1

On the very idea of a frame of reference[1]

Jérôme Dokic & Elisabeth Pacherie

Institut Jean-Nicod (CNRS-EHESS-ENS), Paris

---

0. Introduction

It is widely assumed, both in philosophy and in the cognitive sciences, that perception essentially involves a relative or egocentric frame of reference. In his discussion of a variant of Molyneux's question concerning the relationship between the frames of reference used in particular languages and the frames of reference involved in non-linguistic spatial representations, Levinson has explicitly challenged this assumption. Instead, he argues in favour of the 'neo-Whorfian' hypothesis that the frame of reference dominant in a given language infiltrates spatial representations in non-linguistic, and in particular perceptual, modalities.

Our aim in this paper is to assess Levinson's neo-Whorfian hypothesis at the philosophical level and to explore the further possibility that perception may not just use frames of reference other than the relative one but may also, in some cases at least, be perspective-free[2].

We shall proceed as follows. First, we shall introduce Levinson's variant of Molyneux's question (section 1) and his useful taxonomy of frames of reference (section 2). A logical reconstruction of Levinson's argument will then be given (section 3). The rest of the paper will focus on perceptual and linguistic representations. We shall first discuss four

---

[1] This paper is based on earlier joint work on frames of reference; see Dokic & Pacherie (1999).
[2] We adopt Levelt's terminology (1996) and call perspective-free the representations that do not involve frames of reference and perspective-bound those that do.

important differences between them, which arguably are compatible with Levinson's argument (section 5). Then we shall ask whether, contrary to what Levinson claims, perception is necessarily perspective-bound (section 6). In the next section, the notion of implicit frame of reference is introduced, and we shall examine its relevance to the part of Levinson's argument which concerns the compatibility of frames of reference across modalities (section 8).

1. Generalizations of Molyneux's Question

In 1688, William Molyneux wrote a letter to John Locke posing the following celebrated question, which Locke reproduced in his *Essay Concerning Human Understanding* (1690: II, ix, 8):

> Suppose a man born blind, and now adult, and taught by his touch to distinguish between a cube, a sphere of the same metal, and nighly of the same bigness, so as to tell, when he felt one and the other, which is the cube, which is the sphere. Suppose then the cube and sphere placed on a table, and the blind man to be made to see: *Quaere*, whether by his sight, before he touched them, he could now distinguish, and tell, which is the globe, which the cube.

> Nowadays, this question is emblematic of the issue of whether spatial representations are modality-specific or not. In other words, are spatial representations specific to a modality such as vision or touch, or are there supramodal or amodal ways of representing space?

> In its original formulation, Molyneux's Question was about representations of *shape*, but it can be generalized to representations of *locations* and *directions*. For instance, we might devise the following variant of Molyneux's Question, which deals with the auditory modality:

> Suppose a man born blind, and now adult, and taught to locate in space two auditory sources, one emitting a low note and the other a high note. Suppose then the two objects placed on a table, and the blind man to be made to see: *Quaere*, whether by his

sight and without hearing anything, he could now tell the directions which the notes came from.

In his well-known essay on Molyneux's Question, Evans (1985) in effect tried to reduce the original Question to variants of this kind. Now there is another direction in which Molyneux's Question can be generalized. The questions formulated so far are about spatial representations in *sensory* modalities. However, *non-sensory* modalities also involve spatial representations. In other words, there are conceptual and linguistic representations of space. So we might think of other variants of Molyneux's Question which deal with linguistic representations of locations, such as the following:

Suppose a man born blind, and now adult, is told about the location of a cube and a sphere in front of him. Suppose that the blind man is made to see: *Quaere*, whether by his sight and without any further testimony, he could now tell where is the cube and where the sphere.

One motivation for treating perception, conception and language as *modalities* is that there seems to be a *unifying concept* that can be invoked to describe our representations of locations at all three levels. This is the concept of a *frame of reference*. As Pinker (1997: 262) puts it, 'reference frames are inextricable from the very idea of location'. Indeed, frames of reference are regularly invoked to characterize the structure of spatial maps and the modes of presentation of locations in various modalities.[3]

As far as language is concerned, the expression of spatial relations involves frames of reference which can vary from one language to the other. The issue we would like to discuss concerns the relationship between the frames of reference involved in particular languages and the frames of reference involved in non-linguistic spatial representations, especially in perception. This issue has been addressed in Levinson's work on frames of reference (see Brown & Levinson 1993, 1994; Levinson 1996; Brown & Levinson 2000; Levinson 2001). In particular, Levinson gives Molyneux's Question a new twist by linking it with the following questions:

---

[3] For instance, see the section on frames of reference in Eilan et al. (1993), and especially the introduction by Bill Brewer and Julian Pears. See also Campbell (1994) and Pinker (1997).

Levinson's Questions (1996: 153):

1. 'Do the different representational systems natively and necessarily employ certain frames of reference?'

2. 'If so, can representations in one frame of reference be translated (converted) into another frame of reference?'

Levinson's own answers to these questions are based on a 'neo-Whorfian' hypothesis. The original Whorfian hypothesis (see Whorf 1956) was that our perception and conception of the world are *determined* by the semantic and grammatical organization of our language. On Levinson's weaker, neo-Whorfian hypothesis, the various modalities can in principle operate in their own, proprietary frames of reference, but Whorfian effects arise because of the need for them to coordinate and exchange information. In Levinson's own words (1996: 157), first, 'the frame of reference dominant in a given language 'infiltrates' other modalities, presumably to ensure that speakers can talk about what they see, feel, and so on' and, as a consequence, 'other modalities have the capacity to adopt, or adapt on, other frames of reference, suggesting a yes answer to Mr. Molyneux.'

