Vision, shape, and linguistic description: Tzeltal body-part terminology and object description'

STEPHEN C. LEVINSON

Abstract

The Mesoamerican languages are renowned for shape discriminations. The Mayan language Tzeltal is no exception, and the theoretical implications are here explored for one area of the vocabulary that plays an important role in locative descriptions. Tzeltal body-part terms are "metaphorically" mapped onto parts of inanimate objects strictly according to a complex volumetric analysis of shape. This raises a number of fundamental issues:

a. In what sense is this a "metaphorical" process?

b. What is the relation of the volumetric analysis reflected in the language to the volumetric analysis involved in visual object recognition?

The answers given are the following:

a'. The mapping of body part terms to shapes is not done by any form of creative analogy, but by a precise geometrical algorithm. Contrary to assumptions in the Mesoamericanist and cognitive linguistics literature this then has few of the properties of metaphor.

b'. There is a coincidence between the kind of volumetric analysis involved in visual object recognition and that involved in the application of these terms.

The conclusions from (b') may be far-reaching. According to modularity arguments, linguistic processes should have no access to strictly visual processes. Although the present facts are not decisive, together with other observations they favor models where there is shared linguistic and visual access to the underlying processes of volumetric shape analysis.

The paper suggests that cross-linguistic data might play an important role in general speculations about the relation between different kinds of mental representation, and that the Mesoamerican languages might have a special pertinence to the relation between visual and linguistic representations.

Linguistics 32 (1994), 791-855

0024-3949/94/0032-0791

1. Preamble2

This paper starts from the following intuitive observation: there is something very curiously visual about Mayan languages like Tzeltal, that encode shape and position in verbal roots and body-part assignments, so that for example locative descriptions are replete with contour information.

Why visual? To answer that, we must think a little about the difference between different kinds of representation, visual and linguistic. And there are perhaps two levels of answer: an intuitive level, and a technical level.

On the intuitive level, let us imagine a scene of some objects (e.g. a coffee pot, cups, etc.) on a table; we could represent the scene linguistically, as an English description of what there is and how it is all arranged, or we could represent the scene visually, for example as a drawing. The two representations would have different properties:

a. The description but not the drawing might contain information that is not "in" the scene (e.g. that the coffee pot was given to you by your grandmother).

b. The drawing must be from some perspective (although it could be a plan from above). The verbal description may also have the perspective of the viewer ("the coffee pot is to the left of the cup") but it need not ("the coffee pot is in the middle of the table").

But perhaps the most striking difference would be (c):

c. whereas the drawing musT represent shape and relative sizes and distances, the English description need not, and perhaps even CANNOT. Coffee pots are of diverse shape and size, and so are cups, but there is no easy way in English to describe all that succinctly - but a sketch will do it in a trice.

The point is that English spatial descriptions (like The coffeepot is on the table) are sEmAN-rlcaLV GENERAL over shape, angle, and many other Euclidean aspects of the situations described, whereas a drawing CANNOT BE so (even the sketchiest adumbration of a shape suggests a specific shape). This can't be entirely an accidental feature of English: indeed, it is part of the design properties of language - a finite vocabulary implies semantic generality, and the possible shapes of coffee pots alone might soon exhaust the limits of the lexicon.

This has led some analysts to claim that spatial language is intrinsically topological - in effect, indifferent to most aspects of shape, especially coNToun.' But this does not follow from the design requirements of language alone: we can have a notion of triangle (distinctly untopological), and we can let this be semantically general over a precise range of shapes. Moreover, we may rely on our interlocutors to have some familiar unmarked expectation of a good triangle in mind (e.g. isosceles). (Some theorists would think of such a "good form" as a semantic prototype, but I think of this entirely as a pragmatic matter.)

In a nutshell: with regard to description, language is digital, sketches are analog. But this dichotomy can be eroded; if you have enough pixels, you can have analog shapes in digital form. Suppose we have over 20 distinct roots for describing, for example, different contours of vessels, depending on the relative size of the orifice, or of the base, or the sharpness of curvature of the belly or of the lip. Figure 1 shows how Tzeltal roots (here derived as predicate adjectives) provide a prodigal assortment of such fine-grained distinctions. Given these resources, we can verbally encode with a few roots the shape of our coffee pot with a precision and conciseness approximating that of the sketch. It is partly in this intuitive sense that Tzeltal descriptions are more visual, much more finely attuned to visual discriminations than the counterpart English descriptions.'

But there is also another relevant sense in which Tzeltal could be said to be "visual": shape distinctions are not just possible, they are often mandatory. Now a shape (or contour) is just what is neutralized in topology, and on the widespread view that natural-language spatial description is essentially topological, shape is also (it has been argued) neutralized in English prepositions.' Thus in may presuppose a container, but the expression is indifferent to whether the container is a bottle, a flat bowl, a box, or even a garden. In contrast, exact shape, curvature, and contour are crucial to Tzeltal locative (and indeed many other)

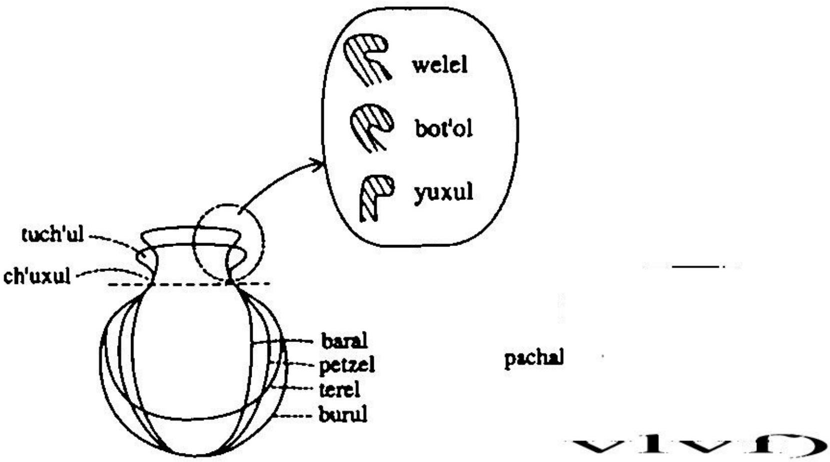


Figure I. Some Tzeltal predicates discriminating pot shapes descriptions. This is guaranteed by two grammatical structures used in the canonical locative description:[6] the first is that the locative predicate must be drawn from a set of many hundreds that put sortal conditions of shape on the subject (or *figure)* and sometimes also the landmark object (or *ground);* the second is that to specify areas on or near the ground, Tzeltal uses "metaphorical" body parts that are mapped onto the object largely on the basis of shape.

That is the direction in which one might give an intuitive affirmative to the question, is Tzeltal curiously VISUAL? But there is another direction. If one looks to the modern theory of vision, one finds there an analysis of visual process based on preoccupations with the shapes of objects, with object-centered points of view, with the partition of objects into parts, and so on, which seems to suggest that perhaps Tzeltal is (compared to English) much more closely connected with and tied into the cognitive stream of operations involved in visual understanding itself. To explain this, we must digress briefly to review modern theories of vision for those unacquainted with them.

2. **Modern theories** of vision

2.1. *Mary's theory*

The "problem of vision" is how the viewer extracts a representation of a stable three-dimensional world from a two-dimensional retinal array that is constantly fluctuating due to eye and head movements. This turns out to be anything but trivial: what we know at a glance about the analysis of a scene into a grouping of objects with size, depth, orientation, and so on cannot be even approximated by the most sophisticated computational analysis of images. A central focus of work in vision has thus come to be, How do we recognize objects, when these can be intrinsically variable in shape in the first place and viewed from any number of angles in the second, are rarely viewed under the same lighting, are typically partially occluded and shaded, with contours not always sharply distinct against other objects?

Theories of vision experienced a revolution in the late *1970s* akin to the Chomskyan revolution in linguistics in the *1960s.* Before that, setting aside the Gibsonian tradition, vision theories had been dominated by a pattern-matching paradigm on the one hand, and a patchwork quilt of unintegrated detailed physiological and psychological investigations on the other. The pattern-matching paradigm assumed that visual object recognition is essentially TOP DOWN, involving the matching of 2D representations:' one has a mental inventory of "snapshots" of objects and tries to match what one sees with what one has "in mind." Successively sophisticated computational models tried to match first gestalt wholes, then features of wholes, and then "structural descriptions" of organized features, but without much success outside very limited domains.

Then in the late *1970s,* David Man in collaboration with physiologists and computational theorists put together an integrative, generative theory of vision. This was resolutely BoTroM up, that is, it assumed that most of the processes involved in object recognition consisted in generalized algorithms directly applied to the visual array, without reference to "world knowledge." And in a tour-de-force of model building, Man integrated many physiological details, for example about the mathematical massaging of the visual signal between the retina and the cortex, into a series of algorithms that yield four distinct levels of visual processing, as follows:

1. "The raw primal sketch," the output of processes extracting, for example, edge segments from the retinal array.

2. "The full primal sketch," connection of edge segments, etc., into a 2D representation of the external contours of objects, similar to a line drawing.

3. "The 2}D sketch," using stereopsis, etc., the extraction of relative depths in a VIEWER CENTERED representation, but without full 3D interpretation of volumes.

4. "The 3D model description," the final segmentation of objects from the background and their interpretation as volumes from an OBJECT cENrsRED perspective. Objects are analyzed into successively finer volumes, each with a central axis along a "generalized cone." They can then be recognized by comparison with an inventory of known objects analyzed in the same style.

What the overall sequences of processes and stages of analysis achieve is passage from a large array of pixels capturing differing levels of light intensity, through an extraction of a line drawing of a scene, to a drawing with half-depth and with objects still embedded in their backgrounds, to the final stage: an extraction of the 3D shape of objects. It is this last stage, which is also the most speculative (and nowadays the most controversial), that is of special interest to us. It involves a switch from the viewer-centered 2}D sketch to an object-centered perspective, in which external spatial frameworks (like verticality, front/back, left/right dispositions) are irrelevant. This is a peculiar process: analysis of retinal images as 3D objects by imaginative location of the analyst in the inner volume of the object itself! For the problem is to reconstruct a volume from a chance view, to do which one must reconstruct the inner geometry of the object that must cast such a view. That inner geometry is of course constant whatever the orientation of the object, hence the irrelevance of external spatial frameworks for this process. Only after recovery of the internal volumetric structure of the object can we recognize the object as, for example, an upside-down bucket viewed with partial foreshortening.

Object recognition in Marr's system is "bottom-up" almost all the way (at least as far as it proves possible); that is, we do not recognize things by comparing 2D views against a mental inventory of such projections for each known objects Rather, without reference to encyclopedic knowledge of the world, we try to reconstruct the entire internal 3D geometry of the object, and then compare that against a mental inventory of 3D objects, just to get the "name" or category. The very nature of the geometrical analysis of single objects helps us to find them in our mental catalogue of object kinds. For example, a human figure is first analyzed in terms of a main axis and its cylindrical volume, then broken down into successively finer subaxes, each with their cylindrical volumes, all the way down to the fingers if visible (Marr 1982: 306). When viewed from a distance or in poor lighting, one will only get the coarser levels of analysis. Now to recognize that shape (i.e. to equate it with one of the types in our mental inventory), at whatever level of discrimination we have managed to discern, we compare it to a hierarchical index of cylindrical shapes in our mental catalogue (Mary 1982: 319). The more detail we can see, the more sure we can be about the subclassification, for example of a human as a female child.' For Man, object recognition is the end-point of the visual process, delivering the information we need about objects into central thought processes, although we also clearly need that information reembedded in a viewer-centered coordinate system so that I can, for example, walk around the table.

Man's theory has set the terms of reference for the theory of vision up to the present time. Nonetheless, of course, there have been many fundamental revisions and many new ideas. We briefly list some new ideas about object recognition (the 3D model part of the theory) as background for what follows.

1. *Shape primitives: the generalized cone.* A generalized cone is a shape generated by sweeping a cross-section of constant shape, but potentially varying size, along its main axis. A startling property of generalized cones is given by *Marr's theorem* (1982: 223): you can get from a silhouette (contour) of an object X to a 3D model only if X is composed of generalized cones (and there is no foreshortening). Man therefore thought that generalized cones are crucial to object recognition, since we

can and do routinely interpret silhouettes, and this interpretation would be impossible on the assumption of some other class of primitive shapes. Marr (1982: 309ff.) went on to suggest generalizations to 2D objects and hollow ones (see also many useful suggestions in Jackendoff and Landau 1992 [1991]).

The linguist observer of the theory of vision should note that the restriction to generalized cones is like restricting the acoustic properties of speech so that a comprehender knows the range of sound classes that could constitute good phonemes. Recent suggestions involve taking the analogy further and suggesting there might just be a fixed inventory of phonemes, that is, a subset of generalized cones. Thus for example the inventory of generalized cones is reduced to 36 primitive ("geons") in Biederman's (1987) account. Or one might assume a search for "semantically" potent shapes, ones that encapsulate causal history (Leyton 1989), and this proves for us a useful auxiliary idea.

2. *Part segmentation.* Marr's account is a little unclear: first, he illustrates (1982: 314-315) an algorithm for segmentation by outline that could proceed independently of axis assignment; later (1982: 317) he shows that axis assignment might involve volumetric analysis. Commentators have explored the former path, trying to decouple segmentation from volumetric analysis (Hoffman and Richards 1984). For the purposes of this essay, Marr's idea that volume and axis might be "chicken and egg" is probably the more fruitful.

3. *Object-centered orientation vs. absolute (e.g. vertical) vs. egocentric orientation.* Marr's claim is that an object's coordinate system is set up PRIOR to object recognition; however, there is some current debate here. Marr allowed that where the principal axis is vertical, the direction would by default be "up" (1982: 310). Leyton (1989) suggests that in natural or flexible objects, shape analysis gives growth and pressure information; since gravity is a constant force in our world, it may be packaged into any implicit analysis of shape.

Designing algorithms that derive 3D models from 2½D or 2D images has in fact proved rather intractable. Partly for this reason, attention has recently turned to the possibility that we store something less than 3D models of objects. The 3D model theory also predicts that we should be equally good at recognizing objects from odd angles, and this proves not so. Could we then really operate with 2D representations, either one or many, and some method of mental rotation, or is a notion of canonical orientation built into 3D representations? We must here leave these current debates to one side (but see Bülthoff 1991; Gibson and Peterson i.p.; and the conclusions to this paper), although they are very pertinent to the central questions about the relation of language to vision since they concern the nature of the final output of the purely visual process.

2.2. *Vision and language*

The theory of vision seems at first sight far removed from linguistic description. Obviously, there is a close connection in the conuao of language in the visual medium, whether in sign language or writing. But equally, in the case of spatial description, there is potentially the very same corrr wr as in visual encoding, namely the description of what I see before me. Here, judging from standard works on the linguistics of spatial description, the level at which the visual processing and the linguistic categorization coincide would be an enriched OBSERVEa-cENTERED perspective view of the visual field."

Only an observer-centered perspective can account for the kinds of notions encoded in the English projective prepositions: *behind the tree* means 'occluded by the tree from the observer's perspective', *at the front of the tree* means 'on the line of sight between the observer and the tree', *to the left of the tree* means '(with the tree in the middle of the observer's visual field) to the left side of the visual field from the observer's point of view', and so on. Many non-Indo-European languages (e.g. the Northwest Coast Amerindian ones) make distinctions in their demonstratives between 'visible' and 'invisible' from the speaker's point of view. Spatial deixis in general is organized around the observer's point of view: to *come here* involves motion to the point of speaking/observing. On the basis of a broad range of evidence of this sort one might suppose that language taps into a final, composite stage of visual processing where objects (now analyzed and recognized by preprocessing in terms of their inner coordinates) are reassembled in a perspective view, oriented with respect to gravity and landscape, with depths calculated from the perspective point of the observer.

It would be in line with a thorough-going modularity of the kind advocated by Fodor (1983) if language had no access to earlier levels of visual processing. But there are tantalizing bits of linguistic evidence that this is not so. For example, Casad and Langacker (1985) report of Cora that different prepositionlike constructions are involved in describing the tail of a dog sticking out so that it is orthogonal to the line of view, vs. aligned with the point of view. This looks like an opposition in shape from the point of view of a two-dimensional visual array. There is also evidence from Tzeltal and other classifier systems that in a number of cases classifiers are applied on the basis of the 2D shape projected to the observer, and not the 3D shape of the object from an object-based coordinate system. For some further examples of such phenomena see section 5.4 below.

To this sort of anecdotal fact, I would now like to add the facts about Tzeltal body-part terminology, as generatively used to segment objects. These would seem to argue forcefully for a direct connection between language on the one hand and on the other the visual processes involved in object recognition BEFORE objects are integrated into the perspectival scene from the viewpoint of the observer." Now one might think that any object naming would be based on such processes - and certainly it must be based on the OUTPUT of such processes (i.e. object recognition) - but in this case we can argue that there is linguistic access to the very processes themselves (i.e. the analysis of the internal volumetric geometry of objects).

This observation is hardly going to shake the foundations of modularity theories, because Marr already allowed that there might be "top-down" feedback in 3D object recognition (for example in dim light, if I think I see a man in the bushes, I may use that information to try to further resolve the visual image). But it does have an important bearing on an architecture for the interaction between modules. Jackendoff and Landau (1992 [1991]: 121) for example argue that vision, touch, language, and other "input" systems all deliver to a central spatial representation system that "is relatively rich in its possibilities for describing object shape; but it is relatively limited in the way it can use object shape to encode spatial relations." They point to the neurological evidence that there are distinct neural pathways for object information (the "what" system) and location information (the "where" system), with the "where" system weak on shape conceptualization. And they hypothesize that this neurological specialization is directly reflected in language, so that many shape distinctions can be made in object names, but not in spatial relators like prepositions. The present observations, if correct, would seem to require at least some modification of the theory. First, the Tzeltal body-part data argue that there is linguistic access not only to the output of the visual process of object recognition, but also to the internal volumetric analysis upon which such recognition depends. Second, the Tzeltal body-part system would seem to show that a rather richer geometric descriptive system can be utilized to specify "where" information, that is, to describe the ground configuration, than Jackendoff and Landau (1992) originally supposed.[12]

3. Body-part system in Tzeltal

3.1. *Locatives in Tzeltal*[3]

Tzeltal locatives have the peculiarity, from the Indo-European perspective, *of* giving a great deal *of* information about the figure (the object to be found) and relatively little about the ground (where to find it)." In English the phrase *opposite the supermarket* adumbrates a scene, since *opposite* presupposes an orthogonal across a separating strip. But Tzeltal has only one vacuous preposition ta, while enforcing a choice between a myriad *of* stative locative predicates with detailed selection restrictions describing the shape and orientation *of* the figure. Thus a minimal locative expression might be, for example,

(1)	pachal	to mexa	bojch
	sitting-of-wide-container	at table	gourd-bowl
	STATIVE ADJECTIVE	PREP PHRASE	SUBJECT
	relation	ground	figure
	'the bowl is sitting on the table'		

If it is desired to be more precise about where in relation to the ground object the figure is, then the ground can be segmented into its parts, labelled by animate body parts on an analogical basis. Then the figure can be said to be in contiguity with that part *of* the ground. Thus:

(2)	waxal	to x-chikin mexa	to p'ine
	standing-of-vertical-cylinder	AT its-ear table	the pot
	STATIVE ADJECTIVE	PREP PHRASE	SUBJECT
	relation	ground	figure
	'the pot is standing at the corner *of* the table'		

It is understanding this application *of body partonymy* that is the focus *of* this paper. However, some *of* the lessons learned from the study *of* the body-part terms may have equal application to the study *of* the positional and other roots that are the base for the stative adjective stems illustrated in the two examples immediately above. The reason is that both body-part terms and the roots *of* such stative predicates encode sensitive specifications *of* shape. The principles underlying these shape specifications are in many cases the same or highly similar, and there are then interesting *sets of* implicational relations between the two sets *of* terms." For example, body-part terms are (I shall claim) essentially oriented to the internal structure *of* objects, and so are most stative predicate roots; however, unlike the body-part terms, the predicate roots often encode in addition an orientation to a wider frame, typically the vertical dimension given by gravity. Thus in the first example, the predicate *pachal,* sortally restricted to hollow objects whose orifice is as wide as the greatest width and whose height does not exceed that width, asserts that the subject (here referring to a bowl) is vertically arranged so that its base (or 'butt') is downward; the contrastive predicate *nujul* asserts that the base ('butt') of the bowl is upward. The predicate *waxal* is contrastive to both, since it describes the vertical orientation of a hollow cylindrical object, whose height is greater than its width, standing on its 'butt'. In this sort of way, the specification of the lexical content of the predicate roots makes essential reference to the body parts of the presupposed class of subjects, and because the application of the body-part terms themselves depends on shape, partly for this reason shape restrictions are imported into the predicates concerned. However, the shape restrictions encoded in body-part terms are perhaps of a coarser grain than those found in the stative predicate roots, so they make a good starting point for the study of shape constraints in the Tzeltal lexicon.

