Frames of Reference and Molyneux's question: cross-linguistic evidence

Stephen Levinson

Contents

Frames of Reference and Molyneux's question: cross-linguistic evidence

1.0 What this is all about[1]

The title of this paper invokes a vast intellectual panorama; yet instead of vistas, I will offer only a twisting trail. The trail begins with some quite surprising cross-cultural and cross-linguistic data, which leads inevitably on into intellectual swamps and mine-fields - issues about how our ‘inner languages’ converse with one another, exchanging spatial information.

To preview the data: first, languages make use of different frames of reference for spatial description. This is not merely a matter of different use of the same set of frames of reference (although that also occurs); it is also a question of *which* frames of reference they employ. For example, some languages do not employ our apparently fundamental spatial notions of 'left'/'right'/'front'/'back' at all; instead they may, for example, employ a cardinal direction system, specifying locations in terms of 'north'/'south'/'east'/'west' or the like.

There is a second surprising finding: the choice of a frame of reference in linguistic coding (as required by the language) correlates with preferences for the same frame of reference in non-linguistic coding over a whole range of non-verbal tasks. In short, there is a cross-modal tendency for the same frame of reference to be employed in language tasks, recall and recognition memory tasks, inference tasks, imagistic reasoning tasks and even unconscious gesture. This suggests that the underlying representation systems that drive all these capacities and modalities have adopted the same frame of reference.

These findings, described in part 2.0, prompt a series of theoretical ruminations in part 3.0. First, we must ask whether it even makes sense to talk of the 'same' frame of reference across modalities or inner representation systems.[2] Second, we must clarify

---

[1] This paper rests on results of joint research, in particular with Penelope Brown on Tzeltal, but also with many colleagues in the Cognitive Anthropology Research Group who have collaboratively developed the research program outlined here; see also Senft 1994, Wilkins 1993, Pederson 1994, Danziger 1994, Hill 1994. I am also indebted to colleagues in the wider Psycholinguistics Institute, who have through different research programs challenged premature conclusions and emboldened others; see e.g. Bowerman, Levelt, Bierwisch in this volume (the debt to Levelt’s pioneering work on the typology and logic of spatial relations will be particularly evident). In addition, John Lucy , Suzanne Gaskins and Dan Slobin have been important intellectual influences; and Bernadette Schmitt and László Nagy have contributed to experimental design and analysis. The contributions, ideas and criticisms of other participants at the conference at which this paper was given have been woven into the text; my thanks to them and the organizers of the conference. Finally, I received very helpful comments on the manuscript from Sotaro Kita, Lynn Nadel, Mary Peterson and David Wilkins, not all of which I have been able to adequately respond to.

[2] I shall use the term 'modality' in a slightly special, but I think motivated, way: when psychologists talk of 'cross-modal' effects, they have in mind transfer of information across sensory modalities (vision, touch, etc.). Assuming that these sensory input systems are 'modules' in the Fodorean sense, we are then interested in how the output of one module, in some particular inner representation system, is related to the output of some other module, most likely in another inner representation system appropriate to another sensory faculty. Thus cross-modal effects can be assumed to occur through

the notion 'frame of reference' in language, and suggest a slight reformation of the existing distinctions. Then we can, it seems, bring some of the distinctions made in other modalities into line with the distinctions made in the study of language, so that some sense can be made of the idea of 'same frame of reference' across language, non-verbal memory, mental imagery, etc. Finally, we turn to the question: why does the same frame of reference tend to get employed across modalities or at least across distinct inner representation systems? It turns out that information in one frame of reference can not easily be converted into another, distinct frame of reference. This has interesting implications for what is known as Molyneux's Question, the question about how and to what extent there is cross-modal transfer of spatial information.

2.0 Cross-modal transfer of frame of reference: evidence from Tenejapan Tzeltal

To describe where something (let us dub it the Figure) is with respect to something else (let us call it the Ground) we need some way of specifying angles on the horizontal. In English we achieve this either by utilizing features or axes of the Ground (as in “the boy is at the front of the truck”) or by utilizing angles derived from the viewer’s body coordinates (as in “the boy is to the left of the tree”). The first solution I shall call an Intrinsic frame of reference, the second a Relative frame of reference (because the description is relative to the viewpoint - from the other side of the tree the boy will be seen to be to the right of the tree). The notion ‘frame of reference’ will be explicated in section 3.0, but can be thought of as labelling distinct kinds of coordinate system.

At first sight, and indeed on close consideration (see e.g. Clark 1973, Milller & Johnson-Laird 1976), these solutions seem inevitable, the only natural solutions for a bipedal creature with particular bodily asymmetries on our planet. But they are not. Some languages use just the first. Some languages use neither of these solutions, but solve the problem of finding angles on the horizontal plane by utilizing fixed bearings, something like our cardinal directions ‘north’, ‘south’, ‘east’ and ‘west’. Spatial descriptions utilizing such a solution can be said to be in an Absolute frame of reference (since the angles are not relative to a point of view, i.e. are not Relative, and are also independent of properties of the Ground object, i.e. are not Intrinsic). A tentative typology of the three major frames of reference in language, with some indication of the range of subtypes, will be found in section 3.0 below. Here I wish to introduce one such Absolute system, as found in a Mayan language.

Tzeltal is a Mayan language widely spoken in Chiapas, Mexico, but the particular dialect described is spoken by at least 15,000 people in the Indian community of Tenejapa; I will therefore refer to the relevant population as Tenejapans. The results reported here are a part of an ongoing project, conducted with Penelope Brown (Brown & Levinson 1993a, 1993b, Levinson & Brown 1994).

2.1 Tzeltal Absolute linguistic frame of reference

---

communication between central, but still sense-specific, representation systems, not through peripheral representation systems specialized to modular processes. But see 4.0 below.

Tzeltal has an elaborate Intrinsic system (see Brown 1991, Levinson in press), but it is of limited utility for spatial description because it is usually only employed to describe objects in strict contiguity. Thus for objects separated in space, another system of spatial description is required. This is in essence a cardinal direction system, although it has certain peculiarities. First, it is transparently derived from a topographic feature: Tenejapa is a large mountainous tract, with many ridges and cross-cutting valleys, which nevertheless exhibits an overall tendency to fall in altitude towards the north-north-west. Hence 'downhill' has come to mean (approximately) north, and 'uphill' designates south. Second, the coordinate system is deficient, in that the orthogonal 'across' is labelled identically in both directions (east, and west); the particular direction can be specified periphrastically, by referring to landmarks. Third, there are therefore certain ambiguities in the interpretation of the relevant words. Despite this however, the system is a true fixed-bearing system. It applies to objects on the horizontal as well as on slopes. And speakers of the language point to a specific direction for 'down', and they will continue to point to the same compass bearing when transported outside their territory. Figure 1 may help to make the system clear.

((Figure 1 about here))

The three-way semantic distinction between 'up', 'down' and 'across' recurs in a number of distinct lexical systems in the language. Thus there are relevant abstract nominals describing the directions, specialized concrete nominals of different roots for describing e.g. edges along the relevant directions, and motion verbs designating ascending (i.e. going south), descending (going north) and traversing (going east or west). This linguistic ramification, together with its insistent use in spatial description, make the three-way distinction an important feature of language use.

There are many other interesting features of this system (Brown & Levinson 1993a). But the essential points to grasp are the following. First, this is the basic way to describe the relative locations of all objects separated in space on whatever scale. Thus if one wants to pick out one of two cups on a table, one would ask for e.g. the uphill one; if one wants to describe where the boy is hiding behind a tree, one would designate e.g. the north (downhill) side of the tree; if one asks where someone is going the answer may be 'ascending (going south)', and so on. Second, linguistic specifications like our 'to the left', 'to the right', 'in front', 'behind' are not available in the language: thus there is no way to encode English locutions like 'pass the cup to the left', 'the boy is in front of the tree', 'take the first turning right', etc.[3] Third, the use of the system presupposes a good sense of direction; tests of this ability to keep track of directions (in effect to dead-reckon), show that Tenejapans, even without visual access to the environment, do indeed maintain the correct bearings of various locations as they move in the environment.

[3] Although there are phrases designating left-hand and right-hand, these are body-part terms with no spatial uses, while body-part terms for 'face' and 'back' are used for spatial description nearly exclusively for objects in contiguity and then on the basis of an Intrinsic assignment, not a Relative one based on the speaker's viewpoint (see Levinson in press).

In short, this linguistic system does not provide familiar viewer-centered locutions like ‘turn to the left’ or ‘in front of the tree’. All such directions and locations can be adequately coded in terms of antecedently fixed, absolute bearings. Following work on an Australian language (Haviland 1993, Levinson 1992b) where such a linguistic system demonstrably has far-reaching cognitive consequences, a series of experiments were run in Tenejapa to ascertain whether non-linguistic coding might follow the pattern of the linguistic coding of spatial arrays.

2.2 Use of an Absolute frame of reference in non-verbal tasks.

2.2.1 Memory and inference

As part of a larger comparative project, we have together with colleagues devised experimental means for revealing the underlying non-linguistic coding of spatial arrays for memory (see Danziger & Baayen 1994). The aim is to find tasks where subjects’ responses will reveal which frame of reference, Intrinsic, Absolute or Relative, has been employed during the task. Here we concentrate on the Absolute vs. Relative coding of arrays. The simple underlying design behind all the experiments reported here can be illustrated as follows. A subject sees an array on a table, call it Table 1, say, an arrow pointing to his right, or objectively to the north. (See Figure 2). The array is then removed, and after a delay, the subject is rotated 180 degrees to face another table, Table 2. Here there are, say, two arrows, one pointing to his right and one to his left - equally, one to the north, and one to the south. He is then asked to identify the arrow like the one he saw before. If he chooses the one pointing to his right (and incidentally to the south), it is clear that he coded the first arrow in terms of his own bodily coordinates, which have rotated with him. If he chooses the other arrow, pointing north (and to his left), then it is clear that he coded the original array without regard to his bodily coordinates, but with respect to some fixed bearing or environmental feature. Using the same method, we can explore a range of different psychological faculties: recognition memory (as just sketched), recall memory (by e.g. asking him to place an arrow so that it is the same as the one on Table 1) and various kinds of inference (as sketched below).

((Figure 2 about here))

We will describe here just three such experiments in outline form (see Brown & Levinson 1993b for further details and further experiments). They were run on at least 25 Tenejapan subjects (depending on the experiment) of mixed age and sex, and a Dutch comparison group of at least 39 subjects of similar age/sex composition. As far as the distinction between Absolute and Relative linguistic coding goes, Dutch like English relies heavily of course on a ‘right’, ‘left’, ‘front’, ‘back’ system of speaker-centered coordinates for the description of most spatial arrays. So the hypothesis entertained in all the experiments is the following simple ‘Whorfian’ conjecture: the coding of spatial arrays - that is the conceptual representations involved - in a range of non-verbal tasks should employ the same frame of reference that is dominant in the language used in verbal tasks for the same sort of arrays. Since Dutch, like English, provides a dominant Relative frame of reference, we expect Dutch subjects to solve all the non-linguistic tasks utilizing a Relative frame of reference. On the other hand, since Tzeltal offers only an Absolute frame of reference for the relevant arrays, we expect Tenejapan subjects to solve the non-linguistic tasks utilizing an Absolute frame of reference. Clearly it is crucial that the instructions for the experiments, or the wording used in training sessions, do not suggest one or another of the frames of reference: instructions (in Dutch or Tzeltal) were of the kind "Point to the pattern you saw before", "Remake the array just as it was", "Remember just how it is", i.e. as much devoid of spatial information as possible, and as closely matched in content as could be achieved across languages.

Recall Memory[4]

Method:
The design was intended to deflect attention from memorizing direction towards memorizing order of objects in an array, although the prime motive was to tap recall memory for direction. The stimuli consisted of two identical sets of four model animals (pig, cow, horse, sheep) familiar in both cultures. From the set of four, three were aligned in a pre-randomized order, all heading in (a randomly assigned) lateral direction on Table 1. Subjects were trained to memorize the array before it was removed, then after a three-quarters of a minute delay to rebuild it 'exactly as it was', first with correction for misorders on Table 1, then without correction under rotation on Table 2. Five main trials then proceeded, with the stimulus always presented on Table 1, and the response required under rotation, and with delay, on Table 2. Responses were coded as 'Absolute' if the direction of the recalled line of animals preserved the fixed bearings of the stimulus array, and as 'Relative' if the recalled line preserved egocentric 'left' or 'right' direction.

Results:
95% of Dutch subjects were consistent Relative coders on at least four out of five trials, while 75% of Tzeltal subjects were consistent Absolute coders by the same measure. The remainder failed to recall direction so consistently. For purposes of comparison across tasks, the data have been analyzed in the following way. Each subject's performance is assigned an index on a scale from 0-100, where 0 represents a consistent Relative response pattern, 100 a consistent Absolute pattern, and inconsistencies between codings over trials are represented by indices in the interval. The data can then be represented by the graph in Figure 3, where subjects from each population have been grouped by 20-point intervals on the index.

((Figure 3 about here))

As the graph makes clear, the curves for the two populations are approximately mirror-images, except that Tenejapan subjects are less consistent than Dutch ones. This may be due to various factors: the unfamiliarity of the situation and the tasks, the 'school' -like nature of task performed by largely unschooled subjects, or to interference from an egocentric frame of reference that is available but less dominant. Only two Tenejapan subjects were consistent Relative coders (on 4 out of 5 trials) . This pattern is essentially repeated across the experiments.

[4] The design of this experiment was much improved by Bernadette Schmitt.

The result appears to confirm the hypothesis that the frame of reference dominant in the language is the frame of reference most available to solve non-linguistic tasks, like this simple recall task.

Recognition Memory[5]

Method:
Five identical cards were prepared: on each there was a small green circle and a large yellow circle. The trials were conducted as follows. One card was used as a stimulus in a particular orientation; the subject saw this card on Table 1. The other four were arrayed on Table 2 in a number of patterns so that each card was distinct by orientation (see Figure 4). The subject saw the stimulus on Table 1, which was then removed, and after a delay the subject was rotated and led over to Table 2. The subject was asked to identify the card which was the most similar to the stimulus. The eight trials were coded as indicated in Figure 3: if the card which maintained orientation from an egocentric point of view (e.g. 'small circle towards me') was selected, the response was coded as a Relative response, while the card which maintained the fixed bearings of the circles ('small circle north') was coded as an Absolute response. The other two cards served as controls, to indicate a basic comprehension of the task. Training was conducted first on Table 1, where it was made clear that sameness of type rather than token identity was being requested.

((Figure 4 about here))

Results
We find the same basic pattern of results as in the previous task, as shown in Figure 5. Once again the Dutch subjects are consistently Relative coders, while the Tenejapans are less consistent. Nevertheless, of the Tenejapan subjects who performed consistently over 6 or more of 8 trials, over 80% were Absolute coders. The greater inconsistency of Tenejapan subjects may be due to the same factors mentioned above, but there is also here an additional factor because this experiment tested for memory on both the transverse and sagittal (or north-south and east-west) axes. As mentioned above, the linguistic Absolute axes are asymmetric: one axis has distinct labels for the two half-lines north and south, while the other codes both east and west identically ('across'). If there was some effect of this linguistic coding on the conceptual coding for this non-linguistic task, one might expect more errors or inconsistency on the east-west axis. This was indeed the case.