2. Three kinds of frames of reference

Since its modern introduction by Gestalt psychology, the notion of frames of reference has been widely used in various disciplines, such as linguistics, psychology, philosophy and the neurosciences. One of Levinson's important achievement is his systematic typological work in surveying and simplifying the various distinctions to be found in the relevant literature. In this section, we shall describe what Levinson sees as the three main kinds of frames of reference, namely intrinsic, absolute and relative frames of reference.

Insert Figure 1 about here

Basically, any frame of reference involves the selection of *reference objects* (the referent and the relatum) and determines the way the spatial relation between them is represented.

The simpler frame of reference, and perhaps the more widespread in natural languages, is the *intrinsic frame of reference*. It allows for the representation of binary relations between referent and relatum. The origin is fixed on the relatum, and the coordinate system is determined by intrinsic properties of the relatum. These properties exploit the relatum's asymmetries and functions, often in a culturally specific way.

The *absolute frame of reference* also allows for the representation of binary relations between referent and relatum. Similarly, the origin is fixed on the relatum, but the coordinate system is determined by fixed bearings in the environment, such as the cardinal directions.

Finally, the *relative frame of reference* allows for the representation of ternary relations among referent, relatum and point of view. The origin of the primary coordinate system is the point of view.[4]

It is important to note that the use of each frame of reference involves different cognitive abilities and resources. First, the use of an intrinsic frame of reference requires the ability to *identify* and *analyze* the relevant relatum – its orientation, functional and dynamical properties (Levelt 1996: 87). Second, the use of an absolute frame of reference requires the ability to *keep track of the relevant fixed bearings* over time (dead reckoning might be necessary). Third, the use of a relative frame of reference requires in particular the ability to *keep track of one's left and right*, which goes hand in hand with the ability to *recognize enantiomorphs across perceptual contexts* (e.g. to recognize a left hand from one scene to another).

As it is now well-known, particular languages do not use the same kinds of frames of reference. As Levinson points out, some languages predominantly use only one kind of frame of reference, which can be the absolute one or the intrinsic one. It appears that if a given language uses a relative frame of reference, it also uses an intrinsic one. Other languages use two kinds of frames of reference, which can be either the relative and the intrinsic ones, or the absolute and the intrinsic ones. Finally, some languages, including English, Dutch and French, use all three kinds of frames of reference.

---

[4] There can be a secondary coordinate system (often involving the "front"/"back" terms) whose origin is the relatum. Moreover, the point of view itself can be fixed or mobile.

3. The case of Tzeltal

Levinson and his collaborators conducted a series of experiments in order to understand the relationship between frames of reference used in linguistic and non-linguistic representations. More precisely, these experiments were designed to test the influence of linguistic frames of reference on the coding of spatial relations in non-linguistic tasks.

The target language of most of these experiments is Tzeltal, a Mayan language spoken in Tenejapa (Chiapas, Mexico). The peculiarity of Tzeltal is that it does not seem to involve a relative frame of reference; it has no words for 'left' and 'right'. Tzeltal involves an absolute frame of reference, which is used for coding the spatial relations between objects separate in space, and an intrinsic frame of reference, which is used only for coding the spatial relations between parts of a single object or between contiguous objects.

Tzeltal's absolute frame of reference is derived from a topographical feature; Tenejapa is a large mountainous tract which tends to fall in altitude toward the north-northwest. In particular, the term we might translate by 'downhill' designates approximately the southern direction, while the term we might translate by 'uphill' designates approximately the northern direction. There is only one word, which we might translate by 'across', designating indifferently the eastern and the western directions. Note that this frame of reference is still used when Tenejapans are transported outside their territory:

Insert Figure 2 about here

In his experiments, Levinson compared the performance of speakers of Tzeltal with that of speakers of Dutch in tasks of recognition and spatial reasoning.

In a typical experiment, the subject sees an arrow on a table. The arrow points to her right, or objectively to the north. The subject is then rotated 180 degrees to face another table. There are two arrows on the second table, one pointing to her right, the other to her left. The subject is asked to identify 'the arrow like the one he saw before' (1996: 114).

Insert Figure 3 about here

The results of these experiments is that Dutch subjects predominantly use a relative frame of reference (i.e. they tend to preserve the orientation of the arrow relative to themselves), whereas Tenejapans predominantly use an absolute frame of reference (i.e. they tend to preserve the orientation of the arrow relative to their environment). In another study, Levinson and Brown (1994) have shown that Tenejapans make no essential use of the notion of left and right in daily life. As a consequence, they are relatively insensitive to left/right inversions of enantiomorphs.

4. A reconstruction of the neo-Whorfian argument

Levinson's experiments have been criticized, for instance by Li & Gleitman (2002). These authors have shown that speakers of languages using the three kinds of frames of reference can choose any of them depending on the context. For instance, if there are visible landmarks around, some native speakers of English have a preference for an absolute frame of reference, just as speakers of Tzeltal. However, it is not our intention to criticize the details of Levinson's experiments. Rather, we would like to reconstruct the main logical steps of Levinson's neo-Whorfian argument in order to assess it at a more conceptual level. We thus aim at a better understanding of the notion of a spatial frame of reference and of the nature of cross-modal transfer of spatial information.

The neo-Whorfian argument

1. Any modality of spatial representation (perception, conception, or language) operates in some *frame of reference*.

2. There is a cross-modal sharing of information. For instance, we can talk about what we have seen, we can gesture about what we have explored by touch, and so on.

3. Such cross-modal sharing of information would be possible under any of the following conditions:

3.a. One can translate spatial information from one frame of reference to another.

3.b. A single frame of reference operates in the relevant modalities.

3.c. Some modalities can adopt various frames of reference depending on the context.

4. Against 3.a: There are constraints on translation, so that one cannot freely convert information from one frame of reference to another.

