



Turning the tables: language and spatial reasoning

Peggy Li, Lila Gleitman*

Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut Street, Suite 400A, Philadelphia, PA 19104, USA

Received 8 December 1999; received in revised form 20 February 2001; accepted 1 February 2002

Abstract

This paper investigates possible influences of the lexical resources of individual languages on the spatial organization and reasoning styles of their users. That there are such powerful and pervasive influences of language on thought is the thesis of the Whorf–Sapir linguistic relativity hypothesis which, after a lengthy period in intellectual limbo, has recently returned to prominence in the anthropological, linguistic, and psycholinguistic literatures. Our point of departure is an influential group of cross-linguistic studies that appear to show that spatial reasoning is strongly affected by the spatial lexicon in everyday use in a community (e.g. Brown, P., & Levinson, S. C. (1993). *Linguistic and nonlinguistic coding of spatial arrays: explorations in Mayan cognition* (Working Paper No. 24). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics; *Cognitive Linguistics* 6 (1995) 33). Specifically, certain groups customarily use an externally referenced spatial-coordinate system to refer to nearby directions and positions (“to the north”) whereas English speakers usually employ a viewer-perspective system (“to the left”). Prior findings and interpretations have been to the effect that users of these two types of spatial system solve rotation problems in different ways, reasoning strategies imposed by habitual use of the language-particular lexicons themselves. The present studies reproduce these different problem-solving strategies in speakers of a single language (English) by manipulating landmark cues, suggesting that language itself may not be the key causal factor in choice of spatial perspective. Prior evidence on rotation problem solution from infants (e.g. Acredolo, L.P. (1979). Laboratory versus home: the effect of environment on the 9-month-old infant’s choice of spatial reference system. *Developmental Psychology*, 15 (6), 666–667) and from laboratory animals (e.g. Restle, F. (1975). Discrimination of cues in mazes: a resolution of the place-vs.-response question. *Psychological Review*, 64, 217–228) suggests a unified interpretation of the findings: creatures approach spatial problems differently depending on the availability and suitability of local landmark cues. The results are discussed in terms of the current debate on the relation of language to thought, with particular emphasis on the question of why different cultural communities favor different perspectives in talking about space. © 2002 Published by Elsevier Science B.V.

Keywords: Turning the tables; Language; Spatial reasoning

* Corresponding author. Fax: +1-610-660-0176.

E-mail addresses: pegs@psych.upenn.edu (P. Li), gleitman@psych.upenn.edu (L. Gleitman).

1. Introduction

Language has means for making reference to the objects, relations, properties, and events that populate our everyday world. It is possible to suppose that these linguistic categories and structures are more-or-less straightforward mappings from a preexisting conceptual space, programmed into our biological nature: humans invent words that label their concepts. This perspective would begin to account for the fact that the grammars and lexicons of all languages are broadly similar, despite historical isolation and cultural disparities among their users.

But having assigned the language identities to underlying conceptual identities, what are we to make of the – also real – differences among languages? Could it be that the situation is symmetrical? That insofar as languages differ from each other, there are corresponding differences in the modes of thought of their users? More marvelously by far, could the linguistic differences be the original causes of distinctions in the way peoples categorize and reason? Benjamin Whorf and Edward Sapir offered a positive answer to such questions. In Sapir's words:

Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression ... the "real world" is to a large extent unconsciously built up on the language habits of the group (E. Sapir as quoted by Whorf, 1941/1956, p. 134).

And Whorf:

Language and culture are constantly influencing each other. But in this partnership the nature of the language is the factor that limits free plasticity and rigidifies channels of development in the more autocratic way (Whorf, 1941/1956, p. 156).

The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented as a kaleidoscopic flux of impressions which has to be organized by our minds – and this means largely by the linguistic systems in our minds (Whorf, 1940/1956, p. 213).

The influence of this position diminished considerably in academic favor during the last half of this just past century, for two main reasons. First, the universalist position of Chomskian linguistics, with its potential for explaining the striking similarity of language learning in children all over the world, captured the imagination of a generation of scholars. Second, a series of experimental studies during the 1960s and 1970s, documenting the independence of hue and brightness perception from linguistic color-naming practices, seemed to settle the controversy in favor of the universalists (see particularly Berlin & Kay, 1969; Heider, 1972; Heider & Oliver, 1972; Jameson & Hurvich, 1978).

Recently, however, a number of discussions and experimental studies have reawakened interest in the question of how language may shape thought. Rightly, the studies of color vision and naming, elegant and compelling as they were, have been judged too topically

narrow as the basis for writing off the position as a whole. Moreover, some recent findings appear to challenge aspects of the original findings and their interpretation (Davies, 1998; Levinson, 2000; Roberson, Davies, & Davidoff, 2000). Finally, effects of language use, or indeed any learning effects at all, would be least likely in peripheral, low-level perceptual processes such as hue discrimination. Therefore they do not constitute a fair test (Lucy, 1996).

Many newer studies have focused on a more central perceptual domain: commonalities and differences in how languages treat spatial properties and relations (Jackendoff, 1996; Talmy, 1978, 1983). To be sure, ultimately linguistic-spatial categories must be built upon a universal perceptual base originating in brain structure shared by all nonpathological humans (Landau & Jackendoff, 1993) with homologies across most of the animal kingdom (Gallistel, 1990, 1992). But within these biologically imposed constraints on the experience that language must describe, there is room for difference. This possibility is instantiated in several cross-linguistic studies of differences in spatial encoding (see the collections of articles in Bloom, Peterson, Garrett, & Nadel, 1996; Gumperz & Levinson, 1996; Pütz & Dirven, 1996). Children appear to find it easy to acquire whatever spatial-linguistic categories their language makes available in its simplest vocabulary and phraseology (Bowerman, 1996a,b; Bowerman & Choi, 2001; Choi & Bowerman, 1991; Choi, McDonough, Bowerman, & Mandler, 1999) and to tailor their speech to accommodate to the domains of applicability of the language-specific terms (Papafragou, Massey, & Gleitman, 2002; Slobin, 1991). Finally, subjects in the psychological laboratory can make implicit use of this language-specific syntax-to-semantics patterning as probabilistic evidence for the interpretation of new words (Imai & Gentner, 1997; for discussion, Fisher & Gleitman, *in press*).

We pursue here the further question of whether linguistic differences in the mapping of space onto language impact the ways that members of a speech community come to conceptualize the nonlinguistic world, as Whorf and Sapir would have it. Do the differences in how people talk create differences in how they think? Several commentators have posited that language distinctions in the spatial domain actually do have strong and significant effects on category formation and deployment in tasks requiring spatial reasoning. For example, in the words of Gumperz and Levinson:

...Consider a language that has no terms for 'in front', 'behind', 'left', 'right', and so on ... preferring instead to designate all such relations, no matter how microscopic in scale, in terms of notions like 'North', 'South', 'East', 'West', etc. Now a speaker of such a language cannot remember arrays of objects in the same way as you and I, in terms of their relative location from a particular viewing angle... Speakers of [such] languages do indeed remember spatial arrays differently in ways that can be demonstrated experimentally and observationally (Gumperz & Levinson, 1996, pp. 26–27).

The experiments to which these commentators are alluding form the point of departure for those we will present in this paper. Levinson and his colleagues have carried out an extensive cross-linguistic research inquiry focusing on the relationship between linguistic patterning and spatial reasoning (Brown & Levinson, 1992, 2000, and other reports that we will cite as we go along). The general finding is that linguistic patterning and spatial

reasoning strategies covary in striking ways, suggesting a causal relation between the two functions. Our immediate purpose in the present paper is to add new evidence that may help in evaluating the direction of causal flow from which these correlations might arise. Perhaps it is the habitual linguistic practice in these communities that determines the relevant modes of thought, as Levinson seems to imply in the quotation above. On the other hand, it could be that cultural differences in modes of thought render certain linguistic usages handier than others, and thus influence their prominence and frequency of use. Perhaps both such mechanisms are at work with, in Whorf's words, "language and culture constantly influencing each other". In any case, it will be of interest to probe for why cultural groups by some implicit but systematic process make their distinctive terminological and spatial-conceptual choices. In the remainder of this introduction, we will summarize the issues and prior findings so as to set the stage for a further experimental review of this topic.

1.1. Cross-linguistically varying spatial categories

There are several ways in which languages express spatial regions and orientations. A major cut is between spatial expressions referring to directions and locations relative to the viewer himself (*deictic* or *egocentric* descriptors such as "left" and "right" also called *body-centered*, *response*, or *relative* in various literatures to which we will refer in this paper) versus externally referenced; that is, relative to landmarks or coordinates outside the observer (*allocentric* descriptors, also called *geocentric*, *place*, *cue-based*, or *absolute*). The allocentric classification is usefully subdivided into systems that refer to the intrinsic properties of external objects ("the foot of the hill", "the front of the house", "the nose of the plane"), and those that refer to local ("near the toolshed") or more global landmarks and regions ("in Cleveland", "east").¹

Allocentric terms like "east" are quite complex. Literally, of course, *east* refers to position relative to a particular landmark: the place whence the sun rises, or where the celestial bodies are first seen at dawn. But in practice – in psychological fact – terms such as east often make immediate reference to more restricted points or areas, e.g. the relation of New York to San Francisco or toward First from Fifth Avenue (for Manhattanites), or to topological features of the local landscape such as "uphill from". An example of this last subtype, described in Brown and Levinson (1992, p. 596), is the Mayan language Tzeltal,

¹ Several cross-cutting distinctions and complexities in how language renders space are ignored here. For the rotation issue on which this article focuses, what comes to the fore is the difference between viewer-referenced (relative egocentric) and externally-referenced (absolute allocentric) systems. To see how flexibly and variously reference can be made to spatial relations within a single language (and even by a single individual), consider the expression "to the left of the chair" uttered by a speaker facing the front of a chair to a listener facing in that same direction. The speaker could be referring to intrinsic properties of the chair, but either taking the chair's own perspective, or taking one's own (the speaker's) perspective; the expression itself is therefore perfectly ambiguous. Interpretation becomes even harder if the speaker and listener are facing the back of the chair or viewing it from different angles. Note also that there are incompatibilities in terminology among investigators, e.g. in some treatments the term *allocentric* refers only to the viewpoint of the other party in a conversation (e.g. "Your left arm looks swollen") rather than to any externally referenced perspective. For important psychological taxonomies of spatial organization and spatial terminology, see Levelt (1996), Levinson (1996), and Newcombe and Huttenlocher (2000).

as spoken by about 15,000 individuals in the area of Municipio Tenejapa, in Chiapas, Mexico. The Tenejapans' village is on a hill:

...there is a system of 'uphill'/'downhill' orientation that is fundamental to the spatial system ... based on the overall inclination of the terrain of Tenejapa from high South to low North, so that [the term for] 'uphill' (and correspondingly, 'downhill'...) [makes] primary reference to the actual inclination of the land ... the terms may be used on the flat to refer to cardinal orientations, or prototypical 'uphill' direction. This system then replaces our use of left/right in many contexts: when there are two objects oriented such that one is to the South of the other, it can be referred to as the 'uphill' object... Now, curiously, this system of North/South alignment is not complemented by a similar differentiation of the orthogonal. There is a named orthogonal ... but the term is indifferent as to whether it refers to East or West; what it really means is 'transverse to the incline'. So there is a three-way distinction.

