

## Chapter 3

---

### Perspective Taking and Ellipsis in Spatial Descriptions

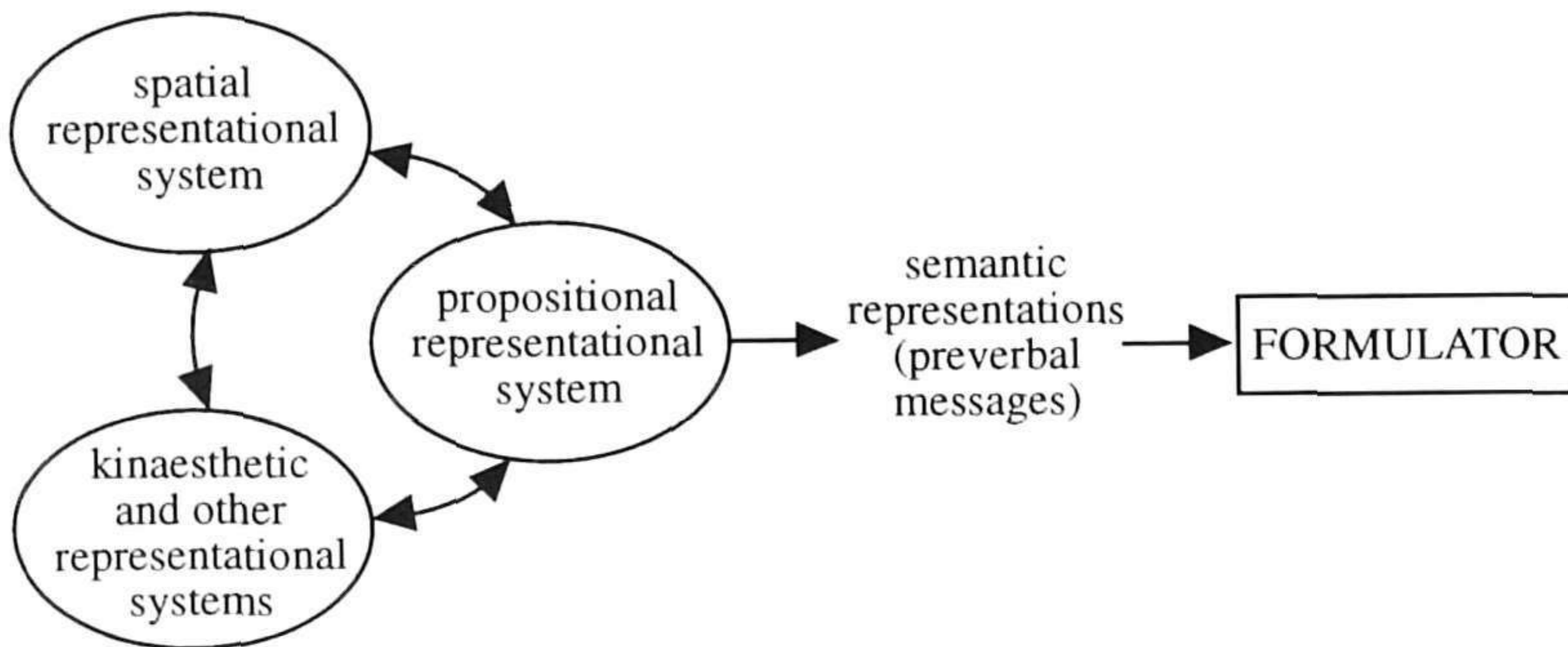
Willem J. M. Levelt

#### 3.1 Thinking for Speaking

There exists happy agreement among students of language production that speaking normally involves a stage of conceptual preparation. Depending on the communicative situation, we decide in some way or another on *what* to express. Ideally, this choice of content will eventually make our communicative intention recognizable to our audience or interlocutor. The result of conceptual preparation is technically termed a *message* (or a string of messages); it is the conceptual entity the speaker will eventually express in language, that is, formulate.

But there is more to conceptual preparation than considering what to say, or *macroplanning*. There is also *microplanning*. The message has to be of a particular kind; it has to be tuned to the target language and to the momentary informational needs of the addressee. This chapter is about an aspect of microplanning that is of paramount importance for spatial discourse, namely *perspective taking*.

In an effort to cope with the alarming complexities of conceptual preparation, I presented a figure in my book *Speaking* (1989) that is reproduced here as figure 3.1. It is intended to express the claim that messages must be in some kind of propositional or “algebraic” format (cf. Jackendoff, chapter 1, this volume) to be suitable for formulation. In particular, they must be composed out of *lexical concepts*, that is, concepts for which there are words or morphemes in the speaker’s language. An immediate corollary of this notion is that conceptual preparation will, to some extent, be specific to the target language. Lexical concepts differ from language to language. A lexical concept in one language may be nonlexical in another and will therefore need a slightly different message to be expressed. To give one spatial example (from Levelt 1989), there are languages such as Spanish or Japanese that treat deictic proximity in a tripartite way: proximal-medial-distal. Other languages, such as English or Dutch, have a bipartite system, proximal-distal. Spanish use of *aquí-ahí-allí* requires to construe distance from speaker in a different way than English use of *here-there*.



**Figure 3.1**

The mind harbors multiple representational systems that can mutually interact. But to formulate any representation linguistically requires its translation into a semantic, “propositional” code (reproduced from Levelt 1989).

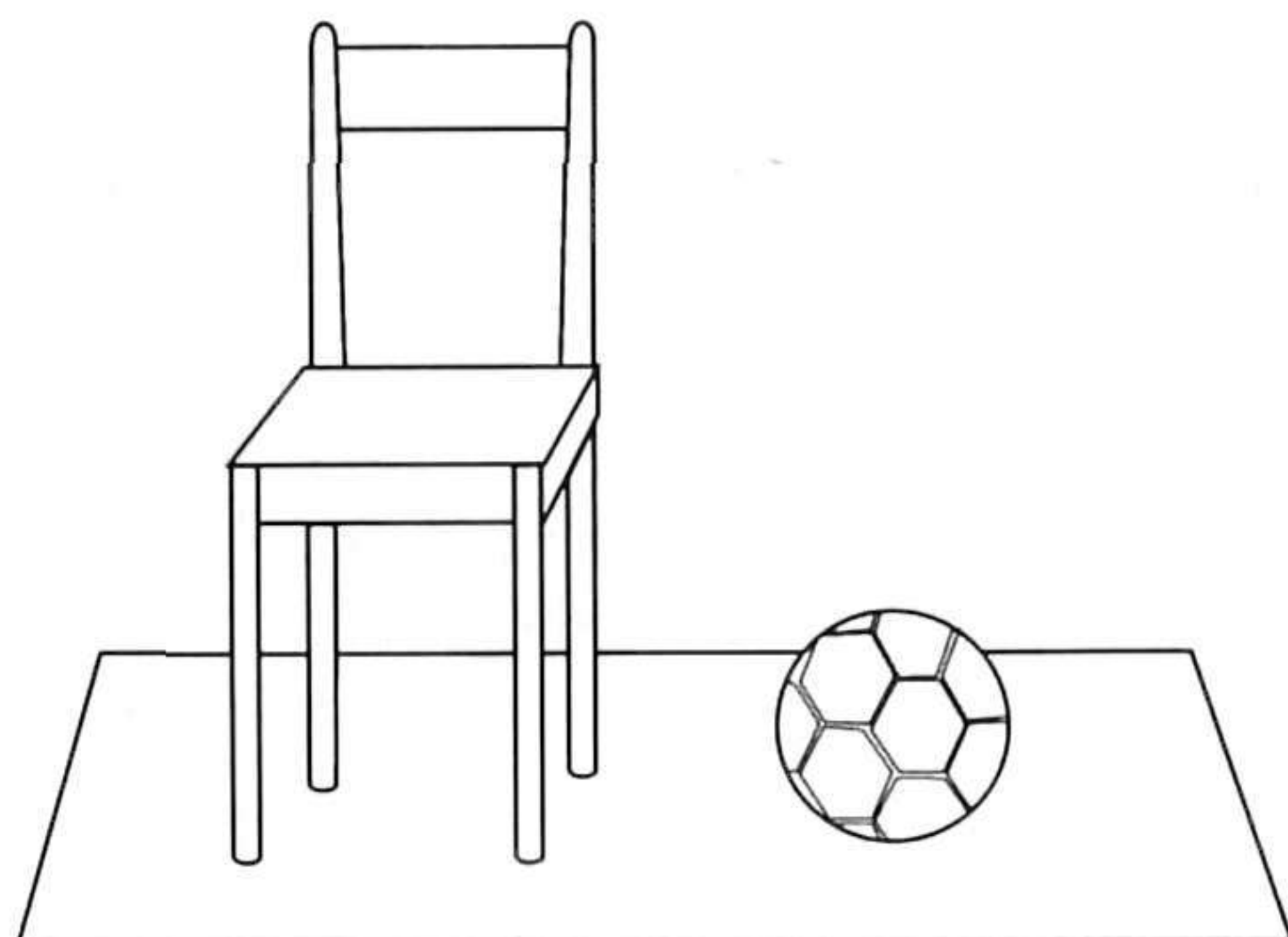
Slobin (1987) has usefully called this “thinking for speaking,” which is an elegant synonym for microplanning.

Thinking for speaking is always involved when we express nonpropositional, in particular spatial, information. Figure 3.1 depicts the notion that when we talk about our spatial, kinesthetic, musical, and so on experiences, we cast them in propositional form. This necessarily requires an act of abstraction. When talking about a visual scene, for instance, we attend to entities that are relevant to the communicative task at hand, and generate predications about these entities that accurately capture their spatial relations within the scene. This process of abstracting from the visual scene for speaking I will call “perspective taking.” Although this term will in the present chapter be restricted to its original spatial domain, it is easily and fruitfully generalized to other domains of discourse (cf. Levelt 1989).

### 3.2 Perspective Taking

Perspective taking as a process of abstracting spatial relations for expression in language typically involves the following operations:

1. Focusing on some portion of the scene whose spatial disposition (place, path, orientation) is to be expressed (Talmy 1983). I will call this portion the “referent.”
2. Focusing on some portion of the field with respect to which the referent’s spatial disposition is to be expressed. I will call this portion the “relatum.”
3. Spatially relating the referent to the relatum (or expressing the referent’s path or orientation) in terms of what I will call a “perspective system.”



**Figure 3.2**

This spatial array can be described in myriad ways, depending on the choice of referent, relatum, and perspective.

Let me exemplify this by means of figure 3.2. One way of describing this scene is

(1) I see a chair and a ball to the right of it.

Here the speaker introduces the chair as the relatum and then expresses the spatial disposition of the ball (to the right of the chair). Hence, the ball is the referent. The perspective system in terms of which the relating is done is the *deictic* system, that is, a speaker-centered *relative* system.<sup>1</sup> When you focus on the relatum (the chair), your gaze must turn to your right in order to focus on the referent (the ball). That is why the ball is to the right of the chair in this system.

Two things are worth noticing now. First, you can swap relatum and referent, as in (2):

(2) I see a ball and a chair to the left of it.

This is an equally valid description of the scene; it is only a less preferred one. Speakers tend to select smaller and more foregrounded objects as referents and larger or more backgrounded entities as relata. Here they tend to follow the Gestalt organization of the scene (Levelt 1989). Second, you can take another perspective system. You can also describe the scene as (3):

(3) I see a chair and a ball to its left.

This description is valid in the *intrinsic* perspective system. Here the referent's location is expressed in terms the relatum's intrinsic axes. A chair has a front and a back, a left and a right side. The ball in figure 3.2 is at the chair's left side, no matter from which viewpoint the speaker is observing the scene. Still another perspective system allows for the description in (4):

(4) I see a chair and a ball north of it.

This description is valid if indeed ball and chair are aligned on a north-south dimension. This is termed an *absolute* system; it is neither relative to the speaker's nor to the relatum's coordinate system, but rather to a fixed bearing.

The implication of these two observations is that perspective is linguistically free. There is no unique way of perspective taking. There is no biologically determined one-to-one mapping of spatial relations in a visual scene to semantic relations in a linguistic description of that scene. And cultures have taken different options here, as Levinson and Brown have demonstrated (Levinson 1992a,b; Brown and Levinson 1993). Speakers of Guugu Yimithirr are exclusive users of an absolute perspective system, Mopan speakers are exclusive users of an intrinsic system, Tzeltal uses a mix of absolute and intrinsic perspectives, and English uses all three systems. Similarly, there are personal style differences between speakers of the same language. Levelt (1982b) found that, on the same task, some speakers consistently use a deictic system whereas others consistently use an intrinsic perspective system. Finally, the same speaker may prefer one system for one purpose and another system for another purpose as Tversky (1991) and Herrmann and Grabowski (1994) have shown.

This freedom of perspective taking does not mean, however, that the choice of a perspective system is arbitrary. Each perspective system has its specific advantages and disadvantages in language use, and these will affect a culture's or a speaker's choice. In other words, there is a *pragmatics* of perspective systems.

In the rest of this chapter I will address two issues. The first one is pragmatics. I will compare some advantages and disadvantages in using the three systems introduced above; the deictic, the intrinsic, and the absolute systems. In particular, I will ask how suitable these systems are for spatial reasoning, how hard or easy they are to align between interlocutors, and to what extent the systems are mutually interactive.

The second issue goes back to figure 3.1 and to "thinking for speaking." I defined perspective taking as a speaker's mapping of a spatial representation onto a propositional (or *semantic*) representation for the purpose of expressing it in language. A crucially important question now is whether the spatial representations themselves are already "tuned to language." For instance, a speaker of Guugu Yimithirr, who exclusively uses absolute perspective, may well have developed the habit of representing *any* spatial state of affairs in an oriented way, whether for language or not. After all, *any* spatial scene may become the topic of discourse at a different place and time. The speaker should then have remembered the scene's absolute orientation. Levinson (1992b) presents experimental evidence that this is indeed the case. On the other hand, I argued above that perspective is free. A speaker is not "at the mercy" of a spatial representation in thinking for speaking. In the strongest non-Whorfian

case, spatial representations will be language-independent, and it is perspective taking that maps them onto language-specific semantic representations. One way of sorting this out is to study how speakers operate when they produce *spatial ellipsis* (such as in *go right to blue and then  $\emptyset$  to purple*, here  $\emptyset$  marks the position where a second occurrence of *right* is elided). I will specifically ask whether ellipsis is generated from a perspectivized or from a perspective-free representation. If the latter turned out to be the case, that would plead for the existence of perspective-free spatial representations.

### 3.3 Some Properties of Deictic, Intrinsic, and Absolute Perspective

Of many aspects that may be relevant for the use of perspective systems I will discuss the following three: (1) their inferential potential, (2) their ease of coordination between interlocutors, and (3) their mutual support or interference.

#### 3.3.1 Inferential Potential

Spatial reasoning abounds in daily life (cf. Byrne and Johnson-Laird 1989; Tversky 1991). Following road directions, equipment assembly instructions, spatial search instructions, or being involved in spatial planning discourse all require the ability to infer spatial layouts from linguistic description. And the potential for spatial inference is crucially dependent on the perspective system being used. In Levelt (1984) I analyzed some essential logical properties of the deictic and intrinsic systems; I will summarize them here and extend the analysis to the absolute system.

**Converseness** An attractive logical property is converseness. Perspective systems usually (though not always) involve directional opposites, such as front-back, above-below, north-south. If the two-place relation expressed by one pole is called  $R$  and the one by the other pole by  $R^{-1}$ , then converseness holds if  $R(A, B) \Rightarrow R^{-1}(B, A)$ . For instance, if object  $A$  is above object  $B$ ,  $B$  will be below  $A$ .

