ELSEVIER

Cognition 84 (2002) 155–188

COGNITION

www.elsevier.com/locate/cognit

Returning the tables: language affects spatial reasoning

Stephen C. Levinson[a,*], Sotaro Kita[a], Daniel B.M. Haun[a], Björn H. Rasch[b]

[a]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*
[b]*Department of Psychology, University of Trier, Trier, Germany*

Received 27 September 2001; received in revised form 2 February 2002; accepted 8 March 2002

Abstract

Li and Gleitman (Turning the tables: language and spatial reasoning. *Cognition*, in press) seek to undermine a large-scale cross-cultural comparison of spatial language and cognition which claims to have demonstrated that language and conceptual coding in the spatial domain covary (see, for example, *Space in language and cognition: explorations in linguistic diversity*. Cambridge: Cambridge University Press, in press; Language 74 (1998) 557): the most plausible interpretation is that different languages induce distinct conceptual codings. Arguing against this, Li and Gleitman attempt to show that in an American student population they can obtain any of the relevant conceptual codings just by varying spatial cues, holding language constant. They then argue that our findings are better interpreted in terms of ecologically-induced distinct cognitive styles reflected in language. Linguistic coding, they argue, has no causal effects on non-linguistic thinking – it simply reflects antecedently existing conceptual distinctions. We here show that Li and Gleitman did not make a crucial distinction between frames of spatial reference relevant to our line of research. We report a series of experiments designed to show that they have, as a consequence, misinterpreted the results of their own experiments, which are in fact in line with our hypothesis. Their attempts to reinterpret the large cross-cultural study, and to enlist support from animal and infant studies, fail for the same reasons. We further try to discern exactly what theory drives their presumption that language can have no cognitive efficacy, and conclude that their position is undermined by a wide range of considerations. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Language; Spatial reasoning; Linguistic relativity

* Corresponding author. Max Planck Institute for Psycholinguistics, P.O. Box 310, NL-6500 AH Nijmegen, The Netherlands.

E-mail address: stephen.levinson@mpi.nl (S.C. Levinson).

0010-0277/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0010-0277(02)00045-8

1. Language and thought in the spatial domain

There seem to be two main currents of speculation about the relationship between linguistic systems and other conceptual systems. One line assumes that language is merely an input/output system for an innately grounded 'language of thought', so that a language either directly reflects an antecedently available pool of universal concepts (Fodor, 1975) or it builds on a rich, core set of 'natural' concepts constituting a universal conceptual base (Landau & Jackendoff, 1993; Pinker, 1994).[1] The other, noting that language is a human prerogative, suggests that the possession of language in general, and specific languages in particular, may reorganize and restructure the underlying cognition even in domains such as space that have been considered 'natural' and 'universal'. The role of language in restructuring thought may then account for some of the special properties of human thinking (Dennett, 1991; Lucy, 1992a; Spelke & Tsivkin, 2001).

There has been a recent resurgence of interest in this second possibility (see, for example, Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, in press). Our own work has been dedicated to exploring this possibility empirically in the spatial domain. Spatial thinking is essential to any animal, and it is a domain where one might expect the strongest biological basis and most conceptual uniformity. But it turns out that there is in fact a great deal of cross-cultural variation in the semantic relations and categories of spatial language. Moreover, in correlation with those language-specific relations and categories, the same or similar distinctions can be shown to play a role in non-linguistic memory and reasoning tasks (Levinson, 1996b; Pederson et al., 1998). We interpret this as evidence in favor of the second position.

In a paper in this issue, Li and Gleitman (in press) try to defend an extreme version of the first position, arguing that thinking is independent of, and impervious to, the details of linguistic coding. "Linguistic systems are *merely* the formal and expressive medium that speakers devise to describe their mental representations" (p. 290, their emphasis), for "linguistic categories and structures are more-or-less straightforward mappings from a preexisting conceptual space, programmed into our biological nature", with the consequence that "all languages are broadly similar"(p. 266). To maintain this thesis, they target our cross-cultural comparison of spatial language and cognition which appears to demonstrate that language and conceptual coding in the spatial domain covary, with the apparent implication that different languages induce distinct conceptual codings (see, for example, Levinson, in press; Pederson et al., 1998). Li and Gleitman's strategy is to try and show that in an American student population they can obtain any of the relevant conceptual codings just by varying spatial cues, holding language constant. They then argue that our findings are better interpreted in terms of ecologically-induced distinct cognitive styles reflected in language. Linguistic coding, they argue, has no cognitive efficacy or cognitive effects – it simply reflects antecedently existing conceptual distinctions.

[1] Thus, Pinker (1994, p. 82) presumes "a universal mentalese", with the corollary "Knowing a language, then is knowing how to translate mentalese into strings of words and vice-versa. People without a language would still have mentalese, and babies and many nonhuman animals presumably have simpler dialects". Landau and Jackendoff (1993, p. 235) argue that the universal properties of spatial conception should directly reflect in language so that we "should find broad similarities in the expression of object and place across languages", especially in closed-class systems of morphemes.

In this response, we concentrate on the specific issues raised by Li and Gleitman's critique of our empirical work. We show that Li and Gleitman did not make the essential conceptual distinctions in this domain – specifically between the different kinds of spatial frames of reference. We try to show that, as a consequence, they have misinterpreted their own experiments. To do this, we replicate their experiments with crucial variants to demonstrate that their results, taken together with our new results, are in fact consistent with the hypothesis of a language-driven preference for conceptual coding. Their attempt to enlist animal and infant studies fails for the same reasons, namely through not making the right distinctions in the studies of frames of reference. The same analytical problems undermine their attempt to reinterpret our large cross-cultural survey. We conclude that our findings stand: there is a demonstrable correlation between the frames of reference used in language and those used in non-linguistic conceptual coding, and the most plausible interpretation is that speaking a specific language can induce specific patterns of non-linguistic conceptualization.

Our response must focus on the fundamental conceptual issues involved in the study of spatial frames of reference, but readers should know that the essential phenomenon that provoked our investigations is the following. In a nutshell: there are human populations scattered around the world who speak languages which have no conventional way to encode 'left', 'right', 'front', and 'back' notions, as in 'turn left', 'behind the tree', and 'to the right of the rock'.[2] Instead, these peoples express all directions in terms of cardinal directions, a bit like our 'East', 'West', etc. Careful investigation of their non-linguistic coding for recall, recognition, and inference, together with investigations of their dead-reckoning abilities and their on-line gesture during talk, shows that these people think the way they speak, that is, they code for memory, inference, way-finding, gesture and so on in 'absolute' fixed coordinates, not 'relative' or egocentric ones (the full details can be found in Levinson (in press), but the studies are now being replicated across the world by other scholars; see, for example, Wassmann and Dasen (1998)). The phenomenon should be of fundamental interest to cognitive science as showing human variability where least expected, and should not be lost sight of in disagreements about its correct interpretation.

We proceed in the following way. First, we outline the distinctions essential to the study of spatial frames of reference, and the reasoning behind our work. We also summarize our results that Li and Gleitman attempt to undermine in their paper (Section 2). We then critically review Li and Gleitman's experiments and the thinking behind them (Section 3). We then empirically show (Section 4), through variants of those experiments, that Li and Gleitman have not correctly interpreted their own results, in large part because they did not make the essential distinctions in frames of reference. Their results seem in fact largely in line with our hypotheses. Then, we further discuss the crucial distinction in frames of reference that Li and Gleitman did not make, and the implications of not making that distinction (Section 5). Finally (in Section 6), we query Li and Gleitman's theoretical

[2] This is not a matter of preferential coding, as Li and Gleitman from time to time imply. These languages simply lack, for example, any straightforward way of coding a ternary relation 'x is to the left of y from vantage point v' as in "The ball is left of the tree". Most of them also provide no coding for the simpler notion 'at my left', or even 'my left side'.

motivation – it is hard to find a plausible theory under which the language one speaks would have no impact on the way one thinks.

2. Cross-cultural studies of frames of reference in language and cognition

The idea of distinct 'frames of spatial reference' is understood to imply the use of underlying *coordinate systems* built on different principles (not to be confused with different origins for the same coordinate system).[3] Although the details vary widely, our linguistic investigations show that there are three main over-arching types or families of such system used in languages, which we have called relative, intrinsic and absolute, the logical and spatial properties of which can be precisely delineated (as sketched below, but see Levinson, 1996b, in press). In the relative frame of reference, objects are located in terms of viewer-centered coordinates based on body axes (left/right/front/back), as in 'The ball is to the left of the chair'. In the intrinsic frame of reference, the location is described in terms of the object-centered coordinates of the reference or landmark object based on 'intrinsic' facets of the object, as in 'The ball is at the chair's front'. In the absolute frame of reference, things are described in terms of coordinates based on fixed bearings or cardinal directions, centered on the reference object, as in 'The ball is north of the chair'. Why these three? Probably they all have bases in internal, innate systems guiding most mammalian behavior (see Burgess, Jeffery, & O'Keefe, 1999, for possible brain bases). These perceptual and motoric representations are at least partially encapsulated, but nevertheless they may provide a repertoire for the development of conceptual systems, a point taken up at the end of this paper.

In an extended systematic comparison involving two score scholars and over 20 languages from over a dozen stocks, we have shown that languages vary greatly in the frames of reference they employ to describe spatial locations (see review in Levinson, 1996a; case study in Levinson, 1996b, in press, for the full details, as well as Pederson et al., 1998, the focus of the Li and Gleitman paper). What is cross-linguistically variable, from a semantic point of view, is (a) the particular conceptual details of each system (for example, the geometry of axes), and (b) the fact that specific languages select from these three frames of reference, using only one, or two or all three of them, variably (see, for example, Levinson, in press; Pederson et al., 1998; Wilkins & Levinson, in press). It was this variation in public, external linguistic representations that prompted our investigations of variation in internal representations, such as those involved in coding experience for memory.

Now, to investigate non-linguistic representations we used many techniques and many kinds of observation regarding people's sense of directionality. For example, we investi-

[3] Rock (1992, p. 404) summarizes the Gestalt definition of a 'frame of reference' as follows: "a unit or organization of units that collectively serve to define a coordinate system with respect to which certain properties of objects, including the phenomenal self, are gauged". The need to distinguish origin from coordinate system becomes especially clear in language, where the same coordinate system (for example, relative) can be used with distinct origins (for example, egocentric vs. allocentric) – see Levinson (1996b) where the many different proposals about distinctions in types of frames of reference are compared, and the synthesis that is being used here is justified in detail.

Table 1
Classifications of frames of reference

	Orientation-free	Orientation-bound	
	Allocentric		Egocentric
	Intrinsic	Absolute	Relative
Description falsified under rotation of viewer	No	No	Yes
Description falsified under rotation of Ground (i.e. reference object)	Yes	No	No

gated how gestural depiction of events is spatially oriented in a number of cultures (Haviland, 1993; Kita, Danziger, & Stolz, 2001; Levinson, in press; Wilkins, in press). We have probed directionality in the memory for real-life events that people have experienced (Levinson, 1997a). We have also examined dead-reckoning and navigational abilities in various cultures (Levinson, 1996c, in press).[4]

In order to further probe non-linguistic conceptual representations in a more controlled fashion, we also exploited the distinct logical and spatial properties that frames of reference have under rotation, and we developed a battery of experiments under which a participant is shown a stimulus on one table, then rotated 180 degrees and, for example, asked to recognize the earlier stimulus from alternates, or remake the first stimulus, on another table (first developed and reported in Levinson, 1992). The battery of tests systematically explored recall, recognition memory, and different kinds of inference (Levinson, 1996b). What such an experiment distinguishes is whether participants are or are not rotating the coordinates with themselves. It thus distinguishes between *egocentric* and *allocentric* reference frames, but it does not precisely distinguish what kind of allocentric reference frame is involved. Allocentric frames of reference include both absolute and intrinsic ones, as explained with care in Levinson (1996b, pp. 148–152). To further distinguish between these two, other tasks or collateral evidence is required. Higher level classifications of the frames of reference are explained in Table 1 (but see riders in Levinson, 1996b).

