Language, culture and cognition

Editor
STEPHEN C. LEVINSON
Max Planck Institute for Psycholinguistics, Nijmegen

This new series looks at the role of language in human cognition —
language in both its universal, psychological aspects and its variable,
cultural aspects. Studies will focus on the relation between semantic and
conceptual categories and processes, especially as these are illuminated
by cross-linguistic and cross-cultural studies, the study of language
acquisition and conceptual development, and the study of the relation
of speech production and comprehension to other kinds of behavior in
cultural context. Books come principally, though not exclusively, from
research associated with the Max Planck Institute for Psycholinguistics
in Nijmegen.

1. Jan Nuyts and Eric Pedersen (eds.) Language and conceptualization
2. David McNeill (ed.) Language and gesture
3. Melissa Bowerman and Stephen C. Levinson (eds.) Language
   acquisition and conceptual development
Covariation between spatial language and cognition, and its implications for language learning

Stephen C. Levinson
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

1 Word learning: the scope of the problem

Much of this volume is concerned with the problem of how children learn the meanings of words, or more exactly morphemes of various kinds. The answers may depend on many factors of course: different kinds of morphemes may be, indeed must be, learnt in different ways, the ways themselves may be opened up by conceptual development within the child over time, and so on. However, we can (not without some danger) abstract away from these divergent factors and ask general questions about the scope of the “induction” or “mapping” problem.

Now in the chapters that immediately precede, those by Bowerman & Choi, de León, and Brown, an issue is raised that has significant bearing on the dimensions of the “mapping problem.” What these chapters show is that, in a number of spatial domains, the kinds of categories that need to be associated with the meanings of words can vary rather drastically across languages, and moreover that, at least by the beginning of systematic speech, there is no evidence of a uniform initial state of the learning machine, i.e. little evidence that children are presuming certain kinds of natural categories, later discarding them in favor of the local idiiosyncrasies. Nor does this picture change when one starts to plumb comprehension before the age of complex utterance production. In short, the semantic categories look almost as variable as the phonological strings onto which they must be mapped.

In this chapter, my central purpose is to lay out some additional facts about adult mental life that seem to compund the problems radically: they seem to raise the stakes against the child’s possibilities of success still higher. The essential finding that I will lay out is that not only, on a crosscultural basis, do we find sound-systems changing, and meaning-systems in radical diversity, but also we find that the adult cognitive operations that underly or support those meanings seem to covary with the linguistic system. Why this should radically alter the picture needs some explanation.

First, I need a pair of linguistic examples that may demonstrate different kinds of semantic variation across languages. Suppose we translate the English description (a) below into Tzeltal (b), the Mayan language further described in Brown’s paper (ch. 17 of this volume).

(a) ENGLISH
Put the bowl behind the box

(b) TZEILTLAL
pach-an-a bojch ta y-anil te karton-e
bowl+put-CAUSE-IMP gourd+bowl at its-down the cardboard-DEIC

The first word of the Tzeltal example is the imperative of the causative of the root pach, which means (roughly) ‘place a bowl-shaped vessel upright on a surface.’ This is one kind of semantic variation we may find across languages: familiar distinctions turning up in unfamiliar places, here bowl-shaped (or hemispherical) being a verbal rather than a nominal component. Similarly, the word bojch denotes not a bowl in general, but one made from a half of a spherical gourd (compare English bowl, which is indifferent to the material from which it is made). These kinds of crosslinguistic contrast suggest that what we may find when we look across languages are repackagings or distinct lexicalizations of the same or at least similar semantic parameters. Now consider karton, a Spanish loan, which in Tzeltal is semantically general across ‘cardboard’ rather than ‘box’; this is a Mayan pattern whereby the basic nominal (other than names for humans, beasts and artifacts) denotes a substance rather than a thing (Lucy 1992). Similarly lo’bal ‘banana stuff’ denotes the tree, the leaves, the roots as well as the fruit (for the cognitive consequences of this Mayan pattern, see Lucy & Gaskins, ch. 9 of this volume). This is arguably somewhat different from a mere shuffling of semantic features, since we have no colloquial notion of ‘banana essence’ or “stuff manufactured by banana genotype”; but in any case it shows that nouns are not necessarily basically words for things – which might seem to be an essential assumption a child must make at the initial stages of language-learning (see Gentner & Boroditsky, ch. 8 of this volume). Finally, consider how English behind is rendered into Tzeltal ta yanil: although these terms may on an occasion of use have descriptive equivalence, these are actually not in the same intensional ballpark at all. English behind is ambiguous (between what I will call an intrinsic and relative – or “deictic” – coordinate system), but is here used in the “deictic” or relative way, so that the utterance means that the box is between the speaker and the bowl. Tzeltal has no term equivalent to either of these meanings of English behind. Tzeltal ta yanil ‘at its down(hill)’ (also ambiguous) here means something quite different, which may be specified in our concepts (not theirs) as lying in the quadrant bisected by the line North 010° – it is a cardinal direction term,
which they think of in terms of a world essentially tilted North(-north-east)wards. These cardinal direction parameters (uphill, downhill, across) show up in a systematic range of vocabulary from motion verbs to words for edges. Clearly, we can't find any such semantic parameters built on the assumption of a tilted world showing up in English. This is a different kind of semantic variation across languages, and this chapter is about the cognitive consequences of such semantic variation.