((Figure 5 about here))

Transitive Inference
Levelt 1984 observed that Relative, as opposed to Intrinsic, spatial relations support transitive and converse inferences; Levinson 1992a noted that Absolute spatial relations also support transitive and converse inferences (see also Levelt this volume). This makes it possible to devise a task where, from two spatial arrays or non-verbal

[5] The design of this experiment is by Eric Pederson and Bernadette Schmitt, building on an earlier design described in Levinson 1992b.

‘premises’, a third spatial array, or non-verbal ‘conclusion’ can be drawn by transitive inference utilizing either an Absolute or a Relative frame of reference. The following task was designed by Eric Pederson and Bernadette Schmitt (and piloted in Tamilnadu by Pederson, see his 1994).

The Design

The design is as follows. The subject sees the first non-verbal ‘premise’ on Table 1, e.g. a blue cone A and a yellow cube B arranged in a predetermined . The top diagram in Figure 6 illustrates one such array from the perspective of the viewer. Then the subject is rotated and sees the second ‘premise’, a red cylinder C and the yellow cube B in a predetermined orientation on Table 2 (the array appearing from an egocentric point of view as e.g. in the second diagram in Figure 6). Finally, the subject is rotated again and led back to Table 1, where he is given just the blue cone A and asked to place the red cylinder C in a location consistent with the previous non-verbal ‘premises’. For example, if the subject sees (‘premise 1’) the yellow cube to the right of the blue cone, then (‘premise 2’) the red cylinder to the right of the yellow cube, when given the blue cone, he may be expected to place the red cylinder C to the right of the blue cone A . It should be self-evident from the top two diagrams in Figure 6, representing the arrays seen sequentially, why the third array (labelled the ‘Relative solution’) is one natural non-verbal ‘conclusion’ from the first two visual arrays.

((Figure 6 about here))

However, this result can only be expected if the subject codes the arrays in terms of egocentric or Relative coordinates which rotate with him. If instead the subject utilizes fixed bearings or Absolute coordinates, we can expect a different ‘conclusion’ - in fact the reverse arrangement, with the red cylinder to the left of the blue cone (see the last diagram labelled ‘Absolute solution’ in Figure 6)! To see why this is the case, consider Figure 7, which gives a ‘bird’s eye view’ map of the experimental situation. If the subject does not use bodily coordinates that rotate with him, the blue cone will be e.g. south of the yellow cube on Table 1, and the red cylinder further south of the yellow cube on Table 2, so the conclusion must be that the red cylinder is south of the blue cone. As the diagram makes clear, this amounts to the reverse arrangement from that produced under a coding using Relative coordinates. In this case, and in half the trials, the Absolute inference is somewhat more complex than a simple transitive inference (involving notions of relative distance), but in the other half of the trials the Relative solution was more complex than the Absolute one in just the same way.

((Figure 7 about here))

Method

Three objects distinct in shape and colour were employed. Training was conducted on Table 1, where it was made clear that the positions of each object relative to the other object - rather than exact locations on a particular table - was the relevant thing to remember. When transitive inferences were achieved on Table 1, the subject was introduced to the rotation between the first and second premises; no correction was given unless the placement of the conclusion was on the orthogonal axis to the stimulus arrays. There were then ten trials, randomized across the transverse and sagittal axes (i.e. the arrays were either in a line across or along the line of vision).

Results
The results are given in the graph in Figure 8. Essentially we have the same pattern of results as in the prior memory experiments, with Dutch subjects consistently Relative coders, and Tenejapan subjects strongly tending to Absolute coding, but with more inconsistency. Of the Tenejapans who produced consistent results on at least 7 out of 10 trials, 90% were Absolute coders (just two out of 25 subjects being Relative coders). The reasons for the greater inconsistency of Tenejapan performance are presumably the same as in the previous experiment: unfamiliarity with any such procedure or test situation, and the possible effects of the weak Absolute axis (the east-west axis lacking distinct linguistic labels for the half-lines). Once again, Tenejapans made most errors, or performed most inconsistently, on the east-west axis.

((Figure 8 about here))

Discussion
The results from these three experiments, together with others unreported here (see Brown & Levinson 1993b) , all tend in the same direction: while Dutch subjects utilize a Relative conceptual coding (presumably in terms of notions like left, right, in front, behind) to solve these non-verbal tasks, Tenejapan subjects predominantly use an Absolute coding system. This is of course in line with the coding built into the semantics of spatial description in the two languages. The same pattern holds across different psychological faculties: the ability to recall spatial arrays, to recognize ones one has seen before, and to make inferences from spatial arrays. Further experiments of different kinds, exploring recall over different arrays and inferences of different kinds, all seem to show that this is a robust pattern of results.

The relative inconsistency of Tenejapan performance might simply be due to unfamiliar materials and procedures in this largely illiterate, peasant community. But as suggested above, errors or inconsistencies accumulated on one Absolute axis in particular. However, since the experiments were all run on one set of fixed bearings, the error pattern could have been due equally to a strong vs. weak egocentric axis (and in fact it is known that the left/right axis - here coinciding with the east-west axis - is less robust conceptually than the front/back axis). Therefore half the subjects were recalled and the experiments rerun on the orthogonal Absolute bearings. The results showed unequivocally that errors and inconsistencies do indeed accumulate on the east-west Absolute axis (although there also appears to be some interference from egocentric axes). This is interesting because it shows that Tenejapan subjects are not simply using an ad hoc system of local landmarks, or some fixed-bearing system totally independent of the language; rather, the conceptual primitives used to code the non-verbal arrays seem to inherit the particular properties of the semantics of the relevant linguistic distinctions.

This raises the sceptical thought that perhaps subjects are simply using linguistic mnemonics to solve the non-verbal tasks. However, an effective delay of at least three quarters of a minute between losing sight of the stimulus and responding on Table 2 would have required constant subvocal rehearsal for the mnemonic to remain available in short term memory. Moreover, there is no particular reason why subjects should converge on a linguistic rather than a non-linguistic mnemonic (like crossing the fingers on the relevant hand, or using a kinaesthetic memory of a gesture - which would yield uniform Relative results). But above all, two other experimental results suggest the inadequacy of an account in terms of a conscious strategy of direct linguistic coding.

2.2.2 Visual recall and gesture

The first of these further experiments concerns the recall of complex arrays. Subjects saw an array of between two and five objects on Table 1, and had to rebuild the array under rotation on Table 2. Up to five of these objects had complex asymmetries, e.g. a model of a chair, a truck, a tree, a horse leaning to one side, or a shoe. The majority of Tenejapan subjects rebuild the arrays preserving the Absolute bearings of the axes of the objects. This amounts to mental rotation of the visual array (or of the viewer) on Table 1 so that it is reconstructed on Table 2 as it would look like from the other side. Tenejapans prove to be exceptionally good at this, preserving the metric distances and precise angles between objects. It is far from clear that this could be achieved even in principle by a linguistic coding: the precise angular orientation of each object and the metric distances between objects must surely be coded visually, and must be rebuilt under visual control of the hands. This ability argues for a complex interaction between visual memory and a conceptual coding in terms of fixed bearings: an array that is visually distinct may be conceptually identical, and an array visually identical may be conceptually distinct (unlike with a system of Relative coding, where what is to the left side of the visual field can be described as to the left). Thus being able to “see” that an array is conceptually identical in Absolute terms to another may routinely involve mental rotation of the visual image. That a particular conceptual or linguistic system may exercise and thus enhance abilities of mental rotation has already been demonstrated for ASL by Emmorey (this volume). Tenejapans appear to be able to memorize a visual image of an array tagged, as it were, with the relevant fixed bearings.

There is another line of evidence that suggests that the Tenejapan Absolute coding of spatial arrays is not achieved by conscious, artificial use of linguistic mnemonics. To show this one would wish for some repetitive, unconscious non-verbal spatial behaviour that can be inspected for the underlying frame of reference which drives it. There is indeed just such a form of behaviour, namely unreflective spontaneous gesture accompanying speech. Natural Tenejapan conversation can be inspected to see whether, when places or directions are referred to, gestures preserve the egocentric coordinates appropriate to the protagonist whose actions are being described, or whether the fixed bearings of those locations are preserved in the gestures. Preliminary work by Penelope Brown shows that such fixed bearings are indeed preserved in spontaneous Tenejapan gesture.[6] A pilot experiment seems to confirm this. In the experiment, a subject, facing north, sees a cartoon on a small portable monitor with lateral action from east to west. The subject is then moved to another room where he retells the story as best he can to another native speaker who

[6] This phenomenon was first noticed for an Australian Aboriginal group by Haviland (1993), who subsequently demonstrated the existence of the same phenomenon in Zinacantan, a neighbouring community to Tenejapa.

has not seen the cartoon. In one condition, the subject retells the story facing north; in another condition the subject retells the story facing south. Preliminary results show that at least some subjects under rotation systematically preserve the fixed bearing of the observed action (from east to west) in their gestures, rather than the direction coded in terms of left or right. (Incidentally, the reverse finding has been established for American English by McCullough, 1993.). Since subjects had no idea that the experimenter was interested in gesture, we can be sure that the gestures record unreflective conceptualization of the directions. Although the gestures of course accompany speech, gestures preserving the fixed bearings of the stimulus often occur without explicit mention of the cardinal directions, suggesting that the gestures reflect an underlying spatial model, at least partially independent of language.

Conclusion from the Tenejapan studies

Putting all these results together we are led to the conclusion that the frame of reference dominant in the language, Relative or Absolute, comes to bias the choice of frame of reference in various kinds of non-linguistic conceptual representations. This correlation holds across a number of ‘modalities’ or distinct mental representations: over codings for recall and recognition memory, over representations for spatial inference, over recall apparently involving manipulations of visual images, and over whatever kind of kinaesthetic representation system drives gesture. These findings look robust and general: similar observations have previously been made for an Aboriginal Australian community which uses Absolute linguistic spatial description (Haviland 1993, Levinson 1992b), and a cross-cultural survey over a dozen non-Western communities shows a strong correlation of the dominant frame of reference in the linguistic system and frames of reference utilized in non-linguistic tasks (see Baayen & Danziger 1994).

3.0 Frames of Reference across modalities

So far, we have acquired some new facts: (i) not all languages use the same predominant frame of reference, (ii) there is a tendency for the frame of reference predominant in the language to remain the predominant frame of reference across modalities, as displayed by its use in non-verbal tasks of various kinds, unconscious gesture, etc. The results seem firm; they appear to be replicable across speech communities.

But the more one thinks about the implications of these results, the more peculiar they seem to be. First, the trend of current theory hardly prepares us for such 'Whorfian' results: the general assumption is rather of a universal set of semantic primes (conceptual primitives involved in language), and the identity or homomorphism of universal conceptual structure and semantic structure. Secondly, ideas about modularity of mind make it seem unlikely that such cross-modal effects could occur. Thirdly, the very idea of the same frame of reference across different modalities, or different internal representation systems specialized to different sensory modalities, may seem incoherent.

In order to make sense of the results, I shall in this section attempt to show that the notion 'same frame of reference across modalities' is, after all, perfectly coherent, and indeed already adumbrated across the disciplines that study the various modalities. This requires a lightning review of the notion 'frame of reference' across the relevant disciplines (section 3.1 and 3.2); it also requires a reformation of the linguistic distinctions normally made (section 3.3). With that under our belts, we can then face up to the peculiarity, from the point of view of ideas about the modularity of mind, of this cross-modal adoption of the same frame of reference (section 4.0). Here some intrinsic properties of the different frames of reference may offer the decisive clue: if there is to be any cross-modal transfer of spatial information, we may have no choice but to fixate predominantly on just one frame of reference.

3.1 The notion 'Spatial frames of reference'

The notion of 'frames of reference' is crucial to the study of spatial cognition across all the modalities and all the disciplines that study them. The idea is as old as the hills: medieval theories of space, for example, were deeply preoccupied by the puzzle raised by Aristotle, the case of the boat moored in the river. If we think about the location of objects as places that they occupy, and places as containing the objects, then the puzzle is that if we adopt the river as frame of reference the boat is moving, but if we adopt the bank as frame, then it is stationary (see Sorabji 1988:187ff for this problem which dominated medieval discussions of space).

But the phrase 'frame of reference', and its modern interpretation originates, like so much else worthwhile, from Gestalt theories of perception in the 1920s. How, for example, do we account for illusions of motion, as when the moon skims across the clouds, except by invoking a notion of a constant perceptual window against which motion (or the perceived vertical, etc.) is to be judged? The Gestalt notion can be summarized as "a unit or organization of units that collectively serve to identify **a coordinate system** with respect to which certain properties of objects, including the phenomenal self, are gauged" (Rock 1992:404, my emphasis).[7]

In what follows, I will myself emphasize that distinctions between frames of reference are essentially distinctions between underlying coordinate systems and not for example between the objects that may invoke them. Not all will agree.[8] In a recent review, philosophers Brewer & Pears (1993) ranging over the philosophical and psychological literature, conclude that frames of reference come down to the selection of reference objects: take the glasses on my nose - when I go from one room to another, do they change their location or not? It depends on the 'frame of reference' -

[7] Rock (1992) is here commenting on Asch & Witkin (1948), which built directly on the Gestalt notions. See also Rock (1990).

[8] One kind of disagreement is voiced by Paillard 1991:471: "Spatial frameworks are incorporated in our perceptual and motor experiences. They are not however to be confused with the *system of coordinates* which abstractly represent them." But this is terminological; for our purposes we wish precisely to abstract out the properties of frames of reference, so we can consider how they apply across different perceptual or conceptual systems.

nose or room.[9] This emphasis on the Ground, or relatum or reference object[10] severely underplays the importance of co-ordinate systems in distinguishing frames of reference, as I shall show below.[11] Humans use multiple frames of reference: I can happily say of the same assemblage (ego looking at car from side, car's front to ego's left): "the ball is in front of the car" and "the ball is to the left of the car", without thinking that the ball has changed its place. In fact, much of the psychological literature is concerned with ambiguities of this kind. I will therefore insist on the emphasis on coordinate systems rather than on the objects or 'units' on which such coordinates may have their origin.

3.2 'Frames of reference' across modalities and the disciplines that study them

If we are to make sense of the notion 'same frame of reference' across different modalities, or inner representation systems, it will be essential to see how the various distinctions between such frames that have been proposed in different disciplines, can be ultimately brought into line. This is no trivial undertaking, because there are a host of such distinctions, and each of them has been variously construed, both within and across the many disciplines (such as philosophy, the brain sciences, psychology and linguistics) that explicitly employ the notion 'frames of reference'. A serious review of these different conceptions would take us very far afield. On the other hand, some sketch is essential, and I will briefly survey the various distinctions in Table 1, with some different construals distinguished by the letters (a), (b), (c).[12]

Table 1: Spatial Frames Of Reference: Some distinctions in the literature

'relative' vs. 'absolute':
(philosophy, brain sciences, linguistics)
(a) space as relations between objects vs. abstract void
(b) egocentric vs. allocentric
(c) directions: relations twixt objects vs. fixed bearings

'egocentric' vs. 'allocentric'
(developmental and behavioural psychology, brain sciences)
(a) body-centred vs. environment-centred

[9] 'When places are individuated by their spatial relation to certain objects, a crucial part of what we need to know is what those objects are. As the term 'frame of reference' is commonly used, these objects would be said to provide the frame of reference' (Brewer & Pears, 1993:25).