Apparently, then, some languages either lack body-centered spatial terminology or restrict its use very narrowly. Moreover, in most cases even a community whose language instantiates both deictic and allocentric terms may heavily favor one of them, at least for small-scale description. Though in familiar brands of American English we *could* say "Give me the spoon that's northeast of your teacup", this sounds pretty ludicrous. We are strongly inclined to say "...to the left of your teacup" instead. Thus, as Whorf pointed out, languages that share semantic and structural resources often still differ in "fashions of speaking", preferred "ways of analyzing and reporting experience which have become fixed in the language" and might affect the everyday modes of thought of their users (Whorf, 1941/1972, p. 148).²

Conceptually, the first step in the cross-linguistic project of Levinson and his colleagues has been to document the striking typological differences in spatial vocabulary by which speakers typically describe the positions of objects in space. Standardized procedures were established for eliciting these spatial descriptions from the native speakers of a broad range of languages, living in several small-scale traditional cultural communities (e.g. Mayans, Austronesians) and two large-scale urbanized communities (Japanese, Dutch). Pederson, Danziger, Wilkins, Levinson, Kita, and Senft (1998) (henceforth PDWLKS) detail this elicitation procedure (called the *man and tree test*; a description and replication for American speakers of English will be reported in Section 2 of the present paper). This procedure identified several different usage preferences across languages, including heavily allocentric and egocentric systems, and some mixed ones.

Once these differing language use patterns were established, the question was whether

² It is these "fashions of speaking", typical linguistic practice as opposed to formal capability, that chiefly interested Whorf himself, and the Brown–Levinson–Pederson group of investigators as well. They write regarding the spatial-terminological studies that we are reviewing here. "Instead of comparing grammar alone, we compare linguistic *practice* – the meaning patterns that consistently emerge from domain-directed interactive discourse of members of a given speech community" (Pederson et al., 1998, p. 565). And more particularly, "Here we have empirically demonstrated that the choice of linguistic frame of reference for projection of coordinates in tabletop space is quite differently distributed across languages – even in a standardized referential and interactive context" (Pederson et al., 1998, p. 574).

and how these impinged on spatial reasoning patterns. To find out, the investigators asked native speakers of these languages to memorize, recall, and reconstruct spatial arrays after being spatially reoriented. The question under experimental review was whether differences in the spatial vocabularies and spatial usage patterns of the subject populations predicted differences in their performance on these spatial identification and reconstruction tasks.

Though there were some procedural differences, most of the rotation experiments described by PDWLKS were variations on a single theme concerning the effects of rotation. The main paradigm was the *animals in a row test*, schematized in Fig. 1, adapted from Brown and Levinson (1993) (further details of this procedure are given in Section 2 of the present paper). The experimental subjects memorize the positions of three animals, arrayed in a line in front of them on a tabletop, the “Stimulus Table” (Panel 1). The animals are then removed from view. After a brief delay, the subjects are turned around or escorted to another table oriented differently (usually, by 180 degrees; the “Recall Table”, Panel 2). There, the subjects are handed the three original animals in random order and asked to position the animals “in the same way as before”.

There are two ways to solve the puzzle. Suppose that the animals were originally displayed with their noses facing north, which happens to be to the subject’s right as s/he faces the Stimulus Table. They can now be set up on the Recall Table still facing north, the allocentric or absolute solution (Panel 3a). Or, as shown in Panel 3b, they can be set up

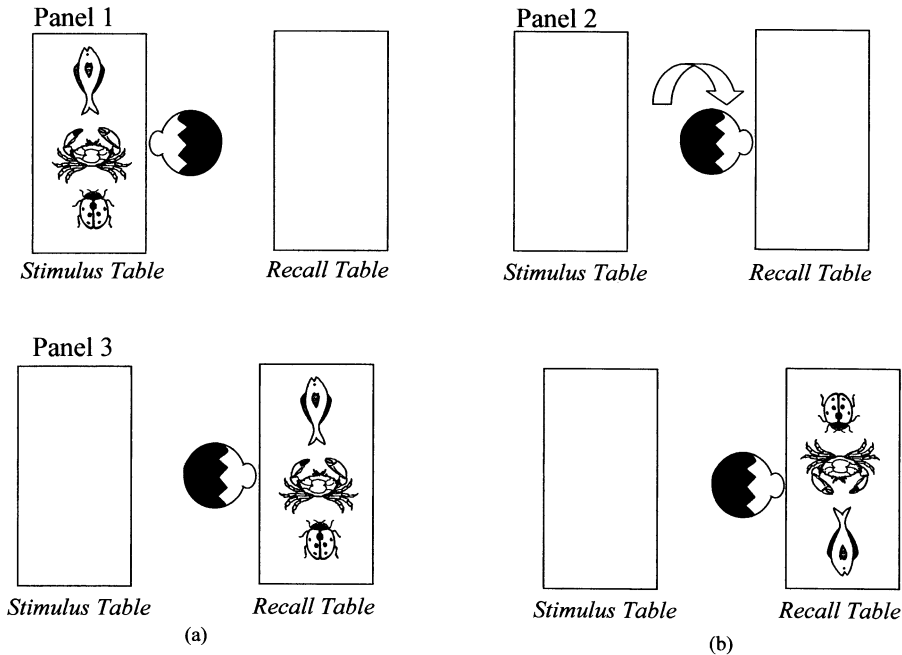


Fig. 1. The Animals in a Row task. Panel 1: Stimulus Table with array of toy animals; Panel 2: rotated 180 degrees at Recall Table; Panel 3a: the “absolute” rearrangement; Panel 3b: the “relative” rearrangement (adapted from Pederson et al., 1998).

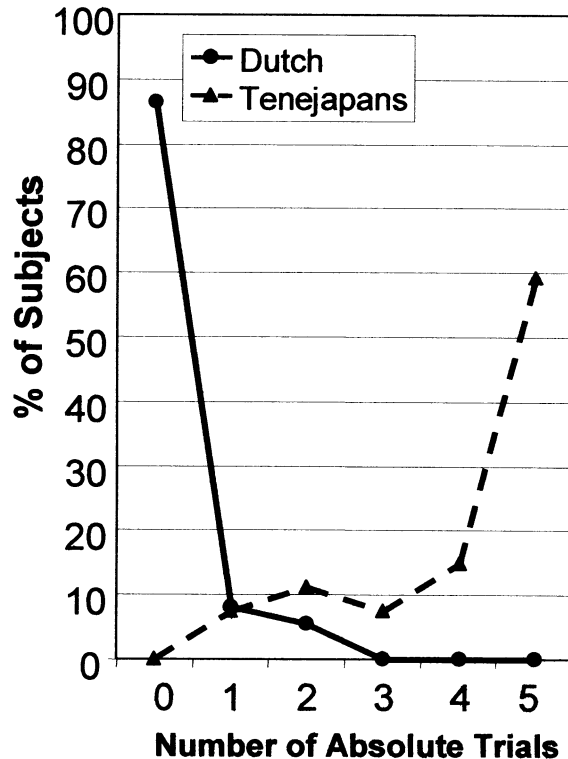


Fig. 2. Proportions of absolute choices for speakers of two languages (adapted from Brown & Levinson, 1993).

on the Recall Table still facing right, the egocentric or relative solution. PDWLKS reported that speakers of languages classified as relative solved the spatial task relatively and vice versa.

Detailed cross-linguistic outcomes of these studies for the Tenejapan Mayan (absolute) and Dutch (relative) subject groups are graphically shown in Fig. 2, adapted from Brown and Levinson (1993). The figure plots the distribution of absolute responses by subjects from the two languages.³ The overwhelming majority of Tenejapan speakers quite consistently rearranged the animals so that after rotation they were heading in the same cardinal direction, whereas Dutch speakers overwhelmingly often rearranged the animals such that if these had been facing left on the first table, they faced left after rotation as well. The difference is that *what is north* does not vary under rotation, while *what is left* certainly does.

³ PDWLKS generally plot percent absolute responses against number of subjects. The resultant graphs can be visually misleading when the number of subjects in various conditions differs by a wide margin. Therefore, we always replot these results by percent rather than number of subjects.

1.2. Summary and a question

Levinson and his colleagues interpret results like those just discussed as showing that this terminological distinction among languages influences spatial reasoning in a dramatic and straightforward way. After all, Fig. 2 (and related results for several other languages reported by PDWLKS) documents a powerful correlation between speaking a language classified as absolute by elicited speech measures (the Man and Tree test) and solving the rotation problem absolutely in the Animals in a Row test. However, it is just as possible to interpret the correlation in reverse; namely, that culturally differing spatial reasoning strategies lead these groups to deploy different terminologies, those that are consistent with their reasoning. And it is possible that some third variable that differed for the subject populations is responsible both for the linguistic difference between them and for their approaches to spatial tasks. Just because linguistic and cultural practice are so often and usefully intermeshed, it is difficult to tell cause from effect.

One way to investigate the direction of causality is by changing the language spoken by one cultural group, maintaining their spatial environments in other regards, to see if their reasoning or categorization behavior changes as well. For example, consider the well-known claim by Whorf (1956) that Eskimo populations have a large number of words differentiating among types of snow.⁴ Would such Eskimo populations be affected in their discrimination of snow types if they continued to live where and as they now do, but came to speak English rather than an Eskimo language? Here we will try to break the confound between language and cultural practice from the other end; namely, by varying the environment in which speakers of a single language carry out a task. The analogy is to testing the Eskimos out of the snow. The purpose is to assess the degree to which the language forms or typical usage patterns hold their users in thrall, as Whorf and Sapir maintained, versus the degree to which local circumstances (“snow all around”) legislate fashions of speech.