Now we are prepared to think about these different kinds of semantic variation more abstractly (indeed, I shall here propose a way of thinking that is far too abstract, returning later to correct the picture). A precondition to mapping words (or other morphemes) onto meanings is of course to isolate the phonological words (or other morphemes). Current research shows that babies are beginning to learn the fundamental algorithms here a year or more before they speak (Cutler 1995). Language-specific factors are involved not only at the level of phoneme inventories, stress, and syllable identification, but also at the level of finding roots obscured by derivational and inflectional morphemes, secondary phonological processes, etc. Meaning may also be involved, but in the limited sense of finding units that “have meaning,” a precondition to asking what meaning they have (Lyons 1968:412). That second question is the mapping problem.

We can imagine three distinct levels, or degrees, of ascending complexity in the mapping problem.

1.1 Degree 1.0 Mapping Problem: mapping known phonological entities onto known semantic or conceptual entities (or mapping language-specific phonological units onto language-independent semantic units)

Once the word-forms^2 (or major meaning-bearing morphemes) have been isolated, the easiest kind of mapping would be from words to preexisting conceptual bundles, i.e. concepts identified independently of language, on nonlinguistic grounds. Fodor (1975) is an influential proponent of such a view.^3

As Quine (1960) pointed out in his celebrated conundrum of the indeterminacy of radical translation, even here, at the simplest level of form–meaning mapping, the mapping problem becomes fundamentally problematic: rabbit accompanied by glimpse of scurrying rabbit could mean ‘scurry,’ ‘white fluff,’ or whatever. The problem seems quite insoluble without powerful general heuristics of some kind. But the more specific the heuristics the more they threaten to compound the problem, for they will then be inconsistent with one another. Consider for example the oft-proposed heuristic designed for learning nominals: let there be a bias towards object-naming, so that the child may assume that 3D constancies of size and shape map onto nominals (see Gentner & Boroditsky's chapter, this volume; see also Bloom (ch. 6), where count nouns are presumed to map to individuals). Such a heuristic will then hinder rather than help with the learning of abstract or relational nominals (like birthday, night, or brother), and interfere with the learning of predicates (again see Gentner & Boroditsky, this volume, and Gleitman 1990 for parallel suggestions about Degree 1 verb heuristics). Clearly a prior form–class identification will help; and one might hope for language-independent patterns, e.g. at least a distinction between function-morphemes and content-morphemes, but even here things look language-specific (see Slobin, ch. 14 of this volume). These problems are well rehearsed in other chapters in this volume (see, e.g., ch. 7 by Carey and ch. 2 by Gopnik).

We can think of the Degree 1 mapping problem as constructing the set of ordered pairs for each correct association between a phonological word and a semantic concept, where the latter are preexisting conceptual bundles. Then we can get a measure of the size of the “design space” (Dennett 1995) or problem space for the child under Degree 1 assumptions: she must select the correct pairings from all the possible pairings, that is, from the Cartesian product of the set of lexemes/morphemes A and the set of meanings/concepts B.

1.2 Degree 2.0 Mapping Problem: mapping known phonological entities onto unknown semantic entities (or mapping language-specific word-forms onto language-specific word meanings, in turn constructed from universal concepts)

It is easy to see that Degree 1 assumptions may underlay the scope of the problem. Even between closely related languages, we cannot expect exactly the same bundling of semantic notions even in basic lexical items and morphemes. Thus Indo-European spatial prepositions tend to encode similar notions; but we are not surprised to find the range of English on split into German auf and an or Dutch aan, om, and op (see Bowerman 1996 on the acquisition of such terms). Once outside the language family, we can be sure to find not only subdivisions of semantic space but also new cross-cutting parameters (see Bowerman & Choi, ch. 16 of this volume). Thus the semantic bundles that the child will find in its first language cannot be predicted in advance.^4 (If they could, what would conceptual development be about anyway? See Carey, this volume.)

We are often blind to the cultural foundations of our basic vocabulary: English brother seems a word that would have pertinence in most children's universes, but few languages have lexical items that denote just that. Tzeltal
jitzi'n, for example, denotes the younger male sibling of a male or the younger female or male sibling of a female, or cousins of specific kinds, and there are distinct words for older siblings. The chapters that precede this section clearly establish that many words that are learnt early have this culture-specific quality. They confirm that the problem facing the child trying to map word-forms to meanings is actually of a higher order than Degree 1. Rather than map words to preexisting meanings, the child must construct the meaning. Where the child does this construction by selecting from a reassuring, finite set of universal underlying atomic concepts, we have a Degree 2 level of difficulty.

Now the mapping problem begins to look extremely difficult: the problem space is now much greater than at the Degree 1 level. Indeed it is beginning to look vast. We can glimpse the size of the problem space by supposing (as a radical idealization) that a possible cultural concept is any combination or permutation of the set of universal concepts B. Then under Degree 2 assumptions the problem space for the child (i.e. the set of all possible pairings of word-forms and possible meanings) is the Cartesian product of A and all the combinations and permutations of B (of some maximum length) out of which vast array of possible mappings the child must select just the small set of actual ones.

