1 The background to the study of the language of space

Stephen C. Levinson and David P. Wilkins

1.1 Spatial language and cognition

Spatial cognition is a fundamental design requirement for every mobile species with a fixed territory or home base. And there is little doubt that it plays a central role in human thinking and reasoning. Indeed, the evidence for that centrality is all around us, in our language where spatial metaphors are used for many other domains, in the obvious cognitive utility of diagrams and tables, and in the special role of place in memory. The idea that space is a fundamental intuition built into our nature goes back at least to Kant (1768), and the idea that our apperception of space is governed by cognitive universals informs much current cognitive science.

But in some ways human spatial cognition is puzzling. First, it is unspectacular – we are not as a species, compared to bees or pigeons, bats or whales, particularly good at finding our way around. Second, human spatial cognition is obviously variable – hunters, sailors and taxi-drivers are in a different league from the ordinary city-dweller. This suggests that many aspects of effective spatial thinking depend on cultural factors, which in turn suggests limits to cognitive universals in this area.

The language of space becomes an important focus of research, then, for a number of reasons. First, it may help to reveal the underlying conceptual structure in human spatial thinking, which may be much harder to extract from an inarticulate species. Naturally, universals of spatial thinking should be reflected in universal conceptualizations in spatial language. Second, and contrastively, the very variability of language promises an interesting insight into the possible cultural variability of spatial thinking. Third, this reasoning presumes a close correlation between spatial language and spatial thinking – essentially, a (possibly partial) isomorphism between semantics and conceptual structure. Where we have linguistic universals, the correlation may be presumed to be driven by cognitive universals. But where we have cultural divergences, language may not so much reflect underlying cognition, as actively drive it.

All this suggests a natural line of research, namely a parallel, independent investigation of spatial language and human spatial thinking. In a concerted

effort over nearly a decade, in a project involving over forty researchers and as many languages, researchers at the Max Planck Institute (MPI) for Psycholinguistics have tried to pursue these parallel investigations in as many cultures of independent tradition as possible. The outcome has been surprising. Human spatial thinking is indeed quite variable, sometimes based on incommensurate conceptual systems. Languages reflect this variability, for semantic distinctions do indeed closely match conceptual structure. Moreover, sometimes there is a good case for supposing that language, and more broadly communication systems, are causal factors in inducing specific ways of thinking about space. These correlations between language and cognition, and the methods employed to probe non-linguistic spatial thinking, are the subject of the companion volume to this book, *Space in language and cognition.*

These findings give the subject of spatial language a new and vital interest. Since linguistic differences can have cognitive consequences, what exactly are the limits to the variation? What kind of semantic typology can be constructed to encompass the variation? If fundamental spatial concepts are not given in advance but vary from language to language, how can children acquire such notions? Is there a conceptual bedrock of spatial ideas on which children build? These and many further fundamental questions arise.

This books deals centrally with linguistic variation in this domain. It illustrates in detail how languages may mismatch on fundamental spatial distinctions. But it also suggests a number of constraints and a restricted inventory of possibilities. It demonstrates a method of controlled comparison which can reveal both recurrent regularities and contrastive differences across languages. In the conclusions to this volume, both universal patterns and axes of variation will be reviewed and illustrated from the material elsewhere in the book.

1.2 Nature of this book

This book collects together in one volume closely comparable descriptions of spatial language in a dozen languages, nearly all from unrelated stocks. It allows one to see more or less at a glance how differently languages may treat a single important semantic domain. Curiously, information of this kind has never before been made available – instead comparisons have focussed on particular parts of speech (like spatial adpositions), or have focussed on the particular resources of an individual European language. Information on spatial description can, of course, be found in grammars, but it is distributed and always incomplete, and one cannot reliably compare one such description with another. In contrast in this book, in order to achieve close comparison, the papers each touch upon a series of key topics, and the researchers have all used a shared set of elicitation techniques. In each case, fieldwork has been undertaken specifically

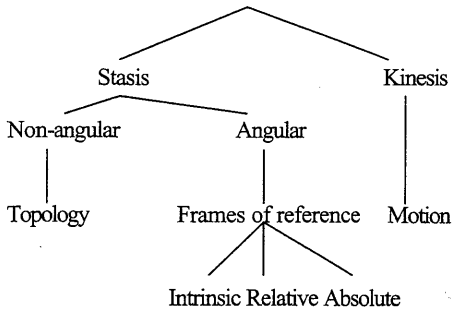

Figure 1.1 Conceptual subdivisions of the spatial domain

to illuminate the issues at hand, and each paper represents a summary of in-depth research, which has been subject to extensive mutual discussion. This kind of collaborative work is rare in the social sciences, and we hope that it will inspire more joint efforts of this kind.

This book therefore provides a unique window on how an important conceptual domain may be coded differentially across languages. For many researchers in linguistics and cognitive science the degree of diversity will come as a profound surprise. On the other hand, the existence of underlying constraints on the spatial imagination is also clearly revealed in the very extent to which close comparison and contrast is possible.

The basis of comparison has emerged from a long-term project on spatial language and cognition at the MPI for Psycholinguistics. The reader will find that the spatial domain has been partitioned into 'topological description', 'motion description' and 'frames of reference'. This partition does not exhaust the domain – spatial deixis, for example, is orthogonal and will be treated in a sister publication – but we have selected these sub-domains because they cover the major themes in the literature. The partition itself reflects major conceptual cleavages in the domain: stasis vs. kinesis on the one hand, and angular vs. non-angular static descriptions on the other (see Figure 1.1).

Leibniz and Newton (through his protégé Clark) had a heated exchange on the essential nature of spatial concepts, Newton insisting that space was an abstract envelope, while Leibniz insisted that it was relational. Most (but not all) natural language descriptions of spatial scenes are Leibnizian – that is, they describe the location or motion of one thing with respect to other things. Thus in a spatial description, something – call it the 'figure' (theme or trajector) – is generally located with respect to something else – call it the 'ground' (or landmark).

The conceptually simplest spatial description simply indicates a spatial coincidence of figure and ground. This is the core concept in the topological

sub-domain, but we can also subsume relations of propinquity, contact and containment – thus English prepositions 'at', 'on' and 'in' are usually considered to lie at the heart of the topological sub-domain (Herskovits 1986).[1] Once figure and ground are separated in space, such non-angular specifications are not of much use – we want to know in which direction from a ground we need to search to find the figure. Some kind of coordinate system now comes into play. One way to specify an angle is to name a facet of the ground and indicate that the figure lies on an axis extended from that facet, as in 'The statue is in front of the cathedral'. We call this the 'intrinsic' frame of reference, since it relies on a prior assignment of 'intrinsic' or inherent parts and facets to objects. Another way to specify an angle is to use the viewer's own bodily coordinates, as in 'The squirrel is to the left of the tree'. This is, of course, useful where an object seems to lack intrinsic facets useful for horizontal discriminations, like trees. A third way to specify angles is to use fixed bearings – independent of the scene – to specify a direction from a ground or landmark, as in 'The coast is north of the mountain ridge'. We call this the 'absolute' frame of reference, because the names and directions of the fixed bearings are fixed once and for all. Although there are many intriguing variants of these three kinds of coordinate systems or 'frames of reference', these three types (intrinsic, relative, absolute) seem to exhaust the major types used in natural languages.