[10] I shall use the opposition 'Figure' vs. 'Ground' for the object to be located vs. the object with respect to which it is to be located, respectively, after Talmy 1983. This opposition is identical to that between Theme vs. Relatum, Referent vs. Relatum, Trajector vs. Landmark, and various other terminologies.

[11] Brewer and Pears (1993) consider the role of co-ordinate systems but what they have to say only increases our puzzlement: "Two events are represented as being in the same spatial position if and only if they are assigned the same co-ordinates. Specifying a frame of reference would have to do with specifying how co-ordinates are to be assigned to events in the world on the basis of their spatial relations to certain objects. These objects provide the frame of reference" (Brewer & Pears 1993:26). This fails to recognize that 2 distinct systems of co-ordinates over the same objects can describe the same place.

[12] There are many good sketches of parts of this intellectual terrain (see e.g. Miller & Johnson-Laird 1976, Jammer 1954, O'Keefe & Nadel 1978), but none of it all.

(Note many egocentres: retina, shoulder, etc.)
(b) subjective (subject-centred) vs. objective

'viewer-centered' vs. 'object-centered' or
'2.5D sketch' vs. '3D models'
(vision theory, imagery debate in psychology)

'orientation-bound vs. orientation-free'
(visual perception, imagery debate in psychology)

'deictic' vs. 'intrinsic'
(linguistics)
(a) speaker-centric vs. non-speaker-centric
(b) centered on speaker or addressee vs. thing
(c) ternary vs. binary spatial relations

'viewer-centred' vs. 'object-centred' vs. 'environment-centered'
= 'gaze-tour' vs. 'body-tour' perspectives
=?'survey perspective' vs. 'route-perspective'
(psycholinguistics)

First, then, **relative vs. absolute** space. Newton's distinction between absolute and relative space has played an important role in ideas about frames of reference, in part through the celebrated correspondence between his champion Clarke and Leibniz, who held a strictly relative view. [13] For Newton, absolute space is an abstract infinite immovable three-dimensional box with origin at the centre of the universe, while relative space is conceived of as specified by relations between objects: psychologically, he claimed, we are inclined to relative notions: "Relative space is some moveable dimension or measure of the absolute spaces, which our senses determine by its position to bodies ... and so instead of absolute places and motions, we use relative ones" (quoted in Jammer 1954:97-8). Despite fundamental differences in philosophical position, most succeeding thought in philosophy and psychology has assumed the psychological primacy of relative space - space anchored to the places occupied by physical objects and their relations to one another - in our mental life. A notable exception is Kant, who came to believe that notions of absolute space are a fundamental intuition, although grounded in our experience through the use of our body to define the egocentric co-ordinates through which we deal with it (Kant 1768; see also van Cleve & Frederick 1991). O'Keefe and Nadel (1978; see also O'Keefe 1993 and this volume) have tried to preserve this Kantian view as essential to the understanding of the neural implementation of our spatial capacities, but by and large psychologists have considered notions of 'absolute' space irrelevant to theories of the naive spatial reasoning underlying language (see Clark 1973, Miller & Johnson-Laird 1976:380). (Absolute notions of space may though be related to cognitive maps of the environment - discussed under the rubric of 'allocentric' frames of reference below.) The distinction between relative and absolute space early on acquired certain additional associations: for example, relative space became associated with egocentric

---

[13] Some notion of absolute space was already presupposed by Descartes' introduction of coordinate systems , as Einstein (1954:xiv) pointed out.

co-ordinate systems, and absolute space with non-egocentric ones (despite Kant 1768),[14] so that this distinction is often confused with the egocentric vs. allocentric distinction (discussed below). Another interpretation of the 'relative' vs. 'absolute' distinction, in relating relativistic space to egocentric space, goes on to emphasize the difference in the way coordinate systems are constructed in absolute vs. relative spatial conceptions: "Ordinary languages are designed to deal with relativistic space; with space relative to the objects that occupy it. Relativistic space provides three orthogonal coordinates, just as Newtonian space does, but *no fixed units of angle or distance are involved, nor is there any need for coordinates to extend without limit in any direction*" (Miller & Johnson-Laird 1976:380, my emphasis). Thus a system of fixed bearings, or cardinal directions, is opposed to the relativistic 'space concept', whether egocentric or object-centered, which Miller & Johnson-Laird (and many other authors, like Clark 1973, Herskovits 1986, Svorou 1994:213) have assumed to constitute the conceptual core of human spatial thinking (Miller & Johnson-Laird 1976:395). But since, as we have seen, some languages use as a conceptual basis coordinate systems with fixed angles (and coordinates of indefinite extent), we need to recognise that some languages utilize what may be appropriately called Absolute coordinate systems. Hence I have opposed Absolute vs. Relative frames of reference in language (see the next section).

Let us turn to the next distinction in Table 1, viz. **egocentric vs. allocentric**. The distinction is of course between coordinate systems with origins within the subjective body frame of the organism, versus coordinate systems centered elsewhere (often unspecified). The distinction is often invoked in the brain sciences, where there is a large literature concerning 'frames of reference' (see e.g. the compendium in Paillard 1991). This emphasizes the plethora of different egocentric co-ordinate systems required to drive all the different motor systems from saccades to arm-movements (see e.g. Stein 1992), or the control of the head as a platform for our inertial guidance and visual systems (again see papers in Paillard 1991). In addition, there is a general acceptance (op. cit. p. 471) of the need for a distinction (following Tolman 1948, and O'Keefe & Nadel 1978) between egocentric vs. allocentric systems. O'Keefe & Nadel's demonstration that *something* like Tolman's mental maps are to be found in the hippocampal cells is well known.[15] O'Keefe's recent work is an attempt to relate a particular mapping system to the neuronal structures and processes (O'Keefe 1993). The claim is that the rat can use egocentric measurements of distance and direction towards a set of landmarks to compute a non-egocentric abstract central origo (the 'centroid') and a fixed angle or 'slope'. Then it can keep track of its position in terms of distance from centroid and direction from slope. This is a 'mental map' constructed through the rat's exploration of the environment, which gives it fixed bearings (the slope) *but just for this environment*. Whether this strictly meets the criteria for an

[14] This was in part due to the British empiricists like Berkeley whose solipsism made egocentric relative space the basis for all our spatial ideas. See O'Keefe & Nadel 1978:14ff.

[15] Much behavioural experimentation on e.g. rats in mazes, has led to classifications of behaviour parallel to the notions of frame of reference: O'Keefe & Nadel's 1978 classification, for example, is in terms of body-position responses (cf. egocentric frames of reference), cue responses (a kind of allocentric response to an environmental gradient) and place responses (involving allocentric mental maps). Work on infant behaviour similarly relates behavioural response types to frames of reference, usually egocentric vs. allocentric (or 'geographic' - see Pick 1988:147ff).

objective, 'absolute' allocentric system has been questioned (Campbell 1993:76-82).[16] We certainly need to be able to distinguish mental maps of different sorts: egocentric 'strip-maps' (Tolman 1948), allocentric landmark-based maps with relative angles and distances between landmarks (more Leibnizian), and allocentric maps based on fixed-bearings (more Newtonian).[17] But in any case, this is the sort of thing neurophysiologists have in mind when they oppose 'egocentric' and 'allocentric' frames of reference.[18]

Another area of work where the opposition has been used is in the study of human conceptual development. For example, Acredolo (1988) shows that, as Piaget argued, infants have indeed only egocentric frames of reference in which to record spatial memories; but contrary to Piaget, this phase lasts only for perhaps the first six months. Thereafter, they acquire the ability to compensate for their own rotation, so that by 16 months they can identify say a window in one wall as the relevant stimulus even when entering the room (with two identical windows) from the other side. This can be thought of as the acquisition of a non-egocentric, 'absolute' or 'geographic' orientation or frame of reference.[19] Pick (1993:35) points out, however, that such apparently allocentric behaviour can be mimicked by egocentric mental operations, and indeed this is suggested by Acredolo's (1988:165) observation that children learn to do such tasks via adopting the visual strategy 'if you want to find it, keep your eyes on it (as you move)'.

These lines of work identify the egocentric vs. allocentric distinction with the opposition between body-centred vs. environment-centred frames of reference. But as philosophers point out (see e.g. Campbell 1993), ego is not just any old body, and there is indeed another way to construe the distinction as one between subjective vs. objective frames of reference. The egocentric frame of reference would then bind together various body-centered coordinate systems with an agentive subjective being, complete with body-schema, distinct zones of spatial interaction (reach, peripheral vs. central vision, etc.). For example, phenomena like 'phantom limbs' or proprioceptive illusions argue for the essentially subjective nature of egocentric coordinate systems.

---

[16] See also Brewer & Pears 1993:29 who argue that allocentric behaviour can always be mimicked through egocentric computations: "Perhaps language .... provides the only conclusive macroscopic evidence for genuine allocentricity".

[17] These distinctions seem rarely properly made in the literature on mental maps in humans. Students of animal behaviour, though, have noted that maps consisting of relative angles and distances between landmarks have quite different computational properties to maps with fixed bearings: in the former, but not the latter, each time landmarks are added to the map, the database increases exponentially (see e.g. McNaughton, Chen & Markus 1990). Despite that, most rat studies fail to distinguish these two kinds of allocentricity, relative and absolute.

[18] Paillard (1991:471ff) has a broader notion of 'frames of reference' than most brain scientists (and closer to psychological ideas): he proposes that there are four such frames subserving visually guided action, all organized around the geocentric vertical: (i) a body frame, presuming upright posture for action; (ii) an object frame, presumably similar to Marr's object-centred system, (iii) a world frame, a Euclidean space inclusive of both body and object, and (iv) a retinal frame, feeding the object- and world-frames. He even provides a rough neural 'wiring diagram' (p. 473).

[19] The age at which this switch to the non-egocentric takes place seems highly task dependent (see Acredolo 1988 who gives 16 months as an end-point; see also Pick 1993, for a route-finding task, where the process has hardly begun by 16 months).

The next distinction on our list, **viewer-centered vs. object-centered**, comes from the theory of vision, as reconstructed by Marr (1982). In Marr's well-known conceptualization, a theory of vision should take us from retinal image to visual object-recognition, and that, he claimed, entails a transfer from a viewer-centered frame of reference, with incremental processing up to what he called the 2.5D sketch, to an object-centered frame of reference, a true 3D model or structural description.[20] Since we can recognize an object even when foreshortened or viewed in differing lighting conditions, etc., we must extract some abstract representation of it in terms of its volumetric properties to match this token to our mental inventory of such types. Although recent developments have challenged the role of the 3D model within a modular theory of vision,[21] there can be little doubt that at some conceptual level such an object-centered frame of reference exists. This is further demonstrated by work on visual imagery, which seems to show that presented with a viewer-centered perspective view of a novel object, we can mentally rotate it to obtain different perspectival 'views' of it, e.g. to compare it to a prototype (Shepard & Metzler 1971; Kosslyn 1980 ; Tye 1991:83-6). Thus at some level, the visual or ancillary systems seem to employ two distinct reference frames, viewer-centered and object-centered.

This distinction between viewer-centered and object-centered frames of reference relates rather clearly to the linguistic distinction between deictic and intrinsic perspectives discussed below: the deictic perspective is viewer-centered, while the intrinsic perspective seems to use (at least partially) the same axial extraction that would be needed to compute the volumetric properties of objects for visual recognition (see Landau & Jackendoff, 1993 and both authors this volume; also Levinson in press). This parallel will be further reinforced by the reformation of the linguistic distinctions suggested in the section below.

This brings us to the distinction between **orientation-bound vs. orientation-free** frames of reference.[22] The visual imagery and mental rotation literature might be thought to have little to say about frames of reference. After all, visual imagery would seem to be necessarily at most 2.5D and thus necessarily in a viewer-centred frame of reference (even if mental rotations indicate access to a 3D description). But recently there have been attempts to understand the relation between two kinds of shape recognition: the process where shapes can be recognized without regard to orientation (thus with no response-curve latency related to degrees of orientation from a familiar related stimulus), and another process where shapes are recognized by apparent analog rotation to the familiar related stimulus. The Shepard & Metzler paradigm suggested that only where handedness information is present (as where enantiomorphs have to be discriminated) would mental rotation be involved, which implicitly amounts to some distinction between object-centered and viewer-centered

---

[20] This leap from a perspective image, or worse a silhouette, is possible (he argued) only by assuming that objects can be analyzed into geometrical volumes of a specific kind (generalized cones); hence 3D models must be of this kind, where principal axes are identified.

[21] Others have suggested that what we store is a 2.5D image coupled with the ability to mentally rotate it (Tarr & Pinker 1989), thus giving our apparent ability to rotate mental images (Shepard & Metzler 1971) some evolutionary raison d'etre. Yet others suggest that object-recognition is achieved via a set of 2.5D images from different orientations (Bülthoff 1991), while some (Rock, Wheeler & Tudor 1989) suggest we have none of these powers.

[22] I am grateful to Eve Danziger for putting me in touch with this work; see Danziger 1994 for possible connections to linguistic distinctions.

frames of reference: discrimination of enantiomorphs depends on an orientation-bound perspective, while the recognition of simpler shapes may be orientation-free.[23] But some recent controversies seem to show that things are not as simple as this (Tarr & Pinker 1989, Cohen & Kubovy 1993). Just and Carpenter (1985) argue that rotation tasks in fact can be solved using four different strategies, some orientation-bound and some orientation-free.[24] Similarly, Takano (1989) suggests that there are four types of spatial information involved, classifiable by crossing *elementary* (simple) vs. *conjunctive* (partitionable) forms with the distinction between *orientation-bound* and *orientation-free*. He insists that only orientation-bound forms should require mental rotation for recognition. However, Cohen & Kubovy (1993) claim that all this makes the wrong predictions since handedness-identification can be achieved without the mental-rotation latency curves in special cases. In fact, I believe that despite these recent controversies, the original assumption - that only objects lacking handedness can be recognized without mental rotation - must be basically correct for logical reasons that have been clear for centuries.[25] In any case, it is clear from this literature that the study of visual recognition and mental rotation utilizes distinctions in frames of reference that can be put into correspondence with those that emerge from e.g. the study of language: Absolute and Relative frames of reference in language (to be firmed up below) are both orientation-bound, while the Intrinsic frame is orientation-free (Danziger 1994).

