14 Patterns in the data: towards a semantic typology of spatial description

Stephen C. Levinson and David P. Wilkins

The chapters in this book present a kaleidoscopic impression of the range of variation in the linguistic treatment of the spatial domain. Each chapter presents a wealth of linguistic detail – what makes the overall exercise special is that, because each description uses the same elicitation devices, we can relatively easily set up fairly precise semantic comparisons in an unusual way. Naturally, the reader may not immediately be able to see the wood for the trees, and many detailed questions naturally arise. For example, one language, like Tiriyó, has a wealth of spatial adpositions, another, like Tzeltal has only one (and that not restricted to space). In contrast, Tiriyó has only one locative verb available for its basic locative construction, while Tzeltal has an extraordinary richness in spatial predicates available for its basic locative construction. Is there a systematic pay-off here? In addition, general questions also come to the fore: behind this variable expression, can we discern universal semantic parameters that might be attributed to general properties of human cognition?

In this final chapter, we have two goals. First, and foremost, we will try to draw the lessons learned from the careful comparative study reflected in these chapters – what are the main patterns that inductively emerge? Second, we will attempt to draw out the implications of this emerging typology of variation for the disciplines that have an important stake in the nature of human spatial cognition.

14.1 Universals and particulars: variation and its limits in semantic typology

As we review the languages surveyed, and draw out the general patterns, one impression that will remain is the extraordinary diversity in both the underlying conceptualizations of spatial distinctions and the manner in which they are coded in specific languages. Direct generalizations are not to be found on a superficial level. Rather, what we will find is that the cross-linguistic patterns can only be extracted on the basis of in-depth study of a reasonable sample of languages. These patterns are sometimes quite abstract – they may, for example, take the form of an underlying hierarchy, which may determine splits in the

coding of different kinds of spatial scenes, but will not predict either the type of coding itself nor, for any one language, where the splits will occur. Another kind of abstract pattern that will emerge is that in any one spatial sub-domain there are a limited set of semantic types – that is, a finite set of conceptual construals of the sub-domain, from which any one language will draw one or more types. Again, these types are only to be found on quite an abstract level, presupposing a real depth of semantic analysis, often with special techniques.

This picture is not in general different from the picture emerging from general linguistic typology, which is mostly dedicated to patterns of morphosyntactic coding. In empirical linguistic typology, simple universal generalizations are not to be found – instead, what we are offered are similar kinds of underlying hierarchies, together with limited series of types. Much of the interest of this work is located in implicational generalizations over types, where for any one language the possession of one type tends to imply the possession of another. What is interesting about the work represented in the current book is that it reveals the same kind of oblique and abstract patterning – underlying hierarchies, types and implicational scales over types. But this work, unlike most typology, is driven by semantic concerns. Indeed, we could say that this book represents the first extended essay in *semantic typology* – and the interest then is that we do not find in semantic typology any simpler pattern than we find in *syntactic typology*.[1] Further, there is no one-to-one correspondence between the semantic types and the syntactic types – we cannot predict the semantic patterns from the syntax, or the syntactic patterns from the semantic patterns.

For many in the cognitive sciences, this will be surprising. The predominant view is that cognitive universals provide a rich, innate representation of the world, which is mapped into the variable surface formats of languages: 'Knowing a language, then, is knowing how to translate mentalese into strings of words and vice versa. People without a language would still have mentalese, and babies and non-human animals presumably have simpler dialects' (Pinker 1994: 82). On this view, semantic analysis of different languages should reveal a single, universal conceptual representation in any domain – and especially in a domain like spatial cognition, essential to the survival of the organism: 'These linguistic categories and structures are more-or-less straightforward mappings from a preexisting conceptual space, programmed into our biological nature . . . This perspective would begin to account for the fact that the grammars and

[1] Traditional morphosyntactic typology and the new semantic typology here proposed have a close relation, but distinct goals and methods. Traditional typology makes recourse to semantic equivalence, where necessary, in order to establish formal equivalence – it uses meaning equivalence to explore formal patterning. Semantic typology does the converse: it uses formal distinctions as clues to the underlying structure of semantic fields, but its goals of course are to explore similarities and differences in semantical concepts. To achieve these goals, new methods are required, as exemplified in this book. Morphosyntactic typology can be done in the library, semantic typology has to be done in the field.

lexicons of all languages are broadly similar' (Li and Gleitman 2000). The idea of a semantic typology then hardly arises – there should be just the one type![2]

But as readers of this volume will have concluded for themselves, nothing like this degree of uniformity of spatial conceptualization is revealed by comparative semantics. Instead, we do indeed need to construct a serious semantic typology, and search for underlying patterns and uniformities on a quite abstract level. That is the message of this book. It does not follow that we cannot talk of semantic universals in the spatial domain – but these are constraints on the way in which a language builds its own conceptualization of a domain. They are, if one likes, more like building regulations than like blueprints. In what follows we will try to draw out these abstract constraints, offering a systematic vision of this new field of semantic typology. We will proceed by taking the three sub-domains that have been the organizing themes for this book – topology, motion, frames of reference – contrasting the patterns exemplified in the languages described in the book, and offering tentative generalizations about underlying patterns.

14.2 Topology

14.2.1 The BLC and a topological similarity space

Recollect that the topological sub-domain concerns the description of situations where figure and ground are in contiguity or close proximity, and angular discriminations are thus not relevant or required. In the introduction we introduced the notion of a *basic locative function* (answers to Where-questions), which allows the identification of a *basic locative construction* (BLC). Further, we introduced a hierarchy of scenes, *the BLC Hierarchy* (Figure 1.2 in Chapter 1, repeated here as 14.1), which is an emergent pattern from earlier work, based on the treatment of over fifty different scenes in a sample of eleven languages, about half of which overlap with the current sample (see Kita and Dickey 1998, Chapter 7). The essential finding was that, in any one language, the BLC may have restricted application over certain scenes; other scenes will be described using contrasting constructions of various kinds. We can use these formal distinctions in the linguistic treatment of spatial scenes as clues to the underlying structure of the semantic field. If one puts these individual language patterns together, one obtains an implicational scale: any language that uses the BLC for scene *i* will also use it for *j*, where *i* is higher in the scale than *j*. Not surprisingly, perhaps, at the core or bottom of the scale is a class

[2] The extreme versions of this doctrine are associated with those, like Fodor, who do not believe in semantic decomposition (see Levinson 1997b and 2003 for assessment of different positions here). Part of the motivation for holding onto semantic uniformitarianism is the belief, expressed by Pinker, Gleitman and others, that without it language learning would be impossible. But, as shown in this volume, semantic diversity is a fact, and the theories need to adjust to the reality (see Bowerman and Levinson 2001 for a range of opinion here).

