of the lamp, as seen from my point of view" and "The ball is in front of the block, as seen from your point of view" (which I call *deictic*). Examples of *extrinsic use* are not given in their paper.

Primary objects can be located inside ("The altar is in the front part of the church.") or outside ("The fountain is in front of the church.") a reference object. In this article, I deal mainly with the latter kind of localization.

In natural language descriptions, not all combinations of primary objects and reference objects are possible. Nobody would utter a sentence such as "The cathedral is behind the

red bicycle." With respect to size, salience, mobility and so on, the reference object and the primary object have to be in a certain relation to each other (Talmy 1983; Schulze 1987). I don't deal with this issue in detail here.

Which Prepositions Can Be Used Intrinsically, Deictically, and Extrinsically?

As stated in the beginning of this article, only the projective prepositions allow intrinsic, deictic, and extrinsic use. In particular, most authors (Dirven 1981; Herrmann 1987; Herrmann et al. 1986; Herskovits 1986; Miller and Johnson-Laird 1976; Saile 1984; Ullmer-Ehrich 1982; Wunderlich 1985; Wunderlich and Herweg 1988) agree that in front of, behind, to the left of, and to the right of belong to this group. Ullmer-Ehrich (1982) classifies them as belonging to the secondary deixis system, in contrast to the terms here and there, which belong to the primary deixis system. In addition to in front of, behind, left of, and right of, Saile (1984) includes the preposition beside.

Miller and Johnson-Laird (1976), as well as Wunderlich and Herweg (1988), include the prepositions over and under. Talmy (1983) includes a group of expressions that describe relations in the vertical axis, that is, on the top of, on the bottom of, on top of, underneath, over, under, above, below, higher than, and lower than.

However, spatial dimensions are not equal. According to many authors (Adorni, Boccalatte, and DiManzo 1981; Clark 1973; Herskovits 1986; Hill 1982; Lyons 1977; Miller and Johnson-Laird 1976; Wunderlich 1985), the vertical dimension has the status of a privileged direction because it is fixed by the gravitation of the earth, whereas in both horizontal dimensions, man can move freely. Thus, the vertical dimension can be conceived of as the primary dimension. Moreover, the human body is asymmetric in the front-back dimension, enabling man to distinguish more easily between front and back than between left and right. This asymmetry permits the less salient distinction between the front-back

dimension as the secondary dimension and the left-right dimension as the tertiary dimension (Hill 1982; Lyons 1977; Saile 1984).

For prepositions describing verticality, deictic use is not possible (Vandeloise 1984). However, we can distinguish two uses based on different concepts of verticality. One is based on the intrinsic top and bottom of the reference object (Clark calls it intrinsic verticality) and, thus, can be regarded as intrinsic use. The other is based on the gravitation of the earth (Clark [1973] calls it geological or gravitational verticality; also see Sondheimer [1974]). The latter case is a kind of extrinsic use in which the earth imposes its orientation on the reference object.

Thus, ambiguities can occur if objects that usually stand upright are lying horizontally, as in the example "There is a fly three inches above the lady's knee" (Clark 1973; Wunderlich 1985; also see Levelt 1986). However, Adorni, Boccalatte, and DiManzo (1981) argue that gravitational verticality is more absolute than intrinsic verticality and, thus, has priority.

In the following sections, I concentrate on spatial relations in the horizontal plane and, thus, on the prepositions in front of, behind, to the left of, and to the right of.

How Is the Reference Frame Determined in Intrinsic Use?

In intrinsic use, the front, back, left, and right regions around a reference object are those regions adjacent to the intrinsic front, back, left, and right sides of the reference object, respectively. Because of the symmetry in the left-right dimension, intrinsic left and right sides of objects are rare (Vandeloise 1984). Because of this symmetry and the dominance of gravitation in the vertical dimension, the determination of intrinsic fronts seems to be the most interesting case. Once the intrinsic front is identified, the back and the left and right sides can be deduced (see the end of this section).

Instead of intrinsic, often the term inherent is used (for example, in Wunderlich and Herweg [1988]). Clark (1973), for instance, speaks of inherent

front (which I call intrinsic front) to differentiate it from the egocentric front (deictic front). Lyons (1977) uses the term in a different way: He distinguishes between inherent orientation and canonical orientation. When a certain orientation is an indispensable characteristic of an object (for example, a top in the case of mountains, buildings, or trees and a front in the case of human beings or animals), he speaks of inherent orientation. When an object is usually but not necessarily oriented in a certain way, he speaks of canonical orientation. Because for us this difference is irrelevant, I subsume both terms under intrinsic orientation. Cresswell (1978) and Dirven (1981) speak of conventional front. Saile (1984) uses the terms objective front (in the case of intrinsic use) and subjective front (in the case of deictic use). Finally, Talmy (1983) speaks of biased parts (front, back, right, left, top, bottom).

In the following paragraphs, criteria for the identification of the intrinsic front are pointed out. According to Miller and Johnson-Laird, the intrinsic front is the side lying in the characteristic direction of motion (for example, for arrows, bullets, and torpedos), the side containing the perceptual apparatus (for example, for people, animals, dolls, and cameras), or the side characteristically oriented to the observer (for example, for cars, chairs, clothing, desks, radios, and mirrors). Examples of objects that have no intrinsic front are tables, vases, trees, blocks, cubes, balls, and stars (Miller and Johnson-Laird 1976; Saile 1984; Wunderlich and Herweg 1988).