Nearly all descriptions of motion also involve Leibnizian reference to landmarks or ground locations (exceptions are statements like 'In the summer the geese fly west', where 'west' is not a place but a direction). Motion is typically specified as motion to (or towards) a 'goal', or from a 'source'. Specification of both (as in 'He went from Antwerp to Amsterdam') determines a unique vector – so one can specify a direction without employing frames of reference. Deictic verbs of motion (as in 'He came late') may specify a goal (or source), namely the place of speaking. Often, though, frames of reference will be employed either exclusively (as in 'In the summer the geese fly west') or as part of, or in addition to, goal or source specification (as in 'He ran off behind the building'). Apart from deictic contrasts, verbs of motion may build in 'attainment of goal' as in 'reach, arrive', or departure from source as in 'leave'. Verbs of motion may also package other semantic material, like manner of motion, and even languages with very restricted verbal inventories seem to have a set of contrastive motion verbs (see the description of Jaminjung in Chapter 3).

There are many other kinds of variation in spatial coding across languages, as the reader will find exemplified in this volume. First, within each of these sub-domains, there are quite variable conceptual distinctions. For example,

[1] 'Topology' is here used with some departure from the well-defined mathematical concept. The term came into linguistic description through Piaget's analysis of the spatial concepts of children and includes a number of spatial relations that are not strictly speaking topological.

the topological relationships encoded in specific languages overlap and cross-cut one another – there is no one-to-one mapping of spatial relators cross-linguistically. In the frame-of-reference domain, not all languages utilize all three frames of reference, and each frame of reference may be instantiated in quite distinct concepts across languages. For example, where languages have a 'left'/'right'/'front'/'back' system used in such expressions as 'behind the tree', 'behind' and 'left' can mean exactly the converse of what they mean in English. And in the motion domain, languages differ in what is conceptually grouped or packaged in motion verbs.

A second major axis of variation is how these concepts are coded linguistically. Existing literature on spatial language gives the impression that the heart of spatial description is generally encoded in a set of contrastive spatial adpositions. Thus in English we use the same kind of prepositional phrases in topology ('in the bowl'), frames of reference ('in front of the building') and motion description ('into the building'). But many languages deploy distinct grammatical and lexical systems in these different domains. Further, some languages have no spatial adpositions. Others have only one general-purpose adposition. Such languages perforce code spatial relations elsewhere in the clause, frequently in the verb, or in local cases, or in special spatial nominals, or in adverbials. In general, most languages distribute spatial information throughout the clause. For example, a topological relation (as in 'The cup is on the table') may often be expressed through the simultaneous deployment of a number of contrastive choices in lexicon and morphology – one may say in effect something like 'The cup table top-AT stands', where 'top' is drawn from a set of contrastive spatial nominals, AT is expressed by case or adposition, and 'stand' contrasts with 'sit', 'hang' and other locative predicates.

There are no simple, hard generalizations about exactly where in the clause different kinds of spatial information are encoded. Nevertheless, as a generalization, one can say that the shape of the figure is normally encoded in locative predicates, and only occasionally in adpositions, while the shape and geometry of the ground is typically coded in adpositions and spatial nominals; the spatial relation between figure and ground may be encoded in locative verbs and case, but is especially to be found in adpositions and spatial nominals.

It is the combination of these two axes of substantial variation – semantic and grammatical – that is illustrated throughout this book. This variation raises the fundamental cognitive questions alluded to in the prior section – how are we to reconcile incommensurable semantic parameters with 'the psychic unity of mankind'? How do children then learn semantical concepts for which they cannot be prepared by independent cognition? The variation also raises a series of questions within comparative linguistics:

- What constraints are there on the *semantic parameters* involved – in short, what does the *semantic typology* of space look like?

As we shall see, despite a great deal of variation, the high-level typology here seems quite constrained. But at a greater level of detail there is sufficient variation to ensure that comparable expressions in different languages scarcely ever have the same meaning and extensional range.

- What constraints are there on the formal expression of these semantic types – what does the morphosyntactic typology of spatial expression look like?

Contrary to the literature, we will find that spatial notions are not universally encoded in specific parts of speech like adpositions or case inflections but are distributed throughout the clause.

- Are the various kinds of conceptual domain in spatial description (as in Figure 1.1) formally distinguished in languages?

As already hinted, the answer is not always, but the distinctions exist often enough to suggest that these domains do mark natural cleavages.

- How much spatial information is coded in language and how much inferred, and are the patterns the same across languages?

What we will find is that although the same kind of pragmatic principles are arguably universally in play, languages do not universally code semantically to the same level of specificity. For example, in many languages the distinction between 'on the table' vs. 'in the bowl' will not normally be coded, but rather left to pragmatic inference from expressions of the kind 'table-LOCATIVE' vs. 'bowl-LOCATIVE'.

1.3 The language sample

It is not possible in a volume of this kind to have sketches from a representative sample of the world's languages – such a book would have perhaps 400 chapters! Instead, what we have collected here is something of an opportunistic sample, which has arisen from the chance the authors have had to work closely together, and thus produce closely matched descriptions of the languages in which they are expert. Nevertheless, it is a happy sample, in the sense that the languages are geographically distributed over five continents, representing cultures with major variations in environment and land use. Both small-scale and large-scale societies are represented, and there is a bias to relatively little-known languages, so that nearly all the material presented here is new, and not to be found properly laid out in existing grammars. Altogether, seven language families are represented, along with two isolates. Some regional and linguistic clusters of languages (Australian and Mayan) allow readers to come to their own conclusions about the importance of areal and genetic factors in semantic typology. Table 1.1 gives some basic details about the languages and their speakers. From a grammatical point of view, the languages offer a wide spectrum of linguistic types. There are languages with most of the predominant word orders:

Table 1.1 *Grammars of space – language sample*

Language	Language affiliation	Country where research was done	Number of native speakers
Arrernte (Eastern and Central)	Australian, Pama-Nyungan	Australia	2,000
Jaminjung	Australian, non-Pama-Nyungan	Australia	100
Warrwa	Australian, non-Pama-Nyungan	Australia	2
Yélî Dnye	Papuan, Isolate	Papua New Guinea	4,000
Kilivila	Austronesian	Papua New Guinea	23,000
Tzeltal	Mayan	Mexico	200,000
Yukatek Maya	Mayan	Mexico	800,000
Tiriyó	Cariban, Taranoan	Brazil, Surinam	2,000
Ewe	Niger Congo, Kwa	Ghana	2,000,000
Tamil	Dravidian	India	70,000,000 (world-wide)
Japanese	Isolate? / Altaic?	Japan	118,000,000
Dutch	Indo-European, Germanic	Netherlands	15,000,000 (in the Netherlands)

PHRASE ORDER IN TRANSITIVE CLAUSES (S=subject, O=Object, V=Transitive verb)

Ewe: SVO

Yélî Dnye: SOV tendency; *Japanese*: SOV [canonical]; *Tamil*: SOV

Tzeltal: VOS [both prefixes and suffixes]; *Yukatek Maya*: VOS; *Kilivila*: VOS

Jaminjung: Free Phrase Order; *Arrernte*: Free Phrase Order [V-final tendency]

Tiriyó: Free Phrase Order

There are languages of both 'head-marking' and 'dependent-marking' types (where S=subject and O=object):

ARGUMENT MARKING ['cross-referencing'] ON VERB/IN VERB PHRASE:

Ewe – No; *Japanese* – No; *Arrernte* – No [optional number marking for subject]

Kilivila – Yes, just S; *Dutch* – Yes (reduced), just S; *Tamil* – Yes, just S [suffix]

Jaminjung: Yes, both S and O; *Tzeltal*: Yes, both S and O; *Yélî Dnye* – Yes, both S and O, by free particles in VP; Tiriyó – Yes, S and O.

From a morphological point of view, within the sample there are languages of isolating vs. agglutinating vs. (mildly) polysynthetic types. And there

are various forms of morphological ergativity vs. morphological nominative-accusative patterns. In short, most of the major formal types of language are represented in the sample.

1.4 Controlled comparison: the stimuli

Cross-linguistic (and more generally, cross-cultural) comparison is fraught with difficulties. Although isolated features or traits may be readily extracted and compared, their value or function depends on the system in which they play a part. But comparing whole systems is like comparing apples and oranges, and anyway is rarely possible. Comparative linguistics and linguistic typology proceed, nevertheless, most confidently across related languages, or in areas where there are intrinsic limits to variation (like phonetics) or where there seem to be strong universals or limited types (as in morphosyntax). Comparative semantics as a systematic enterprise has hardly begun – there are only isolated domains like colour, ethnobotany or kinship where we have any overall idea about patterns of variation across unrelated languages. In these domains, the structure of the natural world (colour and its perception, the differentiation of species, biological reproduction) gives us some 'etic' metalanguage of comparison. An 'etic' metalanguage (coined on the model of 'phonetic' by Pike) is some objective description of the domain which makes maximal discriminations, so that we can specify precisely how a language groups these discriminations within its own 'emic' (cf. 'phonemic') concepts. These groupings are most easily appreciated extensionally, that is, by looking at the range of denotation for a native term; to understand the meaning or intension, we need to look at the kinds of contrasts the terms make with one another.

The semantic domain of space is altogether more complex and abstract than these more referential domains and, as we have seen, is internally differentiated into sub-domains. A simple 'etic' metalanguage is not available. Nevertheless, there are obvious ways in which to proceed. A good sample of unrelated languages will give us a sense of which kinds of discriminations are likely to be made. We can then build these maximal contrasts into a series of spatial 'scenes', and see for any one language whether they are in fact discriminated, and if so how. We can then readily compare these extensional groupings, and then (not quite so readily) explore the intensional principles upon which the groupings are made.

During the course of the space project at the MPI for Psycholinguistics, many specialized stimuli have been developed for exploring spatial language. These include specialized stimuli for eliciting deictic motion verbs, a specific instrument for deciding on the precise semantics of enter/exit verbs, various methods for eliciting demonstratives, stimuli geared to discriminations in contrastive locative verbs, and so forth. All the papers in this volume are informed

by these systematic stimuli and mutual discussions about results. But here we have chosen to focus on three main stimuli, as an illustration of the method and the kinds of comparative results that can thus be obtained.

1.4.1 Topology Series 'Picture-Book'

This stimulus is a book of seventy-one line drawings, 'The Topological Relations Picture Series', to be used in elicitation sessions with three or more native speakers. Each picture shows principally two objects, one of which is designated (by an arrow, or coloured yellow in the original) to be the figure object, the other the ground. The native speaker is asked how one might colloquially answer the question 'Where is the X (the figure object)?', given the kind of association between figure and ground indicated in the picture. This is not intended to be a mechanical elicitation procedure – the investigator may need to choose alternative local items to be found in similar configurations, and a range of answers should be collected, noting which occur in which order, and which are said to be preferred or most normal. Three or more consultants allow some qualitative and quantitative analysis of preferred solutions.

The edition used in the chapters below is the 1993 version from the MPI for Psycholinguistics (the original design is by Melissa Bowerman, with supplementary additions by Penelope Brown and Eric Pederson). The book was specifically designed to investigate the maximal range of scenes that may be assimilated to canonical IN- and ON-relations (and thus includes a number of scenes unlikely to be so assimilated). English, for example, might be held to have a prototype ON-relation at the heart of the preposition *on* (as exemplified in *The cup is on the table*), but many other kinds of spatial relations – like a ring on a finger, a picture on a wall, a shoe on a foot – are assimilated to the same preposition. Not surprisingly, perhaps, even closely related languages like Dutch prefer other contrastive adpositions for many of these scenes. The full set of pictures include spatial relations that contrast on a range of partially overlapping dimensions:

+/– horizontal support
+/– vertical support (hanging)
+/– adhesion
+/– liquid/mastic adhesion
+/– marks on surface
+/– living creature on non-horizontal surface
+/– attachment of projecting figure to ground
+/– attachment by cord
+/– encirclement
+/– envelopment
+/–clothing/adornment