of scenes of the kind: a relatively small, manipulable, inanimate, movable and independent figure object is in close contiguity with a relatively large, relatively stationary (fixed or immobile) ground object – for example, a cup on a table, an apple in a bowl. As these features are varied, so there is an increasing probability that some construction other than the BLC will be employed. For example, the figure may be attached to the ground to an increasing degree – consider a stamp glued onto an envelope, vs. a handle on a pan (or other part-whole relationships). The tighter the attachment, or the more the figure is an intrinsic part of the ground, the less likely we are to find the employment of the BLC. Another dimension of variation is contact. As contact is diminished, and there is increasing space between figure and ground, again the more likely we are to find the BLC avoided for another construction. This hierarchy has proved reasonably robust over a larger sample of languages, nevertheless some problems with it have emerged (some reflected in the details of this book, which do not support, for example, the position of clothing and adornment). It now seems more revealing to view the BLC Hierarchy as an emergent generalization over a complex, multidimensional semantic space. There are quite clearly a number of factors that make a particular figure–ground constellation a good candidate for BLC treatment – on the following dimensions, properties to the left favour a straightforward locative treatment:

1. Close contact ◄——► Separation
2. Independent figure ◄——► Attached figure ◄——► Part-whole configuration
3. Contained figure ◄——► Contained ground
4. Inanimate figure or ground ◄——► Animate figure or ground
5. Relatively small figure compared to ground ◄——► Relatively large figure
6. Stereotypical relation between figure and ground ◄——► unusual, atypical relation
7. Canonical figure (three-dimensional physical object) ◄——► two- or one-dimensional ◄——► negative space (or hole)

A situation like cup on table (Picture 1 in Figure 1.2, Chapter 1) then comes out high on BLC-inducing features, but a situation like an apple on a skewer (Picture 70 in Figure 1.2) scores relatively low on such features (the figure is relatively large compared to the ground, figure and ground are attached, and the relation is not quite stereotypical). In contrast, a stamp stuck on a letter (Picture 3 in Figure 1.2) is relatively more likely to get BLC treatment, because it has only one undesirable feature, namely attachment, and this at least is stereotypical. Thus there can be 'bad' figures (too large, or negative figures like holes or cracks), 'bad' grounds (too small, too animate) and 'bad' relations between figure and ground (too separated, too attached, too non-stereotypical, or preferred relations between figure and ground reversed). One can imagine a feature-optimizing account along Optimality Theory lines, which will increase

Likelihood of other constructions

1. Figure is impaled by Ground
2. Figure is stuck to Ground
3. Figure is 'damage' or negative space (e.g. crack, hole)
4. Figure is part of whole (part of Ground)
5. Figure is adornment or clothing
6. Figure is inanimate, movable entity in contiguity with ground

Greater likelihood of BLC

Figure 14.1 The hierarchy of scenes most likely to get coding in the 'basic locative construction' (BLC)

the chances of coding in the basic locative construction in accord with the optimal collection of features – hence the possibility of abstracting something like a unilinear hierarchy out of a multidimensional space.

This is not a bad account, but we believe that it is still an oversimplification. What is missing is that the BLC occupies a slot in a grammatical as well as a semantic space, and in that grammatical space there are competing constructions. There are thus not only semantic gradations away from good locative scenes, but positive attractors towards other constructions. One specially relevant class of competing constructions are stative resultative constructions, which express the result of an action. Figures that are cultural artefacts used to perform actions with typical results are likely to evoke this competing class of constructions. Hence a skewer through an apple (or a rope tied around a tree, or a ring placed on a finger) is more likely to invoke a resultative construction than a cup on a table, which is more likely to invoke the basic locative construction.

The overall picture then is a multidimensional semantic space, in which scenes can diverge from good locative scenes, and converge with good scenes for competing constructions. We will try to represent a large portion of this space using eight picture stimuli that are distributed across it. The pictures can be arranged in a plane to make simple diagrammatic comparisons between languages possible, as in Figure 14.2. Because the organization of this space is fundamental to topological relations, we will see some of the same spacing of scenes affecting the choice not only between constructions – as between the BLC and rival non-locative constructions – but also between resources or options within the basic locative construction.

Let us now illustrate how the languages in our sample systematically extend their *basic locative constructions* differentially through this space. Figures A1–11 in Appendix 1 represent for each language a mapping of crucial coding-choices onto the same similarity space, but with pictures reduced to verbal descriptions for diagrammatic clarity (in Figure A12 we add a similar diagram for English, for comparative purposes). Take first Arrernte (see Figure A4),

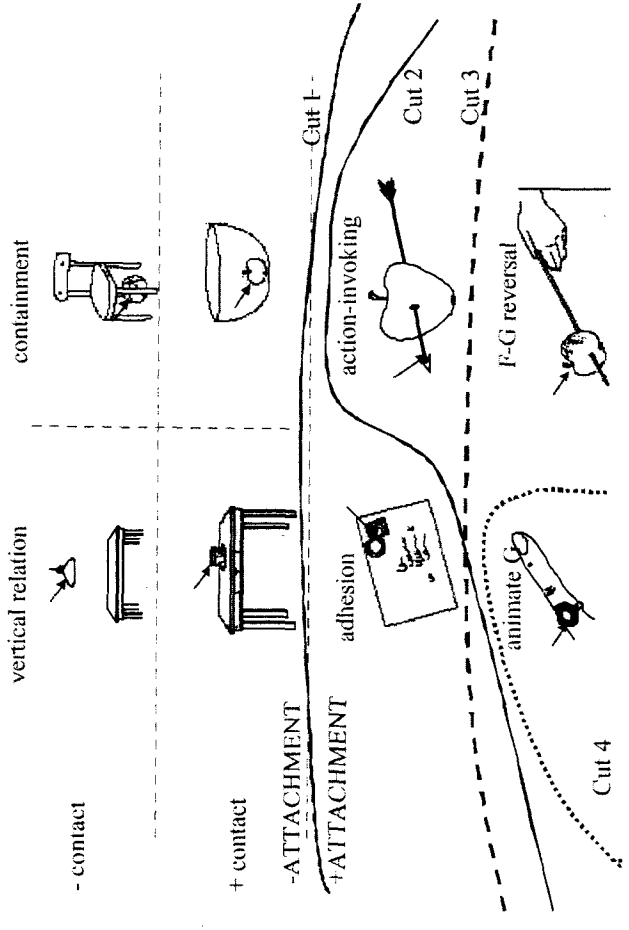

Figure 14.2 A similarity space for topological notions, with the relative positions of the Topology Series pictures (numbers 1, 2, 3, 10, 13, 16, 30, 70), with 'cuts' made by different languages according to whether the BLC is or is not used for portions of the space

where the BLC has a locative case on the ground phrase and an (optional) positional in the predicate. The BLC is not extended to any of the attachment scenes – that is, none of the pictures under Cut (1) in Figure 14.2. All the attachment scenes require a different construction, which treats them as the end result of a prior action (glossing, perhaps, as 'figure has been VERBED'). Note that Kilivila in Figure A3 shows roughly the same pattern. The BLC is here a stative construction with a locative preposition expressing the ground relation, and a positional verb sensitive to the shape and orientation of the figure. Again, this cannot be extended to the attachment scenes, which must have an action description (of the sort 'someone speared the apple'). Since the alternative construction to the BLC in these two languages assimilates all these attachment scenes to an action or its results, we are here seeing these scenes – each of which can be construed as the result of a deliberate action – being attracted to rival action or resultative constructions.