But before proceeding, it is essential to set the locative resources embodied in the body-part terms in the context of a wider set of resources for describing the ground, the region where the figure object is located. The body-part terms have expressions that are close syntactic and functional kin, the relational nouns. For example, *to s-jol mexa* 'at its-head table, i.e. at the head of the table', is clearly closely related to the construction *to s-ba mexa* 'at its-top table, i.e. on top of the table'. But the former construction is a body-part construction, the latter a relational-noun construction." What's the difference?

The answer is that body-part terms belong to a special form class, with distinctive morphology and its own semantic properties, which includes of course the names for human and animal body parts. When these are possessed, in the construction POSSESSIVE-NP I + NP2 (e.g. *s jol* mexa 'its-head table'), NPI is understood to be a physical part of NP2 (for example that part of the table that can be called its 'head'). Now in contrast the relational nouns do not primarily name parts of objects, they name something more abstract: for example a region projected underneath or above an object, a midpoint on a line drawn between two objects (or along the main axis of one object), and so on. In the construction POSSESSIVE-NPI +NP2 (e.g. *s-ba mexa* 'its-top table'), ba is not construed as a part of the table (as perhaps the English gloss may suggest); if the table is upside down, the *sba* is the present upper surface. Or even more clearly in the relational noun construction *y-ajk'ol te'* 'its-upness the tree, i.e. uphill of the tree', the 'uphillness' is not a part, not even a property, of the tree. In short, the possessive construction has a different construal in the case of body-part terms and relational nouns.

The full set of locative relational nouns is given in Table 1 derived from Brown (1991). She notes that *s-ba, s-tz'eel,* and *y-ejtal* can designate parts, but I think it is clear that they do not name intrinsic parts in the same sense as body parts. (I do not think that having cut off the top or side or edge of a table, you could now designate them with the erstwhile relational nouns, although you could do so with the corresponding body parts.)

There is a further striking semantic difference between body parts and relational nouns, which this paper attempts to establish: body-part terms invoke object-internal coordinate systems, relational nouns invoke external spatial frameworks, especially absolute coordinate systems (e.g. verticality, or 'uphill' [south] and `downhill' [north]).

3.2. *The body-part terms*

3.2.1. Isolating the set: morphological and semantic criteria for body-part terms. What constitutes a body-part term in Tzeltal? The question is not idle. First, there are many terms for human body parts that have primary reference to the parts of vegetable bodies: thus one talks of the 'trunk of the arm' *(s-te'el k'ab),* 'whiskers' are referred to by the term for `roots' *(isim),* the flap of the ear is its `leaf' *(y-abenal x-chikin),* even the word for eye *(sit)* may have primary reference to 'seed', and so on. Second, there are a number of terms (the "relational nouns") that behave functionally and syntactically just like body-part terms but do not literally name segments of bodies. Moreover, like body-part terms they play an essential role in locative expressions. If we are interested in locatives, perhaps the bodily origin of terms is irrelevant.

But in fact there are special properties of real body-part terms that prove to be crucial. First there is the question of noun class. To this day, no one has done the lexicographical work required to sort out Tzeltal

Table I. ***Locative relational nouns***

s-ba	**'its top surface or edge'**
y-ajk'ol	**'its upness, its uphill region'**
y-alan	**'its downness, its downhill/underneath region'**
y-anil	**'its underneath'**
y-uti'	**'its inside'**
y-ejtal	**'its bottom surface or edge'**
s-tojol	**'straight ahead of it, on X's sightline'**
y-olil	**'its midline, middle'**
s-glee!	**'its side, horizontal edge'**

nominal classes." But crudely, we can identify a class, let us call it class I, of nominals that may be said to be "inalienably possessed" (Stross 1976). These have the property that they normally occur with the possessive prefix (identical in Tzeltal to the ergative verbal prefixes), but where they do occur unpossessed they must have a -VI suffix (which never occurs in the possessed form)," as illustrated in (3):

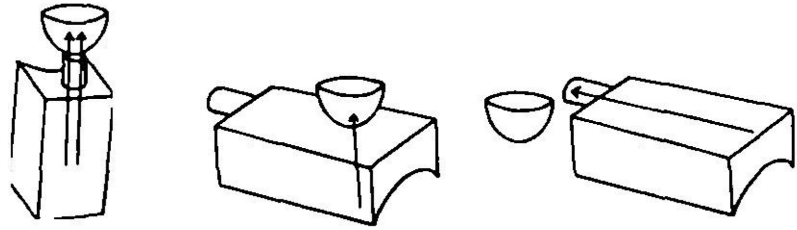
(3) *Nominal class I: inalienably possessed:*
possessor+root *s jol* 'his head'
root+VI *jol-ol* 'head'
Members of the class: nearly all body parts, body products, kinds of soul, kin terms, clothes, etc."
(N.B. some body parts don't belong to this class: *bak* 'bone', *ch'ich* 'blood' take -VI only when possessed: *e.g. s-bake!* 'his bone')

In example (3), the citation form is *jol-ol,* but since this rarely occurs, we will henceforth identify body-part terms by their possessed forms, presuming a third-person possessor (the prefix is s- before consonants, y- before vowels)." This leads to some translational absurdities, as I shall sometimes treat the possessed forms as if they were citation forms, but this will prove harmless enough.

There are a number of other nominal classes that contrast with this class, for example a class whose members can only appear possessed while simultaneously bearing the -VI suffix, or a class that never bears -VI suffixes, etc.

Being a member of the morphological class I illustrated in *(3) will* prove to be a necessary, but not sufficient, condition for belonging to the semantic class (describing, when unpossessed, the internal geometry of objects) with wide locative uses that we are interested in here.

There are a number of other obligatorily possessed nominals that play a crucial role in Tzeltal locatives that we specifically exclude here. They do not meet the morphological condition just described. Nor, on close examination, do they share the same kind of semantics. A good example is the relational noun *s-ba;* it lacks the corresponding unpossessed form **ba-il* or **ba-al* necessary for inclusion in class 121 It does not directly denote a body part, although *sba k'ab* 'tops of the hand/arm' may be used to describe the fingers." The semantic contrast comes out in the locative uses, where *s-ba* means roughly 'on the top of' or 'above', contrasting for example with *s jol* 'on the head of' (as illustrated in Figure *2).* In many cases these will be extensionally equivalent, for example when we are talking about something (e.g. a nail) fixed on top of a vertical pole. But when the pole falls over, the nail is still *to s-jol* 'at its head' but not *to s-ba* 'at its top'. The vertical dimension, and indeed



to s-ba "at / on its top"
to s-jd "at / on Its head'

to s-ba
('ta syol)

to s-lol
('ta s-ba)

Figure 2. *The vertical: relevant to positional nouns, irrelevant to body parts*

orientation generally, is not normally relevant to the assignment of body parts, but it or another absolute coordinate is mostly crucial to the assignment of locative relational nouns like *s-ba* (or the other terms in Table 1). This abstraction away from vertical orientation in the body-part system makes an important contrast with English and the theory of perception based on English usage.[23]

3.2.2. The body part terms. Tzeltal offers fine-grained segmentation of the anatomies of animate entities. The outline in *(4)* gives some idea of the numbers of terms involved and refers the reader to sources where more or less exhaustive lists may be obtained.

(4) Numbers of body-part terms

(i) *human body-part terms: c.* 80 primary terms (Stross 1976), plus many compound terms with "metaphorical" base; for example `nose of breast'=nipple; 'head of leg/foot'=knee; 'neck of leg/foot'=ankle

(ii) *animal body part terms: c.* 57 primary animal/human terms, only 10 restricted to animals (horns, feathers, etc.) (Hunn 1977)

(iii) *plant parts:* 116 terms, of which *21* occur also as human part terms (Berlin et al. 1974)

The diagrams in Figures *3* and 4 show a subset of commonly utilized body-part terms for a human and a bovine body respectively, simply to give an impression of the density of terms in everyday use. The terms on the right-hand side of each figure are the major ones generalized to parts of inanimate objects.

Now these three domains, humans, animals, and plants, provide the sources for the terms applied in the inanimate domain, to describe the

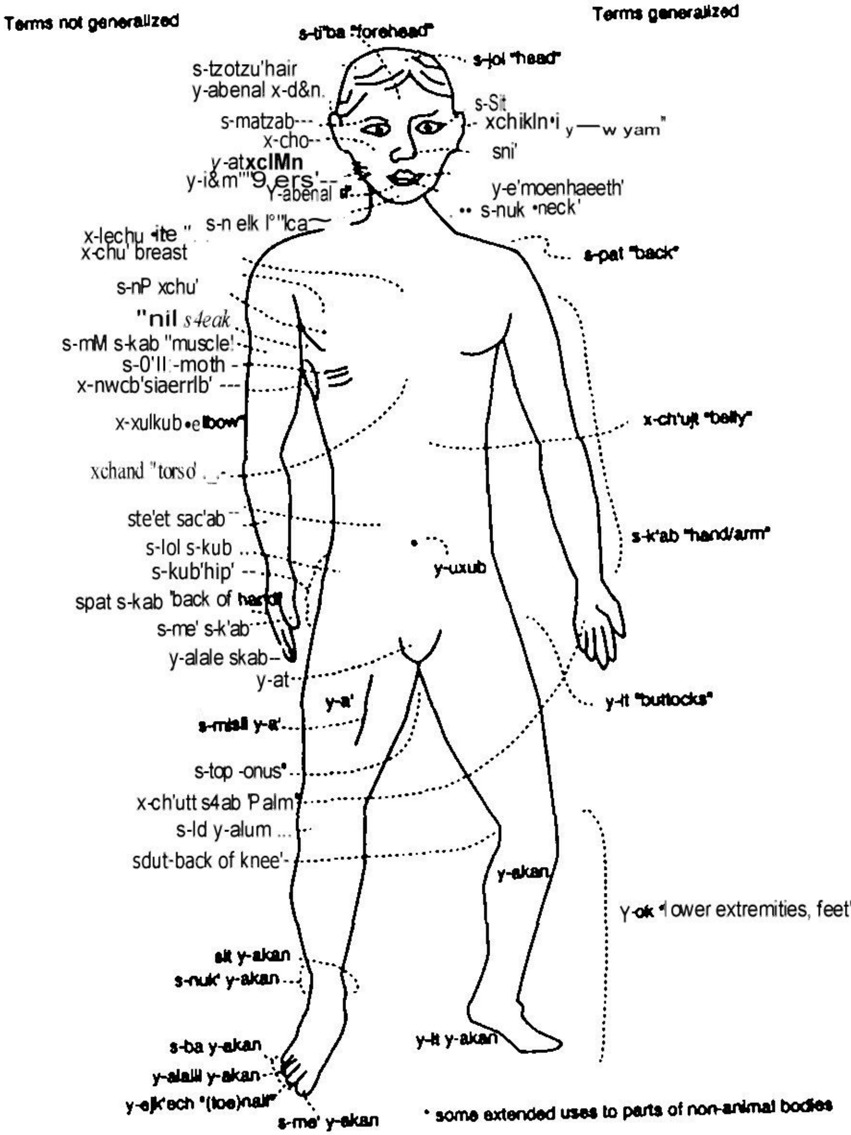


Figure 3. *Major* external *human body-part terms*

parts of nonliving entities. It is this application to inanimate objects that is the focus of this paper. The number of such body-part terms applicable to inanimate objects is very much less, of the order of 20.[24] Thus we have a situation like that expressed in (5), which expresses a

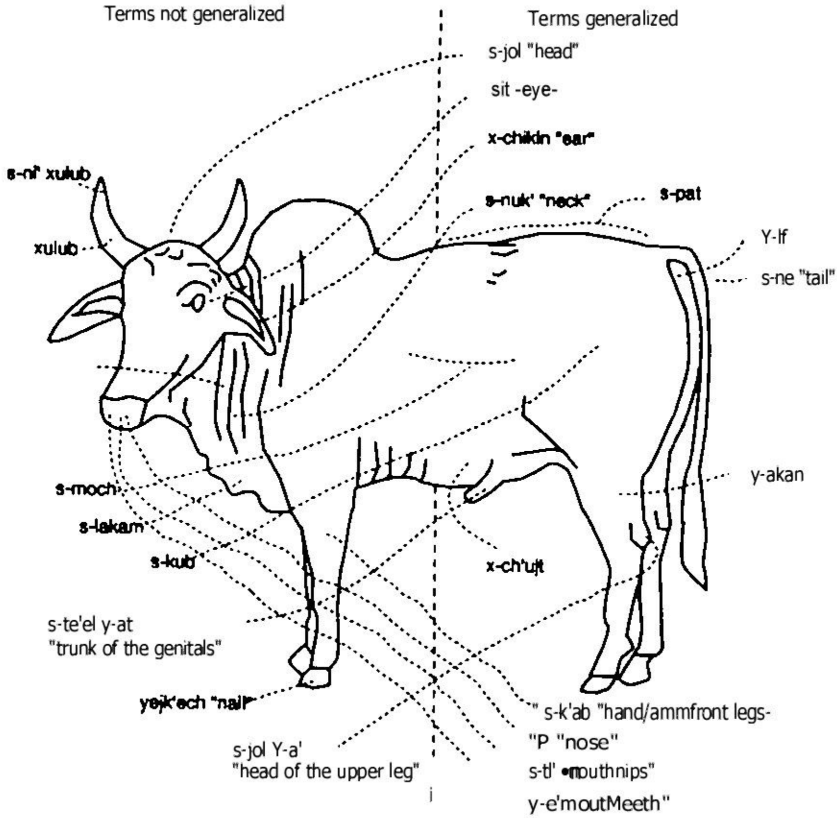


Figure 4. *Major external animal body-part* terms

familiar hierarchy, corresponding perhaps to an implicit four-category ontology:

(5)	*Domain*	*Primary terms*	*Terms transferred to inanimates*
I	human	c. 80	16 terms from either I or II
II	animal	c. 60	1 term clearly from II
III	plant	c. 80	3?
IV	inanimate	c. 20	0?

The 20 or so terms frequently transferred to inanimate objects are given in Table 2. Only 14 of these are central, in the sense that they clearly meet the morphological and semantic criteria and in addition have frequent application to inanimate objects; three more are reasonably frequently used but do not clearly have central reference to animate body

Table 2. *Terms frequently transferred to* inanimate *objects*

Root+VL		Source category	Possessed form	
a. The core set:				
jol-ol	'head'	1,11	to s-jol	'at its head'
pat-il	'back'	1, II, III	to s-pat	'at its back'
ch'ujt-il	*'belly'*	*I, II*	*to xch'ujt*	'at its belly'
akan-il	'foot'	1, II, III	to y-akan	'at its foot'
k'ab-al	'arm'	1, II, III	to s-k'ab	'at its arm'
it-it	'rump'	I, II, III	to y-it	'at its buttocks'
ni'-il	'nose'	1, II, III	to s-ni'	'at its nose'
elaw-il	'face	1, II	to y-elaw	'at its face'
ti'-il	'mouth'	I, II	to s-ti'(il)	'at its mouth'
chikin-il	*'ear'*	*I, II*	*to x-chikin*	'at its ear/comer'
ne-il(?)	'tail'	if	to s-ne	'at its tail'
nuk'-il	*'neck'*	*I, II*	*to s-nuk'*	'at its neck'
ej-il	'teeth'	I, it	to yej	'at its teetht[6]
sit-il	*'eyes'*	I, II	to s-sit	'at its eye"[27]
b. Some cases of unclear kind and origin:				
ok-il(?)	'footing'	I or III?	to y-ok	'at its base[128]
xujk	'flank'	IV?	to x-xujk	'at its side[129]
uxub-il(?)	'navel'	I or IV?..	to y-uxub	'at its navel'[3e]
c. Some cases with marginal extensions to inanimates:				
ak'-ul	'vine'	III	to y-ak'ul	'at its vine[131]
te'-el	'tree'	III	to s-te'el	'at its trunk'[s2]
isim	'whiskers'	I	to y-isim	'at its roots'
(or	'root'	III	to y-isim	'at its whiskers'??)
bak	'bone'	I, II	to s-bakil	'at its hard-core'J[8]

parts and may derive from other sources - of these only *xujk* 'side' is functionally important. The four last examples in (c) don't really belong at all, because they fail the morphological test and have very specialized uses: they modify other descriptions *(s-te'-el* 'trunk'), or describe a texture as much as a part *(y-isim* 'whiskers'), or have to do with internal structure *(s-bak-il)*. They are included here to show that there are a number of additional body-part terms that may occasionally be used in the inanimate domain but play a minimal role in the locative system.

3.3. *Body-part metaphors??*

It seems natural to think of the relation between the animate and the inanimate domains as one of metaphorical transfer. I think that in certain respects, to be made clear presently, this is a thoroughly misleading conception. But first, let us explore its intuitive appeal, which is considerable.

On this analysis there are three source domains, namely (I) human body parts, (II) animal body parts, and (III) plant parts; and one target domain: (IV) inanimate physical objects, corresponding perhaps to a fundamental set of ontological distinctions: (I) humans, (II) beasts (quadrupeds, also feathered bipeds), (III) plants, and (IV) residual set of physical objects.

The "metaphorical" application of the body-part terms to inanimates has none of the flavor of nonce metaphor: these terms just provide, in each case, the correct term for the part of the object in question, often without alternative. How then can we be sure that the metaphorical transfer is not the other way around, from inanimate to animate? For example, given the term *y-akan, which* has equal reference to the handles of implements and to the legs of animals, which is prior?

Clearly, one should appeal to the theory of metaphor, such as it is. Metaphors are typically (a) from more concrete to more abstract domains *(the heart of the theory),* (b) from better understood, better articulated, to less clearly articulated or less precise domains *(a column of smoke).* On those grounds, one expects the linguistic expression in question when applied to the SOURCE domain to have more specific and precise meaning; when applied to the TARGur domain to have more general, more schematic meaning. Here then are some putative tests, to be taken as jointly indicative of which domain is source and which target:

Criteria for identifying source vs. target domains

(i) more exact application: for example *sit* 'eye of animal' vs. 'any marked depression on for example fruit or stick'
(ii) restricted number of identical parts: for example *sjol* 'head', only one for humans, but two possible for boxes, tables, boards, etc. Or, *x-chikin* 'ears': only two for animates, up to four or more for inanimates
(iii) source combines two or more schema, targets choose subset: for example *s-ti'* 'mouth': (i) orifice, (ii) edge or outline of 2D plane, (iii) 3D ring or band (cf. lips), (iv) closure or 'stopper' of orifice:'
cf. 'mouth of fire'=circle of hearth
'mouth of house'=door (orifice, closure)
'mouth of pot'=lid, or lip, or orifice

These criteria would help to establish that the mappings in question are FROM animate domains I, II, and III TO IV, the inanimate objects. But they do not in general help us to decide between the source domains, and especially between I and II. In general, it seems that all three domains may be in play as sources of the relevant "metaphors," providing competing *schema* (sets of configurational and shape ideas) for the interpretation of target shapes and parts of inanimate objects.