Based on this reasoning, Section 2 of this paper presents two experimental studies of the malleability of spatial reasoning strategies *within a single* (English-speaking) *linguistic community*. Specifically, our question is this: if we manipulate the circumstances in which monolingual English speakers solve the spatial rotation problem, will they readily shift between the egocentric and allocentric strategies? An analogue to this within-language kind of test for linguistic determinism comes from the work of Terry Au (1983) who reinvestigated findings from Alfred Bloom (1981). Bloom had studied Chinese speakers' story understanding and come to the conclusion that Chinese speakers gave counterfactual interpretations to counterfactual stories less often than English speakers. Bloom related these findings to the apparent fact that Chinese does not have a grammatical marker of

⁴ We are referring here to Whorf's famous discussion of snow (Whorf, 1940/1956, p. 216). According to Pullum (1991), citing Martin (1986), Whorf was incorrect in thinking that Inuit or Yupik dialects are rich in snow words – various morphological and translation confusions and unchecked references led to the gross inflation of a snow vocabulary no more extensive than that found in English. Still, the oft-mentioned snow case is a handy example of the way that Whorf and his various intellectual descendents have thought about the relations between language and thought. See also Kemmerer (1999) for an important study on language divergence in spatial terminology in the presence of sameness of their speakers on neurophysiological and behavioral measures.

counterfactuality (analogous to the English subjunctive). Au's approach was to vary the stories themselves against the linguistic factors, i.e. to find plausible situations in which specific language and subjects' response characteristics could be prized apart. Among other findings, she showed circumstances in which monolingual Chinese speakers gave significantly more counterfactual responses to counterfactual-containing stories than did monolingual English speakers and Chinese-English bilinguals. She concluded that "the mastery of the English subjunctive is probably quite tangential to counterfactual reasoning". Following Au's lead, we attempt here to study variation in spatial reasoning within a single language community.

Section 3 of this paper then compares these outcomes to prior findings concerning spatial reasoning *in populations for which a linguistic-causal explanation is simply out of the question*: nonlinguistic (animal) and prelinguistic (human infant) groups. These results in hand, Section 4 considers some factors that covary in communities that adopt different spatial terminologies, factors that may bear some of the explanatory burden for differences in spatial reasoning. These include variation in schooling, urbanization, insularity, and geographical cohesion of communities. The quest is for a unified explanation of when and why individuals or populations – be they speakers of one language or another, prelinguistic humans, or members of other species – solve spatial problems in varying ways.

2. Spatial reasoning in varying frames of reference: an experimental review

Each language population in the PDWLKS experiments was tested in its own community, usually under social and geographical frame-of-reference contexts that commonly exist there. For instance, the Tenejapan population was tested on its hill, out of doors, near a largish rectangular house. The Dutch population was tested indoors in a laboratory room. It is reasonable enough to test each group in an environmental setting that it finds natural, that is, to honor the ecological conditions existing for the peoples tested. But the upshot is that two or more factors may have been varying at once – the language and the larger frame of reference (geographical and social) in which the spatial task was to be solved. Which caused the characteristically different behavior of these groups? To find out, we manipulated the spatial circumstances under which our subjects were tested, but we held language constant. Subjects were drawn from a single cultural and linguistic subgroup: monolingual native-English speaking undergraduates at the University of Pennsylvania. The experimental question was whether we could induce Tenejapan-like and Dutch-like spatial reasoning behavior in this population by systematically varying the spatial contexts in which they are tested. All these manipulations employed the Animals in a Row rotation test (Fig. 1), the task for which the PDWLKS research group has reported the most systematic cross-linguistic evidence.

2.1. Experiment 1: typological classification: the Man and Tree test

It is obvious that English speakers, like Dutch speakers, typically describe the positions of nearby objects relatively ("the teacup to your left" rather than "the teacup to your east"). But for completeness, we tested a group of eight English speakers (four pairs, as next

described) using the Man and Tree test developed by the PDWLKS group (most particularly, by Penelope Brown) to find out how they would actually describe and understand small-scale spatial arrays. This number of subject pairs is comparable to the numbers as described in PDWLKS, and the test materials and procedure were designed to follow theirs closely. Subjects were assigned at random to one of the four pairs. Within each pair, each subject was assigned to either the Director or Matcher role. The Director and Matcher sat side by side facing in the same direction, in a laboratory room. The room had a large window through which could be seen the university library across the street. The Director was seated with his right side facing this window. In front of each of the members of the pair was a set of 14 photographs. Director's and Matcher's sets were identical, but were differently arrayed in front of them. Eight of these pictures showed a toy girl in some position relative to a toy umbrella (Fig. 3b). The remaining pictures ("distractors", see a sample of these in Fig. 3a) were of scenes in which spatial relations between objects were not the crucial contrasting factor. A screen divided Director and Matcher such that they could not see each other's photographs or each other's gestures. The procedure was for a Director to choose (any) one of the photographs and describe it so that the Matcher could find his own replica of it in his own array, e.g. "the girl with her back to the umbrella". Back and forth conversation between Director and Matcher was encouraged so that a unique photo was eventually chosen by the two. The procedure was repeated until all photos were described and chosen for all four pairs of subjects.

The question of interest is the spatial language used (chiefly by the Directors, but sometimes also by the Matchers trying to get clarification or detail). Though PDWLKS did not report percentages for their Dutch subjects, they state that these subjects never gave absolute responses and "overwhelmingly" gave relative responses (Pederson et al., 1998, p. 572). Speakers of both languages called absolute and those called relative in PDWLKS also produced some intrinsic descriptions (e.g. "facing me towards the east", for Longgu; Pederson et al., 1998, p. 574). Our English speakers' responses conformed to the Dutch-Japanese pattern: 82% of responses were relative ("girl is to left of the umbrella"), 18% were intrinsic ("girl facing umbrella") and 0% were absolute ("girl to North of umbrella" or "uphill of umbrella").

We conclude that English falls into the PDWLKS *relative* group of languages in the sense of typical language practice (Whorf's "fashion of speaking"). Though its speakers (just like Dutch speakers) are quite capable of speaking of north and west, and though their language provides easy means, they rarely do so in describing objects that are in close proximity in a visual world being co-observed with a listener. PDWLKS' studies of spatial reasoning concentrated experimental attention on language communities which, like the American university students just described and like the Dutch and Tenejapans, were quite consistent as between the relative or absolute choices of description in this task (Pederson et al., 1998, p. 574).

2.2. Experiment 2: spatial reasoning: Animals in a Row

With English-speaking undergraduates drawn from the same population tested in Experiment 1, we replicated the ingeniously designed PDWLKS task now described in further detail (with variations that we will describe as we go along). The materials and

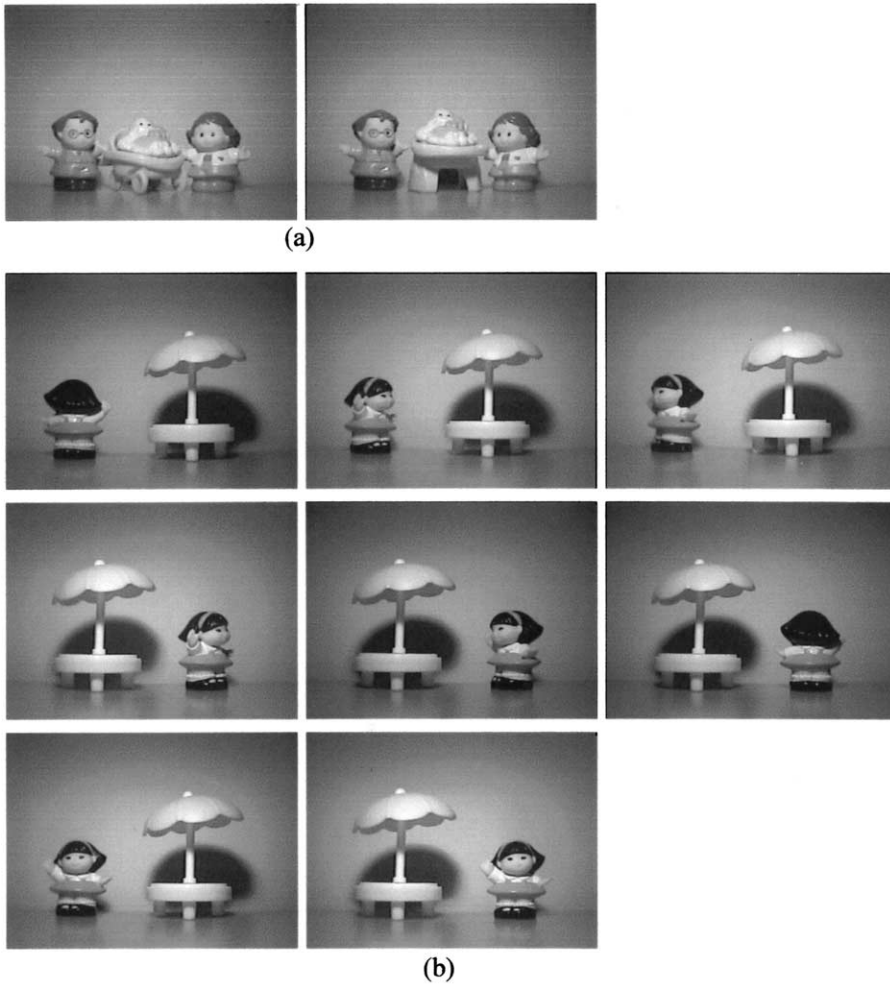


Fig. 3. The Man and Tree task: (a) sample distracter items; (b) test items. In this adaptation of the Brown and Levinson (1993) task we used pictures with an umbrella and a girl rather than a tree and a man.

general procedure are shown in Fig. 1. Subjects were tested individually. The subject was first seated on a swivel chair at the Stimulus Table and asked to study an array of three toy animals (out of five animals in the total test-set of animals) that had been placed there in advance. Subjects were also asked in advance to name each animal to assure that there was a consensual word for each. Because these were a toy dinosaur, lion, dog, rabbit, and elephant, there was no disagreement about the labels. Each toy was symmetrical along its longitudinal axis and was approximately 10 cm long (in one manipulation, described later, we used larger toy animals). Both in practice and test trials, the experimenter always set up the animals so that their noses pointed in the same cardinal direction, either north or south, depending on the trial (equivalently, in the testing circumstances, pointing toward the

subject's left or right). The subject was instructed to study the array for as long as he liked. Now the animals were scooped up by the experimenter. In an initial practice trial, the subject was then handed the three animals and asked to set them up again ("make it the same") on the Stimulus Table. All of our subjects always did this correctly.