Linguists have long distinguished **'deictic' vs 'intrinsic'** frames of reference, because of the rather obvious ambiguities of a sentence like "the boy is in front of the house" (see e.g. Leech 1969:168, Fillmore 1971, Clark 1973, etc.). It has also been known for a while that linguistic acquisition of these two readings of terms like "in front", "behind", "to the side of" is in the reverse direction from the developmental sequence 'egocentric' to 'allocentric' (Pick 1993): 'intrinsic' notions come resolutely earlier than deictic ones (Johnston & Slobin 1978). Sometimes a third term **'extrinsic'** is opposed, to denote e.g. the contribution of gravity to the interpretation of words like 'above' or 'on'. But unfortunately the term 'deictic' breeds confusions. In fact there have been at least three distinct interpretations of the 'deictic' vs. 'intrinsic' contrast, as listed in Table 1: (a) speaker-centric vs. non-speaker-centric (Levelt 1989), (b) centered on any of the speech participants vs. not so centered (Levinson 1983), (c) ternary vs. binary spatial relations (implicit in Levelt 1984 and this volume, to be adopted here). These issues will be taken up in the section below, where we will turn to ask what distinctions in frames of reference are grammaticalized or lexicalized in different languages.

---

[23] As Kant 1768 made clear, objects differing in handedness (enantiomorphs or incongruent counterparts in his terminology), can not be distinguished in an object-centered (or intrinsic) frame of reference, but only in an external coordinate system. See Van Cleve & Frederick 1991, and, for the relevance to Tzeltal, Levinson & Brown 1994.

[24] e.g. the Cube Comparisons Test can be solved by (a) rotation using viewer-centered coordinates, (b) rotation around an object-centered axis imaged with viewer-centered coordinates, (c) rotation of the perspective point around the object, (d) purely object-centered comparisons.

[25] Thus Cohen & Kubovy display deep confusion about frames of reference: they suggest (1993:379) that one can have orientation-free representations of handedness information in an orientation-free frame of reference by utilizing the notion *clockwise*. But as Kant (1768) showed, and generations of philosophers since have agreed (see van Cleve & Frederick 1991), the notion 'clockwise' presupposes an external orientation.

Let us turn now to the various distinctions suggested in the psychology of language. Miller & Johnson-Laird (1976), drawing on earlier linguistic work, explored the opposition between deictic and intrinsic interpretations of such utterances as “the cat is in front of the truck”; and the logical properties of these two frames of reference, and their interaction, have been further clarified by Levelt (1984, 1989 and this volume). Carlson-Radvansky & Irwin (1993:224) summarize the general assumption in psycholinguistics as follows:

> “Three distinct classes of reference frames exist for representing the spatial relationships among objects in the world ....: **viewer-centered frames, object-centered frames, and environment centered** frames of reference. In a viewer-centered frame, objects are represented in a retinocentric, head-centric or body-centric coordinate system based on the perceiver's perspective of the world. In an object-centered frame, objects are coded with respect to their intrinsic axes. In an environment-centered frame, objects are represented with respect to salient features of the environment, such as gravity or prominent visual landmarks. In order to talk about space, vertical and horizontal coordinate axes must be oriented with respect to one of these reference frames so that linguistic spatial terms such as “above” and “to the left of” can be assigned (Miller & Johnson-Laird 1976).”

Notice that on this formulation frames of reference inhere in spatial perception and cognition rather than in language: “above” may simply be semantically general over the different frames of reference, not ambiguous (Carlson-Radvansky & Irwin (1993:242).[26]Thus the corresponding threeway distinction deictic, intrinsic and extrinsic are merely alternative labels for the linguistic interpretations corresponding, respectively, to viewer-centered, object-centered and environment-centered frames of reference.

There are other oppositions that psycholinguists employ, although in most cases they map onto the same triadic distinction. One particular set of distinctions, between different kinds of survey or route description, is worth unravelling because it has caused confusion. Levelt (1989:139ff) points out that when a subject describes a complex visual pattern the linearization of speech requires that we ‘chunk’ the pattern into units that can be described in a linear sequence. Typically, we seem to represent 2D or 3D configurations through a small window, as it were, traversing the array: i.e. the description of complex static arrays is converted into a description of motion through units or ‘chunks’ of the array. Levelt has examined the description of 2D arrays, and found two strategies (this volume): a **gaze tour** perspective - effectively the adoption of a fixed deictic or viewer-centered perspective, and a body or **'driving' tour** - effectively an intrinsic perspective, where a pathway is found through the array, and the direction of the path used to assign ‘front’, ‘left’, etc., from any one point (or location of the window in describing time). Since both perspectives can be thought of as egocentric, Tversky (1991, see also Taylor & Tversky in press, and Tversky, this volume) opts to call Levelt’s intrinsic perspective ‘**deictic** frame of reference' or '**route** description' and his deictic perspective she labels ‘**survey** perspective’.[27] Thus Tversky’s ‘deictic’ is Levelt’s ‘intrinsic’ or non-deictic perspective! This confusion is, I believe, not merely terminological, but results from

---

[26] This view would seem to be subtly different from Levelt’s (1989): see below.

[27] The equation is hers; actually, her survey perspective in some cases (e.g. outside the context of maps) may also relate to a more abstract 'absolute' spatial framework where both viewer and landmarks are embedded in a larger frame of reference.

the failure in the literature to distinguish coordinate systems from their origins or centers (see next section).

There is a final issue of some importance. In psycholinguistic discussions about frames of reference, there seems to be some unclarity, or sometimes overt disagreement, about *at which level* - perceptual, conceptual or linguistic - such frames of reference apply. Thus Carlson-Radvansky & Irwin (1993, quoted above) make the assumption that a frame of reference must be adopted within some spatial representation system, as a precondition for co-ordinating perception and language, whereas Levelt (1989, but see this volume) has argued that a frame of reference is freely chosen in the very process of mapping from perception or spatial representation to language (see also Logan, this volume). On the latter conception, frames of reference in language are peculiar to the nature of the linear, propositional representation system that underlies linguistic semantics: they are different ways of conceiving the same percept in order to talk about it.[28]

The view that frames of reference in linguistic descriptions are adopted in the mapping from spatial representation or perception to language seems to suggest that the perceptions or spatial representations themselves are frame-of-reference-free. But this of course is not the case: there has to be some coordinate system involved in any spatial representation of any intricacy, whether at a peripheral, or sensory, level or at a central, or conceptual, level. What Levelt's results (this volume) or Friederici & Levelt (1990) seem to establish, is that frames of reference at the perceptual or spatial conceptual level do not necessarily determine frames of reference at the linguistic level. This is exactly what one might expect: language is flexible and it is an instrument of communication - thus it naturally allows us e.g. to take the other guy's perspective. Further, the ability to cast a description in one frame or another implies an underlying conceptual ability to handle multiple frames, and within strict limits (see below) to convert between them. In any case, we need to distinguish in discussions of frames of reference between at least three levels: perceptual, conceptual and linguistic, and we need to consider the possibility that we may utilize distinct frames of reference at each level (but see section 4.0 below).

There is much further pertinent literature in all the branches of psychology and brain science, but we must leave it here. It should already be clear that there are many confusingly different classifications, and different construals of the same terms, not to mention many unclarities and many deep confusions in all of this. However, despite this forest of distinctions with obscuring undergrowth, there are some obvious common bases to the distinctions we have reviewed. It is clear for example, that on the appropriate construals, 'egocentric' corresponds to 'viewer-centered' and '2.5D sketch and 'deictic' frames, while 'intrinsic' maps onto 'object-centered' or '3D model' frames of reference, while 'absolute' is related to 'environment-centered', and so on. We should seize these commonalities, especially as in this paper we are concerned with making sense of the notion of the 'same frame of reference' across modalities and

---

[28]The conceptual system is abstract over different perceptual clues, as shown by the fact that astronauts can happily talk about "above and to the left", etc., where one perceptual clue for the vertical (namely gravity) is missing (Friederici & Levelt 1990). Levelt (1989:154-5) concludes that the spatial representation itself does not determine the linguistic description: "there is ...substantial freedom in putting the perceived structure, which is spatially represented, into one or another propositional format".

representational systems. However, before proposing an alignment of these distinctions across the board, it is essential to deal with linguistic frames of reference, which present a troubling flexibility which have led to various confusions.

3.3 Linguistic Frames of Reference in Cross-Linguistic Perspective

Cursory inspection of the linguistic literature will give the impression that the linguists have their house in order. They talk happily of topological vs. projective spatial relators (e.g. pronouns like "in" vs. "behind"), deictic vs. intrinsic usages of projective prepositions, and so on (see e.g. Bierwisch 1967, Lyons 1977, Herskovits 1986, Vandeloise 1991, and psycholinguists Clark 1973, Miller & Johnson-Laird 1976). But the truth is less comforting. The analysis of spatial terms in familiar European languages remains deeply confused,[29] and those in other languages almost entirely unexplored. Thus the various alleged universals should be taken with a great pinch of salt (in fact many of them can be directly jettisoned). One major upset is the recent finding that many languages use an 'absolute' frame of reference (as illustrated in the case of Tzeltal) where European languages would use a 'relative' or viewpoint-centred one (see e.g. Levinson 1992a,b, and Haviland 1993). Another is that some languages, like many Australian ones, use such frames of reference to replace so-called 'topological' notions like 'in', 'on' or 'under'. A third is that familiar spatial notions like 'left' and 'right' and even sometimes 'front' and 'back' are missing from many, perhaps a third of all languages. Confident predictions and assumptions can be found in the literature that no such languages would occur (see e.g. Clark 1973, Miller & Johnson-Laird 1976, Lyons 1977:690).

These developments call for some preliminary typology of the frames of reference that are systematically distinguished in the grammar or lexicon of different languages (with the caveat that we still know little about only a few of them). In particular, we will focus on what we seem to need in the way of co-ordinate systems and associated reference points to set up a cross-linguistic typology of the relevant frames of reference.

In what follows I shall confine myself to linguistic descriptions of static arrays, and I will exclude the so-called 'topological' notions, for which a new partial typology concerning the coding of concepts related to 'in' and 'on' is available (Bowerman & Pederson in prep.).[30] Moreover, I will focus on distinctions on the horizontal plane.

[29] For example, there is no convincing explanation of the English deictic use of 'front', 'back', 'left', 'right': we talk of the cat in front of the tree, as if the tree was an interlocutor facing us, but when we say the cat is to the left of the tree we do not (as e.g. in Tamil) mean the cat is to the tree's left, therefore to our right. The reason is that the facts have always been underdescribed, the requisite coordinate systems not being properly spelled out even in the most recent works.

[30] The so-called 'topological' prepositions or relators have a complex relation to frames of reference. First, note that frames of reference are here defined in terms of coordinate systems, and many 'topological' relators express no angular or coordinate information, e.g. *at* or *near*. However, others do involve the vertical absolute dimension and often intrinsic features, or axial properties, of landmark objects. Thus proper analysis of the 'topological' notions involves partitioning features of them between non-coordinate spatial information, and features of information distributed between the frames of reference mentioned below: thus English *in* (in uses like 'the money in the piggy-bank') is an intrinsic notion based on properties of the Ground object, *under* (in 'the dust under the rug') compounds intrinsic (under-surface, bottom) and absolute (vertical) information, and so forth.

This is not whimsy: the perceptual cues for the vertical may not always coincide, but they overwhelmingly converge, giving us a good universal solution to one axis. But the two horizontal co-ordinates are up for grabs: there simply is no corresponding force like gravity on the horizontal.[31] Consequently there is no simple solution to the description of horizontal spatial patterns, and languages diverge widely in their solutions to this basic problem: how to specify angles or directions on the horizontal.

Essentially, three main frames of reference emerge from this new data as solutions to the problem of description of horizontal spatial oppositions. They are appropriately named 'intrinsic', 'relative' and 'absolute', even though these terms may have a somewhat different interpretation from some of the construals reviewed in the section above. Indeed the linguistic frames of reference potentially *cross-cut* many of the distinctions in the philosophical, neurophysiological, linguistic and psychological literatures, for one very good reason. The reason is that linguistic frames of reference can not be defined by reference to the nature of the origin of the coordinate system (in contrast to e.g. 'egocentric' vs. 'allocentric'). It will follow that the traditional distinction 'deictic' vs. 'intrinsic' collapses - these are not opposed terms. All this requires some explanation.

We may start by noting the difficulties we get into by trying to make the distinction between 'intrinsic' and 'deictic'. Levelt (1989:48-55) organizes and summarizes the standard assumptions in a useful way that illustrates the problem: we can cross-classify linguistic uses according to (a) whether they presume that the coordinates are centered on the speaker (‘Deictic’) or not (‘Intrinsic’), (b) whether the **relatum** or Ground is the speaker or not. Suppose then we call the usage 'Deictic' just in case the coordinates are centered on the speaker, 'Intrinsic' otherwise. This yields, for example, the following classification of examples:

(1) "The ball is in front of me"
Coordinates: **Deictic (i.e. origin on speaker)**
Relatum: Speaker

(2) “The ball is in front of the tree"
Coordinates: **Deictic (i.e. origin on speaker)**
Relatum: Tree

(3) "The ball is in front of the chair (at the chair's front)"
Coordinates: **Intrinsic (i.e. origin not on speaker)**
Relatum: Chair

Clearly it is the locus of the origin of the coordinates that is relevant to the traditional opposition 'intrinsic' vs. 'deictic', otherwise we would group (2) and (3) as both sharing a non-deictic relatum. The problem comes when we pursue this classification further:

---

[31] Except in some places: thus in the Torres Straits, where the trade winds roar through Westward, spatial descriptions can be in terms of 'leeward' and 'windward'. Or where the earth drops away in one direction, as on the edges of mountain ranges, gravity can be naturally imported into the horizontal plane.

(4) "The ball is in front of you"
Coordinates: **Intrinsic (origin on addressee, not speaker)**
Relatum: Addressee

(5) "The ball is to the right of the lamp, from your point of view"
Coordinates: **Intrinsic (origin on addressee)**
Relatum: Lamp

Here the distinction between 'intrinsic' vs. 'deictic' is self-evidently not the right classification, as far as frames of reference are concerned. Clearly, (1) and (4) belong together: the interpretation of the expressions is the same, with the same coordinate systems, there are just different origins, speaker and addressee respectively (moreover, in a normal construal of 'deictic', inclusive of first and second persons, both are 'deictic' origins). Similarly, in another grouping, (2) and (5) should be classed together: they have the same conceptual structure, with a viewpoint (acting as the origin of the coordinate system), a relatum distinct from the viewpoint, and a referent - again the origin alternates over speaker or addressee.