5. Against 3.b: It is not the case that all modalities operate in the same frame of reference.

6. Hence, from 3, 4 & 5, some modalities should be able to adopt various frames of reference depending on the context.

7. Some languages, such as Tzeltal, have only one frame of reference (in the case of Tzeltal, an absolute one).

8. Hence, insofar as one can express linguistically spatial representations acquired from other modalities (2), these modalities should be able to adopt the dominant (perhaps unique) linguistic frame of reference.

The argument is valid, so we have to reject one or more premisses if we do not accept the neo-Whorfian conclusion. In what follows, our targets will be the following:

Contra premiss 1: Spatial information carried by non-linguistic representations need not be based on a frame of reference. In other words, there is a sense in which non-linguistic spatial representations can be perspective-free. This style of objection will be developed in sections 6 and 7, focussing on perceptual representations.

Contra premiss 3: Cross-modal transfer of information need involve neither a unique frame of reference nor translatable frames. The model of translation is misleading to characterize the transfer of information across modalities. This objection will be developed in section 8.

5. Four differences between perceptual and linguistic spatial representations

To begin with, let us mention four interesting differences between perceptual and linguistic spatial representations.

1. Linguistic representation of space is *selective*. In contrast, non-linguistic representation of space always carries *nested information* (Dretske 1981).

In perception, spatial information is always *nested in* other perceptual information. Normally, one cannot perceive a fact about an object without perceiving many other facts about it. Perceptual information is *dense*. This holds for spatial relations. Normally, one does not just perceive the distance between the chair and the bottle; one perceives many other spatial facts involving these objects, such as their spatial relations to other objects or to the background.

On the other hand, conceptual and linguistic representations carry more abstract information and allow a *selective* representation of the scene. For instance, the distance between the chair and the bottle can be specified in language (at least in a coarse way; see below) independently of a linguistic representation of other spatial relations involving these objects.

2. Linguistic representation of space is often *coarse*. Non-linguistic representation of space is *analogue* (Peacocke 1989).

In perception, it seems that any value of the spatial dimensions may enter the fine-grained content of the representation. For instance, within certain psychophysical limits, any distance between two objects can be determinately represented. In contrast, our linguistic description of a visual scene will often be less determinate (unless of course we use precise numerical values).

3.   Any region of a represented scene can be *directly accessed* through perception, whereas some of them can only be *indirectly accessed* through language.

There is an indefinite number of different regions and directions in the perceptual field, but they can all be directly identified. For instance, we do not perceive the region of the visual field which we call 'up to the left' as a product of two localisations: 'up' and 'to the left'. We could express this fact by saying that orientation in perception is homogenous; all egocentric regions and directions are perceived in the same way, i.e. without the mediation of a system of coordinates. This is not to deny that there are privileged dimensions in perception (such as the gravitational axis). The claim is rather that these dimensions do not act as axes *relative to which* regions and directions have to be located.

4.   Often, the origin of a coordinate system cannot be *dissociated from the figures to be located*. In language, such a dissociation is possible.

For instance, if perception represents things from a point of view, namely the perceiver's, the latter is fixed at any given time and cannot be changed (except in imagination). I cannot *perceive* the world from another's point of view. In contrast, I can describe the bottle as being to your left, even if I perceive it as being to my right.

These differences, between perceptual and linguistic spatial representations, are substantial, but arguably they do not compromise the possibility of translation or conversion between non-linguistic and linguistic frames of reference. They do not contradict the general claim that perception and language can use the same *kind* of frame of reference. Let us now turn to other, potentially more significant differences between perception and language as far as spatial representation is concerned.

6. Perception and frames of reference

In what sense, if any, does perception involve a frame of reference? In this section, we shall try to answer this question with respect to each of the three kinds of frames of reference we have distinguished following Levinson. We shall argue that none of the three kinds of frames of reference is essential to perception. We shall further suggest that one possibility is that

perception is in and of itself perspective-free (in the sense that it need not involve any frame of reference). The reason it may appear perspectival is that frames of reference are imposed on the perceptual scene at the conceptual and linguistic levels.[5]

6.1. Intrinsic frames of reference

In *How the Mind Works*, Pinker speaks of frames of reference as 'overlaying the visual field'. In particular, he argues that 'objects themselves can plot out reference frames' (1997: 266-7).

As an example, let us consider the spatial array depicted at the bottom of figure 4 below. One way of *seeing* this array is by representing the bottle as being in front of the chair. It may seem obvious that this example and others of the same kind show that perception is perspective-bound, and in particular that it involves intrinsic frames of reference. Such frames of reference are intrinsic in the sense that they exploit functional or geometrical properties of the perceived objects.

However, it is not clear to us that the notion of a frame of reference is needed to explain the relevant phenomena. There is another way of seeing the array, namely by representing the bottle as being next to the chair. Of course there is a difference between seeing the bottle as being *in front of* the chair and merely seeing it as being *next to* the chair. The latter relation (x is next to y) is perspective-free and requires no analysis of the objects in component parts (it is enough that they be seen as two different wholes), whereas the former relation (x is in front of y) is perspective-bound and requires such an analysis. Yet it may be argued that insofar as we take into account the existence of geometrical and/or functional properties that are perceived *over and above* the spatial relations between elements in the scene, perspective-bound relations can be reduced to perspective-free relations. In the case of the bottle in front of the chair, we perceive the internal structure of the chair, which is spatially asymmetrical, so that we perceive the bottle not merely as being next to the chair, but as having various spatial relations to *parts* of it. To say that we perceive the bottle as in front of the chair is tantamount to saying that we perceive it as being near a particular side of the chair, namely its front.

---

[5] This claim is also made by Levelt (1996), although our arguments for it are different.