To see that rotation of the viewer makes no difference to intrinsic and absolute descriptions as opposed to relative descriptions consider the descriptions below:

(1) Intrinsic: *The ball is at the chair's front.*
(2) Absolute: *The ball is north of the chair.*
(3) Relative: *The ball is to the left of the chair (from my viewpoint).*

[4] A reviewer asks how dead-reckoning could vary with frame of reference, since dead-reckoning is by definition egocentric. The answer is that frames of reference (coordinate systems) are not equivalent to origins, egocentric or otherwise. You could dead-reckon your current position in terms of distances covered on legs of the journey after left and right turns (using an intrinsic or relative frame of reference depending on how the journey is conceived), or you could reckon your position in terms of celestial observations (using an absolute frame of reference).

Assume that each of them is true for the array from a fixed vantage point. Now walk around to the other side of the array: (1) and (2) will stay true, but (3) will now be false.

To dissociate (1) and (2) you need to carry out another rotation: let us now rotate the Ground (or landmark or reference) object, the chair – now (1) is falsified, (2) remains true and (3) remains false. All this is explained at length in Levinson (1996b).

The first rotation, of the viewer, distinguishes egocentric from allocentric coordinate systems; the second rotation, of the Ground object, distinguishes orientation-free vs. orientation-bound allocentric systems.[5] We have thus generally relied on multiple results to disambiguate both the linguistic and non-linguistic picture, for these rotations can be simply applied to investigate the spatial and logical properties of non-linguistic coding for memory and inference. We underscore these points because the Li and Gleitman experiments, described in Section 3 below, confound intrinsic and absolute frames of reference.

Our work has been based on first investigating the frames of reference utilized in the local language, then making a prediction about what frames of reference will *not* occur in non-linguistic coding – for example, we would predict that if a population uses a language where only intrinsic and absolute frames are coded, then members of that population will *not* generally use the relative frame for non-linguistic coding for memory and inference. For this, it will suffice to test the one rotation, that of the viewer. But from the absence of non-linguistic relative coding we cannot make the reverse linguistic prediction: the language may have effectively only absolute, only intrinsic, or both those frames of reference. Our predictions follow the logic of our hypothesis, that language predicts cognitive coding strategies. In contrast, Li and Gleitman want to explore primarily the non-linguistic coding of arrays in context, and for this they must disambiguate between the two allocentric frames of reference, which they failed to do and which we attempt to do for them below with a new experiment.

Only once we have a precise understanding of linguistic coding in the relevant dialect for the precise subject population do we turn to the non-linguistic experimentation. Obviously here great care has to be taken to control the verbal instructions in each native language and make sure that no verbal or non-verbal clue is present to bias the results. Delays, and verbal tasks interposed between stimulus and response, can be utilized to suppress subvocal rehearsal. Because many of these experiments were run in field conditions on uneducated peoples without written languages, they had to be relatively simple. Nevertheless the full battery of tasks involves tests for recognition, recall, inference from motion to path, and transitive inference.

The results from all of these different methods – the study of gesture, way-finding and rotation experiments on memory and inference – converge. It turns out that there are strong correlations between the frames of reference involved in linguistic tasks and those involved in non-linguistic tasks. Our investigation of gesture in various cultures reveals that where languages use predominantly an absolute or cardinal direction system, and do not use relative left/right/front/back axes, gestures preserve correct cardinal directions. For example, in an Australian aboriginal community, two natural tellings of the same story filmed at a 2 year interval preserved every orientation, for example, of a boat rolling over westwards (Haviland, 1993). This suggests that every event is coded in memory for correct

[5] There is a technical literature on these distinctions – see the discussion in Levinson (1996b, pp. 127–134).

fixed orientation. A further investigation of memory of directionality in real-life events confirms this (Levinson, 1997a). We have also examined dead-reckoning and navigational abilities in various cultures, and found that these vary with the predominant frame of reference in the language (Levinson, 1996c, in press).

The experiments involving 180 degree rotation of participants, as explained above, also show a striking correlation between the frames of reference predominant in the languages of the participants and those employed in non-verbal memory and inference tasks. Levinson (1997a) investigated speakers of Guugu Yimithirr, a language that expresses directionality based on the absolute frame of reference (this language does not have linguistic means to express directionality based on the relative frame of reference). It was found that they also used the absolute frame of reference in a number of different non-linguistic experiments based on the rotation logic. Pederson (1995) compared two dialects of Tamil speakers, one of which uses expressions based on absolute (and intrinsic) frame of reference ('absolute speakers'), and the other of which uses expressions based on relative (and intrinsic) frame of reference ('relative speakers'). Different rotation experiments revealed that the absolute speakers are more likely to give non-linguistic responses based on the absolute frame of reference than the relative speakers (see also Pederson et al., 1998). Pederson et al. (1998) compared two languages that use expressions based on the relative frame of reference, Japanese and Dutch, and three languages that use expressions based on the absolute frame of reference, Longgu, Tzeltal and Arrernte. In this study, a rotation experiment that involved recall of a row of three animals was administered. It was found that Japanese and Dutch speakers coded the row of animals based on the relative frame of reference and Longgu, Tzeltal and Arrernte speakers coded the spatial array based on the absolute frame of reference. Li and Gleitman's critique is based on this study, and they use different variations of this experiment, which we shall call Animals-in-a-row. Levinson (in press), which was not available to Li and Gleitman through the timing of publication, summarizes further evidence based on rotation experiments in a larger sample and other studies probing non-linguistic representation of directionality in different cultures.

Thus, the evidence has amassed from numerous cross-cultural studies for systematic covariation between the frames of reference used in language and the frames used in non-linguistic aspects of cognition. In order to rule out other factors that may contribute to the choice of the frames of reference preferred in non-linguistic tasks, we have checked for statistical correlations with literacy, age, sex, or indices of culture-change, and found few such correlations (Levinson, in press; Levinson & Nagy, 1998). For example, there is no general correlation between literacy or years of schooling in our sample – only in the Tamil and Belhare subsamples (peoples in touch with populations who use relative systems) is there any positive correlation of literacy with coding-strategy.

So, if there is a correlation between linguistic frame of reference and non-linguistic frame of reference, which is chicken and which is egg? We reasoned as follows (Levinson, 1996b, 1997b; Pederson et al., 1998):

1. There are neighboring, closely related cultures in similar ecology in which distinct subsets of the linguistic frames of reference are used (for example, three Mayan cultures we have investigated: Mopan, intrinsic only; Tzeltal, absolute and intrinsic; Yukatek,

relative, absolute and intrinsic), so material culture and ecology can not be the only determinant.

2. If you are going to speak a language which, for example, only uses the absolute frame of reference, you will have to code scenes in memory using absolute coordinates. This follows from the fact that the frames of reference are not intertranslatable without ancillary information (Levinson, 1996b). So specific linguistic frames of reference demand specific non-linguistic coding.[6]
3. To get a community-wide consensus, there must be a community-shared source – which suggests language or some other semiotic system (like gesture) as a crucial catalyst.

Hence, we concluded, cautiously, that language is probably the driving force.

In sum, the program, of which the rotation experiments form a part, has been based on the following ingredients:

(1) careful collection of linguistic data according to standardized protocols and communication tasks taken from the community to be tested;

(2) the formulation of hypotheses about non-verbal cognition on the basis of the verbal behavior in (1);

(3) the observation and recording of verbal and non-verbal spatial behavior, including language acquisition (see Brown & Levinson, 2000), gesture and daily way-finding (Levinson, 1996c, in press);

(4) the testing of the hypotheses using the rotation paradigm, the results being interpreted in the light of (1) and (3).

3. The Li and Gleitman experiments

Li and Gleitman suspect that our experimental results are artifacts of the environmental conditions under which they were carried out, and reflect nothing about underlying cognitive differences, let alone linguistic determinism. They imply that our experiments with absolute populations were mostly run outside, and all with relative populations inside. But in fact there is no such confound – some of our strongest absolute results come from populations, such as Aboriginal Australians, tested indoors. For example, the Arrernte data reported in Pederson et al. (1998) were in fact collected in a room without any window (and similarly the Guugu Yimithirr data reported in Levinson (1997a) were collected indoors), while the Tzeltal data Li and Gleitman gloss as 'outdoors' were in fact collected under a low veranda, with restricted visual access not dissimilar to indoors with windows, and similar experiments were carried out indoors with similar results. And we had strong relative results from small ethnic groups tested

[6] A reviewer asks whether it might not be possible to use egocentric imagery to calculate absolute coordinates in real-time when required. Try thinking of your childhood bedroom and now describe without hesitation the location of the door, window, cupboard, etc. in correct cardinal direction terms – this is computationally demanding, if you can do it at all. If you can do it, you have at least one 'fix' to an absolute coordinate – without this you cannot ever recover the correct directions. In actual fact, there is evidence that absolute speakers/thinkers code mental imagery right at the start in cardinal direction terms (Levinson, 1996b, pp. 123–124, 1997a).

outdoors.[7] Not, at least originally, aware of this, Li and Gleitman therefore tested American speakers of English under varying ecological conditions.

Of all of our experiments, Li and Gleitman have chosen to replicate the very simplest ('Animals-in-a-row' – see Levinson (in press) for the many converging results from other experiments) and have gone on to simplify it further. The task in essence consists of presenting participants with a row of three animals on a table, rotating the participants 180 degrees, and making them reconstruct the array on another table so that to their satisfaction it matches the first. In our experiment, it was a crucial part of the design that participants' attention was deflected from the *direction* of the stimulus by being required to memorize the *order and identity* of three toy animals drawn from a larger set of four (Levinson, 1996b, p. 114). It was presented as a memory test, first without rotation, then with rotation (and both accuracy of order and direction were coded), and the participant was walked up to 20 m between stimulus and response.

The details of the Li and Gleitman replication vary from the original, including no translation of the participant and no delay after removal of the stimulus and thus a considerably shorter period for retention in memory,[8] and most importantly, the participant's memory task was reduced by presenting the subject with the three animals used in the stimulus, not the full set of four.[9] As we shall show, memory load can make a major difference (see our Experiment 2 below), and in addition when the experiment becomes too transparent participants may second-guess the experimenter's interest in direction rather than order or type of animal. It seems clear from the Li and Gleitman report that many of their participants were second-guessing the experimenter's intentions – they queried the instructions in a way that suggests that direction was clearly at issue. It is always hard to design a task that is matched across schooled and unschooled subject populations, but this should have been a warning that the task was not sufficiently opaque, for what we are interested in is participants' non-reflective utilization of a spatial coding scheme for memory, not what they think the experimenters think they should be doing. Incidentally, for this reason in our original Dutch experiments we used 40 participants of mixed ages (21–77), sex and occupational background, like the participants in our cross-cultural studies.

Li and Gleitman then go on to see if they can vary the results of the same experiment by varying the environmental conditions – in that case the results would show nothing about language or conceptual predispositions, but only about context. So they ran the task

[7] For example, Bantu Kgalagadi speakers – who speak a language with both Relative and Absolute terms available – used systematic Relative non-verbal coding on some of our memory tasks.

[8] They used a swivel chair to rotate the participant. This just could have the effect that the participant thinks of the whole setup as one location, not as in our experiment two locations separated by arbitrary distance. The predicted effect of that would be to make an intrinsic frame of reference more salient, and that, as we shall show, can partially mimic an absolute one.