The reader may get some intuitive grasp of the alleged mappings by looking at Figures 5, 6, and 7. Figure 5 shows how `nose' may be associated via the human model with a small protuberance, but via the animal model with the idea of the `leading point'. Figure 6 shows how the notion of `mouth', taken equally from human or animal models, encapsulates the ideas of not only orifice, but surrounding edge, ring of material, and 'stopper', the entity closing the orifice. Figure 7 shows how it may be essential to have three distinct source schema, one for each of the animate domains, in order to explain the very different mappings of

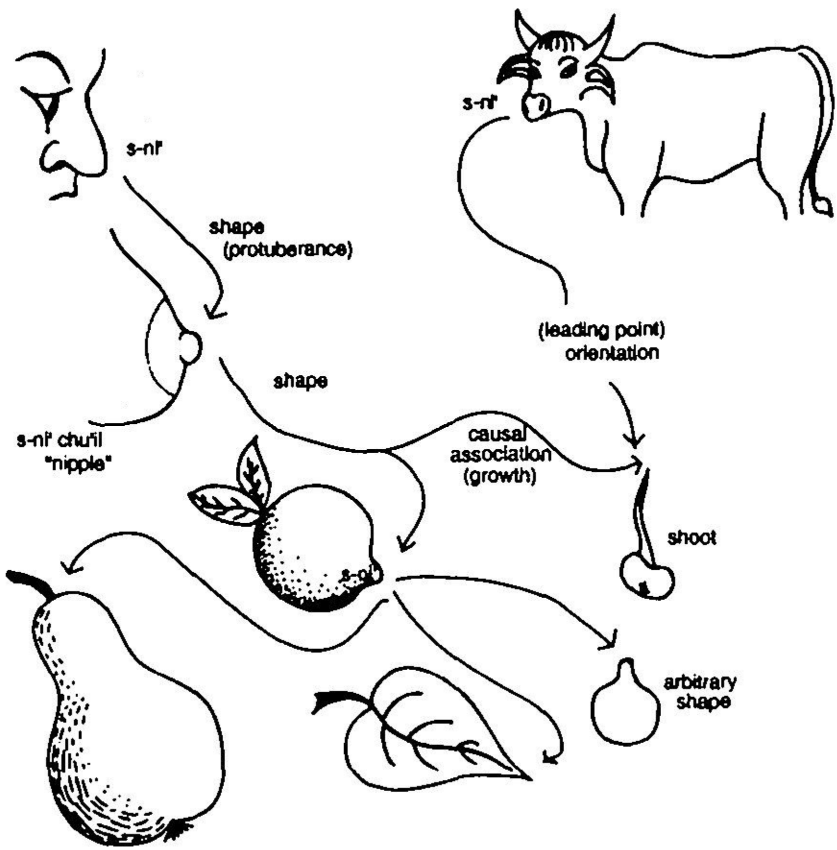


Figure 5. Putative metaphorical extensions of 'nose'

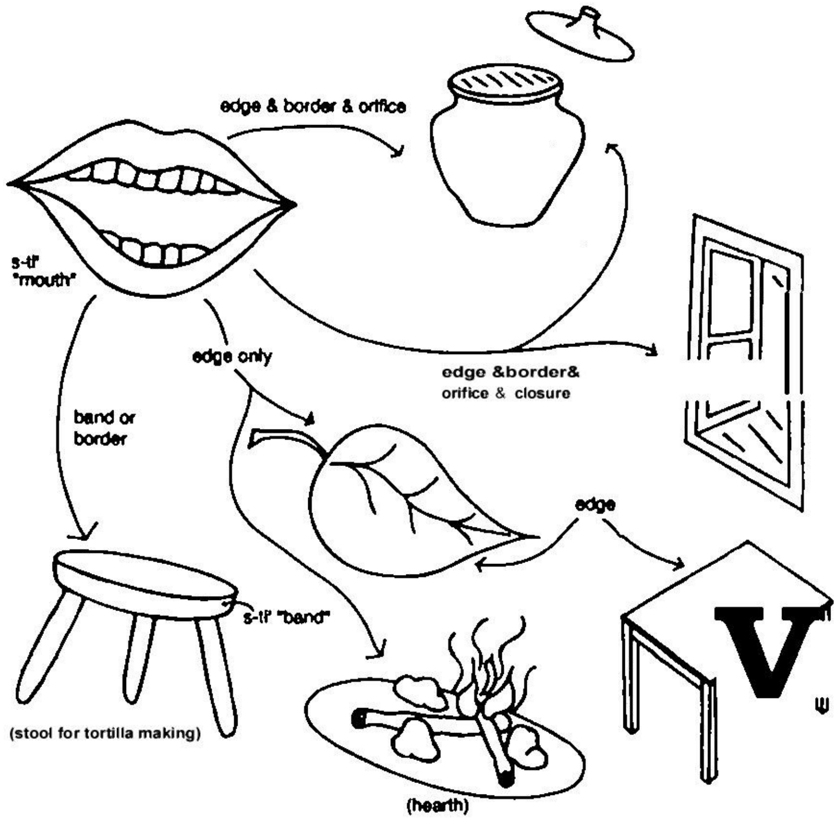


Figure 6. Putative metaphorical extensions of 'mouth'

a term (here `back') to the inanimate domain. One can detect the operative source domain for the inanimate application of the term by seeing whether `back' is interpreted as vertical (human source), horizontal (animal source) or spherical (plant source, where `back' means `skin, bark, peel'). Thus:

s-pat (I) `back (of human)': the far VERTICAL side of object (the object is like a confronting human)
spat (II) `back (of animal)': top horizontal facet
spat (III) `skin of fruit, bark of tree': 2D encircling surface

Although there is not much written about highland Mayan body-part terms (but see e.g. de Leon 1992a, 1992b), this is the intuitive analysis one suspects students of these languages have had in mind. Certainly,

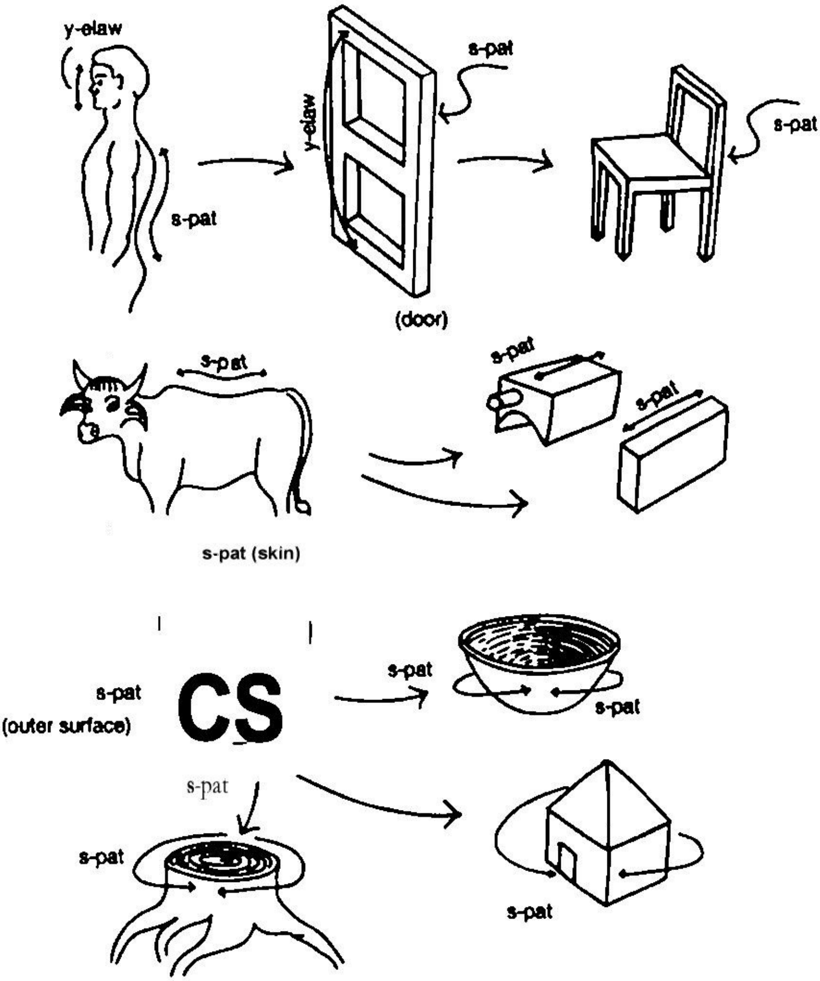


Figure 7. Putative metaphorical extensions of 'back' via three source domains

students of other Mesoamerican languages with similar body-part systems have advanced metaphorical analyses, most recently in the framework of "cognitive semantics" (Brugman 1983; Lakoff 1987: 313ff.; MacLaury 1989)." MacLaury (1989) in particular compares the body-part systems of three Mesoamerican Otomanguean languages, noting that although metaphorical applications differ in detail, the general metaphorical genre is "diagnostic of the Mesoamerican diffusion area." [35] The process at first sight seems natural enough, and if the ideas of metaphorical extension are rather vague, that's in the nature of metaphor: as Lakoff (1987: 314) opines, "we can understand the Mixtec system because we too have the capacity for metaphorical projection of this sort, even though our conceptual system is not conventionally organized in this way. Systems like this are neither rare nor obscure."

But is the metaphorical analysis right? I have some serious doubts. First, if there is an ongoing "fresh" metaphorical process at work, why are the terms that get extended tusT THESE and no others? Why is *y-akan* `foot, lower leg' extended to inanimates and not y-a' `upper leg'? Why *x-ch'ujt* `belly' and not *s-moth* `flank'? Why *x-chikin* `ear' and not *s-nejkel* `shoulder'? If the answer to that is, these are "dead metaphors," then how can one account for the rapid and natural extension to brand new, unfamiliar objects, or familiar objects that come in never-ending varieties of shapes (like stones)? This extension to novel objects moreover is not now like "fresh" metaphor, where one speaker may invent a novel application based on shape, which may be more or less apt. For although there are often different possible rival schemes, informants are quite clear about what is not a possible scheme, even where free metaphorical transfer would suggest its adequacy. One of these limitations is that free metaphorical transfer would naturally involve some kind of Gestalt of the source object, complete with details of its normal position and behavior; thus the process should naturally include notions like *canonical orientation, normal direction of motion, dangerous end,* etc. But these seem not to be involved.

In short, the problem is that the application of body-part terms is a generative process, yet the application of transferred terms is strictly controlled both in the terms that may be borrowed and in the way they may be applied. The controls do not look metaphorical in nature, that is, applied by some loose analogy. They look algorithmic. I shall now try to demonstrate this for the core of the system. In doing so, we shall leave behind the notion of metaphor entirely. Further, we shall abandon all "top-down," knowledge-based processes altogether.

4. How to compute the application of Tzeltal body-part terms

The theory about to be advanced maintains that the core of the Tzeltal body-part terminology is quite strictly an oamcT CENTERED sYSTEM: the terms are applied as if the objects were novel entities encountered in an orientationless void, for example, weightless in outer space! Moreover, the process is "bottom-up," invoking no world knowledge, and thus excluding comparison (metaphorical or otherwise) to other entities.

Rather, the terms are applied on the basis of the internal geometry of the object itself. Only at the margins of the system, when the internal geometry of objects leaves arbitrary decisions open, is an external reference frame or functional knowledge secondarily involved."

In fact all that is needed is the very system of object analysis that Marr and followers have argued is independently required by the visual system for the recognition of objects. Thus the analysis can follow closely the system of volumetric analysis posited in the theory of vision.

In order to apply the terms, a computation of the inner geometrical properties of the object must proceed in a number of distinct stages. These are the following:

i. object segmentation;

ii. finding the main or "model axis," and mapping generalized cone structures along this generating axis;

iii. determining the directedness of the model axis;

iv. applying terms to the two ends of the model axis;

v. locating secondary projections from the model axis, analyzing these secondary axes for angle in relation to the main axis, and finding their associated volumes;

vi. naming these projections on the basis of shape;

vii. naming surface features or protrusions.

We shall take these in turn.

4.1. *Object segmentation*

Object recognition requires analysis of wholes as assemblages of parts - partly because one never sees the whole from all aspects at one time, partly because we may need to guess the existence of an object from seeing a part, partly because nonrigid objects cannot be recognized without understanding the potential movements of the parts. Needless to say, the naming of parts presupposes the same process.

In the theory of vision, it has been shown that for a large range of cases the processed retinal image of an object can be computationally segmented on the basis of (a) the assumption that it is constructed of generalized cones, and (b) some simple heuristics (Mary 1982: 314). The heuristics basically work by looking for regions of sharp concavity along the outline of the object and then segmenting across on the shortest route to another concavity or at least flat surface. Thus we can segment the parts of a donkey on the basis of its silhouette (Marr 1982: 315). Recently, a considerable amount of work has gone into the mathematical basis for this kind of algorithm. It has also been shown that the algorithms correctly predict perceptual reversals; for example in figure-ground reversals (of the Necker Cube, or better, the Rubin vase-face, type) what was concave becomes convex, and therefore we see different parts, as predicted by the theory (see for example Hoffman and Richards 1984).

A similar process must lie behind the segmentation of objects presupposed by the Tzeltal body-part system. We gave line drawings of human and animal bodies to informants and asked them to draw lines at the division of the main parts of the bodies. The divisions were of the kind shown in Figures 8 and 9.[37] Despite the fact that Tzeltal body parts are rather differently segmented than English ones (e.g. there is no distinction between hand and arm), most of these divisions would be predicted by the "split at sharp concavities" rule.[38] With respect to Figure 8, the human figure is thus split across the top of the legs, from armpit to shoulder, across the neck, and at a more sensitive level of discrimination, across the knees. With respect to the bull in Figure 9, the division into tail, forelegs, neck, head, etc., follows the same rule. In both figures, there are some internal divisions that could not be predicted from these silhouettes, for example flank and belly for both man and bull, although in the case of the man, belly would be recoverable as a concave or convex shape in a side view." However, it is the segmentation of inanimate objects that we are particularly interested in here, and we have as yet administered no comparable task in that domain.

4.2. *The primary or "model axis"*

4.2.1. *Finding the main or "model axis."* If we follow Marr's original ideas, the model axis - or main internal coordinate of an object - is found *by* attempting to map a generalized cone onto the segments output *by* the segmentation process. As mentioned, generalized cones are shapes generated *by* constant cross-sections of potentially varying size drawn along the generating axis (e.g. a vase shape), and they have the interesting property of being 3D shapes recoverable from a 2D silhouette.

The top row (a) of Figure 10 shows three simple shapes that might be found among household objects in Tenejapa (e.g. a tin or pill bottle, a clay pot, the flat cylinder of the base of a frying pan).' The generating axes of the relevant generalized cones are as shown by the arrows. Where more than one generating axis could be drawn, the main axis will be assigned to that which generates the generalized cone with the greatest volume. In general this will be the longest axis of the object (as in the can and pot), but sometimes it may be a short axis (as in the case of the frying pan, where the handle would form an axis of less volume).

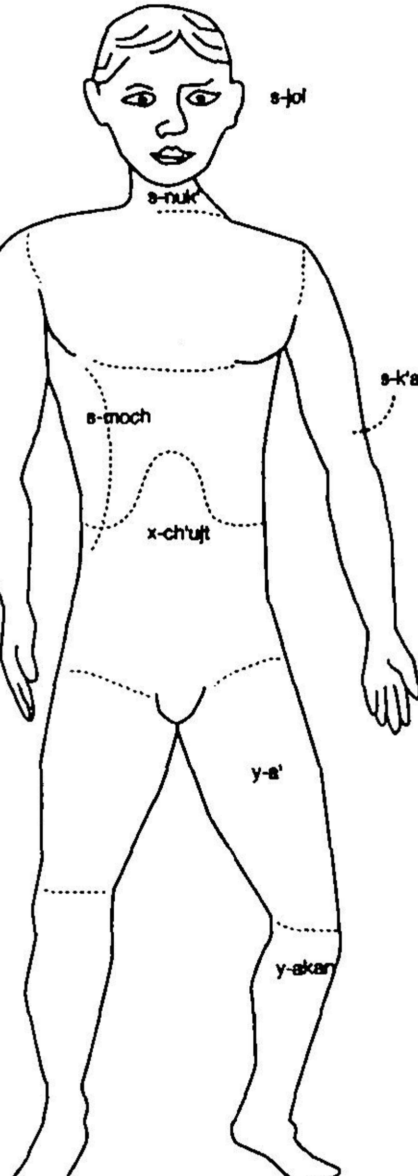


Figure 8. Native-speaker segmentations of the hwnan body

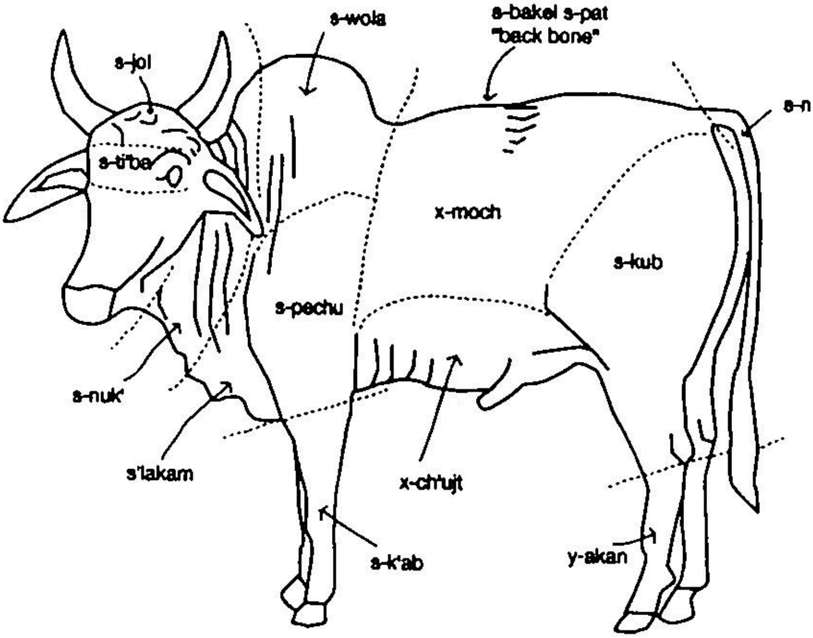
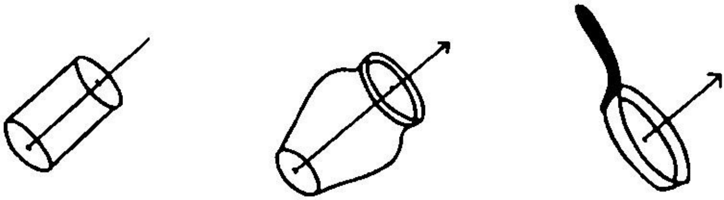


Figure 9. Native-speaker segmentations of the animal *body*

In the case of 2D objects, the main axis is assigned to the longest axis of symmetry, that is, the longest axis about which the shape reflects. The second row (b) of Figure 10 illustrates how this would apply to the leaf of a vine, the leaf of a fern, and a plank (treated for these purposes as two-dimensional like a strip of paper).