Perhaps not all words are as problematic as this – Gentner & Boroditsky (this volume) argue that relational nominals and predicates have a linguistic relativity of this sort, but other nominals may allow a way into a language by offering a Degree 1 transparency. But nevertheless it is clear that amongst the early words a child masters are indeed words of Degree 2 complexity.

1.3 Degree 3.0 Mapping Problem: mapping L-specific word-forms onto L-specific word meanings, given non-universal working concepts

Degree 2 mapping problems look staggering difficulty. Yet children succeed. Perhaps they succeed at something even harder: Degree 3 mapping problems. Under a Degree 3 mapping problem, the kinds of semantic units mapped onto words are not simply a combination of a set of universal concepts B; instead they are combinations of culture-specific concepts – the set B cannot be taken to be identical for all languages and cultures. This degree of problem would arise if there is no guaranteed commonality between adult everyday working nonlinguistic concepts and the infants' naturally attainable concepts (concepts based on innate predispositions or shared terrestrial fate). Then what seems salient to the child may not be at all what the adult has in mind.

We've already met an example of the kind of fundamental linguistic variation which might constitute a Degree 3 problem, namely the difference between English behind and Tzeltal ta yanil. Adults in these two speech communities (say, Boulder and Tenejapa) don't have the choice to use the other system – their linguistic systems provide them no easy way to speak in the other manner. Nor, as I shall show below, is the other way of thinking easily accessible to them. This is because the two systems, while using fundamentally different coordinate systems, offer roughly equivalent functionality. Of course, this doesn't rule out the possibility that at some yet further, deeper level of decomposition, the two coordinate systems use the same universal primitives; but it does make the point that this universal level is not the level of everyday concepts which the child can tap into, not only through the observation of language but also through many other kinds of observed behavior; but the child has no privileged access to that culturally variable level by virtue of matching innate predispositions.

In case there is still unclarity about the distinction between Degree 2 and Degree 3 problems, consider it another way. Degree 1 problems involve finding correspondences between just two levels – the word-forms and the innate concepts corresponding directly to the meanings. Degree 2 problems consist of mappings between three levels: word-forms, word meanings, and the universal semantic primes from which they are formed. Degree 3 problems consist of mappings between four levels: word-forms, word meanings, semantic parameters, and the universal conceptual primes (if any) underlying the culture-specific semantic parameters. The more intermediate levels are between universal concepts and language-specific word-forms, the more the child actually has to construct in the way of entities to be mapped onto the other, and the more different possible mapping relations arise.

The learning challenge raised by Degree 3 problems differs from that raised by Degree 2 problems in both quantitative and qualitative ways. Firstly, the problem space – the set of possible meanings the child may construct – is now either infinite (if there are no conceptual primes), or at least truly vast (if there are, as we may assume). The vastness of the problem space can be assessed along the following lines: the set B of culture-relative everyday working concepts forms one conceptual level, a level of composite concepts or semantic parameters which are themselves recombinable to make semantic units (word meanings). For example, Tzeltal yanil might have a meaning specified in terms of a quadrant bisected at a fixed bearing – our N 010°, but their parallel lines on a notional inclined plane. So far this seems just like a Degree 2 problem. But these B-level concepts are culture-relative, and are attributable, let us suppose, just because they are in turn combinations and permutations of elements of set C, a large universal inventory of atomic concepts from which those macro-concepts at level B are constructed. Then the problem space (i.e. the set of possible pairings of word-forms A and
meanings B, ultimately strings at level C) is vastly increased because B is not a fixed inventory of meanings. And from that utterly vast space of possibilities, the child must select just the correct assignment of meaning to each and every morpheme.

Why care about the size of the problem space? Sure, in really small spaces, one might hit on the right answer by chance, and a few clues will reduce the chance to certainty. But once it is really large (as it already is in Degree 2, and arguably even in Degree 1, problems), does it matter how large it is? Well, consider two points. First, one can search even large spaces efficiently with simple algorithms (witness the game of "Twenty Questions"), but really vast spaces are another matter. When the filing cabinet gets too large, it's quicker to rewrite the document than to try to find it: construction is a better strategy than search. At that point, Fodor's imagined stock of innate meanings will be a hindrance rather than a help. A second point is that as the search space becomes vaster, the heuristics must get better. This is the paradox of linguistic relativity: cultural variation is not, as commonly thought, in opposition to nativism—the more variation (and thus the greater the problem space), the more we would need in the way of pregiven heuristics and constraints to find correct solutions to the mapping problem, as Sperber (1996) has pointed out.

Secondly, the difficulty is qualitative. Consider the Gricean theory of meaning (Grice 1957): S says something X intending addressee A to figure out that S intended X to have a specific mental effect E on A, by virtue of a shared mentality that makes it salient that S might mean E by X. Some such picture seems to be assumed by those who emphasize the shared intentional and attentional structure lying behind language acquisition (see Tomasello, ch. 5 of this volume). But suppose there is no such shared mentality—how can communication get off the ground? Only if A can build a model of what S thinks A would find salient. This is the fundamental difficulty posed by Degree 3 problems: the child must somehow discern the conceptual parameters that the adult is using to construct the semantic distinctions that show up bundled in morphemes.