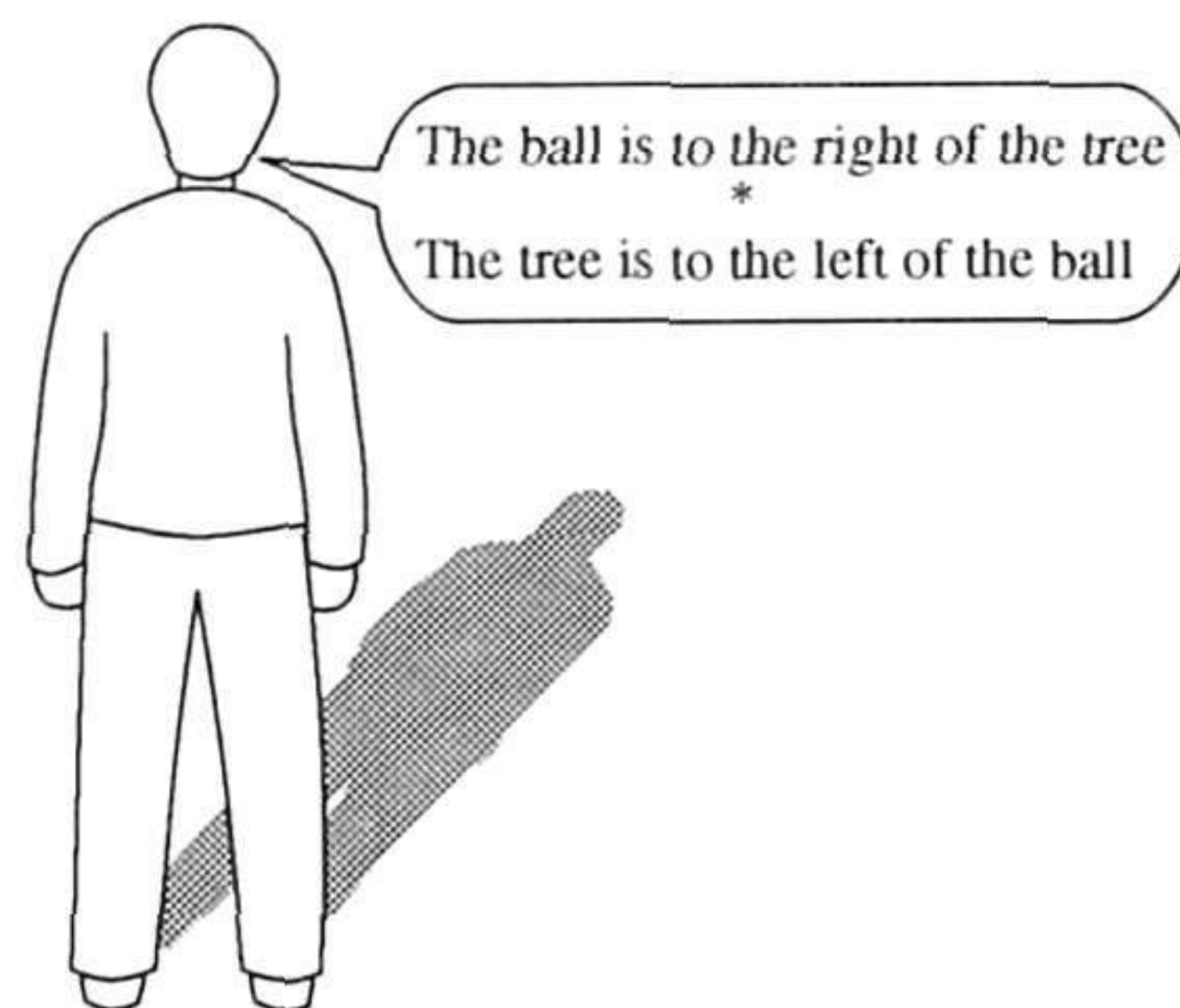
Converseness holds for the deictic system and for most cases<sup>2</sup> of the absolute system, but not for the intrinsic system. This is demonstrated in figure 3.3. Assuming that it is about noon somewhere in the Northern Hemisphere with the sun shining, the shadows of the tree and ball indicate that the ball is east of the tree. Using this absolute bearing, the tree must be west of the ball, where west is the converse of east. Converseness also holds for the (three-place) deictic relation. From the speaker's point of view, the ball (referent) is to the right of the tree (relatum), which necessarily implies that the tree (referent) is to the left of the ball (relatum). But it is easy to violate converseness for the intrinsic system. The ape can be on the right side ("to the right") of the bear at the same time the bear is on the right side ("to the right") of the

## Absolute

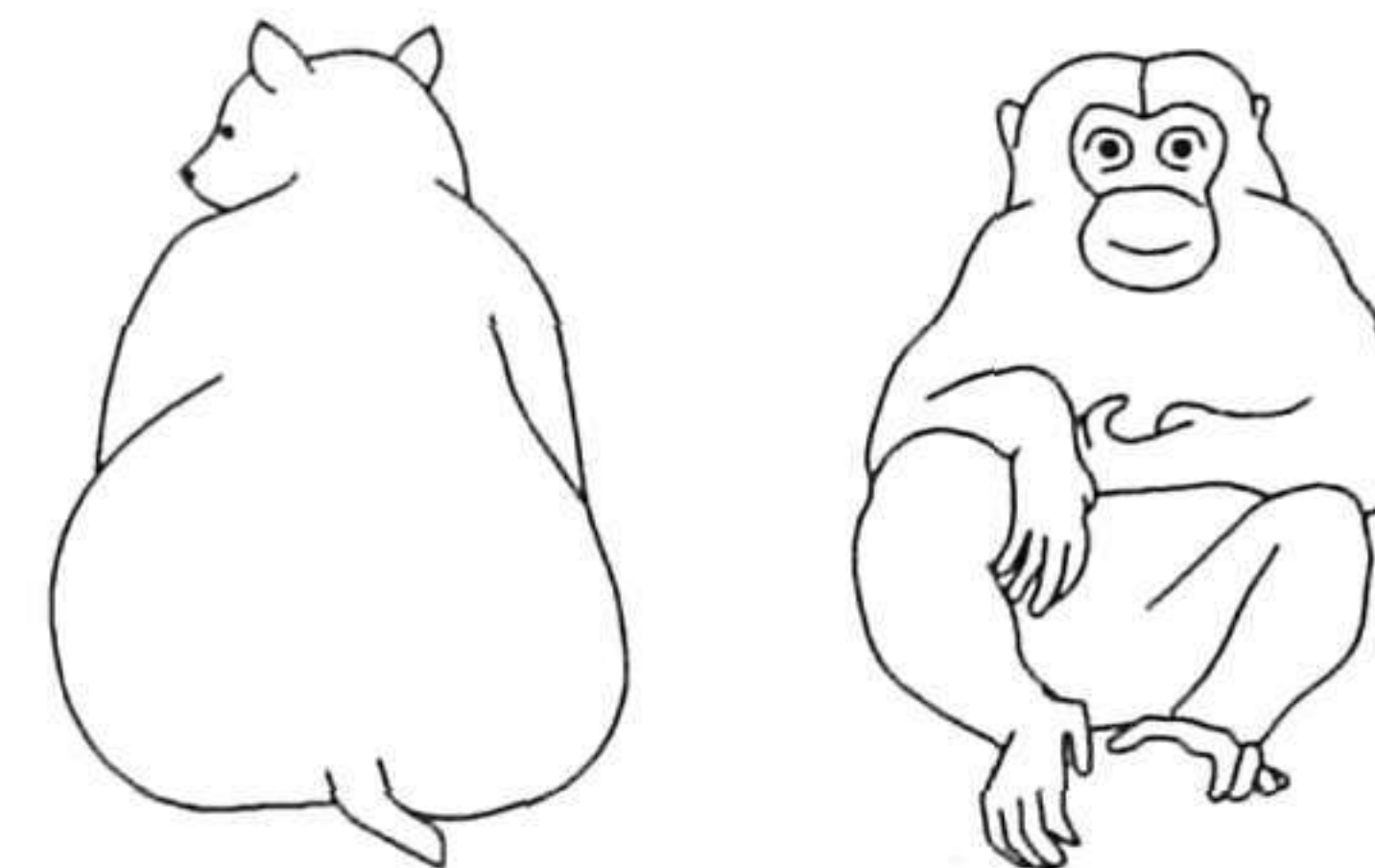


The ball is East of the tree  
\*  
The tree is West of the ball

## Deictic



## Intrinsic



The ape is to the right of the bear  
\*  
The bear is to the right of the ape

**Figure 3.3**

Converseness holds for the absolute and deictic systems, but not for the intrinsic system.

ape. It is therefore impossible to infer the relation between relatum and referent from the relation between referent and relatum in the intrinsic system, which is a major drawback for spatial reasoning.

**Transitivity** Transitivity holds if from  $R(A, B)$  and  $R(B, C)$ , it follows that  $R(A, C)$ . This is the case for the absolute and deictic systems, but not for the intrinsic system. This state of affairs is demonstrated in figure 3.4. The flag, tree, and ball scene depicts the transitivity of “east of” in the absolute system and of “to the right of” in the deictic system. For the intrinsic system it is easy to construct a case that violates transitivity. This is the case for the bear, cow, and ape scene. The user of an intrinsic system cannot rely on transitivity. From  $A$  is to the right of  $B$ , and  $B$  is to the right of  $C$ , one cannot reliably conclude that  $A$  is to the right of  $C$ , and so forth. Hence one cannot create a chain of inference, using the previous referent as a relatum for the next one.

These are serious drawbacks of the intrinsic system. Converseness and transitivity are very desirable properties if you want to make inferences from spatial premises. And spatial reasoning abounds in everyday discourse, for instance, in following route directions, in jointly planning furniture arrangements or equipment assembly, and so on. I will shortly discuss further drawbacks of the intrinsic system for spatial reasoning.

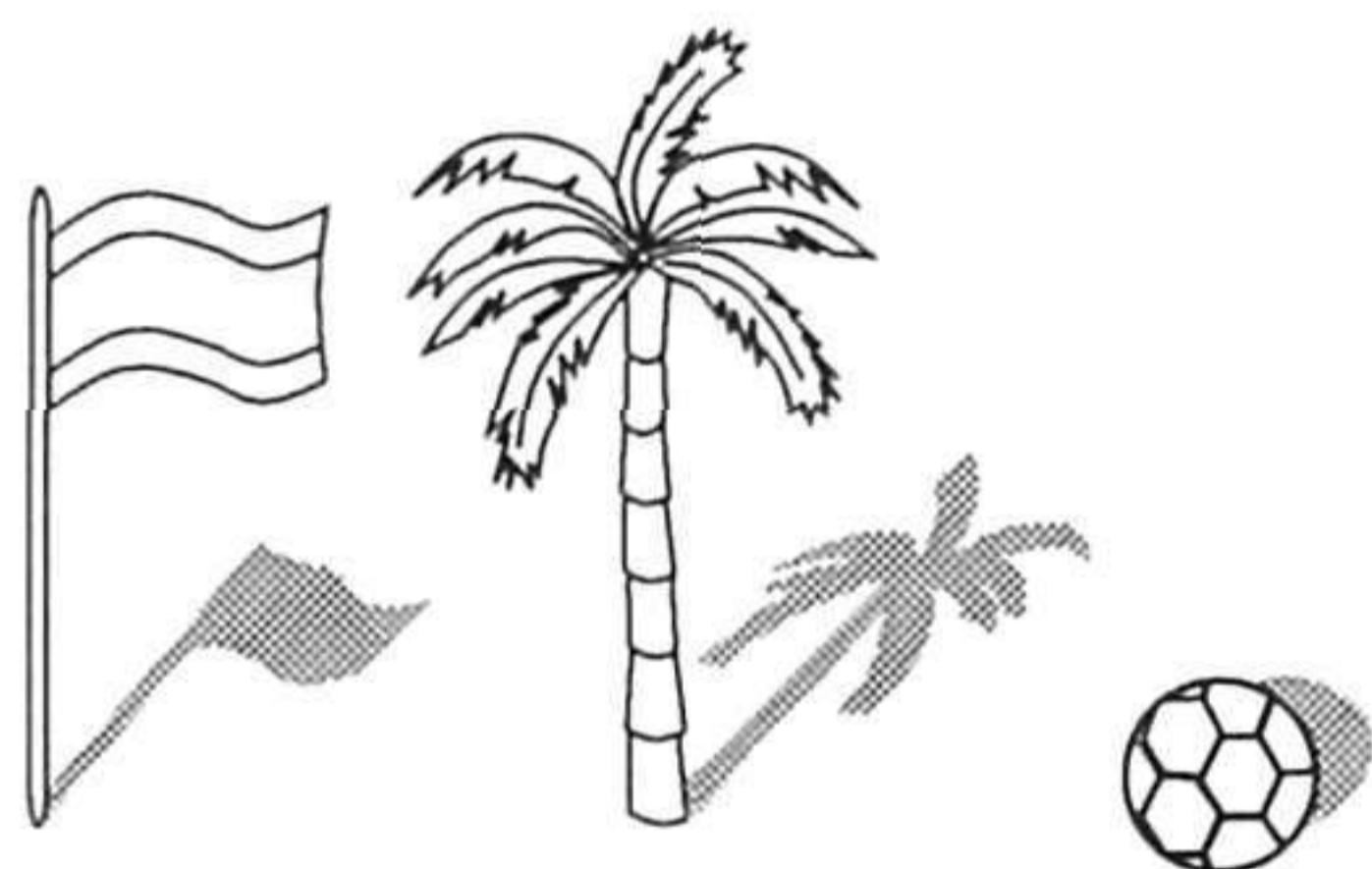
### 3.3.2 Coordination between Interlocutors

It is more the exception than the rule that interlocutors make explicit reference to the perspective system they employ in spatial discourse (for references and discussion, see Levelt 1989, 51). Usually there is tacit agreement about the system used, but not always. An example of nonagreement turned up in an experiment where I asked subjects to describe colored dot patterns in such a way that other subjects would be able to draw them from the tape-recorded descriptions. An example of such a pattern is presented in figure 3.5. Subjects were instructed to start at the arrow. It turned out that most subjects used deictic perspective. A typical deictic description of this pattern is the following:

Begin with a yellow dot. Then one step up is a green dot and further up is a brown dot. Then right to a blue dot and from there further right to a purple dot. Then one step down there is a red dot. And left of it is a black one.

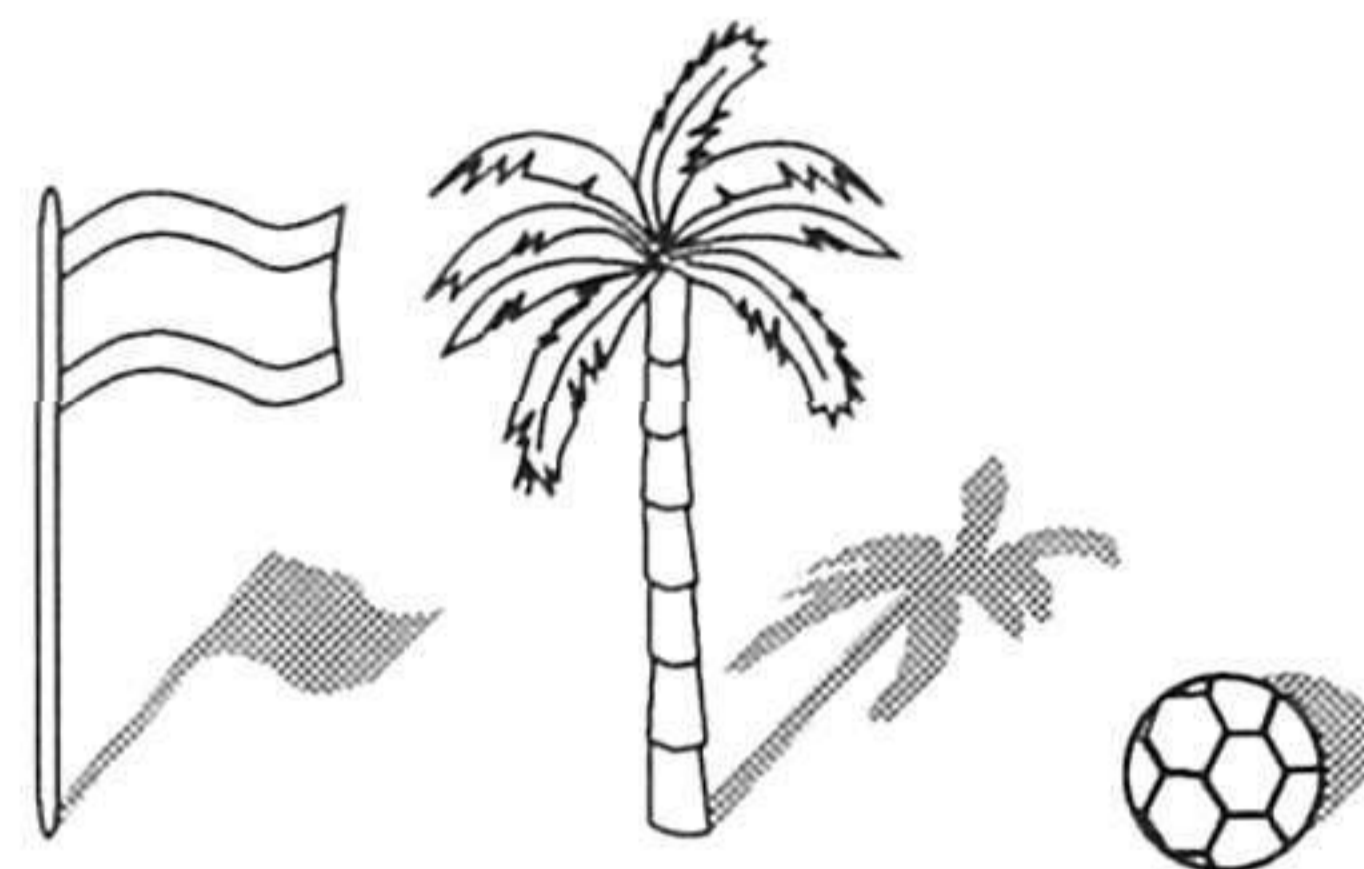
Although the dot pattern was always flat on the table in front of the subject, moves toward and away from the subject were typically expressed by vertical dimension terms (*up*, *down*). This is characteristic for deictic perspective, because it is viewer-centered. It essentially tells you where the gaze moves (see Levelt 1982b; Shepard and Hurwitz 1984). For the pattern in figure 3.5, the gaze moves up, up, right, right,