4.2.2. *Finding the directedness of the model axis.* The next problem is to assign a direction to the model axis, which may be thought of as a directed arc. The importance of this assignment is just this: the directedness of the arc is the intrinsic orientation of the object in question. There are various ways in which this assignment might be achieved, but I shall appeal to an interesting theory of the causal interpretation of shape developed by Leyton *(1989)*.[41] This is a natural extension of the analysis of wholes into parts by the segmentation process mentioned earlier, and it relies on the same basic processes. It is a mathematical theory of how all outline shapes can be reduced to sequences of sharp vs. mild convex or concave curves, and how each of these primitives has a natural interpretation as the result of a causal process. Thus for example a sharp convex fingerlike shape, with concavities on either side where it

(a) Major longitudinal axis that will generate a generalized cone



(b) 2D **shapes: main axis of symmetry**

Figure 10. *Finding* **the model (or main) axis**

joins the main body of an object, will be interpreted as a growth point, a protrusion; while a flattened convex curve (without bounding concavities) will be interpreted as a "squashing," a process of an inner force yielding to a greater outer pressure (see top of Figure 11). Parts of objects (animate or inanimate) can be identified with protrusions, and as Leyton *(1989: 363)* puts it, "A part is a PROCESS. It is a TEMPORAL or CAUSAL concept."[42]

The algorithm for finding the head (or arrow) of the directed arc should basically identify the head with that end of the model axis that ends in a protrusion or the sharpest convexity. Equally, the head of the arc may be assigned to the end of the axis opposite the flattest, most "squashed" end. This will invariably give the right result, as in row (a) of Figure 11, except possibly for hollow objects to be discussed later. But what happens where there is symmetry of both ends of the model axis, as in row (b) of the figure? In this case, either the axis is assigned two

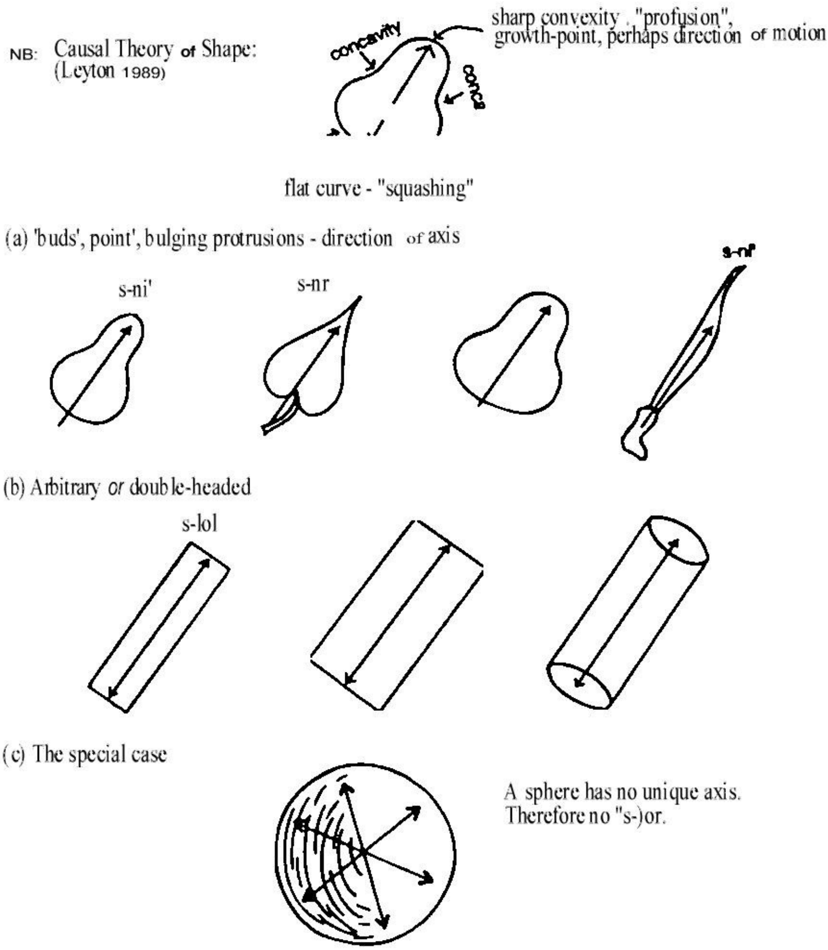


Figure 11. Finding the direction of the model axis

heads, that is, it is bidirectional, or an arbitrary decision is made. Only where an arbitrary decision is to be made may information from outside the object itself be brought to bear - specifically, by relating the head of the arc to vertical "up" as given by gravity. But this is not obligatory, or even especially common.

There is a special case of importance: the sphere. This is the one shape where there are infinitely many "longest" axes in all directions. Consequently a sphere has no unique model axis (any will do to generate it as a generalized cone), and no direction or intrinsic orientation can be assigned.

Hollow objects appear to lie outside Leyton's generalizations." How to assess a hollow end? Suppose the model axis has been assigned such that it generates a generalized cone one of whose ends terminates in a void. The void itself (at the axis point at least) has no shape; therefore, attention should be paid to the shape of the other, solid, end. If this is flattened, assign the head of the arc to the hollow end (by the corollary described above); if pointed, assign the head of the arc to the solid (pointed) end. This would predict that a hollow paper cone would have its head at the point, but an opened tin can should have its head at the open cavity." A hollow cylinder open at both ends (like a roll of fencing wire) should be assigned bidirectionality or an arbitrary head. As far as I can determine, these are the correct predictions.

4.2.3. Assigning terms to the ends of the model axis: y-it vs. s jol/s-ni'/s-ti'. We now have for each object a model axis and its two distinct ends. Linguistic labels are now assigned to the ends in a very straightforward way, as indicated in Figure 12. The base of the arc, opposite to the head of the "arrow," is almost invariably assigned the term *y-it* 'buttocks'. This fits well with the flattened convex shape that Leyton (1989) associates with the interpretation "squashing"; but regardless of the presence of such a shape, the term is the same. Thus a knife has *a y-it* at the end of the handle, opposite the point, regardless of the shape of the handle. And once the head of the arc of a plank, or carton, or other reversible shape has been assigned, *y-it* is the reverse end even though the shape is square or cubic."

The naming of the head of the arc is a little more complex. There is a default term *s jol* 'head', overridden by terms determined by specific shapes of the extremity. Where the extremity is pointed, or has the sharp convexity of a Leyton "protrusion," the term applied is invariably *s-ni'* 'nose'. Where there is a protrusion, but with more gently curved, circular outline and only minor concavities on either side of the outline, the proper term is *sjol* 'head'. Where the head of the arc lies in a hole, a void, because the object is hollow like a container without a lid, the term is invariably *s'ti'* 'mouth'." (Where the head ends in a wide flattened rectangular or oval surface, the term seems to be *y-elaw* 'face', but I am short on data here.) Otherwise, if none of these shape conditions apply, the term reverts to *s jol*, now without any shape commitments. Thus the head of the axis of a board, or a table, or a rafter is simply *s jol*.

Now note that this analysis correctly predicts that a sphere, because it has no unique model axis, has no *s jol* and no *y-it*. Any analysis that

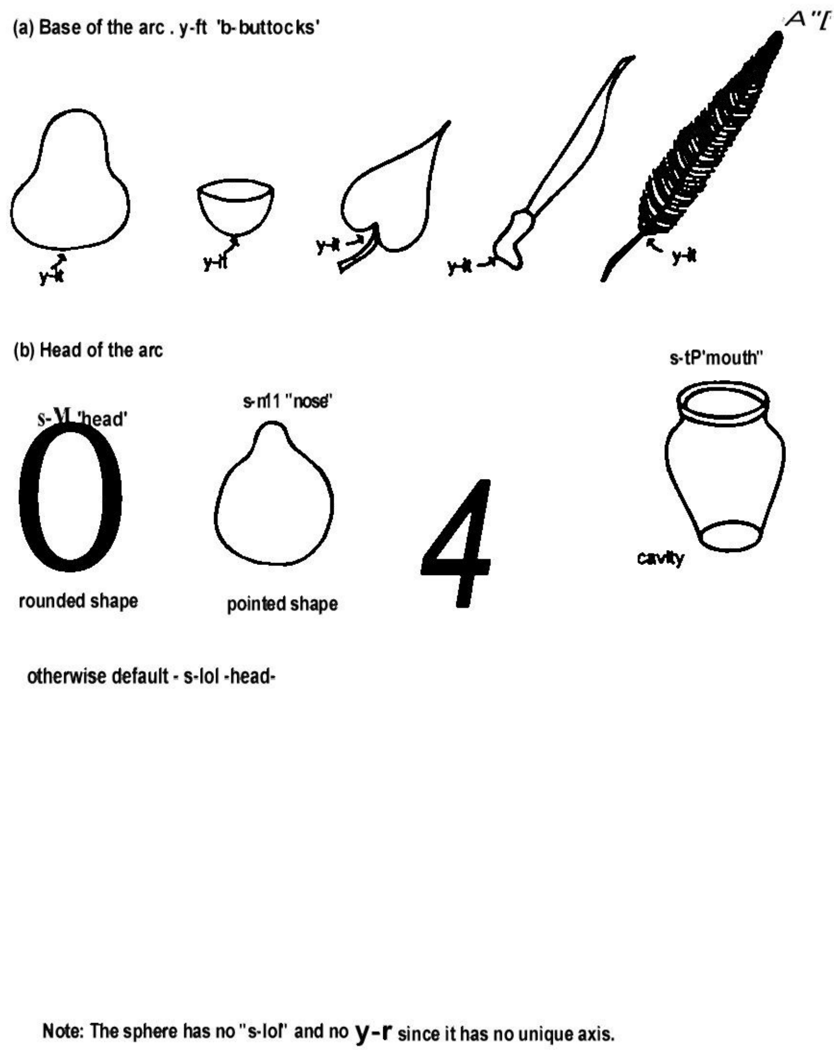


Figure 12. *Naming the ends of the model axis*

brings in non-object-centered coordinates like gravity will incorrectly predict that a sphere has in effect a top and a bottom - but in Tzeltal it doesn't, or rather it has no *syol* and no *y-it*.

At this point we should mention that for all Tzeltal object-part terms there is an ambiguity over whether the terms label points/edges/surfaces, on the one hand, or volumes on the other. My assumption is that for the central volume of an object (i.e. the volume generated along its model axis) the terms primarily label points, edges, and surfaces and not volumes. They can then be extended to yield volumetric parts, for example a stone inside a bucket can be said to be *to y-it* `at its -butt' even though it is on the inner surface, and this is because the term here has a volumetric meaning. In contrast, terms labelling axes projecting from the model axis primarily label the associated volumes, as we shall see.

The reader should note that although terms like *s-ni' will* continue to be glossed `nose', etc., on this analysis they are stripped of their bodily associations: *s-ni'* simply labels a pointed protrusion at the head of the main axis of an object, while *y-it* `buttocks' simply means base of the model axis, devoid of colorful allusions.

4.3. *Finding and naming sides of the central volume*

So far we have sketched the outlines of an algorithm for finding and naming the ends of the central volume of an object. The central volume of an object is often the only volume, and in any case it is always crucial to object recognition. Thus it is important now to assign names to further facets of the central volume. To do this, we must decide which is the main secondary axis, roughly orthogonal to the central axis. Having found that, we then have two designated facets, those in which the orthogonal axis terminates. These can then be named according to shape.

4.3.1. *Finding the orthogonal axis.* The conceptual problem here is that the orthogonal axis can be drawn at any of the 360 degrees around the central axis, as illustrated in Figure 13(a). To fix the angle, we need a procedure. The procedure is to find that section through the object that would yield two paired symmetrical halves or rather REFLECTIONS about a line. The line of the section can then be taken to be the orthogonal axis. This is illustrated in Figure 13(b), which shows, first, such a section through *a tz'amal-te'* `traditional stool' and a machete. Note that for a generalized cone with circular cross-section like a pot, there will be no unique solution. For flat objects like a leaf or board, we can get two symmetrical parts (reflections of each other) by cutting the flat surface along the model axis into two strips.

4.3.2. *Finding the direction of the subsidiary arc and naming the sides.* The ends of the secondary orthogonal axis now point to two main facets of the object, which are thus orthogonal to the secondary

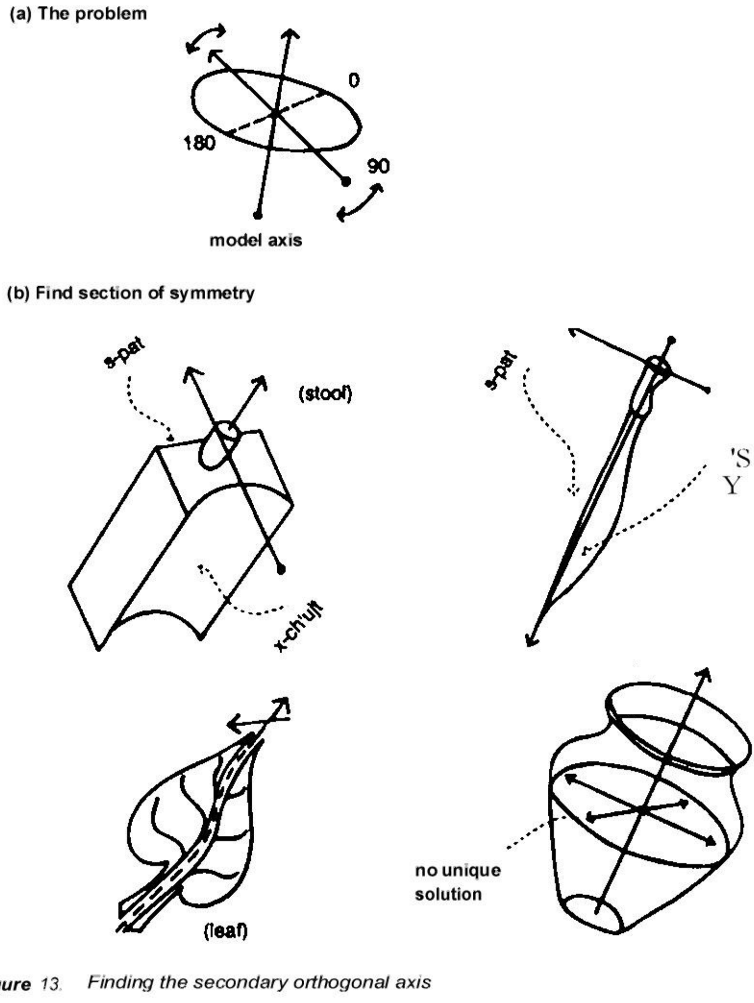


Figure 13. Finding the secondary orthogonal axis

axis itself. We can now proceed to assign the directedness of the secondary arc (to find the head of the arrow, as it were)."

The direction of the subsidiary orthogonal arc is assigned thus. Find the surfaces pointed to by each end of the arc. The head or arrow of the arc is that end that points to that facet that has the flatter surface and that is also less complex, i.e. has fewer further subfeatures. Thus in Figure 13(b), the head of the arc is toward the surface of the *tz'amal-te'* `stool' facing away from us in the picture (this is the flat sitting surface; see also Figure 14). Similarly the back edge of a knife or machete, although only a thin surface, is also pointed to by the head of the arc because the outline is less curved than the knife-edge, and also flatter across its narrow width.

What happens where there is no unique secondary axis, as when the section of the generalized cone is circular? Well, then, there can be no unique head to the secondary arc either.

What happens when the orthogonal axis points to two more or less identical surfaces? This will often happen in the case of (effectively) flat objects, for example sheets of paper or planks of wood. In this case, just as with the model axis, we can have bidirectional axes, or an arbitrary decision. Again, in a similar way to the orientation of the model axis, secondary factors may enter into such an arbitrary decision, factors of a non-object-centered sort: in this case deictic factors occasionally intrude into Tzeltal (so that `surface away from me'=head of secondary arc).

Now we can get down to naming. The head of the secondary orthogonal arc is invariably called *s-pat* `its back'. This predicts of course that the flatter surface, or the surface with fewer features, is called *s-pat.*[48] Thus the *tz'amal-te'* `stool' in Figure 13(b) clearly has *a spat* facing away from us (see also first figure in Figure 14). Similarly the back edge of a knife or machete, although only a thin surface, is *spat,* because the outline is less curved than the knife-edge. Tenejapan houses have minimal features, mostly lacking windows, but they do have a door, so the less featured side is what we would call the back, and it can be designated *s-pat.* (In actual fact there are rival conceptualizations of houses, about which more later.) Even something as seemingly symmetrical as a chile pepper can have *a spat,* being the smoother, less-curved surface, lacking a curving point (see Figure 14 top row). In short, once one has hold of the flatter, less-featured surface (conceivable as the `head' of the secondary arc), one has the object's *spat.*

Now we need to name the opposing surface, at the other end (the base) of the arc. Here the name depends on the shape (reversing the pattern on the model axis where the name of the head of the arc depends on shape, but not the bottom of the arc). If this surface is (compared to s-pat) relatively curved, whether concave or convex, it is named *x-ch'ujt* `belly', as shown in Figure 14(b). If an object has two flat sides, and one has been arbitrarily designated *spat* (facet at the head of the arc), then the other is *y-elaw* `face': thus a door has *a spat* and *y-elaw* but this cannot be predicted by, for example, which side faces the outside - it will vary on occasions of use. Sometimes, but not necessarily, the assignment will be determined by deictic criteria (see Figure 15[b]). It is possible that some flat objects have a tendency to a fixed "intrinsic" assignment of *spat* and *y-elaw,* for example a leaf. Texture may be relevant here:

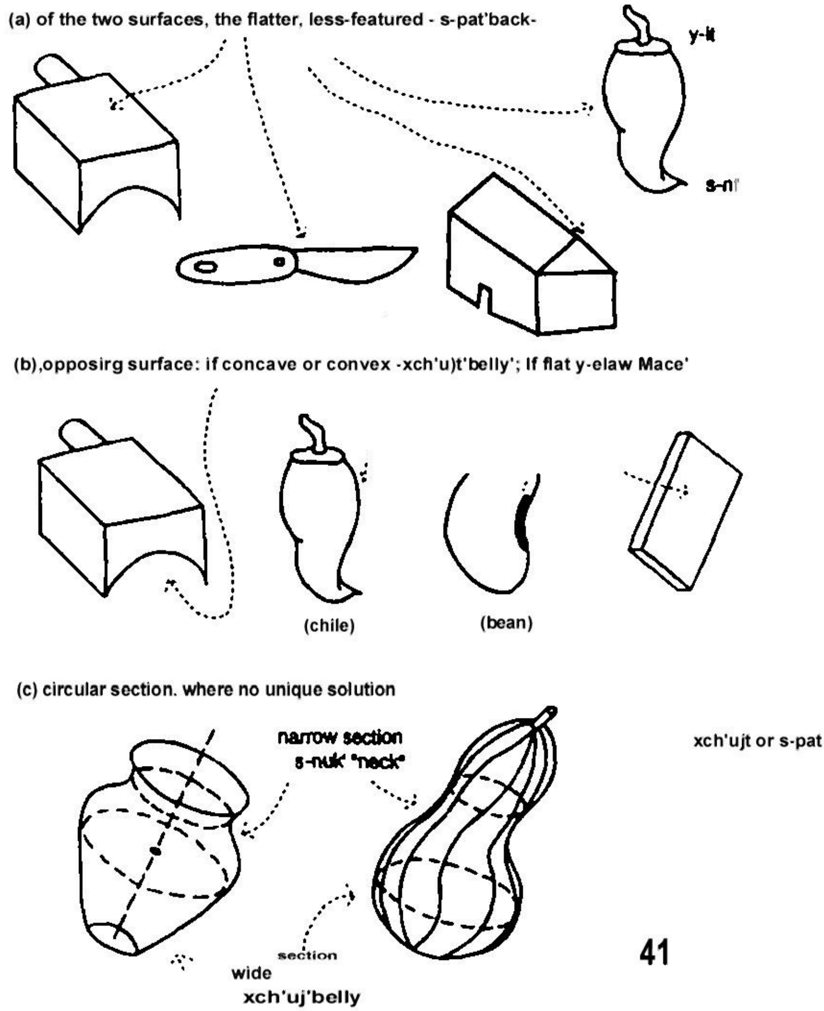


Figure 14. *Naming the surfaces at the ends of the secondary axis*

the less textured surface is perhaps preferably the *y-elaw;* but there is informant variation here, some preferring *spat.* As with doors, the variation suggests underlying arbitrariness with resolution by secondary characteristics; indeed deictic factors may also be involved here, as one normally encounters a leaf facing the *y-elaw* (Figure *15).*

What happens with objects of circular cross-section, like pots, gourds, and buckets, where there is no unique way of assigning an angle or

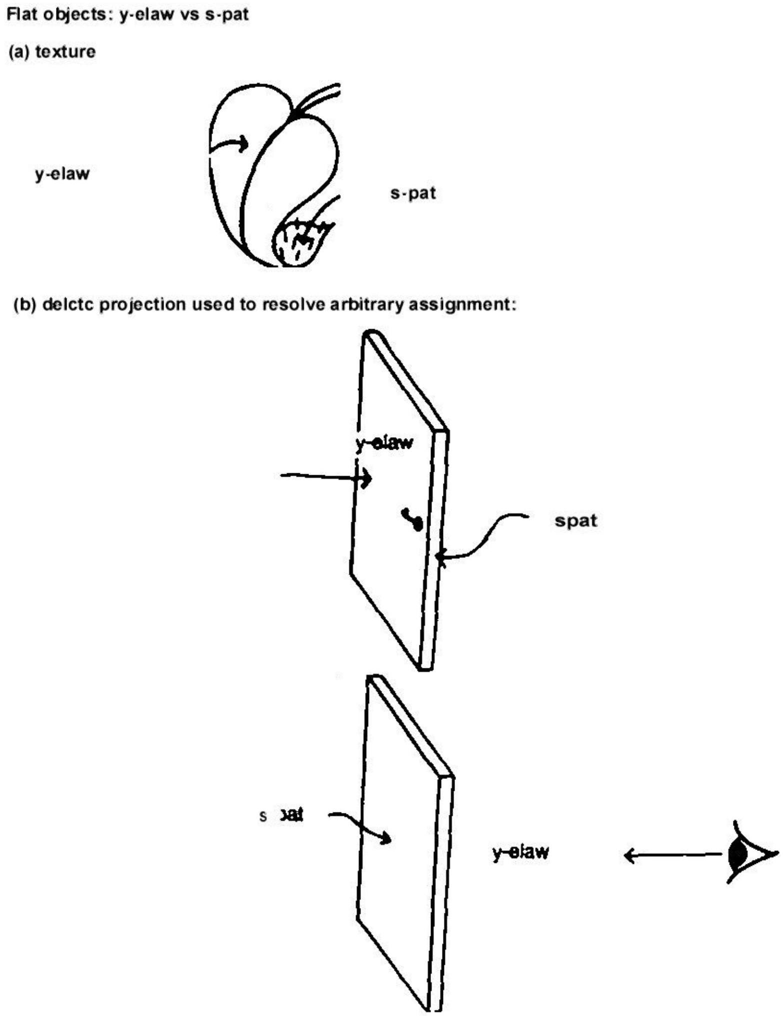


Figure 15. *Some additional motivating factors where the analysis yields only arbitrary solutions for flat objects*

direction to the orthogonal arc? In that case again an arbitrary naming decision must be made, and if the cross-section is of constant or constantly increasing or decreasing size along the model axis, then either *x-ch'ujt* or *spat* may be used to refer to the entire cylindrical surface. On the other hand if the cross-section increases and decreases, the possibility of a distinction between a wide and a narrow cross-section arises, and the former is called *x-ch'ujt* 'belly', the latter *s-nuk'* 'neck', as shown in Figure 14(c).