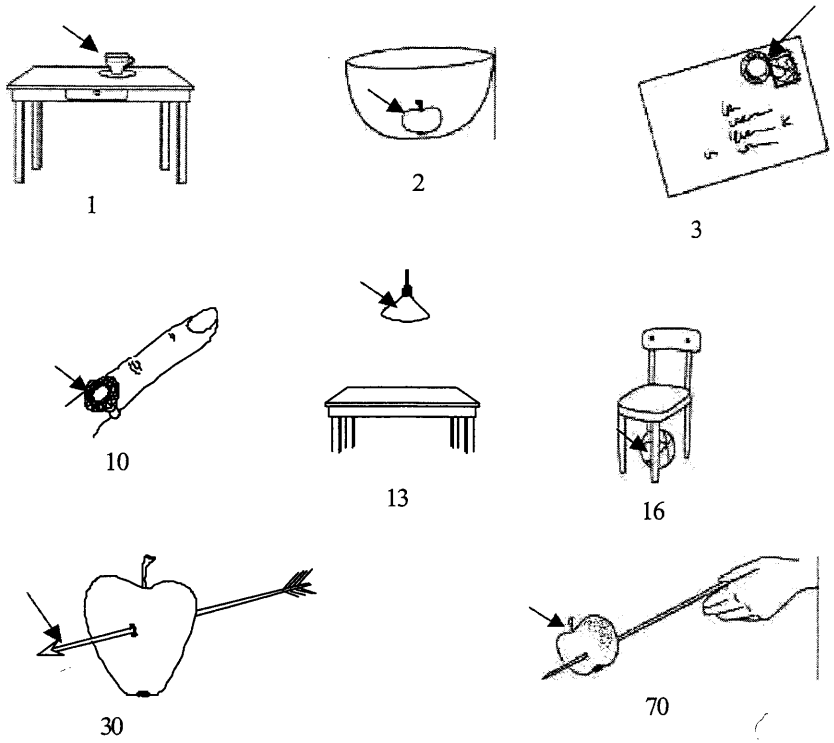

Figure 1.2 Set of pictures from the 'Topological Relations Picture Series'

+/– complete containment
+/– partial containment
+/– containment in liquid or mass
+/– containment in encircling boundary
+/– attachment by piercing
+/– negative spaces (holes, cracks)
+/– vertical non-contact (above)
+/– behind
+/– in front of
+/– under
+/– next to

For reasons of space, we have chosen just eight of these pictures to form a set over which the languages represented in each chapter can be compared. They are reproduced in Figure 1.2, with their original numbers (Pictures 1, 2, 3, 10, 13, 16, 30, 70). Authors of the chapters below occasionally mention other pictures, and the full set can be found in Appendix 4 at the end of the book. The pictures were selected on the basis of a prior study which showed

that these represent maximally different scenes from the point of view of the differentiation of spatial adpositions.[2] They include both canonical IN- and ON-relations, and then some other relations allowing some maximal contrasts between, for example, contact and non-contact, or attached vs. non-attached, as well as what happens in figure/ground alternations. For reasons that are discussed in Section 1.5.1 below, it is interesting to see how freely placed objects contrast with attached ones, and how such special spatial relations like figure piercing ground, or figure as personal adornment, are dealt with in spatial descriptions. Experience shows that languages differ greatly in the extent to which these more specialized situations are assimilated to central topological codings.

1.4.2 The Men and Tree Space Game

Structured elicitation sessions using controlled stimuli as in the picture-book described above are not the only way in which controlled information can be obtained about spatial description. An often more revealing method is to structure an interaction between native speakers over a set task. In the Space Games series, a native speaker 'director' describes a stimulus to a native speaker 'matcher', who is screened off from the director in such a way that the matcher can find the stimulus from a set of contrasting stimuli, randomly arranged. Director and matcher know that both of them have the same full set of stimuli, they know they are both facing the same direction, and they know they must find descriptions adequate to identify the stimuli in the absence of shared vision. The director freely describes the stimulus, and the matcher queries the description, until both parties feel convinced that, although they have no visual contact, they have identified the same stimulus. Such games can involve photo-photo matching, as in the game described here, or photo-object matching, or object-object matching. Matching can require recognition (as in the game described here), or construction, as in the Tinkertoy game where a director has a model that the matcher must construct again from pieces (see chapter 6 on Kilivila).

The Men and Tree photo-matching series was developed specifically to investigate frame-of-reference choice. The core set of contrasts from one of these games (Men and Tree Game 2) is illustrated in Figure 1.3 (the game includes another six photos that act as controls). There are six photos (here reproduced as line drawings) of a toy tree and toy man in various positions. The structured oppositions involve both alternations in relative position (which we call *standing* relations) – tree to visual left of man, or tree to visual right of man – and alternations in the orientation of the man (which we call *facing* relations) – facing left, facing right, facing the viewer, or facing away from the viewer. In the chapters of this book, descriptions will focus on just three of these, labelled

[2] The study was by Eric Pederson and Melissa Bowerman, and remains unpublished.