[9] In this, they followed a deviation from the standard elicitation method for the study reported in Pederson et al. (1998), which was used on the first population to be tested (Tzeltal), as reported in Brown and Levinson (1993, p. 14). We took this simplification in this initial testing (before the standardization of the method), fearing that this unschooled Mayan population would not manage the memory task. It subsequently became clear on all other tasks that we need not have worried. All other populations reported in Pederson et al. (1998) were run with the standardized procedure of offering four animals to the participant for reconstruction.

indoors with blinds up or blinds down, inside vs. outside (Experiment 2a), and indoors with or without additional local environmental cues (Experiment 2b). What they found was that such environmental manipulations seemed to vary the results: the more 'outdoorsy' the setting (windows open, or out in the park), the more 'absolute' the results, and in inside conditions, strong table-top cues could be seen to bias the results in either direction, absolute or relative. More precisely, in the inside but blinds up condition in Experiment 2a, they reported mixed relative/absolute responses (although there was no statistically reliable difference between the blinds up/blinds down conditions), and in the outside condition they reported similar mixed absolute/relative responses, now significantly different than the indoor/blinds down condition. In the internal cue situation (Experiment 2b), they appeared to obtain relative responses when the cues were placed at say the left of each table, but absolute responses when they were placed at say the north end of each table.

On the assumption that their American participant population uses predominantly relative and intrinsic linguistic coding for all conditions (not just the one tested), then the results in Experiment 2b are in fact not unexpected on the assumption of a correlation between frames of reference available in language and those predominant in cognition, in ways that we will explain. But the results in Experiment 2a in the outdoor condition are more puzzling from our point of view.

4. Some more experiments: probing the Li and Gleitman results

We set out to try and find out why Li and Gleitman got the responses they got, and we conducted three sets of experiments. A first step was to try and replicate their results. Since our Dutch data as reported in Pederson et al. (1998) had been obtained under a 'blinds up' condition over six different rotation experiments (see Levinson, 1996b, in press, for the description of the other tasks), all without any shred of evidence of absolute coding tendency, we saw no chance of being able to replicate the Li and Gleitman finding under that condition.

But clearly we needed to see if we could replicate the outside condition. Like Li and Gleitman, we chose a location in the center of campus, and one where there are strong environmental cues to direction. We administered three different tasks for each participant. The first one is the Animals-in-a-row task, in which we sought to replicate the Dutch result reported in Pederson et al. (1998) in an outdoor setting. We used the method as described in Pederson et al. (1998). And, the participants had to chose three animals out of the four offered to reconstruct an array, unlike Li and Gleitman's experiment in which the same three animals were given to the participants for reconstruction.

The second task was the 'Motion-maze' task (Pederson & Schmitt, 1993). In this task, the memory for directionality is embedded in a larger task; thus it reduces the chance of participants second-guessing the purpose of the experiment, and increases the chance of participants falling back on their habitual default frame of reference. The maze task requires the participant to observe the movement of a toy man, then under rotation to recognize the path traversed from within a maze-like diagram containing both absolute and relative possibilities (see Fig. 1). It has been shown that speakers of absolute languages

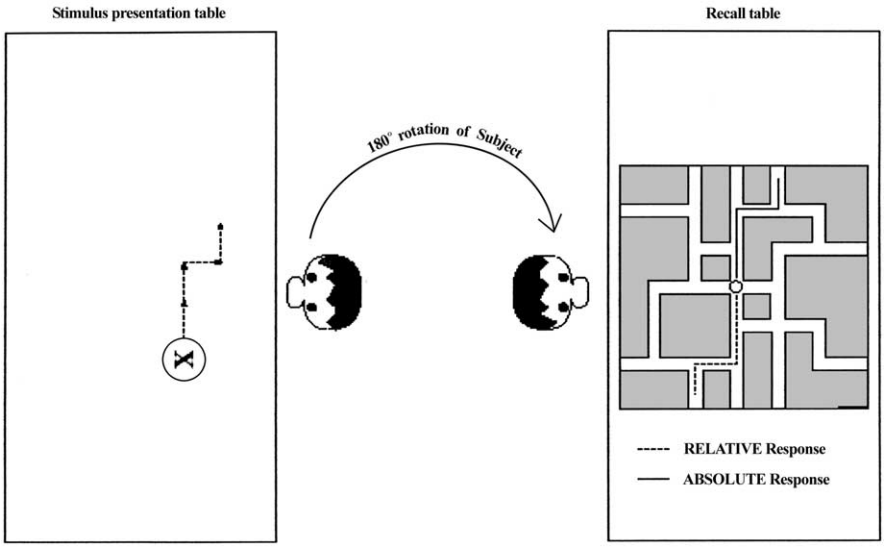

Fig. 1. Motion-maze task.

(Tzeltal and Arrernte) recognize the path based on the absolute frame of reference, and speakers of relative languages (Dutch and Japanese) recognize the path based on the relative frame of reference (Levinson, 1996b, in press).

The results from the above two 'outdoor' tasks will be compared to the Dutch results for the same tasks run under the 'indoor blinds-up' condition in earlier studies (Levinson, in press; Pederson et al., 1998). In those earlier studies, the methods were the same as those in the current experiments, except that in the original studies we used participants of mixed ages (21–77), sex, and occupational background, while the current experiments use a student population like Li and Gleitman's.

The last task to be administered was a linguistic task, requiring the verbal distinction between two lateral mirror-image photos. This was to probe the linguistic frame of reference *in the outdoor condition* (note that the linguistic data Li and Gleitman report were collected not outdoors, but indoors in a room with a view to the outside through a window).

4.1. Experiment 1: three tasks in an 'outdoor' condition

4.1.1. Method

4.1.1.1. Site The experiment was administered in a large open space outside the university canteen, an area familiar to all students. The north–south/east–west grid layout of the campus is particularly evident at this location. Buildings surrounding the location provide overwhelming directional cues. To the east is a large tower block (the tallest not only in the university, but also in the city). To the west is the only café in the university. To the south

is the building for the main university canteen. To the north is the main library of the university.

4.1.1.2. Layout Two tables were placed 4 m apart so that one table was north of the other. Participants stood between the two tables and were rotated 180 degrees in walking from one to the other. The stimulus arrays were place along an east–west axis. The participants started at the stimulus presentation table and then turned and walked to the recall table.

4.1.1.3. Participants Twenty local university students were recruited at the experiment site. They received 8.5 guilders for participating in the experiment. All the participants were tested individually. Each participant did the following three tasks.

4.1.1.4. Animals-in-a-row task The first task to be administered was Animals-in-a-row. This task was developed by Levinson and Schmitt (1993).

4.1.1.4.1. Material Stimulus arrays were created from a set of four plastic animals (pig, horse, cow, and sheep) from the Duplo™ series for infants. Their shapes are symmetrical along their head-to-tail axis. The four animals have distinct colors and shapes. The sizes range from 5 to 7 cm from the head to the tail, and they are all 2.5 cm wide and 3–4 cm tall.

4.1.1.4.2. Procedure All participants were tested individually. A session consisted of a few training and practice trials followed by five experimental trials. For all trials, the experimenter set up a row of three animals from the four available on the presentation table. The animals were facing either the participant's left or right, along an east–west axis. The animals were separated from each other by roughly 6 cm.

The participant was told to remember the animals just as they were. The participant was allowed to look at the stimulus array as long as he or she liked. The participant was asked whether he or she was ready, and if he or she said yes, the array was removed. For the initial practice trial(s), the participant was immediately given the four animals, and was asked to rebuild the row of animals in the same way on the stimulus presentation table without any rotation of the participant's body. Note that he or she had to choose the three appropriate animals out of the four given to rebuild the array. The direction, the order, and the identity of the animals were corrected if necessary. The procedure was repeated until the participant's performance became consistent.

In the experimental trials, the participant was told that he or she would do the same thing, but that this time they should reconstruct a row of animals elsewhere. Again, three out of the four animals were placed in a row on the presentation table, and the participant was asked to remember them just as they were. After the participant indicated readiness, the animals were removed. The participant was required to wait for 30 s,[10] and then walk to the recall table. Here, the experimenter offered the four animals to the participants, and asked him/her to rebuild the row. No correction was made to the participant's response. All presentations were along the left–right axis. The order and direction of the stimulus array were semi-randomized.

Throughout the experiment, none of the instructions contained any words denoting

[10] This delay reduces the chance of direct recall from short-term memory (visual scratch pad or auditory loop). Participants were allowed to look at whatever they wanted. The participant and the experimenter did not converse during this period. There was additional delay and visual input resulting from walking between two tables.

spatial directions or locations. If a reference to a location or direction became necessary during the training, deictic terms ('here', 'this') and pointing gestures were used.

4.1.1.4.3. Coding Responses were coded for either absolute (actually allocentric) or relative (here, egocentric) direction in which the animals were facing when rebuilt. As well as direction, the sequence of animals was also recorded in order to screen out trials that were especially poorly remembered. When the location of all the animals was wrong, namely, when the array cow–sheep–horse was rebuilt as sheep–cow–horse, the trial was not considered.

4.1.1.5. Motion-maze This was the second task to be administered. This task was developed by Pederson and Schmitt (1993).

4.1.1.5.1. Procedure All participants were tested individually. A session consisted of a few practice trials followed by five experimental trials. For all trials, the experimenter demonstrated a motion along a path by a plastic toy man (about 5 cm tall) moved manually but precisely on the presentation table. A small cross (about 1 cm by 1 cm) printed on a circular piece of paper (about 5 cm in diameter) was placed on the presentation table, and it served as a starting point of the toy man.

Before the demonstration of motion, the experimenter said, "Now this little man is going to go for a walk from this cross. Watch carefully because I want you to remember how he goes". Then, the experimenter walked the toy man from the starting-point cross. The motion was scaled to a particular path on the maze, which the participant did not see during the presentation of the paths. The paths consisted of straight segments that were either along a right–left axis (which was also an east–west axis) or a front–back axis. The paths for practice trials had one or two segments (the paths for experimental trials had two or three segments, as in Fig. 2). The experimenter produced 'footstep' sound effects as the man was moved to emphasize the distance between turns. The motion was repeated twice (or until the participant indicated readiness). Then, the man and the paper with a cross

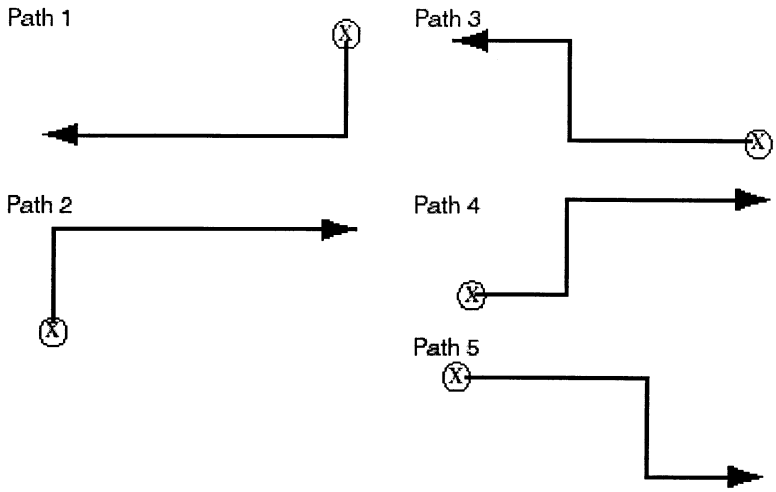

Fig. 2. The paths to be remembered in the Motion-maze task.

were removed, and a maze printed on 27 cm by 27 cm paper was put on the table. The participant was asked where the man would end up on the maze if he had followed the precise path previously demonstrated. The maze consisted of complex connected paths which ran either along a left–right axis or a front–back axis, and which led to eight possible end points. The participant either pointed at or named the label for one of the eight possible end points. During the practice trials, the participant did not rotate his/her body between the stimulus presentation and the recall on the maze. Throughout the experiment the participant did not manipulate the toy man. Two practice trials were administered, and if necessary more practice trials were run, until the participant correctly matched previously seen motion to recognized path.