Problems of Degree 3 kind may seem insoluble in principle, since no heuristics are likely to be workable in an infinite or utterly vast problem space. But as a matter of fact it seems that children do attain concepts of a culture-specific kind, which are not just trivial combinations of concepts that are culture-independent. I will spell out some examples of adult concepts of this kind below, and the chapter by Brown provides some corresponding information about child acquisition of such concepts in the same culture. We will then return at the end of this chapter to consider just how such concepts may be learnable after all. Meanwhile we should note that presumably not all linguistic meanings are of this kind: perhaps there are

Degree 1, Degree 2, and Degree 3 problems in different areas of the vocabulary (again, see Gentner & Boroditsky, this volume).

In the child language literature it has for long been thought that the target problem was of Degree 1 difficulty. Many of the heuristics proposed by, e.g., Markman (1989, 1992), Gleitman (1990), and others, are aimed at this target, which is already extremely challenging. However, as many chapters in this volume suggest (see e.g. Bowerman & Choi, Gaskins & Lucy, Gentner & Boroditsky, de León, Brown), the problem—at least in part—actually appears to be of at least Degree 2 difficulty. The particular focus of this chapter is to suggest that even that underestimates the problem, which may even be of Degree 3 difficulty: for I will produce evidence that adult nonlinguistic cognition varies, not randomly, but in line with the coding of the adult language. This does radically change the picture: instead of children having to find meanings in a cognitive space which can be presumed to be shared between themselves and adults, children actually have to construct progressively an adult-like cognitive space in which the conceptual parameters which will enter word meanings must be discovered.

2 Cultural variation in spatial frames of reference

2.1 Variation in language

I turn now to sketch the finding that adult cognition may vary in line with the language spoken. Our data come from the crosscultural study of spatial cognition. As made clear by Spelke & Tzvkin in ch. 3 of this volume, there is in human spatial cognition a strong background of common mammalian inheritance in spatial abilities. For this reason, spatial conception is one of the areas in which we might least expect to find significant variation across cultures. Indeed, there is no shortage of pronouncements in the linguistics and psycholinguistics literature to the effect that the kinds of spatial notion we find in our own languages are more or less inevitable, due either to our common biological inheritance or our shared terrestrial existence, or both (see e.g. H. H. Clark 1973; Miller & Johnson-Laird 1976).

After working in about twenty field sites around the world, our research group at the Max Planck Institute for Psycholinguistics has come to the conclusion that this picture simply is not right (see, e.g., Pederson, Danziger, Wilkins, Levinson, Kita, & Senft 1998). There is substantial variation in the semantic parameters employed in languages for spatial description, even if the variation can often be seen to be restricted within certain types. I will report now on one of these dimensions of variation, namely variation in coordinate systems, or "frames of reference" as they are called in the psychological literature (see Levinson 1996a).
I shall restrict myself to discussing the specification of angles on the horizontal, that is, how languages specify the location of one thing, the Figure or object to be located, as positioned at a specific angle or direction from another, the Ground or landmark object. (Amongst the many other kinds of spatial concept therefore ignored here are, for example, the so-called topological notions like "at," as well as deictic notions like "there," where no angular specification is given, which therefore lie outside the scope of these remarks.)

Finding angles on the horizontal is a non-trivial task: there simply is no overwhelming force on the horizontal as there is gravity on the vertical – we have to invent or construct coordinate systems and apply them. In a tradition that goes back at least to Kant (1768), it has been assumed that languages will inevitably make central the egocentric, person-based coordinates of the kind exemplified in our "left/right/front/back" terms. This turns out not to be the case. Instead, languages make differential use of three quite different strategies for solving the problem of specifying angles on the horizontal, which we will call the Intrinsic, the Relative, and the Absolute frames of reference. In the Intrinsic frame of reference, designated facets of the Ground object are used to specify an angle, which in turn can be used to specify a search domain in which the Figure will be found, as in "The ball is in front of the chair" (see figure 19.1, top). In the Relative frame of reference, the body planes of the viewer can be utilized to extract a coordinate system, such that one can say, e.g., "The ball is to the right of the chair" (figure 19.1, middle). In the Absolute frame of reference, fixed bearings based ultimately on such things as celestial, meteorological, or landscape constancies can be used to specify that, e.g., "The ball is north of the chair" (figure 19.1 bottom).

These systems may seem familiar enough, but the familiarity can be misleading. Note for example that north of the chair is not a locution we would normally use, because, for one thing, most English speakers will not at any one moment (especially inside a house) know where North is. For another, we reserve all such locutions for geographical, not table-top or intimate, space. Secondly, each of these frames of reference is a large genus, with many language-specific distinct species. For example, there are Intrinsic systems that are anything but familiar; thus, whereas for us the "front" of a television or a book or a building is defined in terms of the functional properties of those objects, for speakers of Tzeltal the "face" of an object is defined by strict geometry (see Levinson 1994). Similarly, our Relative system could be different: whereas for us a man hiding behind a tree from us could be said to be "behind the tree," for a Hausa speaker he would be "in front" of it (Hill 1982); and in conditions where we would say the man is "to the left of the tree," other languages may prefer "to the right of the tree," etc.
having mapped a left and right upon the tree. Finally, Absolute systems are again various: many have abstract bearing systems askew from our North/South/East/West (see e.g. Haviland 1979; Levinson 1992a, 1996b), others use bearings derived from the stars, the winds, or landscape features like drainage systems, which may form no quadrant. Yet other systems, like the Austronesian ones, may mix landmark axes (whose directions therefore vary with travel) with true fixed-bearing systems (based, e.g., upon monsoons; see Levinson 1992b, 1996b for references). Thus each of these types is actually a great family of systems, the members of which can be radically different from one another.