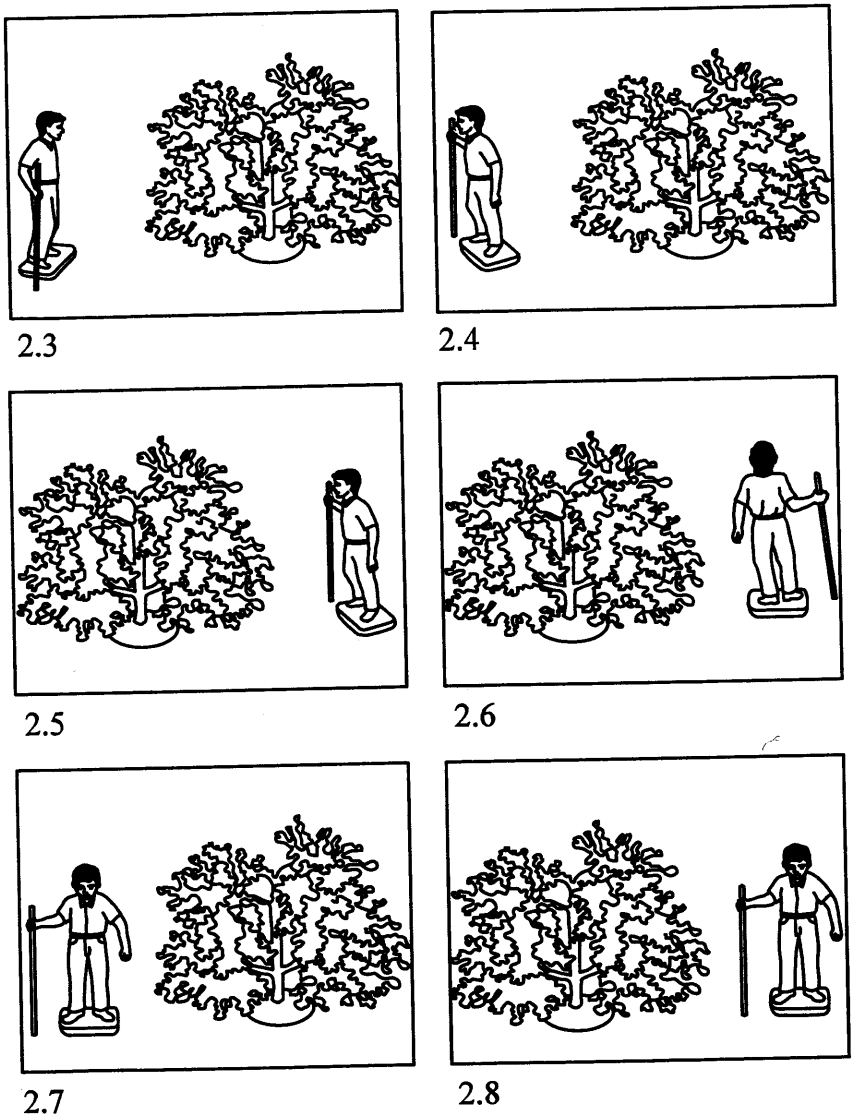

Figure 1.3 Men and Tree Game 2

2.3, 2.4 and 2.5 in the figure. The discourse that results from the game can be transcribed and queried, and can also be systematically coded for comparative purposes. A method of coding for this particular game is described in Pederson et al. 1998. The coding method allows one to isolate expressions that can be said to be *functionally equivalent*, in that they make the same distinctions, and

further, to isolate the propositional content used to make the functional distinction. For example, 'man to left of tree' may contrast with 'man to right of tree'; propositions in terms of 'man to south of tree' and 'man to north of tree' may make the same functional contrast, but involve different semantic parameters (or propositions) – in this case, different frames of reference.

A number of other 'games' of this sort have been employed by the authors of the chapters below to arrive at their general conclusions about how spatial description works in the languages in question. For example, another game (the Route Directions task) was specifically devised to elucidate frames of reference in motion description, and involved a director describing the motion of a toy man through a model landscape in such a way that the matcher could emulate it in an identical landscape.

1.4.3 The Frog Story

As an example of the stimuli that may be used to obtain motion descriptions, we have chosen the 'Frog Story' to exemplify different patterns of motion description across languages. The story comes from the wordless picture-book *Frog, where are you?* by M. Mayer (1969), published as a first book for children. It has been successfully used as a stimulus in the study of the development of narrative skills in Western children by Berman and Slobin (1994; the full set of pictures is published there as an appendix). This study has revealed major differences across languages in the way in which complex motion scenes are coded linguistically. The Berman and Slobin procedure (1994: 20) is to present the picture-book to children, who leaf through the twenty-four pages, and then retell the story to an interlocutor as they leaf through the book again. The story is recorded and transcribed in the normal way.

As a stimulus for cross-cultural research the Frog Story has certain limitations – as Wilkins has pointed out (see Berman and Slobin 1994: 21–2), it presupposes many details of Western semiotic conventions. In many of the cultures reported on in the chapters below, picture-books have no currency at all, and straightforward narratives are not always obtainable. Still, the very fact that it has been used in well over fifty different languages makes it an invaluable point of comparison. Except where noted below, the Frog Story retellings are by adults to other native speaker adults who have not seen the book.

For the purposes of this book, as an illustration of complex motion description, we have chosen four pictures that detail a crucial event in the story (what Slobin calls a *journey* – see §1.5.2 below), where a boy (the hero of the book) is picked up on the antlers of a deer and, with his dog running beside, is taken to a cliff and dumped over the cliff into a pond. This allows us to compare how such complex events are coded, how manner and path (or trajectory) are expressed, how source and goal are specified, and how simultaneous vs. sequential events

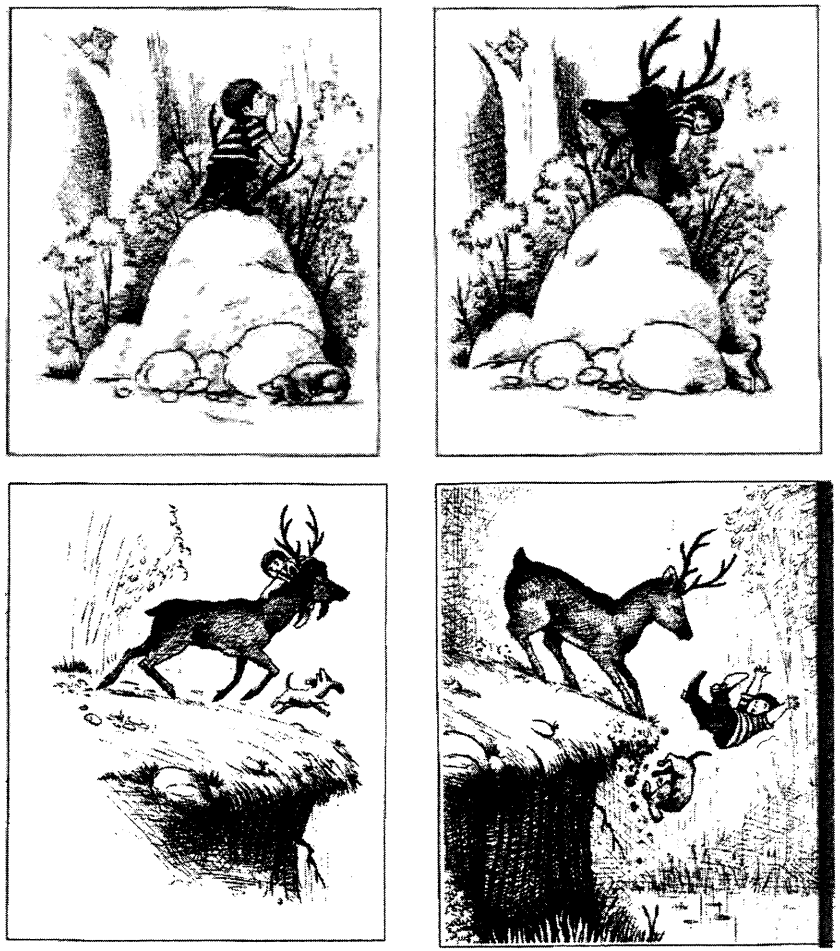

Figure 1.4 The cliff scene from the Frog Story

are coded. More detailed analyses of these descriptions in Arrernte and Tzeltal can be found in Wilkins (1997b), and Brown (2000).

1.4.4 Other elicitation tools

A number of other, more carefully designed elicitation devices for motion semantics are referred to in the chapters below. One of these is the 'COME' and 'GO' Questionnaire, a series of scenes devised to elucidate deictic distinctions in motion verbs. This questionnaire provides a series of twenty scenes,

discriminating, for example, motion to vs. towards vs. obliquely towards the deictic centre. The questionnaire and some results are described in Wilkins and Hill 1995. Another such tool is the ENTER/EXIT elicitation film designed by S. Kita, where motion vs. change of state are precisely distinguished. Some of the interesting contrasts here are exemplified in the Japanese chapter below (see also Kita 1999, Senft 1999b).