Turning now to cubes and parallelepipeds (rectangular solids), where the sides have unequal dimensions, the analysis so far will give arbitrary *s jot* and *y-it* (or double s jot) along the major axis, *y-elaw* and *spat* along the major subsidiary orthogonal. This leaves two sides, which are both given the term *x-xujk,* which names an object part and not, it seems, a body part (it may have application to bodies, but no part of a body is so uniquely named). In a rectangular solid with a lengthened axis but square cross-section, the sides may all be labelled *x-ch'ujt* 'belly' (especially if they bulge slightly), because just like the endless 'sides' around a pot *(x-ch'ujt),* there is no unique solution to a location of the secondary orthogonal axis (see Figure 16).

Thus on this analysis the assignment of facets and their names is determined strictly by the geometry of the object. Where the geometry offers two alternatives, we expect two naming possibilities. Where it offers a continuous range of possibilities, as around the circumference of a pot, we expect a name for the continuous surface. And so on.

It is instructive now to compare the assignment *of spat* by the theory of metaphorical transfer sketched in Figure 7. To account for the various interpretations, we needed there to refer to three distinct source domains or *schema,* the vertical human back (to get for example the vertical facet of a door), the horizontal back of a quadruped (to get the application to the horizontal surface of a stool) and the spherical outer surface of a fruit or the cylindrical outer surface of a tree (to obtain the application to the walls of a bowl or pail). The problem with that account is that it offers many competing interpretations of for example a squat upright cylinder or ovoid (e.g. a tuna or herring tin) - is the *s-pat* a part of the vertical side, the top surface, or the entire vertical surface? In short, how does one know which model to invoke?

Now one can find competing interpretations or labellings of objects, but it is hard to relate these to the alleged competing metaphorical schema. For example, one or other of the flat facets of a door or plank may be either its *s-pat* or its *y-elaw;* the reversibility of the object offers no intrinsic analysis. Often, but not always, this may be resolved by the importation of deictic criteria, as sketched in Figure 15(b). A case that might be of more interest is the analysis of the house. It does seem that more than one analysis is available. Does the metaphorical account predict these? Figure 17 gives two possible analyses following the intrinsic object-coordinated system being advanced here. The top diagram shows the shape of a traditional Tenejapan house, roughly cubic with a pyramidal roof. In our terms the shape is generated by taking a square cross-

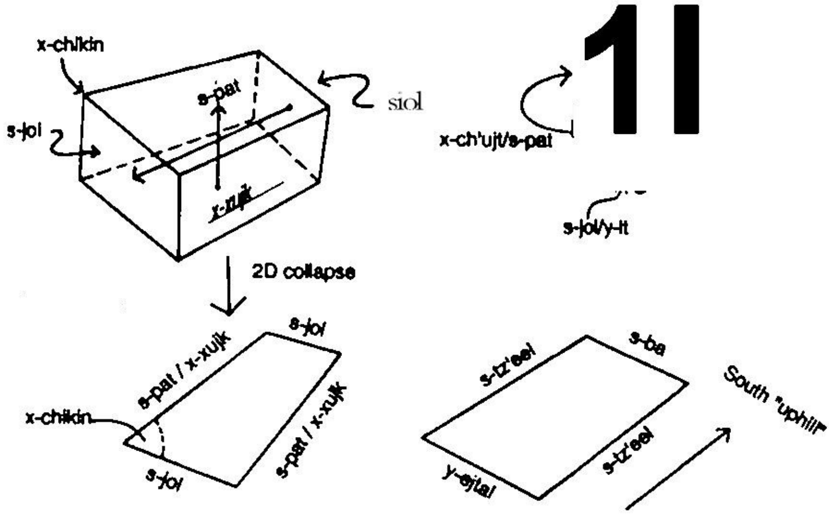


Figure 16. ***Parallelepipeds, rectangles, as collapsed parallelepipeds, relational nouns for parts of rectangles***

section of constant size along a central axis, and then systematically decreasing the size to the vanishing point (the apex of the house). The model axis is therefore (now from a non-object-centered perspective) upright, and the secondary orthogonal axis is horizontal. The account correctly predicts that the roof is *s jol,*[49] and that all the sides can be thought of as *spat* (since there is an undecidable symmetry around the central axis)." If one decides to focus on the door as a significant feature, then by virtue of having more features that side will be the *y-elaw,* and the opposite side the *s-pat,* with the corollary that the two remaining sides will be *x-xujk* (see Figure 17[b]).[51]

Now the metaphorical account does not make these predictions clearly. The upright s jot will suggest that the human model or source must be involved; but then it is difficult to account for the fact that the fruit or vegetable model must be simultaneously invoked to account for the outer surface being *spat* on all sides. The human model fares better on the alternative set of terms in the second figure (with 'face' and 'back' assigned pretty much as in English). But then why can we not invoke the animal model, so obtaining a horizontal main axis (which would fit

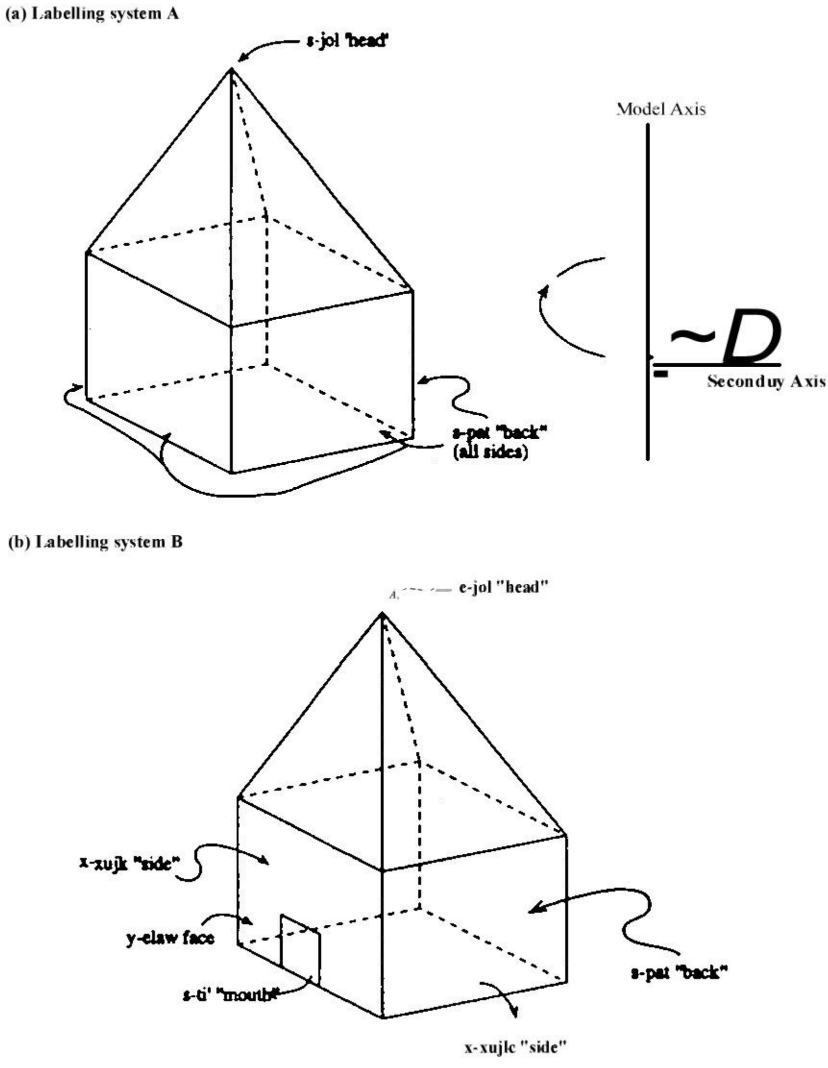


Figure 17. *Two alternative systems of labels for house parts*

with the more rectangular design of modem houses) with a 'head' at one end, 'buttocks' at the other, and 'back' along the ridge? Overall the metaphorical account adds nothing, suggests alternatives that do not occur, and indeed fails to explain the alternation between the two systems of labelling.

4.4. *Secondary volumes and their axes*

We have now finished with the central volume of an object. The next problem is to locate and name secondary volumes, by finding the axis of the main volume of the secondary part, although sometimes - where the secondary axis is very small relative to the main axis of the object - the analysis can be circumvented directly by questions of shape. Where there is more than one significant secondary volume, the location of their attachments relative to the main axis becomes pertinent.

The axis of a projecting part may come off the model axis at any angle; nevertheless orthogonal projections are frequent enough to be the main thing to look for.

4.4.1. *Small protrusions.* If the axis of a protrusion is small relative to the main axis of an object, and the volume is also small relative to the main volume (less, say, than a twentieth), names can be directly assigned on the basis of shape. The shape must necessarily have an outline that is a convexity bounded by two concavities, hence if the protrusion has three-dimensional depth the term *s-ni'* 'nose' is appropriate; if it is flattened in section the appropriate term is *x-chikin* 'ear' (see Figure 18[a]).

Now note that the metaphorical theory of body parts will here be in some trouble. For on our analysis, an object may have a pointed end to its main axis, and thus a s-ni''nose' on those grounds; now it may acquire another *s-ni'* by way of a label for a small rounded protrusion. But none of the metaphorical models or sources (human and animal bodies or plant parts) normally have more than one s-ni''nose'. The same goes for *x-chikin* 'ear': an object may have one, two, three, four, or more 'ears'. This is the normal word for the handle (lug) of a cup or water pot; traditional water pots *(k'ib)* have three pierced lugs around the neck through which ropes are attached. It is hard to see how the metaphorical theory can account for these sorts of usages, except by invoking further source schema incorporating the very objects that include these multiple parts.

4.4.2. *Large projections.* Let us (arbitrarily) define a large projection as one whose long axis is at least a fifth of the length of the model axis, radiating out from the main volume. If this projection has any significant volume, it is almost without exception labelled *y-akan* 'lower leg'. For example, the handle of a frying pan or western-style cooking pot, or a projecting part of an arbitrary plasticine shape as in Figure 18(b), are all labelled *y-akan.* On the other hand a thin projection, for example a cord out of an electrical appliance, or a cordlike projection from an

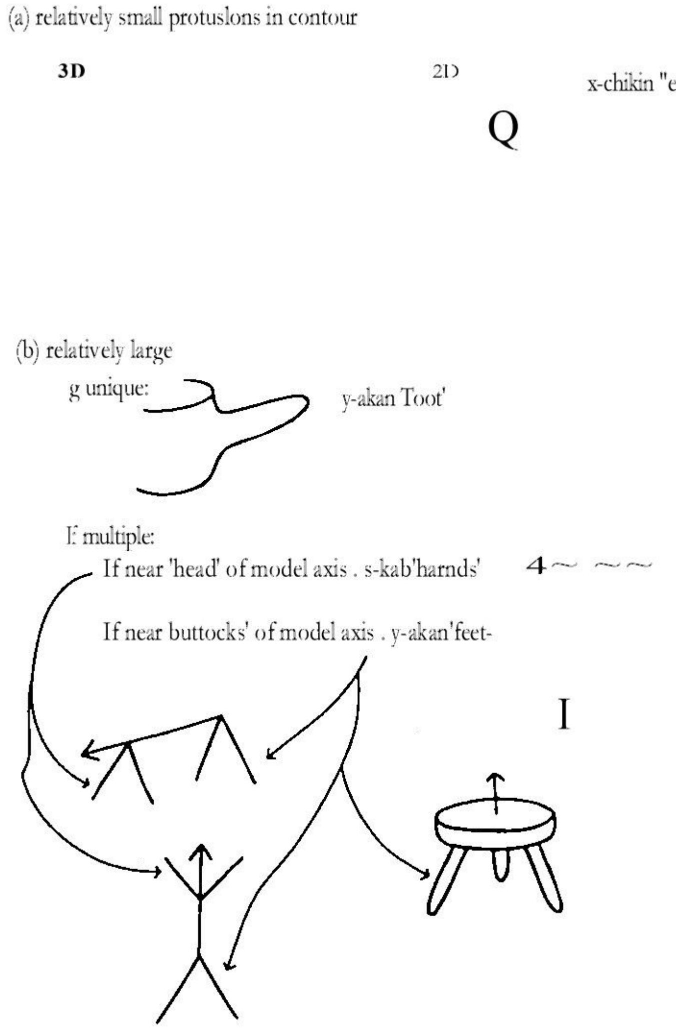


Figure 18. *Naming secondary or branching volumes*

arbitrary plasticise blob, may be called *s-ne* 'tail' (this term is relatively rarely invoked in the description of physical objects).

What if there are multiple stout projections? Here the orientation relative to the model axis becomes relevant. If there are multiple projections near the head of the arc or the "arrow" of the model axis, they are labelled *s-k'ab,* 'hand/arm', if near the base of the arc, they are *y-akan* `lower leg'. Hence the branches of a tree are *s-k'ab,* the legs of a chair are *y-akan.* Note that this analysis, which makes no appeal to the vertical or horizontal orientation of the model axis, correctly predicts the usage of these terms, even in the animate domain, since a quadruped in Tzeltal is not indeed such a thing: it has two `hands' and two `feet', regardless of its present orientation! (see Figure 18[b]).

4.5. *Remaining terms*

We have described the heart of the system, and the use of the most common body-part terms for inanimate objects. But a number of terms in Table 2 have not been described: here are some short notes.

4.5.1. The parts of surfaces, and junctions between surfaces. Many objects have flat more or less rectangular surfaces, for example the sides of planks or of houses, the tops of tables or of cement patios, the floors of houses. These surfaces can be thought of as having parts, which thus play an important role in locatives. To assign parts they must be thought of as 2D objects, or more exactly as 3D objects collapsed into one plane as suggested in Figure 16, and then the normal rules apply: the longest axis is the model axis, thus assigning *s-jol* `head' to at least one and often both ends of a rectangle. The edge around any object conceived of as a surface plane, whether an empty void (e.g. mouth of a jar) or a solid surface (e.g. a table edge), can be called its *s-ti'* `mouth' all the way around." Conceiving of a rectangle as a neutralization of a three-dimensional parallelepiped into a two-dimensional rectangle, the edges or edge areas orthogonal to the model axis may be called either *s-pat.* or *x-xujk.* The corners are labelled *x-chikin* 'ears'; this is appropriate because as we have seen this term labels 2D protrusions from 3D volumes, so the usage makes sense given a collapse of three dimensions onto two.

x-xujk is also sometimes used for corner: the basis of this usage is perhaps indicated in usages like *to x-xujk s jol mexa* 'to the side of the head of the table' (i.e. in a corner of the table), indicating an orthogonal line to the main axis terminating in *s jol* (recollect that *x-xujk* in 3D models is orthogonal to both the model axis and the *spat/ yelaw* axis). However, *x-xujk* also refers to an edge in a 3D model (unlike *s-ti',* which refers to an edge in a 2D model). As mentioned, *x-xujk is* not a proper body part, in the same nominal class as the terms we have been investigating. It has a counter-part *tz'eel* 'side, edge', from a positional root meaning 'edgewise', which is also frequently used to denote both 2D and 3D edges on an orthogonal to the main axis.

The middle of a planar surface is designated *y-olil* 'middle'. This is also not a body-part term, although it may be a true object-part term. However, it probably has prime reference to the mid-point of a 1D line and refers to a 2D or 3D object through extension.

4.5.2. *Surface features.* The remaining terms *s-sit* 'eye', *y-uxub* 'navel', and *y-ej* 'teeth' can be taken to refer to surface features. A small round surface marking is normally *s-sit,* a slightly larger dimple *y-uxub,* although I doubt there is any precise distinction. An edge that is formed by two planes at such an acute angle that it is truly 'sharp' can be labelled *y-ej.* Thus the sharp edge of a knife, axe or machete is *y-ej* 'its teeth', and this term preempts the assignment *of x-ch'ujt (i.e.* the opposing term to *s-pat)* but just on the very line of the sharp edge, leaving the axis-determined term to apply to the region along the edge.

4.5.3. *Some problems for the analysis.* The keen-eyed reader will no doubt spot loopholes in the analysis so far attempted. But I should draw attention to those I know of, together with some object labellings that escape the current analysis.

One crucial issue is whether coordinates outside the object can really be avoided. There are two such coordinates that, on the present theory, enter in a secondary role, as optional methods of deciding between arbitrary object analyses. The one is the vertical dimension given by gravity; where the model axis is so aligned, it is natural (but not necessary) to assign the vertex or head of the arc to the upward end of the object. But this I stress is only possible where the intrinsic shapes of the apexes of the arc fail to determine the head of the arc. The second is the egocentric ("deictic") angle of viewing: again, just where the internal coordinate system fails to decide the apex or head of the secondary arc orthogonal to the model axis, it is possible (but distinctly optional) to assign the head of the arc as pointing away from viewer. This will achieve the effect of *spat* meaning 'the surface on the far side of the object from the viewer', like English *behind.* This is not the main use of the term, and there is no such option where the object is intrinsically assigned *a spat.* I view these effects as marginal to this system - the optional importation of external coordinates when the internal system fails to determine a unique solution.

Another point is that there are objects for which the present account possibly gives the wrong predictions. A possible counter-class is constituted by agricultural tools. The *s jol,* the 'head' of an axe or a hoe is the working metal end, the sharp edges are *y-ej* 'teeth', the base of the handle *y-it,* the body of the handle is *y-akan.* The *y-akan will* normally exceed the head of the tool in length and may do so in volume, though perhaps rarely in mass. Is it thought of as a secondary projection to the main body constituted by the head of the tool? If so, then knowledge of function (or at least knowledge of mass) and not purely bottom-up geometry is required to label tools." Since this forms a small (albeit important) fixed class of local objects, the failure of the account (if such it is) may not matter too much. But it may token a much larger range of cases where knowledge about function intrudes in the designation of object parts (which we have partly already acknowledged in the assignment of *y-ej* to sharp edges).

Finally, I should mention that for complex objects, I assume that the process can somehow be applied recursively. That is, having found the model axis, and a substantial secondary axis, one may find a tertiary axis off the secondary one. In this case one will need to have assigned directed arcs to the secondary axis, which I presume will proceed as for the model axis, and the secondary axis will consequently have a 'nose' (or 'head') and 'butt' and 'back', etc. I presume speakers could be pressed into doing such an analysis, for example if they had to label a complex Lego or Tinker Toy construction for another person. But we presently lack the relevant data.