In general, perceptual analysis of an object in terms of geometrical and/or functional properties selects sides of the object, thus creating further terms for perceivable spatial relations. It follows that the differences that are sometimes described as perspectival differences really are objective differences in the represented scene itself. Similarly, the asymmetry between figure and relatum, which is characteristic of the intervention of a frame of reference, has not been shown to have perceptual significance over and above what is perceptually represented. The claim that perception is perspective-free is not affected by such examples of so-called intrinsic frames of reference.

6.2. Relative frames of reference

The claim that perception uses a relative frame of reference lies deep in a whole tradition of thought, from Kant to phenomenology and recent analytic philosophy. Here are some quotations from recent work:

> Our perceptual field has an orientational structure, a foreground and a background, an up and down. [...] This orientational structure marks our field as essentially that of an embodied agent (Taylor 1978: 154; quoted by Evans 1982: 156).

> [Let us] reflect upon how we might specify the spatial information which we imagine the perception to embody. The subject hears the sound as coming from such-and-such a position, but how is the position to be specified? Presumably in *egocentric* terms (he hears the sound as up, or down, to the right or to the left, in front or behind). These terms specify the position of the sound in relation to the observer's own body; and they derive their meaning in part from their complicated connections with the subject's *actions* (Evans 1982: 155).

It may seem like a plain phenomenological fact that we always perceive objects as being to the left, to the right, above or below. However, we think that another of Levinson's great merits is to have shown, or at least suggested, that what we consider as a necessary condition of perception is in fact quite contingent. As Brown and Levinson put it:

Kant was wrong to think that the structure of spatial regions framed on the human frame, and in particular the distinctions based on left and right, are in some sense essential human intuitions (1994: 9).

In particular, speakers of Tzeltal do not possess the notions of left and right. Of course, the fact that they lack these notions does not yet show that they do not *perceive* the world as being left-right oriented. However, once it has been shown that our cognitive scheme is not necessarily sensitive to distinctions based on left and right, the claim that perception must involve a relative frame of reference loses much of its force.

What are the arguments for the claim that perception must involve a relative frame of reference? One argument is that we need such a frame to recognize enantiomorphs across perceptual contexts. However, the case of Tzeltal shows that we can perceive the world while being relatively indifferent to enantiomorphs. Moreover, recognition across perceptual contexts might involve representations that are not strictly perceptual. The claim that at some level we *encode* perceptual scenes using a relative frame of reference does not show that perception itself is perspective-bound.

Another common argument in favour of the claim that perception must involve a relative frame of reference concerns the connection between perception and action. Evans writes that 'egocentric terms are the terms in which the content of our spatial experiences would be formulated, and those in which our immediate behavioural plans would be expressed' (1982: 154). However, it is not clear that perception has to use a relative frame of reference in order to *distinguish* directions in the perceptual scene, such as the left/right direction as opposed to the right/left direction. In each perceptual context, the relevant distinction can be drawn in *demonstrative* terms, for instance as the direction that goes from *here* to *there*. In general, acting on a location in the perceptual scene can be direct, in the sense that it does not rely on an explicit identification of the acted upon location as opposed to other locations in the scene.[6]

In conclusion, although we naturally *specify* the contents of perception in egocentric terms, perception itself need not involve a relative frame of reference. To this extent we agree

---

[6] For further development of this idea, see Dokic (2003).

with Levinson. However, his point is that since the Tenejapans do not have a relative frame of reference in their language, their perception is not structured in terms of such a frame. Our conclusion is rather different. Apart from the remarkable fact that Tenejapans are by and large indifferent to enantiomorphs, their perceptual phenomenology might be pretty much the same as ours. Despite the fact that our language does possess a relative frame of reference, it is quite possible that, just like speakers of Tzeltal, many of our perceptions are not structured in terms of such a frame, and can be perspective-free.

6.3. Absolute frames of reference

In his *Philosophical Remarks*, Wittgenstein famously claimed that visual space involves a frame of reference:

> We can also say visual space is an oriented space, a space in which there is an above and below and a right and left.

Despite the use of terms like 'above', 'below', 'right' and 'left', who would seem to belong to *relative* frames of reference, Wittgenstein insists that the visual frame of reference is not relative to anything, but is *absolute*:

> And *this* above and below, right and left have nothing to do with gravity or right and left hands. It would, e.g., still retain its sense even if we spent our whole lives gazing at the stars through a telescope. (1975: §206)

Interestingly, Russell defended a similar view in *Human Knowledge*:

> At every moment, what is in the centre of my visual field has a quality that may be called 'centrality'; what is to the right is 'dexter', what to the left 'sinister', what above 'superior', what below 'inferior'. These are *qualities* of the visual datum, not relations (1948: 316).[7]

---

[7] On Russell's theory, see Casullo (1986).

Wittgenstein gives the following argument for the claim that visual space involves an absolute frame of reference:

Couldn't we imagine a visual space in which we would only perceive spatial relations, but not absolute positions? [...] I don't believe we could.

In visual space there is absolute position and hence also absolute motion. Think of the image of two stars in a pitch-black night, in which I can see nothing but these stars and they orbit around one another (1975: §206).

Is this argument cogent? Wittgenstein's point holds at the level of what he called a 'phenomenological language', which he contrasts with a 'physical language'. Phenomenologically speaking, visual space is exactly how it appears to us; there is no relevant distinction between visual space as it appears to us and visual space as it really is. As a consequence, if we can see two stars orbiting around one another without seeing any relational change in the scene, visual space must be absolute. However, it does not follow that physically speaking, it is absolute. The stars actually change their spatial relations to at least parts of our body (including parts of the retina). There is more to the scene than what appears to us. Our perceptual systems can be sensitive to relational changes without representing them explicitly. As situated perceivers, we can perceptually represent the motion of an object without representing the necessary changes in its spatial relations to some bodily sides. From the point of view of a physicalist theory of visual space, Wittgenstein's example shows at best that we can perceive motion in visual space independently of a *representation* of relational changes. It does not show that perception involves an absolute frame of reference.