Now the experiment proper began. The experimenter again set up three of the animals on the Stimulus Table, as the subject watched. The subject studied this new array as long as she liked, followed by a 30 s delay. The subject was then swiveled on her chair 180 degrees to face the Recall Table, which was empty, and handed the three animals in random order. The instructions to the subject were very simple: "Make it the same". This procedure was repeated (of course, changing the particular animals and their arrangement on each trial) with each subject five times. Sometimes subjects asked for clarification (of what we meant by "the same"). The experimenter blandly responded "Just make it the same" and, improbably enough, the subject then always said "OK" and carried out the task.⁵ To repeat, as schematized in Fig. 1, there are two correct ways that subjects could reconstruct the array on the Recall Table. With rare exceptions, each of our subjects was consistent across the five trials as to the absolute versus the relative choice. We now describe three conditions under which (different) subject groups of monolingual English speakers were tested.

2.2.1. Experiment 2a: landmarks in the reference world beyond the tabletop (Blinds-Down/Blinds-Up; Outdoors)

This experiment and those that follow altered the context in which subjects carried out the Animals in a Row task by adding implicit landmark or bearing cues of various kinds, both on and outside the experimental tabletops. For after all, the results found for the Dutch and Tenejapan subjects might be attributable to the differential availability of such landmark information in a laboratory room versus in some other spatial setting, e.g. an outdoor environment or a cluttered room. This landmark factor varied in the studies reported by PDWLKS. This is because it is often inappropriate or infeasible to test members of certain communities in the sterile laboratory rooms familiar to European and American undergraduates. For example, the Tenejapans whose data are summarized in Fig. 2 were tested out of doors, on their hill, in front of a large building.

2.2.1.1. Subjects, materials, and procedure Forty subjects, tested individually, participated, ten in each of two indoor conditions and 20 in an outdoor condition. In the indoor conditions, subjects were tested in a laboratory room which was essentially featureless except for a floor-to-ceiling window forming one side of the room. This room was where we had previously tested the four pairs of director–matchers. The testing tables were set up relative to the window in such a way as to be similar to the

⁵ Small procedural details differ between the original Brown and Levinson task and the one reported here: these authors (see also PDWLKS) had some distance between the tables, to avoid their subjects' sometime tendency to look forward to the Recall Table while they were at the Stimulus Table. With our undergraduate subjects, it was sufficient to tell them not to look toward the other table. So the Brown and Levinson subjects were walked, rather than swiveled, between the two tables. This walk-time added a few seconds between training and test, so the delay for Brown and Levinson subjects might have been slightly longer than for our subjects. And of course, our toys were different from their toys.

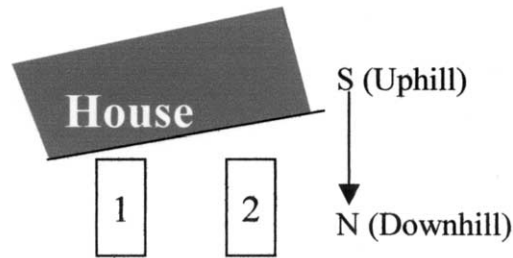


Fig. 4. Test environment for the Tenejapan subjects (following Brown & Levinson, 1993).

placement of the house that had been visible to Brown and Levinson’s Tenejapan subjects (Fig. 4), as shown in Fig. 5. Ten subjects were tested in this room with the blinds pulled down so that they could not see what lay behind the window. For another ten subjects, the blinds were in their raised position. Under this latter condition, the subjects if they looked toward the window would view the familiar sight of the university library that lay across the street from the laboratory. No mention of the blinds or of landmarks was made to any of the subjects. Finally, 20 subjects were tested outdoors in a grassy area about an acre in size on the university campus, ranged around which were both tall buildings (a student high-rise apartment house) and short ones (a large house and a church; Fig. 6).

2.2.1.2. Results The results in terms of proportions of absolute responses are shown in Fig. 7. The subjects in the Blinds-Down condition behaved much as the Dutch subjects in Brown and Levinson (see Fig. 2 for comparison). The subjects in the Blinds-Up condition and the Outdoor condition behaved differently, yielding a U-shaped distribution that lies somewhere between the prior Dutch and Tenejapan results. Even with the small number of subjects tested in each indoor condition, the difference between Blinds-Up and Blinds-

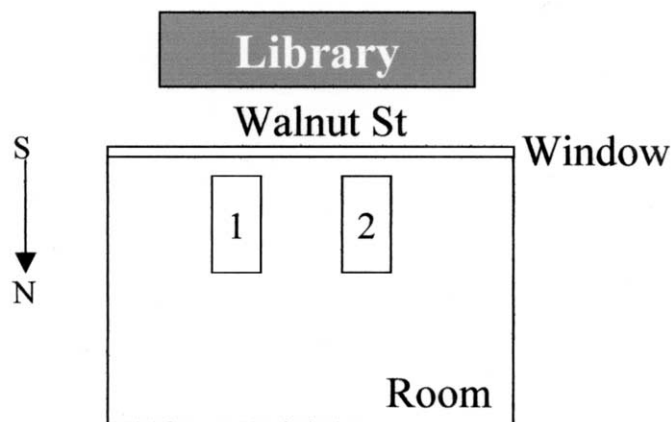


Fig. 5. Test environment for the indoor conditions of Experiment 2a. The University Library is to the south of the tables just as was the house in Fig. 4.

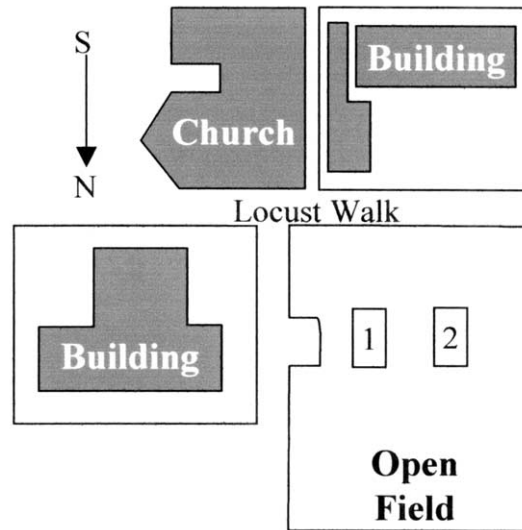


Fig. 6. Test environment for the Outdoor condition of Experiment 2a. Buildings are to the south and southeast of the grassy test area.

Down subjects approaches significance using the same evaluative instrument used in the prior studies by Brown and Levinson (Mann–Whitney U -test, $P = 0.056$). For the outdoor group with its larger number (20) of subjects, a reliable difference between their behavior and that of the Blinds-Down condition was obtained ($P = 0.035$).⁶

In addition to recording subjects' spatial solution to the problem, we also noted whether they asked for clarification during the test or mentioned it during a post-interview. We found that in the Blinds-Down condition (closest to Brown and Levinson's Dutch study), only in 20% of cases did subjects ask for clarification during training ("Should I put them still going the same way, or still going to my left?"). In the other, landmark-rich, conditions, the ambiguity of the task leapt out at the subjects not only in their response styles (the bimodal result shown in Fig. 7) but also in the 70% instance of spontaneous queries during the training phase. Though explicit querying seems to have been rare in the traditional populations, it did sometimes occur (Wilkins, e-mail communication, June 13, 2000; Levy, e-mail communication, November 27, 2000).

⁶ Additional variants of the outdoor condition were run. In the version reported in the text, we used very large towels on the grass (instead of tables) and larger animal toys (approximately 26 cm in size) to try to keep the experimental items and the surrounding landmarks similar in scale to that of the two indoor tabletop conditions. But then lest it had been animal size rather than changed landmarks that accounted for the result, we ran yet another ten subjects outdoors in front of the library with the small animals, i.e. they just had a closer-up view of the library than did the subjects in the Blinds-Up indoor condition. And again about half the subjects opted for the relative solution and half for the absolute solution to the rotation task. Collapsing across the 30 subjects in these two outdoor variants, their difference from the subjects in the Blinds-Up condition is again significant ($P = 0.032$).

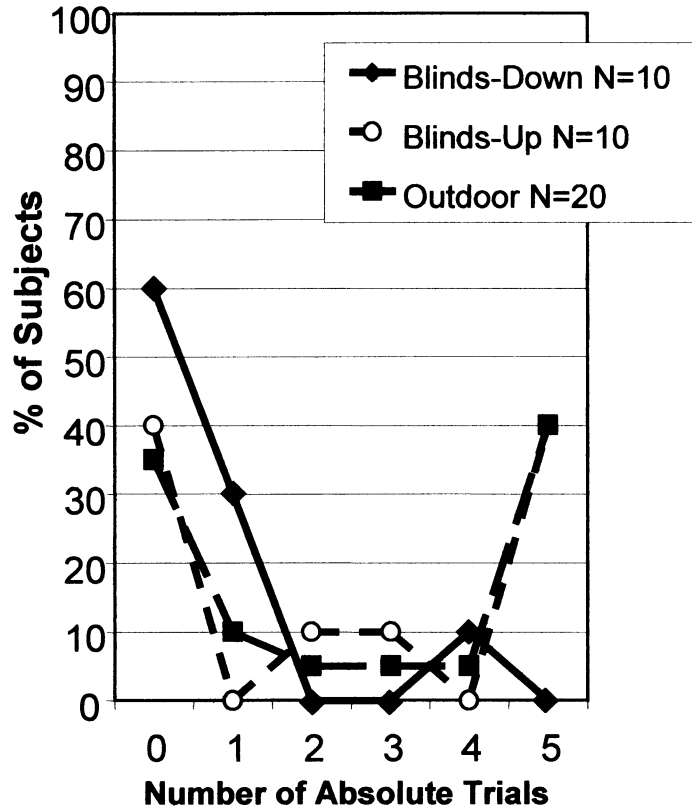


Fig. 7. Proportions of absolute choices in Blinds-Down, Blinds-Up, and Outdoor conditions. Subjects in the Blinds-Down condition predominantly chose the relative response. In the other conditions, subjects were divided in their preference for either all absolute responses or all relative responses.