1.5 Patterns of variation

In the conclusions to this book, the reader will find a systematic comparison of the patterns of variation exemplified in the languages described in this book. But here it will be useful to preview some of the themes and provide some comparative terminology to aid the reading of the individual chapters. Each chapter touches on the three sub-domains mentioned earlier – topology, motion verbs and frames of reference – and we will take these in turn.

1.5.1 Topology

When comparing spatial language it is essential of course to compare like with like, and specifically to specify *functional equivalents*. Since all languages appear to have Where-questions, we can use this as a functional frame: we will call the predominant construction that occurs in response to a Where-question (of the kind 'Where is the X?') the *basic locative construction* or BLC for short. (Note that this expression is a shorthand for 'the construction used in the basic locative function' – constructions can have different functions.) Locative descriptions, of course, occur outside the Where-question context, as in a guidebook description of the kind *The Cathedral stands at the heart of the old city, overlooking the Rhine*. Notice that such a sentence would be odd indeed as an answer to a *Where*-question, which is more likely to be something of the kind *It's in the central square*, where the locative verb is *be* and the location is given in terms of a concrete landmark. For English, then, the BLC is of the form NP BE PP, where the first NP (noun phrase) is the figure, and the PP (prepositional phrase) expresses the ground, as in *The apple is in the bowl.*

Different languages have quite different structures in their BLCs. Some, of course, have no prepositions, or adpositions, using case marking and/or spatial nominals instead (as in 'square-LOCATIVE' or 'square middle' or 'square middle-LOCATIVE'). Some languages have no locative verb, assimilating the BLC to nominal predication, but more often there are a number of locative verbs to chose from. Many languages have a small set of locative verbs or *positionals*, often related to posture verbs like 'stand', 'sit', 'lie', but also often including predicates like 'hang'. These then contrast and their usage is usually determined by the shape and function of the subject (the figure NP), under certain

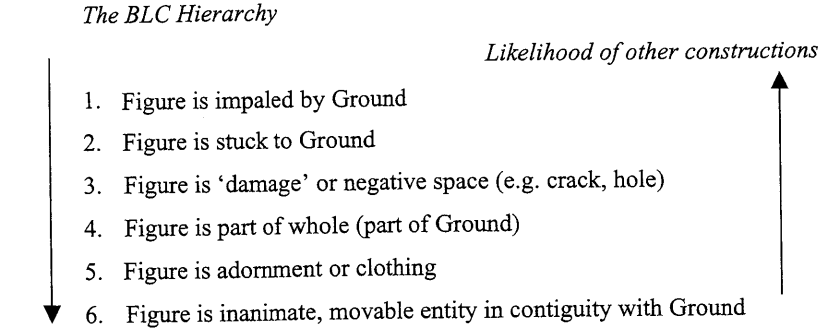

Figure 1.5 The hierarchy of scenes most likely to get BLC coding

orientational constraints (see, e.g., Chapter 5 on Yélî Dnye). Other languages have a much larger set of *dispositional* predicates used in the BLC, where the precise orientation and disposition of the subject with respect to the ground is the crucial determinant of choice (see, e.g., Chapter 7 on Tzeltal).

The BLC is thus constructed from distinct form classes – adpositions, nominal predicates, case inflections, locative verbs – according to the language. These choices are themselves influenced not only by semantic factors but also by systematic pragmatic factors. In many cases the BLC may be abbreviated. This is not merely ellipsis (as in *Where's the cup – On the table*), but a systematic way of indicating that figure and ground are in a canonical or stereotypical relation, as in the use of the locative case without further specification (as in 'The cup table-AT', where this will be understood as 'The cup is table-top-AT'). Pragmatics provides some theory for understanding these alternations (Levinson 2000a), although as a practical matter it is not always easy to decide whether the BLC has a reduced form, expanded in certain circumstances, or has an underlying expanded form, reduced in certain circumstances.

Even in response to Where-questions, languages generally deploy a number of quite different constructions. Identifying the BLC relies on the notion of a prototypical kind of scene – e.g. a moveable object on a restricted surface. Speakers of many languages will not use their BLCs to describe, for example, a ring on a finger, or a crack in a vase, or a spike through an apple – they may use other specialized constructions or resultative constructions (as in 'The spike has been speared through the apple'). In fact, it turns out that spatial scenes can be ordered in what we shall call *the BLC Hierarchy* according to the likelihood that they will be encoded using the BLC. A portion of the hierarchy is depicted in Figure 1.5. Linguistic theorizing about topological relations has suggested that spatial relations are concentrated in spatial relators – typically

adpositions – which have a limited kind of semantic content (Talmy 1985, Landau and Jackendoff 1993, Svorou 1994). As we have just seen, spatial information is in fact potentially distributed across the clause, some languages putting all the burden in the locative verb, others in case (as in Finnish). The semantic content is also not nearly as predictable as these accounts suggest. Landau and Jackendoff suggest, for example, that such semantic content is abstract and axial, while Talmy suggests it is abstract but topological rather than Euclidean. In fact, as we shall see, the information can be very specific and language-particular, reflecting cultural preoccupations. Look out, then, in the chapters below, for such specificities as the 'aquatic' ground, or distinctions between different kinds of container built into locative verbs!

1.5.2 Motion

As a first approximation, we can say that motion involves spatial change, although, as we will see, perhaps not all change of spatial relations involves motion. Change involves time, and dynamic change over time is the typical province of verbs. There has been a great deal of linguistic theorizing about the nature of the semantic content of verbs in general, and verbs of motion in particular (see Frawley 1992, Chapter 4, for a useful survey). Here we will review a number of recurring themes – the typology of lexical packaging in motion verbs, the underlying notions of path and manner, the tendency in languages for motion verbs to constitute minor form classes, the way in which source and goal are encoded, and constraints on the complexity of motion components that can be packaged within the single clause.