But the main point here is simply that not all languages use all three systems. There are languages that use almost exclusively only Intrinsic or only Absolute coordinates (see, respectively, Danziger 1996 and Levinson 1992a). Many languages use a combination of all three systems, or just two of them; in fact the only combination that doesn’t seem to occur is Relative without Intrinsic. This variation in the basic set of coordinate systems available in a language is a fundamental dimension of semantic variation at the most abstract level, and was scarcely to be expected given the current trends of thinking in the cognitive sciences. As already indicated, most languages can be expected to differ further in the way these major types are instantiated (a) in terms of their conceptual anchors (e.g. whether a fixed bearing is determined by notional inclined planes or monsoons), (b) as a set of semantic categories (and of course formal categories too), and (c) in terms of mapping rules for deriving secondary coordinate systems of various kinds (see Levinson 1996a, b for discussion) – but this variation, important as it is, is at a different level of detail. The surprise value of the findings is the high-level variation in the fundamental kinds of coordinate system employed.

2.2 Variation in adult cognition in line with language

An interesting observation is that the distinct underlying frames of reference employed in semantic systems require different cognitive underpinnings. This is self-evident in the case of cardinal directions: if at this moment you cannot accurately point to North, let alone to where you were born, you simply don’t have the necessary “dead-reckoning” system constantly operating in your conceptual background. “Dead-reckoning” implies knowing where you are by virtue of knowing how far in each direction you have traveled: it allows you to estimate straight-line directions and distances to a range of familiar locations. The properties of such a system are non-trivial. Gallistel (1990) provides an outline of the mathematical routines required. Many species, from ants, to bees, to birds, are thought to have these routines hard-wired. But the extreme cultural variability in human populations (see Levinson 1996c) makes it probable that in those cultures where humans are good dead-reckoners, they are so by virtue of learnt “software.” In fact, the cross-cultural evidence we have seems to show that these abilities correlate with language: languages that primarily provide absolute coordinate systems for spatial description require constant dead-reckoning; consequently their speakers can consistently point to a range of locations from a novel place, while speakers of languages which predominantly use relative coordinates are highly inconsistent or inaccurate or both (Levinson 1996c). Thus speaking a language that requires absolute specifications forces a constant background computation of direction.

The possibility is then raised that semantic differences in this area run more than skin-deep. Against this background it is interesting to ask what are the cognitive implications of linguistic specializations in this area – e.g. what happens to the way we think about space if we are speakers of a language which uses only (or at least primarily) Absolute and not Relative frames of reference, or vice versa? Is there, despite the linguistic variation (and despite the fact that one might need additional computations to speak in terms of fixed bearings), just one basic human way to think about spatial arrays? Or, to put it another way, do the language-specific semantic notions employed in the language one speaks match the conceptual notions one uses to, e.g., solve nonlinguistic spatial problems?

To explore these issues empirically, we can exploit properties of each of these frames of reference, in particular the fact that these three different coordinate systems have different properties under rotation. Using the Relative system, the speaker rotates “left” and “right” with himself: thus if X was “left” of Y on a table, and the speaker goes around to the other side, X will now be “right” of Y. But using the Absolute system, the speaker’s rotation has no effect on spatial description: if X was north of Y it remains so, regardless of speaker position or rotation. On the other hand, if the array – e.g. ball X to the “left” and “north” of chair Y – rotates while the speaker remains constant, both Absolute and Relative designations must change; however Intrinsic designations (e.g. “The ball is at the chair’s front”) are invariant to such external coordinates – the coordinate system is based within the assemblage.

Thus we can with relative ease devise nonlinguistic tasks that will reveal which kind of coordinate systems subjects utilize to solve them. So now we can explore the question: if a language L provides for a particular kind of array just one natural frame of reference, is this frame of reference also the one employed by speakers of that language when performing nonlinguistic tasks on similar kinds of array?
Let me illustrate with a very simple nonlinguistic task, utilized to distinguish whether subjects coded for recall memory using an Absolute or a Relative frame of reference (if they used neither, they failed the task, in the sense that they produced random or arbitrary results). Subjects were shown three toy animals in a row, asked to memorize them, and after a short delay turned around 180°, led over to another table, and asked to make the same assemblage again. If they preserved the left–right direction and ordering of the animals under rotation, they were clearly using body-centered or Relative coordinates, while if they preserved the, say, northwards orientation and North–South ordering of the row, they were clearly using some kind of coordinate system locked to the larger environment or to abstract bearings (like “North”). Figure 19.2 illustrates the task diagrammatically.

On the basis of many informal observations, we made the simple prediction that if a language provides only or primarily an Absolute frame of reference for the description of such arrays, then speakers of that language when performing nonlinguistic tasks would also employ the same frame of reference; and conversely we expected speakers of languages favoring the Relative frame of reference to employ that same coordinate system in their nonlinguistic memory. The results are broadly in line with the predictions. For example, just comparing two contrasting speech communities – Tzeltal-speaking Tenejapans, who utilize an Absolute system in language to describe arrays of such a kind, and Dutch-speaking subjects, who of course use a Relative linguistic system like English – one obtains the results illustrated in figure 19.3. The abscissa here is an index of “Absolute” responses; the more trials performed using Absolute coordinates the higher the score, the more performed using Relative ones the lower the score. As is immediately evident, the Dutch were consistent Relative encoders on this memory task, while the Tenejapans were Absolute encoders. In short, subjects appear to memorize spatial arrays using a coding system isomorphic with the language they speak.