2.2.1.3. Discussion In the relatively featureless Blinds-Down condition, we replicated for our urban English speakers the Brown and Levinson (1993) Dutch finding in the rotation task: speakers of a language community that favors “relative” egocentric terminology overwhelmingly chose the body-centered solution of the tabletop spatial task. This is so despite the fact that English (and Dutch) speakers do have words for cardinal directions, and use them for describing larger-scale spatial contexts, and belong to a culture that often uses maps. That the egocentric solution was usually the only one that crossed the subjects’ minds under these testing circumstances is further documented in their usual failure to comment on the ambiguity of the task. But in landmark-rich contexts (Blinds-Up and Outdoors), a larger orientational framework outside of the tables coexists with the framework provided in tabletop space. How is the subject to decide which frame of reference to employ and, therefore, which side of the second table corresponds to a given side of the first table? The toy animals can under rotation be placed *in the same direction relative to the spatial layout defined by the landmarks surrounding the tabletops* or *in the same direction relative to one’s own body position facing a tabletop*.

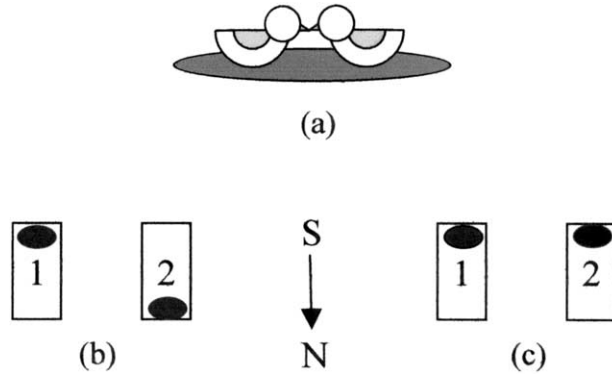


Fig. 8. Landmark cues. (a) The duck toy; (b) Relative duck placement, before and after rotation; (c) Absolute duck placement, before and after rotation.

Accordingly, about half the subjects in each manipulation now opted for the egocentric (“relative”) solution and half for the allocentric (“absolute”) solution, yielding the U-shape distribution of scores shown in Fig. 7. Consistent with the idea that the visible landmarks made subjects more aware of the availability of two possible spatial frameworks, about 70% of them in these latter conditions, as opposed to 20% in the Blinds-Down condition, asked for (but did not receive) clarification before they carried out the task.⁷

2.2.2. Experiment 2b: landmark cues in tabletop space (*Absolute/Relative Ducks*)

In the manipulations just described, we looked at the effects of ambient landmark cues in the visible world beyond the tabletops. And we found that the presence of such cues weakened the bias of our subjects toward egocentric responses, as compared with that bias for subjects who were tested in the featureless (Blinds-Down) condition. Specifically, their performance formed a bimodal distribution with some subjects now “Dutch-like” and others “Tenejapan-like” in their strategy. We here ask how malleable these group performances really are. Can landmark information if it is salient enough more completely determine the degree to which a single population solves spatial problems from an egocentric versus allocentric perspective? To find out, we now examined the effects of matched absolutely versus relatively placed landmarks.

2.2.2.1. Subjects, materials, and procedure

Forty new subjects participated in this experiment, 20 in each of two conditions. Each of these subjects was tested individually in the original laboratory room used for the indoor variants of Experiment 1. Now the blinds were always up. As shown in Fig. 8, a little toy stood on the Stimulus Table, to the right/south side of the subject him/herself. This toy was placed there before

⁷ It is of some interest to understanding these subjects’ state of mind that in a post-test interview, several of those who had not queried the ambiguity before or while doing the task now raised the question. Putting together all (pre-test, test, and post-test) queries about the interview, these totaled 40% in the two indoor conditions but 80% in the Outdoor condition. Again this suggests that salient landmark cues jog the subject to consider the solution that is disfavored in the no-landmark setting.

the subject entered the testing room, and it remained there, unmentioned and unmoved, for all five trials. As the figure shows, it was a pair of kissing styrofoam ducks on a paper lake (a longitudinally symmetrical toy, approximately 17 cm long). Because of its symmetrical construction, this toy had no inherent back or front and so was not amenable to intrinsic encoding of its own but was a potential landmark.

As usual, the subjects' task was to memorize the positions of the three line-up animals (as in Fig. 1) on the Stimulus Table. No line-up animal was closer than 15 cm to the kissing duck landmark, to avoid its being interpreted as part of the stimulus array. When subjects swiveled to the Recall Table, they always saw an exact replica of the kissing duck to one side of it, placed there in advance. For half the subjects (the egocentrically-biasing, or Relative Ducks group), the replica was always on the side of the Recall Table to the right of the subject. For the other half (the allocentrically-biasing, or Absolute Ducks group) the replica was always on the south side of the Recall Table. Now, as in the earlier experiments, the subject was handed the set of line-up-animals and asked to "make it the same as before".

2.2.2.2. *Results* The outcomes of this manipulation are graphically shown in Fig. 9. The

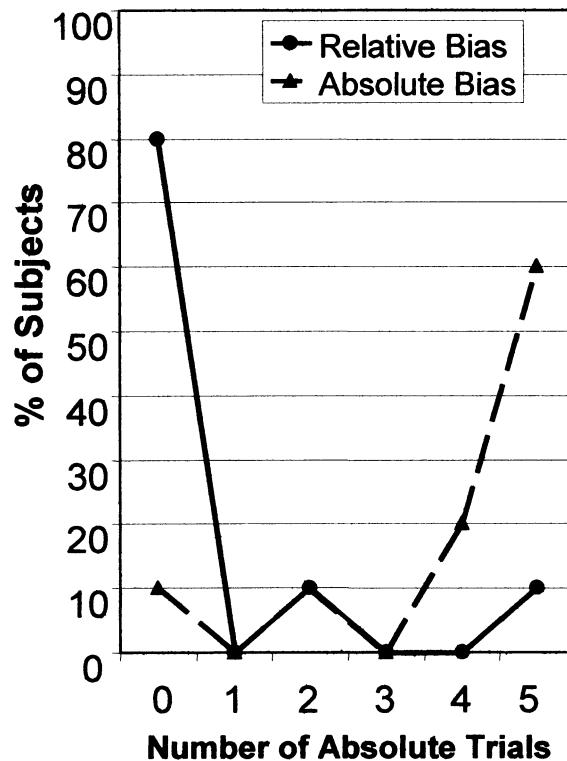


Fig. 9. Proportions of absolute choices in Experiment 2a (Relative versus Absolute Ducks). The subjects' preference for absolute or relative response is predicted by the positioning of the ducks.

results in the Relative Ducks condition were exactly like those obtained by Brown and Levinson (1993) for their Dutch subjects. And the results in the Absolute Ducks condition were exactly like those they obtained for the Tenejapans (compare Figs. 9 and 2). To evaluate these findings statistically, we again performed the Mann–Whitney *U*-test and found a highly reliable difference for these two conditions of the present experiment ($P = 0.003$).

2.2.2.3. Discussion The results of these variations on the original experiments suggest once more that the subjects' problem is to decide which side of the second table corresponds to a given side of the first table. In the prior experimental variants, it was shown that subjects used the landmark cues in the world beyond the tabletop to make the choice, where such landmark cues were made available (Blinds-Up and Outdoors conditions). In the present variants, the placement of the kissing duck trivially directs this choice within the frame of reference of the tabletop itself.⁸

3. Rotation tasks with nonlinguistic and prelinguistic populations

So far we have seen that the relative/absolute strategies for the rotation task can be reproduced within a single language community. This tends to vitiate the claim that specific language features – say, the spatial-terminological difference between Tzeltal and Dutch – are the underlying cause, or the sole underlying cause, of the original effects. We had so designed these experiments as to expose an alternative explanation of the original results. As we showed, the monolingual subjects solved this task differently depending on the presence and strength of the landmark cues made available to them. Thus, it may not be the nature, even the *linguistically learned nature*, of an individual to solve this kind of spatial problem in some particular way. Rather, the choice may be a function of the cues made consistently available in the environment. It is possible that such

⁸ For completeness, we ran another ten subjects in a single-table condition. After memorizing the stimulus array, these new subjects were walked around 180 degrees and seated at the opposite side of the same table (rather than being swiveled around 180 degrees to face a new table) and, after the usual 30 s delay, were asked to recreate the original array. Notice that the problem in this variant is formally the same one as before: to decide what is spatially “the same place” after a 180 degree rotation. And external landmark cues were unavailable as in the “Blinds-Down” condition. However, in fact there are crucial differences in the problem with which the subjects are now faced. They have no correspondence problem to solve. The subject's initial position at the table during testing polarizes the table, i.e. the subject herself is the landmark identifying which side of the table is which. Walking around the table, the subject by continuous updating can keep track of her position relative to that original polarized configuration. Thus, the instruction to recreate the original array is unambiguously interpretable as being within the original framework (the same tabletop), and for a subject who knows where she now is situated with respect to that original framework. (We particularly thank Randy Gallistel for discussion of these points.) Predictably, then, under this condition nine of the ten subjects behaved “absolutely”. One further lesson from this finding is that results in rotation tasks can vary in response to apparently small changes in procedure and so the results from any single procedure should be interpreted with due caution. Recent findings in the spatial literature are to the effect that subjects' response style in rotation tasks is extremely responsive to particular test conditions, e.g. whether a subject moves with respect to a stationary array versus whether the array moves relative to a stationary observer, whether the observer passively observes an orientation change or caused it himself, and whether he visually observes the reorientation as it happened or was blindfolded (Wang & Simons, 1999).

factors entered into the results reported by Brown and Levinson as well, as the landmark cues differed for their Dutch and Tenejapan subjects (though this cannot be known for sure without systematic retesting of those populations under appropriately controlled conditions). But even if so, it is certainly possible to suppose that, while landmark cues are variables that materially influence spatial reasoning, so are language variables such as the “habit” or “practice” of saying *west* rather than saying *left*. Here we pursue these issues further by reviewing studies of spatial reasoning by two populations with no language at all; namely, laboratory animals and prelinguistic human infants. Under rotation conditions closely analogous to those we have just reported, these populations too have been shown to vary in their spatial reasoning as a function of the strength and stability of landmark cues. This may point the way to a unified picture of when creatures solve rotation problems allocentrically versus egocentrically.