These criteria are adopted or only slightly changed by many other authors (Bennett 1975; Bürkle, Nirmaier, and Herrmann 1986; Hill 1982; Lyons 1977; Sondheimer 1974, 1976). According to Lyons, confrontation (that is, the canonical encounter) and the canonical direction of locomotion serve to identify the canonical front. For Bennett, the inherent front is the side that normally leads when the object moves (for example, for busses) or the side which is most frequently seen (for example, for stations). According to Bürkle, Nirmaier, and Herrmann, as well as Sondheimer, the

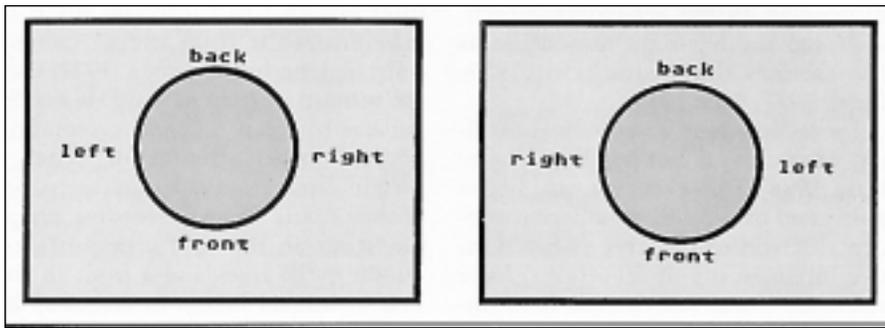


Figure 1. (a) Reference Object Seen from Outside;
(b) Reference Object Seen from Inside.

intrinsic front is defined either by means of anthropomorphic criteria (the side with perceptual apparatus, for example, for cameras, or the side that leads in motion, for example, for locomotive engines) or familiarity conventions (the way things are worn, for example, for clothes, or handled, for example, for cupboards or typewriters, or the direction in which one moves with them, for example, for cars).

In addition to the way objects are used (for example, main access to buildings, rooms, or open spaces) and anthropomorphic metaphors (that is, resemblance to human beings, for example, with respect to normal direction of motion, facial features, or other striking features), Herskovits (1986) states symmetry as a criterion for determining the intrinsic fronts of objects.

Conflicts can occur if two or more criteria contradict each other. Fillmore (1971) argues that in the case of animals, the criterion based on the perceptual apparatus outweighs the criterion based on the direction of motion because we speak of crabs as moving sideways, not as having their heads on the sides of their bodies.

Once the intrinsic front of a reference object is identified, the other sides can be assigned in two ways, depending on whether the reference object is thought of as seen from the outside (see figure 1a), as in the case of radios, mirrors, desks, cupboards, or buildings, or the inside (see figure 1b), as in the case of chairs or clothing (Miller and Johnson-Laird 1976).

If the intrinsic front is the same, regardless of whether the reference object is seen from the inside or the

outside, the difference doesn't have any effect on the assignment of the front and back regions because the front region is always adjacent to the intrinsic front and the back region to the intrinsic back of the reference object. However, the difference does affect the determination of the left and right regions because if the reference object is seen from the outside, the left side is clockwise from the front (see figure 1a), whereas if it is seen from the inside, the left side is counterclockwise from the front (see figure 1b) (Levelt 1986; Miller and Johnson-Laird 1976; Saile 1984; Sondheimer 1974, 1976). Human beings and other animate beings (with a perceptual apparatus that defines their fronts) as reference objects are also treated as if they were seen from the inside (Miller and Johnson-Laird 1976; Sondheimer 1974).

The situation shown in figure 1b would be classified by Herskovits as a coincidence situation because observer and reference object are assumed to coincide (here, the base coordinate axes are in basic order). In contrast to this situation, in an encounter situation, observer and reference object are thought of as being opposite each other (see figure 1a). This distinction is independent of the distinction between intrinsic and deictic use (Herskovits 1986).

Conflicts in identifying the left and right sides of the reference object can occur in those cases in which the reference object can be seen from the inside, as well as the outside, as in the case of cars. The intrinsic front of a reference object might be different, depending on whether it is seen from the outside or inside (as in the case of

churches or theaters) (Herskovits 1986). However, this ambiguity doesn't pose any big problems because the reference object would usually be thought of as seen from the outside if the primary object was located outside it and from the inside if the primary object was located inside it.

How Is the Reference Frame Determined in Extrinsic Use?

If a reference object doesn't have an intrinsic front (for example, a tree), or its intrinsic orientation isn't used for establishing the frame of reference, a front can still be contextually induced or projected on it (Hill 1982; Miller and Johnson-Laird 1976; Saile 1984; Vandeloise 1984; Wunderlich and Herweg 1988). In this case, Miller and Johnson-Laird speak of it as an accidental front. An accidental front can be induced on an object in different ways. According to Hill, Vandeloise, and Miller and Johnson-Laird, an object can acquire an accidental front through other objects in its vicinity (for example, the front of a tree standing in a yard in front of a house will be the side facing the street [Miller and Johnson-Laird 1976]).

Wunderlich uses the criterion of accessibility: The front of a reference object is the side that is accessible the soonest inside a container, the side which is accessible the soonest at a material boundary, the side which is accessible the soonest by a particular moving primary object, or the side which is accessible the soonest because of the reference object's own actual motion (Wunderlich and Herweg 1988; Wunderlich 1986; Herskovits 1986; Levelt 1986). The first two conditions of Wunderlich can be subsumed under Miller and Johnson-Laird's criterion of contextual induction through other objects in the vicinity.

While a criterion for determining intrinsic fronts was the way objects usually move or are worn, a criterion for the contextual induction of fronts can be the way objects actually move or are worn in the particular situation (Wunderlich and Herweg 1988; Herskovits 1986).

Finally, the vertical axis, determined by the gravitation of the earth,

can also serve to impose sides (in this case, tops and bottoms) on objects. In Which Prepositions Can Be Used Intrinsically, Deictically, and Extrinsically? this projection of top and bottom sides by means of gravitation was called geological or gravitational verticality.

Herskovits (1986) doesn't distinguish between intrinsic use and the kind of extrinsic use described here. She subsumes intrinsic sides and contextually induced sides (where a certain side is intrinsic to an object other than the reference object) under privileged directions.

In the kind of extrinsic use that can be conceived of as deixis at phantasma (see the beginning of this article), the reference frame is determined by an analogy to deictic use (see the next section).

How Is the Reference Frame Determined in Deictic Use?

In deictic use, the front, back, left, and right regions around a reference object are those regions adjacent to the deictic front, back, left, and right sides of the reference object, respectively.

In deictic use, too, the reference frame can be orientated in two ways. One way is to have the (deictic) front facing the point of view. The opposite side of the reference object is its back (see figure 2a) (Bennett 1975). This constellation is the way the front and back sides are assigned to an object in the Indo-European and many other languages that belong to the group of facing languages (Hill 1982; Levelt 1986); the reference object is treated "as if it were the other person in a canonical encounter, a person facing directly towards the speaker" (Clark 1973, p. 45).

In Hausa, an aligning language (Hill 1982; Levelt 1986), the opposite way (called the tandem principle) is chosen: The reference object can be thought of as looking in the same direction as the speaker. Thus, the back is facing the point of view, and the front is on the far side of the reference object (see figure 2b) (Hill 1982; Talmy 1983; Wunderlich 1985; Wunderlich and Herweg 1988).