We might therefore be tempted to just alter the designations, and label (1), (2), (4) and (5) all 'deictic' as opposed to (3) 'intrinsic'. But this would be a further confusion. First, it would conflate the distinct conceptual structures of our groupings (1) and (4) vs. (2) and (5). Secondly, the conceptual structure of the coordinate systems in (1) and (4) is in fact shared with (3). How? Consider: "the ball is in front of the chair" presumes (on the relevant reading) an intrinsic front, and uses that facet to define a search domain for the ball; but just the same holds for "the ball is in front of me/you".[32] Thus the logical structure of (1), (3) and (4) is the same: the notion "in front of" is here a **binary** spatial relation, with arguments constituted by the Figure (or referent) and the Ground (or relatum), where the projected angle is found by reference to an intrinsic or inherent facet of the Ground object. In contrast, (2) and (5) have a different logical structure: "in front of" is here a **ternary** relation, presuming a viewpoint V (the origin of the coordinate system), a Figure and Ground, all distinct.[33] In fact, these two kinds of spatial relation have quite different logical properties, as demonstrated elsewhere by Levelt (1984, this volume), but only when distinguished and grouped in this way (more in a moment). Let us dub the binary relations '**intrinsic**', but the ternary relations '**relative**' (because the descriptions are always relative to a viewpoint, in contradistinction to 'absolute' and 'intrinsic' descriptions).

To summarize then, the proposed classification is:

(1) "The ball is in front of me"
Coordinates: **Intrinsic**
Origin: Speaker

---

[32] The reader may feel that the notion of 'front' is different for chairs and persons (and so of course it is), and in particular that 'in front of me' is somehow more abstract than 'in front of the chair'. But notice that we could have said "at my feet" or "at the foot of the chair" - here 'foot' clearly means something different in each case, but shares the notion of an intrinsic part of the relatum object.
[33] The importance of the distinction between binary and ternary spatial relators was pointed out by Herrmann 1990.

Relatum: Speaker

(3) "The ball is in front of the chair (at the chair's front)"
Coordinates: **Intrinsic**
Origin: Chair
Relatum: Chair

(4) "The ball is in front of you"
Coordinates: **Intrinsic**
Origin: Addressee
Relatum: Addressee

(2) "The ball is in front of the tree"
Coordinates: **Relative**
Origin: Speaker
Relatum: Tree

(5) "The ball is to the right of the lamp, from your point of view"
Coordinates: **Relative**
Origin: Addressee
Relatum: Lamp

(6) "John noticed the ball to the right of the lamp" (or
"For John, the ball is in front of the tree")
Coordinates: **Relative**
Origin: Third person (John)
Relatum: Lamp (or Tree)

Note that use of the Intrinsic system of coordinates entails that Relatum (Ground) and Origin are constituted by the same object (the spatial relation is binary, between F and G), while use of the Relative system entails that they are distinct (the relation is ternary, between F, G and viewpoint V). Note too that whether the center is deictic, i.e. whether the origin is speaker (or addressee) or not, is simply irrelevant to this classification. This is obvious in the case of the grouping of (1), (3) and (4) together. It is also clear that although the viewpoint in Relative uses is normally speaker-centric, it may easily be addressee-centric or even centered on a third party as illustrated in (6). Hence *deictic and intrinsic are not opposed;* instead we need to oppose (a) coordinate systems 'intrinsic' vs. 'relative', on the one hand, and (b) origins 'deictic' and 'non-deictic' (or, alternatively, ego-centric vs. allocentric) on the other. Since frames of reference are coordinate systems, it follows that in language, frames of reference cannot be distinguished according to their characteristic, but variable, origins.

I expect a measure of resistance to this reformation of the distinctions, if only because the malapropism 'deictic frame of reference' has become a well-worn phrase. How, the critic will argue, can you define the frames of reference if you no longer employ the feature of deicticity to distinguish them? I will expend considerable effort in that direction in section 3.3.2. But first we must compare these two systems with the third system of coordinates in natural language, namely Absolute frames of reference. Let us review them together.

3.3.1 The three linguistic frames of reference

As far as we know, and according to a suitably catholic construal, there are exactly three frames of reference grammaticalized or lexicalized in language (often, lexemes are ambiguous over two of these frames of reference, sometimes expressions will combine two frames,[34] but often each frame will have distinct lexemes associated with it).[35] Each of these frames of reference encompasses a whole family of related but distinct semantic systems.[36] It is probably true to say that even the most closely related languages (and even dialects within them) will differ in the details of the underlying coordinate systems and their geometry, the preferential interpretation of ambiguous lexemes, presumptive origins of the coordinates, etc. Thus the student of language can expect that expressions glossed as, say, intrinsic 'side' in two languages will differ considerably in the way in which 'side' is in fact determined, how wide and how distant a search domain it specifies, etc. With that caveat, let us proceed.

Let us first define a set of primitives necessary for the description of all systems.[37] The application of some of the primitives is sketched in Figure 9, which illustrates three canonical exemplars from each of our three main types of system. Minimally, we need the primitives in the following table, the use of which we will illustrate in passing:

((Figure 9 about here))

[34] For example, the Australian language Guugu Yimithirr has (derived) lexemes meaning 'north side of', 'south side of', etc. , which combine both intrinsic and absolute frames of reference in a single word. Less exotically, English "on" as in "the cup on the table" would seem to combine absolute (vertical) information with topological information (contact) with intrinsic information (supporting planar surface).

[35] This point is important: some psychologists have been tempted to presume, because of the ambiguity of English ‘in front’ etc., that frames of reference are imposed on language by a spatial interpretation, rather than being distinguished semantically (see e.g. Carlson-Radvansky & Irwin 1993).

[36] We know one way in which this tripartite typology may be incomplete: some languages use conventionalized landmark systems which in practice grade into Absolute systems; however, there are some reasons for thinking that landmark systems and fixed-bearings systems are distinct conceptual types.

[37] I am indebted to many discussions with colleagues (especially perhaps Balthasar Bickel, Eric Pederson, David Wilkins) over the details of this scheme, although they would not necessarily agree with this particular version.

Table 2: Inventory of Primitives

1. System of labelled angles:
i.e. labelled arcs specified by coordinates around origin (language specific) ; such labels may or may not form a fixed armature or template of oppositions.

2. Coordinates:
(a) coordinates may be polar, by rotation from a fixed x-axis, or rectangular, by specification of two or more axes;
(b) one primary coordinate system C can be mapped from origin X to secondary origin X_2, by the following transformations:
-translation,
-rotation
-reflection
(and possibly by combination)
to yield a secondary coordinate system C_2.

3. Points:

F	=	Figure or referent with center point at volumetric center Fc
G	=	Ground or relatum, with volumetric centre Gc, and with a surrounding region R
V	=	viewpoint
X	=	origin of the coordinate system, X_2 = secondary origin
A	=	Anchor point, to fix labelled co-ordinates
L	=	designated landmark

4. Anchoring system: using -
A = Anchor point e.g. within G OR V; in landmark systems A = L.
'Slope' = fixed bearing system, yielding parallel lines across environment in each direction

Combinations of these primitives yield a large family of systems which may be classified in the following tripartite scheme:

1. Intrinsic frame of reference:
Informally, this frame of reference involves an object-centered coordinate system, where the coordinates are determined by the 'inherent' features, sidedness or facets of the object to be used as the Ground or relatum.

The phrase 'inherent features', though widely used in the literature, is misleading: such facets have to be conceptually assigned according to some algorithm, or learned on a case-by-case basis, or more often a combination of these. The procedure varies fundamentally across languages. In English, it is (apart from "top" and "bottom", and special arrangements for humans and animals) largely functional (see e.g. the sketch in Miller & Johnson-Laird 1976:403), so that the 'front' of a TV is the side we attend to, while the 'front' of a car is the facet that canonically lies in the direction of motion, etc. But in some languages, it is much more closely based on shape: e.g. in Tzeltal the assignment of sides utilizes a volumetric analysis very similar to the object-centered analysis proposed by Marr in the theory of vision, and function and canonical orientation is largely irrelevant (see Levinson in press).[38] In many languages the morphology makes it clear that human or animal body- (and occasionally plant-) parts provide a prototype for the opposed sides: hence we talk about the fronts, backs, sides, lefts and rights (and in many languages heads, feet, horns, roots, etc.) of other objects.[39] But whatever the procedure in a particular language, it relies primarily on the conceptual properties of the object: its shape, canonical orientation, characteristic motion and use, etc.

The attribution of such facets provides the basis for a coordinate system in one of two ways. Having found, e.g. the 'front', this may be used to anchor a ready-made system of oppositions 'front', 'back', 'sides', etc.[40] Alternatively, in other languages, there may be no such fixed armature as it were, each object having parts determined e.g. by specific shapes; in that case, finding 'front' does not predict the locus of 'back', etc., but nevertheless determines a direction from the volumetric centre of the object through the 'front' which can be used for spatial description.[41] In either case, we can use the designated facet to extract an angle, or line, radiating out from the Ground object, within or on which the Figure object can be found ("the statue in front of the town hall").

The geometrical properties of such intrinsic coordinate systems vary cross-linguistically. Systems with fixed armatures of contrastive expressions generally require the angles projected to be mutually exclusive (non-overlapping), so that in the intrinsic frame of reference (unlike the relative one) it makes no sense to say "The cat is to the front and to the left of the truck"). Systems utilizing single parts make no such constraints (cf. "The cat is in front of, and at the foot of, the chair"). In addition, the metric extent of the search domain designated (e.g. how far the cat is from the truck) can vary greatly: some languages require Figure and Ground to be in contact, or visually continuous, others allow the projection of enormous search domains ("in front of the church lie the mountains, running far off to the horizon"). More often, perhaps, the notion of a region, an object's penumbra as it were, is relevant, related to its scale.[42]

---

[38] Thus the 'face' of a stone may be the bottom surface hidden in the soil, as long as it meets the necessary axial and shape conditions.

[39] We tend to think of human prototypes as inevitably the source of such prototype parts. But such anthropomorphism may be ethnocentric: e.g. in Mayan languages plant-parts figure in human-part description (see Laughlin 1975, Levinson, in press).

[40] Thus Miller & Johnson-Laird 1976:401, thinking of English-speakers: "People tend to treat objects as six-sided. If an object has both an intrinsic top and bottom, and an intrinsic front and back, the remaining two sides are intrinsically left and right...". Incidentally, the possession of 'intrinsic left/right' is perhaps an indication that such systems are not exclusively object-centered (since left and right cannot be distinguished without an external frame of reference).

[41] For a nice contrast between two apparently similar MesoAmerican systems, one of which is armature-based, the other based on the location of individual facets, see MacLaury (1989) on Zapotec, and Levinson in press on Tzeltal.

[42] See the notion of intrinsic region in Miller & Johnson-Laird 1976. They suggest this may be linked to perceptual contiguity within 10 degrees of visual arc (1976:91), but that this perceptual notion of region has a conceptual counterpart which combines perceptual and functional information about the region of social or physical interaction of one body with another (1976:387-8).

More exactly:
An Intrinsic spatial relator R is a binary spatial relation, with arguments F and G, where R typically names a part of G. The origin X of the coordinate system C is always on (the volumetric centre of) G. An intrinsic relation R(F, G) asserts that F lies in a search domain extending from G on the basis of an angle or line projected from the centre of G, through an anchor point A (usually the named facet "R"), outwards for a determined distance. F and G may be any objects whatsoever (including ego), and F may be a part of G. The relation R does not support transitive inferences, nor converse inferences (see below).

Coordinates may or may not come in fixed armatures. When they do, they tend to be polar: e.g. given that facet A is the 'front' of a building, clockwise rotation in 90 degree steps will yield 'side', 'back', 'side'. Here there is a set of four labelled oppositions, with one privileged facet A: given A, we know which facet 'back' is. Since A fixes the coordinates, we call it the Anchor point. But coordinates need not be polar, or indeed part of a fixed set of oppositions: e.g. given that facet B is the 'entrance' of a church, Gc its volumetric centre, we may derive a line B-Gc (or an arc with angle determined by the width of B) - thus 'at the entrance to the church' designates a search area on that line (or in that arc), with no necessary implications about the locations of other intrinsic parts, 'front', 'back', etc. Since A determines the line, we call A the 'Anchor point'.

2. Relative frame of reference:
This is roughly equivalent to the various notions of viewer-centered frame of reference mentioned above (e.g. Marr's 2.5D, or the psycholinguists' 'deictic'). But it is not quite the same. It presupposes a ‘viewpoint’ V (given by the location of a perceiver in any sensory modality), and a Figure and Ground distinct from V. It thus offers a triangulation of three points, and utilizes co-ordinates fixed on V to assign directions to Figure and Ground. English "the ball is to the left of the tree" is of this kind of course. Since the perceptual basis is not necessarily visual, calling this frame of reference 'viewer-centered' is potentially misleading, but perhaps innocent enough. Calling it deictic, however, is potentially pernicious, because the 'viewer' need not be ego, and need not be a participant in the speech event - cf. "Bill kicked the ball to the left of the goal". Nevertheless, there can be little doubt that the deictic uses of this system are basic (prototypical), conceptually prior and so on.

The coordinate system, based on viewer V, seems generally to be based on the planes through the human body, giving us an 'up'/'down', 'back'/'front' and 'left'/'right' set of half-lines. Such a system of coordinates can be thought of as centered on the main axis of the body and anchored by one of the body-parts (e.g. my chest). In that case we have polar coordinates, with quadrants counted clockwise from 'front’ to 'right' , 'back' and ‘left’ (Herskovits 1986). Although the position of the body of viewer V may be one criterion for anchoring the coordinates, the direction of gaze may be another, and there is no doubt that Relative systems are closely hooked into visual criteria. Languages may differ in the weight given to the two factors, e.g. the extent to which occlusion plays a role in the definition of 'behind'.

But this set of coordinates on V is only the basis for a full Relative system; in addition a secondary set of coordinates is usually derived by mapping (all or some of) the coordinates on V onto the relatum or Ground object G. The mapping involves a transformation which may be 180 degree rotation, translation (movement without rotation or reflection) or arguably reflection across the frontal transverse plane. Thus "The cat is in front of the tree" in English entails that the cat F is between V and G (the tree), because the primary coordinates on V appear to have been rotated in the mapping onto G, so that G has a 'front' before which the cat sits. Hausa (Hill 1982) and many other languages *translate* rather than rotate the coordinates, so that a sentence glossing 'the cat is in front of the tree' will mean what we would mean in English by 'The cat is behind the tree'. But English is also not so simple, for rotation will get 'left' and 'right' wrong in English: "The cat is to the left of the tree" has left on the same side as V, not rotated. In Tamil, the rotation is complete: thus just as front and back are reversed, so are left and right, so the Tamil sentence glossed 'The cat is on the left side of the tree' would (on the relevant interpretation) mean 'The cat is on V's right of the tree'. To get the English system right, we might suppose that the coordinates on V should be *reflected* over the transverse plane, as if we wrote the coordinates of V on a sheet of acetate and flipped it over in front of V and placed it on G. This will get 'front', 'back', 'left' and 'right' at least in the correct polar sequence around the secondary origin. But it may not be the correct solution, since other interpretations are possible, and indeed more plausible.[43] But the point to establish here is that a large variation of systems is definable, constituting a broad family of Relative systems.