But note that Yukatek in Figure A5 crosses over this division to include the stamp-on-letter case within its stative, basic locative construction (with predicate *yáan* and prepositional ground phrase). It extends the BLC to Cut (2) in the diagram above (Figure 14.2), assigning the rest of the scenes to a resultative construction. Note that although a stamp stuck on a letter can be construed as the result of an action, it is also a normal place for the location of a stamp, unlike the skewer through the apple – stereotypical locations are more likely to be accommodated within the BLC, as we have seen. Tzeltal in Figure A6 goes a bit further, and encompasses the skewer through the apple within its BLC, which is formed with a rich inventory of stative positional predicates. Because of this rich inventory, Tzeltal can directly code within its BLC a great number of complex scenes which depict the result of actions. Still, Tzeltal does not code the cases with odd figure–ground relations within its BLC, that is the items beneath Cut (3) in Figure 14.2 – the one involving an animate ground, and the other a figure–ground reversal.

Japanese in Figure A2 is in some ways just like Arrernte and Kilivila but also illustrates the more complex patterns just reviewed. In Japanese the basic locative construction involves a stative locative verb with postpositions marking the ground. The top half of our similarity space is covered by this construction, just as in Arrernte or Kilivila. But four of the pictures lie outside the scope of this construction – that is, native speakers prefer to use other kinds of construction for these scenes. A 'middle' (quasi-passive) construction is used for the stamp-on-postcard (glossing say, 'The stamp is stuck on the postcard'), and a resultative construction for the skewer through the apple and the apple spiked on a skewer (i.e. for both these scenes which strongly invoke action schema). In a not uncommon pattern, one scene escapes easy description altogether – namely the ring on finger, where a reversal of figure and ground is necessary to express the relation. Because this scene is described with a construction most removed from the BLC, we have marked it out with 'Cut (4)' in Figure 14.2.

(This is because animates – especially humans – are too prominent in many languages to form good grounds: in such languages one would rather say 'The man is wearing a hat' than 'The hat is on the man'.)

All the rest of the languages – Dutch, Yélî Dnye, Jaminjung, Tiriyó, Warrwa, Ewe and English of course – allow the use of their BLCs over the entire range of scenes.

In this sort of way, then, we can generate a clear hierarchy of scenes, such that if a scene to the left is described with a language's BLC, all scenes to the right will be too:

Implicational hierarchy across topological space:

Animate-Ground > *Figure-Pierced* > *Ground-Pierced* > *Adhesion* > *Core-Scenes*

ring on finger > apple on skewer > arrow in apple > stamp > cup on table, fruit in bowl, lamp over table, ball under chair

The interest to semantic typology of these discriminations made by constructional alternates is that (a) these cuts preserve the similarity space (not isolating opposite corners as it were), and (b) they indicate a core set of topological relations with increasing divergences from the core in various directions.

The actual hierarchy generated here would need to be tested against a much larger sample of languages and stimuli, but the procedure is clear. Despite the limitations of the current sample, the reader will find that many of these same patterns recur when we take into account distinctions marked by differential use of resources within the basic locative construction. Note that because the space is multidimensional, we can expect some 'ties' or branching structure to emerge in the hierarchy.

14.2.2 The similarity space and contrasts within the BLC constructional alternates and adpositions

Let us now turn to the use of constructional alternates within the BLC. A good example is provided by Arrernte, as illustrated in Figure A4. As noted, the top four scenes (ON-, IN-, OVER- and UNDER-scenes) fall within the Arrernte BLC. But, as the diagram in Figure 14.2 makes clear, this space is differentiated by '+/– Contact between figure and ground'. In Arrernte the +contact scenes are indicated by the 'part-whole construction', where a spatial nominal in apposition to a ground nominal is marked with the locative case. In contrast, lack of contact is marked by the 'relative location' construction, in which the ground takes an ablative case and the spatial nominal a locative (for ON vs. OVER, the distinction is roughly 'table superadjacent-AT' vs. 'table-FROM

superadjacent-AT', i.e. the figure is in the vertical space away from the table). This example reminds us of the importance of constructional meaning in the spatial domain. Arrernte and many languages also offer other constructional alternates – for example, Yélî Dnye allows the abbreviation or truncation of the full BLC with systematic meaning difference. The full form, with postposition and positional verb, is the normal BLC; dropping the postposition is possible just in case the speaker intends to convey that the spatial array has stereotypical dispositions. This kind of alternation can be treated as pragmatic, since it conforms to Gricean or iconic considerations (Levinson 2000).

Another interesting perspective on this semantic space can be had by considering how distinctions are made within it by lexical choices within the BLC. First, taking adpositions alone, note how the space is differentially fractionated in our language sample (this can be easily seen by glancing over the figures in Appendix 1, where the range of adpositions or their functional equivalents is indicated in Venn-like diagrams). Some languages of course have no adpositions, like Jaminjung or Arrernte. Tzeltal has just one preposition *ta* which covers all the six scenes covered by the BLC; Yukatek uses just two, one general preposition *ti'* and another *ich(il)* reserved for scenes with proper containment. Japanese uses three, conflating OVER and ON, but distinguishing IN and UNDER; Kilivila uses four prepositions distinguishing OVER, ON, IN and UNDER. English most naturally uses five prepositions to cover the scenes – *over*, *under*, *on*, *in and through*. Tiriyó, which includes all eight scenes within its BLC, offers as few as six adpositions, or as many as seven – this is because, as described in the paper, Tiriyó exhibits an unusual hierarchical structure amongst its adpositions. Dutch requires seven prepositions to cover the eight scenes – conflating only cup on table and stamp on letter under *op*, coding surface contact. Yélî Dnye also requires seven postpositions, but it conflates only stamp-on-letter with ring-on-finger, under *p:uu*, coding adhesion.