# Absolute

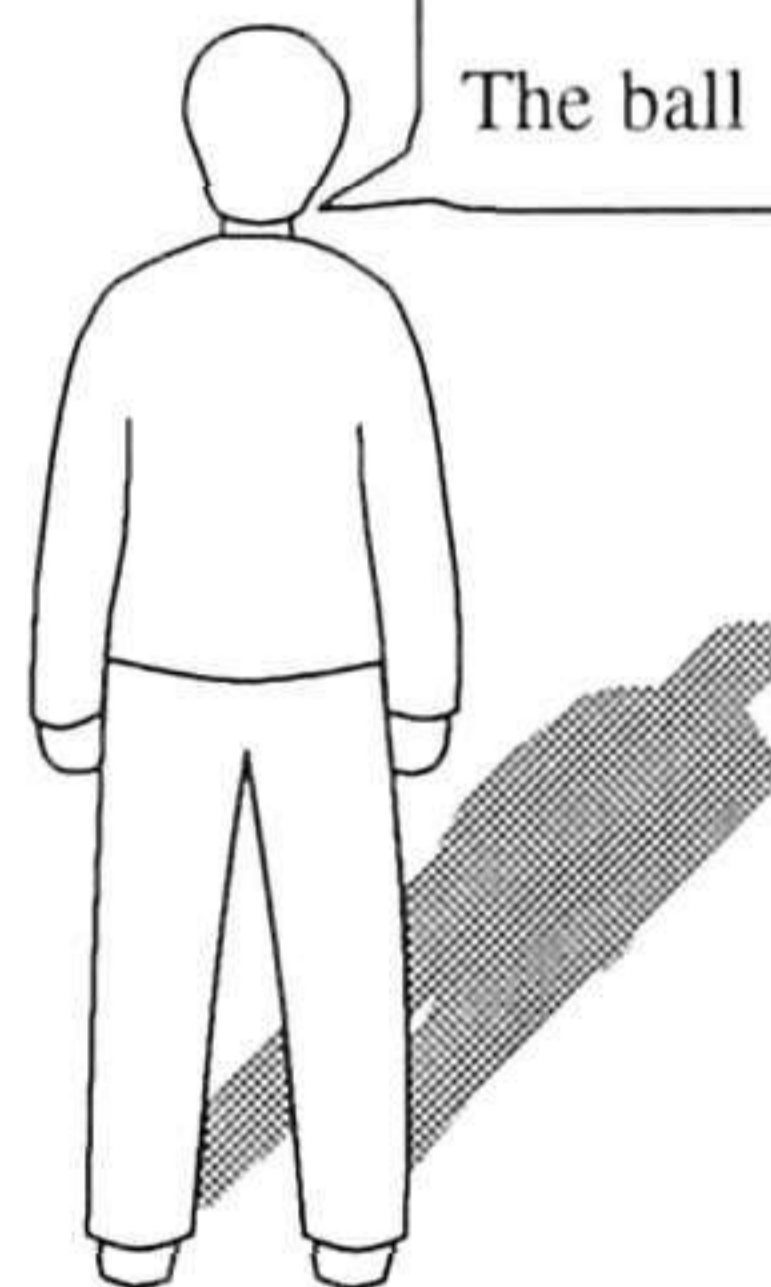


The ball is East of the tree  
 The tree is East of the flag  
 \*  
 The ball is East of the flag

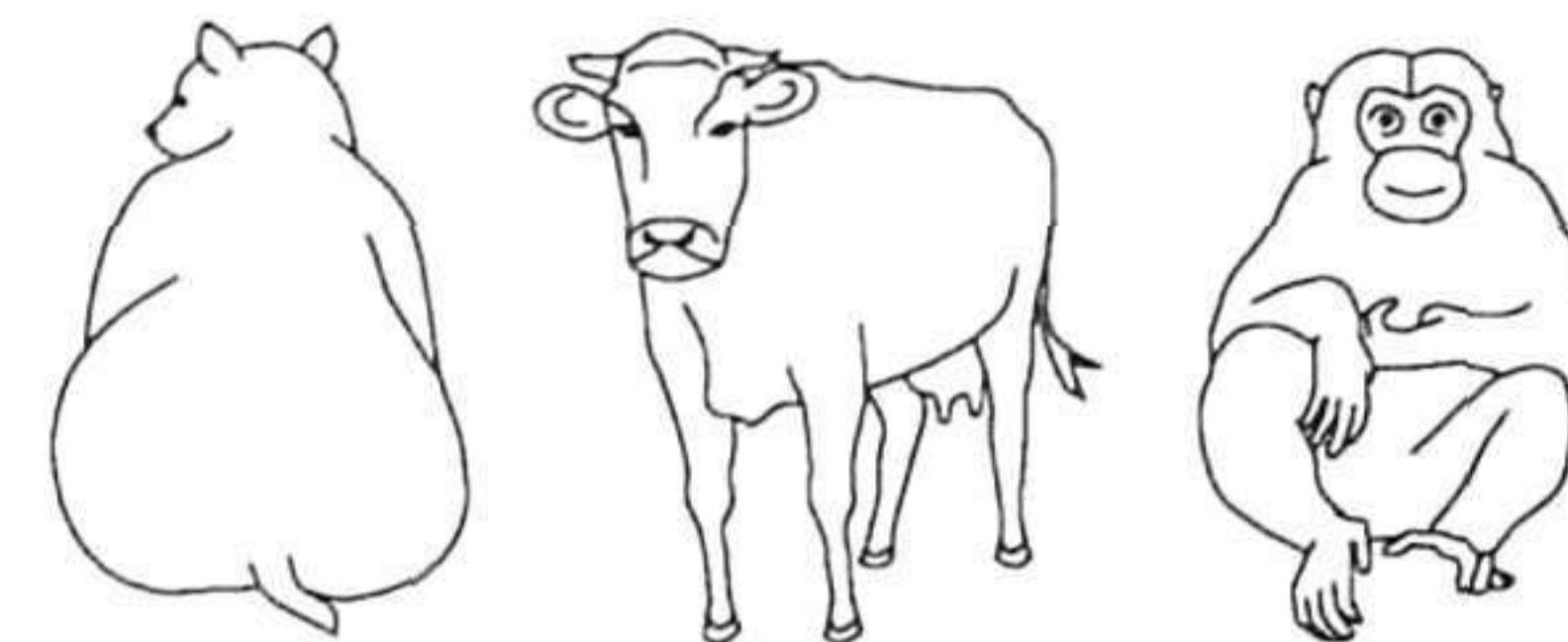
# Deictic



The ball is to the right of the tree  
 The tree is to the right of the flag  
 \*  
 The ball is to the right of the flag



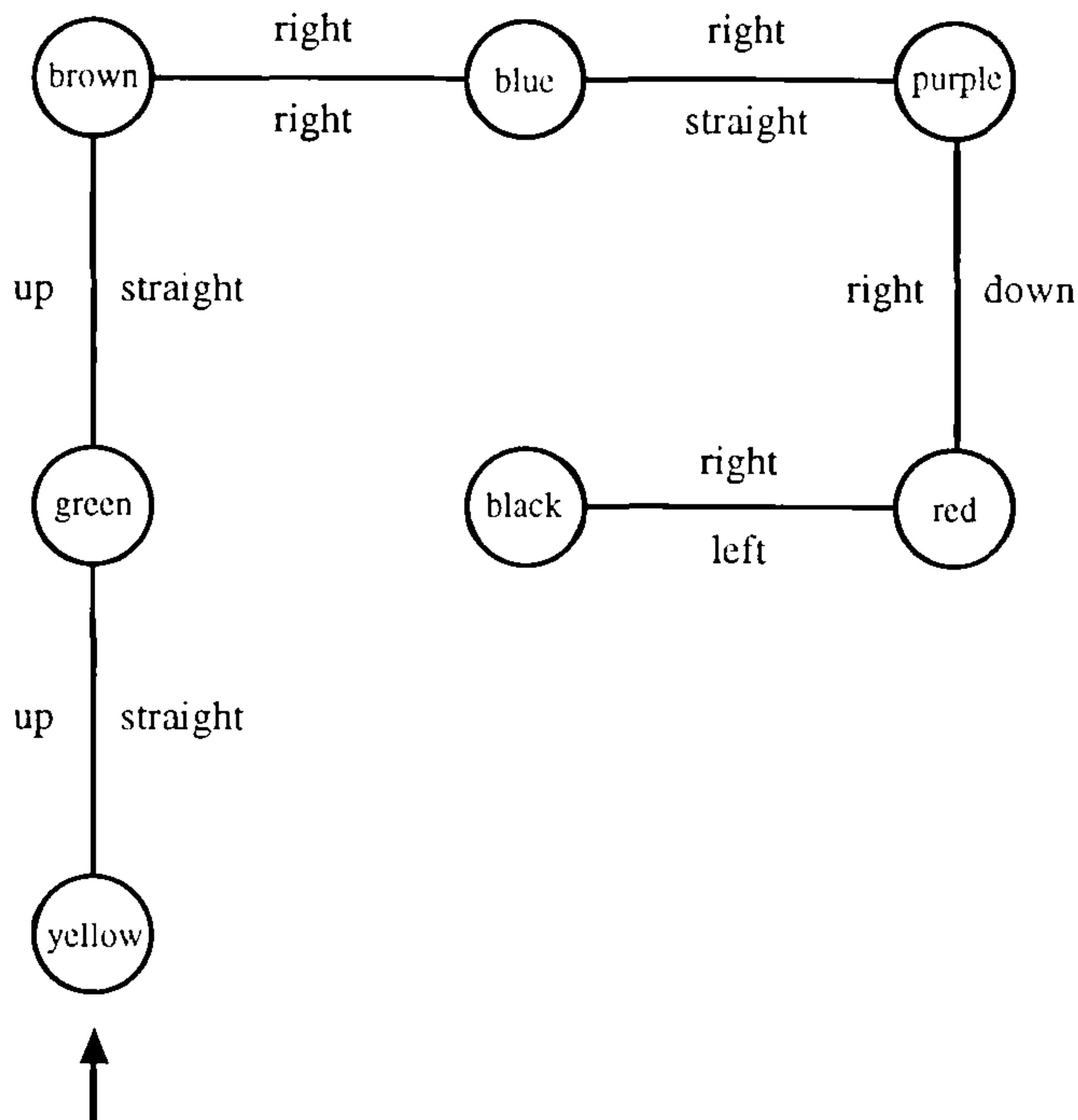
# Intrinsic



The ape is to the right of the bear  
 The bear is to the right of the cow  
 \*  
 The ape is to the left of the cow

**Figure 3.4**  
 Transitivity holds for the absolute and deictic systems, but not for the intrinsic system.





**Figure 3.5**

Pattern used in a spatial description task. The nodes were colored (here replaced by color names). On the outside of the arcs are the dominant directional terms used in deictic descriptions; on the inside, the ones used dominantly in intrinsic descriptions.

down, and left. These directional terms in the description are depicted at the exterior side of the pattern. Notice that all terms would have been different if the pattern had been turned by 90 degrees.

But other subjects used the intrinsic system. They described the scene as if they were moving through it or leading you through it. This is a typical intrinsic description.<sup>3</sup>

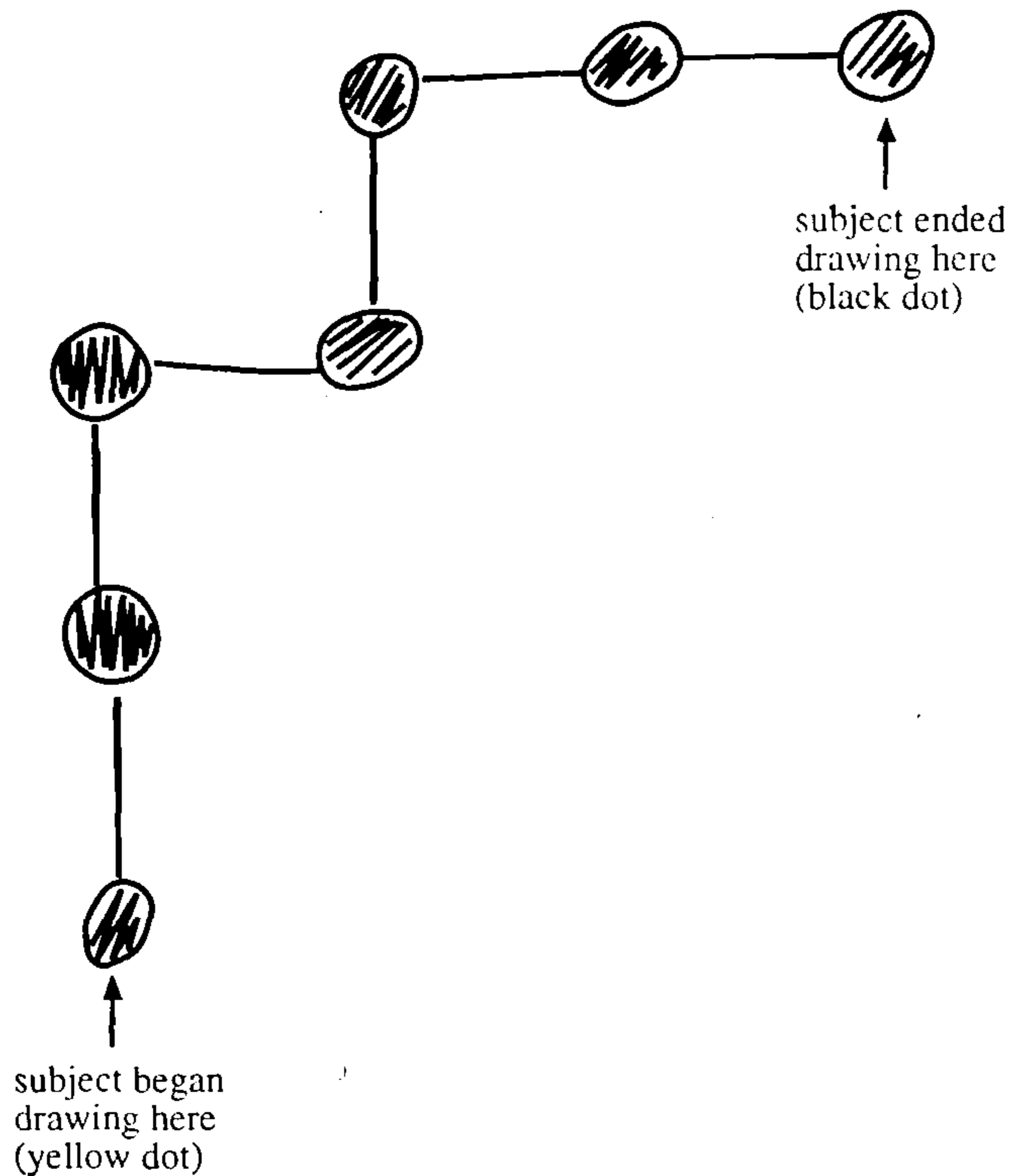
You start at a yellow point. Then go straight to a green dot and straight again to brown. Now turn right to a blue dot and from there straight to a purple dot. From there turn right to red and again right to a black dot.

There are no vertical dimension terms here. The description is not viewer-centered, but derives from the intrinsic directions of the pattern itself; the directional terms

would still be valid if the pattern were turned by 90 degrees. The interior of figure 3.5 depicts the directional terms used in this intrinsic description.

When I gave the deictic descriptions to subjects for drawing, they usually reproduced the pattern correctly. But when I presented the intrinsic description, subjects' drawings tended to be incorrect, and systematically so. Most reproductions are like the one in figure 3.6, which is a typical example. What has happened here is obvious. The listener tacitly assumes a deictic perspective and forces the intrinsic description into this deictic Procrustean bed. The incongruent term *straight* is interpreted as "up." This, then, is a case of failing speaker/hearer coordination.