In the experimental trials, a new maze, as shown in Fig. 3, printed on 27 cm by 27 cm paper, was placed on the recall table (there was no maze on the presentation table). The procedure was the same as the practice trials except for the following two points. First, after the presentation of a path, the participants were asked to wait for 30 s, and then turned around and walked to the recall table to respond. Secondly, no feedback was given to their responses. At the beginning of the experimental trials, the participants were informed that there would be multiple trials, and the toy man might end up in the same destination more than once and some of the eight possible destinations might never be reached by the toy man. If the participants could not remember the path, they were allowed to go back to the

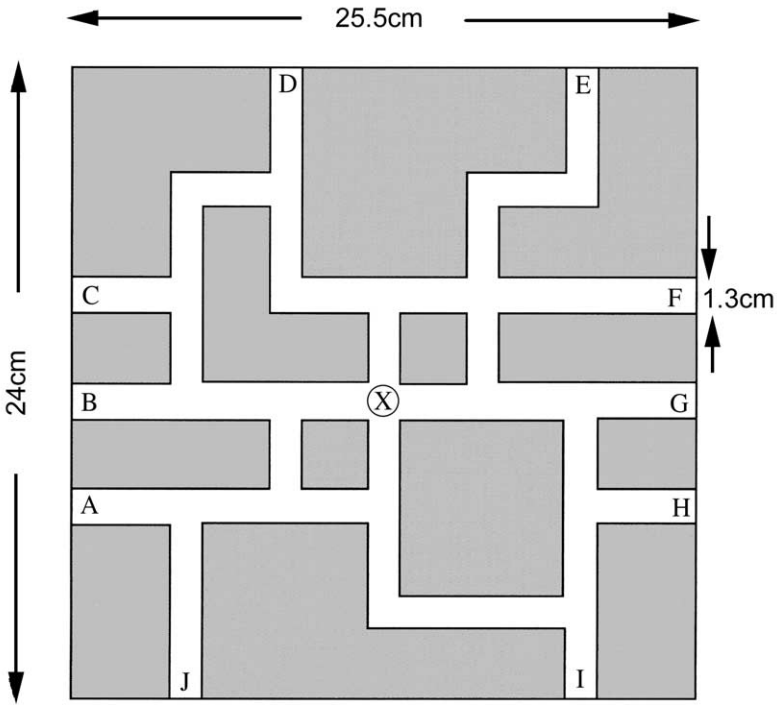

Fig. 3. The maze on which the path was recalled in the Motion-maze task.

presentation table and see the motion again. This procedure was repeated five times, once for each of the five paths in Fig. 2.

As in the previous experiment, none of the instructions contained any words denoting spatial directions or locations. If a reference to a location or direction became necessary during the training, deictic terms ('here', 'this') and pointing gestures were used.

4.1.1.5.2. Coding For each demonstrated motion, there were in fact two possible but different solutions embedded in the maze of paths – one correct if the path was coded in absolute terms, and one correct if coded in relative terms (due to the subjects' rotation, these paths ended up in distinct end points – see Fig. 1 for illustration). The response was coded as 'relative' if the end location was A for Path 1, F for Path 2, C for Path 3, F for Path 4, and H for Path 5. The response was coded as 'absolute' if the end location was F for Path 1, A for Path 2, H for Path 3, A for Path 4, and C for Path 5.

4.1.1.6. Linguistic elicitation After the above two tasks, a linguistic task was administered. The participant was shown two lateral mirror-image photos (two out of the six photos in Fig. 3 in Pederson et al. (1998), with a man to the left of a tree vs. a tree to the left of a man), which were arranged on the east–west axis. They were asked to describe each photo so that somebody else could tell which picture was being described.

4.1.2. Results

(a) The results for the Animals-in-a-row task are displayed in Fig. 4. In the figure, zero absolute response implies that all five trials were coded relative, except for one participant in the outdoor condition who gave four relative responses and a response that was neither absolute nor relative. The outdoor condition is compared to the 'indoor blinds-up' condi-

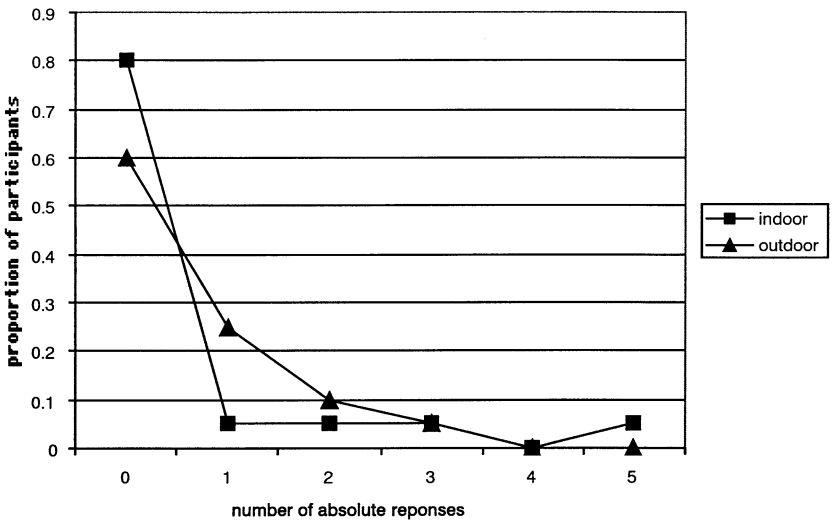

Fig. 4. Direction of animals in the Animals-in-a-row task with Dutch participants: Indoor and Outdoor conditions (a total of five responses from each participant are coded either relative or absolute).

tion. The data from the indoor blinds-up condition consist of 20 participants drawn randomly from the 38 Dutch participants in the earlier study reported in Pederson et al. (1998).[11] As is immediately clear, under both conditions a significant majority of the participants had predominantly (three or more) relative responses: Indoor (Binomial, $p < 0.01$), Outdoor (Binomial, $p < 0.01$). The mean number of absolute responses for the indoor condition was 0.55 out of five trials (*SD* 1.31), and that for the outdoor condition was 0.60 out of five trials (*SD* 0.88). There was no significant difference in the mean number of absolute responses between participants in the outdoor condition and the indoor condition (Mann–Whitney *U*-test, $U = 168$, $P = .40$). The difference is not significant with the *t*-test, which is more sensitive, either (*t*-test, $t = 0.14$, $df = 38$, $p = .89$).[12]

(b) The results from the more exacting Motion-maze task are shown in Fig. 5. Once again in the figure, zero absolute response implies that all five trials were coded relative, except for one participant in the outdoor condition who gave four relative responses and a response that was neither absolute nor relative. The outdoor condition is again compared to the 'indoor blinds-up' condition. The data for the indoor blinds-up condition consist of ten participants drawn randomly from the Dutch participants in Levinson (in press). Under both conditions, all the participants had predominantly (three or more) relative responses: Indoor (Binomial, $p < .01$), Outdoor (Binomial, $p < .01$). The mean number of absolute responses for the indoor condition was 0.05 out of five trials (*SD* 0.22), and that for the outdoor condition was 0.25 out of five trials (*SD* 1.11). The mean numbers of absolute responses do not significantly differ between the outdoor and indoor conditions (Mann–Whitney *U*-test, $U = 200$, $p = .99$). The difference is not significant with the *t*-test, which is more sensitive, either (*t*-test, $t = 0.78$, $df = 38$, $p = .44$).

(c) The descriptions of the photographs were analyzed in terms of the key expression that encodes the spatial relationship between the man and the tree, such as "to the right of". This analysis revealed that all participants used relative coding (but no absolute coding) in language in the outdoor condition. The result is consistent with what has been reported about Dutch speakers in Pederson et al. (1998) and Brown and Levinson (1993) and congruent with the choice of relative coding in the Animals-in-a-row task and the Motion-maze task.

4.1.3. Discussion

In both the Animals-in-a-row and the Motion-maze experiments, 95% of the participants gave more relative responses than absolute responses in the outdoor and indoor conditions. In accord with our predictions, this matches the choice of linguistic frame of reference for the description of tabletop spatial relationships both indoors and outdoors (as reported in Pederson et al., 1998). We do not find anything like the qualitative difference between the results for the indoor and outdoor conditions that Li and Gleitman report – their results showed a bimodal distribution, in which 35% of their participants gave no

[11] Two of the 40 participants were excluded in that earlier experiment because they responded 'monodirectionally', namely, using one fixed direction of response regardless of the direction of the stimulus.

[12] Reconfirming the non-significance with the *t*-test for this result and the results for Figs. 5 and 8 was suggested by one of the reviewers.

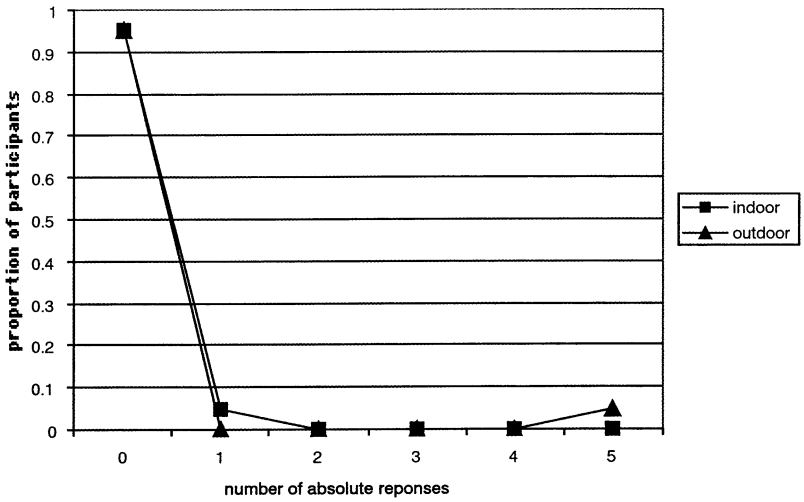

Fig. 5. Motion-maze task with Dutch participants: Indoor and Outdoor conditions (a total of five responses by each participant are coded either absolute or relative). Note: the two lines overlap at zero, two, three, and four absolute responses.

absolute responses and 40% of them produced only absolute responses (see their Fig. 7). In short, we have been unable to replicate their results.

We will return to ask why we got such different results. But first, let us note that there are non-significant trends in the data that could be interpreted as in accord with Li and Gleitman's hypothesis that landmark cues in outdoor settings lead to the use of the absolute frame of reference. In both the Animals-in-a-row task and the Motion-maze task, the outdoor condition indeed yielded a higher (but non-significant) mean number of absolute responses. But it would be premature to conclude that this supports the Li and Gleitman hypothesis. Firstly, in the case of the Motion-maze, the trend is entirely contributed by one individual, who produced only absolute responses. (Incidentally, there is also such a person in the indoor condition in the Animals-in-a-row task.) Secondly, the outdoor condition introduces additional confounds: the outdoor experiments were run (following Li and Gleitman) in the center of campus amidst the distractions of passers-by, and direction errors in a relative frame of reference will get coded as absolute in these tasks. The pattern of responses – a slight depression in relative performance – is entirely compatible with this interpretation. Why is the effect more pronounced in the Animals than in the Motion-maze task? Three independent variables need to be remembered in the Animals task (identity, order and direction of animals), and only one path in the Motion-maze task. Hence, the response in the Animals task may be more fragile under distraction.