Not all languages have terms glossing 'left', 'right', 'front', 'back'. Nor does the possession of such a system of oppositions guarantee the possession of a Relative system. Many languages use such terms in a more or less purely intrinsic way (even when they are primarily used with deictic centers): i.e. they are used as binary relations specifying the location of F within a domain projected from a part of G (as in "To my left", "In front of you", "At the animal's front", "At the house's front", etc.). The test for a Relative system is (a) its utilizability with what is culturally construed as an object without intrinsic parts,[44] (b) whether there is a ternary relation with viewpoint V distinct from G, such that when V is rotated around the array, the description changes (see below). Now, languages that do indeed have a Relative system of this kind, also tend to have an Intrinsic system sharing at least some of the

---

[43] It may be that 'left' and 'right' are centered on V, while 'front' and 'back' are indeed rotated and have their origin on G. Evidence for that analysis comes from various quarters. First, some languages (like Japanese) allow both the English and Hausa style interpretations of 'front', while maintaining 'left' and 'right' always the same, suggesting that there are two distinct subsystems involved. Secondly, English 'left' and 'right' are not clearly centered on G, since something can be to the left of G but not in the same plane at all (e.g. "the mountain to the left of the tree"), while 'front' and 'back' can be centered on G, so that it is odd to say of a cat near me that it is in front of a distant tree. Above all, there is no contradiction in "the cat is to the front and to the left of the tree". An alternative analysis of English would have the coordinates fixed firmly on V, and give 'F is in front of the tree' an interpretation along the lines 'F is between V and G' ('behind' glossing 'G is between V and F'). My own guess is that English is semantically general over these alternative interpretations.

[44] Note that e.g. we think of a tree as unfeatured on the horizontal dimension, so that it lacks an intrinsic front, while e.g . some Nilotic cultures make the assumption that a tree has a front, away from the way it leans.

same terms.[45] This typological implication, apart from showing the derivative and secondary nature of Relative systems, also more or less guarantees the potential ambiguity of 'left', 'right', 'front', 'back' systems (although they may be disambiguated syntactically, as in 'to the left of the chair' vs. 'at the chair's left'). Some languages that lack any such systematic Relative system, may nevertheless have encoded the odd isolated Relative notion, like 'F is in my line of sight towards G'.

Relative systems that clearly use secondary coordinates mapped from V to G suggest that these mappings are by origin a means of extending the intrinsic frame of reference to cases where it would not otherwise apply (and this may suggest that the intrinsic system is rather fundamental in human linguistic spatial description[46]). Through projection of co-ordinates from the viewpoint V, we assign pseudo-intrinsic facets to G, as if trees had inherent fronts, backs and sides.[47] For some languages, this is undoubtedly the correct analysis: the facets are thus named and regions projected with the same limitations that hold of intrinsic regions.[48] Thus many Relative systems can be thought of as *derived intrinsic* ones - systems that utilize Relative conceptual relations to extend and supplement Intrinsic ones. One particular reason to so extend Intrinsic systems is their extreme limitations as regards logical inference of spatial relations from linguistic descriptions: Intrinsic descriptions support neither transitive nor converse inferences, but Relative ones do (Levelt 1984, this volume, and see below).[49]

Although from a perceptual point of view a frame of reference like the Relative one seems entirely fundamental, from a linguistic point of view it is not. In fact it is entirely dispensable. Western children learn this kind of system very late (mastering 'projective' left and right only by 11 or 12). Many languages simply do not employ this frame of reference at all,[50] or only in marginal uses of 'intrinsic' or 'absolute' lexical items. That means such languages have no way of expressing notions like 'in front/behind/to the left/right/side of the tree' as determined by the location of a 'viewer' or speaker. This will probably come as a bit of a shock to psychologists, who have, on the basis of familiar languages, confidently predicted its universality (Clark 1973, Miller & Johnson-Laird 1976, Takano 1989, etc.).

More exactly:

A Relative relator R expresses a ternary spatial relation, with arguments V, F and G, where F and G are unrestricted as to type, except that V and G must be distinct.[51] The primary coordinate system always has its origin on V; there may be a secondary coordinate system with origin on G. Such coordinate systems are normally polar, for

---

[45] But some languages encode Relative concepts based directly on visual occlusion or the absence of it, which do not have Intrinsic counterparts (as S. Kita has pointed out to me).

[46] As shown by its priority in acquisition (Johnson & Slobin 1978). On the other hand, some languages hardly utilize an intrinsic frame of reference at all (see e.g. Levinson 1992b on an Australian language).

[47] I owe the germ of this idea to Eric Pederson.

[48] This does not seem, once again, the right analysis for English 'left'/'right', since F and G need not be in the same plane at all ('the tree to the left of the rising moon'), and intuitively 'to the left of the ball' does not ascribe a left-facet to the ball.

[49] Although transitivity and converseness in Relative descriptions hold only on the presumption that V is held constant.

[50] Conversely, other languages like Tamil use it in more far-reaching ways.

[51] F may be a part of G, as "the bark on the left (side) of the tree".

example, 'front', 'right', 'back' and 'left' may be assigned by clockwise rotation from 'front'. Coordinate systems built primarily on visual criteria may not be polar, but be defined e.g. by rectangular coordinates on the two dimensional visual field (the retinal projection) so that 'left' and 'right' are defined on the horizontal or x axis, and 'front' and 'back' on the vertical or y axis ('back' has (the base of) F higher than G and/or occluded by G).

Terms that may be glossed 'left' and 'right' may involve no secondary coordinates, although they sometimes do (as when they have reversed application from the English usage). Terms glossed 'front' and 'back' normally do involve secondary coordinates (but compare the analysis in terms of vectors by O'Keefe this volume). Secondary coordinates may be mapped from primary origin on V to secondary origin on G under the following transformations: rotation, translation, and (arguably) reflection.[52]

Typological variations of such systems include: degree to which a systematic polar system of coordinates is available, degree of use of secondary coordinates, type of mapping function (rotation, translation, reflection) for secondary coordinates, differing anchoring systems for the coordinates (e.g. body-axis vs. gaze), differing degrees to which visual criteria (like occlusion, or place in retinal field) are definitional of the terms.

3. Absolute frame of reference

Amongst the many uses of the notion 'absolute' frames of reference, one refers to the fixed direction provided by gravity (or the visual horizon under canonical orientation). Less obviously of psychological relevance, the same idea of fixed directions can be applied to the horizontal. In fact, many languages make extensive, some almost exclusive, use of such an absolute frame of reference on the horizontal. They do so by fixing arbitrary fixed bearings, 'cardinal directions', corresponding one way or another to directions or arcs that can be related by the analyst to compass bearings. Speakers of such languages can then describe an array of e.g. a spoon in front of a cup, as 'spoon to north of cup' (etc.) without any reference to the viewer/speaker's location.

Such a system requires that persons maintain their orientation with respect to the fixed bearings at all times. People who speak such languages can be shown to do so - e.g. they can dead-reckon current location in unfamiliar territory with extraordinary accuracy, and thus point to any named location from any other (Lewis 1976, Levinson 1992b). How they do so is simply not known at the present time, but we may presume that a heightened sense of inertial navigation is regularly cross-checked with many environmental clues.[53] Indeed, many such systems are clearly abstractions and

---

[52] Rotation will have 'front' towards V, and clockwise (looking down on G) from 'front': 'right', 'back', 'left' (as in Tamil). Translation will have 'back' towards V, and clockwise from 'back': 'left', 'front', 'right' (as in Hausa). Reflection will have 'front' towards V, but clockwise from 'front' : 'left', 'back', 'right' (as in English, on one analysis). The rotation and translation cases clearly involve secondary polar coordinates on G. The reflection cases can be reanalyzed as defined by horizontal and vertical coordinates on the retinal projection, or can be thought of (as seems correct for English) as the superimposition of two systems, the left-right terms involving only primary coordinates on V, and the front-back terms involving rotated secondary coordinates on G.

[53] Environmental clues will not explain such heightened dead-reckoning abilities outside familiar territory, which seem to exist. I presume that such people have been socialized to constantly compute

refinements from environmental gradients (mountain slopes, prevailing wind directions, river drainages, celestial azimuths, etc.).[54] These 'cardinal directions' may therefore occur with fixed bearings skewed at various degrees from, and in effect unrelated to, our 'north', 'south', 'east' and 'west'. It perhaps needs emphasizing that this keeping track of fixed directions is, with appropriate socialization, not a feat restricted to certain ethnicities, races, environments or culture types as shown by its widespread occurrence (in perhaps a third of all human languages?) from MesoAmerica, to New Guinea, to Australia, to Nepal. No simple ecological determinism will explain the occurrence of such systems, which can be found alternating with e.g. Relative systems, across neighbouring ethnic groups in similar environments, and which occur in environments of contrastive kinds (e.g. wide open deserts and closed jungle terrain).

The conceptual ingredients for such systems are simple: the relevant linguistic expressions are binary relators, with Figure and Ground as arguments, and a system of co-ordinates anchored to fixed bearings, which always have their origin on the Ground. In fact, these systems are the only systems with conceptual simplicity and elegance. For example, they are the only systems that fully support transitive inferences across spatial descriptions: Intrinsic descriptions do not do so, and Relative ones only do so if viewpoint V is held constant (Levelt 1984). Intrinsic systems are dogged by the multiplicity of object types, the differing degrees to which the asymmetries of objects allow the naming of facets, and the problem of 'unfeatured' objects. Relative systems are dogged by the psychological difficulties involved in learning left/right distinctions, the complexities involved in mapping secondary coordinates, and because they are often developed from Intrinsic ones, display ambiguities across frames of reference (like English 'in front of'). The liabilities of Absolute systems are not, on the other hand, logical but psychological: they require a cognitive overhead, namely the constant background calculation of cardinal directions, together with a system of dead reckoning that will specificy for any arbitrary point P which direction P is from ego's current locus (so that ego may refer to the location of P).

Absolute systems may also show ambiguities of various kinds. First, places of particular sociocultural importance may come to be designated by a cardinal direction term, like a quasi-proper name, regardless of their location with respect to G. Secondly, where the system is abstracted out of landscape features, the relevant expressions (e.g. 'uphill' or 'upstream') may either refer to places indicated by relevant local features (e.g. local hill, local stream), or to the abstracted fixed bearings, where these do not coincide. Thirdly, some such systems may even have Relative interpretations (e.g. 'uphill' may imply further away in my field of vision; cf. our interpretation of 'north' as top of a map, etc.).

---

direction as a background task, by inertial navigation with constant checks with visual information and other sensory information (e.g. sensing wind direction). But see Baker (1989) who believes in faint human magnetoreception.

[54] Note that none of these environmental bases can provide the cognitive basis of abstracted systems: once the community has fixed a direction, it remains in that direction regardless of fluctuations in local landfall, drainage, wind source, equinox, etc., or even removal of the subject from the local environment. Thus the environmental sources of such systems may explain their origins, but do not generally explain how they are used, or how the cardinal directions are psychologically 'fixed'.

One crucial question with respect to Absolute systems is how, conceptually, the coordinate system is thought of. It may be a polar system, as in our 'north', 'south', 'east' and 'west', where north is the designated anchor and east, south, west found by clockwise rotation from north.[55] Other systems may have a primary and a secondary axis, so that e.g. a 'north'/'south' axis is primary, but it is not clear which direction, north or south, is itself the Anchor.[56] Yet other systems favour no particular primary reference point, each half-axis having its own clear Anchor or fixed central bearing.[57] Some systems like Tzeltal are 'degenerate', in that they offer two labelled half-lines (roughly 'north', 'south'), but label both ends of the orthogonal with the same terms. Even more confusing, some systems may employ true abstracted cardinal directions on one axis, but landmark designations on the other, guaranteeing that the two axes do not remain orthogonal when arrays are described in widely different places. Thus on Bali, and similarly for many Austronesian systems, one axis is determined by monsoons, and is a fixed, abstracted axis, but the other is determined by the location of the central mountain, and thus the bearing varies continuously when one circumnavigates the island. Even where systematic cardinal systems exist, the geometry of the designated angles is variable. Thus, if we have four half-lines based on orthogonal axes, the labels may describe quadrants (as in Guugu Yimithirr), or they may have e.g. narrower arcs of application on one axis than the other (as appears to be the case in Wik Mungan[58]). Even in English, though we may think of North as a point on the horizon, we also use arcs of variable extent for informal description.

More precisely:
An Absolute relator R expresses a binary relation between F and G, asserting that F can be found in a search domain at the fixed bearing R from G. The origin X of the coordinate system is always centered on G. G may be any object whatsoever, including ego or another deictic centre; F may be a part of G. The geometry of the coordinate system is linguistically/culturally variable, so that in some systems equal quadrants of 90 degrees may be projected from G, while in others something more like 45 degrees may hold for arcs on one axis, and perhaps 135 degrees on the other. The literature also reports abstract systems based on star-setting points, which will then have uneven distribution around the horizon.

Just as Relative relators can be understood to map designated facets onto Ground objects (thus "on the front of the tree" assigns a named part to the tree), so Absolute relators may also do so. Many Australian languages for example have cardinal edge roots, then affixes indicating e.g. 'northern edge', etc. Some of these stems can then only be analyzed as an interaction between the Intrinsic facets of an object and Absolute directions.

---

[55] Due no doubt to the introduction of the compass in medieval times. Before, maps typically had east at the top, hence our term 'orient oneself', showing that our use of polar coordinates is older than the compass.

[56] Warlpiri may be a case in point. Note that such a system may be based on a solar compass, but since solstitial variation makes it necessary to abstract an equinoctial bisection of the seasonal movement of the sun along the horizon, it is less confusing to fix the system by reference to a mentally constituted orthogonal.

[57] Guugu Yimithirr would be a case in point, since there are no elicitable associations of sequence or priority between cardinal directions.

[58]. The Wik Mungan system (another Aboriginal language of Cape York) was described by Peter Sutton in a presentation to the Australian Linguistics Institute, Sydney, June 1992.

3.3.2 The 'logical structure' of the three frames of reference

We have argued that as far as language is concerned we must distinguish frame of reference *qua* coordinate system from e.g. deictic centre *qua* origin of the coordinate system. Still, the sceptical may doubt that this is either necessary or possible.

First, to underline the necessity: each of our three frames of reference may occur with or without a deictic center (or egocentric origin). Thus for the Intrinsic frame, we can say " the ball is in front of me" (deictic center); for the Absolute frame we can say "the ball is north of me"; and of course in the Relative frame, we can say "the ball is in front of the tree (from ego's point of view)". Conversely, none of the three frames need have a deictic center: thus in the Intrinsic frame one can say "in front of the chair", in the Absolute frame "north of the chair", and in the Relative frame "in front of the tree from Bill's point of view". This is just what we should expect given the flexible nature of linguistic reference - it follows from Hockett 's (1960) design feature of displacement, or Bühler's (1982 (1934)) concept of transposed deictic center.

Second, we need to show that we can in fact define the three frames of reference adequately without reference to the opposition deictic vs. non-deictic center or origin. We have already hinted at plenty of distinguishing characteristics of each of the three frames. But to collect them together, first consider the logical properties. The Absolute and Intrinsic relators share the property that they are binary relations whereas Relative relators are ternary. But Absolute and Intrinsic are distinguished in that Absolute relators define asymmetric transitive relations (if F1 is north of G, and F2 is north of F1, then F2 is north of G), where converses can be inferred (if F is north of G, G is south of F). The same does not hold for Intrinsic relators, which hardly support any spatial inferences at all without further assumptions (see Levelt 1984, and this volume). In this case, Absolute and Relative relators share logical features, since Relative relators support transitive and converse inferences provided that viewpoint V is held constant.