Coordination failures can be of different kinds. In this example the listener tacitly assumes one perspective system where the speaker has in fact used a different one. Our deictic and intrinsic systems are subject to this confusion because many of the



**Figure 3.6**

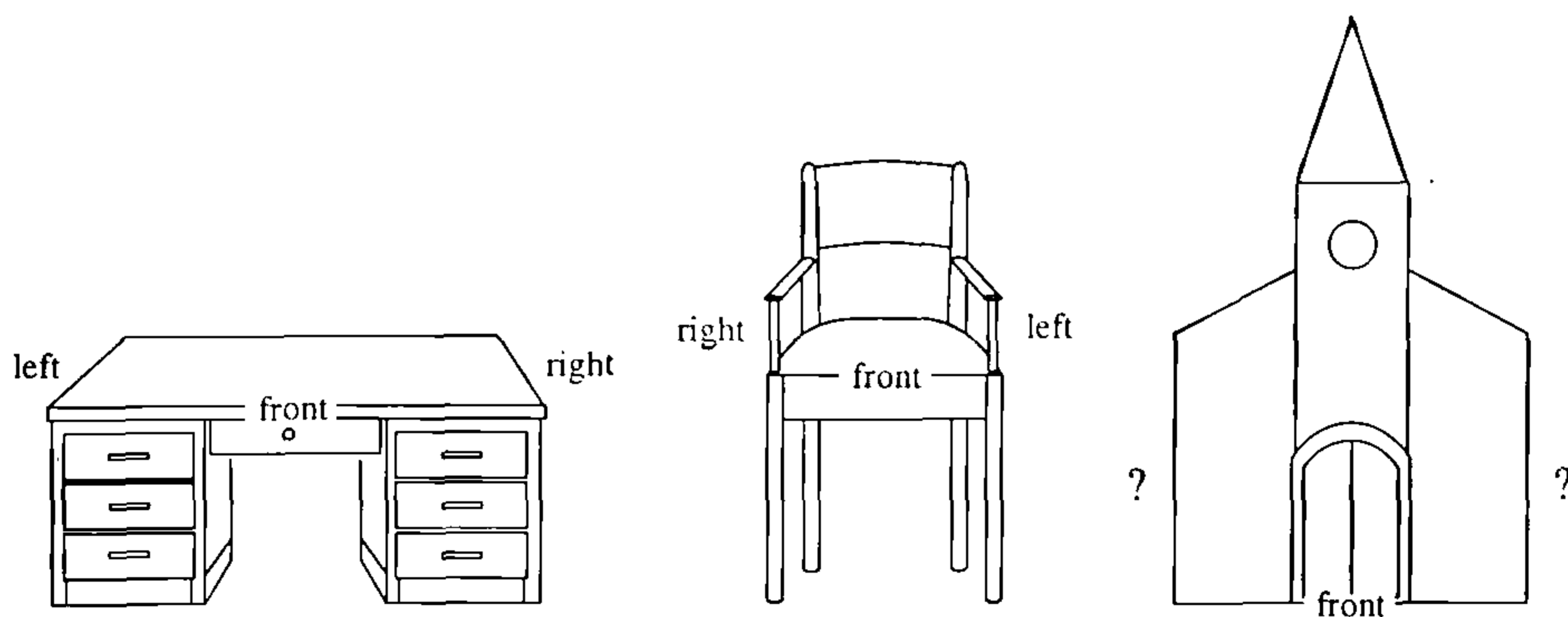
A subject's reconstruction of the pattern in figure 3.5 from its intrinsic description.

dimensional terms are the same or similar in the two systems. But also within the same perspective system coordination failure can arise.

For the deictic system, a major problem in coordination is that the system derives from the speaker's viewpoint, that is, the speaker's position and orientation in the scene. And because the viewpoints are never fully shared, there is continuous switching back and forth in conversation between the coordinate systems of the interlocutors. The interlocutors must keep track of their partners' viewpoints throughout spatial discourse.

This contrasts with the intrinsic and absolute systems, which are speaker-independent. The intrinsic system, however, requires that the interlocutor is aware of the relatum's orientation. The utterance *the ball is to the right of the chair* can only effectively localize the ball for the interlocutor if not only the chair's position is known, but also its orientation. In a perceptual scene, therefore, the intrinsic system requires *recognition* of the **relatum** on the part of the listener, not only awareness of its localization.

The felicity of speaker/hearer coordination in the intrinsic system is, therefore, crucially dependent on the shared image of the relatum. First, coordination in the intrinsic system is only possible if the relatum is oriented. Any object that does not have an intrinsic front is excluded as a base for the front/back and left/right dimensions (Miller and Johnson-Laird 1976). Second, frontness is an *interpretative* category, not a strictly visual one. There is no visual feature that characterizes both the front of a chair and the front of a desk (see figure 3.7a–b). These properties are *functional* ones, derived from our characteristic uses of these objects, and these uses



**Figure 3.7**

The alignment of an object's left, front, and right side does not depend on its spatial, but on its functional, properties.

can be complex. What we experience as the front side of a church from the outside (figure 3.7c) is its rear or back from the inside. Still worse, the alignment of an object's front, left, and right is not fixed, but dependent on its characteristic use (compare the alignments for chair and desk in figures 3.7a and 3.7b); it may even be undetermined or ambiguous (as is the case for the church in figure 3.7c).

Not all intrinsic systems share all of these problems. Levinson (1992a) was able to show that speakers of Tzeltal are much more vision-bound in deriving the intrinsic, orientation-determining parts of objects than English or Dutch, which tend to use a more functional approach. Still, the use of intrinsic perspective always requires detailed interpretation of the relatum's shape, and this has to be shared between interlocutors. These problems do not arise for the deictic and absolute systems.

So far we discussed some of the coordination problems in utilizing the deictic or the intrinsic system. What about speaker/hearer coordination in terms of an absolute system? Here, the interlocutors must agree on absolute orientation, for instance on what is north. Even if such a main direction is indicated in the landscape as a tilt or a coastline, dead reckoning will be required if successful spatial communication is to take place in the dark, in the fog, farther away from one's village, or inside unfamiliar dwellings (Levinson 1992b). The only absolute dimension that is entirely unproblematic is verticality, for which we have a designated sensory system (and even this one can nowadays be tampered with; see Friederici and Levelt 1990 for some experimental results in outer space). So even an absolute system is not without its drawbacks in spatial communication.

### 3.3.3 Interaction between Perspective Systems

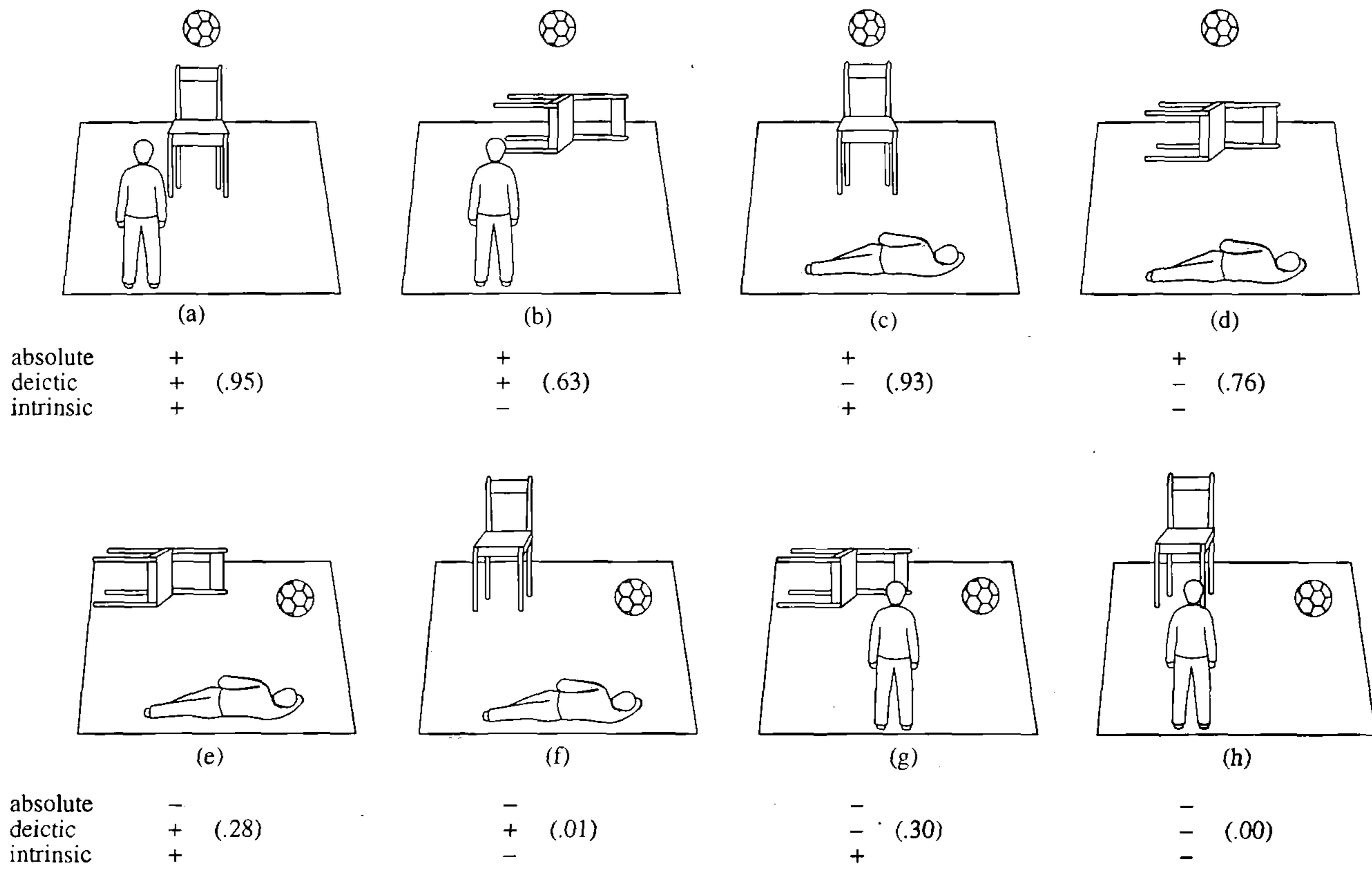
When language users have access to more than a single perspective system, additional problems arise. A first problem already appeared in the previous section. Interlocutors must agree on a system, or must at least be aware of the system used by their partners in speech. This mechanism failed in the network description task in figure 3.6. Various factors can contribute to the establishment of agreement. One important factor is the choice of a default solution. Depending on the communicative task at hand, interlocutors tend to opt for the same solution (Taylor and Tversky 1996; Herrmann and Grabowski 1994). In addition, a speaker's choice of perspective is often given away by the terminology typical for that perspective. When a speaker uses terms such as *north* or *east*, the chosen perspective cannot be deictic or intrinsic. And there are more subtle differences. I have mentioned the presence of vertical dimension terms in deictic directions in a horizontal plane and their total absence in intrinsic directions (the relevant data are to be found in Levelt 1982b). Hence, for these descriptions, presence or absence of vertical dimension terms gives away which perspective system is being used. Surprisingly, the subjects in my experiment completely

ignored this distinctive information when they drew patterns such as in figure 3.6. There are still other linguistic cues. When you say *The chair is on Peter's left*, you are definitely using the intrinsic system, and so is the Frenchman who says *la chaise est à la gauche de ma soeur* (Hill 1982), or the German who utters *Der Stuhl ist zu ihrer Linken* (Ehrich 1982). I am not familiar with any empirical study about the effectiveness of such linguistic cues in transmitting the speaker's perspective to the listener.

Two problems that arise with multiple perspectives are *alignment* and *preemption*. Different perspectives may or may not be aligned in a particular situation, and if they are not aligned, one perspective may gain (almost) full dominance, more or less preempting the other perspectives. This is most easily demonstrated from the use of vertical dimension terms, such as in *A is above/below B*. The basis for verticality is different in the three systems under consideration. In the absolute system verticality is determined by the direction of gravity. In the intrinsic system it is determined by the top/bottom dimension of the relatum. In the deictic system it is probably determined by the direction of your retinal meridian (Friederici and Levelt 1990). In any perceptual situation these three bases of verticality may or may not coincide. Let us consider situations where there is a ball as referent and a chair as relatum and there is an observer/speaker.<sup>4</sup> The ball can now be *above* the chair with respect to one, two, or all three of these bases. The eight possibilities that arise are depicted in figure 3.8.<sup>5</sup>

The appropriateness of saying *the ball is above the chair* varies dramatically for the depicted speaker in the eight scenes. This we know from the work by Carlson-Radvansky and Irwin (1993), who put subjects in the positions depicted in figure 3.8 and asked them to name the spatial relation between the referent and the relatum. Although the scenes were formally the ones in figure 3.8, they varied widely in the objects depicted and in backgrounds.<sup>6</sup> Figure 3.8 shows the percentage of "above" responses for each configuration. Clearly, absolute perspective is quite dominant here (scenes *a–d* are "above" cases in absolute perspective). But in the absence of absolute *above*, intrinsic *above* keeps having some force, whether or not it is aligned with deictic *above* (scenes *e* and *g*, respectively). Deictic *above* alone, however, (scene *f*) is insufficient to release "above" responses. More generally, the deictic dimension does not seem to contribute much in any combination. But further work by the same authors (Carlson-Radvansky and Irwin 1994), in which reaction times of judgments were measured for the same kind of scenes, showed that all three relevant systems contribute to the reaction times. The three systems mutually facilitate or interfere, depending on their alignment. In addition, the reaction times roughly follow the judgment data in figure 3.8. The fastest responses are for *above* in absolute perspective, followed by intrinsic and then deictic *above* responses.

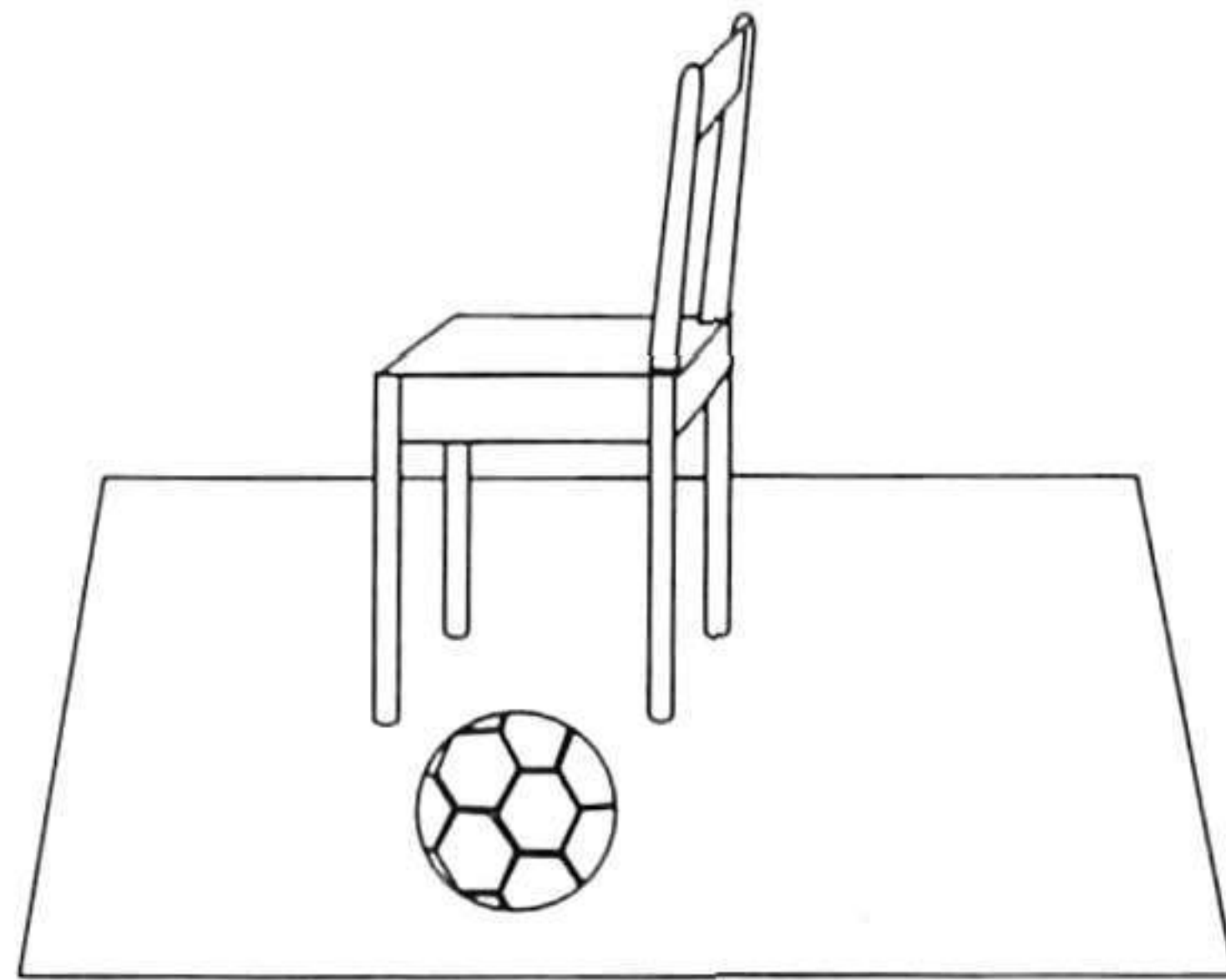
These findings throw a new light on a discussion of my "principle of canonical orientation" (Levelt 1984) by Garnham (1989). I had introduced that principle to