We have carried out a research group parallel examinations in a large range of languages (see e.g. Levinson & Nagy 1997; Pederson et al. 1998). What we find is that where languages rely primarily on an Absolute or a Relative frame of reference for the description of arrays of a similar sort, then we obtain just the same kind of results as we have just reported for the Tenejapans on the one hand and the Dutch on the other. Thus there appears to be a robust tendency for the mental coding of arrays to follow the pattern in the subject’s language.

A single task is only a diagnostic. But in many of these fieldsites we have been able to run a whole battery of tasks. The idea here was to see whether the effects are only “skin-deep” as it were, or whether if we give subjects tasks involving differing cognitive capacities – recall, recognition, pattern matching, inference – they nevertheless persist in preferring the frame of reference enshrined in and encouraged by their language. For example, in Tenejapa Penelope Brown and myself have conducted a series of tasks involving propositional coding of spatial arrays for recall, recognition, and inference. In the inference task (designed by E. Pederson), the subject sees a spatial “premise” of a kind that might be coded “blue cylinder to the left of yellow cube,” or conversely “blue cylinder to the north of yellow cube.” He or she is then rotated 180°, and sees a second spatial “premise,” which might be coded as, say, “yellow cube to the left of red cone” or “yellow cube to the south of red cone.” The subject is then rotated back to the starting position and asked to draw the nonverbal conclusion from the two premises by arranging the red cone with respect to the blue cylinder. Absolute coders do it one way, Relative coders another (work it out – or see Levinson 1996a). In addition to these kinds of tasks which clearly require a coordinate system but which require no precise metric retention, we also carried out tasks which would seem to require visual memory for metric distances and angles. For example, we asked subjects to memorize a heap of objects, and then after rotation to recall and rebuild the assemblage: Dutch subjects
Covariation between spatial language and cognition

recall the array as seen from the original vantage point, Tenejapans remake the array as if they had seen it from the other side (i.e. they appear to make a mental rotation; see Levinson 1996a). Finally, we have observed unself-conscious gesturing accompanying speech. Where the gesturer speaks a language preferring Absolute coordinates, as in Tenejapan Tzeltal, regardless of rotation the speaker uses absolute coordinates when gesturing, i.e. points in the correct direction to the places or events described (even when, if less consistently, this was a virtual event like a movie stimulus); in contrast Dutch speakers when rotated also rotate their coordinates with them. In short, if one finds a clear preference for one coordinate system in the semantics, one turtle supporting linguistic behavior as it were, then it is turtles all the way down!

This consistent behavior across a range of tasks or behaviors is fundamentally interesting. It appears to show that the coordinate system pre-

dominant in the subject's language comes to work its way into a number of distinct kinds of representational systems, so that we end up with an isomorphism across semantic representations, nonlinguistic propositional representations, imagistic representations, and even kinesthetic ones. Indeed, it is probably this which makes it possible to speak about what we remember, see, or feel (see Levinson 1996a).

These are the sorts of findings that seem to substantiate the view that the kind of problem facing the child learning a language is more formidable than had been imagined. Instead of being able to assume that an adult will share the same fundamental "take" on a scene, with a consequent likely mapping of words to things or relations, the possibility arises that adult conceptual classifications will be initially quite unfathomable. Two table settings opposite one another with fork on left and knife on right may appear to be the same array from the perspective of the diners in a Relative perspective, and appear systematically opposed (or contrary) representations from the perspective of Absolute diners (after all, one fork points, say, west and is north of the knife while the other points east and is south of the knife). If there is a fundamental mismatch in child and adult cognition, how does the child even begin to make the first tentative steps towards cracking the local linguistic code? 

3 Solving the impossible

At the beginning of this chapter, I outlined a ranking of learning problems from what I called Degree 1 to Degree 3, according to the size of the problem space within which the child must find the solution. The adult language and cognition data I have now described suggest that the child really does face Degree 3 problems in some areas of the vocabulary.

Now the suspicion must arise that all this amounts to a reductio ad absurdum: there must be something wrong with this depiction. If the Degree 1 Mapping Problem, the Quinean conundrum, is already so difficult, then apparent demonstrations that the child's problem is actually Degree 2, let alone Degree 3, can be discounted: if these demonstrations were correct, language would be unlearnable. Ergo, there must be some flaw in those demonstrations. After all, even powerful heuristics and constraints may not be enough to help one find an atom in a haystack. What use would be innate domain-specific theories, special constraints on expected word meanings, etc., if the problem space is impossibly vast?