This is already sufficient to distinguish the three frames. But we may add further distinguishing factors. Certain important properties follow from the nature of the anchoring system in each case. In the Intrinsic case we can think of the named facet of the object as providing the Anchor, in the Relative case we can think of the viewpoint V on an observer, and the anchor being constituted by e.g. the direction of his front or his gaze, while in the Absolute case either one or more of the labelled fixed bearings establishes a conceptual 'slope' across the environment, thus fixing the coordinate system. From this certain distinct properties under rotation emerge as illustrated in Figure 10.[59] These properties have a special importance for the study of non-linguistic conceptual coding of spatial arrays, as they allow systematic experimentation (as illustrated in part 1; see also Levinson 1992b, Brown & Levinson 1993b, Pederson 1993, 1994, Danziger (ed.) 1993).

((Figure 10 about here))

---

[59] I am grateful to David Wilkins, and other colleagues, for helping me to systematize these observations.

Altogether then we may summarize the distinctive features of each frame of reference as in Table 3; these features are jointly certainly sufficient to establish the nature of the three frames of reference independently of reference to the nature of the origin of the coordinate system.

Table 3: Summary of properties of different frames of reference

	INTRINSIC	ABSOLUTE	RELATIVE
relation:	binary	binary	ternary
origin on:	Ground	Ground	Viewpoint V
anchored by:	A within G	'slope'	A within V
transitive:	no	yes	yes if V const
constancy under rotation of:			
whole array:	yes	no	no
viewer:	yes	yes	no
Ground:	no	yes	yes

We may conclude this discussion of the linguistic frames of reference with the following observations:
(a) Languages use, it seems, just three frames of reference: Absolute, Intrinsic, and Relative;
(b) Not all languages use all frames of reference: some use predominantly one only (Absolute or Intrinsic; Relative seems to require Intrinsic), some use two (Intrinsic and Relative, or Intrinsic and Absolute), while some use all three;
(c) Linguistic expressions may be specialized to a frame of reference, so we cannot assume that choice of frame of reference lies entirely outside language, e.g. in spatial thinking, as some have suggested. But spatial relators may be ambiguous (or semantically general) across frames, and often are.

3.3.3 Realigning Frames of reference across disciplines and modalities
We are now at last in a position to see how our three linguistic frames of reference align with the other distinctions in the literature arising from the consideration of other modalities (as listed in Table 1 above). The motive, recollect, is to try to make

sense of the very idea of 'same frame of reference' across modalities, and in particular from various kinds of nonlinguistic thinking to linguistic conceptualization.

An immediate difficulty is that, by establishing that frames of reference in language should be considered independent of the origin of the coordinate systems, we have opened up a gulf between language and the various perceptual modalities, where the origin of the coordinate system is so often fixed on some egocentre. But this mismatch is in fact just as it should be: language is a flexible instrument of communication, designed (as it were) so that one may express other persons' point of view, take other perspectives and so on. At the level of perception, origin and coordinate system presumably come pre-packaged as a whole, but at the level of language, and perhaps more generally at the level of conception, they can vary freely and combine.

So to realign the linguistic distinctions with distinctions made across other modalities, we need to 'fix' the origin of the coordinate system so that it coincides, or fails to coincide, with ego in each frame of reference. We may do so as follows. First, we may concede that the Relative frame of reference, though not necessarily egocentric, is prototypically so. Second, we may note that the Intrinsic system is typically, but not definitionally, non-egocentric. Thirdly, and perhaps most arbitrarily, we may assign a non-egocentric origin to the Absolute system. These assignments should be understood as special subcases of the uses of the linguistic frames of reference.

If we make these restrictions, then we can align the linguistic frames of reference with the other distinctions from the literature as in Table 4. [60]

Table 4: Aligning Classifications of Frames of Reference

If we fix whether the origin = Ego, or not as follows:

INTRINSIC	ABSOLUTE	RELATIVE
Restriction: *- ego*	*- ego*	*+ ego*
Object -Centered	Environment -Centered	Viewer- -Centered
intrinsic- perspective		deictic- perspective
3D Model		2.5D Sketch
allocentric	allocentric	egocentric
orientation- free	orientation- bound	orientation- bound

[60] This table owes much to the work of Eve Danziger (see e.g. her 1994).

Notice then that, under the restriction concerning the nature of the origin,
(a) Intrinsic and Absolute are grouped as allocentric frames of reference, as opposed to the egocentric Relative system;
(b) Absolute and Relative are grouped as orientation-bound, as opposed to Intrinsic which is orientation-free.
This correctly captures our theoretical intuitions: in certain respects Absolute and Intrinsic viewpoints are fundamentally similar - they are binary relations which are viewpoint independent, where the origin may happen to be ego but need not be: they are allocentric systems which yield an ego-invariant picture of the 'world out there'. On the other hand, Absolute and Relative frameworks are fundamentally similar on another dimension, because they both impose a larger spatial framework on an assemblage, and thus specify its orientation with respect to external coordinates: thus in an Intrinsic framework it is impossible to distinguish enantiomorphic wholes, while in either of the orientation-bound systems it is inevitable.[61] Absolute and Relative frameworks presuppose a Newtonian or Kantian spatial envelope, while the Intrinsic framework is Leibnizian.

The object-centered nature of the Intrinsic system hooks it up to Marr's 3D model in the theory of vision, and the nature of the linguistic expressions involved suggest that the Intrinsic framework is a generalization from the analysis of objects into their parts: a whole configuration can be seen as a single complex object, so that we can talk of the leading car in a convoy as 'the head of the line'. On the other hand, the viewer-centered nature of the Relative framework connects it directly to the sequence of 2D representations in the theory of vision. Thus the spatial frameworks in the perceptual systems can indeed be correlated with the linguistic frames of reference.

Let us summarize: I have sought to establish that there is nothing incoherent in the notion 'same frame of reference' across modalities or inner representation systems. Indeed, even the existing distinctions that have been proposed can be seen in many detailed ways to correlate with the revised linguistic ones, once the special flexibility of the linguistic systems with respect to origin is taken into account. Thus it should be possible, and intellectually profitable, to formulate the distinct frames of reference in such a way that they have cross-modal application. Notice that this view conflicts with the views of some that frames of reference in language are imposed just in the mapping from perception to language via the encoding process. On the contrary, I shall presume that any and every spatial representation, perceptual or conceptual, must involve a frame of reference, e.g. retinotopic images just are, willy nilly, in a viewer-centered frame of reference.

But at least one major problem remains: it turns out that the three distinct frames of reference are 'untranslatable' from one to the other, which throws further doubt on the idea of correlations and correspondences across sensory and conceptual representational levels. This brings us to Molyneux's question.

4.0 Molyneux's question

[61] See van Cleve & Frederick 1991 for discussion of this Kantian point. For the cross-cultural implications, and a working out of the place of Absolute systems in all this, see Danziger 1994.

In 1690 William Molyneux wrote John Locke a letter posing the following celebrated question: if a blind man, who knew by touch the difference between a cube and a sphere, had his sight restored, would he recognize the selfsame objects under his new perceptual modality or not?[62]

The question whether our spatial perception and conception is modality specific is as live now as then. Is there one central spatial model, to which all our input senses report, and from which instructions can be generated appropriate to the various output systems (touch, movement, language, gaze and so on)?

There have of course been attempts to answer Molyneux directly, but the results are conflicting: on the one hand, sight-restored individuals take a while to adjust (Gregory 1987:94-6, Valvo 1971), monkeys reared with their own limbs masked from sight have trouble relating touch to vision when the mask is finally removed (Howard 1987:730f), and touch and vision are attuned to different properties (the tactile sense is more attuned to weight and texture than shape, etc.; Klatsky & Lederman 1993); on the other hand, human neonates immediately extrapolate from touch to vision (Meltzoff 1993) and the neurophysiology suggests direct cross-wirings (Berthoz 1991:81; but cf. Stein 1992), so that some feel that the answer to the question is a "resounding 'yes'" (Eilan 1993:237). More soberly, it seems that there is some innate supramodal system observable in monkeys and infants, but it may be very restricted, and sophisticated cross-modal thinking may even be dependent on language.[63]

Here I want to suggest another way to think about this old question. Put simply, we may ask whether the same frames of reference can in principle operate across all the modalities, and if not, whether at least they can be translated into one another.

What we should mean by 'modality' here is an important question. In what follows I shall assume that corresponding to (some of) the different senses, and more generally to input/output systems, there are specialized 'central' representational systems, e.g. an imagistic system related to vision, a propositional system related to language, a kinaesthetic system related to gesture, and so on (cf. Levelt 1989, Jackendoff 1991, etc.). Our version of Molyneux's question then becomes:
(a) do the different representational systems natively and necessarily employ certain frames of reference?
(b) if so, can representations in one frame of reference be translated (converted) into another frame of reference?

Let us discount here the self-evident fact that certain kinds of information may perhaps in principle be modality-specific: e.g. spatial representations in an imagistic mode must, it seems, be determinate with respect to shape, while those in a propositional mode need not, and perhaps, cannot be so.[64] Similarly, the haptic-kinaesthetic modality will have available direct information about weight, texture,

[62] The problem is discussed in Locke, Essay on Human Understanding, book II, ix, 8. The question was brought back into philosophical discussion by Gareth Evans, and many of the papers in Eilan, McCarthy & Brewer (1993) explicitly address Molyneux's question.
[63] See e.g. Ettlinger 1987:174: "language serves as a cross-modal bridge".
[64] The issue may be less clear than it at first seems; see Tye, 1991:5-9.

tactile warmth and three-dimensional shape which we can only guess at from visual information (Klatsky & Lederman 1993), while the directional and inertial information from the vestibular system is of a different kind again. All this would seem to rule out a *single* supra-modal spatial representation system: What hybrid monster would a representation system have to be to record such disparate information? All that concerns us here is the compatibility of *frames of reference* across modalities.

So, first, let us take the question of translatability across frames of reference. This is the easier question, and the answer to it offers an indirect answer to the first question. Here there is a striking, but on a moment's reflection, self-evident fact: one cannot freely convert information from one framework to another. Consider for example an array, with a bottle on the ground at the (intrinsic) front side of a chair; suppose too that you view the array from a viewpoint such that the bottle is to the right of the chair, and as it happens the bottle is also north of the chair (see Figure 11). Now I ask you to remember it, and suppose you 'code' the scene in an Intrinsic frame of reference: 'bottle in front of chair', discarding other information. It is immediately obvious that from this Intrinsic description you cannot later generate a Relative description - if you were viewing the array so that you are facing one side of the chair, then the bottle will be to the left of or to the right of the chair - depending on your viewpoint. So without a 'coding' or specification of the locus of the viewpoint V, one cannot generate a Relative description from an Intrinsic description. The same holds for an Absolute description: knowing that the bottle is at the front of the chair will not tell you whether it is north or south or east or west of the chair - for that you will need ancillary information. In short, one cannot get from an Intrinsic description - an orientation-free representation - to either of the orientation-bound representations.

What about conversions between the two orientation-bound frameworks? Again, it is clear that no conversion is possible: from a Relative description or coding 'the bottle is to the left of the chair' you do not know what cardinal direction it lies in, nor from 'the bottle is north of the chair' can one derive a viewpoint-relative description like 'to the left of the chair'.

((Figure 11 about here))

In fact, the only directions in which one may convert frames of reference is, in principle, from the two orientation-bound frames to the orientation-free one.[65] For if the description of the orientation of the Ground object is fully specified, then one can derive an Intrinsic description: e.g. from the Relative description 'the chair is facing to

---

[65] This possibility of getting from a Relative representation to an Intrinsic one may help to explain the apparent inconsistency between our findings here and Levelt's (this volume). The apparent inconsistency is that in Levelt's task subjects who made ellipses always presupposed an underlying uniform spatial frame of reference, even when their spatial descriptions varied between Relative and Intrinsic, thus suggesting that frames of reference might reside in the mapping from spatial representation to language rather than in the spatial representation itself. But the data are compatible with an analysis whereby the spatial representation is itself in a Relative frame of reference, and the mapping is optionally to an Intrinsic or Relative description, as Levelt acknowledges. The mapping from Relative to Intrinsic is one of the two mappings in principle possible between frames of reference, as here described, whereas a mapping from Intrinsic spatial representation to linguistic Relative representation would be in principle impossible. This would seem to explain all the data that we currently have in hand.

my right and the bottle is to the right of the chair in the same plane' one may in principle get to the Intrinsic specification 'the bottle is at the chair's front', and similarly from the Absolute description 'the chair is facing north and the bottle is to the north of the chair'. Normally, though, the orientation of the Ground object is irrelevant to the orientation-bound descriptions, so this remains only a translation in principle. Translations in all other directions are in principle 'out', i.e. impossible.

This simple fact about translatability across frames of reference may have far-reaching consequences. Consider for example the following syllogism:

(1) Frames of reference are incommensurable (i.e. a representation in one framework is not freely convertible into a representation in another);
(2) Each sense utilizes its own frame(s) of reference: e.g. while vision primarily uses a viewer-centered frame, touch arguably uses primarily an object-centered frame, based on the appreciation of form through three-dimensional grasping.
Ergo:
(3) Representations from one modality (e.g. haptic) cannot be freely translated into representations in another (e.g. visual).

The syllogism suggests then that the answer to Molyneux's question is negative - the blind man upon seeing for the first time will not recognize by sight what he knew before by touch. More generally, we will not be able to exchange information across any internal representation systems that are not based on one and the same frame of reference.

I take this to be a counter-intuitive result, a clearly false conclusion, in fact a *reductio ad absurdam*. The fact of the matter is that we can indeed form mental images of contour shapes by touch alone, we can gesture about what we have seen, we can talk about, or draw, what we have felt with our fingers and so on. Since premise (1) seems self-evident true, we must then reject premise (2), the assumption that each sensory modality or representational system operates exclusively in its own primary, proprietary frame of reference. In short, either the frame of reference must be the same across all modalities or representational systems in order to allow the cross-modal sharing of information, or each must allow more than one frame of reference.

Intuitively, this seems the correct conclusion. On the one hand, peripheral sensory systems may operate in proprietary frames of reference: e.g. low-level vision may know only of 2D retinotopic arrays, while otoliths only know of a gravitational frame of reference. But, on the other hand, at a higher level, visual processing seems to deliver 3D analyses of objects as well as 2D ones. Thus when we (presumably) use the visual system to imagine rotations of objects, we project from 3D models (Intrinsic) to 2.5D (Relative) ones, showing that both are available. Thus more central, more conceptual, levels of representation seem capable of adopting more than one frame of reference.