Talmy (1985) influentially proposed a major typological dichotomy between different kinds of motion coding in languages: verb-framed vs. satellite-framed. The typology rests on a dissection of the components in a motion event into (a) the figure, i.e. the thing moving, (b) the ground, specifying source or goal of motion, or both, (c) the path or trajectory of the motion, (d) manner of motion, (e) the predicated event itself (other elements are the site or medium in which the motion takes place, and the means or instrument of motion). Thus in *The bird flew up into a tree*, the figure is the bird, the ground is the tree, the path is expressed by *up into*, and the predicated motion together with manner of motion is expressed by *flew*. Talmy's typology rests on a simple observation: languages tend either to package the path with the predication, as in Spanish *entrar* 'to go in', *salir* 'to go out', *cruzar* 'to go across', leaving manner to an additional clause or gerund, or alternatively to package the predication with manner, leaving the path to be expressed in 'satellites' as in the English particles in *run in, crawl up, climb down*. Although the two types clearly do capture major differences in the way in which motion is packaged in languages, the

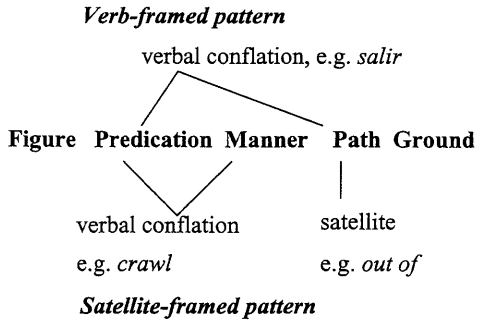

Figure 1.6 Talmy's (1985) typology of path encoding

typology has been subject to critique and revision.[3] A simple difficulty is that many languages allow both kind of packaging (as in English *go in* vs. *enter*), requiring Talmy to discern what he calls the 'characteristic mode of expression' (thus English is satellite-framed, with Romance loans displaying the contrary type in a minority, but many languages resist this kind of easy conclusion). More problematic is what exactly is to count as a satellite, since many different form classes may carry path or trajectory information – are deverbal directionals as in the Mayan languages satellites or verbs (see the chapters on Tzeltal and Yukatek below)? Some languages have very restricted inventories of verbs, but supplement them with preverbs or coverbs – see, for example, the chapters on Warrwa and Jaminjung below – and it is then no longer clear how to apply the typology.

Another doubt is raised by the notion of path. The core of a motion event might be thought to be displacement of the figure in space along a trajectory, where this trajectory constitutes the path. But careful analysis suggests that in some languages the displacement of the figure over time along a trajectory is not actually what is coded by motion verbs. We tend to think that motion must be conceptualized as *translocation*, that is as a durative event involving passage through an indefinite series of points in space over time. But there are other possibilities, with different Aktionsarten, and differential focus on figure–ground relations. In fact, on the basis of the kind of work reviewed in the chapters below (and see especially the chapter on Japanese), we will propose in the final chapter a new semantic typology of motion conceptualization.

Incidentally, although much of the dynamic component of motion events tends cross-linguistically to be encoded in verbs, this is not exclusively so. Many languages have special constructions that indicate 'motion while doing' or 'motion with purpose'. In the languages detailed below, Arrernte provides a case where there is an elaborate array of fifteen alternative categories, indicated

[3] See Frawley 1992, Schultze-Berndt 2000, Talmy 2000, Slobin n.d.

by suffixes attached to non-motion verbs, encoding motion components such as 'do upwards while VERB-ing' or 'VERB while going back'. Such rich 'associated motion' categories may be an isolated areal feature, but many of the other languages exemplified in this book have more limited categories of this sort encoded elsewhere than in the verb root.

So far, we have been concentrating primarily on the semantics of the verb, and different kinds of lexical packaging of the verb in cross-linguistic perspective. But for comparative purposes we need to consider larger units of motion description, what Slobin (1996) calls a *journey*: an extended, complex path that can include 'milestones' and subpaths each with sources and goals, possibly situated in different media. For example, the Frog Story scene pictured in Figure 1.4 was described by an English-speaking five-year-old as *He threw him over a cliff into a pond*, or by a nine-year old as *He [the deer] starts running and he tips him off over a cliff into the water. And he lands.* (Slobin 1996: 202). Slobin points out that this kind of accumulation of prepositional phrases is vanishingly rare in Spanish Frog-stories, where only one prepositional phrase per clause tends to occur. Slobin analyses this as a stylistic feature induced by structural facts. But in some of the languages described below there seem to be hard grammatical constraints on the number of ground-specifying phrases: thus both Yélî Dnye and Yukatek seem to allow at most one such phrase per clause – specification of both source and goal will require two clauses of the kind 'He left the source, and arrived at the goal.'[4] Further, it will turn out that the actual coding of source and goal is cross-linguistically variable, being sometimes coded on these adjuncts, sometimes coded in the verb, and sometimes both. In the final chapter we will propose a typology of this kind of variation.

Finally, another interesting dimension of variation concerns the extent to which languages use the same resources in the description of motion vs. stasis. Again, Talmy has suggested that they universally tend to do so, since static locatives are derivative from or modelled on motion descriptions. Thus in English, *He went out of the office* is very similar in structure to *He is out of the office*. But some languages make very fundamental distinctions between the two domains. Tzeltal, for example, uses quite different resources in the two domains – even frame-of-reference information has different coding in stasis vs. motion. Further comparisons on this dimension will be found in the final chapter of the book.

1.5.3 Frames of reference

As already sketched above, once a figure object is removed in space from a relevant ground object or landmark, it becomes pertinent to specify a direction,

[4] This contradicts assumptions in the literature that all languages permit both source and goal to be simultaneously encoded in the clause. See, e.g., Frawley 1992: 173.

or angle, relative to the landmark in which the figure may be found. Such angular or directional specifications of location require some form of coordinate system. Natural languages seem to employ only polar coordinates, specifying a direction by rotation around a ground object. As mentioned, there seem to be only three major abstract types: intrinsic, relative and absolute. These have different logical and rotational properties, which make the distinctions quite clear.

Consider, for example, a spatial array of the following kind: a toy man is placed at the front of a toy truck on a rotatable board. In the case of the relative and absolute frames of reference, the angular distinctions are mapped onto the scene from outside it, using the observer's own axes (as in 'The man is to the left of the truck') in the relative frame, and fixed absolute bearings (as in 'The man is to the north of the truck') in the absolute frame. Now if we rotate the board, the description of the scene will change – the man is now to the right of the truck, or to the south of it. But in the intrinsic frame of reference the angles are found by naming a designated facet of a landmark or ground object (like 'at the front of') within the scene to be described, and if the whole scene is rotated the description may stay the same (as in 'The man is at the front of the truck'). The intrinsic frame is thus sometimes said to be 'orientation free', while the other two frames are 'orientation bound'. However, the latter also differ in their rotational properties – if the describer walks around the scene to the other side, the relative description changes (now 'The man is to the right of the truck') but the absolute description remains the same (the man is still 'to the north of the truck').

These fundamental semantic differences justify the typology into three main types (see Levinson 1996b, 2003 for additional properties). Incidentally, although the three main types had been distinguished on the vertical dimension by psychologists interested in perception, it was not until the comparative work exemplified in this volume that it became clear that these types also structure the linguistic distinctions on the horizontal plane – partly because languages systematically using the absolute frame of reference on the horizontal had not before been properly described.