The assignment of left and right regions to the reference object doesn't

follow the proper rules of canonical encounter (Clark 1973). The application is not reversed but maintained. This constellation is expressed more appropriately by the term Spiegelbildprinzip (mirror principle) used by Wunderlich and Herweg (1988) (see figure 2a).

Herskovits (1986) uses a different term for the same phenomenon. Her term encounter situation (where the base axes of the coordinate system are in mirror order) denotes the same relationship between speaker and reference object as the term mirror principle (see figure 2a).

When Are Prepositions Used Intrinsically, Deictically, or Extrinsically?

Sentences such as "The post office is behind the church" are ambiguous because they can be used either deictically, with the point of view being omitted, or intrinsically (Bennett 1975; Clark 1973; Jackendoff 1987; Lyons 1977; Saile 1984; Sondheimer 1976). Apart from ambiguities between intrinsic and deictic use, ambiguities among deictic uses (between speaker or hearer as the point of view [Klein 1983]), among intrinsic uses (between the different criteria in determining the intrinsic front and the left and right regions [Sondheimer 1976]), between intrinsic and extrinsic use (because an object with an intrinsic front can also acquire an accidental front [Wunderlich and Herweg 1988; Sondheimer 1976]), and even among extrinsic uses (between different extrinsic orientations imposed on the reference object) can occur.

Ambiguities between deictic and intrinsic use can be avoided by explicitly stating the point of view (using expressions such as from here) or reformulating the sentence in a way that it can only be interpreted intrinsically. Examples in German and French are "Das Taxi steht, von mir/dir aus gesehen, hinter dem Lieferwagen" (deictic) versus "Das Taxi steht an der Rückseite des Lieferwagens" (intrinsic) (Saile 1984), "C'est ma soeur a gauche de Jean" (deictic) versus "C'est ma soeur a la gauche de

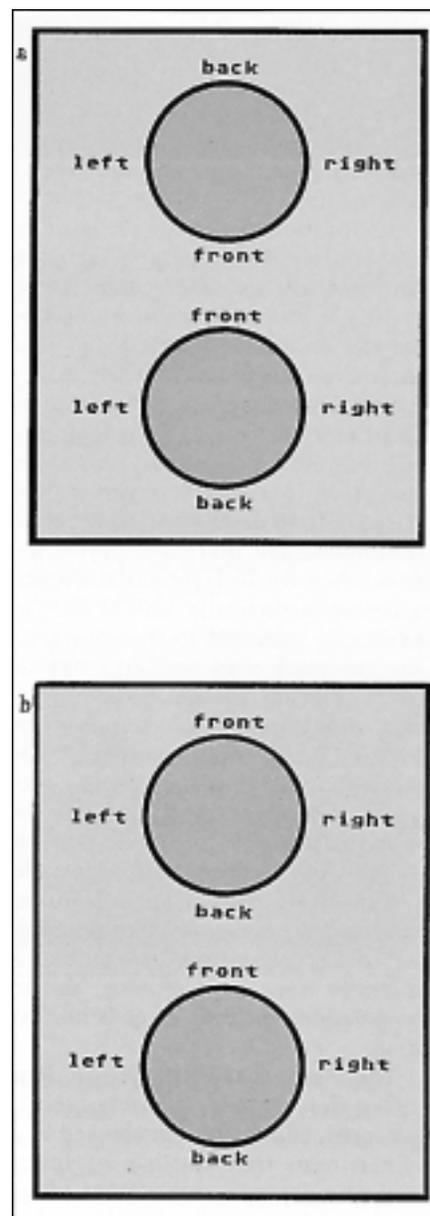


Figure 2.
(a) The Mirror Principle;
(b) The Tandem Principle.

Jean" (intrinsic) (Hill 1982), and "Ronald est devant l'arbre." (deict. or intr.) versus "Ronald est en face de l'arbre" (intrinsic) (Vandeloise 1984). Some suggestions for avoiding ambiguities in the use of locative expressions in man-machine communication are given in Sondheimer (1974, 1976).

Miller and Johnson-Laird claim that ambiguities between deictic and intrinsic use are resolved in the following way: Intrinsic interpretations usually dominate deictic ones. Thus, deictic use has to be explicitly marked using expressions such as "from my point of view" or "as I am looking at it." For interpretations, this dominance of intrinsic use means that "people first determine whether the landmark has intrinsic parts. If it does, they try to interpret the spatial relation intrinsically unless they are explicitly informed to the contrary. If the landmark does not have intrinsic parts relevant to the spatial indication, they must rely on context to provide a deictic interpretation. If both strategies fail, they may ask for more explicit information" (Miller and Johnson-Laird 1976, p. 398f; also see Saile 1984). Levelt (1986), states that the intrinsic system is predominant under certain conditions, in particular when the reference object has an intrinsic front or is moving, and the prepositions in front of or behind are used.

The claim that intrinsic use dominates deictic use is not generally accepted. Ullmer-Ehrich showed in an experiment that the use of spatial concepts depends on the kind of text produced. In room descriptions, which are often structured as a gaze tour, the deictic system is predominant. Ambiguity between deixis and intrinsics is avoided "by using temporal in place of spatial expressions, which - due to the underlying tour format of the descriptions - provide the desired singularity of descriptions" (Ullmer-Ehrich 1982, p. 247f).

According to Levelt (1982), individual preferences determine whether the intrinsic or deictic system is used. In contrast to Levelt, Bürkle, Nirmaier, and Herrmann (1986) found through experimenting that the selection of deictic versus intrinsic use is

not a matter of individual preferences, but that it depends on general conditions of the communication situation. The deictic system is predominant; it is the default or canonical case and can be modified within certain limits by the communication situation. If the perceptual field is not shared by speaker and listener, the predominance of deictic use is less striking. Deictic use is often marked by the explicit mention of the point of view (speaker). The selection of the point of view also depends on the kind of listener. If the listener is specifically in need of information (like a child) or authorized to receive exact information (like a professor), the amount of listener-oriented localizations (that is, the listener as point of view) increases (Bürkle, Nirmaier, and Herrmann 1986; Herrmann 1987). Vandeloise (1984) calls this adaptation transfer by intermediate person and views it as a manifestation of the principle of cooperation.