Here then is the first part of the answer to our puzzle. Representational systems of different kinds, specialized to different sensory modalities (like visual memory) or output systems (like gesture and language), may be capable of adopting different frames of reference. This would explain how it is that Tenejapans, or indeed Dutch subjects, can adopt the same frame of reference when utilizing different representation systems - those that generate gesture, those involved in tasks requiring visual memory, those involved in making spatial inferences, as well as those involved in speaking.

But to account for the facts described in part 1, it will not be sufficient to establish that the same frame of reference can in principle be used across different kinds of internal representation system, those involved in nonverbal memory, gesture and language, etc. To account for those facts, it will necessary to assume that individual subjects do indeed actually utilize the same frame of reference across modalities. But now we have an explanation for that apparent fact: for the non-translatability across frames of reference requires individuals to stabilize their representational systems within a limited set of frames of reference. For example, if a Tenejapan sees an array, and remembers it only in terms of a viewer-centered framework, he will not later be able to describe it - the language simply fails to provide a systematic viewer-centered frame of description. Thus the facts that
(a) frameworks are not freely convertible,
(b) languages may offer restricted frameworks as output,
(c) it may be desirable to describe any spatial experience whatsoever at some later point,
- these conspire to require that one codes spatial experiences at the time of experience in whatever output frameworks one's dominant language offers.

Conclusions

This paper began with some quite unexpected findings: languages can differ in the set of frames of reference they employ for spatial description; moreover, the options in the language seem to dictate the use of frames of reference in non-linguistic tasks - there seems thus to be a cross-modal tendency to fix on a dominant frame of reference. This raises a number of fundamental puzzles: what sense does it make to talk of 'same frame of reference' across modalities, or psychological faculties of quite different kinds? If it does make sense, why should it be so? What light does the phenomenon throw on how spatial information is shared across the senses, across the various 'input' and 'output' devices?

I have tried to sketch answers to these puzzles. The answers converge in two kinds of responses to Molyneux's question, 'do the senses talk to one another?'. The first kind of response is an empirical argument:
1. The Frame of Reference dominant in one's language 'infiltrates' other modalities, presumably to ensure that we can talk about what we see, feel, etc.
2. Therefore other modalities have the capacity to adopt, or adapt to, other frames of reference, which suggests a 'Yes' answer to Mr Molyneux.

The second kind of response is an *a priori* argument:
1. Frames of Reference cannot freely 'translate' into one another;
2. Therefore, if the modality most adaptive to external influences, namely language, adopts one Frame of Reference, the others must follow suit;
3. To do this, all modalities must have different Frames of Reference available, or be able to 'annotate' experiences with the necessary ancillary information, which suggests a 'Yes' answer to Mr Molyneux.

Actually, an affirmative answer to Molyneux's Question is evidently required - otherwise we couldn't talk about what we see. What is deeply mysterious is *how* this cross-modal transfer is achieved. The untranslatability across frames of reference greatly increases the puzzle. It is in this light that the findings with which we began - the standardization of frames of reference across modalities in line with the local language - now seem not only less surprising, but actually inevitable.

References

Acredolo, L. (1988) Infant mobility and spatial development. In J. Stiles-Davis, M. Kritchevsky & U. Bellugi (eds.), *Spatial cognition: Brain bases and development* (pp. 157-166). Hillsdale, NJ: Lawrence Erlbaum.

Asch, S. E. & Witkin, H. A. (1948) Studies in space orientation II: perception of the upright with displaced visual fields and with body tilted. *Journal of Experimental Psychology*, 38, 455-77. Reprinted Journal of Experimental Psychology, General, 1992, Vol 121, No. 4, 407-18.

Baayen, H. & Danziger, E. (eds.) (1994). *Annual Report of the Max Planck Institute for Psycholinguistics, 1993*. Nijmegen.

Baker, M. (1989) *Human navigation and magnetoreception*. Manchester: University of Manchester Press.

Berthoz, A. (1991) Reference frames for the perception and control of movement. In J. Paillard (ed.), *Brain and space* (pp. 81-111). Oxford: Oxford Science Publications.

Bickel, B. (1994) Spatial operations in deixis, cognition, and culture: Where to orient oneself in Belhare. Working paper no. 28, Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, Nijmegen.

Bierwisch, M. (1967) Some semantic universals of German adjectivals. *Foundations of Language*, 3, 1-36.

Bowerman, M. & Pederson, E. (in prep.) Typology and universals of ‘topological’ spatial relators. Max Planck Institute for Psycholinguistics, Nijmegen. (CHECK IN PROOF)

Brewer, B. & Pears, J. (1993) Frames of reference. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 25-30). Oxford: Blackwell.

Brown, P. (1991) Spatial conceptualization in Tzeltal. Working Paper 6, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.

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

Brown, P. & Levinson, S.C. (1993b) Explorations in Mayan Cognition. Working paper no. 24. Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics , Nijmegen.

Bühler, K. 1982 (1934). The deictic field of language and deictic words. In Jarvella, R. & Klein, W. (eds.) *Speech, Place and Action*, New York: Wiley, pp. 9-30.

Bülthoff, H.H. (1991) Shape from X: Psychophysics and computation. In M.S. Landy & J.A. Movshon (eds.), *Computational models of visual processing*. Cambridge, MA: MIT Press.

Campbell, J. (1993) The role of physical objects in spatial thinking. N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 65-95). Oxford: Blackwell.

Carlson-Radvansky, L.A. & Irwin, D. A. (1993) Frames of reference in vision and language: Where is above? *Cognition*, 46, 223-244.

Clark, H.H. (1973) Space, time, semantics and the child. In T.E. Moore (ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

Cohen, D. & Kubovy, M. (1993) Mental rotation, mental representation, and flat slopes. *Cognitive Psychology* 25:351-382.

Danziger, E. (ed.) (1993) Cognition and space kit version 1.0. Cognitive Anthropology Research Group, Nijmegen.

Danziger, E. (1994) As fresh meat loves salt: The logic of possessive relationships in Mopan Maya. Working Paper no. 30, Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, Nijmegen.

Eilan, N. (1993) Molyneux's question and the idea of an external world. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 236-255). Oxford: Blackwell.

Eilan, N., McCarthy, R. & Brewer, B. (1993) *Spatial representation: Problems in philosophy and psychology*. Oxford: Blackwell.

Einstein, A. (1954) Introduction to *Concepts of space. The history of theories of space in physics* by M. Jammer (1954). Cambridge, MA: Harvard University Press.

Ettlinger, G. (1987) Cross-model sensory integration. In R. Gregory (ed.), The *Oxford Companion to the Mind* (pp. 173-174). Oxford: Oxford University Press.

Fillmore, C. (1971) Towards a theory of deixis. Paper read at Pacific Conference on Contrastive Linguistics and Language Universals, January 1971, University of Hawaii. Mimeo.

Studies in linguistic semantics. New York: Holt, Rinehart & Winston.

Friederici, A. & Levelt, W.J.M. (1990) Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception & Psychophysics*, 47(3), 253-266.

Gregory, R.L. (1987) *Oxford companion to the mind*. Oxford: Oxford Univ. Press.

Haviland, J.B. (1993) Anchoring and Iconicity in Guugu Yimithirr pointing gestures. *Journal of Linguistic Anthropology*, 3(1), 3-45.

Herrmann, T. (1990). Vor, hinter, rechts und links: das 6H-Modell. In Zeitschrift für Literaturwissenschaft und Linguistik 78, 117-140.

Herskovits, A. (1986) Language and spatial cognition: An interdisciplinary study of the prepositions in English. *Studies in natural language processing*. Cambridge: CUP.

Hill, C. (1982) Up/down, front/back, left/right: a contrastive study of Hausa and English. In Weissenborn, J. & Klein, W. (eds.) *Here and there: cross-linguistic studies on deixis and demonstration*. Amsterdam: Benjamins, pp. 11-42.

Hill, D. (1994) Spatial configurations and evidential propositions. Working paper no. 25, Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, Nijmegen.

Hockett, C. F. (1960) The origin of speech. *Scientific American* 203:89-96.

Howard, I.P. (1987) Spatial coordination of the senses. In R.L. Gregory (ed.), The *Oxford Companion to the Mind* (pp. 727-32). Oxford: Oxford University Press.

Jackendoff, R. (1991) Parts and boundaries. *Cognition*, 41, 9-45.

Jackendoff, R. (this volume)

Jammer, M. (1954) *Concepts of space. The history of theories of space in physics*. Cambridge, MA: Harvard University Press.

Johnston, J.R. & Slobin, D. (1978) The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, 6, 529-545.

Just, M. & Carpenter, P. (1985) Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92(2), 137-172.

Kant, E. (1768) Von dem ersten Grunde des Unterschiedes der Gegenden im Raume. Translated as 'On the first ground of the distinction of regions in space' in Van Cleve, J. & Frederick, R.E. (eds.) (1991) *The philosophy of right and left: Incongruent counterparts and the nature of space* (pp. 27-34). Dordrecht: Kluwer.

Klatsky, R.L. & Lederman, S.J. (1993) Spatial and nonspatial avenues to object recognition by the human haptic system. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology*, pp. 191-205.

Kosslyn, S. M. (1980) *Image and Mind*. Cambridge, MA: Harvard University Press.

Landau, B. & Jackendoff, R. (1993) "What" and "where" in spatial language and spatial cognition. *Behavioural and Brain Sciences*, 16, 217-265.

Laughlin, R. (1975). *The Great Tzotzil Dictionary of San Lorenzo Zinacantán*. Washington, DC: Smithsonian.

Leech, G. (1969) *Towards a semantic description of English*. London: Longmans.

Levelt, W.J.M. (1984) Some perceptual limitations on talking about space. In A.J. van Doorn, W.A. van der Grind & J.J. Koenderink (eds.), *Limits in perception* (pp. 323-358). Utrecht: VNU Science Press.

Levelt, W.J.M. (1989) *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W.J.M. (1994) This volume.

Levinson, S.C. (1983) *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S.C. (1992a) Primer for the field investigation of spatial description and conception. *Pragmatics* 2(1), 5-47.

Levinson, S.C. (1992b) Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. Working paper no. 13, Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, Nijmegen.

Levinson, S.C. (1994) Vision, shape and linguistic description: Tzeltal body-part terminology and object description. Special volume of *Linguistics*, 32(4), 791-856.

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

Lewis, D. (1976) Route finding by desert aborigines in Australia. *Journal of Navigation*, 29, 21-38.

Logan, G.D. (1994) This volume.

Lyons, J. (1977) *Semantics* (vol. 1&2). Cambridge: CUP.

Marr, D. (1982) *Vision*. New York: Freeman.

MacLaury, R. (1989) Zapotec body-part locatives: Prototypes and metaphoric extensions. *International Journal of American Linguistics*, 55(2), 119-154.

McCullough, K. E. (1993) Spatial information and cohesion in the gesticulation of English and Chinese speakers. Paper presented at the Annual convention of the American Psychological Society.

McNaughton, B., Chen, L. & Markus, E. 1990. 'Dead reckoning', landmark learning and the sense of direction: a neurophysiological and computational hypothesis. *Journal of Cognitive Neuroscience*, 3 (2): 191-202.

Meltzoff, A.N. (1993) Molyneux's babies: Cross-modal perception, imitation and the mind of the preverbal infant. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 219-235). Oxford: Blackwell.

Miller, G.A. & Johnson-Laird, P.N. (1976) *Language and perception*. Cambridge, MA: Harvard University Press.

O'Keefe, J. (1993) Kant and the sea-horse: An essay in the neurophilosophy of space. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 43-64). Oxford: Blackwell.

O'Keefe (1994). This volume

O'Keefe, J. & Nadel, L. (1978) *The hippocampus as a cognitive map*. Oxford: Clarendon Press.

Paillard, J. (ed.) (1991) *Brain and Space*. Oxford: Oxford Science Publications.

Pederson, E. (1993) Geographic and manipulable space in two Tamil linguistic systems. In A.U. Frank & I. Campari (eds.), *Spatial information theory*. Berlin: Springer Verlag.

Pederson, E. (1994) Language as context, language as means: Spatial cognition and habitual language use. Working paper no. 26, Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, Nijmegen.

Piaget, J. & Inhelder, B. (1956) *The child's conception of space*. London: Routledge & Kegan Paul.

Pick, H.L. Jr. (1988) Perceptual aspects of spatial cognitive development. In J. Stiles-Davis, M. Kritchevsky & U. Bellugi (eds.), *Spatial cognition: Brain bases and development* (pp. 145-156). Hillsdale, NJ: Lawrence Erlbaum.

Pick, H.L. jr. (1993) Organization of spatial knowledge in children. In N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial representation: Problems in philosophy and psychology* (pp. 31-42). Oxford: Blackwell.

Pinker, S. (1989) *Learnability and cognition*. Cambridge, MA: MIT Press.

Rock, I. 1990. The frame of reference. In I. Rock (ed.), *The legacy of Soloman Asch* (pp. 243-68). Hillsdale, NJ: Lawrence Erlbaum,.

Rock, I. 1992. Comment on Asch & Witkin's "Studies in space orientation II", J. of Exp. Psych. General: 121(4):404-6.

Rock, I., Wheeler, D. & Tudor, L. (1989) *Can* we imagine how objects look from other viewpoints? *Cognitive Psychology*, 21, 185-210.

Senft, G. (1994) Spatial reference in Kilivila. The Tinkertoy matching games - A case study. *Language and linguistics in Melanesia*, 25, 98-99.

Shepard, R.N. & Metzler, J. (1971) Mental rotation of three-dimensional objects. *Science*, 171, 701-703.

Sorabji, R. (1988) *Matter, space and motion: Theories in antiquity and their sequel*. London: Duckworth.

Stein, J.F. (1992) The representation of egocentric space in the posterior parietal cortex. *Behavioral and Brain Sciences*, 15(4), 691-700.

Svorou, S. (1994) *The grammar of space*. Amsterdam: Benjamins.

Takano, Y. (1989) Perception of rotated forms: A theory of information types. *Cognitive Psychology*, 21, 1-59.

Talmy, L. (1983) How language structures space. In H. Pick & L. Acredolo (eds.), *Spatial orientation: Theory, research and application* (pp. 225-282). New York: Plenum Press.

Tarr, M. & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, 21, 233-282.

Taylor, H.A. & Tversky, B. (in press) Perspective in spatial descriptions.

Tolman, E.C. (1948) Cognitive maps in rats and men. *The Psychological Review*, 55(4), 189-208.

Tversky, B. (1991) Spatial mental models. *The Psychology of Learning and Motivation*, 27, 109-145.

Tversky, B. (this volume)

Tye, M. (1991) *The imagery debate. Representation and mind*. Cambridge, MA: MIT Press.

Valvo, A. (1971) Sight-restoration after long term blindness: the problems and behavior patterns of visual rehabilitation. New York.

Van Cleve, J. & Frederick, R.E. (eds.) (1991) *The philosophy of right and left: Incongruent counterparts and the nature of space*. Dordrecht: Kluwer.

Vandeloise, C. (1991) *Spatial prepositions: A case study from French*. Chicago: University of Chicago Press.

Wilkins, D. (1993) From part to person: Natural tendencies of semantic change and the search for cognates. Working paper no. 23, Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, Nijmegen.