In this section, we have tried to show that neither intrinsic nor relative nor absolute frames of reference are essential to perception. We have also tentatively argued that perceptual representations can be perspective-free. This contradicts premiss 1 of Levinson's argument (section 4), according to which spatial representations must operate in some frame of reference.

7. Implicit frames of reference

Our account of Wittgenstein's example might raise the following objection. We have argued that perception need not involve a relative frame of reference (section 6.2). At the same time, we have criticised Wittgenstein's claim that visual space is absolute, by pointing out that represented locations are individuated at a pre-representational level by their relations to bodily sides. This seems to involve a contradiction. If perceived locations are individuated by their relations to the perceiver's body, perception does involve a relative frame of reference after all.

To answer this objection, it is necessary to distinguish two notions of frames of reference. Perception need not involve a relative frame of reference in the sense that what it represents would have to be described by means of egocentric terms such as 'left' and 'right'. Perception is not necessarily egocentric in this sense. However, it can be said to be egocentric in a second, different sense. When we perceive objects and other spatial entities, we *exploit* our *actual* spatial situation relative to them. This relative situation need not be represented as such in perception. All perception is *implicitly* egocentric; it represents only *local* objects, regions, and directions, but it need not represent them *as* local. The implicit egocentricity of perception is well described by Perry in the opening passage of his essay 'Thought without representation':

> I see a cup of coffee in front of me. I reach out, pick it up, and drink from it. I must then have learned how far the cup was *from me*, and in what direction, for it is the position of the cup relative to me, and not its absolute position, that determines how I need to move my arm. But how can this be? I am not in the field of vision: no component of my visual experience is a perception of me. How then can this experience provide me with information about how objects are related to me? (1993: 205).

In the next section, we shall try to go some way toward alleviating the perplexity Perry expresses here. In the meanwhile, the point is that even 'selfless' perception exploits at a pre-representational level one's spatial situation relative to the perceived scene. Perception represents things from a point of view which itself is not normally represented in our experience.

We can use the notion of a frame of reference to capture the second sense of the claim that perception is egocentric. Perception involves a frame of reference centred on a point of view, namely the perceiver's body, and whose axes are defined on the basis of our bodily axes. However, it is important to realise that this notion of a frame of reference is very different from the one used by Levinson. According to Levinson, a relative frame of reference allows for the representation of ternary relations among referent, relatum and point of view. In contrast, in order for something to be perceived, neither our own body nor its spatial situation relative to what is perceived must be represented. Levinson uses the notion of a frame of reference in such a way that the referent, the relatum and the point of view have to be explicitly represented. Perception is egocentric in another sense. It must be possible to exploit spatial relations between parts of our body and the world in order to perceive anything, but this is a condition of possibility of spatial represent*ing*, not something that is itself spatially represent*ed*. If there is a frame of reference here, it is *implicit*, not part of what is explicitly represented.[8] To use a linguistic analogy, implicit frames of reference pertain to the *syntax* of perception, and only indirectly to its semantics.[9]

So we can maintain the claim that perception need not involve any explicit intrinsic, relative or absolute frame of reference and in that sense can be perspective-free. However, the possibility of implicit frames of reference has an interesting bearing on Levinson's claim about the commensurability between different kinds of frames of reference. To this last issue we now turn.

8. The cross-modal flow of information

According to Levinson, the cross-modal flow of spatial information presupposes shared frames of reference: 'we will not be able to exchange information across internal representation systems that are not based on one and the same frame of reference' (1996: 155). Indeed, a crucial premiss of the neo-Whorfian argument is that one cannot freely convert

---

[8] The notion of an implicit frame of reference is related to Campbell's notion of an intensional egocentric frame of reference. As he writes, 'when the subject is identifying places egocentrically, he cannot be thought of as doing so by first identifying a physical thing, himself, through a body image, and then identifying places by their relation to his body.' (1994: 13). In other words, the egocentric localisation of a position in space does not depend on a prior identification of a body and does not presuppose an explicit representation of oneself as a term of a spatial relation to the position.

[9] For further thought on these lines, see Pacherie (2003).

information from one frame to another. Levinson gives the following summary of the compatibilities and incompatibilities of frames of reference (Levinson 1996: 154):

Insert Figure 3 about here

For instance, the representation 'bottle in front of chair', involving an intrinsic frame of reference, cannot be used to ground the representation 'bottle to right of chair', involving a relative frame of reference. In general, the only directions in which conversion is possible are from the two orientation-bound frames (relative and absolute) to the orientation-free one (intrinsic), and then only if the orientation of the ground object is fully specified.

We have two remarks on this premiss of Levinson's argument. The first is fairly obvious, but perhaps worth reminding. When Levinson argues for untranslatability across frames of reference, he explicitly asks us to 'discard other information'. It is only if we do so that we cannot generate, for instance, the relative description 'bottle to right of chair' from the intrinsic description 'bottle in front of chair'. Such limitations occur only when other information is discarded. In contrast, appropriately nested representations are convertible. For instance, one may use a relative or an intrinsic mode of representation and nevertheless encode cardinal directions systematically enough to allow for a translation into an absolute mode of representation.

The point is relevant to the case of Tzeltal. As we saw, Tenejapans keep track of absolute directions, but it does not follow that they perceive or even conceive the world with the help of an absolute frame of reference. Keeping track of absolute directions enables them to talk about what they perceive and conceive, by converting perceptual and conceptual information into absolute linguistic descriptions. There is no need to suppose that their language influences the way in which representations are coded at the perceptual and conceptual levels. As far as translatability constraints between frames of reference are concerned, perceptual and conceptual representations might involve any non-relative frame of reference, or indeed no frame of reference at all.