5. Conclusions: object-centered orientation in Tzeltal

5.1. *Metaphor vs. intrinsic geometry*

Let us sum up. The theory advanced claims that Tzeltal body parts are applied to objects almost purely in terms of their internal geometry, without essential reference to any external coordinates (such as the vertical dimension, or horizontal angles determined with respect to the viewer). Further, there is no necessary reference to encyclopedic knowledge: the derivation of the names for parts is potentially entirely "bottom up," by calculation of intrinsic shape. Hence there is no mystery in the ease and naturalness with which new shapes can be assigned names.

This theory contrasts greatly with that implicit or explicit in other treatments of such phenomena, namely the view that these part names are metaphorically applied from (especially) the human and animal domains. For a metaphorical analysis presupposes a relation between *two Gestalten,* brought into correspondence as wholes, so that the articulations or segmentations of the source can be used to supply labelled segmentations to the target. Such a theory is "top down," requiring in the analysis of object A knowledge about another object B. Nor does it seem to require any kind of thorough-going geometrical analysis: when we see the Big Dipper as a frying pan, we are essentially pattern-matching. (At least if it is geometry, it's topology, not the Euclidean stuff we find in the implicit geometry of Tzeltal body-part terms.)

The advantages of the analysis in terms of intrinsic geometry are numerous, and a number have already been mentioned in passing. The argument against the metaphorical analysis can be summarized as follows:

i. The metaphor analysis requires multiple *schema* or source domains (e.g. human, animal, and vegetable sources for 'back'), as otherwise (as a matter of fact) the configurations of parts of the targets will not prove matchable.

ii. Consequently, it will be relatively hard to find attested uses that cannot be "generated" by a metaphorical analysis (and in this respect it is hard to disprove such a theory), because it is always possible for the analyst to choose between the most appropriate source schema, and indeed invent others.

iii. Nevertheless, despite that, it is difficult to find plausible source schema that can account for the strange numbers and dispositions of parts - for example all the branches of a tree are 'hands', and it is commonplace for objects to have two 'heads', multiple 'noses', four or more 'ears', etc.

iv. Further, a metaphorical analysis will hugely overgenerate predicted uses beyond what is attested. Multiple competing source schemas should predict alternative analyses of objects, each of which is aligned to the corresponding source; but in fact, the alternative analyses offered by informants fail to match the predicted ones - rather they support the geometrical analysis (see e.g. remarks on houses above);

v. Analogies and metaphors map most naturally from relatively simple and concrete sources to relatively complex and abstract target domains. This suggests that objects with few features (e.g. geometrical shapes, or a cylinder of plasticine with a small protrusion) should prove difficult to analyze into named parts, as they are simpler than, and give too little structure to engage with, any of the source schema. But in fact the parts of simple objects are easily and directly labelled.

vi. Analogy or metaphor typically proceeds by fruitful mapping between two domains: all parts and aspects of the source configuration may be put into correspondence with the target domain. This predicts that all body-part terms should be exploitable, while in fact less than a score are. It also predicts exploitation of Gestalt aspects of the source domain like vertical orientation of the human frame, while in fact canonical orientation plays little or no role in the application of body-part terms to objects.

vii. If for all these reasons we reject an account in terms of "fresh" metaphor or generative analogy, we might still hope to retreat to an account in terms of "dead" or conventionalized metaphor. But then we have no account of the central phenomenon, namely the generative application of body-part terms without hesitation to the parts of novel objects or to objects of indefinitely varying shape.

viii. Finally, there is an argument from simplicity. Metaphor is a mysterious, complex process; elementary geometry is not; an account in terms of the latter, if feasible, should be preferred.

If metaphor is not central (or even peripheral) to our understanding of the spatial and object-segmenting uses of these terms, how does one account for the fact that nevertheless these Axe body-part terms applied elsewhere? Before settling graciously for a diachronic account (such as to be found in Heine et al. 1991) in terms of an ultimate (if now irrelevant) metaphorical origin, two cautions are in order. First, if you're spoiling for a fight, you could reverse the direction of argument, and say that humans have *y-akan* 'lower legs' because that follows from the meaning *of y-akan* (major secondary axis attached at one basal node to model axis). Second, there is areal evidence for a source VIA semantic calques from other languages (see section 5.2). Thus it is possible that there never was a strictly metaphorical origin at all."

It is perhaps worth dwelling a little further on the point in (iv), namely that a metaphorical analysis fails to account for the fact that only a handful of the 100 or so available terms from the three source domains can be used to label parts of inanimate objects. If the process is metaphorical, and one sees a pair of sharp projections as horns, why can't one call them *xulub* 'horns' (presuming for example a bull model, bulls being the prototype Tenejapan livestock)?[55] One can't, or at least, only fancifully, that is, metaphorically! And why use *s ni*''nose' as the prototype protrusion, and not say *x-chu'* 'breast'?[56] Why ignore all the metaphorical possibilities of shoulders, chins, chests, cheeks, and so on? How would one block the application of such terms under any free metaphorical process?

On the object-centered, geometrical account, there is no reference to rich sets of vocabulary in parallel domains. There is simply a need for whatever terms are required by the particular geometrical distinctions. For each object, we need terms for each end of the model axis, between one and four for the facets related to the orthogonal axis (one for circular cross-sections, two for flat objects, four for parallelepipeds), three for secondary protrusions, and an indeterminate number of surface features (three mentioned above). Allowing some choice according to shape for the ends of the axes, we have terms enough in Table 2. Why TIm5E particular 15 or 20 terms out of the 100-odd terms for parts of humans or animals or vegetables? Well, why not? One might be able to go a little bit further and show that once you've chosen body-part terms to select from, only some provide the necessary clear shapes, connections, and axes: for example *y-akan* 'lower leg' has only one connection, at the knee, *y-a'* 'upper leg' has two, and would not be of the same general utility; *s-moch* 'flank' might be useful (cf. *x-xujk*), but it will not be carved out by the algorithm that segments at concavities.

Spelling out the geometric account is of course tedious, but that should not be mistaken for complexity. Locating a couple of axes, applying some heuristics for directionality, choosing terms for apexes according to simple shape oppositions - that's child's play. And so indeed it is: children appear to (begin to) learn the system very early: de Leon (1991 b) finds children using the Tzotzil system well before three years of age (and I assume that the Tzotzil system is at least similar to the neighboring Tzeltal one). In contrast, making complex metaphors may not be junior business at all."

I have tried to establish the plausibility of a geometric account of the body-part terms. It would need refining and testing. One attraction is that it makes a great many testable predictions. For example, since what is posited is a sequence of geometrical constructions for each object, we may expect a conceptual and temporal priority of model axis assignments over secondary axis assignments: it should, for example take informants longer to judge misassignments of terms on minor axes. Similarly, we would expect the model axis assignment to have a primacy in acquisition, and thus the correct use of *s jol* (head of arc) and *y-it* (base of arc) to proceed the correct use of *spat* (surface at head of secondary axis) and *y-elaw* (flat surface at base of secondary arc). Although these terms may be produced early because of their importance in adult speech, their correct use might be harder for the child to understand than the directly shape-based terms like *sni'*, *y-akan,* and *x-chikin.* Above all, the account makes precise predictions about the correct partonymy of any random novel object.

5.2. So *what for the analysis of Mayan?*

5.2.1. *Tzeltal.* I have tried to establish that Tzeltal might be said to be a richly "visual" language in two senses. First: it makes you notice and report shapes. Second: the way it makes you think about shapes is just the way the visual system makes you think about shapes.

In the case of the body-part system, that amounts to saying that to speak the language you need to be able to do instant visual geometry of the kind described. You can't say the equivalent of 'Watch out for that snake ON THE Top of that stone' without deciding on the geometrical properties of the stone, since the relevant term might be *spat, y-elaw, s-ni',* or *s jol* (well, there are ways around the decision, but that's another story). You can't make those decisions without actually seeing the stone.

The Tzeltal body-part system is, as earlier described, important for locative descriptions, because it is the simplest way to build detailed specifications about where the figure (the object to be located) is with respect to the ground object. Essentially this is done by partitioning the ground object into its parts and saying where the figure is with respect to the closest or enclosing part (e.g. 'at the ear [corner] of the table', 'at the lower leg of the fence-post', 'at the head of the door', 'at the bottom of the pot', etc.). This system is pretty strictly interpreted: *to y-it p'ine* 'at its-butt the bean-pot, i.e. at the butt of the bean-pot' means that the. figure is to be found contiguous with (very close if not actually touching) that facet of the pot labelled its butt. It does not, for example, mean 'underneath, but not touching'. Similarly, something on my head can be said to be *to j jol,* 'at my head', but not floating two inches above it (that would be *to y-ajk'ol j jol* 'at the aboveness of my head'). Most of the body parts have this very strict interpretation.[58] But at least two terms allow a little more latitude, that is, they allow the figure to be in some region around the designated face. *x-xujk is* one such term, allowing some distance between figure and ground, a bit like English 'to the side of' (although *x-xujk,* as mentioned above, doesn't really belong to the same nominal class). The other is *s-pat,* which can be used to designate a region of some extent projected off the facet so named. It is perhaps not irrelevant that (as mentioned) *spat* may be assigned deictically if there is no intrinsic determination. However, the regional use of *spat* occurs both when the facet is intrinsically determined and when it is deictically determined. In any case, this is the body-part term that is most semantically "prepositional," in the sense that it potentially labels a region and not just a facet.

It is entirely in keeping with the intrinsic geometric nature of the Tzeltal body-part terms applied to inanimate objects that they essentially just label parts, not (with the exceptions just mentioned) regions. To describe regions around objects another set of relational nouns is used, which allow one, for example, to specify some object as 'uphill' (equating with south) or 'downhill' (north) of another (see Brown 1991 for overview).

But although the Tzeltal body-part terms have some real importance to locative description, that is not perhaps the main reason to be interested in an analysis of this sort. For these terms are the poor cousins, as it were, of an enormously rich system of Tzeltal shape descriptors, encoded in the couple of hundred positional roots, and also in many transitive verb roots. As mentioned in section 1, perhaps half of these roots describe object shape and flexion without reference to external frames of reference - i.e. they describe the internal geometry of the subject just like the body-part terms do for their possessors. Not all Mayan languages seem to have body parts that work as they do in Tzeltal locatives, but perhaps all have positional roots with these kinds of shape specifications. In that case, understanding these poor cousins, as well as their richer positional cousins, may be a good road into one of the core areas of Mayan semantics.[59]

Brown (1991: 36ff., this issue) has pointed out that there are nontrivial connections between body parts and positionals.[60] For example, the description of the content of positional roots often requires reference to body parts: they may specify that arguments be standing on their heads, or lying on their sides, or on their backs face up. But the connection may be a little deeper. For example, what do a man on his haunches, a traditional stool placed upright on its `butt', and a lump of dough shaped like an upright artillery shell have in common? They're all described with the positional adjective *jukul*, and they all have their *y-it* 'buttocks' on the ground and their *s jol* 'heads' in the air. Or a slightly more complex example: roots of 'standing' are a little puzzling: there's a different positional adjective (i) for humans (or four-footed animals) standing on their 'hind' legs, as opposed to trees *(tek'el* and *tekel)*,[61] (ii) another one for four-footed animals on all fours, feathered bipeds and the traditional stool *(kotol)*,[62] (iii) for the Western-type chair, the harp, and stationary vehicles *(chotol)*, etc. Analysis in terms of oriented model axes may supply the key: (i) applies to things with the major arc vertically 'up'; (ii) to things with the major arc horizontal; (iii) to things with a less clear orientation - the flat plane at the top of the chair curves up the back (*y-elaw)*, while the traditional harp is a complex three-footed soundbox with vertical projection (whose body parts I do not know). One now needs to reexamine the data.

There may be a second connection: from body parts to directionals and motion verbs. Naturally, humans and animals have canonical directions of motion. But bottles and feathers do not. When describing static arrays of objects, informants sometimes talk in terms of things 'coming (this way)', 'going (away from me)'. This seems to be quite basic to Tzotzil descriptions of states (see de Leon 1991b); I believe it's much less central in Tzeltal, but it occurs. It seems that in Tzeltal, but not in Tzotzil, it would be natural to talk of something static 'coming toward me' only if the head of the model axis is pointing to the speaker - for example if the bottle is lying down with its mouth *(s-ti)* toward the viewer, or the feather is so arranged that its tip *(s-ni) is* heading toward the speaker. Again, this now needs checking.

5.2.2. The language family and the linguistic area. Body-part metaphors are a Mesoamerican areal feature central to spatial description in many of the languages of the region. They have also been supposed to be universally, or at least very generally, the sources of prepositions and their associated content (Svorou 1986; Heine 1989; Bowden 1991). It is 1mportant therefore to see that (a) the sense of "metaphor" may be rather misleading, (b) it is possible to use body parts for the description of objects in very different ways - attributing the parts on the basis of internal geometry (like Tzeltal) or external orientation or conventional assignment (like English *top, bottom, front, side, back).* Such systems may look superficially alike but be very different indeed.

To underline this point it is worth contrasting Zapotec body-part locatives, as described by MacLaury (1989). The system consists of a core of seven parts in an array of fixed locations, which might be conceived of as a sort of fixed vertical armature surrounding an object. Thus when an object is rotated within this armature, the terms for the parts must change. This is a system based not on object orientation, but extrinsic orientation, and indeed seems (compared to Tzeltal) to make relatively few shape-based distinctions.

Despite this, Zapotec, Nahuatl, Tzeltal, and no doubt many other Mesoamerican languages of unrelated stocks utilize a similar core set of body-part terms oFrEN WITH ALMOST EXACTLY THE SAME SHAPE APPLICATIONS. For example, the term for 'mouth' is associated with cavity, edge, and lid in all three languages." Quite clearly, not only are these not fresh metaphors, they are not independent inventions. Presumably these systems are semantic calques based on high-prestige languages during various periods of imperial extension (whether Natuatl, lowland Mayan, or further back in time). Indeed, the systems may offer some interesting insights into political prehistory. But the point to be made here is that despite these striking similarities, the application of the systems seems to be on an entirely different basis (object orientation vs. extrinsic orientation) in the case of Tzeltal vs. Zapotec (see MacLaury 1989 also for comparison with Mixtec).

5.3. *Object-centering in typologies of spatial systems*

The object-centered nature of Tzeltal body-partonymy may be theoretically quite instructive.

In reasoning about English prepositions, there is much talk of a "topological" vs. "projective" distinction: *in, on, at* are "topological," *in front of, behind, to the left of* are "projective." The terms are Piagetian, quasi rather than truly mathematical, and play an important role in the theory of acquisition: topological notions and terms are said to precede projective ones. In fact, though, the data is rather less clear: the sequence in English is roughly *up, down > in > on > under > next to, beside > (non-deictic) in back of/in front of> (deictic) in front of/in back of.* As Melissa Bowerman has pointed out, that makes absolute verticality (not the topology of the assemblage) earliest developmentally; then comes topological containment; then proximity; then object-internal geometry; then viewer-centric perspective. This is none too clear a sequence, especially in English where the notions are not kept apart *(on top of* might refer to the intrinsic top or the vertically assigned top). Tzeltal keeps things (relatively) separate: verticality is in the relational nouns, object-centered perspectives are encoded in the body parts, and viewer perspective in the verbs of motion, as shown in Figure *19.* For all sorts of purposes, whether cross-linguistic typology, the diachronic development of adpositions, or the study of acquisition, we might do well to attend to the Tzeltal family of systems, which offer a much finer-grained set of distinctions."

Levelt *(1984),* thinking just of the object-centered vs. viewer-centered opposition in English or related languages, asks the question, why two independent systems for spatial description? He suggests that each has its own advantages: the object-centered description does not presuppose that the addressee has the same point of view as the speaker; while the viewer-centered system makes it much easier to reason spatially, because notions like *to the left of* (used deictically) are transitive, while deictic *in front* and *behind* are converses. To these observations we should add that the Tzeltal use of absolute coordinates (a bit like north and south) has the functional advantages of Born systems: such systems are transitive ('X is uphill of Y' and ' W is uphill of X' allow the inference ' W is uphill of Y'), yet do not presuppose a particular viewing point. Although Tzeltal has both intrinsic and absolute systems,[65] there are languages that have only the absolute ones, which suggests that Levelt's hypothesis is essentially correct."

In all this, there lurks one striking puzzle: in the theory of vision, we start off with a viewer perspective, and all the processing goes into abstracting out a non-viewer-centered, object-centered perspective. But in the theory of human development, it appears that English-speaking children at least start off with an object-centered perspective, supplemented with the absolute coordinates `up'/`down', and slowly learn how to adopt a viewer-centered perspective. A paradox? Not if you believe in

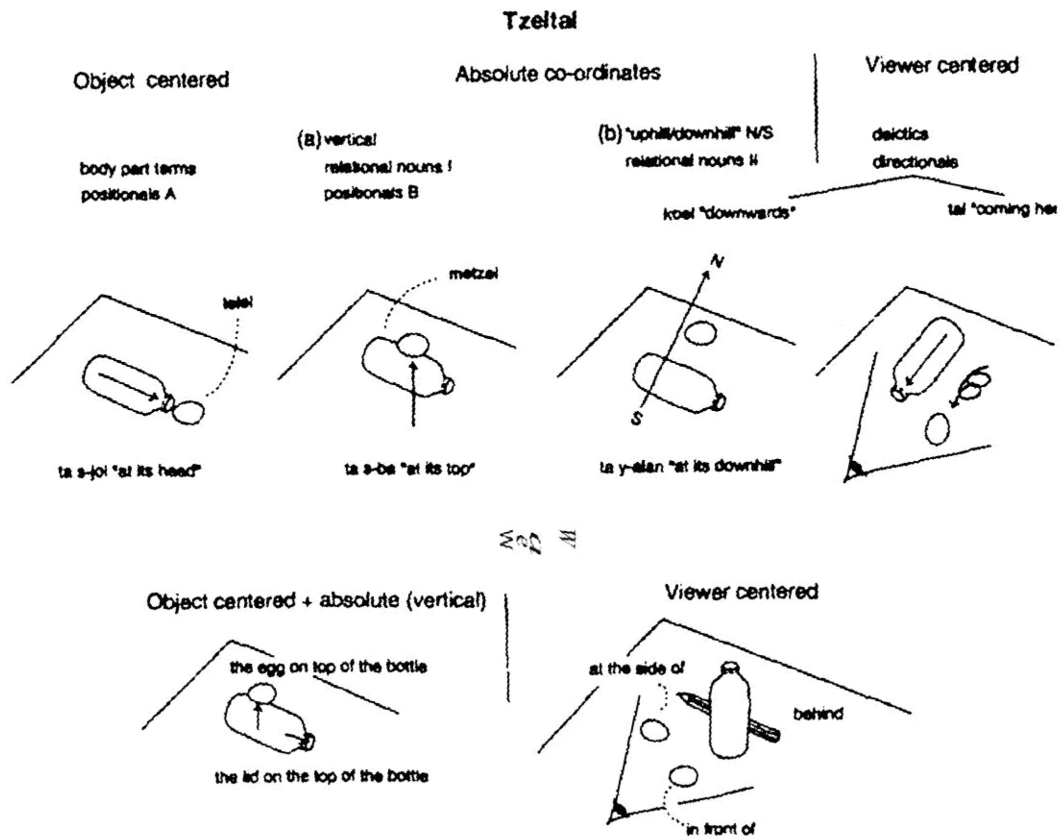


Figure 19. *Tzeltal vs. English: distinctions between object-centered, viewer-centered, and absolute* <

modularity - then it's just what you'd expect: the vision system dumps object-centered perspectives into the central processor where language can get at them; language has little access to earlier stages of vision processing, since they are encapsulated in a `module'. Thus the child must reconstruct *de novo* on the conceptual level what is given from the start, but inaccessible, on the perceptual level." This is precisely what Piaget so influentially sought to establish.

The paradox in the Tzeltal case, however, may not be so comfortably banished. First, although we are only beginning to collect acquisition data for Tzeltal, de Leon's (1991a, 1991b, 1992, in press) studies of the acquisition of the closely related language Tzotzil, seem to point in a different direction. Both absolute anchoring and deictic or viewer-centered trajectories appear to be early acquired in Tzotzil. Second, the Tzeltal data argue for something less than full encapsulation (or inaccessibility) of the visual processes. This brings us full circle, back to a consideration of the relation between language and vision.