How can one explain the discrepancy between our study and that by Li and Gleitman? One possibility is simply that the subject pool Li and Gleitman used in the University of Pennsylvania is much more heterogeneous than our pool of subjects in the University of Nijmegen – students no doubt come from all over the States and beyond, but Li and

Gleitman apparently screened their subject pool, which they characterize as "a single cultural and linguistic subgroup" (p. 13), so this explanation seems unlikely.[13]

The second more plausible explanation is that Li and Gleitman's simplified task was simply too transparent to their participants,[14] who attempted to second-guess the intentions of the investigator. This interpretation is supported by the fact that 70% of their participants in the blinds-up and outdoor conditions asked the experimenter which of the two solutions they should choose, showing that they were aware of both. There are obvious ways to test this explanation. In our original version of the Animals task, the focus of participants was deflected from direction to identity and order – they had to recall which three of four animals were lined up in which order, with direction as an implicit variable, but the Li and Gleitman simplification (just three animals) and coding of direction without order lost this aspect of the task. The other tasks we have used in our large cross-cultural sample further background direction by embedding memory for direction in, for example, a reasoning task (see Levinson, 1996b). A further manipulation is to increase the memory load further, for example by placing a fourth animal in the sagittal plane, so that the participant has to memorize an array on two axes (for example, cow, pig, horse in a line heading *left* across the direction of view and sheep in *front*, occluding pig) – then a real absolute response under 180 degree rotation will have the line heading *right* with the sheep *behind* the row, occluded by the pig.[15] Both embedding of direction-coding in a larger task and increasing the difficulty of the task should avoid the meta-awareness displayed by Li and Gleitman's subjects – with the consequence we confidently predict, that their subjects would act just like ours. Since unfortunately we cannot replicate their results, we are unable to test this further. Instead we will turn to examine their Experiment 2b, and show that, contrary to their assumptions, this has nothing to do with an absolute frame of reference.

4.2. Li and Gleitman's 'duck pond' experiment

In explaining their motivations for their Experiment 2b, Li and Gleitman make clear that they think that an absolute system is all about landmarks. True absolute systems have nothing to do with landmarks – the geometry of such systems does not consist of lines converging on a landmark, instead it has infinite parallel lines constituting an abstract 'slope' across an environment (see Levinson, 1996b, Fig. 4.9). Most cardinal direction systems are abstractions off landscape features or off meteorological or celestial features, but they are indeed abstractions. For example, although the Tenejapan Tzeltal system names South as 'uphill', 'uphill' remains 'uphill' on the flat – it is a cardinal direction

[13] There are anecdotal reports that Midwestern Americans utilize more cardinal directions in *both* language and cognition than East Coast residents, though no studies have been conducted. Such a mix could in principle lead to Li and Gleitman's bimodal distribution, but again seems to have been ruled out by screening of subjects.

[14] Furthermore, if Li and Gleitman's participants were psychology students (they do not say), then clearly the experimenter's goals may have been clearer to them than to our participants from random faculties, who were recruited at the site of the experiment.

[15] Pilots in Tenejapa (Mexico) and Hopevale (Australia) show that absolute-speaking populations will do this. We didn't use this manipulation for the reasons explained – our tasks were run on some populations who had had no schooling whatsoever, and we simplified all tasks to the minimum.

system in disguise.[16] At night, in an alien city, facing a device never seen before (namely a sink with two taps), one Tenejapan asked another, "Which is the hot tap, the uphill (southern) or the downhill (northern) one?". They maintain a constant sense of absolute orientation, presumably by running a continuous background computation of egocentric heading with respect to abstract bearings, integrating multiple internal and external cues to achieve this.[17] This is the phenomenon that we are trying to capture.

So what are the characteristics of a landmark system? Well, it no doubt depends on the system. Some of them cover a vast territory and operate very much like absolute systems (Austronesian inland/sea systems or Alaskan upstream/downstream systems are of this type, see Levinson, 1996a). Others are local, and are more like very large intrinsic arrays: if I have a mental map of the internal arrangements of a large building like a library or city administration, but can't orient this map in a larger landscape, I am operating with an 'orientation-free' representation as in the intrinsic frame in Table 1. Notice that for intrinsic coding either the ground (or landmark) object or the figure (object to be located) must have intrinsic features, as in "The animals are facing the pond".

Li and Gleitman set out to ask in their Experiment 2b "Can landmark information, if it is salient enough, completely determine the degree to which a single population solves spatial-problems?". As 'landmarks' they used 'duck ponds', big colorful symmetrical objects. They placed one of these on both the stimulus and response tables of the same Animals task as before: in their 'relative' condition they placed the duck ponds always to the participants' right on both tables; for the 'absolute' condition, they placed the ducks always to the south of both tables (and thus with left/right alternation under rotation). The results were that under the 'absolute' condition, participants lined up the animals facing the duck ponds, and in the relative condition they did the same, with the animals in the reverse direction.

One has to note immediately that these are obviously not 'landmarks' in any normal sense, since identical objects are replicated in different locations (you don't expect to have clones of the local cathedral on neighboring streets!), and the landmark objects are clearly relatively small and movable. Rather, they will be interpreted by participants as part of the scene to be replicated. What participants clearly did was use the large, bright objects as an orientational cue – they were treating the whole assemblage, both duck ponds and animals, as one array to be reproduced. What kind of coordinate system is involved in maintaining the internal arrangements of an array while its orientation is varied? An orientation-free frame of reference of course – what we call an intrinsic frame of reference (see Levinson, 1996b, pp. 147ff). So what Li and Gleitman actually tested was whether they could bias

[16] Interestingly, Tzeltal children seem to key into the abstract nature of the system relatively early, and they do not seem to pass through a stage of using landmarks on the way (Brown, 2001; Brown & Levinson, 2000).

[17] A reviewer asks how they do this. Unfortunately, we do not really know – verbal protocols suggest that dead-reckoning of current position by keeping track of turns and distances traversed is involved, and that many environmental cues are constantly used to correct accumulated errors. But *that* these peoples maintain such a 'mental compass' is not in doubt. For some of the groups we have tested, by transporting individuals to unfamiliar locations, the ability to point to unseen locations is quite spectacular, exceeding the accuracy of, for example, 'homing pigeons' initial flight paths over similar distances (see Levinson, in press, Chap. 6). This accurate sense of direction correlates with the use of absolute frame of reference in language.

participants between the two frames of reference predominantly used in English, namely the intrinsic and the relative, and they found they could. We would never have doubted that they could do so (see Tversky, 1996 for the long tradition of research here). We only predict that a true absolute frame of reference, if absent from ordinary language usage for these kinds of 'table top' contexts, is hardly accessible to these participants for non-linguistic conceptual coding for similar arrays.

Can we directly demonstrate that an intrinsic frame of reference is what is involved, and that the rival frame is relative? We needed first to replicate Li and Gleitman's finding, then vary the conditions, and this is what we did. We performed two experiments. Our Experiment 2 first replicated Li and Gleitman's Experiment 2b under its so-called 'absolute' (our intrinsic) condition, with an extra condition to test whether we could induce a relative frame of reference while maintaining cues biasing to their 'absolute' frame of reference. Our Experiment 3 was designed to demonstrate that an intrinsic frame of reference, not an absolute one, is really what is at stake.

4.3. Experiment 2

We followed the procedure and setting of Li and Gleitman's 'absolute' condition in their Experiment 2b. It is a version of Animals-in-a-row with a pair of identical 'landmarks' ('duck ponds') given on both the presentation and recall table. We investigated the choice of frames of reference under two conditions with different memory load. One condition is the Three Animal condition, which precisely replicates the absolute condition of Li and Gleitman's Experiment 2b. At the recall table, participants are given just the three animals used in the array on the presentation table. Thus, the participants have to remember only the order and direction of animals, but not the identity of the animals used in each trial. The other condition is the Four Animal condition, in which the participants have to choose the three out of four possible animals at the recall table, according to the animal types used in the stimulus. This is equivalent to our Experiment 1, and adds slightly to the load on recall memory.

Our prediction was as follows. Dutch subjects have two frames of reference (intrinsic and relative) available in language, and thus use just these two available frames also for conceptual coding. However, both earlier linguistic and non-linguistic tasks had suggested that for Dutch speakers the relative frame of reference is predominant (see, for example, Levelt, 1996, p. 99). We therefore expected the very salient 'duck pond' cues to bias towards the intrinsic frame, but the increased memory load to bias in the other direction, toward the more habitual relative frame.

4.3.1. Method

4.3.1.1. Material Stimulus arrays were created using the same four plastic animals as in Experiment 1. A pair of identical 'duck ponds' were used as 'landmark' cues on the stimulus table and the recall table. Just as in Li and Gleitman's experiment, they were roughly circular (about 20 cm in diameter) and longitudinally symmetrical, and had prominent bright colors, consisting of two yellow toy ducks fixed on a blue surface representing a pond.

4.3.1.2. Setting and layout The setting and the layout for the experiment were recreated as closely as possible to the 'absolute' condition in Li and Gleitman's Experiment 2b.[18] The stimulus presentation table and the recall table were aligned to a north–south axis, and were close enough to each other so that the participant could swivel his or her chair 180 degrees to face the recall table (as in Li and Gleitman's corresponding experiment). One of the duck ponds was placed in advance on the stimulus presentation table so that it was on the participant's right side when facing the table. On the recall table, the other duck pond was also placed in advance of all trials, but now on the left side of the participant when facing the table.

4.3.1.3. Participants Twenty student participants were recruited from the Max Planck Institute participant pool. The participants were different from those in the other experiments reported in this paper. They were paid 8.5 guilders each for their participation.

4.3.1.4. Procedure Half of the participants were randomly assigned to the Three Animal condition, and the other half to the Four Animal condition. The procedure was essentially the same as that for 'Animals-in-a-row' in our Experiment 1, except that the delay after the removal of the stimulus was 15 s.

(a) In the Three Animal condition, the procedure was essentially equivalent to that of Li and Gleitman's Experiment 2b. In this condition, three animals were lined up on the stimulus presentation table, and the same three animals were given to the participants at the recall table to reconstruct the array.

(b) In the Four Animal condition, three animals were lined up on the stimulus presentation table, and four animals were given to the participant at the recall table. Thus, the participant had to choose the relevant three animals out of the four given, according to the animals used in the stimulus.

4.3.1.5. Coding Five experimental responses were coded for either the intrinsic or relative direction in which the animals were facing when rebuilt. The sequence of animals was also recorded, to screen out the trials that were especially poorly remembered. When the location of all the animals was wrong (for example, when the array cow–sheep–horse is rebuilt as sheep–cow–horse), the trial was not considered.

4.3.2. Results

In the Three Animal condition, we obtained just the results that Li and Gleitman did, namely the direction of recall was cued by the 'duck pond'. But in the Four Animal condition, with the greater memory load, participants ignored the 'duck pond' cues, and reproduced the animals in a relative way, i.e. preserving left/right orientation. The results are contrasted in Fig. 6 – in this figure we label what Li and Gleitman called 'absolute' codings as 'intrinsic' ones, for reasons that will become clear. Note that in the figure, zero intrinsic response implies that all five trials were coded relative, and vice-versa, thus

[18] A difference was that our experiment was carried out in a room without any window, while in Li and Gleitman's version the room had windows and blinds were up. However, here this variable was incidental and not a controlled condition in their experiment.

the full five intrinsic responses imply that for those participants no relative responses were produced. The mean number of intrinsic responses in the Three Animal condition was 3.8 (*SD* 2.04) out of five trials, and that in the Four Animal condition was 1.0 (*SD* 1.89). The difference between the two means is significant (Mann–Whitney *U*-test, $U = 19$, $p < .01$).