Our second remark is more general, and concerns Levinson's model of cross-modal sharing of spatial information. Information flows from a source representational state to a

target representational state, i.e. from one represent*ing* to another. Now Levinson seems to assume that the target representational state can only exploit what is explicitly represented in the source state. Thus, he uses the image of translation to capture what is involved in cross-modal flow of information. However, Levinson's assumption is unwarranted. Even if a source representational state has a content which cannot be translated into a target representation involving a given frame of reference, there can be enough information associated with the state (although not part of its content) to ground a reliable perspective-bound representation. In other words, relevant information which is not explicitly represented in the source state can be implicitly nested in, or associated with it.

Suppose for instance that I perceive a bottle next to a chair. We have argued that my perceptual representation might be perspective-free; in particular, I do not need to perceptually represent the bottle and the chair *as* bearing different spatial relations to parts of my body. As a consequence, there is no way of translating what is represented in perception ('bottle next to chair') into a relative description (either 'bottle right to chair' or 'bottle left to chair'). It does not follow that no relative description can be grounded on perception. The perceptual state supervenes on various mechanisms that carry information about spatial relations between elements in the scene and bodily sides. Such information is not part of what is perceptually represented, but can be exploited by a cognitive system in order to produce the perspective-bound representation 'bottle right to chair'. Perhaps a conceptual representation based on perception can be sensitive not only to what is perceptually represented in perception, but also to the *mode* or *manner* of perceiving the scene.

In a nutshell, the cross-modal flow of information requires *reliable* transitions between representational states. A reliable transition can be sensitive to more than what is explicitly represented in the source state.[10] In particular, it can be sensitive to information implicitly nested in such a state or carried by underlying mechanisms. Levinson's model of translation is too crude to characterize the transfer of information from one mode of representation to another. Information can be exchanged even if there is an incompatibility at the level of content or what is explicitly represented.

---

[10] This point also holds of other kinds of transition. For instance, the transition from the belief that it's raining to the belief that one believes that it's raining can be reliable (it is an instance of reflection), and is sensitive both to what is explicitly represented (that it's raining) and to the mode of the representation (the fact that what is represented is believed). For an account of such a transition, see Peacocke (1999).

From these two remarks, we conclude that there is no good argument from translation limitations to neo-Whorfianism, and that Levinson's third premiss (section 4) does not hold.

9. Conclusion

Although we agree with Levinson's negative point that perception need not essentially make use of a relative frame of reference, we do not think he succeeds in establishing his positive neo-Whorfian claim. There are two alternatives to the view that perception essentially involves a relative frame of reference. One is that perception does indeed essentially involve a frame of reference, but that this frame can also be absolute or intrinsic. The other is that perception can be perspective-free, in the sense that it need not involve any explicit frame of reference. Levinson considers only the first alternative and claims that perception tends to adopt the frame of reference dominant in the language of the perceiver. We argued that the second alternative is at least as plausible as the first one. We also argued that once the existence of implicit, as opposed to explicit, relative frames of reference in perception is taken into account, the transfer of information across modalities can be explained without appealing to Levinson's neo-Whorfian hypothesis.

REFERENCES

Bloom, P., Peterson, M. A., Nadel, L., & Garrett, M. F. (eds), 1996, *Language and Space*, Cambridge (Mass.): MIT Press.

Brown, P. & Levinson, S. C., 1993, ''Uphill' and 'downhill' in Tzeltal', *Journal of Linguistic Anthropology* 3 (1), p. 46-74.

Brown, P. & Levinson, S. C., 1994, 'Immanuel Kant among the Tenejapans: Anthropology as Empirical Philosophy', *Ethos* 22 (1), p. 3-41.

Brown, P. & Levinson, S. C., 2000, 'Frames of spatial reference and their acquisition in Tenejapan Tzeltal', in L. Nucci, G. Saxe & E. Turiel (eds), *Culture, thought, and development*, Mahwah, NJ: Erlbaum, p. 167-197.

Campbell, J., 1994, *Past, Space and Self*, Cambridge (Mass.): MIT Press.

Casullo, A., 1986, 'The Spatial Structure of Perceptual Space', *Philosophy and Phenomenological Research* 46, p. 665-671.

Dokic, J., 2003, 'L'espace de la perception et de l'imagination', in J. Bouveresse & J.-J. Rosat (eds), *Philosophies de la perception. Phénoménologie, grammaire et sciences cognitives*, Paris: Odile Jacob, p. 77-99.

Dokic, J. & Pacherie, E., 1999, 'Percevoir l'espace et en parler. Une nouvelle version du problème de Molyneux', ~~*Voir*~~ 19, p. 20-33.

Dretske, F., 1981, *Knowledge and the Flow of Information*, Oxford: Blackwell.

Eilan, N., McCarthy, R. & Brewer, B., 1993, *Spatial Representation. Problems in Philosophy and Psychology*, Oxford: Blackwell.

Evans, G., 1982, *The Varieties of Reference*, Oxford: Clarendon Press.

Evans, G., 1985, 'Molyneux's Question', in *Collected Papers*, Oxford: Clarendon Press.

Levelt, W. J. M., 1996, 'Perspective Taking and Ellipsis in Spatial Descriptions', in P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (eds), p. 77-107.

Levinson, S. C., 1996, 'Frames of Reference and Molyneux's Question: Crosslinguistic Evidence', in Bloom, P., Peterson, M. A., Nadel, L., & Garrett, M. F. (eds), p. 109-169.

Levinson, S. C., 2001, "Covariation between spatial language and cognition, and its implications for language learning", in M. Bowerman & S. C. Levinson (eds), *Language acquisition and conceptual development*, Cambridge: Cambridge University Press.

Li, P. & Gleitman, L., 'Turning the tables: language and spatial reasoning', *Cognition* 83, 2002, p. 265-294.