5.4. *Vision and language again*

What exactly is the relation of linguistic representations to visual ones? We observed at the outset (section 2.2) that much, but not all, spatial language is deictic, that is, the relevant representations are often viewer-centered (or 2.5D) rather than object-centered as in Marr's 3D visual representations. Even the so-called topological prepositions, like English *on,* not to mention the projective prepositions like *under,* tend to incorporate environmental directions, especially gravitational, and are thus not strictly object-centered in the Marr sense." All this would seem to make Jackendoff's (1987: 196ff.) view that Marr's 3D level is the right interface between language and vision, and indeed the basis for all our spatial thinking, rather implausible as a general solution .[69] Even with the extensions outlined in Jackendoff and Landau (1992 [1991]) it is not clear how object-centered and viewer-centered information is to be integrated. Instead, perhaps, a general intermodal level of spatial information that is observer-centered and environmentally grounded seems to be what is needed for locomotion and other spatial tasks as well as linguistic description of the familiar English kind.

What is interesting about the Tzeltal body-part system is that, contrary to that last assumption, it is a genuinely object-centered or 3D system, almost completely free of viewpoint perspective or environmental axes. It shows us what a 3D spatial linguistic system would look like, and it is positively un-English. Moreover the linguistic system taps in, not just to the output of a 3D object-recognition system, but also to the internal primitives of such an object-centered process of analysis (i.e. to many internal axes, and specific shape contours). That is what makes the Tzeltal body parts important evidence about the nature of the language/vision interface.

This paper has tried to suggest that cross-linguistic data might throw some light on the interface between language and vision in the area of shape and space representation. It would be nice if such data could be used in a principled, knock-down way to decide between alternative analyses in the theory of vision, and vice versa. Unfortunately, such implications are not straightforward for a number of reasons, which are worth spelling out. There are at least two theoretical points on which a stand must be taken before we can make any such inferences:

1. *Modularity as encapsulation.* If visual processes are truly encapsulated in the Fodorean sense, and thus inaccessible to any other process, any linguistic descriptions anchored, for example, to 2D retinal images might really argue for a 2D, not a 3D, output of the visual module; but if something less than full encapsulation is imagined, no such inferences are valid.

2. *Redundancy and recapitulation.* We might indeed have full modularity, and yet behavior that mimics access to nonencapsulated processes. For example, Piaget and Inhelder (1956) argue that although children must have early perceptual geometric analysis, yet they only acquire such knowledge on the conceptual level through a long process of learning. Adults then exhibit a conceptual handling of the observer perspective (as in *to the left of*) that might erroneously suggest direct access to the observer perspective intrinsic to visual processing. Thus parallel representations might be painfully built up in distinct encapsulated modules. Let us call this duplication of representational processes "recapitulation" (in the sense that what is already constructed in one module has to be rebuilt in another). Is such redundancy plausible? Current answers are likely only to reflect theoretical predilection, although acquisition studies may indeed ultimately offer deciding evidence here.

If one takes the position that vision is not totally encapsulated, and that recapitulation is not to be invoked without special evidence, the Tzeltal body-part system offers some *prima facie* support for a 3D level of visual processing. It is worth now alluding to the fact that Tzeltal also offers some evidence of access to a 2D level of processing.

There is a set of positional stative adjectives that describe a long object as having the property of, variously, being (a) upright (vertical), (b) horizontal, (c) inclined at a steep angle to the vertical, (d) inclined at a roughly 45-degree angle to the vertical, etc. However, these predicate expressions also apply to objects arranged in front of the viewer in such a way that the (a) expression also applies to a horizontal object pointing directly away from the viewer, the (b) expression applies to an object across the line of sight, the (c) expression applies to a horizontal object at a slight angle from the line of sight, the (d) expression to one at a greater angle from the line of sight, etc. How can the same expression mean `vertical' and 'lying horizontal in the line of sight'? The generalization clearly is that from the point of view of the 2D retinal projection, the patterns are the same, as illustrated in Figure 20.

Here is another case. Tzeltal has some expressions that are used adverbially to denote 'uphill' and 'downhill'. Due to the overall fall of the steep terrain toward the north, these terms also denote an abstract 'south' and 'north' respectively. The system is the subject of another paper (Brown and Levinson 1993); suffice it to say here that it is essentially based on absolute coordinates or angles (not on a viewer- or object-centered perspective). But what is curious is that there is an additional use of the same terms: 'uphill' refers to a region further away from the speaker, 'downhill' as closer to the speaker, in the immediate visual field (up to say two meters). Again, this usage is explained on the simple grounds that what is further away from the speaker is higher in the 2D retinal projection, the visual array: thus it is natural that 'uphill' is extended to mean 'further from the viewer' (see Figure 21). Natural, that is, provided that linguistic processes have access to 2D visual processes.

What should we conclude? One possibility is to consider all these facts about Tzeltal as arguments tending in one (highly controversial) direction: vision is not fully encapsulated.` Tzeltal then exploits this, gaining access to 2D and 3D processes at various levels, and making linguistic

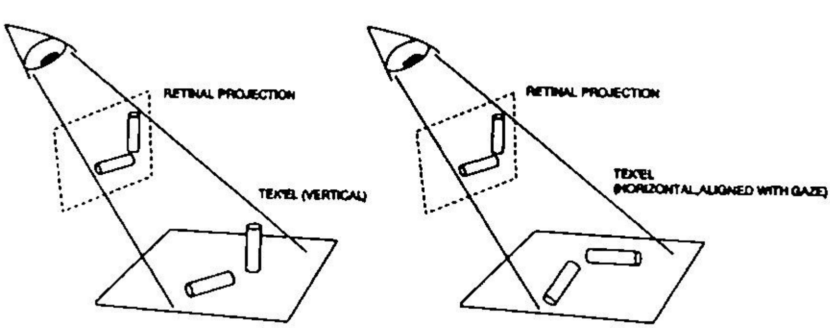


Figure 20. *Predicates collapsing the horizontal and the vertical on the basis of 2D equivalence*

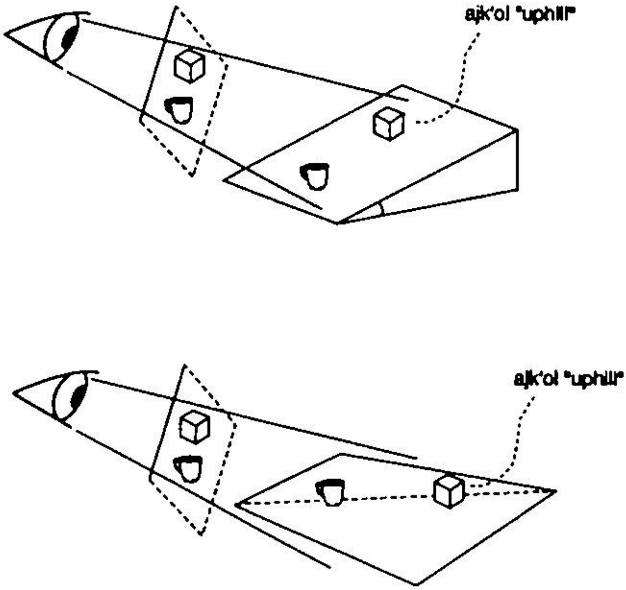


Figure 21. *Relational nouns 'uphill/downhill' collapsing horizontal and vertical on the basis of 2D equivalence*

discriminations sometimes extending over both 2D and 3D representations (as in the 'uphill' case), sometimes specific to a very detailed level of 3D processing (as in the body-part system). But to force this conclusion through would require detailed evidence against a recapitulation theory, of a sort that could only perhaps come from acquisition studies.

Received 28 April 1993
Revised version received
15 April 1994

Max Planck Institute for
Psycholinguistics, Nijmegen

Notes

1. Paper delivered to the Workshop on Spatial Description in Mayan Languages, Nijmegen, Feb. 10-13, 1992, and thereafter circulated in a longer version than the current article as a working paper of the Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics. At the time, this was an exploitation of Marr's 3D model for linguistic purposes independent of the important paper by Ray Jackendoff and Barbara Landau (1992 [1991]; Landau and Jackendoff 1993), and indeed we have made different things of it. Comparing the views would take another paper. Suffice it to say that they want to make a general point about how little of the

axial geometry of objects, let alone contour information, figures in spatial relators as opposed to object names; while I thought that I happened to have found almost by chance (though prompted by Jackendoff 1987) a ready made metalanguage for an exotic spatial relator form class in a Mayan language, which might help one more generally understand the shape preoccupations of Mesoamerican languages. There is, it turns out, some tension between the views: (a) I thought the interest of Marr's 3D model was precisely how uniquely appropriate it was to Tzeltal body parts and how inappropriate for English prepositions (or indeed other parts of speech), while J & L hope to extend it into a general spatial model; (b) Tzeltal seems to allow more shape specification in spatial relators than J & L supposed would occur – on the other hand as far as the body part terminology is concerned, the differences are perhaps more of degree (number of object axes involved) than of kind (direct contour specification, although that also occurs). I have not attempted here any serious adjudication between these positions, which would in any case require a better articulation of theory and a better framework for linguistic comparison than is currently available (but see Brown, this issue); instead I have made numerous cross references to their work. I am also grateful to both authors for conversational clarification of their views. Other debts are mentioned in the next note. Correspondence address: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.

2. This paper forms part of an ongoing joint project with Penelope Brown concerning the Tzeltal system of spatial description, with data collected jointly in 1990, 1991, and 1992–1993 and singly by her in 1992. Summaries of the work may be found in Brown and Levinson (1990, 1992, 1993), Levinson and Brown (1994), and especially Brown (1991). I am doubly indebted to Penelope Brown here, since she proved a harsh, if ultimately sympathetic, critic of these ideas. Lourdes de *Leon* by consistently worrying about body parts in Tzotzil (see de Leon 1991b, 1992a, 1992b, 1993) made me turn my attention to the Tzeltal ones. The paper also owes much to discussion with Melissa Bowerman, John Haviland, Wolfgang Klein, Pim Levelt, John Lucy, and Dan Slobin (although they may not recognize their contributions). John Haviland and David Irwin gave me detailed comments on a draft; as did Mary Peterson, but at a point when I could only insert covering footnotes. Thanks to them all for bibliographical leads, and to Penelope Brown for checking some of the Tzeltal data in the field in 1992 on my behalf. Incidentally, the data on which this paper is based consist of (a) many hours of videotaped elicitation about the names of parts of objects, together with videotapes of spatial tasks requiring verbal interaction between two native speakers, (b) field notes, (c) some text searches in a large corpus of Tzeltal texts (Brown n.d.) (but more needs to be done). Unfortunately, only after writing the paper am I clear about the crucial further data that would be needed to test rigourously the hypotheses here advanced. That is the lot of the anthropological linguist: more work is always in order.

3. See Talmy (1983). Similarly, Jackendoff and Landau (1992) and Landau and Jackendoff (1993) emphasize the poverty of shape information in the prepositions of familiar European languages.

4. Of course this is at best an impressionistic demonstration of the richness of Tzeltal shape discrimination. Jackendoff and Landau (1992 [1991]) estimate that there are 10,000 names for things, applicable on the basis of shape, in the average English working vocabulary. But the Tzeltal lexical richness for shape is of a different kind: (a) the expressions denote shape properties, not names for things; (b) the shape specificity is built into roots – but the total root vocabulary of Tzeltal is relatively small, say of the order of 3000, which yields by powerful derivational morphology a working

vocabulary of say 30,000 stems. Perhaps a quarter of those roots have shape constraints as fundamental elements of meaning.

5. See again Talmy (1983) (but for an insistence on the Euclidean geometry involved in English prepositions, see Herskovits 1986). Jackendoff and Landau (1992) emphasize the lack of shape information in spatial prepositions but describe this in terms of limited use of the (nontopological) axial structure of object geometry.
6. For much more detail, and also some ways of avoiding these choices, see Brown (1991). For the role of shape in expressing spatial relations of containment and support in Tzeltal, see Brown (1992, this issue).
7. DavidIrwintellsmehefindsthischaracterizationinapposite:"bottom up"emphasizes a relatively data driven mode of processing, "top down" a relatively expectation driven mode of processing. Clearly, pattern matching techniques could be relatively "bottom up"; but they are "top down" to the extent that they are template driven. Mary Peterson adds that one main divide between Man and the theorists before him was Marr's emphasis (wrong, in her view) on the necessity of deriving a 3D model before object representations can be accessed.
8. There is, I think, some confusion concerning Marr's ideas about the "modularity" of the visual system. Modularity can imply (a) bottom up reasoning only; (b) no inter action between subcomponents. Marr's 21D level is modular in the (a) sense, nonmodu lar in the (b) sense. Mar` argues that the 21D representation is derived purely bottom up; he implies that much of 3D object recognition is also bottom up but allows that recognition might be used to improve the image (1982: 321). Fodor (1983: 93) argues further that (Roschian) basic object recognition is purely bottom up, in the sense that the inventory of shapes we compare objects to is also pregiven. See discussion by Stillings in Garfield (1987: 392ff.).
9. This analysis might invite a radical nativism, which will have Roschian "basic level objects" (and the associated taxonomic hierarchy) delivered as a natural level of vision analysis; see Fodor (1983). But no such conclusion is necessary of course.
10. It may be that this view is not *retinocentric*, not strictly with coordinates radiating out from the viewer, but rather with a normalized set of parallel coordinates heading in the same direction as the viewer. Such a conversion from the retinocentric to a parallel line coordinate system is supposed to take place at the end of the visual processing at the 21D level (see Man 1982: 285). Because of saccadic eye movements, a visual represen tation has in any case to be in some way externally anchored or normalized to a location relative to the body.
11. Ray Jackendoff (personal communication) queries the assumption here that the object centered representations of individual objects are a distinct level from the representations of scarfs. It seems to me, however, that they must be distinct: 3D representations do not encode viewpoint, but a visual representation of a scene must do so. When we recognize a cup next to a teapot, we somehow combine 3D and 21D information into a higher level representation. One way to think of these two kinds of information is to relate them to the distinct neural pathways known as the "what" and the "where" systems (Landau and Jackendoff 1993); but we lack a theory about how these two distinct kinds of information are put back together into our seamless experi ence of space. In any case, as Jackendoff remarks, Mar and his descendants have given us no theory of scenes working out such a theory will be essential to understanding the vision/language interface.
12. The Tzeltal body part system involves use of more axes than, for example, English prepositions (but not perhaps more than English dimensional adjectives) but, in addi

Lion, names specific shapes at the ends of axes (pointed, rounded, flat, gently convex, etc.), as explained below.

13. This section is a brief and incomplete exposition; we Brown (1991, 1992), Levinson (1991a, 1991b, and references therein) for more detail.

14. Talmy (1983) claims that closed-class spatial expressions universally specify characteristics of the ground and not the figure. Jackendoff and Landau (1992: 109ff.) concur but emphasize "how coarsely both the figure and reference objects appear to be represented" in terms of shape. Tzeltal appears to challenge such observations (see Levinson 1991 b; Brown this issue), but much depends on what we are to count as "closed-class" form classes or "spatial relators."

15. See Brown (1991: 36ff., this issue).

16. I am here using the terms "relational noun" and "body part" contrastively; in much of the Mayanist literature the former is used inclusively of the latter. The terminology is immaterial: the point that is important is that in Tzeltal (although not perhaps in neighboring languages) the two form classes can be distinguished, although at a higher level of analysis they share most syntactic properties.

17. I know of two attempts to outline noun classes in Tzeltal. Brent Berlin in a manuscript of 1961 sketches eight classes for Tenejapan Tzeltal: (1) taking plural -etik directly, (2) taking -Vl-etik plurals, (3) obligatorily prefixed with x- or j- (mostly names), (4) obligatorily taking -Vl when unpossessed (mostly our class I), (5) taking plural -tak (e.g. kin terms), (6) deverbal nouns, (7) place names, (8) pronouns and perhaps relational nouns. Kaufman (1971: 106) isolates six classes in the distinct dialect of Aguacatenango Tzeltal. Our class I probably corresponds to Kaufman's class 3; he distinguishes 3a (essentially kin terms) as taking plural -ab when possessed, 3b as taking -iak. Probably the most useful lead is Haviland's (1988) correction of Laughlin's Tzotzil noun classes.

18. Penelope Brown tells me that when unpossessed, the meaning of the term is generic, 'X in general'; for example jolol'heads in general; the class of heads'.

19. This list is almost certainly not as ad hoc as it seems; but that would require some Whorfian exegesis. Stross (1976: 244-245) fists also bilil 'name', betil 'debt', ti unibil 'planting-seed', akie'al'staff of office', we'elil'food/meal', ch'ulelal'soul', labil'animal spirit companion'. In short: a person's physical and metaphysical parts and social identity. Class I nouns are further subdif erentiated, however, by other morphological criteria; for example, kin terms take irregular plural inflections, so that the body-part terms could possibly be fractionated out as a coherent subclass (see note 13).

20. Before a palatal x or ch, s- becomes palatalized x-; before a root beginning in s or x, the prefix is naturally not audibly detectable and will usually not be marked here. Before j the form of the third person prefix is y-, with the sequence y-j realized as y.

21. There is the possessed form s-ba-al 'contents', but if this is the same root, then it clearly shows that s-ba doesn't belong to the class that must drop -Vl when possessed.

22. So catalogued by Stross (1976); our informants preferred other designations, for example yalalil sk'ab 'children of the hand/arm, i.e. fingers', sme'sk'ab 'mother of the hand/arm, i.e. thumb'.

23. Levelt (1984), considering English, correctly pointed out that English terms like in front even when used intrinsically presuppose that the object is in canonical position (relative to the vertical) relative to the frame of orientation. Another interpretation of his careful observations would simply be that the object-centered perspective and the vertical absolute coordinate are deeply interwoven in English - just as they are precisely disentangled in Tzeltal.

24. Of course, there may be additional technical terms for the parts of specialist artefacts,

but we exclude consideration of these here, for we are interested in the generative partitioning of objects as used in locative expressions, and such technical terms play no role here.

25. *The application of "metaphorical" extensions between domains I, II, and III is considerable, and it is often by no means clear which domain should be thought of as originating the term. For example, the term for 'eye' may derive from I or II or even III (where it means 'seed, fruit'). Sometimes there are good reasons for supposing the origin in one domain, for example s-ne 'tail' from II, y-(j) abenal'leaf' from III. Terms that may ricochet between I, 11, and III, but do not reach IV, are not properly examined here.*
26. *Stross (1 976: 25 7) gives the forms e'id and y-e; we hear an aspiration and give the root the form ej.*
27. *The term for fruit is sit (no form sill!). There is a widespread Mesoamerican areal association of 'eye' and 'fruit' or 'seed'. In the light of this, it is possible that originally there was a vegetable metaphor for body part here.*
28. *y-ok may have primary reference to the base of a plant, rather than to the supporting limbs of humans and animals, to which it also applies, meaning 'lower extremities'. If so, we would expect the absence of the form okil; our data has a gap here (as does Stross 1 976: 260). On the other hand note the expression y-ok sk'ab 'lower extremities and hands, i.e. limbs'.*
29. *Clearly not an animate body part; we doubt there is a form xujkil. Semantically, the term behaves more like a relational noun; informants deny there is any part of the body to which the term intrinsically refers. Stross (1 976: 266) concurs: he glosses the term 'side/summit' and notes that it has secondary application to the body in the phrase xujk sitil 'side of the eyes, i.e. temple'. On the other hand, there is a bona fide body-part terra xujkubil/xujkub'elbow'. Note the expression pat xujk'backside'—'neighboi .*
30. *y-uxub may in fact have primary reference to 'knot' (in wood) and metaphorical reference to the navel. Our data is unclear whether there is any form uxubil, as predicted if it were a genuine body part. Stross (1 976) gives muxuk'-il for the navel. Berlin et al. (1 974 : 76) assume that it is a genuine body part.*
31. *Used like me 'tail' to describe attached flexible cords or the like.*
32. *Note that this is not a primary part term at all by semantic or morphological criteria: it is the root for 'tree' with a -VI suffix in the possessed form. Nevertheless it occurs in compound body-part terms, like steel sk'ab'trunk of the hand/arm, i.e. foream.*
33. *This term belongs to an alienable morphological class, since it takes -VI only when possessed. It is the normal term for heartwood, bone, etc. The term for 'seed, pit', bak', has a "glottalized" final segment but may be etymologically related. There are a range of -VI forms with specialized meanings, for example s-bake! 'core of corn cob'.*
34. *For a general account of the metaphorical nature of the transfer from body parts to spatial relations, with special attention to African languages, see Heine et al. (1 991 : 45ff., 123l1f).*
35. *The extent to which Mesoamerica constitutes a linguistic area in Emeneau's sense especially with respect to spatial description is an interesting theme (see de Leon and Levinson 1 992).*
36. *See section 4.5.3 below. Mary Peterson points out to me that, nevertheless, this marginal functional information may be important to the understanding of the process of term application, since it may need to take priority over shape analysis. To achieve this, it may be necessary to process in parallel, checking structural representations of familiar objects as wholes while simultaneously proceeding with shape analysis of parts.*
37. *The task was administered by Penelope Brown on my behalf, to whom many thanks.*

The names of the parts in the figures were as given by informants. The figures reproduced are the original stimuli but the divisions have been redrawn by me.