4.3.3. Discussion

This experiment establishes that the result in Li and Gleitman's Experiment 2b is replicable (unlike their Experiment 2a) – but we think that the participants used the intrinsic frame of reference to code the array. The result for the Four Animal condition is interesting. It shows that despite the prominent cues, what we suppose to be an intrinsic result is fragile: as soon as the memory load is upgraded slightly, it appears that participants revert to their habitual, predominantly relative way of coding spatial scenes. The above result also throws light on Li and Gleitman's Experiment 2a: as we suggested above, we predict that if they upgrade the memory load, participants will not be able to engage in the second-guessing behavior that we suspect underlies their 'absolute' result, and will react in a relative way.

4.4. Experiment 3: the 'duck pond' experiment under 90 degree rotation

We now attempt to show experimentally what we have already argued conceptually, namely that the Li and Gleitman 'absolute' condition is nothing of the kind, but just an intrinsic condition. To do that, we need to use the 'orientation-free' character of intrinsic arrays (see Table 1), so we ran the same 'duck pond' experiment as in their Experiment 2b

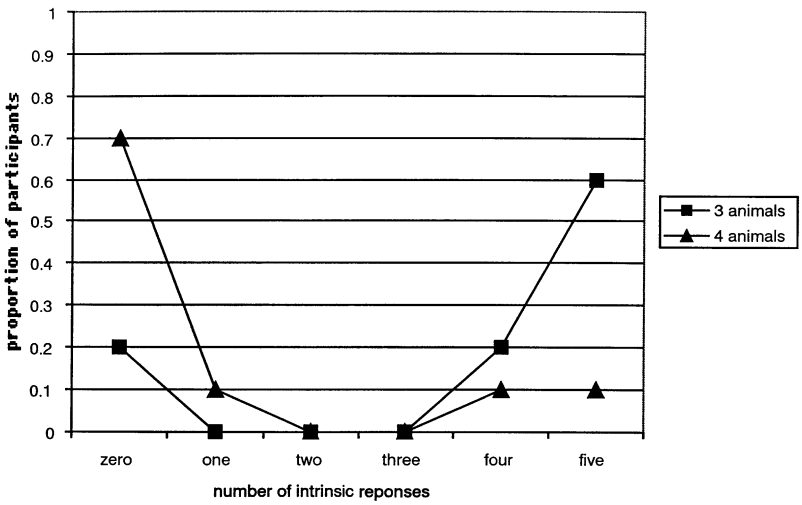

Fig. 6. Animals-in-a-row task with duck pond 'landmarks' with Dutch participants: Three and Four Animal conditions (a total of five responses from each participant are coded either intrinsic or relative). Note: the two lines overlap at two and three intrinsic responses.

Condition 1, but under a 90 degree rather than a 180 degree rotation. Then we can compare the two conditions.

Let us clarify the reasoning. An intrinsically-coded array is orientation-free in the sense that only its internal arrangement has to be preserved – in this case animals facing towards or away from the 'duck pond'. Both an intrinsic and absolute solution can look the same under 180 degree rotation – that is, the participant may be thinking "animals facing duck pond" (intrinsic) or equally "animals facing north" (absolute). The intrinsic and absolute solutions can become separated under any rotation, but since the intrinsic solution by definition can be in any direction, it will tend to be oriented by local ecological factors, like the main axis of the table, and viewpoint-preserving factors, like egocentrically transverse vs. sagittal arrangement. Thus, under 180 degree rotation with a duck pond at one end of the table and the main axis of the table in the egocentric transverse, they will tend to align. But if we now put the recall table at 90 degrees to the stimulus table, the absolute solution will require a sagittal alignment away from the participant in response to a transverse stimulus, while the intrinsic solution is likely to be influenced by ad hoc factors, like the main axis of the table or preservation of the transverse viewpoint. Thus, the two frames of reference should now separate. Our prediction of course is that what Li and Gleitman are calling an absolute response is in fact coded intrinsically by participants like theirs or ours.

4.4.1. Method

The material, the setting, and the procedure were identical to the Three Animal condition in Experiment 2. The layout was the only difference. The stimulus presentation table and the recall table were arranged at a 90 degree angle. Thus, the participant swiveled the chair 90 degrees rather than 180 degrees, as in Fig. 7 (the layout with 180 degree rotation in Experiment 2 is also shown but dotted for comparison).

4.4.1.1. Participants Ten participants were recruited from the Max Planck Institute participant pool. The participants were different from those in any other experiments reported in this paper. They were paid 8.5 guilders each for their participation.

4.4.1.2. Coding Five experimental responses were coded for either intrinsic (towards or away from the duck pond) or absolute direction (fixed compass bearing) in which the animals were facing when rebuilt. The sequence of animals was also noted to allow poorly remembered trials to be discarded (for example, when the array cow–sheep–horse was rebuilt as sheep–cow–horse). Note that the location of the duck pond on the recall table was such that a relative response was not possible.

4.4.2. Results

The results are depicted in Fig. 8, which charts the 90 degree condition against the matching 180 degree condition (i.e. the Three Animal condition in our Experiment 2). Along the *x*-axis we now have number of *intrinsic* trials, that is the trials which preserve a direction headed to or away from the 'duck pond' cue. In the figure, the full five intrinsic responses imply that for those participants no absolute responses were produced in the case of the 90 degree condition, and that no relative responses were produced in the case of

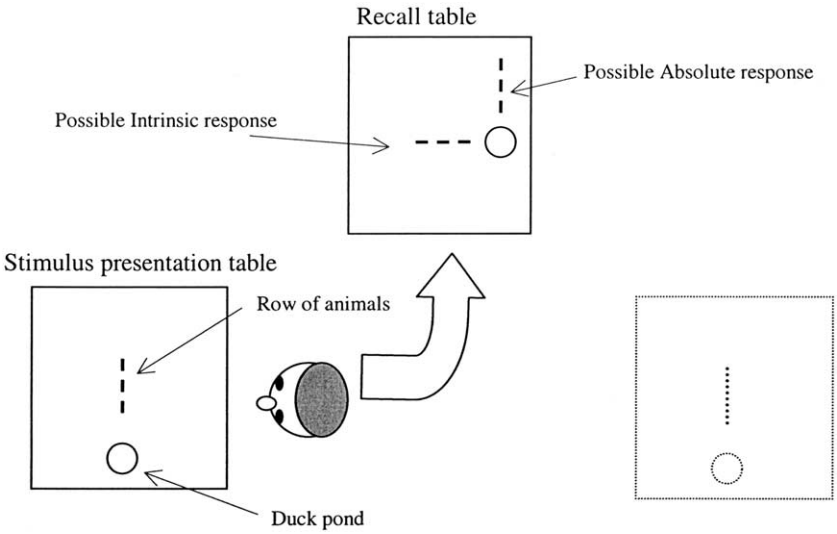

Fig. 7. The layout of Experiment 3.

the 180 degree condition. It is clear that in the 90 degree condition the great majority of trials did NOT align sagittally (which would have allowed an absolute interpretation), but were oriented intrinsically. The majority of the participants had three or more intrinsic responses (Binomial, $p = .11$).

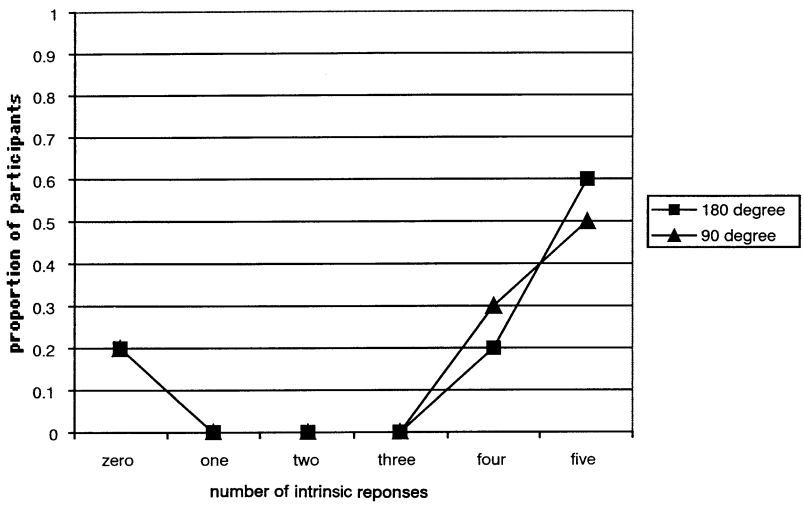

Fig. 8. Animals-in-a-row task with duck pond 'landmarks' with Dutch participants: 180 degree and 90 degree conditions (a total of five responses from each participant are coded either intrinsic or absolute). Note: the 180 degree condition in this figure plots the same data as the Three Animal condition in Fig. 6. The two lines overlap at zero, one, two, and three intrinsic responses.

The mean number of intrinsic responses in the 90 degree condition was 3.7 (*SD* 2.00) out of five trials, and that in the 180 degree condition was 3.8 (*SD* 2.04). There is no significant difference between the two means (Mann–Whitney *U*-test, $U = 46$, $p = .74$). The difference is not significant with the *t*-test, which is more sensitive, either (*t*-test, $t = 0.11$, $df = 18$, $p = 0.91$). This strongly suggests that behavior under both conditions comes from the same source: an intrinsic coding. If you pool the participants from both conditions, a significant majority of participants had three or more intrinsic responses (Binomial, $p < .01$).

4.4.3. Discussion

The result of Experiment 3 makes it clear that what Li and Gleitman called 'absolute' responses in their Duck-on-tables experiment (their Experiment 2b) and the predominant responses in the Three Animal condition in our Experiment 2 were in fact intrinsic responses. For our Dutch participants it takes low memory load, together with prominent local cues which can easily be construed as forming a single array with the test objects, to induce a switch from the relative to the intrinsic frame of reference.

Both English and Dutch are languages which (in most dialects anyway) offer two frames of reference in common parlance: namely both intrinsic and relative. Of these two, relative is predominant. For example, in an abstract description task – neutral over real scale or real objects – Levelt (1996, p. 99) found that less than 25% of Dutch participants were verbally consistent intrinsic coders, and Li and Gleitman report similar figures. Still, both frames of reference are perfectly colloquial. Thus, on the hypothesis that language correlates with and influences cognition, we would predict that both frames of reference may be used in non-verbal coding, with the relative frame predominant. The results from the two duck pond experiments taken together indeed indicate once again a language–cognition correlation, here in terms of linguistically favored frame of reference and most robust frame of reference in memory. As a reviewer points out, to establish that this correlation has a causal interpretation would take a further demonstration, and a first step would be to show that cross-culturally linguistic preference always correlates with a default, robust frame of reference in memory.

It is nevertheless important to re-emphasize that Dutch and English speakers switching between the intrinsic and relative frames of reference is compatible with the hypothesis. Thus, contrary to what Li and Gleitman argue, "showing that speakers of a single language …can be induced to vary in their spatial reasoning strategy by changing the circumstances of test" (p. 290) does not constitute counter-evidence to the hypothesis under investigation.

5. Distinguishing the absolute and intrinsic frames of reference

A fundamental problem with Li and Gleitman's study is that they do not make the necessary distinctions between frames of reference. They consistently equate our 'absolute' frame with the higher-order classification 'allocentric' (pp. 268–270), thinking that the 'intrinsic' frame of reference is a kind of 'absolute'. But in fact, as we must now show,

the intrinsic frame and the absolute frame have crucially different properties. First, they have quite different logical properties. As Levelt (1989) has pointed out, the intrinsic frame of reference does not support transitive inference, while the relative and, we may add, the absolute ones do. The inference "Abel is north of Beth, Beth is north of Cain, therefore Abel is north of Cain" is valid. But the corresponding inference when interpreted intrinsically, "Abel is at Beth's left, Beth is at Cain's left, therefore Abel is at Cain's left" is invalid – it will be true only if Abel, Beth and Cain happen to be facing the same way. Second, as pointed out in the previous section, the rotational properties of absolute and intrinsic codings are fundamentally distinct: absolute codings of arrays are made in terms of fixed bearings that have nothing to do with the array itself, while intrinsic codings are based on array-internal relationships, and are hence invariant to the rotation of the whole array. There is no sense, then, in which the intrinsic frame is a kind of absolute frame.[19] Further, and crucially for the matter in hand, there is no translation possible from intrinsic coding to absolute or relative coding (or from relative to absolute) – that is, there is no way to convert information from, for example, an intrinsic or relative coding to an absolute one (at least, without ancillary information; see Levinson, 1996b, pp. 152–158 for the demonstration). It is this lack of inter-translatability between frames of reference that guarantees a congruence between linguistic coding and the coding people use in non-linguistic memory.