3.1. *Place learning and response learning in the T-maze*

During the 1940s and 1950s, experimental psychologists posed the question whether animals are naturally “response learners” or “place learners” in regard to navigating through space to reach a goal (e.g. Tolman, Ritchie, & Kalish, 1946). This issue, seen as a crucial one for understanding the fundamental cross-species nature of learning, was operationalized within experimental contexts that are formally analogous to the rotation task as designed by Brown and Levinson and by PDWLKS and as replicated in the experiments of the present paper.⁹ Typically, rats were trained to find food at one leg of a maze. Then either the rat or the training maze itself would be rotated 90 or 180 degrees. If the animal continued to turn in the trained direction on the maze even after rotation, he was judged to have learned only a motor *response* (say, “Turn left to get food”). If instead he was responsive to extra-maze cues and thus made a different turn on the rotated trials, he was assumed to have learned something more sensible; namely, to turn toward the *place* where the reward was previously found (say, “Turn north to get food”). Notice that the *place learning/response learning* distinction is in this regard operationally equivalent to the *allocentric/egocentric* (or in PDWLKS’ terminology, *absolute/relative*) distinction in the tabletop rotation task.¹⁰

What kind of learners did the rats turn out to be? In fact, neither the place nor response learning characterization of rat learning was conclusively shown to be the correct one for this task. As with the experiments we have just reported, the findings were best understood not by considering possible inherent tendencies in an animal to make the same turn or to go to the same place. Rather, the animals were responsive to the information provided to them (Blodgett & McCutchan, 1947). In a definitive review of this literature, Restle (1957, p. 226) wrote:

There is nothing in the nature of a rat which makes it a “place” learner or a

⁹ We are indebted to Henry Gleitman (personal communication) for pointing out to us the relevance of the animal literature in the present regards. Thanks also to Robert Rescorla for related discussion.

¹⁰ In these experiments, as in the human baby experiments reported in the next session, there is a right and a wrong answer, i.e. the food pellets are to be found either to the left or to the North, as the experimenter arbitrarily decrees. Subjects in the rotation studies were not rewarded for one kind of response or the other.

“response” learner. A rat in a maze will use all relevant cues, and the importance of any class of cues depends on the amount of relevant stimulation provided as well as the sensory capacities of the animal. In place-response experiments, the importance of place cues depends on the amount of differential extra-maze stimulation.

Thus, several thousands of rats gave their lives for very little gain, as the place-learner/response-learner distinction was apparently misdrawn to start with. When provided with sufficiently rich and stable landmark cues, any self-respecting rodent will use them and thus be classified as a place learner. In the absence of landmark cues, the disoriented rat is dependent solely upon his kinesthetic sense to update the egocentric information, and so repeats his prior response.

3.2. *Infant spatial reasoning*

Human infants’ spatial reasoning seems to be malleable in much the same way as just described for rats. Of particular interest is a series of demonstrations with 6–11-month-olds from Acredolo and her colleagues (Acredolo, 1978, 1979; Acredolo & Evans, 1980; see also Reiser, 1979; and Bloch & Morange, 1997 for a review). Acredolo’s experiments with infants, details aside, formally reproduce the animal maze-learning experiments just discussed. In a typical experiment, infants are seated in their mothers’ laps on a moveable chair facing a round table painted black on its right side and white on its left. They are trained by viewing something interesting which recurrently becomes visible, say, to their left: maybe some toy will rise from a well in the tabletop there, or maybe a smiling lady will appear there and talk to the child. The infants, chair and all, are next rotated 180 degrees. Now they are tested to see whether they expect (indexed by first reaching direction) the interesting thing to reappear in the same egocentric (“relative”) position, i.e. to their left, versus in the same allocentric (“absolute”) position (here called the *objective* strategy) where it had first appeared. In this paradigm only the objective strategy will actually get the baby its reward. The presence and strength of landmark cues is varied.

Overall, young babies exhibit an egocentric bias, though this materially weakens for this task between 6 and 11 months of age. And all these babies are able and willing to make use of landmark cues, though for the youngest these have to be made very salient and striking (e.g. colored stars and flashing lights). One manipulation (Acredolo & Evans, 1980) is of greatest relevance for interpretation of our own findings in the rotation tasks. Acredolo tested 9-month-olds in unfamiliar landmark-free environments (a small laboratory enclosure), unfamiliar environments containing potential landmarks (an office in all its everyday clutter), and a familiar environment with potential landmarks (the child’s home). The babies’ spatial reasoning styles significantly changed across these conditions: responses in both the laboratory and the landmark-rich unfamiliar environments were very predominantly egocentric (or *relative*, using PDWLKS terminology, or *response*, using the terminology of the rat laboratory). This tendency was dramatically reversed in the familiar home environment and became overwhelmingly objective (*absolute/place*).

4. General discussion

As we have just seen, nonlinguistic populations including rodents and human infants solve spatial rotation problems differently depending on the availability and salience of landmark cues. This was also true of our American subjects who were induced to adopt the allocentric strategy under closely analogous conditions: when landmark cues of sufficient potency salience were provided either on or beyond the tabletop – outdoor landscapes and kissing ducks were our adult equivalent of Acredolo’s flashing lights.

Such results in no way threaten the cross-linguistic findings from PDWLKS nor diminish their interest, but they do bear on how these facts should be interpreted. There are two main issues here in want of an explanation. The first is why in the neutral or landmark-poor condition members of different language groups tend to interpret the tabletop task (“make it the same”) differently. The second is why different communities often show a general bias – Whorf’s “fashions of speaking” – toward allocentric or egocentric spatial usage. We now take up these two topics in turn.

4.1. Interpretation under ambiguity

Several recent studies show that speakers (including young language learners) can perform classification tasks similarly despite striking cross-linguistic distinctions in relevant terminology. For instance, novel displays are sorted similarly into *object* versus *substance* categories whether (as in English) or not (as in Japanese) the language spoken distinguishes the two classes formally, such as in the English distinction between count and mass noun (e.g. *a dog* versus *water*; *some dogs* versus *some water*; Imai & Gentner, 1997; Soja, Spelke, & Carey, 1991). However, in the same tasks, language-specific effects on classification are often observed when two conditions are simultaneously met: first, when the stimulus array falls somewhere near the midline on the classificatory dimension; and second, when the task instruction is linguistic or involves some linguistic label (even a nonsense word). Thus, when shown a complex articulated object or else a puddle of handcream-like mush (each referred to as “this blicket” or “my blicket”), Japanese and American children and adults agree well in assigning these to the object or substance categories, respectively, as shown in a subsequent sorting task.¹¹ But for very simply-shaped stimuli (e.g. a kidney-bean shaped piece of colored wax), English speakers are more likely to assign them to the object class than are Japanese speakers. Why should this be? After all, “this blicket” is a locution that applies equally well to mass and count nouns (*this dog*, *this water*). But there is a relevant difference. The distributional facts about English are such that count nouns are many times more numerous than mass nouns, and so a new word heard – all other things held equal – is more likely count than mass. Because the mass–count linguistic description maps roughly onto the conceptual distinction between thing and stuff, if English-speaking listeners use distributional patterning to assign *blicket* to the count–noun class, they will also be more likely to guess that the

¹¹ Specifically, subjects are shown new items and asked (again using locutions that are linguistically neutral in the relevant regard) which are also blickets. Choosing more items shaped the same as the standard (but of different substances) is taken as evidence that the original was taken to be some heretofore unknown *object* type; choosing more items made of the same stuff as the standard (but of different shape) indexes a *substance* interpretation.

phrase “this blicket” describes a kidney-shaped object than that it describes colored wax. More generally, when subjects are offered ambiguous linguistic descriptions for an ambiguous stimulus (e.g. something whose construal as object or substance is indeterminate or close to it), they often fall back on implicitly known form-to-meaning patterns in their language as a clue to the communicative intent of the speaker (for related discussion and findings, see Landau & Gleitman, 1985; Naigles & Terrazas, 1998). Such results are consistent with many demonstrations in the psycholinguistic laboratory that individuals are highly sensitive to lexical patterning, and implicitly recruit such probabilistic knowledge to resolve linguistic (syntactic) ambiguities as well (e.g. Snedeker, Thorpe, & Trueswell, 2001; Trueswell & Kim, 1998).

The effects uncovered by the PDWLKS group are subject to a similar interpretation, as several commentators have pointed out (Fisher & Gleitman, in press; Levelt, 1996; Newcombe & Huttenlocher, 2000). Subjects in the tabletop task are told only to “make it the same” after the rotation, a blatantly ambiguous command (though the ambiguity is of course mitigated when properties associated with the stimulus array such as the landmark cues of Experiment 2b are informative or suggestive). Many subjects in laboratory-savvy populations like our American undergraduate one will now protest and ask for clarification; much more rarely (and usually after rather than during the test) participants from the traditional populations may also voice their recognition of the ambiguity. But in this task the experimenter does not then disambiguate her utterance (“just do it however you think...”) so the subjects are left to guess her intent. The pragmatics of language use in English can now serve as an implicit cue: likely if one American is speaking to another about object placement in this small-scale apparatus, it will be in the style of making spatial reference common in the community – left–right, not east–west. For the Tenejapans, of course the opposite likelihood obtains.

Such *effects of language on language* are widespread and well known but they do not imply effects of language on thought in any recognizable sense of linguistic determinism. The subjects have not been forced or guided into a mode of spatial reasoning about space by features of their language. Rather, in a situation that makes either of two available reasoning strategies reasonable, the subject makes a pragmatically sensible guess as to which one the speaker probably had in mind and is talking about.

4.2. *Culturally shared fashions of speaking*

Granting now that we shifted the frame of reference for spatial reasoning by our English-speaking subjects by varying the strength and reliability of landmark cues, how could this bear on the general tendency of human populations to favor particular spatial terminologies? For after all, we must explain more than why Americans or Tenejapans select a particular strategy of reasoning on the fly in the ambiguous tabletop experiments. The phenomenon that the Levinson–Pederson group have tellingly and informatively exposed is that communities of speakers choose among potential linguistic resources and *regularly* (or “*habitually*”) prefer to say either “the spoon to the north of your teacup” or “the spoon to the left of your teacup” even when their language contains both kinds of lexical item. Speakers in different communities have, as Whorf termed it, a preferred “*fashion of speaking*” of spatial relations. This work has been justly acclaimed for the

enhanced perspective it is providing on language–culture relations. At a minimum it stands as a welcome antidote to much familiar linguistic and psychological inquiry that assumes by default that all peoples are about the same as, say, a community of Ivy League sophomores in a Philadelphia dormitory. We now make a preliminary attempt to understand how these different cultural choices come into existence.