These results suggest that speaker-oriented deictic use is easiest for the speaker, whereas listener-oriented deictic use is easier for the listener. Intrinsic use might take on an intermediate position (Miller and Johnson-Laird 1976), although it also seems plausible that for the speaker, it is as difficult as the listener-oriented deictic system and, for the listener, as difficult as the speaker-oriented deictic system. However, the simplicity with which locative expressions involving deictic or intrinsic use can be interpreted also depends on the situation (how the speaker and listener are located in relation to each other) (Herrmann 1987) and the context (how preceding locative expressions were used) (Ullmer-Ehrich 1982). Herrmann, Bürkle, and Nirmaier (1987) showed in experiments that the cognitive effort for listener-oriented deictic use depends on the relative positions of the listener and the primary object with respect to the speaker's location.

Herskovits (1986) gives a list of examples for criteria that can determine the speaker's choice of the frame of reference (that is, the front, back, left, and right region of the reference object) as well as a few heuristics for selecting among them. Her specification is not restricted to the

distinction between intrinsic and deictic use (which is not explicitly dealt with) but is concerned with the determination of the frame of reference in general.

The effect of unusual perception situations (for example, supine position or absence of gravity) on the choice of the reference frame is dealt with in Bayer and Marslen-Wilson (1986) and Levelt (1986). For a discussion of the different factors that influence the choice between intrinsic and deictic use, see Levelt (1986).

The System CITYTOUR

The system CITYTOUR (Andre et al. 1986, 1987) is part of the Visual Translator project (VITRA), which deals with the relationship between natural language and vision. Experimental studies in the design and implementation of an interface between image-understanding and natural language systems are being carried out. The aim is to develop a system for the natural language description of image sequences. The results of this study can be applied to make the output of image-understanding systems more easily accessible for human users.

CITYTOUR¹ operates in the domain of city and traffic scenes. It answers natural language questions (in German) about the spatial relations between objects in a scene and about the movements of objects. An example of a scene is shown in the right window in figure 3. It is a part of the city center of Saarbrücken. Other scenes are the campus of the University of Saarbrücken and the traffic scene "Durlacher Tor" in Karlsruhe. In the latter case, the input data for CITYTOUR are actually provided by a vision system developed at the Fraunhofer Institut für Informations und Datenverarbeitung in Karlsruhe (Sung and Zimmermann 1986; Schirra et al. 1987).²

CITYTOUR mainly deals with spatial relations between nonmoving objects (called static objects) and between a moving (called dynamic object) and a static object. In the system CITYTOUR, static objects are represented by the following characteristics: the center of gravity³; the closed polygon; the delimitative rectangle; and the intrinsic

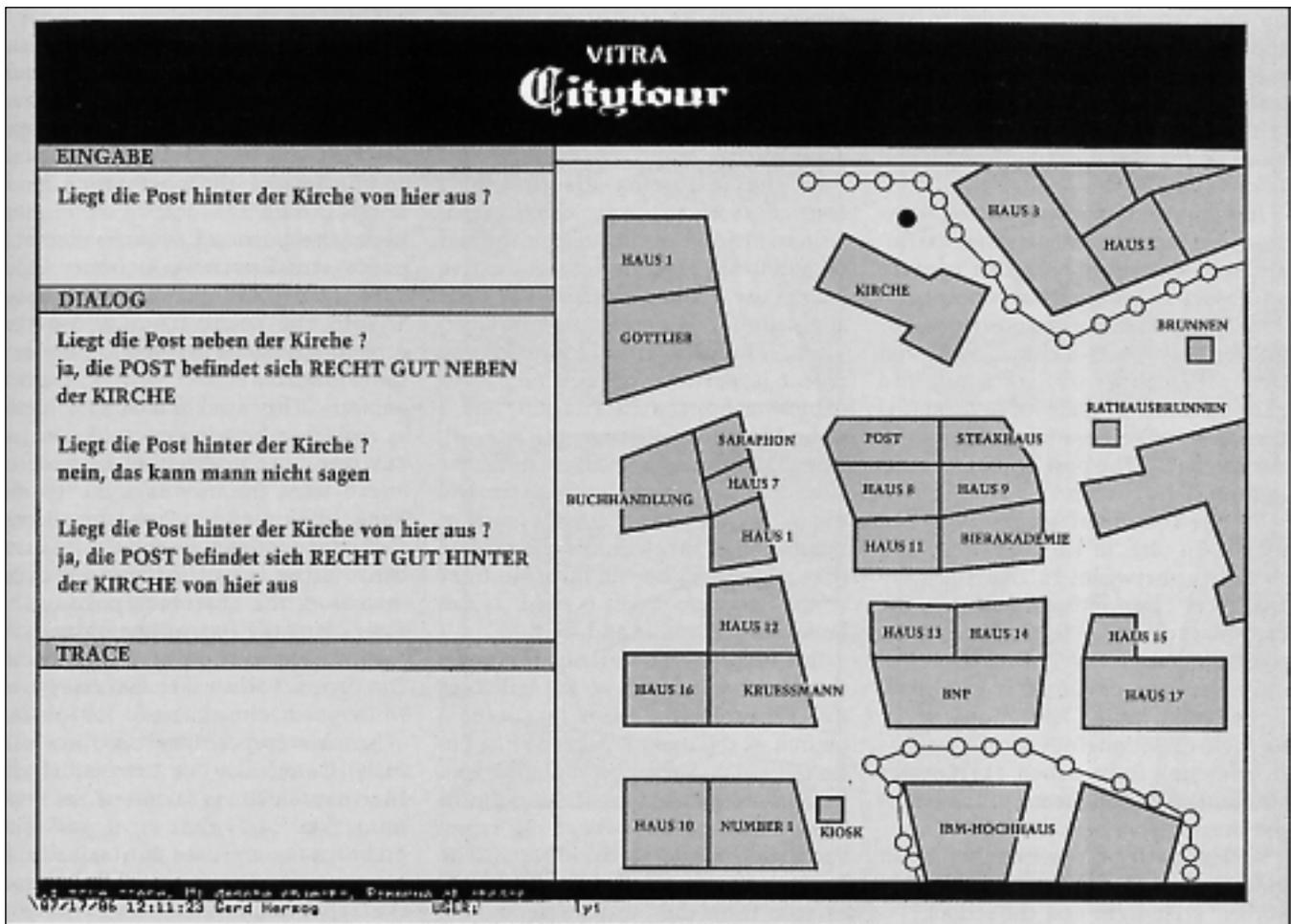


Figure 3. CITYTOUR's windows on the screen.

front, which determines the intrinsic orientation. Dynamic objects are represented as a sequence of space-time-coordinate pairs, called trajectory. So far, only the two dimensions in the horizontal plane are considered (as in city maps). The vertical dimension is ignored.