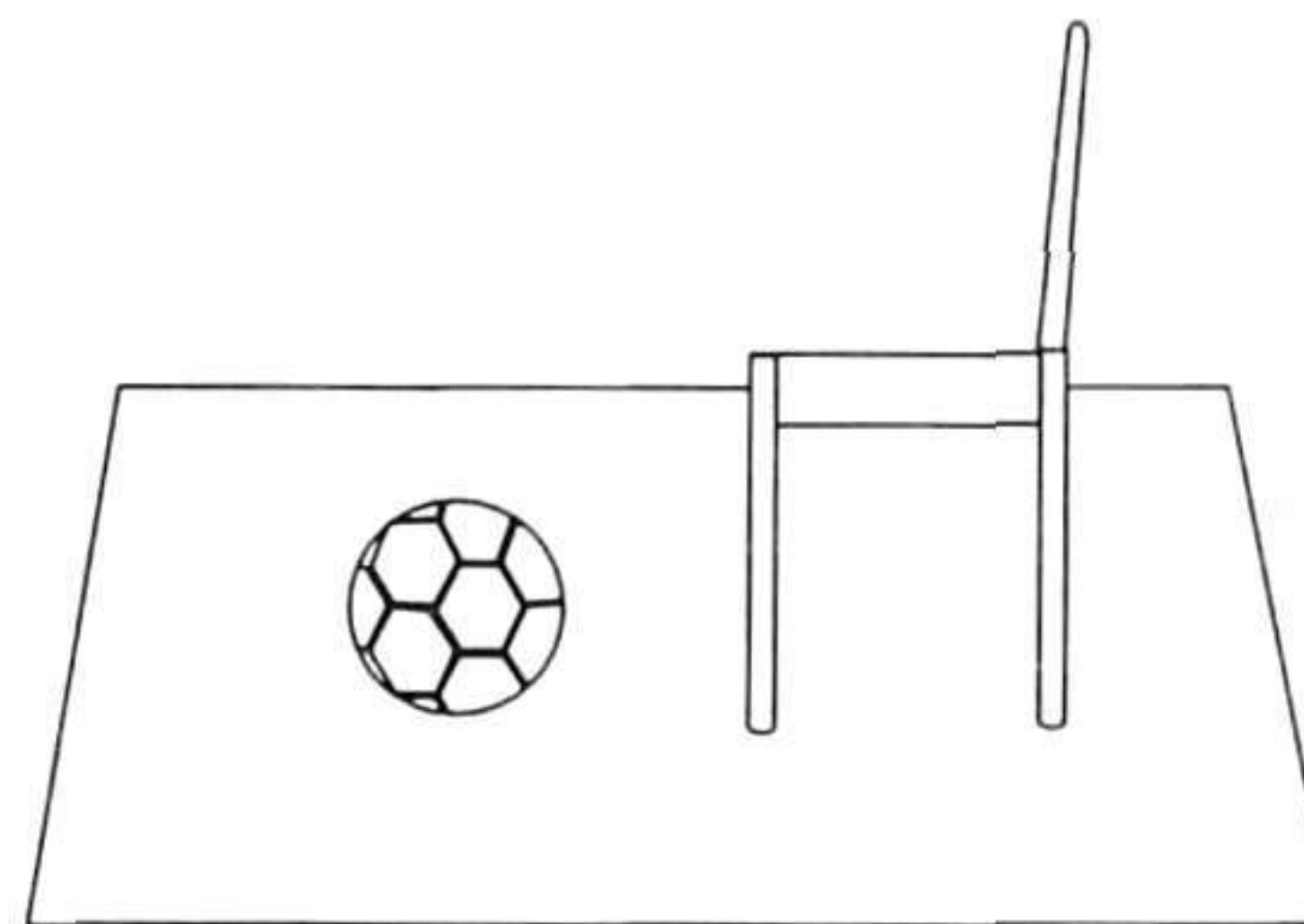
**Figure 3.8**  
 The ball is *above* the chair from all three perspectives (a), from two perspectives (b, c, e), from just one perspective (d, f, g), or from no perspective (h). The numbers in brackets are the percentage of "above" responses obtained in a study by Carlson-Radvansky and Irwin (1993) for scenes with the same formal properties as these.

"The ball is to the left of the chair."

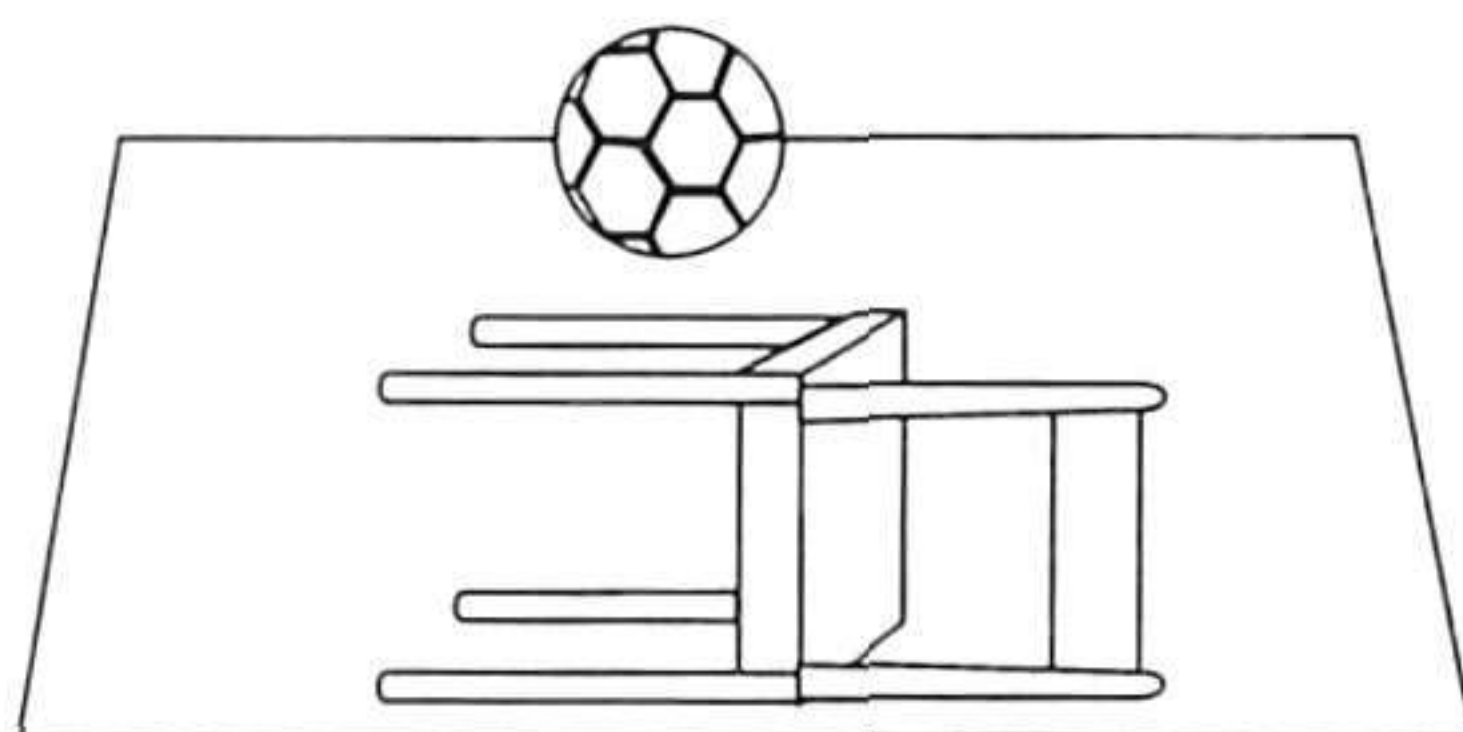
"The ball is in front of the chair."



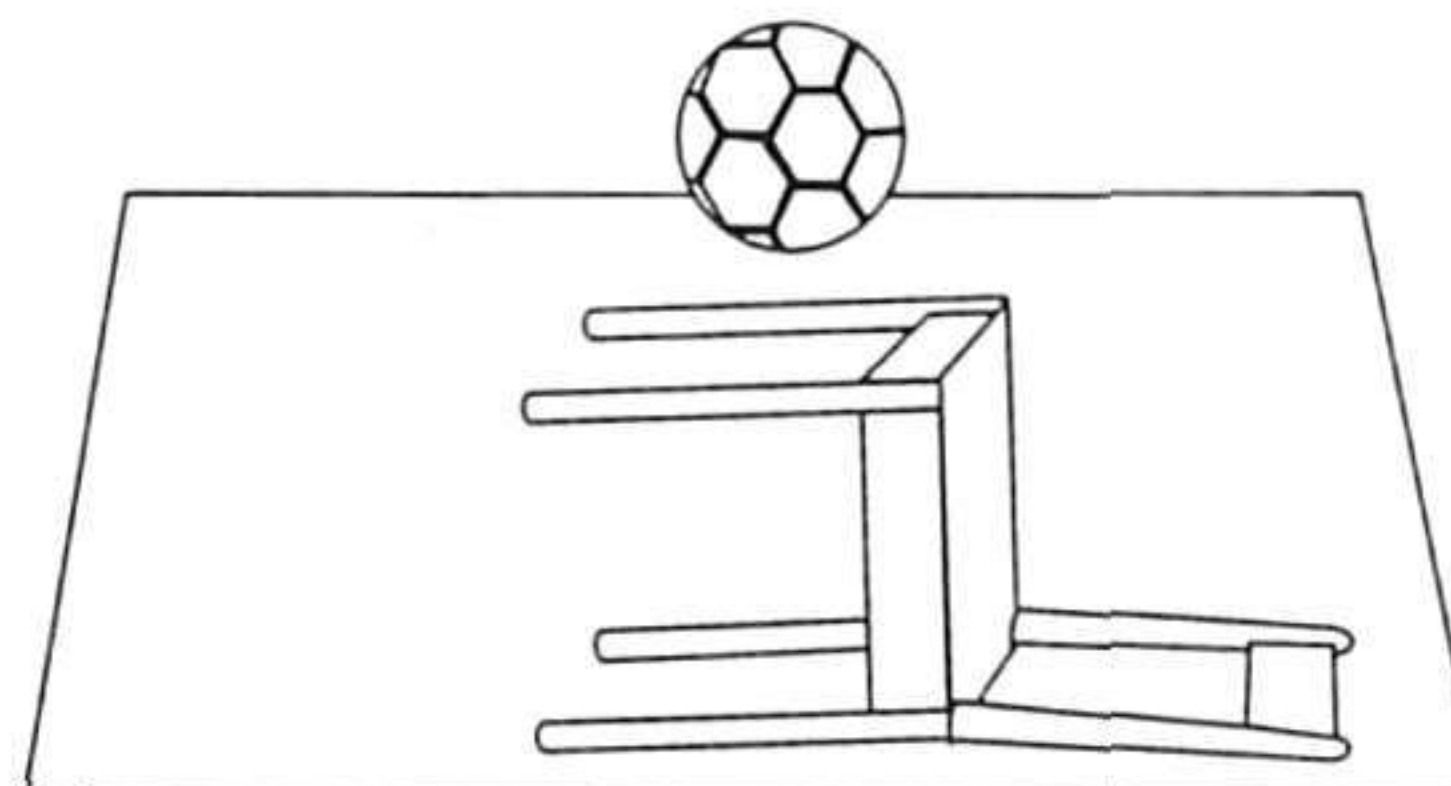
a.



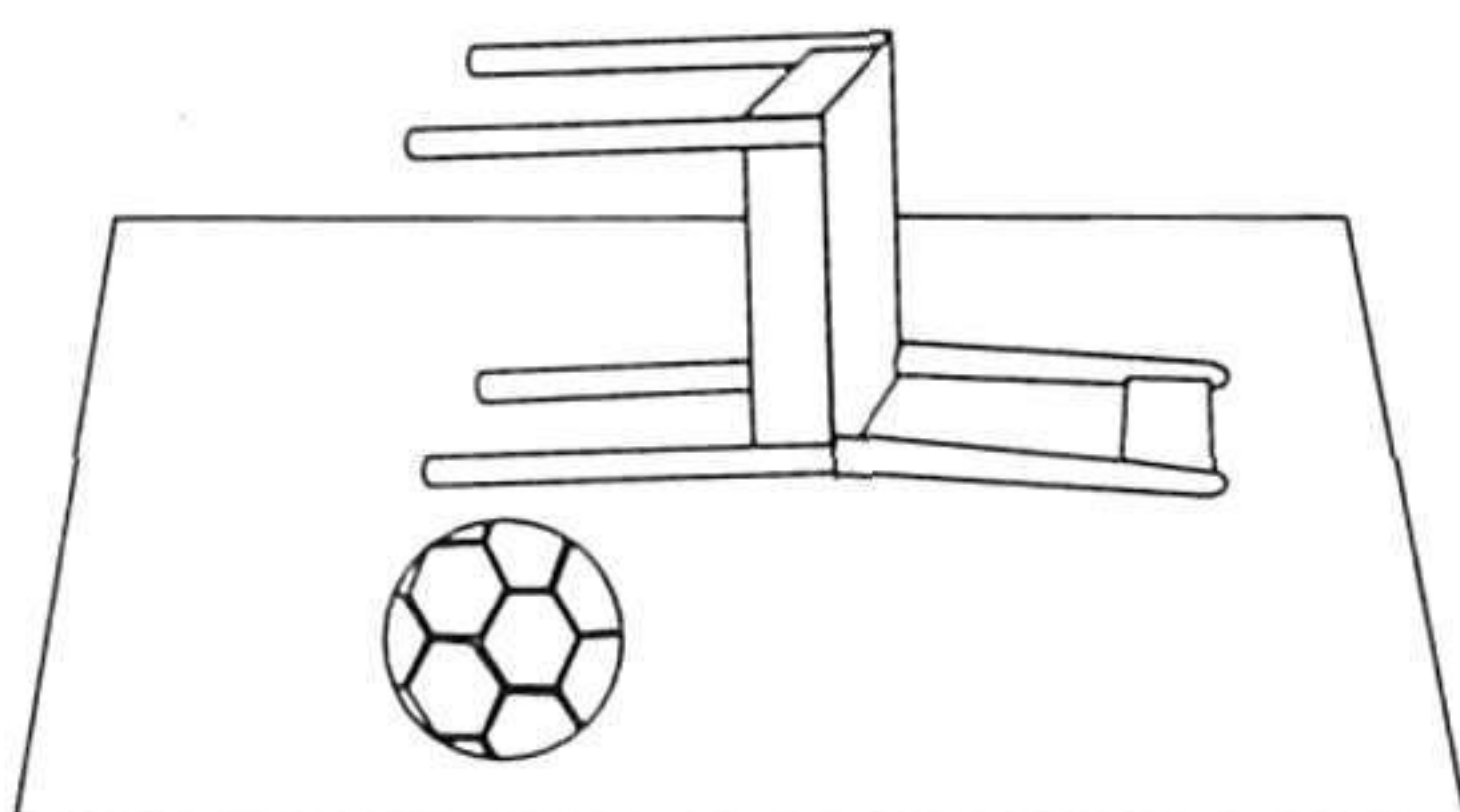
d.



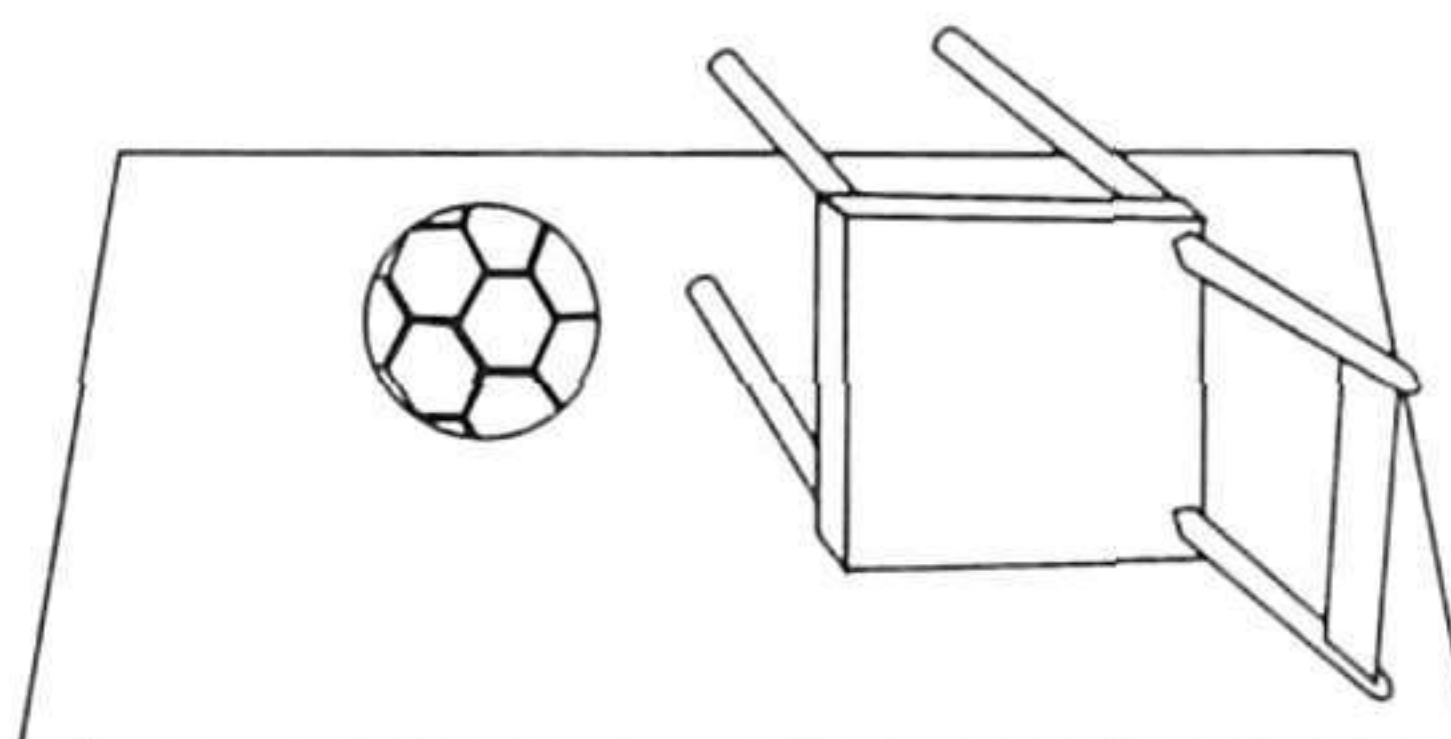
b.



e.



c.



f.