From Li and Gleitman's conflation of the intrinsic and absolute frames of reference numerous confusions follow. First, as we have shown, they misinterpret their own experimental findings. Second, they think that the presence of 'landmarks' as a cue is the defining characteristic of 'absolute' frames of reference, whereas in fact absolute systems proper make no use of a system of landmarks. A landmark system presupposes a radial geometry – if I left my car facing towards the tower, that doesn't tell me which side of the tower to look for it (unlike remembering that I left it north of the tower). Another feature of a landmark system is that the system applies only in a delimited area. Take the example of expressions, *uptown*, *downtown*, and *crosstown*, used in Manhattan Island of New York City. According to our informants from New York City, the application of these terms is strictly limited to directions and locations on Manhattan Island.[20] For example, once you cross Brooklyn Bridge from Manhattan into Brooklyn, suddenly the same absolute directions cannot be referred to by these terms at all. These expressions are thus analogous to the word *front* in expressions like *the front row of a theater* within the intrinsic frame of reference: outside the theatre, the frame is irrelevant, just as are *uptown*, *downtown* when you cross over

[19] Nor is there a sense in which the absolute frame is a kind of intrinsic one, a possibility raised by a reviewer, who questions whether an absolute frame is not simply an intrinsic frame where the 'ground' is the local terrain. This may indeed be the right characterization of landmark systems (a point taken up below), but for all the logical and rotational reasons just explained, it cannot be a correct analysis of a true absolute system using abstract fixed bearings or cardinal directions.

[20] Thanks to Jennie Pyers and Aida Radican for sharing their insights about how people describe directions and locations in New York City.

the river to Brooklyn – the terms are allocentric but not absolute, as Li and Gleitman suggest.[21]

Absolute systems presuppose a conceptual 'slope', or series of infinite parallel lines across the environment. You can't walk around such a conceptual slope, in the way you can walk around a tower. The two systems have long been distinguished in studies of navigation: absolute systems are involved in dead-reckoning, landmarks in piloting, and they involve quite different procedures (Gallistel, 1990). Third, Li and Gleitman therefore imagine that the linguistic and conceptual systems under investigation as 'absolute' by our project are entirely familiar to English speakers, who, they suggest, could at the drop of a hat say "Give me the spoon that's northeast of your teacup" (p. 7). They can't because they can't routinely compute it, anymore than they can instantly give you their telephone numbers in binary code. But the 'absolute' language populations we have been interested in do routinely use such statements, can instantly compute them, and remember everything of whatever scale in terms of the locally relevant conceptual slope, as can be shown not only through memory experiments but also by examining their unconscious gestures during speaking. This is a truly interesting phenomenon, of considerable importance to our understanding of the 'psychic unity' of the species, and nothing is gained by shoving it under a terminological rug.

This conflation of absolute and intrinsic frames of reference vitiates the relevance of Li and Gleitman's discussion of the animal and infant literature. They suggest that humans are just like rats, in that rats show sensitive use of the best spatial cues, thus being absolute- or relative-coders on demand. But the literature does not support this. First, it is false that "when provided with sufficiently rich and stable landmark cues, any self-respecting rodent will use them" (p. 22) – this was conclusively shown by Cheng and Gallistel, for rats are only attuned to geometrical information, ignoring color, pattern and other rich landmark cues (see Gallistel, 1990, Chap. 6). Human use of landmarks is strikingly different in its multi-modality, rat-like behavior being rapidly superceded in infancy, just as language is being acquired and plausibly linked to its multi-modal semantics (Spelke & Tsivkin, 2001). Secondly, rats show no absolute fixed-bearing sense of direction as far as is known – they may use landmarks to form a 'centroid'

[21] English also of course has the words *north*, *south*, *east*, and *west*, which are defined in terms of an absolute frame of reference. However, there are plenty of signs that in ordinary American or British parlance these are not used as their counterparts are in languages like Guugu Yimithirr or Tzeltal, which lack a relative frame of reference. First, they are scarcely ever used on a smaller than geographic scale (a boy who used them inside a house was thought worthy of a note in *Science* in 1931). Second, these terms are more likely to invoke mental representations based on the relative and intrinsic frames of reference, perhaps in accord with how they are acquired, for it seems likely that these notions are acquired largely through the practice of map reading, which involves the convention of north being up, west being left, etc., i.e. representations within the relative frame of reference. For example, when an American English speaker in New York hears a statement such as *Laos is west of Vietnam*, she does not imagine a direction west of where she is, but rather a direction to the left on an imaginary map. Furthermore, some speakers may have another layer of meaning, based on the intrinsic frame of reference, defined by a network of local landmarks. This supports knowledge such as *If I drive down Broadway from Times Square to Harlem, I am heading north*. This kind of directional knowledge is strictly local as we saw in the expressions such as *uptown* in Manhattan. Thus, for all spatial computational purposes, the cardinal direction terms may be based on the relative and intrinsic frames of reference for many English speakers.

from which other locations can be calculated, but this will change as each new landmark is discovered (O'Keefe, 1993). Thus, rats have allocentric systems of orientation, but not (as far as is known) absolute ones of the kind at stake here. In contrast, numerous arthropods and bird species do have absolute senses of direction, utilizing in-built polarized light receptors and magnetoreception – 'hardware' apparently denied to terrestrial mammals (Hughes, 1999), but successfully mimicked in 'software' by humans of certain groups, or by technology in others (Levinson, in press). So, once again, nothing is gained by conflation of distinct frames of reference in the study of animal cognition. And the same goes for the study of infant orientation, where the proper frame of reference distinctions could be most helpful. But here again, there is not the slightest evidence from the literature for genuine absolute responses in Western infants – even landmark-cued allocentric behavior being perhaps derived from transformations of egocentric information (Pick, 1993, p. 35).

Finally, Li and Gleitman advance the hypothesis that the results of our cross-cultural studies could be explained by supposing that the small-scale, unschooled, traditional societies who use absolute systems share familiar landmarks, because in effect they live together. This idea is not in accord with the ethnography (for example, our hunter-gatherer groups are far-flung wanderers, the Tenejapans do not live "in a village on a hill" but have a dispersed settlement pattern over a large territory), nor could it be determinative since there are lots of small, localized human groups who do not use absolute systems of spatial reckoning. But the main reason the hypothesis will not fly is that landmark cues do not play any special role in absolute systems like the Tzeltal or Arrernte systems. If you transport individuals from these communities out of their familiar territories, their 'downhill' or 'north' remains anchored to the same fixed bearing (in our compass degrees) that it always had (see Levinson, 1996c, in press for the experiments).

6. Conclusions

Our critique of the Li and Gleitman paper is based on the following points:

1. Li and Gleitman did not make the fundamental conceptual distinctions between frames of reference, conflating 'absolute' and 'intrinsic' frames of reference.
2. As a result they have misinterpreted their own results: they have *not* discovered that they can systematically induce American students to code absolutely – what they have shown in their Experiment 2b is that they can bias them to switch between their own language-correlated frames of reference, intrinsic and relative. We showed this by a simple 90 degree rotation variant (our Experiment 3). All in all, no environmental manipulations shake our Dutch speakers, at least, out of the two frames of reference available in their language.
3. In the outdoor condition in our indoor-vs.-outdoor experiment, we did not replicate the bimodal distribution with two equally high peaks for participants with predominant relative responses and those with absolute responses, which Li and Gleitman obtained in their Experiment 2a. We think they only got the results they got because they simplified the experiment to the point where participants were second-guessing the

experimenter's intention. It would be interesting to see if their results in the outdoor condition could be replicated with a full battery of tasks, including the Motion-maze and the Transitivity task (reported in Brown & Levinson, 1993; Levinson, 1996b), in which directionality as the issue at stake is much more opaque, being embedded in a more complicated task.

4. Their paper was based on erroneous assumptions about our findings: there was no conflation of variables 'outdoor condition' and 'absolute language' as imagined (some of our absolute results were obtained in a room without any windows, and some of our relative ones were obtained in outdoor conditions). Nor incidentally are the other confounds we are accused of valid for the larger study.[22] Li and Gleitman make a number of further erroneous assumptions, which we cannot correct here, about the ecological and ethnographic backgrounds of the populations we have investigated – the reader should see Levinson (in press) for accurate information. Here we have concentrated on just one issue, the proper analysis of frames of reference and how to experimentally investigate them.

As far as we can see then, our original hypothesis still stands. Not all languages make use of all frames of reference, and the differential use in language predicts the use in non-linguistic tasks. For the reasons we listed in Section 2, we think the correlation suggests that language influences the choice of frames of reference in non-linguistic tasks. But there remains a puzzle: where do the three abstract types of frames of reference come from? As mentioned at the outset, there are many innate neurological and physiological bases governing the relation of an organism to its environment, and these no doubt provide rudiments for the three frames of reference. But these are low-level perceptual and motoric systems, and it is quite another thing to have these available at a conceptual level.

Landau and Gleitman (1985) suggest that 'natural categories' for lexicalization can be recognized during language acquisition because they should display four crucial properties: (i) they should be learnt early in development (well before, say, age 3); (ii) in the course of learning, one should not be able to detect attempts to construe the relevant terms in other, but related, ways; (iii) they should be universally coded in all languages in the 'core vocabulary'; (iv) even under poor input conditions (as where the child has perceptual deficits), they should nevertheless be learnable. By such criteria, there is no evidence that any of the frames of reference are 'natural categories' at a conceptual level – (iii) does not hold for a start, and all the acquisition evidence points to relatively late learning. Western children learn the intrinsic frame first, but this is not mastered in production till nearly 4 years of age (Johnston & Slobin, 1979, p. 538) and relative 'left'/'right' not till as late as 11 (Piaget, 1928; Weissenborn & Stralka, 1984). Interestingly, Tzeltal children learn the absolute system at least as early as the intrinsic system, but again not before 4 years of age,

[22] For example, Li and Gleitman suggest a confound with literacy (pp. 287–288). As we stated above, we had earlier checked for statistical correlations with literacy, age, sex, or indices of culture-change, and found very little (Levinson, in press; Levinson & Nagy, 1998). There is no general correlation between literacy or years of schooling in our sample.

and the system is not mastered fully till about age 7, and they never learn a linguistic relative system (Brown & Levinson, 2000). The pattern is one of slow development right through middle childhood: higher level conceptual representations are constructed late in ontogenesis, in accord with experience and language exposure.

Putting all the facts together, the best account seems to be along the lines of the Karmiloff-Smith (1992) 'representational redescription', whereby during the course of development innate predispositions are progressively reworked into higher level conceptual representations in response to environmental input, so that they become available for a broader range of computation. We argue that language is part of such environmental input driving representational redescription. How does this work?