In addition to the issue of intrinsic, extrinsic, and deictic use, degrees of applicability of spatial prepositions and the semantics of path prepositions were investigated within the framework of CITYTOUR. Because it seems inappropriate to assume a fixed borderline that separates the area in which a spatial relation (for example, "x is behind the church") holds from the area in which it doesn't hold, degrees of applicability of spatial prepositions are calculated. The degree depends on the size of the reference object and the distance

between the primary object and the reference object. The degrees of applicability are expressed by Fuzzy values (LeFavre 1977). In natural language answers, they are described by linguistic hedges (such as directly or approximately). (See Herskovits [1986], who discusses graded concepts).

As examples of path prepositions, along and past were examined. In order for the relation along to hold between a trajectory and a static object, the trajectory has to follow the contour of the object much more closely than in the case of the relation past (Dirven 1981). Therefore, in order to be appropriate, the representation for the static object needs to be more detailed in the case of along than in the case of past. In CITYTOUR, the closed polygon is used for the calculation of along and the delineative rectangle (which is a rougher representa-

tion) for past.

This short outline discussed those aspects of CITYTOUR which won't be dealt with in this article. A more detailed description can be found in Andre et al. (1986, 1987).

Of those AI systems which deal with spatial relations, not many dwell on the distinction between intrinsic and deictic use. In fact, most of them don't take the observer's position into account and, thus, can only cope with intrinsic use. For example, the vision system Alven, which recognizes motion concepts and utilizes spatial concepts such as rightward, explicitly dispenses with the participation of the observer in the description process and, thus, cannot distinguish between intrinsic and deictic use (Tsotsos et al. 1980). Also, in Badler's (1975) system, which analyzes and describes object movements in scenes, only intrinsic

object sides are stored and can be used for localizations. The same holds true for the space planning and design system Wright. This system uses the position relations at-front, at-back, at-left, and at-right, which are defined with respect to object-centered coordinates of the design units (Baykan and Fox 1987).

In contrast to these systems, CITYTOUR can cope with intrinsic, extrinsic, and deictic use. It can operate with the following spatial concepts:

1. The projective prepositions *vor* (in front of), *hinter* (behind), *links von* (left of), *rechts von* (right of), and *neben* (beside) in their locative (static) use, as in “Die Post ist direkt hinter der Kirche” (“The post office is directly behind the church.”)

2. The relational prepositions *bei* (near), *an* (at), *in* (in), *auf* (on), and *zwischen* (between) in their locative use, as in “Der Springbrunnen ist auf dem Marktplatz” (“The fountain is on the marketplace.”)

3. The prepositions *vor*, *hinter*, *links neben*, *rechts neben*, and *neben* in their directional (dynamic) use, as in “Sie ging links neben das Theater von hier aus” (“She went to the left of the theater from here.”)

4. The path preposition *vorbei* (past), as in “Er ging an der Schule vorbei” (“He went past the school.”)

5. Combinations of *vor*, *hinter*, *links an*, *rechts an*, and *neben* with *vorbei*, as in “Sie ging vor der Kirche vorbei” (“She passed in front of the church.”)

6. The path prepositions *entlang an* and *entlang* (along), as in “Er ging am Warenhaus entlang” (“He went along the department store.”) and “Sie ging die Hauptstrasse entlang” (“She went along Main Street.”)

7. The verbs *abbiegen* (turn off) and *einbiegen* (turn into), as in “Sie bog von der Churchillstrasse in die Kennedystrasse ab” (“She turned off from Churchill Street into Kennedy Street.”)

8. The verbs *anhalten* (stop) and *anfahen* (start off), possibly with location specifications of the corresponding events using relational prepositions, as in “Er hielt an der Ampel an” (“He stopped at the traffic lights.”) (Schirra et al. 1987)

Intrinsic, deictic, and extrinsic use

is possible in cases 1, 3, 5, and 8. Thus, questions such as “Where is the supermarket, as seen from the town hall?” can be answered “The supermarket is behind the post office, as seen from the town hall.” However, this type of question is rare in human conversation.

In the following discussion, I restrict myself to the projective prepositions in front of, behind, to the left of, and to the right of in their locative (static) use. Now, I describe how their applicability is determined in CITYTOUR: The area around a reference object is partitioned into four half-planes—a front, a back, a left, and a right half-plane (figures 4a, b, c, d). The half-planes are aligned to the sides of a delineative rectangle around the reference object. In the case of intrinsic use, the delineative rectangle that is oriented by the intrinsic front of the reference object is used, as can be seen in figures 4a and 4b.

In CITYTOUR, the intrinsic fronts are the prominent fronts of the buildings (for example, the west facade of a church or the main entrance of a public building). They have to be defined by the user when the scene is built. They then become part of the representation of the static objects. The intrinsic fronts are always thought of as seen from the outside. Hence, the left and right sides are assigned as shown in figure 1a.

For deictic use, the delineative rectangle that is oriented by the observer's position is taken (figures 4c and 4d). It is then determined in which half-plane(s) the primary object is located so that the appropriate spatial preposition can be selected for the description.

Both speaker (that is, the user) and hearer (that is, CITYTOUR itself) are assumed to be at the same location inside the scene, the location referred to as the observer's position. It is marked on the screen as a large dot (figure 1) and can be moved during the session. The speaker is assumed to be asking the hearer, who knows more about the city (for example, the guide of a sightseeing tour), about the locations of buildings, streets, and so on. The questions are answered by CITYTOUR either with respect to the intrinsic orientation of the buildings and so

on, or with respect to the observer's position.

Thus, for the actual user of the system, we have the case of analogical deixis (Klein 1978, 1982, 1983; Rauh 1984; Sennholz 1985) because when the user asks, “Ist die Bank links von der Post von hier aus?” (“Is the bank to the left of the post office from here?”), with the deictic expression here, the point of view is not the user's actual position in front of the screen but the observer's position within the scene (that is, on the screen). The same is true for the user's interpretation of the system's answers. This kind of analogical deixis can also be viewed as a case of “extrinsic use in disguise” because in effect, what the user asks is, “Ist die Bank links von der Post vom Betrachterstandort aus gesehen?” (“Is the bank to the left of the post office as seen from the observer's position?”), thus a kind of deixis at phantasma.