What factor or factors could lead to the long-term preference for geocentric terms and – perhaps – an associated tendency to approach spatial problems allocentrically? Leaving aside the very special T-maze/tabletop rotation problems, some recent evidence suggests that creatures from ants to humans navigate via momentary egocentric representations that are updated as they move (Wang & Spelke, 2000). If that is so, it seems odd to find human communities that rely heavily on allocentric representations, at least linguistically.

We will begin by thinking about global distinctions among the cultural communities whose terminology has been studied in the cross-language project of the PDWLKS group. Thereafter, we will suggest, speculatively to be sure, that a single factor – geographical cohesion in community life – plays a major role in predicting why social groups develop preferences in their everyday terminology for referring to regions and directions in space. The findings of the PDWLKS group suggest several kinds of communality among the six linguistic populations that they classified via results of the man-and-tree test and subsequently tested in the rotation task. Three of these languages (Tzeltal, Longgu, and Arrandic) are spoken in traditional, largely unschooled, linguistic-cultural communities and the results for all three followed the allocentric pattern. Two other languages (Dutch and Japanese), both spoken by technologically advanced and schooled populations, are those that yielded the egocentric pattern. And we have now added English, another schooled population that yields the egocentric pattern. The sixth language that these investigators studied, Tamil, is an important mixed case that we will discuss below.

Thus, at a minimum three potentially explanatory factors covaried in the communities that showed the absolute bias in solving the rotation problem: they are (1) less schooled, (2) more insular (isolated from other groups) and geographically cohesive than the egocentrically-biased communities, and (3) their elicited spatial terminology in the man-and-tree test is “absolute”. The Whorfian bet is that this third factor, linguistic divergence, holds the explanatory key to reasoning strategies (“...language is not merely a reproducing instrument for voicing ideas but rather is itself the shaper of ideas, the program and guide for the individual’s mental activity”; Whorf, 1940/1956, p. 212). We suggest that the within language and “no language” studies in Sections 2 and 3 of this paper bring this language-based explanation into question for the case of spatial reasoning. Now we consider the other two covarying factors.

4.2.1. Schooling and literacy

There has been extensive discussion of the fact that unschooled and schooled populations often behave quite differently in experimental tasks. John Lucy speaking from within the anthropological linguistic community provides a convincing review, concluding that literacy and school performance have considerable effects on both “patterns of thought” and “language socialization practices for the inculcation of cultural world-view” (Lucy, 1996, p. 57). Discussions of the relations of language categories and speech styles to school practice go back in their modern form at least to Vygotsky (1978) and are the

subject matter of a vigorous and continuing linguistic anthropological tradition (see also Bernstein, 1971; Greenfield, 1972; Schieffelin & Ochs, 1986; Scribner & Cole, 1973). A problem in interpreting the cross-linguistic results on spatial reasoning, then, is that in all the cases on which PDWLKS reported, they have been confounded with important cultural variables such as schooling.

Pederson (1995) (see also PDWLKS) made an interesting attempt to unconfound the linguistic and cultural variables that covaried in the studies that we have discussed thus far by testing individuals from two communities, both of which spoke the same language, Tamil. Tamil has the formal resources for both the absolute and relative spatial reference terms. All the same, the two Tamil-speaking communities varied in their habitual use, one community preferring the absolute and the other preferring the relative terms for describing small-scale regions and directions (as measured, among other ways, by the Man and Tree test). Rotation experiments on these two populations with the same language but different typical use patterns might exclude the alternative culture-difference (as opposed to language-difference) explanation that we have just suggested. However, the results of this manipulation turned out to be quite unclear. The animals-in-a-row experiment yielded no reliable difference between the two subpopulations of Tamil speakers. But a substitute experiment concerning spatial strategies, informally reported in PDWLKS, yielded a reliable difference in the predicted direction: the Tamil population that habitually uses relative terms solved the spatial task relatively, and vice versa. However, this new experiment in practice suffered from exactly the same confound it was designed to remove. The more urbanized Tamil population (residents of the city of Madurai) was the one exhibiting the relative bias both in speech and in spatial reasoning (i.e. they behaved more like the Dutch and Japanese). The more rural Tamil-speaking subpopulation showed the absolute bias that characterized the rural, insular, geographically cohesive Tenejapan Tzeltal population.

Summarizing, the results of the animal-in-a-row studies, including our own, seem to cohere as arising in the presence of supralinguistic cultural differences. These differences predict both the favored linguistic terminology and the spatial reasoning patterns, a distinctly non-Whorfian generalization. But why should traditional, or insular, or less schooled groups prefer geocentric spatial terminologies and schooled urbanized cultures prefer body centered ones? “Culture” is too vague and undifferentiated an explanation of the spatial reasoning phenomena. As we next discuss, all these findings are suggestive of a more specific solution: shared landmark cues vary in their stability and familiarity across communities, much as they do in the various conditions of our experiments, and under the controlled environmental conditions in which nonhuman and preliterate humans have been tested for spatial reasoning strategies under rotation.

4.2.2. Geography, cohesion, and the egocentric–allocentric distinction

Recall that the allocentrically-responding Tenejapans reside in a small village located on the side of a hill, and employ a three-term spatial terminology that alludes to the spatial coordinates of this hill, “downhill” (e.g. “tree standing downhill of man”), “uphill”, and “across-the-hill” or, as it were, “athwart”. It is interesting that spontaneous descriptions from speakers of other spatial-absolute languages studied by this research group again make reference to what must be highly salient local landmarks, e.g. “man stands in ‘land of

soft sand” (Hai//om), “tree standing on side towards sea” (Longgu; these quotes from native speech appear in Pederson et al., 1998, p. 568). As Pederson (1995) points out, it would not be surprising at all if persons living in a small and geographically coherent layout would rely heavily on mutually known local landmarks to orient themselves in space.

Pederson (1995, p. 55) attempted to counter this landmark explanation of absolute solutions in these spatial tasks by testing his villagers away from their homes and immediate work places: “I always tried to select an environment for testing, such as the front of the headman’s storehouse, which would be reasonably removed from the subjects’ daily life.”). But as we have demonstrated, mildly unfamiliar landmarks (such as the Outdoor condition of the present experiments, see Fig. 6) are useable cues and increase allocentric responses. Moreover, previously unknown landmarks, even such adventitious ones as kissing styrofoam ducks (Fig. 8), markedly affect spatial reasoning in adults (though apparently not in Acredolo’s youngest rotated infants). Regardless of the language spoken, it seems, people will use landmark cues when these are available. It does not even have to be people: in this task, rats too will take advantage of landmarks so as to locate food pellets if only these cues are made rich and reliable enough.

Summarizing, we can best understand the findings of the cross-linguistic studies (and the within-language Tamil study) by noting a cross-cutting factor: the geographical and interpersonal cohesion of a society. There seems to be no consensual “uphill” that can serve as a reference point in the very large and shifting communities in which linguistically interacting English, Dutch, or Japanese speakers generally find themselves. “You’ll find the railroad station just northeast of the Drexel University parking lot” is not too useful a direction to give the British tourist who has just arrived in Philadelphia. Body-orientation is the obvious alternative (or auxiliary) in establishing momentary spatial coordinates (“Go down to the corner, turn left and walk three blocks”). In contrast, as PDWLKS appear to show, people who live in a small, mutually familiar, geographical area typically use its local landmarks to devise a spatial coordinate system that makes reference to its stable features (“uphill”, “inland”, etc.). This is so even when the traditional populations have the formal linguistic resources for encoding both absolute and relative spatial terminology (for an important discussion of Longgu usage from this perspective, see Hill, 1995).

Of course the present authors do not know too much about traditional unschooled cultural groups who live in faraway places. Large disparities between investigator and investigated make it difficult to interpret either naming practices or experimental responses across these cultural divides. Indeed, PDWLKS rightly caution us not to add new languages to the sample without being well-acculturated anthropologists on the site. Luckily one does not have to go all the way to Chiapas or Papua-New Guinea to find communities that favor landmark-based spatial terminology: one of us is a native of a highly urbanized culture whose members live and work all crammed together on a skinny little island, about 16 miles long, at the mouth of the Hudson River; namely, Manhattan Island. Culturally diverse (some would even say “literate”) as this community is, its residents share a small, stable, geographical landscape, rich in mutually known landmarks. Likely this is why their terminology for locations in the community is absolute and – like

that of the Tenejapans – typically makes do with only three terms in habitual use: *uptown*, *downtown*, *crossdown*.¹²

In sum, the causal engine both for the engrained spatial reasoning styles and the fashions of speech that we find in different communities may well be a derivative of their ambient spatial circumstances. Whatever these circumstances are, communities of humans will develop terminology to fit. So there is bound to be a strong bond between language and thought, one that is hard to break or even weaken (except, perhaps, in the artificial neutrality of the toys-on-tabletops situation). Commentators from Whorf on forward have properly been struck by apparent connections between the cognitive style and contents of a culture and the linguistic devices in habitual use in that culture. As we have asserted throughout, there is no factual doubt about these correlations. The unresolved issue is the direction of causation. Do the Eskimos think so well about snow because they have a lot of snow words, as Whorf improbably maintained? That would certainly be a lucky coincidence for the Eskimos! But do we actually “dissect nature along lines laid down by our native languages” (Whorf, 1956, p. 213)? Or, conversely, do Eskimos have a lot of snow words because they think so well about snow? Intuitions certainly differ in this case as to which is cause and which is effect. But what we have argued here is that one will never find out by studying Eskimos in the snow.

For the spatial reasoning issue, we tried to break this language–circumstance confound both in the experiments presented and in several other experiments cited and discussed. The language-caused (“linguistic determinism”) solution was, in our view, vitiated by showing that speakers of a single language – or no language at all, such as rats and human infants – can be induced to vary in their spatial reasoning strategies by changing the circumstances of test; specifically, by supplying or withholding landmark information. These findings lead us to reject the explanation offered by PDWLKS for their findings. They assert: “The linguistic system is far more than an *available* pattern for creating internal representations; to learn to speak a language successfully *requires* speakers to develop an appropriate mental representation which is then available for nonlinguistic purposes.” (Pederson et al., 1998, p. 586). Elegant as this statement is, we do not find it defensible on the evidence. Rather, it seems to us, one would want to turn it on its head: linguistic systems are *merely* the formal and expressive medium that speakers devise to describe their mental representations and manipulations of their reference world. Depending on the local circumstances in which human beings find themselves, they select accordingly from this linguistically available pool of resources for describing regions and directions in space.