14.2.3 The cup-on-table scene: adpositional *conflations in the similarity space*

This fractionated picture of overlapping contrasts is not what the literature on spatial language might lead one to expect. Johnston and Slobin (1979), for example, on the basis of acquisition of a wide range of European languages (including Turkish), conclude that IN and ON concepts, that is vertical support vs. proper containment, are universally available and are amongst the earliest learnt by children. But such notions are mostly linguistically conflated with other notions. Consider, for example, the treatment of the **cup-on-table** scene across our sample. Tzeltal conflates under the same adposition the cup-on-table scene with five other scenes out of the eight, and Yukatek covers three other scenes. Note that English too assimilates three other scenes from our eight to *on*

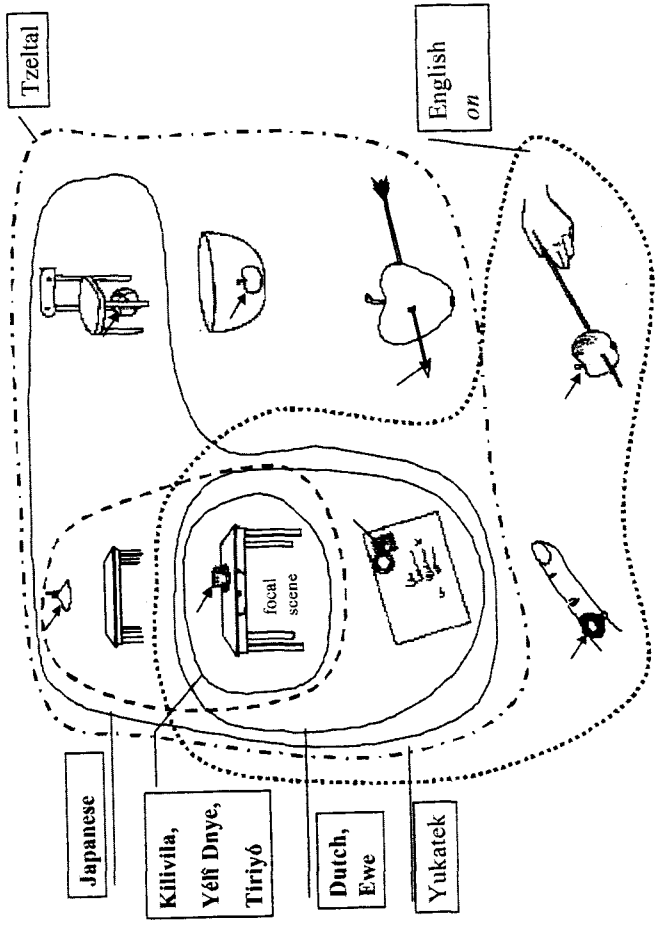

Figure 14.3 The extensional range of adpositions which include the cup-on-table scene

relations – a stamp is also *on* a letter, a ring is *on* a finger, an apple *on* a skewer. Dutch and Ewe are more discriminating but still conflate non-vertical support scenes like stamp-on-letter to the canonical vertical support scene cup-on-table. Japanese conflates OVER and ON. In fact, only Kilivila, Yélî Dnye and Tiriyó isolate the horizontal support scene in their adpositions. Figure 14.3 succinctly shows how variable these inclusions of other scenes are in the adpositional semantics of other languages – the scenes are the same, in the same arrangement, as in Figure 14.2. The diagram, following the conventions in Bowerman (1996), should be read as follows: each Venn-circle represents an extensional conflation under one adposition in one or more languages named alongside. One thing that is immediately obvious looking at this diagram is that we clearly have a good approximation to a similarity space – conflations of scenes under one adposition always suck in neighbouring scenes, not isolated distant ones in the semantic space. Another point to note is that these inclusion relations already suggest a componential analysis of the underlying space. A putative analysis might go as follows, taking the most semantically general adpositions first, and incrementing semantic features as we need them to restrict the denotations. The Tzeltal general preposition includes all scenes except those that invoke action schema with special figure–ground relations – call this 'static location'. Yukatek follows the same pattern, except that it also has a preposition with more semantic content covering proper containment (so with components 'static location' and 'proper containment'), which then probably pragmatically pre-empts the more general preposition just for IN-scenes. English has the next most general adposition *on*, adding a 'contact' requirement between figure and ground – the extensional range is further restricted by pre-emptive *in* and *through*. Japanese relaxes the contact condition but adds a requirement of figure being 'vertically positioned above ground', thus subsuming just the ON and OVER scenes. In contrast, Dutch and Ewe add a requirement of something like 'surface-to-surface contact', covering both scenes of immediate superposition without attachment (cup-on-table) and those with attachment (stamp-on-letter). Finally, Kilivila, Tiriyó and Yélî isolate the ON-scene by adding a condition of vertical support without attachment. Obviously, one cannot move mechanically from an extensional analysis of the kind displayed in the figure to an intensional analysis of the meanings of the terms – but as we have said, the extensional analysis is already very suggestive and essentially constrains such an intensional analysis. But those in search of universal concepts will see that it is only at this underlying componential level that we are likely to find them – there is no linguistic evidence that ON is a universal concept (otherwise it would show up everywhere), but there are certainly strong suggestions that concepts like 'contact', 'vertical relation' and 'horizontal support' are better candidates for such a status.

As mentioned, not all languages have adpositions, and languages that do often have additional means to mark spatial distinctions. If we take these additional factors into account and see how they treat the cup-on-table scene, we will obtain, of course, additional (perhaps redundant) patterns of conflation in this semantic space. Arrernte, for example, uses spatial nominals (of the kind 'superadjacent') with case to make such discriminations – in this domain, taking cup-on-table as the focal scene, it patterns like Japanese in conflating the scenes cup-on-table with lamp-over-table. Jaminjung, which has a very restricted verb inventory, uses two main verbs, glossing 'be' and 'have' in the BLC, supplemented by coverbs, which make fine spatial discriminations. Here we find the reverse pattern from that in Japanese: UNDER and IN are conflated, but OVER and ON are distinguished. Despite the fact that the distinctions are sometimes being made in different parts of speech, the Australian languages reveal some interesting commonalties in semantic patterning. For example all three languages in the sample (Warrwa, Arrernte, Jaminjung) conflate IN and UNDER in a pattern that has been shown to have a wide areal diffusion (see Evans and Wilkins 2000) – Arrernte and Warrwa make the conflation in spatial nominals, Jaminjung in coverbs. This illustrates another utility of the comparative method we are here exemplifying – many areal features are semantic and cannot be extracted by looking at linguistic forms alone.