In the generation of descriptions, CITYTOUR follows the strategy of Miller and Johnson-Laird (1976) (see *When Are Prepositions Used Intrinsically, Deictically, or Extrinsically?*). Intrinsic use is treated as the unmarked case; that is, if possible, prepositions are used intrinsically. If intrinsic use is not possible because the reference object doesn't have a prominent front, the deictic system can be used without explicit mention. If intrinsic use is possible, the deictic system can be used but has to be marked explicitly by means of the phrase *von hier aus* (from here).

CITYTOUR, as well as other systems that distinguish between deictic and intrinsic use, at the moment evades most of the problems mentioned in this article. Contextual induction of fronts is not yet considered, and intrinsic fronts are not determined by the system but are defined beforehand.

Kautz accepts the orientation of objects as primitive (Kautz 1985). In Adorni, Boccalatte, and DiManzo (1981), the privileged direction of an object is indicated by the x-axis of a particular Cartesian triple associated with the object. The storage of intrinsic orientations in memory is in accordance with Saile's (1984) view that in human memory, intrinsic fronts, regardless of how they were

determined, are stored in the internal lexicon.

Extrinsic use in the sense that the point of view doesn't coincide with the speaker's or hearer's location is rarely dealt with in the relevant literature and is not considered in any other natural language system apart from CITYTOUR. Dynamic uses of in front of, behind, to the left of, and to the right of as well as combinations of these with other prepositions such as past, are also rarely dealt with but can be handled by CITYTOUR.

So far, CITYTOUR can only cope with some of the aspects discussed in this article. Extended versions of the system could possibly include the use of contextual induction of the reference frame using the actual direction of motion of a dynamic object in order to extend the set of possible reference objects to dynamic objects, for example, cars (see How Is the Reference Frame Determined in Extrinsic Use?) (Herskovits 1986; Wunderlich and Herweg 1988).

Another plausible extension could allow for the intrinsic fronts not to be predefined but to be detected by the system using the position of the buildings with respect to streets and squares as well as a more exact representation of the objects, for example, including the location of entrances and outdoor stairs. For this purpose, however, more experimental data about the way people determine intrinsic fronts are needed. Another extension that is planned for the future is the incorporation of the third dimension in order to handle spatial relations in the vertical axis.

Conclusions

Apart from differences in terminology and classification, there is relatively broad consent to the issue of intrinsic and deictic use of spatial prepositions in the linguistic and psychological works, with the exception of Herskovits (1986), who introduces a different system for the classification of the uses of spatial prepositions. Extrinsic use, however, is rarely dealt with in the relevant literature and is not classified the way we do it. Disagreement prevails concerning the issues which prepositions can be used

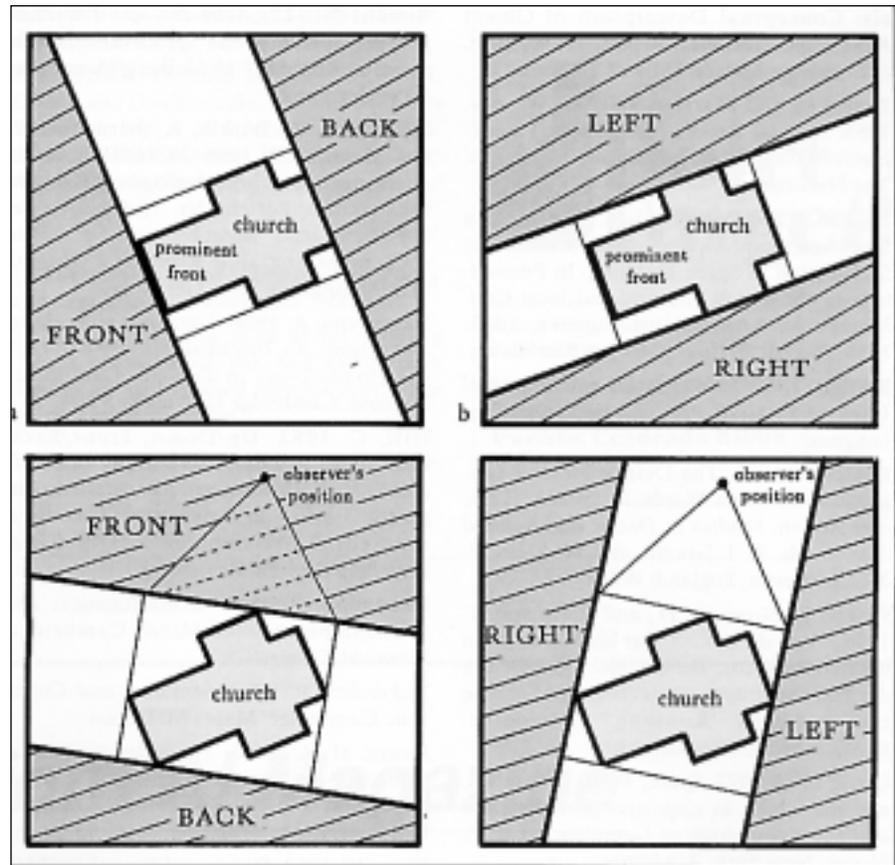


Figure 4. (a) The Front and Back Half-Planes in Intrinsic Use; (b) The Left and Right Half-Planes in Intrinsic Use; (c) The Front and Back Half-Planes in Deictic Use; (d) The Left and Right Half-Planes in Deictic Use.

deictically and intrinsically, in particular, whether prepositions describing spatial relations in the vertical axis can be used deictically, and whether intrinsic or deictic use is the unmarked case, that is, dominates in language understanding and production.

Working AI systems, so far, haven't tried to incorporate many of the results of the linguistic and psychological approaches. Thus, much remains to be done in the area of computational semantics of projective prepositions.

Acknowledgments

The work reported in this article was supported by the Special Collaborative Program on Artificial Intelligence and Knowledge-Based Systems (SFB 314) of the German Science Foundation (DFG), project N2: VITRA. I thank Jürgen Allgayer, Guido Bosch, Sergei Nirenburg, Victor Raskin, Jörg Schirra, Dagmar Schmauks, and Wolfgang Wahlster for their helpful comments on earlier versions of this article. The system CITYTOUR was designed and implemented by Elisabeth André, Guido Bosch,

Gerd Herzog, Thomas Rist, and Ingrid Wellner.