38. David Irwin's response to these animate segmentation diagrams was that they were more suggestive of butchery than visual segmentation! But a glance at universals of body partition is instructive (Andersen 1978): with rare exceptions (like Tarascan 'nose + forehead'), human body part terms reveal a segmentation based on junctions between articulated elements, in line with segmentation by concavity. Andersen notes too that it is shape, not function, that underlies the naming systems, as revealed in etymology and acquisition.
39. This perhaps argues for a human model here, but see below where the 'belly' term is shown to be applied without reference to a particular metaphorical source.
40. Incidentally, Tenejapan Tzeltal speakers are conservative with respect to material culture; many standard Western goods (tinned foods, plastic bottles, etc.) are not in daily evidence. An exception is bottled Pepsi-Cola, which has a ritual significance and is omnipresent.
41. I am indebted to Melissa Bowerman for drawing my attention to this theory.
42. Leyton's theory could be augmented by parallel observations about really sharp convexities, that is, points. These are growing and moving points and in general indicate the direction of motion for good physical reasons, to do for example with whiplash in winds in the case of plants and aero- or hydrodynamics in the case of animals, birds, and fish.
43. Jackendoff and Landau (1992) address the issue of the proper treatment of hollow entities in conceptual structure.
44. I am a little bit worried that this may not correctly assign parts to drinking horns, which may have both a s-ni"nose' and a s-ti"mouth', but this needs checking. Perhaps a drinking horn is biheaded, like planks, etc. One clue is that de Leon found Tzotzil informants had alternative conceptualizations of a pop bead, an item with a clear s-ni'-like protrusion at one end and an equal-sized hole at the other: some informants called (in the Tzotzil equivalent terms) the protrusion a s-ni' and the hole a y-it ('buttocks'), while others called the hole a s-ti' and the protrusion a y-it – in short, they couldn't decide which way to run the axis.
45. There is one apparent exception to this rule, which follows a general tendency: some specialist words preempt the general-purpose labels. The exception here is the term y-*ok*, not much used, which in the domain of animal partonymy refers to 'undercarriage, lower extremities'. There is some doubt that it is a genuine body-part term, as it perhaps does not belong to nominal class I (see note 28 to Table 2). In any case, it seems to presume that the base of the model axis is vertical, orthogonal to, and "planted in" the ground – thus the junction between a tree and the earth or a mountain and a valley can be called *y-ok*. When a tree is felled, it no longer has *a y-ok*, only *a pit*. My analysis of this, and similar phenomena elsewhere, is that *y-ok* is a term belonging to other, non-object-centered systems of description, applied secondarily after primary parts have been labelled.
46. On this analysis the term s-ti' refers to the gap or void itself. This is correct for one use. But it also refers to an edge – this nomenclature is generated quite independently. Many terms on this analysis are polysemic or even homonymic. But that seems right: a door and a door frame and a door void are not the same thing. To account for win the same phonological form has these contents, one may indeed need to refer to metaphor in etymology; but the original metaphor may not even have taken place in a Mayan language at all (see section 5.2.2).
47. It would be possible to assign terms on the basis of shape without this step, but

retaining it will (a) maintain the symmetry of the account between model and subsidiary axes, (b) correctly suggest an order of importance between the two main facets determined on the basis of the secondary axis.

48. Why make the head of the arc point to an object's 'back' and not its 'face'? An important observation is that the 'back' facet is dearly more basic than the opposing facet, as shown by (a) the various names of the latter, and (b) the tendency to choose the 'back' label when only one facet can be named. It is interesting that this conceptual priority holds also for English *back* over *front* when applied intrinsically to objects (Levine and Carey 1982: 6541.), contrary to naive expectation. Applying the same reasoning, though, I should perhaps reverse the direction of the head of the model axis, since the unitary term *y it* 'butt' applies to what I called the base of the arc. But the conceptual priority of pit to its various opposed terms (s ni', *s jol,* etc.) is intuitively less clear.
49. From my drawings, which outline the frame of the house, one might expect *s ni'* for the apex, but it is in fact well rounded by thick thatch, hence *sjol is* the appropriate term.
50. The fact that the main axis is vertical fits the positional predicate used for houses, P. Brown points out to me: they are said to be *waxal,* standing, as of vertical cylinders (e.g. a bottle).
51. This downplays the functional element in this naming pattern of course: the front of the house gives public access, etc. In general, the geometrical account pushes pure structure to the limit, but of course this does not rule out additional or even preemptive functional factors.
52. In some cases, 2D rectangles offer the possibility of some additional terms for sections of the perimeter: for example a concrete coffee drying patio oriented north south may have the northern 'downhill' edge labelled *y ejta(,* the southern 'uphill' edge labelled s *ba* using these relational nouns. These terms then preempt *s ti'* on the relevant facets, leaving s ti'to apply to the two remaining sides. Or the remaining sides may be labelled *ti ee! (a* non body part term). It would be easy but, I believe, mistaken to think of these terms *(s ba, s ti,* etc.) as forming a contrast set they belong to two different systems, as sketched at the bottom of Figure 16.
53. John Haviland questioned, on the basis of neighbouring Tzotzil, whether the present account would work for the traditional stone *met ate,* a rectangular stone with pitted upper surface, almost horizontal on three knoblike feet protruding from the under surface. It turns out that the account works on at least one application of the terms: the main axis of the rectangle has *s jot* at either end, the secondary axis points to two surfaces *s pat* (the under surface) and *y elaw* (the upper one), with the edges labelled either *s ti'* (emphasizing a 2D interpretation) or *xujk* (emphasizing a 3D one), while it has three *y-akan* or 'feet'. This is in line with the predictions, although on both my account and the metaphorical account one might not expect 'feet' to emerge from the surface labelled' back'. But alternative nomenclature has double pit instead of sjol (or one of each), and *s-k'ab* ('hand') or *x cku'* ('breast') for the knoblike feet (where one might on shape have expected even s ni'). It is perhaps not surprising, on the present account, that some such central cultural objects exhibit a more conventional and arbitrary (if variable) nomenclature. Perhaps this may be more problematic for a metaphorical account, where terminology for central cultural objects might be taken to be prime exemplars of correct metaphorical application.
54. A middle way is suggested to me by Mary Peterson: a process of visual analogy, searching for structural similarity, might account for the origin of the system. If such a process operated pairwise over volumes, it might allow, for example, for multiple 'noses' or 'ears' on main volumes. Such an account wouldn't, I think, do as a syn-

chronic analysis of the Tzeltal system, because 'noses' or 'ears' can be at the 'butt' end of main volumes.

55. This requires explanation: Tenejapans don't, at least normally, keep pigs or sheep, unlike neighboring peoples. They keep bulls. No cows (or only as an innovation). They buy in the male calves from neighboring peoples and sell them out as mature adults or, better, eat them. Bulls are beautiful. And the proper stuff for ritual meals.

56. There would be no scruples; these are a no nonsense people who have clan names glossing 'dog shit', 'snake ass', and the like, and talk freely about human parts. I know of only one use of the term x chu' as a part name the feet of a *metate may* be so called (see note 53 above).

57. But caution is in order! To test the received view that metaphorical abilities are late acquired, Gentner (1977) studied English speaking preschool children' abilities to analogically map body parts onto objects. She concluded that four year olds (the youngest children tested) have sophisticated metaphorical abilities, and indeed that such abilities may play a role in language acquisition from the outset. However, there is also strong evidence that children learn the intrinsic parts of objects well before three years old, that is, before they can correctly label them (Levine and Carey 1982). This latter evidence I take to be very suggestive that a geometric analysis is also operative in English intrinsic 'front', 'back', 'top', 'bottom' assignments (although unlike Tzeltal these also involve reference to the absolute vertical axis).

58. The facts are actually rather complex: the distance a reference (figure) object can be from the specified part of a ground object depends on contextual factors. What is clear is that as the figure object is progressively moved away from the ground object so that it becomes closer to another object, speakers will always switch the ground to the second object even if it is much less prominent, thereby indicating a distaste for elastic regions around the original ground. The result is that the acceptable distance increases with the lack of availability of any other reference object, and with the inapplicability of the relational nouns that do primarily describe regions. See Brown (1991) for details.

59. This contrast between the 20 odd body parts with generalized locative uses and the 200 odd positional predicates could be held to vindicate the claim in Landau and Jackendoff (1993) that *where* specifications tend to be shape impoverished (Ray Jackendoff, personal communication). However, as Brown (this issue) argues, it is not clear that one can disentangle locative functions from the 200 odd positional roots in question.

60. John Haviland (work in progress) has shown this also for the closely related language Tzotzil: by starting with a subset of positional roots delimited by derivational potential, he has isolated a class that makes essential reference to positioning of body parts.

61. Informants distinguish *tekel (of* trees) from *tek'el* (of humans), but not always consis tently enough for one to be really sure there is a distinction here (P. Brown thinks there is).

62. Also applicable to human "standing" on hands and knees.

63. I am indebted to Josh Antonio Flores for information about the Nahuatl patterns. See also Campbell et al. (1986) on areal patterns.

64. Brown (1992 and this issue) describes how topological relations expressed in English with topological prepositions am (partially) encoded in Tzeltal.

65. From the point of view of coordinate systems, or "frames of spatial reference," Tzeltal has only two systems: absolute and intrinsic coordinates. The viewer centered systemm as encoded in verbs of motion (or deictic demonstratives) is coordinate free: it specifies motion (or location) in any direction toward *or away* from speaker.

66. See Levinson (1986, 1992). Levelt, though, was incorrect to think that all languages have just the two basic systems exhibited by English.
67. An alternative interesting conception might be that during development what start out as fully encapsulated "modules" must be slowly interconnected, allowing increased interaction. Thus "unnatural" access to visual processes must be slowly learned. This view would be consistent with the view that only a modular theory of mind can explain the initial speed of learning, while not subscribing to a strictly modular theory of adult cognition. Normally, of course, the reverse presumption is made, namely that child hood plasticity hardens into adult modularity (see e.g. Levelt 1989: 105).
68. For experimental evidence that verticality offers preferred interpretations to object centered ones in English see Carlson Radvansky and Irwin (1993). Incidentally, they found that when viewer centered and environment centered interpretations conflict (as when ego is lying down), the latter predominate.
69. Jackendoff tells me, though, that he conceives of 3D descriptions of objects as coNSTrr *uENTS* of perspectivally viewed scenes.
70. Another possibility is that the facts all argue for strictly (modular) visual processes ceasing well before a 3D level of analysis (as e.g. in Btllthoff 1991).

References

Andersen, Elaine S. (1978). Lexical universals of body part terminology. In *Universals of Human Language,* vol. 3: *Word Structure, J.* Greenberg (ed.), 335 368. Stanford, CA: Stanford University Press.

Berlin, Brent (1968). *Tzeltal Numeral Classifiers: A Study* in *Ethnographic Semantics.* The Hague: Mouton.

Breedlove, Dennis; and Raven, Peter H. (1974). *Principles of Tzeltal Plant Class cation.* New York: Academic press.

and Kaufman, Terrence; with Maffi, Luisa (1990). *Un diccionario basico del Tzeltal de Tenejapa, Chiapas, Mexico.* Procomith Project, University of California, Berkeley.

Biederman, Irving (1987). Recognition by components: a theory of human image under. standing. *Psychological Review* 94(2),115 147.

Bowden, John (1991). Grammaticalization of locatives in Oceanic languages. Unpublished MA thesis, University of Auckland.

Brown, Penelope (1991). *Spatial Conceptualization in Tzeltal.* Cognitive Anthropology Research Group Working Paper 6. Nijmegen: Max Planck Institute for Psycholinguistics.

(1992). The INs and ONs of Tzeltal locative expressions: speaker's perspective and the semantics of static descriptions of location. Paper presented at the Workshop on Spatial Description in Mayan Languages, February.

(n.d.). Computerized corpus of Tzeltal texts. Unpublished data.

and Levinson, Stephen C. (1990). *"Recentering" in Mayan Spatial Description, with Special Reference to Tzeltal.* Cognitive Anthropology Research Group Working Paper. Nijmegen: Max Planck Institute for Psycholinguistics.

and Levinson, S. C. (1992). 'Left' and 'right' in Tenejapa: investigating a linguistic and conceptual gap. In *Space* in *Mesoamerican Languages,* Lourdes de Leon and Stephen C. Levinson (eds.), 590 611. Special issue of *Zeitschrift f Phonetik and Kommunikationsforschung.* Berlin: Akademie Verlag.

; and Levinson, Stephen C. (1993). 'Uphill' and 'downhill' in Tzeltal. *Journal of Linguistic Anthropology 3(1),* 46 74.

Bruce, Vicky; and Green, Patrick (1985). *Visual Perception*. Hillsdale, NJ: Erlbaum.
Brugman, Claudia (1983). The use of body part terms as locatives in Chalcatongo Mixtec. *Survey of California and Other Indian Languages 4*, 235 290.
BUlthoff, Heinrich H. (1991). Shape from X: psychophysics and computation. In *Computational Models of Visual Processing*, Michael S. Landy and J. Anthony Movshon (eds.), 305 330. Cambridge, MA: MIT Press.
Campbell, Lyle; Kaufman, Terrence; and Smith Stark, Thomas (1986). Meso America as a linguistic area. *Language* 62(3), 530 570.
Carlson Radvansky, Laura; and Irwin, David (1993). Frames of reference in vision and language: where is above? *Cognition 46*, 223 244.
Casad, Eugene H.; and Langacker, Ronald W. (1985). 'Inside' and 'outside' in Cora grammar. *International Journal of American Linguistics* 51(3), 247 281.
de LeOn, Lourdes (1991 a). *Space Games in Tzotzil: Creating a Context for Spatial Reference*. Cognitive Anthropology Research Group Working Paper 4. Nijmegen: Max Planck Institute for Psycholinguistics.
(19916). Strategies in the acquisition of spatial orientation in Tzotzil. Draft paper, Cognitive Anthropology Research Group. Nijmegen: Max Planck Institute for Psycholinguistics.
(1992a). Locative body parts and geographic anchoring in Tzotzil acquisition. Paper presented at the Child Language Conference, Stanford University, April.
(1992b). Body parts and location in Tzotzil: a case of grammaticalization. In *Space in Mesoamerican Languages*, Lourdes de Le6n and Stephen C. Levinson (eds.), 570 589. Special issue *of Zeitschrift f Phonetik and Kommunikationsforschung* 45(6). Berlin: Akademie Verlag.
(1993). Shape, geometry and location: the case of Tzotzil body part terms. In *CLS 29: Papersfrom the Parasession on Conceptual Representations*. Chicago: CLS.
; and Levinson, Stephen C. (eds.) (1992), *Space* in *Mesoamerican Languages*. Papers from the conference on Space in Mesoamerican Languages at the Cognitive Anthropology Research Group. Special issue of *Zeitschrift far Phonetik, Sprachwissenschaft and Kommunikationsforschung 45(6)*, 527 641. Berlin: Akademie Verlag.
Fodor, Jerry (1983). *Modularity of Mind*. Cambridge, MA: MIT Press.
Garfield, Jay (ed.) (1987). *Modularity in Knowledge Representation and Natural Language Understanding*. Cambridge, MA: MIT Press.
Gentner, Deirdre (1977). Children's performance on a spatial analogies task. *Child Development 48*, 1034 1039.
Gibson, Bradley S.; and Peterson, Mary A. (i.p.). Does orientation independent object recognition precede orientation dependent recognition? Evidence from a cuing paradigm. *Journal of Experimental Psychology: Human Perception and Performance*.
Haviland, John B. (1988). It's my own invention: a comparative grammatical sketch of Colonial Tzotzil. In *The Great Tzotzil Dictionary of Santo Domingo Zinacantdn*, Robert Laughlin (ed.), 79 121. Washington: Smithsonian Institution Press.
Heine, Bernd (1989). Adpositions in African languages. *Linguistique Africaine*, 77 127.
; Claudi, Ulrike; and HUnnemeyer, Frederike (1991). *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
Herskovits, Annette (1986). *Language and Spatial Cognition*. Cambridge: Cambridge University Press.
Hoffman, Donald D.; and Richards, Whitman (1984). Parts of recognition. *Cognition* 18, 65 96.
Humphreys, Glyn; and Bruce, Vicky (1989). *Visual Cognition*. Hillsdale, NJ: Erlbaum.
Hunn, Eugene (1977). *Tzeltal Folk Zoology*. New York: Academic Press.

Jackendoff, Ray (1987). *Consciousness and the Computational Mind.* Cambridge, MA: MIT Press.

; and Landau, Barbara (1992 [1991]). Spatial language and spatial cognition. In *Languages of the Mind,* Ray Jackendoff, 99 124. Cambridge, MA: MIT Press.

Kaufman, *Terrence* (1971). *Tzeltal Phonology and Morphology.* University of California Publications in Linguistics 61. Berkeley: University of California.

Lakoff, George (1987). *Women, Fire and Dangerous Things.* Chicago: University of Chicago Press.

Landau, Barbara; and Jackendoff, Ray (1993). 'What' and 'where' in spatial language and spatial cognition. *Behavioural and Brain Sciences* 16, 217 265.

Levelt, Willem J. M. (1984). Some perceptual limitations on talking about space. In *Limits of Perception,* Andrea L van Doorn, Wim A. van der Grind, and Jan J. Koenderink (eds.), 323 358. Utrecht: VNU.

(1989). *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

Levine, Susan; and Carey, Susan (1982). Up front: the acquisition of a concept and a word. *Journal of Child Language 9,* 645 657.

Levinson, Stephen C. (1986). The semantics/pragmatics/kinetics of space in Guugu Yimidhirr. Paper delivered to symposium at the University of Bamberg.

(1991a). Figure and ground in Mayan spatial description: Tzeltal locative descriptions. Paper delivered to the conference "Time, Space and the Lexicon," November, Max Planck Institute for Psycholinguistics, Nijmegen.

(1991b). Relativity in spatial conception and description. Paper delivered at Symposium no. 112, "Rethinking linguistic relativity," May 3 11, Ocho Rios, Jamaica. Wenner Gren Foundation for Anthropological Research. (Cognitive Anthropology Research Group Working Paper 1. Nijmegen: Max Planck Institute for Psycholinguistics.)

(1992). Primer for the field investigation of spatial description and conception. *Pragmatics 2 (l),* 5 47.

; and Brown, Penelope (1994). Immanuel Knot among the Tenejapans: anthropology as empirical philosophy. *Ethos* 22(l), 3 41.

Leyton, Michael (1989). Inferring causal history from shape. *Cognitive Science* 13, 357 387.

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

Mart, David (1982). *Vision.* New York: Freeman.

Piaget, Jean; and Inhelder, Baerbel (1956). *The Child's Conception of Space.* London: Routledge and Kegan Paul.

Pinker, Steven (1985). *Visual Cognition.* Cambridge, MA: MIT Press.

Stross, Brian (1976). Tzeltal anatomical terminology: semantic processes. In *Mayan Linguistics,* vol. 1, Marlys McClaran (ed.). UCLA: American Indian Studies Center.

Svorou, Soteria (1986). On the evolutionary paths of locative expressions. *Proceedings of the Berkeley Linguistics Society* 12, 515 527.

Talmy, Len (1983). How language structures space. In *Spatial Orientation: Theory, Research and Application.* Herbert Pick and Linda P. Acredolo (eds.), 225 320. New York: Plenum.