Despite the fact that there are from a logical and rotational point of view only three main types of frame of reference, there is nevertheless within each of the three main types a great deal of variation in conceptualization and coding. This is because these directional properties can be constructed in rather different ways. Let us take the three frames of reference one by one and examine the kind of internal variation they exhibit.

The intrinsic frame of reference requires some kind of partitioning of the ground object or landmark into named facets, from which search domains can be projected. All languages provide at least some such segmentations, and nearly all use them in spatial descriptions. English or Dutch does this by a complex

mixture of criteria – the 'front' of a truck is the direction in which it moves, the 'front' of a television the side one watches, the 'front' of a building the side one normally enters, and so on. These criteria thus include canonical orientation of object, functional orientation, normal direction of motion, characteristic orientation of the user, etc. (see Miller and Johnson-Laird 1976: 400–5). But some languages partition objects by more consistent criteria – for example, Tzeltal uses almost exclusively the internal geometry of the ground object (according to its longest axes and the shapes of sides – see the chapter below and Levinson 1994). Interestingly, Tzeltal largely ignores orientation with respect to the vertical, while many languages make this fundamental, what is 'top' becoming 'bottom' upon rotation. There are thus many fundamentally different ways in which this assignment of parts or facets to an object can be achieved. Despite these arbitrary complexities, children seem to master these notions surprisingly early.

The relative frame of reference involves a mapping from the observer's own axes (front, back, left, right) onto the ground object, so that, for example, one can say 'The cat is in front of the tree' by deriving a front for a tree from the observer's front – in this case, clearly, by assigning a front to the tree as if the tree was a confronting interlocutor. These mappings are complex, involving a triangulation of figure, ground and viewer, and they can be made in different ways – in some languages the 'front' of the tree is the far side of the tree (as in the well-known Hausa case, Hill 1982), and in others, what we would call the left side of the tree is the right! There are at least three distinct types of such mappings attested, and languages may mix them (for the details see Levinson 1996b, 2003). An additional source of complexity is that some languages, like English, use the same terms like 'front' and 'left' in both the intrinsic and relative frames of reference. Thus 'The tree is to the left of the man' may be ambiguous: it may mean that the man is facing us, with the tree at his left hand, and thus to our right (an intrinsic interpretation), or it may mean that the tree is in the left visual field regardless of the man's orientation (a relative interpretation). Some languages reduce the ambiguity, either structurally (requiring, e.g., a possessive like 'the man's left' for the intrinsic interpretation), or by procedural rule (as in Kilivila where an intrinsic interpretation takes priority over a relative one wherever the ground has inherent named sides). These systematic interactions between the intrinsic and the relative frame of reference are thus further sources of variation.

The absolute frame of reference in ordinary language use requires fixed bearings that are instantly available to all members of the community. English has a word for 'north', but few Englishmen can effortlessly and reliably point to north, and it does not figure in normal discourse about small-scale spatial relations. Nor do we have clear conventions about what range of horizontal arc will count as north. But there are many communities where conventional

fixed arcs are established and instantly available to all competent speakers of the local language. Such a system can then make the relative frame of reference irrelevant and unnecessary, and there are thus many languages which do not employ a relative 'front', 'back', 'left', 'right' system. Absolute coordinates can be based on many different sources – solar compass, sidereal motion, wind directions, river drainage, mountain slopes, and many of these show up in language systems. For example, in this volume, the Tenejapan Tzeltal system is transparently based on mountain slope, and the Jaminjung system on river drainage. More abstract systems, as exemplified by Arrernte in this volume, are probably based on a fusion of different cues, e.g. solar compass and prevailing winds. What is essential about such systems, if they are to function in everyday communication on a range of scales, is that speakers have internalized the fixed directions so that, for example, in an unfamiliar building in the dark, they still know where the named directions lie.

A major dimension of variation concerns the selection from this inventory of three main types of frames of reference. Although some languages use all three, most languages make do with two frames of reference in everyday communication – in particular, many use either the relative or the absolute frame but not both. The intrinsic frame of reference is nearly always present, at least in some residual form. Where more than one frame of reference is available, each may have restrictions on its use – for example in Tenejapan Tzeltal, once objects are substantially separated in space, the intrinsic frame is dropped in favour of the absolute one. Scale may also be a relevant factor, so that objects on a table top may be described in a different frame from houses in a village. Where all three frames of reference are available, one can expect scale differences to play a role in which frame is normally used in which circumstances (although the restriction of the absolute frame to large-scale space is perhaps a European association).

In summary, then, frame-of-reference coding in language can vary on many dimensions. Although there appear to be only three available frames, a language may draw on only one or two of them, each of them can be constructed in quite different ways, and usage of them may be combined and constrained in restricted ways.

1.6 Conclusion

We hope in this introduction to have given the reader sufficient background to read the individual chapters within a comparative perspective. In the conclusions to this book, we provide a detailed summary of some of the major patterns of variation exemplified in the twelve languages for which detailed chapter-length sketches are given. Because contrastive cases are compared in the conclusions, readers may like to use the conclusions as a road-map to help them navigate

the chapters. In that case, readers may like to go straight to the conclusions, get an idea of the variations in the specific spatial sub-domains, and then go back to the chapters, or, alternatively, they may prefer to read the chapters for their own conclusions. Either way, we guarantee that no reader of this volume will come away without a much deeper appreciation of the richness and surprising variation of this important semantic domain.

Language, culture and cognition

Editor:
Stephen C. Levinson
Max Planck Institute for Psycholinguistics, Nijmegen

This 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 behaviour in social context. Books come principally, though not exclusively, from research associated with the Max Planck Institute for Psycholinguistics in Nijmegen, and in particular the Language and Cognition Group.

Books in this series:
1. Jan Nuyts and Eric Pederson (eds.) *Language and Conceptualization*
2. David McNeill (ed.) *Language and Gesture*
3. Melissa Bowerman and Stephen C. Levinson (eds.) *Language Acquisition and Conceptual Development*
4. Gunter Senft (ed.) *Systems of Nominal Classification*
5. Stephen C. Levinson *Space in Language and Cognition*
6. Stephen C. Levinson and David P. Wilkins (eds.) *Grammars of Space*

Grammars of Space

Explorations in Cognitive Diversity

Editors:

Stephen C. Levinson

Max Planck Institute for Psycholinguistics

David P. Wilkins

Center for Aphasia and Related Disorders, VANCHCS, Martinez, California

CAMBRIDGE UNIVERSITY PRESS 2006