References

- Adorni, G.; Bocalatte, A.; and DiManzo, M. 1981. Object Representation and Spatial Knowledge: An Insight into the Problem of Men-Robots Communication. In Proceedings of the Seventh Conference of the Canadian Man-Computer Communications Society, 37-45: Canadian Man-Computer Communications Society.
- André, E.; Bosch, G.; Herzog, G.; and Rist, T. 1987. Coping with the Intrinsic and Deictic Uses of Spatial Prepositions. In Artificial Intelligence II, Proceedings of the Second International Conference on Artificial Intelligence: Methodology, Systems, Applications, Varna, Bulgaria, eds. P. Jorrand and V. Sgurev, 375-382. Amsterdam: North Holland.
- André, E.; Bosch, G.; Herzog, G.; and Rist, T. 1986. Characterizing Trajectories of Moving Objects Using Natural Language Path Descriptions. In Proceedings of the Seventh European Conference on Artificial Intelligence, vol. 2, 1-8: European Coordinating Committee for Artificial Intelligence.

- Badler, N. I. 1975. Temporal Scene Analysis: Conceptual Description of Object Movements, Technical Report, TR-80, Dept. of Computer Science, Univ. of Toronto.
- Bayer, J., and Marslen-Wilson, W., eds. 1986. Annual Report No. 7, Max-Planck-Institut für Psycholinguistik, Nijmegen, The Netherlands.
- Baykan, C. A., and Fox, M. S. 1987. An Investigation of Opportunistic Constraint Satisfaction in Space Planning. In Proceedings of the Tenth International Joint Conference on Artificial Intelligence, 1035-1038. Los Altos, Calif.: Morgan Kaufmann.
- Bennett, D. C. 1975. Spatial and Temporal Uses of English Prepositions. London: Longman.
- Bühler, K. 1982. The Deictic Field of Language and Deictic Words. In Speech, Place, and Action. Studies in Deixis and Related Topics, eds. R. J. Jarvella and W. Klein, 9-30. Chichester, England: Wiley.
- Bürkle, B.; Nirmaier, H.; and Herrmann, T. 1986. "Von dir aus..." Zur hörerbegrenzten lokalen Referenz, Bericht Nr. 10, Arbeiten der Forschergruppe "Sprechen und Sprache im sozialen Kontext," Heidelberg/Mannheim, West Germany.
- Clark, H. H. 1973. Space, Time, Semantics, and the Child. In Cognitive Development and the Acquisition of Language, ed. T. E. Moore. New York: Academic.
- Cresswell, M. J. 1978. Prepositions and Points of View. *Linguistics and Philosophy* 2: 1-41.
- Dirven, R. 1981. Spatial Relations in English. In Kasusgrammatik und Fremdsprachendidaktik, eds. G. Radde and R. Dirven, 103-132. Trier, West Germany: Wissenschaftlicher Verlag.
- Ehrlich, V. 1982. Da and the System of Spatial Deixis in German. In Here and There. Cross-Linguistic Studies on Deixis and Demonstration, eds. J. Weissenborn and W. Klein. Philadelphia: John Benjamins.
- Fillmore, C. J. 1971. Toward a Theory of Deixis. Paper presented at Pacific Conference on Contrastive Linguistics and Language Universals, Univ. of Hawaii.
- Hays, E. 1987. A Computational Treatment of Locative Relations in Natural Language, Technical Report, MS-CIS-87-31, LINC LAB 58, Dept. of Computer and Information Science, Univ. of Pennsylvania.
- Herrmann, T. 1987. Sprachpsychologische Beiträge zur Partnerbezogenheit des Sprechens, Bericht Nr. 37, Arbeiten der Forschungsgruppe Sprache und Kognition am Lehrstuhl Psychologie III der Universität Mannheim, West Germany.
- Herrmann, T.; Bürkle, B.; and Nirmaier, H. 1987. Zur hörerbegrenzten Raumreferenz: Hörerposition und Lokalisationsaufwand, Bericht Nr. 12, Arbeiten der Forschergruppe "Sprechen und Sprachverstehen im sozialen Kontext," Heidelberg/Mannheim, West Germany.
- Herrmann, T.; Bürkle, B.; Nirmaier, H.; and Mangold, R. 1986. VOHLILIRE: Untersuchungen zur hörerbegrenzten Objektlokalisierung, Bericht Nr. 7, Arbeiten der Forschergruppe "Sprechen und Sprachverstehen im sozialen Kontext," Heidelberg/Mannheim, West Germany.
- Herskovits, A. 1986. Language and Spatial Cognition. An Interdisciplinary Study of the Prepositions in English. Cambridge, England: Cambridge University Press.
- Hill, C. 1982. Up/Down, Front/Back, Left/Right. A Contrastive Study of Hausa and English. In Here and There. Cross-Linguistic Studies on Deixis and Demonstration, eds. J. Weissenborn and W. Klein. Philadelphia: John Benjamins.
- Jackendoff, R. 1987. Consciousness and the Computational Mind. Cambridge, Mass.: MIT Press.
- Jackendoff, R. 1983. Semantics and Cognition. Cambridge, Mass.: MIT Press.
- Kautz, H. A. 1985. Formalizing Spatial Concepts and Spatial Language, Technical Report, CSLI-85-35, Center for the Study of Language and Information, Stanford Univ.
- Klein, W. 1983. Deixis and Spatial Orientation in Route Directions. In Spatial Orientation: Theory, Research, and Application, eds. H. Pick and L. Acredolo, 283-311. New York: Plenum.
- Klein, W. 1982. Local Deixis in Route Directions. In Speech, Place, and Action: Studies in Deixis and Related Topics, eds. R. J. Jarvella and W. Klein, 161-182. Chichester, England: Wiley.
- Klein, W. 1978. Wo ist hier: Präliminarien zu einer Untersuchung der lokalen Deixis. *Linguistische Berichte* 58.
- Lakoff, G. 1987. Women, Fire, and Dangerous Things: What Categories Reveal about the Mind. Chicago: University of Chicago.
- Langacker, R. 1986. Foundations of Cognitive Grammar, vol. 1. Stanford, Calif.: Stanford University Press.
- LeFaivre, R. 1977. Fuzzy Reference Manual, Computer Science Dept., Rutgers Univ.
- Levelt, W. J. M. 1986. Zur sprachlichen Abbildung des Raumes: Deiktische und intrinsische Perspektive. In Perspektiven auf Sprache. Interdisziplinäre Beiträge zum Gedenken an Hans Hörmann, ed. H.-G. Bosshardt, 187-211. Berlin: de Gruyter.
- Levelt, W. J. M. 1982. Cognitive Styles in the Use of Spatial Direction Terms. In Speech, Place, and Action: Studies in Deixis and Related Topics, eds. R. J. Jarvella and W. Klein, 251-268. Chichester, UK: Wiley.
- Lyons, J. 1977. Semantics, vol. 2. Cambridge, England: Cambridge University Press.
- Miller, G. A., and Johnson-Laird, P. N. 1976. Language and Perception. Cambridge, England: Cambridge University Press.
- Nirenburg, S., and Raskin, V. 1987. Dealing with Space in Natural Language Processing. In Proceedings of the 1987 Workshop on Spatial Reasoning and Multi-Sensor Fusion, eds. A. Kak and S.-S. Chen, 361-370. Los Altos, Calif.: Morgan Kaufmann.
- Rauh, G. 1984. Aspekte der Deixis II: Deiktische Dimensionen und die Verwendung deiktischer Ausdrücke, A 123, Duisburg, West Germany: Linguistic Agency, University of Duisberg.
- Reichenbach, H. 1947. Elements of Symbolic Logic. New York: MacMillan.
- Saile, G. 1984. Sprache und Handlung. Braunschweig, West Germany: Vieweg.
- Schirra, J. R. J.; Bosch, G.; Sung, C.K.; and Zimmermann, G. 1987. From Image Sequences to Natural Language: A First Step toward Automatic Perception and Description of Motions. *Applied Artificial Intelligence* 1: 287-305.
- Schulze, R. 1987. The Perception of Space and the Function of Prepositions in English: A Contribution to Cognitive Grammar. In Perspectives on Language in Performance, eds. W. Lörcher and R. Schulze. Tübingen, West Germany: Narr.
- Sennholz, K. 1985. Grundzüge der Deixis. Bochum, West Germany: Brockmeyer.
- Sondheimer, N. K. 1976. Spatial Reference and Natural-Language Machine Control. *International Journal of Man-Machine Studies* 8: 329-336.
- Sondheimer, N. K. 1974. English as a Command Language for Machines and the Semantics of "Left" and "Right," Milwaukee Symposium on Automatic Control (and Autonomous Computing), Milwaukee, Wisc.
- Sung, C.-K., and Zimmermann, G. 1986. Detektion und Verfolgung mehrerer Objekte in Bildfolgen. In Mustererkennung 1986, 8. DAGM-Symposium, Paderborn, West Germany, ed. G. Hartmann, 181-184. Berlin: Springer.
- Talmy, L. 1983. How Language Structures Space. In Spatial Orientation: Theory, Research, and Application, eds. H. Pick and L. Acredolo, 225-282. New York: Plenum.
- Tsotsos, J. K.; Mylopoulos, J.; Covey, H. D.; and Zucker, S. W. 1980. A Framework for Visual Motion Understanding. *IEEE Transactions on Pattern Analysis and*