**Figure 3.9**

According to the principle of canonical orientation, the ball can be intrinsically *to the left of* the chair in (a) and (c), but not in (b). It can be intrinsically *in front of* the chair in (d) and (f), but not in (e).

account for certain cases where the intrinsic system is “immobilized” when it conflicts with the deictic system. Because the principle is directly relevant to the present discussion of alignment and preemption, I cite it here from the original paper:

***Principle of canonical orientation*** For the intrinsic system to refer to a relatum’s intrinsic dimension, that dimension must be in canonical position with respect to the perceptual frame of orientation of the referent.

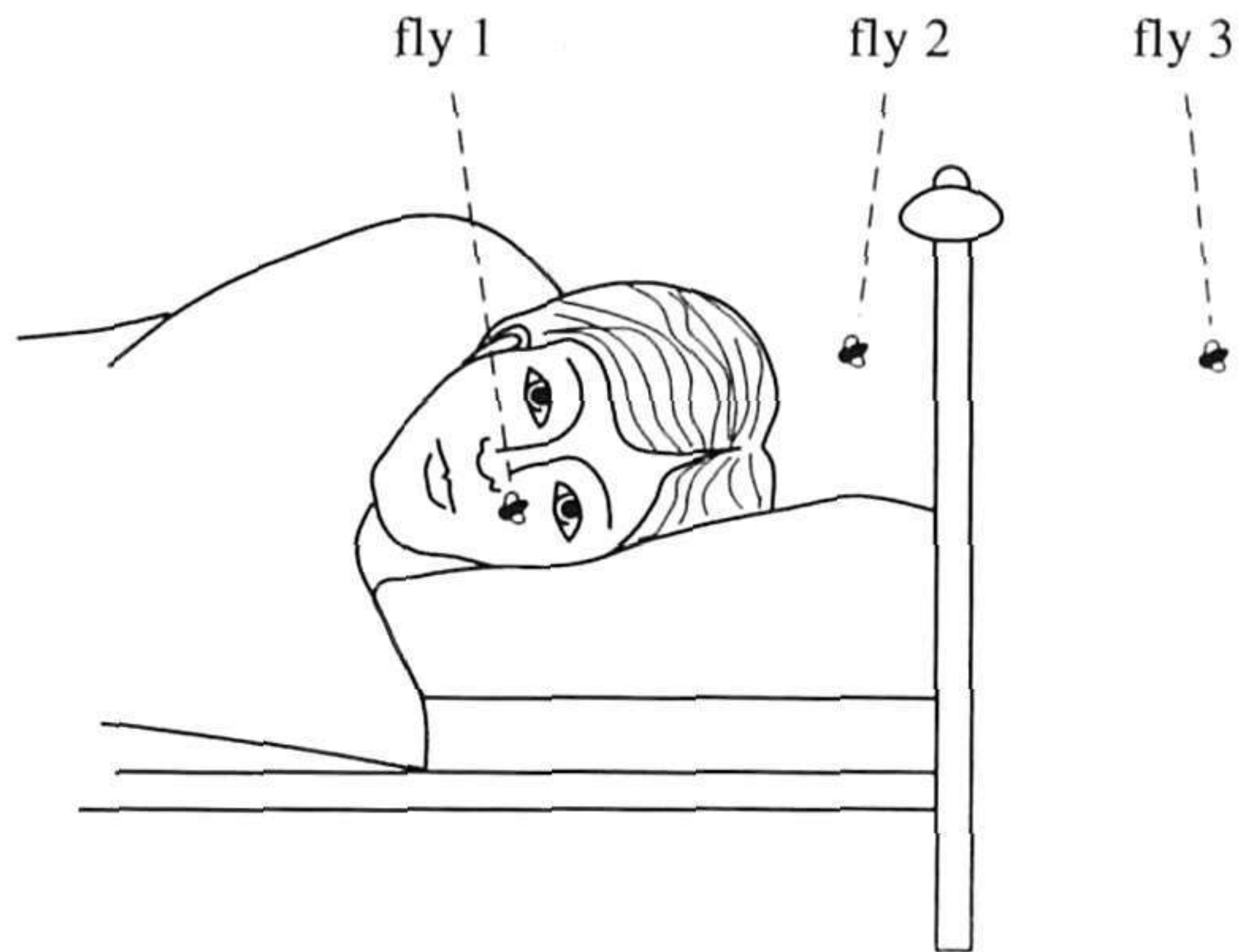
The principle of canonical orientation is easily demonstrated from figure 3.9. Cases *a*, *b*, and *c*, in the left-hand side of the figure, refer to the intrinsic description *the ball is to the left of the chair*.

According to the principle of canonical orientation this is a possible description in (*a*). The description refers to the relatum’s intrinsic left/right dimension. That dimension is in canonical orientation to the relatum’s perceptual frame. The perceptual frame for the chair’s orientation is in this case the normal gravitational field. The chair is in canonical position with respect to this perceptual frame. In particular, the chair’s left/right dimension has a canonical direction, that is, it lays in a plane that is horizontal in the perceptual frame. However, the description is virtually impossible in (*b*). Here the left/right dimension of the chair (the relatum) is not in canonical position; it is not in a horizontal plane, given the perceptual frame. Finally and surprisingly, it is for many native speakers of English acceptable to say *the ball is to the left of the chair* in case of (*c*). Here the chair is not in canonical position either, but the chair’s left/right dimension is; it is in a horizontal plane of the perceptual frame. Hence the principle of canonical orientation is satisfied in this case.

The state of affairs is similar for the intrinsic description *the ball is in front of the chair*. This description is fine for (*d*). It is, however, virtually unacceptable for (*e*), and this is because the front/back dimension of the relatum (the chair) is not in a canonical, horizontal plane with respect to the perceptual frame. Although in (*f*) the chair is not in canonical position, its front/back dimension is. Hence the description is again possible according to the principle, which agrees with intuitions of many native speakers of English to whom I showed the scene (the formal experiment has never been done, though).

Why does the principle refer to “the perceptual frame of orientation of the referent,” and not just to “the perceptual frame of orientation”? In figure 3.9 it is indeed impossible to distinguish between these two. The perceptual frame of the ball is the visual scene as a whole. Its orientation, and in particular its vertical direction, determines whether some dimension of the relatum (the chair) is in canonical position. More generally, a referent’s perceptual frame of orientation will normally be the experienced vertical, as it derives from vestibular and visual environmental cues, and





**Figure 3.10**

According to the principle of canonical orientation, fly 1 can be intrinsically *to the left* of John's nose, and fly 2, but not fly 3, can be *above* John's head (reproduced from Levelt 1984).

will be the same for referent and relatum. But there are exceptions in which a dominant visual Gestalt adopts the function of perceptual frame for the referent. This can happen in the scene of figure 3.10, which is reprinted here from Levelt (1984).

In that paper I argued that it is not impossible in this case to say about fly 2 in the picture: *there is a fly above John's head* even though the top/bottom dimension of John's head is not in canonical orientation. And this is in agreement with the principle. To show this, let us consider the figure in some more detail, beginning at the location of fly 1. Here John's face is a quite dominant background pattern which may become the perceptual frame of orientation for the fly. In that case, the principle of canonical orientation predicts that it is appropriate to say, *there is a fly to the left of John's nose*. This is because the intrinsic left/right dimension in which the fly is spatially related to John's nose is canonically oriented with respect to the perceptual frame. It is in a plane perpendicular to the top/bottom dimension of the face. And fly 2 may similarly take John's face as its perceptual frame, because it is so close to it. If this is a subject's experience, then it is appropriate to say *there is a fly above John's head*, according to the principle. The experimental findings by Carlson-Radvansky and Irwin (1993; cf. figure 3.8g) now confirm that this can indeed be the case.<sup>7</sup> Fly 3 is further away from John's head and does not naturally take John's head as its perceptual frame of reference. Hence it is less appropriate here to say it is "above" John's head. Notice that in these three cases John's head itself has the bed and its normal gravitational orientation as its perceptual frame. Hence the perceptual frame of the referent can be different from the larger perceptual frame in which the relatum

is embedded. In other words, there can be a hierarchy of frames, and it is not necessarily the case that the referent and the relatum share a frame.

Garnham (1989) challenged the principle of canonical orientation. Although he agreed with the intuitions concerning the scenes in figure 3.9, he rejected those with respect to figure 3.10. That allowed him to ignore the distinction between the referent's and the relatum's perceptual frame and to formulate a really simple principle, the "framework vertical constraint," which says that "no spatial description may conflict with the meanings of *above* and *below* defined by the framework in which the related objects are located." But the results by Carlson-Radvansky and Irwin (1993) for scenes *e* and *g* in figure 3.8 contradict this because, according to Garnham, *above/below* derives in this case from the normal gravitational framework. Hence there is a conflict between the meaning of *above* in this framework and the description *the ball is above the chair*, which should make this description impossible according to his constraint, but it does not. The findings are, however, in agreement with the principle of canonical orientation because the experiments involved cases such as the one just discussed for fly 2 in figure 3.10.

Garnham's critique of my 1984 formulation of the principle can, in part, be traced back to a vagueness of the term *canonical position*. It does not positively exclude the following strict interpretation: the dimension on which the intrinsic location is made should coincide with the same dimension in the perceptual frame. This is obviously false, as Garnham (1989) correctly pointed out. For instance, "if a vehicle is parked across a street, a bollard [traffic post] to the intrinsic right of the vehicle can still be described as to its right" (p. 59), even if the perceptual frame for the bollard is given by the street (whose right side is opposite to the vehicle's right side). The only tenable interpretation of "canonical position" is a weaker one:

***Canonical position*** The top-down dimension of the relatum is in canonical position if it coincides with the vertical dimension of the referent's perceptual frame. The left-right and front-back dimensions of the relatum are in canonical orientation if and only if they are in a plane that is perpendicular to the vertical in the referent's perceptual frame.

With this further specification, then, the principle of canonical orientation seems to be in agreement with intuition and with experimental data. If in a scene canonical orientation does not hold, the intrinsic system is evaded by the standard average European (SAE) language user; it is preempted by the deictic or by the absolute system.<sup>8</sup>

In this section I have discussed various properties of perspective systems that are of pragmatic significance. We have seen that systems differ in inferential potential and

in their demands on coordination between interlocutors. We also have seen that if one system is dominant, concurring systems are not totally dormant in the speaker's mind. Their rivalry appears from the kind and speed of a subject's spatial judgments, and the outcome depends on quite abstract properties of the rivaling systems, as is the implication of the principle of canonical orientation.

### 3.4 Ellipsis in Spatial Expressions

Perspective taking is one aspect of our thinking for speaking. When we talk about spatial configurations, we create predications about spatial properties of entities or referents in the scene. These predications usually relate the entity to some relatum in terms of some perspective system. In short, the process of perspective taking maps a spatial representation onto a propositional or semantic one. The latter is the speaker's *message*, which consists of lexical concepts, that is, concepts for which there are words in the speaker's target language.

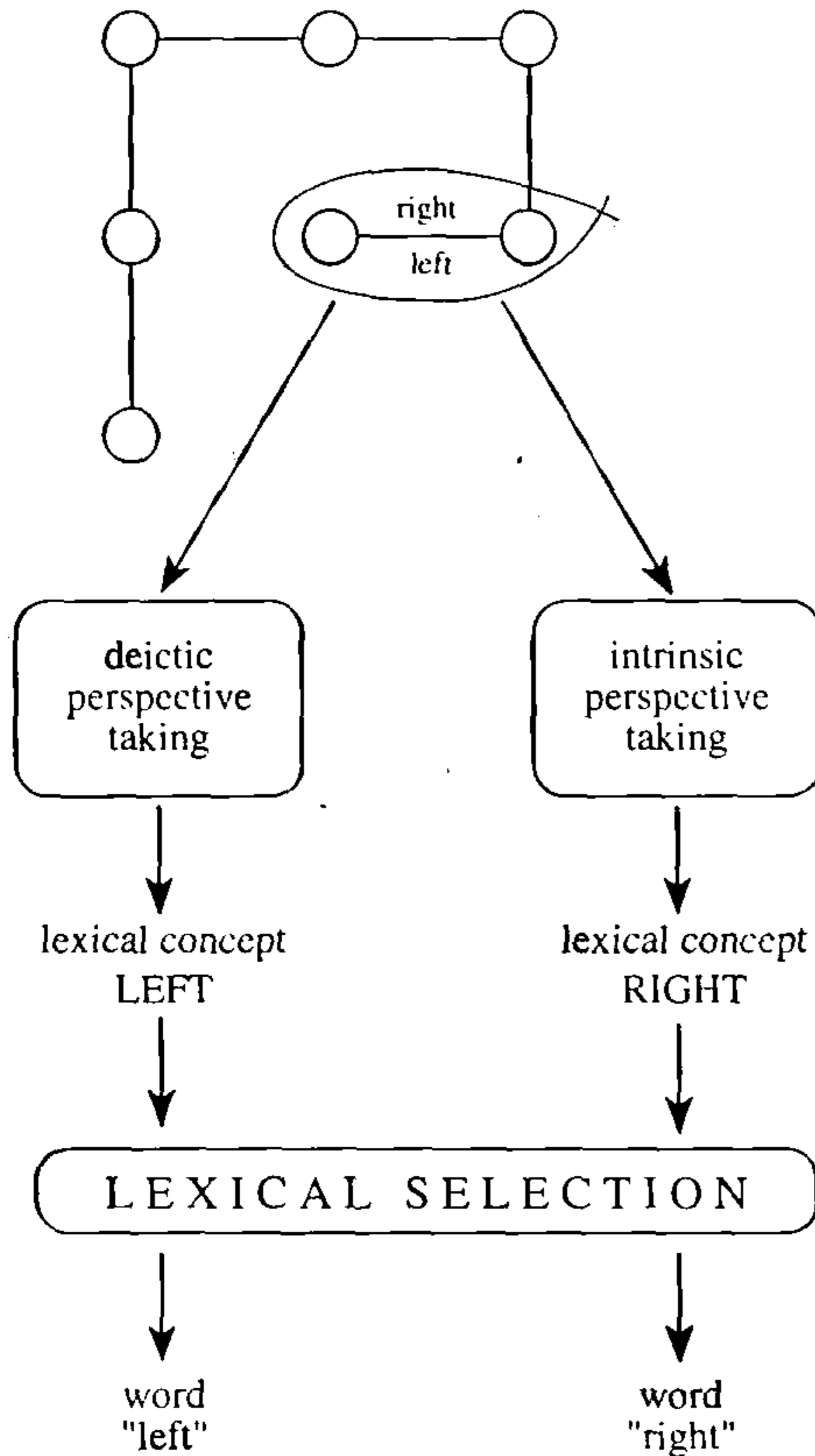
This state of affairs is well exemplified in figure 3.5. The same pattern is expressed in two systematically different ways, dependent on the speakers' perspectives. Figure 3.11 represents one critical detail (circled) of this example. Depending on the perspective taken, the same referent/relatum relation is expressed as *left* or as *right*. Figure 3.11 expresses that the choice of lexical concept (and ultimately of lexical item) depends on the perspective system being used, that is, on thinking for speaking. It is important to be clear on the underlying assumption here. It is that the spatial representation is itself perspective-free; it is neither intrinsic nor deictic. This assumption may or may not be correct, and I will return to it below.

The issue in this section is whether spatial ellipsis originates *before* or *after* perspective taking. In other words, does the speaker decide not to mention a particular feature of the spatial representation, or rather, does the speaker decide not to express a particular lexical concept? In the first case we will speak of "deep ellipsis"; in the latter case, of "surface ellipsis" (roughly following Hankamer and Sag 1976 on "deep" and "surface anaphora").

Compare the following two descriptions from our data. Both relate to the encircled trajectory in the left pattern of figure 3.12, plus the move that precedes it. The first description is nonelliptic with respect to the directional expression, the second one is elliptic in that respect.

Full deictic: "Right to yellow. Right to blue. Finished."