14.2.4 *The cup-on-table scene: conflations and distinctions by* locative predicates *within the similarity space*

This brings us to the discrimination patterns encoded in locative predicates. This is an almost totally neglected subject in linguistics (but see Ameka and Levinson in preparation) – looking at the handbooks would give the impression that predicates never play any essential semantic role in spatial description. But the facts are contrary – in some languages locative predicates are in opposition and carry much semantic load. Where relevant we have marked these verbal discriminations in the chart for each language in Appendix 1. Take Warrwa, for example: within the basic locative construction it opposes a BE (*-nga-*), a HAVE (*-ba-*) and a PIERCE verb (*-ra-*) – the BE verb extension has exactly the same coverage as the Yukatek basic locative construction as a whole (that is it takes in all the topmost scenes in Figure 14.2 down to Cut 2); the HAVE verb isolates the animate ground scene (ring-on-figure), and a 'pierce' verb absorbs the two skewer scenes. Thus, within the BLC of one language (Warrwa), we are getting the same kind of discriminations that other languages may make by varying the construction employed.[3] If we take the rather different kinds of distinction within the locative predicate made in Dutch, which has contrasting positional

[3] Jaminjung shows a fainter but similar pattern: a HAVE verb is optional and preferred just in the ring-on-finger scene, and a PIERCE verb can, but need not, be used for the skewer through the apple.

verbs, again we find some of the same basic cuts across the similarity space that we found when considering the mapping of the basic locative construction vs. alternative constructions in other languages. For example, Dutch *zitten* 'sit' marks off just the same set of scenes that lie outside the BLC in Arrernte and Kilivila (as can be seen in Figure A7).

There are two main types of system of contrasting locative predicates – one type is based on a small class of three to five contrasting 'posture' or positional verbs (e.g. drawn from verbs glossing 'sit', 'stand', 'lie', 'hang'), a type exemplified in this volume by Dutch, Yélî Dnye, Arrernte and Kilivila. Another type has a much larger set, ranging from a dozen to a hundred contrasting spatially descriptive predicates. Tzeltal exemplifies this type in this book. The predicates in the first, small-set, type act as a kind of nominal or sortal classification of the figure, while the locative predicates of the large-set type more precisely describe properties of the figure–ground relationship. The small-set type usually supplements other means of making spatial discriminations in adpositions (as in Yélî Dnye), or in case-marked spatial nominals (as in Arrernte), so that the interaction between the ground-marking system and the predicate-marking system yields a cross-cutting, fine-grained classification of spatial scenes – see the charts for Dutch (A7) and Yélî Dnye (A8) in particular (in other cases, as in Kilivila, the positional predicates may yield additional conflations rather than additional distinctions).

If we again focus just on the canonical ON-scene (cup-on-table), for languages with such contrastive spatial predicates, we can overlay the predicates which include the focal scene and see what patterns of conflation emerge as in Figure 14.4. Note that, once more, with the possible exception of the Yélî Dnye pattern, the similarity space holds up well – the five other languages at least conflate only strictly contiguous scenes. Dutch and Yélî Dnye are both languages with rich adpositions and with small-class positional verb systems that are sensitive to shape and orientation – Dutch *staan* 'stand' isolates our focal scene, since it collocates with a figure object oriented on its base, while Yélî Dnye *kwo* 'stand' collocates both with long or vertical axes (cup and arrow) and with projections from the ground (in the ring-on-finger picture, the ring is shown with a large projecting jewel). Tzeltal, on the other hand, exhibits a different pattern: it is a language which makes no spatial discriminations in adpositions (having as we have seen just one general one, see Figure 14.2 above) but can distinguish each of the six scenes it admits to its basic locative construction by selection from one of about a hundred verb roots, which code very precise spatial configurations. Nevertheless our focal scene, the cup-on-table, can be conflated with the fruit-in-bowl scene. This is because the predicate *pachal* collocates with *either* a figure *or* a ground which is bowl-shaped, an example of the way in which locative predicates of this large-class type are highly sensitive to overall figure–ground configurations.

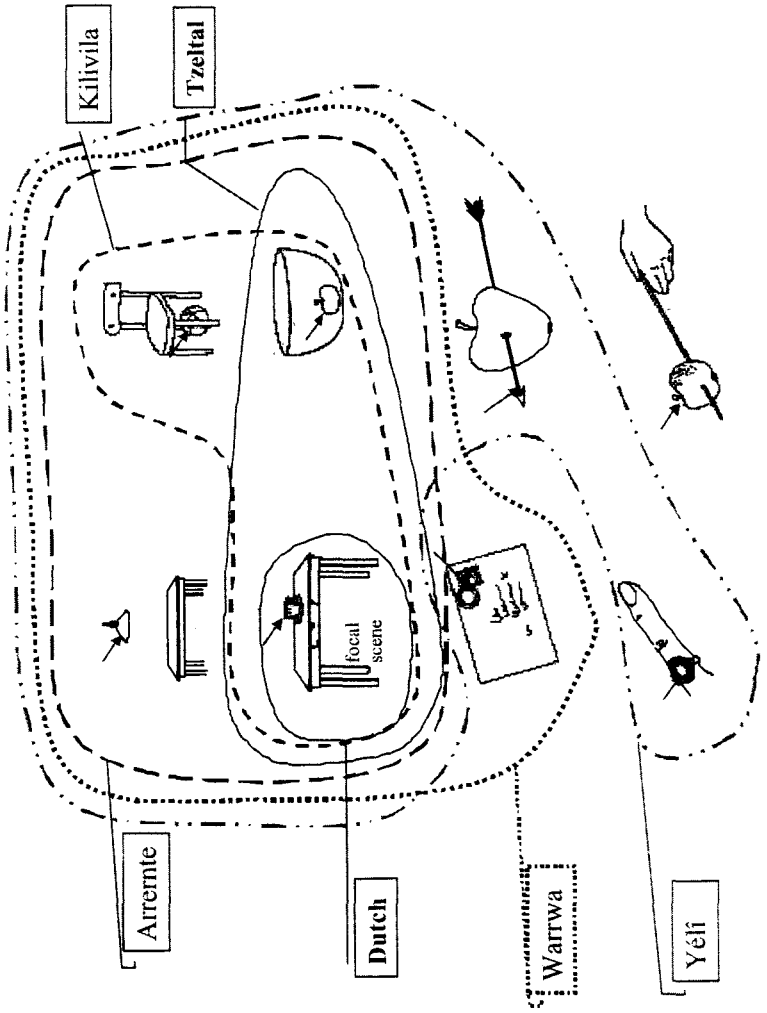

Figure 14.4 Extensional range of contrastive spatial predicates which include the canonical ON-scene

Let us sum up, by way of conclusions to this section. The topological sub-domain is usually judged to be the most fundamental part of spatial language and cognition, for this is the area earliest mastered by children learning European languages at least (see Brown and Levinson 2000 for a constrasting pattern in Tzeltal language acquisition). Despite this, what this section has shown is that there are no simple, surface universals in this sub-domain – for example no universal coding of prototype ON or IN categories of spatial relationship. Nevertheless, controlled comparison reveals that the sub-domain is orderly, constituting a structured similarity space. This similarity space is therefore repeatedly validated or respected by quite different kinds of linguistic distinction – distinctions between the basic locative construction on the one hand and rival constructions on the other, distinctions between constructional resources within the basic locative construction, distinctions within adpositional systems and distinctions made by contrastive locative predicates. Underlying this similarity space seem to be intensional notions like the following, ordered roughly from most general to most specific: stasis, contact, containment, vertical positioning, surface-to-surface contact, adhesion, horizontal supporting surface. These are the notions that are candidate universals in this area, not ON or IN concepts, which are themselves compositionally constructed (see Levinson and Meira 2003 for a more systematic demonstration).