Machine Intelligence 2(6).

Ullmer-Ehrich, V. 1982. The Structure of Living Space Descriptions. In *Speech, Place, and Action: Studies in Deixis and Related Topics*, eds. R. J. Jarvella and W. Klein, 219-249. Chichester, England: Wiley.

Vandeloise, C. 1984. Description of Space in French. Ph.D. diss., Dept. of Linguistics, Univ. of California at San Diego.

Wunderlich, D. 1986. Raum und die Struktur des Lexikons. In *Perspektiven auf Sprache. Interdisziplinäre Beiträge zum Gedenken an Hans Hörmann*, ed. H-G. Bosshardt, 212-231. Berlin: de Gruyter.

Wunderlich, D. 1985. Raumkonzepte. Zur Semantik der lokalen Präpositionen. In *Nach-Chomskysche Linguistik. Neuere Arbeiten von Berliner Linguisten*, eds. T. T.

Ballmer and R. Posner, 340-351. Berlin: de Gruyter.

Wunderlich, D., and Herweg, M. 1988. Lokale und Direktionale. In *Handbuch der Semantik*, eds. A. V. Stechow and D. Wunderlich. Königstein, West Germany: Athenäum. Forthcoming.

Notes

1. The system is implemented in Fuzzy and Zeta Lisp on a Symbolics 3600.
2. So far, only the data about moving objects are provided; the data about static objects are still entered manually with the aid of a graphics editor.
3. Strictly speaking, it is the center of gravity of the delineative rectangle.

Announcing

Readings from AI Magazine

Volumes One Through Five

The First Five Years: 1980-1985

Edited with a Preface by Robert Englemore

AAAI is pleased to announce forthcoming publication of *Readings from AI Magazine*, the complete collection of all the articles that appeared during *AI Magazine's* first five years of publication. Within this 650 page indexed volume, you will find articles on AI, written by the foremost practitioners in the field—articles that earned *AI Magazine* the title "journal of record for the artificial intelligence community." And now, this collection of classics from the oldest and premier publication devoted to the entire field of artificial intelligence is available in one large, paperbackbound desktop reference.

Subjects Include:

- Automatic Programming
- Distributed Artificial Intelligence
- Games
- Infrastructure
- Learning
- Natural Language Understanding
- Problem Solving
- Robotics
- Computer Architectures
- Education
- General Artificial Intelligence
- Knowledge Acquisition
- Legal Issues
- Object Oriented Programming
- Programming Language
- Simulation
- Technology Transfer
- Discovery
- Expert Systems
- Historical Perspectives
- Knowledge Representation
- Logic
- Partial Evaluation
- Reasoning with Uncertainty
- Surveys

To be published Fall, 1988. Prepublication price (good through 31 August 1988): \$60.00 (postpaid); \$74.95 postpaid thereafter. Approximate 650 pages. Illus., appendix, author-title-subject index. ISBN 0-932236-00-0 Available only from the American Association for Artificial Intelligence. 445 Burgess Drive, Menlo Park, California 94025.