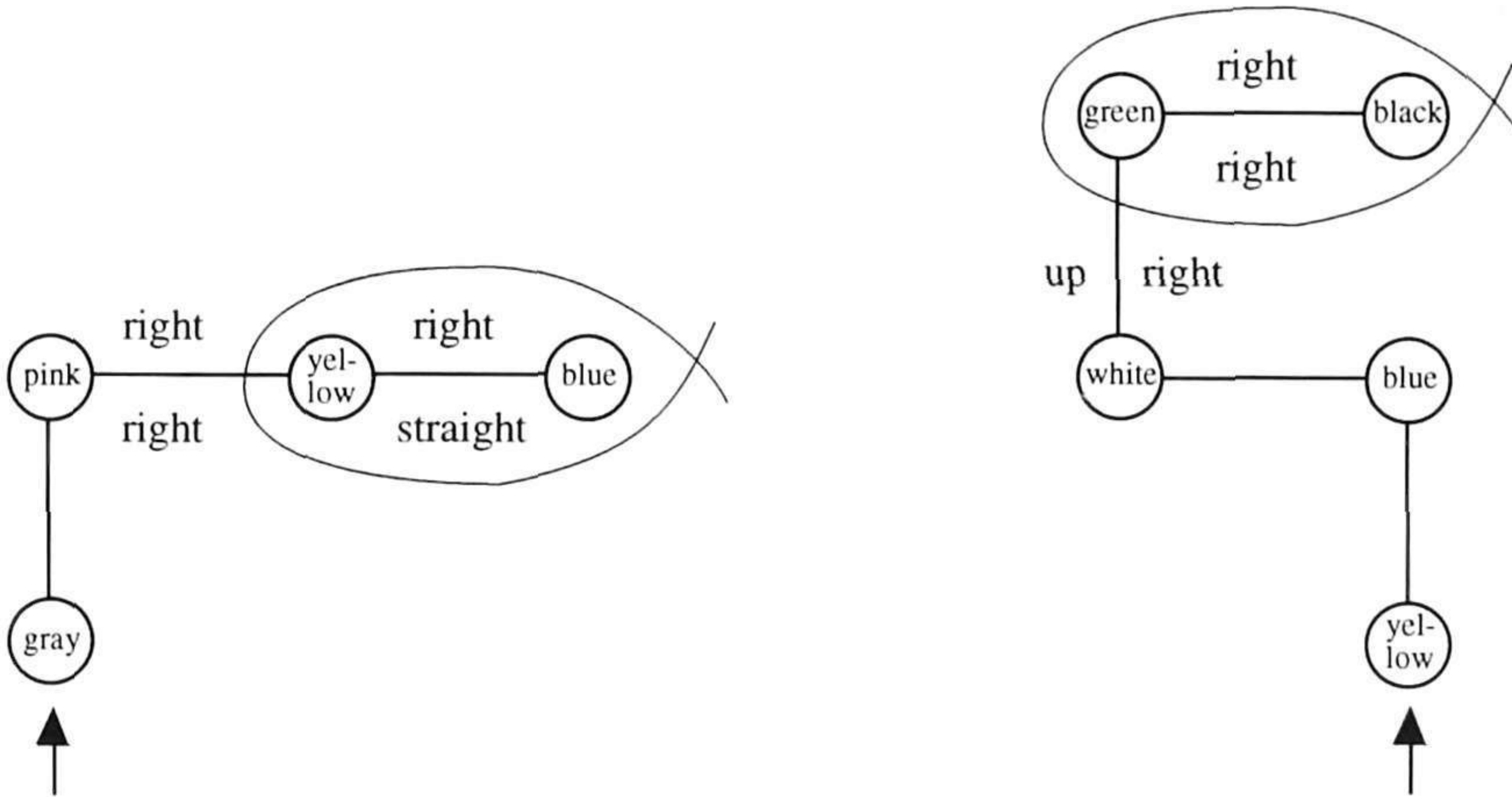
Elliptic deictic: "From pink we go right one unit and place a yellow dot. One, er, one unit from the yellow dot we place a blue dot."



**Figure 3.11**

Perspective taking is involved in deriving a semantic or “propositional” representation from a spatial representation.

The crucial feature of the latter, elliptic expression is that it contains no spatial term that relates the blue dot to the (previous) yellow one. How does the speaker create this ellipsis? There are, essentially, two possibilities. The first one is that the speaker in scanning the spatial configuration recognizes that the new visual direction is the same as the previous one. Before getting into perspective taking, the speaker decides not to prepare that direction for expression again. This is deep ellipsis. The second possibility is that the speaker does apply deictic perspective to the second move, thus activating the lexical concept RIGHT a second time. This repeated activation of the concept then leads to the decision not to formulate the lexical concept a second time,



**Figure 3.12**

Deictic and intrinsic descriptions for two patterns. Can the last spatial term (*right, straight*) be deleted?

that is, not to repeat the word *right*. This is surface ellipsis. These two alternatives are depicted in figure 3.13.

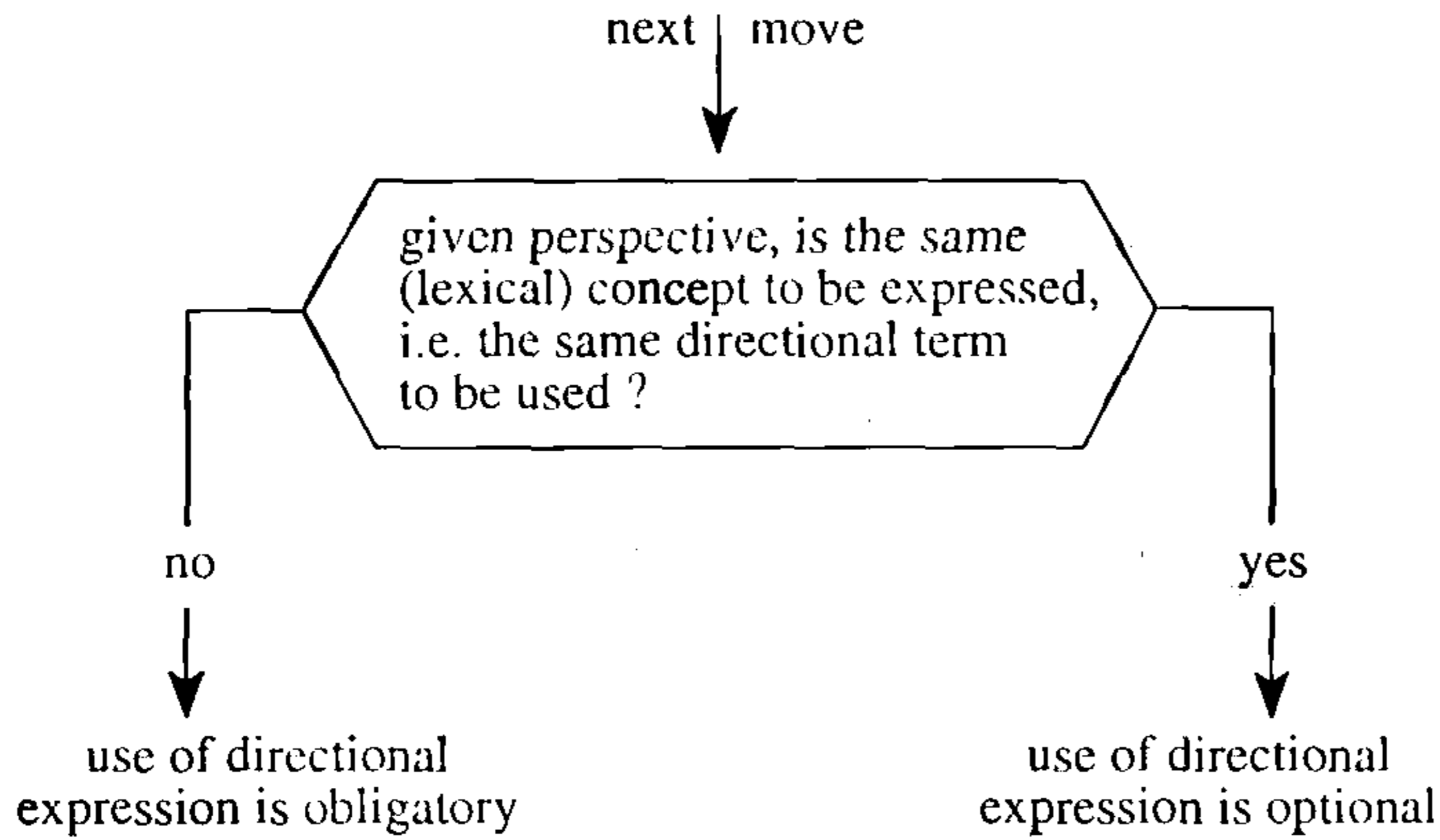
The alternatives can now be distinguished by observing what happens in descriptions from an intrinsic perspective. Here is an instance of a full intrinsic description of the same trajectory:

Full intrinsic: “Then to the right to a yellow node and straight to a blue node.”

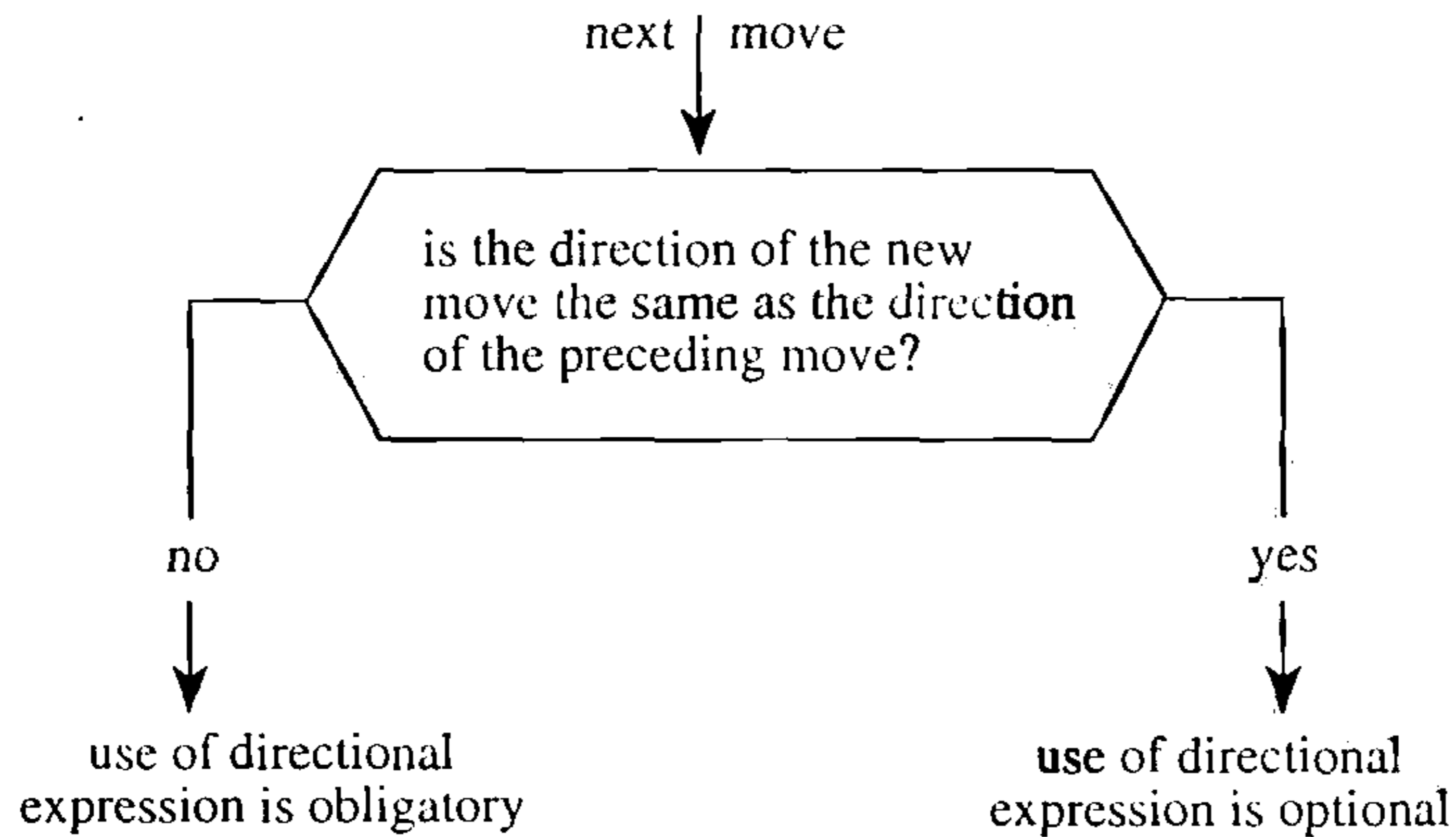
Can the same state of affairs be described elliptically? This should produce something like: *Then to the right to a yellow node and to a blue node*. The answer is not obvious; intuitions waver here. In case of deep ellipsis this should be possible. Just as the previous deictic speaker, the present intrinsic one will scan the spatial scene and recognize that the new direction is the same as the previous one and the speaker may decide not to prepare it again for expression; it is *optional* to mention the direction. But in case of surface ellipsis the intrinsic speaker has a problem. In the intrinsic system the direction of the first move is mapped onto the lexical concept **RIGHT**, whereas the direction of the second move is mapped onto **STRAIGHT**. Because the latter is not a repetition of the former, it has to be formulated in speech. In other words, the condition for surface ellipsis is not met for the intrinsic speaker; it is *obligatory* to use a directional expression.

This state of affairs can now be exploited to test empirically whether spatial ellipsis is deep or surface ellipsis. Does ellipsis occur in intrinsic descriptions of this kind? If

MODEL 1 "Surface ellipsis" (ellipsis is perspective- dependent)



MODEL 2 "Deep ellipsis" (ellipsis is perspective- independent)



**Figure 3.13**

Surface ellipsis versus deep ellipsis. Is it reiterating a lexical concept or a spatial direction that matters?

so, we have an argument for deep ellipsis. And we can create an alternative case where surface ellipsis is possible for intrinsic descriptions, but not deep ellipsis. An example concerns the encircled trajectory in the right pattern of figure 3.12. A normal full intrinsic description of this trajectory (plus the previous one) is

Full intrinsic: Then right to green. And then right to black.

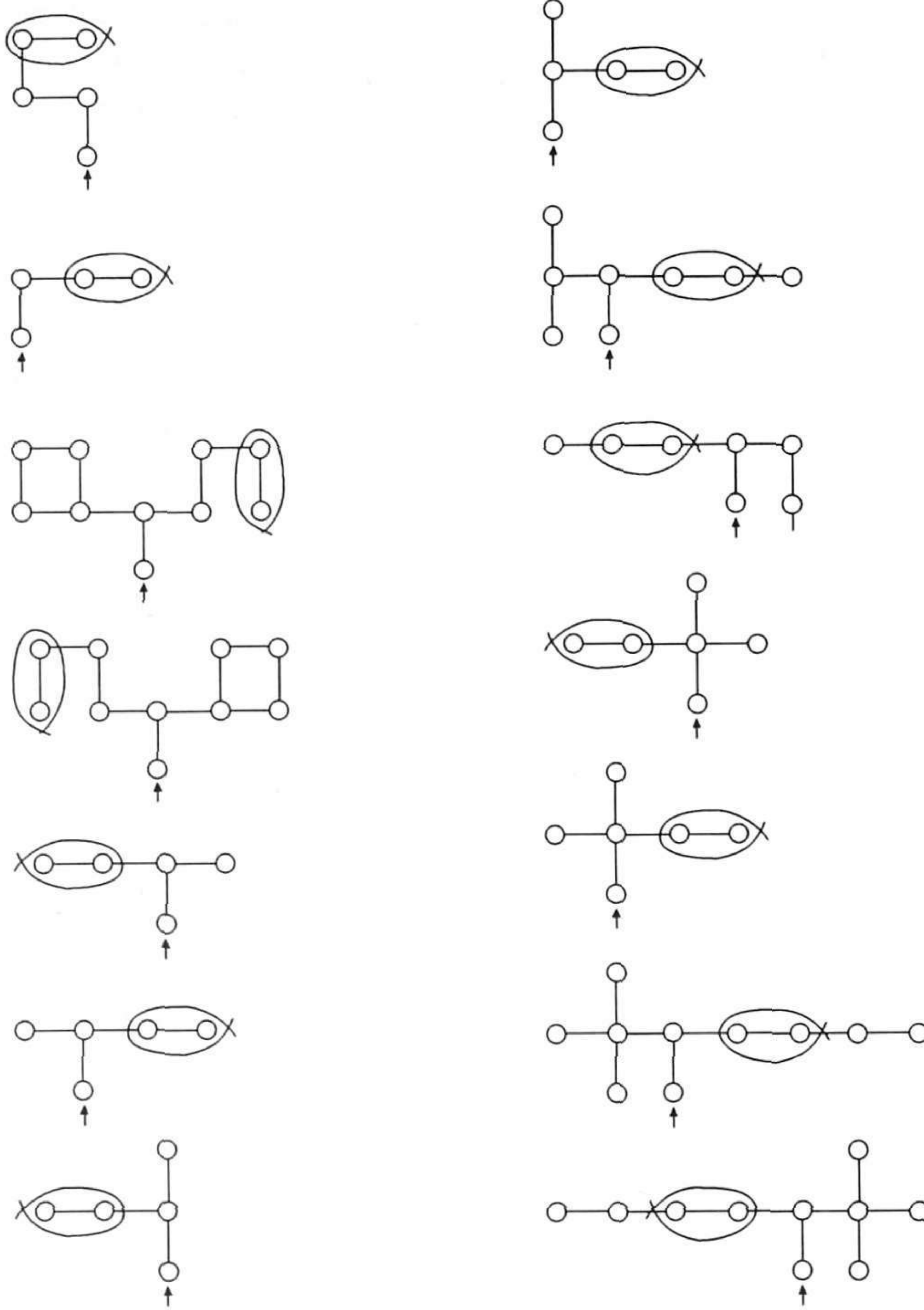
Is surface ellipsis possible here, producing “Then right to green. And to black” or some similar expression? That is an empirical issue. It should be clear that neither deep nor surface ellipsis is possible in a deictic description of this pattern. Take this full deictic description from our data:

Full deictic: From white we go up to a green circle. And from the green circle we go right to a black circle.