Finally, we have clearly seen that this semantic information is variably packaged across languages, and distributed right across the clause of the BLC. The information is essentially information about the figure (especially shape and orientation), information about the ground (especially shape and medium) and information about the precise nature of the spatial relation between the two. How widely this can be distributed in different constructional, morphological and lexical forms throughout the clause is evidenced in each of the chapters – consider again the Arrernte sentence 'the cup is on top of the table' in Fig. 14.5.

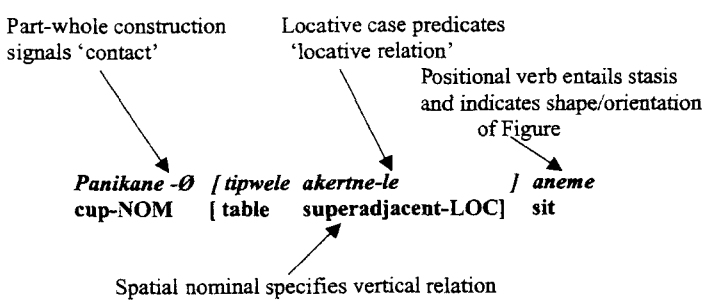

Figure 14.5 Distribution of topological information in an Arrernte clause

14.3 Motion

As we outlined in Chapter 1, there are a number of detailed topics that arise under the rubric of motion description viewed in cross-linguistic perspective:

(a) the typology of semantic packaging in the verb;
(b) the underlying semantical notions of path and motion itself;
(c) the form classes in which such concepts are coded, both verb subclasses and other form classes;
(d) the way in which source and goal are coded;
(e) the way in which all these resources are globally deployed in the clause or beyond to construct an overall depiction of a 'journey' or complex motion path.

Let us take up first the issue of semantic packaging within the verb, within the typological scheme suggested by Talmy (1985, 2000), namely the opposition between satellite-framed and verb-framed languages (see the review in the introductory chapter to this volume). Table 14.1 provides a summary of a number of relevant features of the dozen languages represented in this book. Recollect that Talmy observed that, as a matter of cross-linguistic generalization, 'path' or direction of motion does not conflate with 'manner' into a single verb root (e.g. there should be no verb root meaning 'to go downwards running').[4] This suggests that languages dichotomize into the verb-framed type, where the 'path' or direction of motion is coded in the motion verb, vs. the satellite-framed type, where the 'path' is coded in peripheral, sister constituents to the verb ('satellites'), thus allowing (but not requiring) manner distinctions to conflate into the verbal root. Following the typology, languages are expected to have a 'characteristic tendency' in one direction or the other, thus tolerating a few exceptional verb meanings; further, the presence of manner verbs alone is not diagnostic, only the general conflation of manner and motion, to the exclusion of path.

How well does this typology fare on our sample? A first point to note is that of the languages which can be clearly assigned one way or the other, the great majority are verb-framed – in fact only Dutch, like English, comes out clearly satellite-framed. This suggests that the Germanic satellite-framed pattern may be very restricted typologically. A second point to note is that the typology is in serious trouble with some of the languages – Warrwa, Jaminjung, Kilivila and Yukatek at least – and there are problems of assignment, or other troublesome details (e.g. concerning the coding of manner), in a number of others (Tzeltal, Yélî and Tamil). Take Jaminjung: its small overall set of verbs makes the detection of a 'characteristic pattern' already problematic. But most importantly both path and manner are typically indicated in coverbs – if one treats the coverbs as satellites, then it is unexpected to find manner and

[4] *Climb* might seem to be an exception, meaning 'go up, crawling', but since one can say *climb down*, the path element may be suggested from a meaning more like 'crawl on vertical surface'.

Table 14.1 *Summary of motion coding properties in the sample*

	Verb-framed or satellite-framed	Manner-of-motion resources	Source/Goal ground marking
Arrernte	• Verb-framed • Associated motion inflection also a source of Path encoding	• Very few manner-of-motion verbs • Not commonly coded in clause • If coded, then typically by adverb(ial)	• Case marking: ablative versus either allative or dative [depending on class of motion verb] • Both Source & Goal may occur in a clause
Jaminjung	• Not applicable • Only 7 verbs of locomotion out of a total of approx. 30 verbs in language • 'Fall' is a verb of change of locative relation	• Manner and direction of motion always expressed by coverbs • Path also characteristically coded in coverbs	• Case marking: ablative vs. allative • Both Source & Goal may occur in a clause • Some 'goal' types encoded in coverbs (.e.g. 'enter.water')
Warrwa	• Not applicable • Only 11 simple motion verbs out of an attested set of 61 verbs in language	• Both manner and path characteristically encoded in preverbs in a compound verb construction (CVC)	• Casemarking by postpositional enclitics: two ablatives versus allative or locative for goal • Some 'goal' types encoded in preverbs
Tamil	• Verb-framed • Extensive chaining of converb forms for complex motion events	• Largish class of manner-of-motion verbs • These verbs do not entail translational motion • Special construction of manner verb + go/come to code manner+translation	• Functional equivalent of ablative is complex construction: Noun-LOC +*ire*-Converb Form • Dat marks goal • *viTu* 'leave' as auxiliary codes motion ceases at goal
Kilivila	• Hard to place due to extensive use of serialization, and rich lexicalization of both path and manner	• Large class of manner verbs	• Goal marked by LOC if specific (goal interpretation derived from event expression) • Goal marked by 'to' if general • Verbs often incorporate source and or goal notions (go.down.to.beach)
Yélî Dnye	• Verb-framed (although not easy to identify a 'characteristic' lexicalization form)	• Rich set of motion verbs (often with very specific meanings)	• NPs get the interpretation of source or goal from the verb, not from any marking in NP • Only one reference location (source or goal) overtly expressed per clause