Notions that are linguistically labeled need to be acquired due to their language specificity. Not all absolute linguistic systems are the same. In Tzeltal, 'downhill' means the quadrant centered on N 345° independent of the local slope, whereas in Guugu Yimithirr, 'north' means the quadrant centered on N 17° (and in other systems axes may not be orthogonal, nor arcs of 90 degrees; see Levinson, in press). Nor are all relative linguistic systems conceptually identical. 'In front of' in Hausa semantically unifies a part of what English *front* means (as in 'in front of me') and a part of what English *behind* means (as in 'behind the tree') (Levinson, in press).

These notions are acquired from matchings of language to situations, where the analysis of those situations may be given either in earlier acquired notions or in simple universal computational primitives (axes, angles, vectors) which have bases in perceptual-motor systems. During this process, a particular type of representation, say a spatial representation based on the absolute frame of reference, is repeatedly employed in the service of the linguistic system. This leads not only to the eventual acquisition of the lexicalized notion, but also to the general privileged status of the representational system that supports the notion. In other words, the representational system becomes readily available for all conceptual purposes, both linguistic and non-linguistic. In this fashion, representational redescription bootstraps us from simpler notions to complex, culture-specific wholes, and it also makes a particular type of representation readily available for conceptual purposes. Under this view, universals do not lie in the exact character of the higher conceptual systems, but just in the fact that expressions of frames of reference in various languages seem to belong to the three main abstract types (absolute, relative and intrinsic), suggesting universal low-level perceptual systems as an ultimate source.

This of course is not at all the view that Li and Gleitman are trying to defend. Let us examine their post-Fodorean doctrine in a little more detail. It has two parts: (a) the idea that all our linguistic categories are direct projections from pre-existing biologically-determined concepts (p. 266), and hence "all languages are broadly similar"(p. 266); (b) the idea that linguistic coding can have no cognitive efficacy or cognitive effects. We think the (a) part is, taken literally, clearly untenable – is the claim *really* that Japanese honorifics, Russian aspect, Bantu noun classes and French gender are biologically-determined concepts, and that American Sign Language is *really* "broadly similar in grammar and lexicon" to American English? We suspect that most such opinions are ill-informed about the range of linguistic diversity – it is, for example, extremely difficult to find even a

few shared concepts that all languages lexicalize (many languages do not, for example, lexicalize terms equivalent to our 'red', 'father', 'if' or 'earth').

But when it comes to (b), we doubt that many scholars would on reflection agree that linguistic coding has no cognitive effects. Even Fodor was – while promulgating an extreme variant of (a) – careful to deny (b):

> I am not committed to asserting that an articulate organism has no cognitive advantage over an inarticulate one. Nor … is there any need to deny the Whorfian point that the kinds of concepts one has may be profoundly determined by the character of the natural language one speaks. … there is no principled reason why the experiences involved in learning a natural language should not have a specially deep effect in determining how the resources of the inner language are exploited. (Fodor, 1975/1992, p. 389)

A great range of commonsense observation and the whole of the history of science shows that the character of a representation system can make a profound computational difference – witness, for example, Arabic numerals over Roman ones. A language provides its learners with a rich but unique representation system, which affords some cognitive operations, enforces others, and inhibits the development of yet further notions. Careful studies have established, for example, that having a verbal color distinction makes a difference to color perception (Kay & Kempton, 1984), or that speaking a language with vs. without number distinctions has effects on the likelihood of perceiving and memorizing quantities in the world (Lucy, 1992b). In the spatial domain, an absolute fixed-bearing system radically changes the computational character of mental maps (McNaughton, Chen, & Markus, 1990). Recent work in Li and Gleitman's own territory suggests that language may play a key role in human cognitive development (see, for example, Spelke and Tsivkin (2001) and other papers in Bowerman and Levinson (2001)). Finally, to understand how language could have an effect on cognition, no outlandish mechanisms need be supposed. To drive a car, you need to acquire new motoric and cognitive skills. To speak Tzeltal, you'll need to be able to do base-20 math in the head, since it has a vintegesimal number system, and more relevantly, you'll constantly need to maintain a mental compass, since 'downhill' denotes a quadrant based on c. N 345°, for without that notion you can't describe where anything is. Further, a central mechanism responsible for the cognitive efficacy of language is provided by one of the corner-stones of cognitive psychology, namely Miller's coding theory of short-term memory limitations (Cowan, 2001; Miller, 1956). Languages are prodigal providers of the 'chunks', the recodings, that get us around the bottleneck of short-term memory limitations (as Miller pointed out; see also Levinson, 1997b). Language-specific chunks thus come to play a central role in our thinking. Resistance to this humble truth – linguistically-motivated categories pervade, change and facilitate our thought – is puzzling. Overall, then, we can see neither empirical basis nor theoretical reasoning in favor of the particular position that Li and Gleitman espouse.

What is the alternative? There are no doubt many views consistent with the evidence in hand, but our own position is the following. Languages differ greatly in the semantic distinctions they make. Speakers of these languages can be shown to code for memory and inference in non-linguistic tasks in a manner congruent with those language-specific

distinctions. Consequently, we suppose that the non-linguistic representation systems used in memory and inference are systematically influenced by the language spoken. We do not think this so surprising, for semantic distinctions require cognitive support, and maintaining a code for memory congruent with the language one speaks will facilitate speaking about any retrieved memory. Such a position is of course consistent with the existence of linguistic and cognitive universals, and with many language-independent aspects of cognition, but it suggests that the languages we inherit and, in a minor way, contribute to, provide us with a wealth of concepts that we would not otherwise have arrived at. The transformative power of those accumulated concepts can be seen both in conceptual development in the individual and in the history of cultures.

Acknowledgements

We thank David Wilkins for his original suggestions regarding the experiments, and Carlien de Witte, Menno Jonker, and Wilma Jongejan for helping in data collection. We would also like to acknowledge two anonymous referees who insisted on the clarification of our arguments.

References

Bowerman, M. & Levinson, S. C. (Eds.) (2001). *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.

Brown, P. (2001, June). *Cultural factors in learning an 'absolute' spatial system*. Paper presented at the meeting of the Piaget Society, Berkeley, CA.

Brown, P., & Levinson, S. C. (1993). *Linguistic and non-linguistic coding of spatial arrays: explorations in Mayan cognition* (Working Paper No. 24). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute.

Brown, P., & Levinson, S. C. (2000). Frames of spatial reference and their acquisition in Tenejapan Tzeltal. In L. Nucci, G. Saxe & E. Turiel (Eds.), *Culture, thought, and development* (pp. 167–197). Hillsdale, NJ: Erlbaum.

Burgess, N., Jeffery, K., & O'Keefe, J. (1999). *The hippocampal and parietal foundations of spatial cognition*, Oxford: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24* (1), 87–154.

Dennett, D. (1991). *Consciousness explained*. Boston, MA: Little, Brown & Co..

Fodor, J. (1975). *The language of thought*. New York: Crowell.

Fodor, J. (1992). How there could be a private language. In B. Beakley & P. Ludlow (Eds.), *The philosophy of mind* (pp. 385–391). Cambridge, MA: MIT Press Reprinted from *The language of thought*, by J. Fodor, 1975, New York: Crowell.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gentner, D., & Goldin-Meadow, S. (Eds.). (in press). *Language in mind: advances in the study of language and thought*. Cambridge, MA: MIT Press.

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

Hughes, H. C. (1999). *Sensory exotica: a world beyond human experience*. Cambridge, MA: MIT Press.

Johnston, J. R., & Slobin, D. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language*, *6*, 529–545.

Karmiloff-Smith, A. (1992). *Beyond modularity: a developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, *86*, 65–79.

Kita, S., Danziger, E., & Stolz, C. (2001). Cultural specificity of spatial schemas, as manifested in spontaneous gestures. In M. Gattis (Eds.), *Spatial schemas in abstract thought* (pp. 115–146). Cambridge, MA: MIT Press.

Landau, B., & Gleitman, L. (1985). *Language and experience: evidence from the blind child*. Cambridge, MA: Harvard University Press.

Landau, B., & Jackendoff, R. (1993). 'What' and 'Where' in spatial language and spatial cognition. *Behavioral & Brain Sciences. 16*, 217–238.

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

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (Eds.), *Language and space* (pp. 77–108). Cambridge, MA: MIT Press.

Levinson, S. C. (1992). *Language and cognition: cognitive sequences of spatial description in Guugu Yimithirr* (Working Paper No. 13). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute.

Levinson, S. C. (1996a). Language and space. *Annual Review of Anthropology*, *25*, 353–382.

Levinson, S. C. (1996b). Frames of reference and Molyneux's question: cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Levinson, S. C. (1996c). *The role of language in everyday human navigation* (Working Paper No. 38). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute.

Levinson, S. C. (1997a). Language and cognition: the cognitive consequences of spatial description in Guugu Yimithirr. *Journal of Linguistic Anthropology*, *7* (1), 98–131.

Levinson, S. C. (1997b). From outer to inner space: linguistic categories and non-linguistic thinking. In J. Nuyts & E. Pederson (Eds.), *With language in mind: the relationship between linguistic and conceptual representation* (pp. 13–45). Cambridge: Cambridge University Press.

Levinson, S. C. (in press). *Space in language and cognition: explorations in linguistic diversity*. Cambridge: Cambridge University Press.

Levinson, S. C., & Nagy, L. (1998). *Look at your southern leg: a statistical approach to cross-cultural field studies of language and spatial orientation*. Unpublished working paper, Max Planck Institute of Psycholinguistics, Nijmegen.

Levinson, S. C., & Schmitt, B. (1993). Animals in a row. In *Cognition and Space Kit Version 1.0* (pp. 65–69). Nijmegen: Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics.

Li, P., & Gleitman, L. (in press). Turning the tables: language and spatial reasoning. *Cognition*.

Lucy, J. (1992). *Language diversity and thought*, Cambridge: Cambridge University Press.

Lucy, J. (1992). *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*, Cambridge: Cambridge University Press.

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

Miller, G. (1956). The magical number seven, plus or minus two. *Psychological Review*, *63* (2), 81–97.

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

Pederson, E. (1995). Language as context, language as means: spatial cognition and habitual language use. *Cognitive Linguistics*, *6*, 33–62.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language*, *74*, 557–589.

Pederson, E., & Schmitt, B. (1993). Eric's maze task. In *Cognition and Space Kit Version 1.0* (pp. 73–76). Nijmegen: Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics.

Piaget, J. (1928). *Judgment and reasoning in the child*, London: Routledge.

Pick Jr., H. L. (1993). Organization of spatial knowledge in children. In N. Eilan, R. McCarthy & B. Brewer, *Spatial representation* (pp. 31–42). Oxford: Basil Blackwell.

Pinker, S. (1994). *The language instinct*. New York: Morrow.

Rock, I. (1992). Comment on Asch & Witkin's 'Studies in space orientation II'. *Journal of Experimental Psychology: General*, *121* (4), 404–406.

Spelke, E., & Tsivkin, S. (2001). Initial knowledge and conceptual change: space and number. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 70–100). Cambridge: Cambridge University Press.

Tversky, B. (1996). Spatial perspective in descriptions. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (Eds.), *Language and space* (pp. 463–492). Cambridge, MA: MIT Press.

Wassmann, J., & Dasen, P. (1998). Balinese spatial orientation: some empirical evidence of moderate linguistic relativity. *Journal of the Royal Anthropological Institute (MAN)*, *4*, 689–711.

Weissenborn, J., & Stralka, R. (1984). Das Verstehen von Missverstaendnissen: Eine ontogenetische Studie. *Zeitschrift fuer Literarturwissenschaft und Linguistik, 14 (55), 113–134.*

Wilkins, D. P. (in press). Arrernte pointing. Why pointing with the index finger is not a universal (in socio-cultural and semiotic terms). In S. Kita (Ed.), *Pointing: where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum.

Wilkins, D. P., & Levinson, S. C. (in press). *The grammars of space*. Cambridge: Cambridge University Press.