¹² One could object here that the Manhattan population is really wildly transient and therefore does not “really” have stable landmarks that could be used by all its residents and massive numbers of visitors. That is true. But the trick is not caring. We have the following story from a Swedish tourist entering Manhattan via the George Washington Bridge (which hits the island’s west flank at approximately 175th street): “We saw the signs, one labeled ‘Uptown’, the other ‘Downtown’. We knew we were expected in ‘Midtown’ but this did no good at all and we were lost in the Bronx for two hours.” The moral here is that natives may rely on cues that are unusable by visitors to their island home.

5. Final thoughts

The Whorf–Sapir debate has been joined, since its inception 50 or so years ago, as a series of local skirmishes: does language influence thought about hue? Numerosity? Causality? Objecthood? Temporality? Counterfactual reasoning? Or, as in the present studies, spatial organization? No such particularized investigation can really be decisive as to whether interesting effects of language on thought can be found in other, so far unstudied, domains, or under other, so far unarticulated, versions of the theory. All the same, limited findings do have a way of contingently influencing our broader beliefs. For example, PDWLKS suggested that their obtained rotation results for language and space provided “reason for optimism” that related effects of language on thought would be found in other domains. Turning the tables, we take the present finding of the robustness of human thought to variation in linguistic usage patterns to be an optimistic one. All languages have the formal and expressive power to communicate the ideas, beliefs, and desires of their users. From this vast range of possibilities, human communities select what they want to say and how they want to say it. This stance is at its core the same one that explains why the Elizabethans habitually used terms for falconry and we do not, and why English-speaking vacationers at Aspen and Vail find it natural to develop terms like *sugar*, *powder*, and *granule* to amplify their heretofore impoverished means for discussing the state of the snow on the slopes. In the end, it's the thought that counts.

Acknowledgements

We thank Henry Gleitman, Robin Clark, Randy Gallistel, Pim Levelt, Liz Spelke, the members of the CHEESE seminar, and two anonymous reviewers for important commentary and suggestions. We thank Kathleen Hoffman, Leslie Koo, Aharon Charnov, Lorraine Daley, Jill Kleczko, and Aurit Lazerus for helping in data collection. We especially thank Kathleen Hoffman for getting us started on the project. The research was supported by a predoctoral Fellowship to Peggy Li from the Institute for Research in Cognitive Science, University of Pennsylvania, and by NSF Grant # SBR8920230 and NIH Grant # 1R01HD375071.

References

- Acredolo, L. P. (1978). Development of spatial orientation in infancy. *Developmental Psychology*, *14* (3), 224–234.
- Acredolo, L. P. (1979). Laboratory versus home: the effect of environment on the 9-month-old infant's choice of spatial reference system. *Developmental Psychology*, *15* (6), 666–667.
- Acredolo, L. P., & Evans, D. (1980). Developmental changes in the effects of landmarks on infant spatial behavior. *Developmental Psychology*, *16* (4), 312–318.
- Au, T. K. (1983). Chinese and English counterfactuals: the Sapir-Whorf hypothesis revisited. *Cognition*, *15*, 155–187.
- Berlin, B., & Kay, P. (1969). *Basic color terms*. Berkeley, CA: University of California Press.
- Bernstein, B. (1971). *Class, codes, and control, Theoretical studies toward a sociology of language*, Vol. 1. London: Routledge & Kegan Paul.
- Bloch, H., & Morange, F. (1997). Organising gestures in external space: orienting and reaching. In N. Foreman &

- R. Gillett (Eds.), *Spatial cognition in the child and adult. A handbook of spatial research paradigms and methodologies* (pp. 15–39), Vol. 1. Hove: Psychology Press/Erlbaum.
- Blodgett, H. C., & McCutchan, K. (1947). Place vs. response learning in the simple T-maze. *Journal of Experimental Psychology*, 37, 412–422.
- Bloom, A. H. (1981). *The linguistic shaping of thought: a study in the impact of language in thinking in China and the West*. Hillsdale, NJ: Erlbaum Associates.
- Bloom, P., Peterson, M. A., Garrett, M. & Nadel, L. (1996). Language and space Cambridge, MA: MIT Press.
- Bowerman, M. (1996). Learning how to structure space for language: a crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space. Language, speech, and communication* (pp. 385–436). Cambridge, MA: MIT Press.
- Bowerman, M. (1996). The origins of children's spatial semantic categories: cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity. Studies in the social and cultural foundations of language, No. 17* (pp. 145–176). Cambridge: Cambridge University Press.
- Bowerman, M., & Choi, S. (2001). Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge: Cambridge University Press.
- Brown, P., & Levinson, S. (1992). 'Left' and 'Right' in Tenejapa: investigating a linguistic conceptual gap. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 45 (6), 590–611.
- Brown, P., & Levinson, S. C. (1993). *Linguistic and nonlinguistic coding of spatial arrays: explorations in Mayan cognition* (Working Paper No. 24). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics.
- 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). Mahwah, NJ: Erlbaum.
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language specific lexicalization patterns. *Cognition*, 41, 83–121.
- Choi, S., McDonough, L., Bowerman, M., & Mandler, J. (1999). Early sensitivity to language-specific spatial terms in English and Korean. *Cognitive Development*, 14, 241–268.
- Davies, I. R. L. (1998). A study of colour grouping in three languages: a test of linguistic relativity hypothesis. *British Journal of Psychology*, 89 (3), 433–452.
- Fisher, C., & Gleitman, L. R. (in press). Breaking the linguistic code: current issues in early language learning. In: C. R. Gallistel (Ed.), *Stevens' handbook of experimental psychology. Learning, motivation, and emotion* (3rd ed., Vol. 3). New York: Wiley.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gallistel, C. R. (1992). *Animal cognition* Cambridge, MA: MIT Press.
- Greenfield, P. M. (1972). Cross-cultural studies of mother-infant interaction: towards a structural-functional approach. *Human Development*, 15 (2), 131–138.
- Gumperz, J. J., & Levinson, S. C. (1996). Introduction to Part I. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity. Studies in the social and cultural foundations of language, No. 17* (pp. 21–36). Cambridge: Cambridge University Press.
- Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, 93 (1), 10–20.
- Heider, E. R., & Oliver, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, 3 (2), 337–354.
- Hill, D. (1995). *Finding your way in Longgu: geographical reference in a Solomon Islands language* (Working Paper No. 21). Nijmegen: Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics.
- Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, 62, 169–200.
- Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space. Language, speech, and communication* (pp. 1–30). Cambridge, MA: MIT Press.
- Jameson, D., & Hurvich, L. M. (1978). Dichromatic color language: "reds" and "greens" don't look alike but their colors do. *Sensory Processes*, 2 (2), 146–155.
- Kemmerer, D. (1999). "Near" and "far" in language and perception. *Cognition*, 73 (1), 35–63.

- Landau, B., & Gleitman, L. R. (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 and Brain Sciences*, 16 (2), 217–265.
- Levelt, J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space. Language, speech, and communication* (pp. 77–107). Cambridge, MA: MIT Press.
- Levinson, S. C. (1996). Frames of reference and Molyneux’s question: cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space. Language, speech, and communication* (pp. 385–436). Cambridge, MA: MIT Press.
- Levinson, S. C. (2000). Yéll dnye and the theory of basic color terms. *Journal of Linguistic Anthropology*, 10 (1), 1–53.
- Lucy, J. A. (1996). The scope of linguistic relativity: an analysis and review of empirical research. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity. Studies in the social and cultural foundations of language, No. 17* (pp. 37–69). Cambridge: Cambridge University Press.
- Martin, L. (1986). Eskimo words for snow: a case study in the genesis and decay of an anthropological example. *American Anthropologist*, 88 (2), 418–423.
- Naigles, L., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: influences of language and syntax. *Psychological Science*, 9 (5), 363–369.
- Newcombe, N., & Huttenlocher, J. (2000). *Making space: the development of spatial representation and reasoning*. Cambridge, MA: MIT Press.
- Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, ‘n’ roll: the representation of motion in language and thought. *Cognition* (in press).
- 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., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language*, 74 (3), 557–589.
- Pullum, G. K. (1991). *The great Eskimo vocabulary hoax and other irreverent essays on the study of language*. Chicago, IL: University of Chicago Press.
- Pütz, M. & Dirven, R. (1996). *The construal of space in language and thought*. New York: Mouton de Gruyter.
- Reiser, J. J. (1979). Spatial orientation of six-months-old infants. *Child Development*, 50, 1078–1087.
- Restle, F. (1957). Discrimination of cues in mazes: a resolution of the place-vs.-response question. *Psychological Review*, 64, 217–228.
- Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, 129 (3), 369–398.
- Schieffelin, B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163–246.
- Scribner, S., & Cole, M. (1973). Cognitive consequences of formal and informal education. *Science*, 182, 553–559.
- Slobin, D. (1991). Learning to think for speaking: native language, cognition, and rhetorical style. *Pragmatics*, 1, 7–25.
- Snedeker, J., Thorpe, K., & Trueswell, J. (2001). On choosing the parse with the scene: the role of visual context and verb bias in ambiguity resolution. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children’s induction of word meaning: object terms and substance terms. *Cognition*, 38 (2), 179–211.
- Talmy, L. (1978). Figure and ground in complex sentences. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (Eds.), *Universals of human language* (pp. 625–649). Stanford, CA: Stanford University Press.
- Talmy, L. (1983). How language structures space. In H. Pick & L. Acredolo (Eds.), *Spatial orientation: theory, research, and application* (pp. 225–282). New York: Plenum Press.
- Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning: II. Place learning versus response learning. *Journal of Experimental Psychology*, 35, 221–229.
- Trueswell, J. T., & Kim, A. (1998). How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. *Journal of Memory and Language*, 39, 102–123.

- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, 70, 191–210.
- Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition*, 77 (3), 215–250.
- Whorf, B. L. (1940/1956). Science and linguistics. In J. B. Carroll (Ed.), *Language, thought and reality: selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press. (Reprinted from *Technology Review*, 42(6), 229–231, 247–248).
- Whorf, B. L. (1941/1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought and reality: selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press. (Reprinted from L. Spier (Ed.), *Language, culture, and personality: essays in memory of Edward Sapir*. Menasha, WI: Sapir Memorial Publication Fund).
- Whorf, B. L. (1941/1972). The relation of habitual thought and behavior to language. In P. Adams (Ed.), *Language in thinking: selected readings*. Middlesex: Penguin Books. (Reprinted from L. Spier (Ed.), *Language, culture, and personality: essays in memory of Edward Sapir*. Menasha, WI: Sapir Memorial Publication Fund).