	Verb-framed or satellite-framed	Manner-of-motion resources	Source/Goal ground marking
Tzeltal	• Verb-framed (although there is a large set of directionals which are satellite-like in contributing path info)	• Relatively few manner-of-motion roots • Manners of motion can be indicated by derivational machinery	• One general preposition marking oblique phrases • Only one reference location (source or goal) expressed per clause • Interpretation as 'source' or 'goal' from verb and/or directional and/or context
Yukatek	• Problematic – if forced to assign, then verb-framed, BUT this would force us to redefine the notion of 'path' that is typically used. • 'Path' verbs do not entail durative locomotion along an extended spatial trajectory, but only punctual change of location. (i.e. the characteristic lexicalization form for motion verbs is non-durative, and 'oriented locomotion' is pragmatically inferred)	• Manner of motion primarily lexicalized in active intransitive verbs (only inactive and transitive motion verbs assign source or goal (or 'transit') readings to ground-denoting adjuncts) • By themselves such verbs do not entail change of location	• Grounds of motion events expressed by adverbials • Ground-denoting adjuncts do not reflect the 'path' of the motion event • There is no formal reflex of the 'source'/'goal' distinction • General prepositions, relational nouns, toponyms, etc. only get interpreted as 'source' or 'goal' based on verb semantics • Strict grammatical rule, only one ground-denoting adjunct per clause
Tiriyó	• Verb-framed	• Manner only rarely conflated with motion in the verb • Manner commonly indicated through adverbial, ideophone, or nominalized verbs	• Rich set of goal-indicating postpositions – often with specific meanings (e.g. into water; to the half) • Smaller set of ablative and perlative postpositions
Dutch	• Satellite-framed	• Large set of manner-of-motion verbs	• Source and goal marked by a set of distinct (and path-entailing) prepositions • The set of satellites (which may often occur attached initially to the verb, or after the ground NP) overlaps significantly with the set of prepositions

path both encoded in the same large or open-class category. Besides, as in Warrwa, coverbs play a much more multivaried role than Germanic closed-class verbal particles (like *up* in *go up*). Kilivila raises another kind of problem – it has a rich inventory of verbs encoding both path and manner, and serialization of these verbs makes complex verbs, which encode both path and manner.

Yukatek raises yet another kind of problem, because it challenges the very notion of 'path' as a durative progress through space – we take this issue up below, but the point here is that the element conflated into Yukatek motion verbs may be a different notion from what is conflated in other languages. Yukatek incidentally has good-sized inventories of both path-conflating verbs and manner-conflating verbs, but they are clearly of different classes, which is a possible solution to the constraint against manner and path in the same root. And indeed the Yukatek manner-of-motion verbs can be shown not to encode change of location. Other languages in the sample are problematic for other reasons: Tzeltal, for example, has verbal 'directionals' derived from motion verbs which carry path information (verbal satellites?), but is otherwise verb-framed. Many of the languages have good inventories of both manner-encoding and path-encoding verbs (e.g. Tamil, Yélî Dnye), and in some of these languages at least, the detection of a 'characteristic pattern' of conflation is problematic.

Thus the Talmy typology as it stands, despite having proved useful to understanding subtypes of European languages, does not clearly apply to a worldwide sample, and much more work needs to be done in this area to develop a wider set of types of verbal packaging. We need a better understanding of the underlying components of motion conceptualization, before we can get much further with a typology of how these are differentially conflated in different language types.

Incidentally, manner of motion in verbal semantics (as in 'run', 'crawl', etc.) should properly be distinguished from the conflation of instrument ('ride', 'sail', 'drive') and medium (as in 'float', 'swim'), as Frawley (1992) points out. In our sample there is evidence that languages treat these elements rather differently. Most unexpected, perhaps, are the 'aquatic' postpositions of Tiriyó, marking in, into and out of liquid. The same expression of liquid medium turns up in Jaminjung coverbs. Similarly, Yélî Dnye has a general motion-cum-'inhabit' verb *m:ii*, which means 'move in or inhabit the characteristic medium of the species', thus meaning 'fly' of birds, 'swim' of fish, etc. It also has very specific motion-by-instrument verbs, distinguishing, for example, 'go by sailing a canoe' from 'go by punting a canoe'. Other languages code medium as an adjunct, as in Arrernte 'go in air' or 'go in water' to convey flying and swimming, respectively – such expressions only implicate the characteristic manner of moving through the medium.

Jaminjung, like Tzeltal, has a rich set of posture-encoding elements, but whereas in Tzeltal these are in verb roots usually used in stative form, in Jaminjung they are in coverbs, allowing the composition of complex motion+posture predicates. An additional feature that should properly be distinguished from 'manner' *sensu stricto* is 'speed', which need not entail a change of manner (cf. English *hurry*). Talmy suggested that 'speed' is never grammaticalized as a morphosyntactic feature, but Arrernte offers a counterexample to this with its associated motion suffix indicating 'go speedily/hurriedly'.

Returning to the need for a better understanding of the semantic components involved in motion events, one crucial notion here is the notion of motion itself. As mentioned in Chapter 1, we tend to presume that motion will be conceived of as *translocation*, that is, as a durative displacement of the figure along a continuous trajectory over time. This view entails a certain *Aktionsart* or inherent temporal structure, with predictions about interactions with time adverbials and aspect. But this durative conception of motion does not correspond to the Aktionsart of, for example, the motion verbs like 'enter', 'descend' in Yukatek, which have an inherent punctual change-of-location content (see §14.5 of the Yukatek chapter; see also remarks on 'enter' and 'exit' in the Japanese chapter). These sorts of fact alert us to the cross-linguistic variability of the very concept of motion.