Locke, J., 1989 [1690], *Essai concernant l'entendement humain*, trad. franç. P. Coste, Paris: Vrin.

Pacherie, E., 2003, 'Modes de structuration des contenus perceptifs visuels', in J. Bouveresse & J.-J. Rosat (eds), *Philosophies de la perception. Phénoménologie, grammaire et sciences cognitives*, Paris: Odile Jacob, p. 263-289.

Peacocke, C., 1986, 'Analogue content', *Proceedings of the Aristotelian Society*, suppl. vol. 60, p. 1-17.

Peacocke, C., 1999, *Being Known*, Oxford: Oxford University Press.

Perry, J., 1993, 'Thought Without Representation', in *The Problem of the Essential Indexical, and Other Essays*, Oxford: Oxford University Press.

Pinker, S., 1997, *How the Mind Works*, New York: W.W. Norton & Company.

Russell, B., 1948, *Human Knowledge*, Allen and Unwin: London.

Taylor, C., 1978, 'The Validity of Transcendental Arguments', *Proceedings of the Aristotelian Society*, 79, p.151-165.

Whorf, B. L., 1956, *Language, thought, and reality,* ed. J. B. Carroll, Cambridge (Mass.): MIT Press.

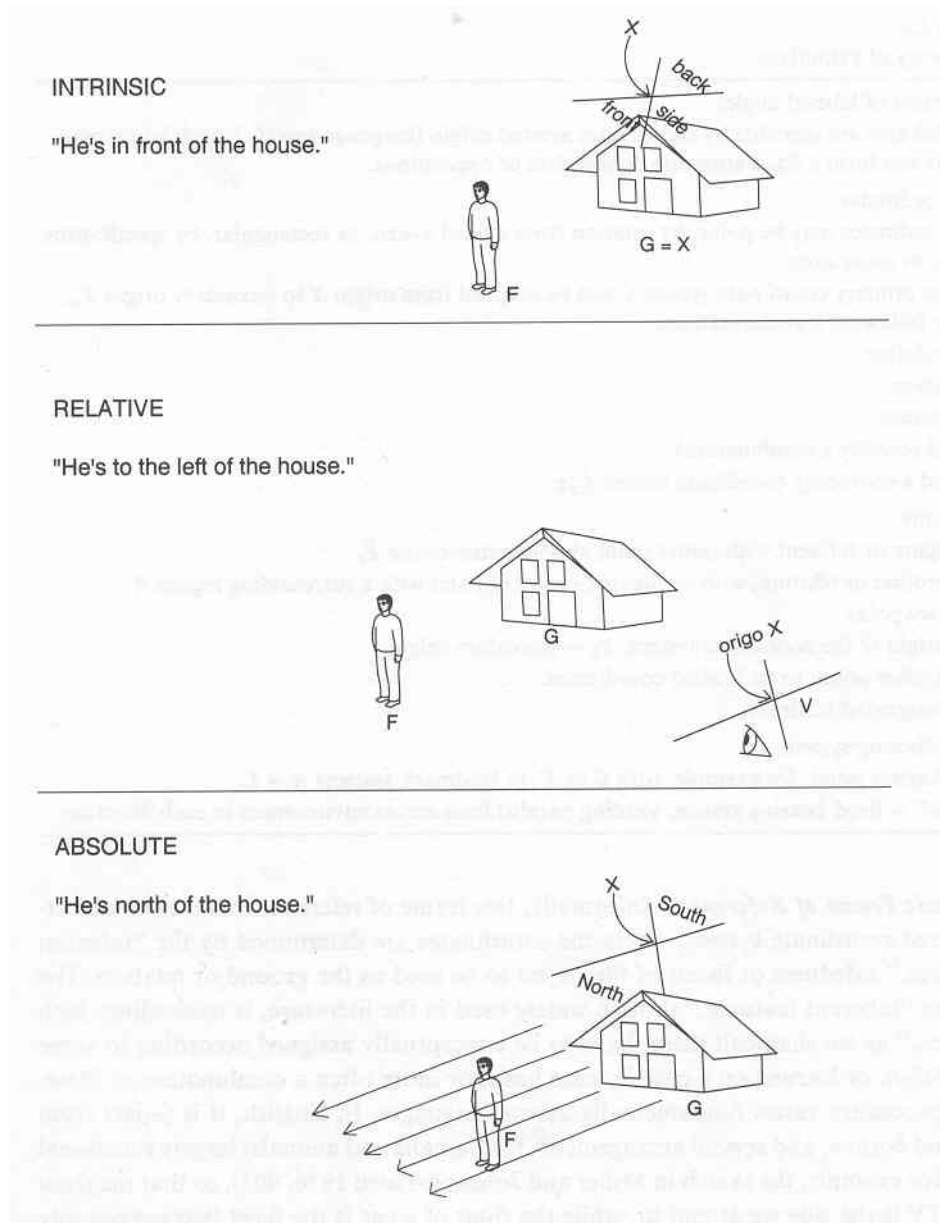Wittgenstein, L., 1975, *Philosophical Remarks*, ed. R. Rhees, Blackwell: Oxford.

INTRINSIC

"He's in front of the house."

RELATIVE

"He's to the left of the house."

ABSOLUTE

"He's north of the house."