In a sense, this is indeed correct: namely, the problem is underdescribed. There are at least two fundamental heuristics that lie outside the mapping space problem as described. Firstly, even the Degree 1 problem presupposes that the child has a theory of meaning, perhaps along the lines
sketched by Grice (1957). Grice's theory of meaning situates meaning-recognition within the sphere of intention-recognition, and brings the interactional context of joint focus of intention, mutual salience, etc., to bear on the problem (see e.g. Schiffer 1972). It is easily empirically demonstrable that such contexts allow determine solutions to infinite mapping problems (see e.g. Schelling 1960; H. H. Clark 1996), thus beating the mathematical odds (see Levinson 1995a): if you and I will both get $1,000 if we think of the same number without communication, we can easily achieve this "telepathy." These issues are addressed by Tomasello (1995:116, and ch. 5 of this volume). Now, I have suggested that the infant might be in a state of cognitive mismatch with the adult, because the adult is thinking along lines consonant with the language he or she masters, yet to be learnt by the infant. This of course makes it harder to achieve "telepathy." On the other hand, the very fact that the adult's thinking is isomorphic to the categories of the language provides a whole barrage of nonlinguistic, behavioral clues to the nature of those categories. For example, many Absolute speakers arrange their world in such a way that environmental constancies reflect fixed bearings: windbreaks always to the East, or hearths to the South, in just the same way that Relative speakers arrange their environments so that forks are left of knives, cars drive consistently on the left or right, knobs turn things on clockwise, etc. Adult cognitive isomorphism to language presents the child with countless behavioral clues to the underlying cognition.11

Secondly, there is a temporal succession within the mapping problem that is often omitted in discussions. Once the child has cracked some word meanings, a pattern for a specific language begins to reveal itself, allowing the assimilation of accumulative clues which reinforce language-specific mapping solutions (see Choi & Bowerman 1991:107; Brown, ch. 17 of this volume). The importance of the temporal succession of word learning is underlined by the demonstration by Linda Smith in this volume (ch. 4) that even quite simple connectionist or associationist models of the learning process will exploit such emerging patterns. Temporal succession might provide a solution in another way: the child might progress in orderly sequence from Degree 1 to Degree 2 to Degree 3 solutions. This could happen in two ways: (1) she might initially think all words have Degree 1 solutions, only later learning that some have Degree 2 solutions, and yet later revising her opinion and noting that some even have Degree 3 solutions; alternatively, (2), the child might be able to detect differences in the size of the problem spaces associated with different words (e.g. perhaps she can tell that some words are inscrutable) – then she might assign Degree 1 solutions only to Degree 1 words, using them first, and only later have a shot at Degree 2 words, aiming for Degree 2 solutions, and so on. This is essentially Gentner & Boroditsky's view (ch. 8 of this volume), who suggest children learn concrete nouns first, verbs second.

It is worth mentioning a third set of powerful heuristics, often omitted from discussions of this kind (but see E. V. Clark 1993 on the principle of contrast, and Merriman, Maratiza, & Jarvis 1995:151ff. for a survey of other suggestions). These are pragmatic principles which appear to govern the structure of a vocabulary field, so that, e.g., alternate terms may be assumed to contrast in one of a small number of ways, partially guessable by the formal properties of the expressions (Levinson 1995b). Such heuristics may be readily learned and transferred from one learning task to another.

While these three kinds of heuristic or additional information significantly restrict the kinds of hypotheses that might be entertained by the child, one must concede that they do not alone show how Degree 2 or 3 problems could in principle be solved. But what they do show is that the discussion of the problem in these terms is not sufficient, is perhaps even a misleading way to think, and that there may be any number of further restricting heuristics that have so far been neglected in the analysis of the child's learning problem. For this reason, I do not think that there is any good implausibility argument to the effect that the child cannot really be successfully navigating Degree 2 or Degree 3 problem spaces.

The findings announced in the chapters by Bowerman & Choi, de León, and Brown all point clearly towards the conclusion that the problem space for the child is at least of the Degree 2 kind, and there are hints (perhaps especially clear in the chapters by Brown and de León) that it is actually of Degree 3. Direct corroboration that the problem is indeed of the Degree 3 kind comes not only from our work reported above on adult cognition, but also from the work of Gaskins & Lucy (ch. 9 of this volume), which includes studies of middle childhood.

The Gaskins & Lucy chapter suggests another interesting possibility. Perhaps children construe Degree 3 problems as if they were Degree 2 problems: that is, they map directly from the deep conceptual primes to the word meanings without going through a level of culture-relative composite semantic parameters. This would be a bit like treating every instance of the number 20 as a long sequence of Is, 11111111111111111 (corresponding, e.g., to all my fingers and toes), rather than as a multiple of 10s. You can still do simple arithmetic if you haven't grasped the decimal system, it's just hard work. Only in middle childhood, so such an account might go, do children remap the system from three levels (as in Degree 2 problems) into four (as in Degree 3 problems), incorporating a level of everyday macro-concepts. It is only at such a point that they think like adults think, and can come to acquire the more complex parts of lexical semantics (for a review of
such later reorganizations of previously acquired semantic and morphosyntactic structures, see Bowerman (1982).