Surface ellipsis is impossible here because “right” is not a repetition of the previous directional term (“up”). Deep ellipsis is impossible because the trajectory direction is different from the previous one. Hence, if we find ellipsis in such cases, we will have to reject both models.

In an experiment reported in Levelt (1982a,b) we had asked 53 subjects to describe 53 colored dot patterns, among them those in figure 3.12. I will call the circled moves in these patterns “critical moves” because the surface and deep models make predictions about them that differ critically for deictic and intrinsic descriptions in the way just described. Among the test patterns there were 14 that contained such critical moves; they are given in figure 3.14. I checked all 53 subjects to determine whether they made elliptic descriptions for any of these 14 critical trajectories. I removed all subjects who did not have a consistent perspective over these 14 critical patterns; a subject’s 14 pattern descriptions should either be all deictic or all intrinsic. This left me with 31 consistent deictic subjects and 13 consistent intrinsic ones,<sup>9</sup> and hence with  $44 \times 14 = 616$  pattern descriptions to be checked. In this set I found a total of 43 cases of ellipsis.<sup>10</sup> These are presented in table 3.1.

The table presents predictions and results under both models of ellipsis. For each critical move I determined whether a directional term would be obligatory or optional (i.e., elidable) under the model in deictic and in intrinsic descriptions (such as I did above for the critical moves of the patterns in figure 3.12). Hence there are four cases per model. The table presents the actual occurrence of ellipsis for these four cases within each model. It should be noticed that the two models make the same predictions with respect to deictic descriptions; if use of a directional term is obligatory under the surface model, it is also obligatory under the deep model and vice versa. But this is not so for the intrinsic descriptions.



**Figure 3.14**

Fourteen test patterns containing “critical moves,” including the two example patterns of figure 3.12. Each test pattern includes either the one or the other example pattern as a substructure (though rotated in two cases). The critical moves are circled.



**Table 3.1**

Distribution of Elliptical Descriptions under Surface and Deep models of Ellipsis

Model	Surface ellipsis			Deep ellipsis		
	deictic	intrinsic	Total	deictic	intrinsic	Total
Description is						
Directional term is						
obligatory	1	18	19	1	0	1
optional	24	0	24	24	18	42
Total	25	18	43	25	18	43

If a model says “obligatory,” but ellipsis does nevertheless occur, that model is in trouble. How do the two models fare? It is immediately obvious from the table that the surface model is out. Where it prescribes obligatory use of a directional term, there are no less than 18 violations among the intrinsic descriptions (i.e., cases of ellipsis) and one among the deictic descriptions, for a total of 19. That is almost half our sample. In contrast, the deep model is in good shape; there is only one deictic description that violates it.<sup>11</sup> All other deictic and all intrinsic descriptions respect the deep model.

These findings show that the decision to skip mentioning a direction is really an early step in thinking for speaking. It precedes the speaker’s application of a perspective; the speaker’s linguistic perspective system is irrelevant here. The decision is based on a visual or imagistic representation, not on a semantic (lexical-conceptual) representation (see figure 3.11). This is, probably, the same level of representation where linearization decisions are taken. When we describe 2-D or 3-D spatial patterns (such as the patterns in figure 3.14 or the layout of our living quarters), we must decide on some order of description because speech is a linear medium of expression. The principles governing these linearization strategies (Levelt 1981, 1989) are non-linguistic (and in fact nonsemantic) in character; they relate exclusively to the image itself.

But these very clear results on ellipsis create a paradox. If ellipsis runs on a perspective-free spatial representation, spatial representations are apparently *not* perspectivized. But this contradicts the convincing experimental findings reported by Brown and Levinson (1993) and by Levinson (chapter 4, this volume), which show that when a language uses absolute perspective, its speakers use *oriented* (i.e., perspective-dependent) spatial representations in nonlinguistic spatial matching tasks. For instance, the subject is shown an array of two objects A and B on a table, where A is (deictically) left of B (hence A-B). Then the subject is turned around 180° to another table with two arrays of the same objects, namely, A-B and B-A, and then asked to

indicate which of the two arrays is identical to the one the subject saw before. The “absolute” subject invariably chooses the B-A array, where A is deictically to the *right* of B. What the subject apparently preserves is the absolute direction of the vector  $\overrightarrow{AB}$ . A native English or Dutch subject, however, typically produces the deictic response (A-B). Hence spatial representations *are* perspectivized already, in the sense that they follow the dominant perspective of the language even in nonlinguistic tasks, that is, where there is no “thinking for speaking” taking place.<sup>12</sup>

How to solve this paradox? One point to note is that the above ellipsis data and Brown and Levinson’s (1993) data on oriented spatial representations involve different perspectives, and the ellipsis predictions are different for different perspectives. As can be seen from table 3.1, columns 1 and 4, the same predictions result from the deep and the surface model under deictic perspective. The two models can only be distinguished when the speaker’s perspective is intrinsic (cf. columns 2 and 5); violations under deictic perspective could only show that *neither* model is correct. In this respect, absolute perspective behaves like deictic perspective. If a speaker’s perspective is absolute, the deep and surface models of ellipsis make the same predictions; if two arcs have the same spatial direction or orientation, the corresponding lexical concepts will be the same as well (e.g., both north, or both east).

In other words, ellipsis data of the kind analyzed here can only distinguish between the deep and surface models if the speaker’s perspective is intrinsic. One could then argue that Brown and Levinson’s findings show that absolute and deictic perspective are “Whorfian,” that is, a property of the spatial representation itself. If, in addition, the intrinsic system is not Whorfian in the same sense, the above ellipsis data would be explained as well.

The problem is, of course, why intrinsic perspective should be non-Whorfian. After all, speakers of Mopan, exclusive users of intrinsic perspective, will profit from registering the position of foregrounded objects relative to background objects that have intrinsic orientation. If at some later time the scene is talked about from memory, that information about intrinsic position will be crucial for an intrinsic spatial description. But if we discard the option of excluding intrinsic perspective from “Whorfianness,” the paradox remains.

More important, it seems to me, is the fact noted in the introduction that perspective is linguistically free. There is no “hard-wired” mapping from spatial to semantic representations. What we pick out from a scene in terms of entities and spatial relations to be expressed in language is not subject to fixed laws. There are preferences, for sure, following Gestalt properties of the scene, human interest, and so on, but they are no more than preferences. Similarly, we can go for one perspective or another if our culture leaves us the choice, and this chapter has discussed various reasons for choosing one perspective rather than another, depending on communica-

tive intention and situation. It is correct to say that Guugu Yimithirr speakers can choose from only one, absolute perspective, but that does not obliterate their freedom in expressing spatial configurations in language. The choice of referents, relata, spatial relations to be expressed, the pattern of linearization chosen when the scene is complex, and even the decision to express absolute perspective at all (e.g., *A* is north of *B*, rather than *A* is in *B*'s neighborhood) are prerogatives of the speaker that are not thwarted by the limited choice of perspective. As all other speakers, the Guugu Yimithirr can attend to various aspects of their spatial representations; they can express in language what they deem relevant and in ways that are communicatively effective. This would be impossible if the spatial representation dictated its own semantics. Hence, Brown and Levinson's (1993) important Whorfian findings cannot mean that spatial and semantic representations have a "hard-wired" isomorphism. A more likely state of affairs is this. A culture's dominant perspective makes a speaker attend to spatial properties that are relevant to that perspective because it will facilitate (later) discourse about the scene. In particular, these attentional biases make the speaker register in memory spatial features that are perspective-specific, such as the absolute orientation of the scene. This does not mean, however, that an ellipsis decision must make reference to such features. That one arc in figure 3.12 is a continuation of another arc is a spatial feature in its own right that is available to a speaker of any culture. Any speaker can attend to it and make it the ground for ellipsis. In other words, the *addition* of perspective-relevant spatial features does not preempt or suppress the registration of other spatial properties that can be referred to or used in discourse.

### 3.5 Conclusion

This chapter opened by recalling, from Levelt (1989), the distinction between macroplanning and microplanning. In macroplanning we elaborate our communicative intention, selecting information whose expression can be effective in revealing our intentions to a partner in speech. We decide on what to say. And we linearize the information to be expressed, that is, we decide on what to say first, what to say next, and so forth. In microplanning, or "thinking for speaking," we translate the information to be expressed in some kind of "propositional" format, creating a semantic representation, or message, that can be formulated. In particular, this message must consist of lexical concepts, that is, concepts for which there are words in the target language. When we apply these notions to spatial discourse, we can say that macroplanning involves selecting referents, relata, and their spatial relations for expression. Microplanning involves, among other things, applying some perspective system that will map spatial directions/relations onto lexical concepts.

The chapter has been largely about microplanning, in particular about the pragmatics of different perspective systems. It has considered the advantages and disadvantages of deictic, intrinsic, and absolute systems for spatial reasoning and for speaker/hearer coordination in spatial discourse. It has also considered how a speaker deals with situations in which perspective systems are not aligned.

“Thinking for speaking” led, as a matter of course, to the question whether this perspectival thinking is just for speaking or more generally permeates our spatial thinking, that is, in some Whorfian way. The discussed recent findings by Levinson and Brown strongly suggest that such is indeed the case. I then presented experimental data on spatial ellipsis showing that perspective is irrelevant for a speaker’s decision to elide a spatial direction term. Having speculated that the underlying spatial representation might be perspective-free, contrary to the Whorfian findings, I argued that this is paradoxical only if the mapping from spatial representations onto semantic representations is “hard-wired.” But this is not so; speakers have great freedom in both macro- and microplanning. There are no strict laws that govern the choice of relatum and referent, that dictate how to linearize information, and so forth. In particular, there is no law that the speaker must acknowledge orientedness of a spatial representation (if it exists) when deciding on what to express explicitly and what implicitly. There are only (often strong) preferences here that derive from Gestalt factors, cultural agreement on perspective systems, ease of coordination between interlocutors, requirements of the communicative task at hand, and so on.

Still, it is not my intention to imply that anything goes in thinking for speaking. Perspective systems are interfaces between our spatial and semantic “modules” (in Jackendoff’s sense, chapter 1, this volume), performing well-defined restricted mapping operations. The interfacing requirements are too specific for these perspective systems to be totally arbitrary. But much more challenging is the dawning insight from anthropological work that there are only a few such systems around. What is it in our biological roots that makes the choice so limited?

## Notes

1. I am in full agreement with Levinson’s taxonomy of frames of reference (here called “perspective systems”) in chapter 4 of this volume. The main distinction is between relative, intrinsic, and absolute systems, and each has an egocentric and an allocentric variant. The three perspective systems discussed here are relative egocentric (= deictic), intrinsic allocentric, and absolute allocentric. The relative systems are three-place relations between referent, relatum, and base entity (“me” in the deictic system); the intrinsic and absolute systems are two-place relations between referent and relatum.

2. Brown and Levinson (1993) present the case of Tenejapan, where the traverse direction in the absolute system is not polarized, that is, spanned by two converse terms; there is just one

term meaning “traverse.” Obviously, the notion of converseness is not applicable. The notion of transitivity, however, *is* applicable and holds for this system (see below in text).

3. Barbara Tversky (personal communication) has correctly pointed out that Bühler (1934) would treat this case as a derived form of deixis, “Deixis am Phantasma,” where the speaker *imagines* being somewhere (for instance in the network). There would be two speakers then, a real one and imaginary one, each forming a base for a (different) deictic system. This is unobjectionable as long as we do not confound the two systems. But Bühler’s case is not strong for this network. It is not essential in the route-type description that “I” (the speaker in his imagination) make the moves and turns. If there were a ball rolling through the pattern, the directional terms would be just the same. But a ball doesn’t have deictic perspective. What the speaker in fact does in this description is to use the last directed *path* as the relatum for the subsequent path. The new path is straight, right, or left from the current one. Hence it is the *intrinsic* orientation of the current path that is taken as the relatum.

4. I am ignoring a further variable, the listener’s viewpoint/orientation. Speakers can and often do express spatial relations from the *interlocutor’s* perspective, as in *for you, the ball is to the left of the chair*. Conditions for this usage have been studied by Herrmann and his colleagues (cf. Herrmann and Grabowski 1994).

5. Here I am considering only one case of nonalignment, namely, a 90° angle between the relevant bases. Another case studied by Carlson-Radvansky and Irwin (1993) is 180° nonalignment.

6. Carlson-Radvansky and Irwin do not discuss item-specific effects, although it is likely that the type of relatum used is not irrelevant. It is the case, though, that their statistical findings always agree between subject and item analyses. Another point to keep in mind is that the experimental procedure may invite the development of “perspective strategies” on the part of subjects, and occasionally the employment of an “unusual” perspective.

7. Carlson-Radvansky and Irwin included several scenes that were formally of the same type as scene (g) in figure 3.8, among them the one in figure 3.9 with fly 2.

8. There is, however, no reason why this should also hold in other cultures. Stephen Levinson (personal communication), for instance, has presented evidence that the principle does not hold for speakers of Tzeltal, who can use their intrinsic system when the relatum’s critical dimension is not in canonical orientation. But the Tzeltal intrinsic system differs substantially from the standard average European (SAE) intrinsic system (see Levinson 1992a). What is intrinsic *top/bottom* in SAE is “longest dimension” or the “modal axis” of an object in Tzeltal; the former, but not the latter, has a connotation of verticality.

9. These numbers differ from those reported in Levelt (1982b) because the present selection criterion is a different one.

10. My criterion for ellipsis was a strict one. There should, of course, be no directional term, but there also should be no coordination that can be interpreted as one directional term having scope over two constituents, as in *From pink right successively yellow and blue* or *A road turns right from pink and meets first yellow and then blue*. I have excluded all cases where subjects mention a line on which the nodes are located.

11. The case occurs in a deictic description of the fourth pattern down the first column in figure 3.14. It goes as follows. *From there left to a pink node. And from there to a green node.* This obviously violates both models of ellipsis. I prefer to see it as a mistake or omission.

12. The discussion that follows in the text is much inspired by discussions with Stephen Levinson.

## References

- Brown, P., and Levinson, S. C. (1993). Linguistic and nonlinguistic coding of spatial arrays: Explorations in Mayan cognition. Working paper no. 24, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.
- Bühler, K. (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer. A major part on deixis from this work appeared in translation in R. J. Jarvella and W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics*, 9–30. Chichester: Wiley, 1982.
- Byrne, R. M. J., and Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language*, 28, 564–575.
- Carlson-Radvansky, L. A., and Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, 46, 223–244.
- Carlson-Radvansky, L. A., and Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language*, 33, 646–671.
- Ehrich, V. (1982). The structure of living space descriptions. In R. J. Jarvella and W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics*, 219–249. Chichester: Wiley.
- Friederici, A. D., and Levelt, W. J. M. (1990). Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception and Psychophysics*, 47, 253–266.
- Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition*, 31, 45–60.
- Hankamer, J., and Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*, 7, 391–426.
- Herrmann, T., and Grabowski, J. (1994). *Sprechen: Psychologie der Sprachproduktion*. Heidelberg: Spektrum.
- Hill, A. (1982). Up/down, front/back, left/right: A contrastive study of Hausa and English. In J. Weissenborn and W. Klein (Eds.), *Here and there: Crosslinguistic studies on deixis and demonstration*, 13–42. Amsterdam: Benjamins.
- Levelt, W. J. M. (1981). The speaker's linearization problem. *Philosophical Transactions of the Royal Society, London, B95*, 305–315.
- Levelt, W. J. M. (1982a). Linearization in describing spatial networks. In S. Peters and E. Saarinen (Eds.), *Processes, beliefs, and questions*, 199–220. Dordrecht: Reidel.
- Levelt, W. J. M. (1982b). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella and W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics*, 251–268. Chichester: Wiley.

- Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. van Doorn, W. van de Grind, and J. Koenderink (Eds.), *Limits of perception: Essays in honour of Maarten A. Bouman*, 323–358. Utrecht: VNU Science Press.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA.: MIT Press.
- Levinson, S. C. (1992a). Vision, shape, and linguistic description: Tzeltal body-part terminology and object description. Working paper no. 12, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.
- Levinson, S. C. (1992b). Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. Working paper no. 13, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.
- Miller, G. A., and Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
- Shepard, R. R., and Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, 18, 161–193.
- Slobin, D. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, and H. Filip (Eds.), *Berkeley Linguistics Society: Proceedings of the Thirteenth Annual Meeting*, 435–444. Berkeley: Berkeley Linguistics Society.
- Talmy, L. (1983). How language structures space. In H. Pick and L. Acredolo (Eds.), *Spatial orientation: Theory, research, and application*. New York: Plenum Press.
- Taylor, H. A., and Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language* (in press).
- Tversky, B. (1991). Spatial mental models. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, vol. 27, 109–146. New York: Academic Press.