Figure 1: Intrinsic, absolute, and relative frames of reference.
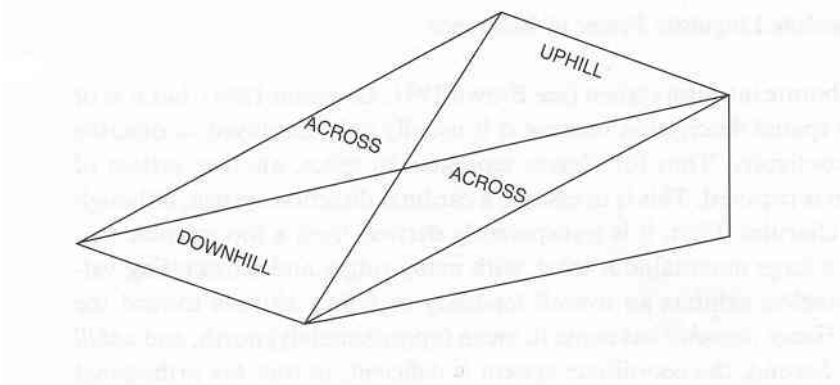
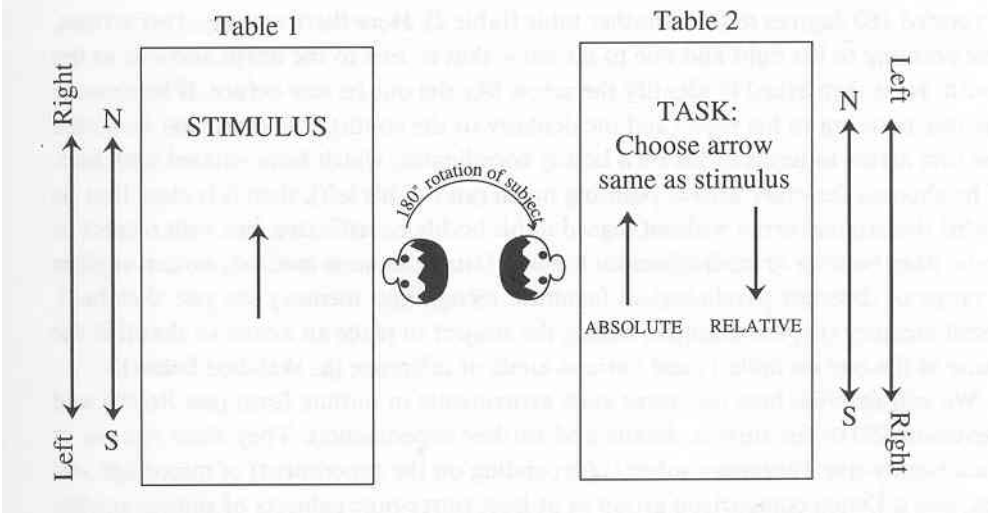**Figure 2**: Topographical representation; from Fig. 4.1 of Levinson 1996: 112.

Figure 3: Design of some experiments; from Figure 4.2. of Levinson 1996: 113.
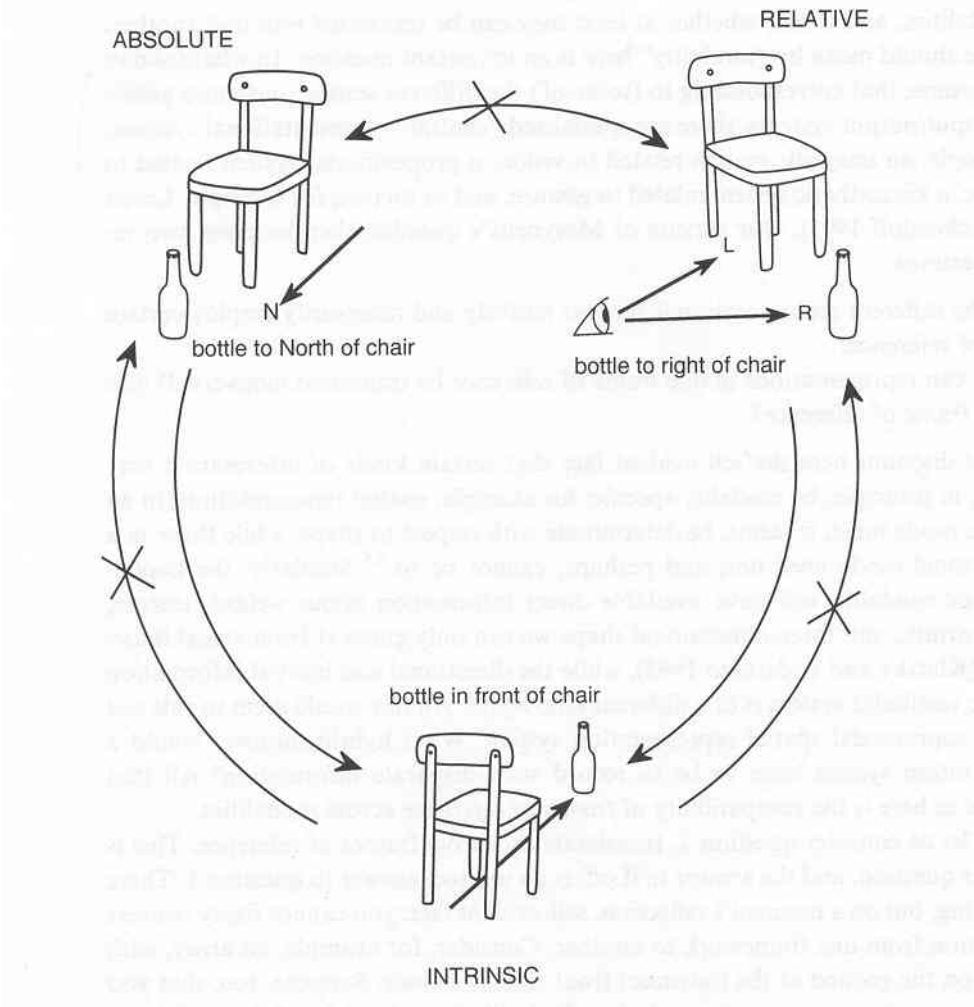
Figure 4: Untranslability across frames of reference; from Figure 4.11 of Levinson 1996: 154.