Indeed, a number of the spatial concepts discussed here are not acquired early, and in various ways fail to meet the criteria outlined by Landau & Gleitman (1985:178) for "natural" or innate categories. Such "natural" categories should be learned early, alternative construals should hardly arise, they should be universal in the core vocabulary, and they should be learnable under poor input conditions. For example, Relative concepts like "front," "back," "left," "right" are not universal. Nor are they learnt early: Relative as opposed to Intrinsic "front" is mastered only by 3% of English-speaking children up to 4;4 (Johnston & Slobin 1979), and of course problems with "left" and "right" are notorious (as Piaget pointed out, Relative "left," as in "ball left of tree," is not fully mastered till as late as age 11). By these criteria, it is the Intrinsic concepts ("front," "back," "between," etc.) that the English-speaking children find relatively "natural." But we do not find this same picture across languages—that is part of the message of the chapters by de León and Brown (this volume): e.g. Absolute semantic concepts may precede Intrinsic ones. This cross-cultural variability in what is most easily accessible to the child suggests that many linguistic categories are simply not natural in any straightforward sense at all: they have to be learnt from instances of usage. Sure, they may be built out of underlying "natural" concepts, and moreover the range of variation may be limited. But the point is that languages construct concepts that otherwise might not have been. And that is precisely the added cognitive value of language: it provides "un-natural concepts," complex conceptual wholes which connect across natural capacities (see Dennett 1991; Spelke & Tsivkin, ch. 3 of this volume), and which can be processed as units in working memory, thus vastly increasing the power of our mental computations (Levinson 1997). This picture is radically opposed to the standard line in child language research, which assumes that language rests directly on the fundamentals of preexisting categories—that is, that the learning problems are of Degree 1. On the new view, when a child learns a language she is undergoing a cognitive revolution, learning to construct new macro-concepts. These macro-concepts which are part of our cultural baggage are precisely the contribution of language to our thinking. Language invades our thinking because languages are good to think with.

NOTES

1 I am most grateful for comments by Penelope Brown and Melissa Bowerman on a draft of this chapter.

2 I'll talk henceforth in terms of "words" despite the perilous ambiguities between word-form, lexeme, type vs. token meanings, etc., and despite the fact that what we are really interested in is morphemes, simply in the interests of a colloquial style.

3 Thus Fodor (1980:151): "A theory of the conceptual plasticity of organisms must be a theory of how the environment selects among the innately specified concepts. It is not a theory of how you acquire concepts, but a theory of how the environment determines which parts of the conceptual mechanism in principle available to you are in fact exploited."

4 For a brief, trenchant philosophical critique of Fodor's position, see Putnam 1988:ch. 1.

5 For an insightful analysis of the apparently similar but in fact very different sibling terms in Japanese and Korean, see Matsumoto (1995:30-31). And Deutsch, Wagner, Burchardt, Schulz, & Nakath (ch. 10 of this volume) gives another exotic case in our own nurseries.

6 We can, indeed I would argue we must (Levinson 1997), interpolate a semantic level S between the concepts B and the morphemes A. We might hope that units of S will map one-to-one onto morphemes A, and therefore that this extra level does not increase the size of the problem space; but see discussion of Degree 3 problems.

7 Of course, we may assume that children only have access to a subset of adult concepts; the question here is whether adult cognition is fundamentally shifted from such "natural concepts" in everyday thinking.

8 Another example may help to distinguish proximate culture-specific semantic dimensions from their ultimate underlying conceptual foundations. Take the meaning of the word December; it is the twelfth month (of roughly thirty days) in our year. Compare the Tzeltal tz'un; it is the eighteenth month (of exactly twenty days) of their year; December and tz'un overlap referentially, but the concepts aren't at all the same. It would be fatuous to claim that both December and tz'un are somehow preexisting concepts, waiting to be named (but see Fodor 1983); even the underlying concepts on which they immediately rest are distinct, since our number system is to base 10 and theirs to base 20, and their year is five days shorter than ours. Still, deep down, the notion of day and number are shared, and out of these ingredients we can derive different number systems and calendrical cycles. But clearly the meaning of the word tz'un is a whole complex cultural concept, built on other cultural concepts, that in turn may be built on conceptual bedrock.

9 To be useful locations in conversation, there has to be an automaticity to comprehension (I once seriously risked an off-road vehicle in quicksand because my Australian aboriginal navigator said "Quick swerve North!" and it simply took me too long to compute the response). Less obvious, perhaps, is that people who don't routinely use Relative left/right tend to take a long time to figure out which is their left side (the Imperial Russian armies are said to have been drilled with straw tied to their left legs).

10 I have here downplayed what other scholars might have played up: namely that on the account given there are only three major types of spatial frame of reference that the child must choose between. Could this be a kind of parameter switch, inbuilt, which the child only has to learn to set? I would rather look at it
differently. Perhaps these three types of frame of reference are innately given to us, perhaps in connection with different cognitive and perceptual modalities (see Levinson 1996a); but they are only huge classes of possible geometric solutions. The local system is still always one of an indefinite array of possibilities. Talk of parameter setting will not solve the learning problem, although constraints on variation do of course restrict the problem space.

11 Would these cultural reflexes alone be sufficient for the child to build culture-specific concepts, without language input? I think this is unlikely to any significant degree. The reason is that it is communication which provides the focus of attention on the exact intentional background to behaviors (see H. H. Clark 1996; Tomasello, ch. 5 of this volume): learning a word is like a parlor miming-game – a coordination problem involving a signal whose issuance carries a warrant that there are just enough clues to find the correct solution. It is interesting to note that alien observers (e.g. anthropologists) have often failed to realize that the community they were living in utilized an Absolute frame of reference.

REFERENCES