On the basis of the kind of description in the chapters above, we can suggest tentatively that there are perhaps three rather different styles of conceptualization involved in the coding of motion events cross-linguistically, as in Figure 14.6. (Contrasting, incidentally, with this set of notions is another one might call 'internal motion' or 'manner of motion without change of locative state' – this is what is apparently coded in the manner verbs of Yukatek and Tamil.) In this typology, *translocation*, that is a durative event involving passage through an indefinite series of points in space over time, is only one possibility. Motion can instead be thought of solely as a change of state without transitional phases: at time t_1 figure F is in state S_1, at time t_2 F is in state S_2 – what happened in between may be immaterial. The simplest case of this is to think of motion as *change of location*: at time t_1, figure F is at the source S, at time t_2, F is no longer at S; alternatively, at time t_1 F is not yet at goal G, at time t_2 F is now at G. This kind of analysis suggests that motion verbs of this kind should never collocate with both a source and a goal. Bohnemeyer and Stolz argue that Yukatek motion verbs show both these tell-tale symptoms – punctual aspect, and no specification of both source and goal. For verbs with this kind of semantics, how the figure got from source to goal is not relevant – details of the trajectory, the manner of motion, the medium and the instruments involved are out of focus as it were. Languages that code motion semantics in verbs as change of location in this way are thus not likely to fuse manner (as in *crawl*), or medium (as in *swim*) or instrument (as in *drive*) into a genuine motion verb. Again in Yukatek, there are verbs meaning, for example, 'swim' or 'fly', but these do not take a source or goal specification (a location will instead be understood

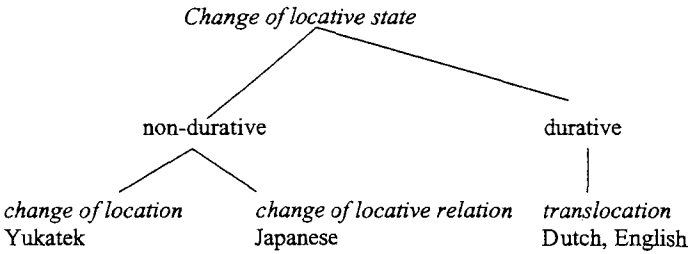

Figure 14.6 Three types of motion conceptualization

as the place within which the swimming or flying activity takes place). The facts in Tzeltal and Yélî Dnye are at least suggestive of a similar analysis.

A further possible subtype of change-of-state semantics for motion verbs would make change of location itself not a necessary part of the semantics. Instead, all that would be required is a *change of locative relation*: at time t_1 figure F is in locative relation R_1, while at time t_2 F is in locative relation R_2. Consider a ball outside a ring; a few moments later it is inside the ring. Although translocation of the ball is perhaps the most obvious way to achieve this change of state, an alternative is to move the ring. Some languages treat these two different translocations as the *identical* motion event – Japanese provides a clear case. Kita devised a motion stimulus to test this situation: paired films showing a circle moving inside a square boundary, and a square boundary moving to enclose a circle, were described in the same way. The same results were found for Tzeltal. More intriguingly, in a general triads test, which pitted identical path of motion against identical end-result of motion, both Japanese and Tzeltal speakers considered the same end-result pairs more similar than the same trajectory pairs, suggesting a general pattern of motion conceptualization. The same linguistic result has now been had from a number of languages using the ENTER–EXIT stimulus film designed by Kita (1999). This shows a man walking into a room from various angles (allowing different deictic codings), and this is contrasted with scenes in which the man dissolves outside the room and reappears inside the room ('beaming in' as it were). Dutch or English speakers will avoid 'going in' or 'entering' locutions, since the actual trajectory is unseen; but Japanese and Tzeltal speakers find 'enter' locutions fine. There is thus clear empirical support for motion verbs in some languages having this third type of semantics, encoding change of locative relation without trajectory.

It turns out that motion verbs in a language are not necessarily of a consistent type, although a predominant type of semantic content may often be discerned. From the details in the chapters a case can perhaps be made for a hierarchy of some kind of across 'motion' verbs, with a change-of-state semantics

more often involved with boundary-crossing verbs, like 'enter', and a translocation semantics involved with basic motion verbs like 'go', roughly perhaps as follows:

likelihood of type of semantic content:
'exit'> 'enter', 'cross' > 'go'

change of state ⟷ *translocation*

It is perhaps for this reason that, counterintuitively, 'exit' verbs also seem frequently to mean 'appear' (as in Tzeltal, and, as on one reading, in Arrernte) – they are especially likely to code change of state without path! The evidence for this hierarchy is that in some languages, like Tzeltal, 'go' clearly encodes durative translocation, but 'enter' and 'exit' code change of state without necessary translocation, while in others, like Arrernte only the 'exit' form encodes change of state without translocation, the 'enter' form having a translocation semantics (indicating that effective doublets, or functional antonyms can in fact have a rather different semantics). The data from Yukatek suggest that a more detailed hierarchy might be sustainable, with an ordering between our three types of motion construal:

	Change-of-locative-relation	**Change-of-location**	**Translocation**
Yukatek	'enter'/ 'exit'	'come'/'go'	
Tzeltal	'enter'/'exit'		'come', 'go'
Arrernte	'exit'		'enter', 'come', 'go'

This brings us to the question of distinct form-classes of motion verbs within languages. Although languages tend to have various minor form classes of verb, what pre-theoretically we may want to call motion verbs are perhaps rarely all of the same class (a point made by Lucy 1994). What one tends to find is that there is a more restricted class, which may include only or predominantly motion verbs (perhaps including some abstract 'path' verbs like perception verbs), and which may have special morphosyntactic properties (e.g. forming auxiliaries or directionals in Mayan languages like Tzeltal). The core class of motion verbs will typically include the deictic motion verbs ('come', 'go', 'return here', etc.) if they are coded lexically – although in many cases (as in Yélî Dnye) these deictic distinctions are made with additional 'hither'/ 'thither' morphemes or the like. One general finding is that such deictic coding is usually one-way: languages typically encode motion towards the deictic centre, but leave the 'away from deictic centre' meaning to pragmatic contrast (see Wilkins and Hill 1995) – as in many languages in the sample including Arrernte and Jaminjung. In any case, this core deictic class of motion verbs may fall into a different verb class from other good motion verbs – thus in Arrernte we need to distinguish three motion classes, with different argument structure: (1) core 'deictic' verbs,

with three argument slots, subject, source and goal, (2) oriented motion verbs like 'fall', with two arguments, and (3) manner verbs with a single nominative argument. In other languages, like Yukatek, verbs that on translation equivalence might be thought to be motion verbs, like the manner verbs, in fact do not encode external change of locative state at all.

Languages with a very small closed-class set of verbs of all types, like Jaminjung (*c.* 30) and Warrwa (*c.* 60), are very instructive here. Jaminjung, for example, has a minor form class of just seven motion verbs, two of which are intransitive (glossing 'come' and 'go'), the rest transitive (glossing 'leave', 'approach', 'follow', 'take' and 'bring'). All of these seem to have a 'translocation' type of semantics. Another minor form-class is centred around locative relations and includes locative 'be', and apparent motion verbs glossing 'fall', 'throw' and 'put'. Schultze-Berndt argues in Chapter 3 that a verb like the 'fall' verb in fact encodes 'enter into a configuration with ground', and the trajectory up or down, and even the fact of motion, are not essential preconditions to its use. In short, verbs of this class belong to our change of state type, and probably the change-of-locative-relation type, like Japanese 'enter'. These languages with small sets of verbs, where the exhaustive partitioning of all the verbs in the language into different subclasses is relatively clear, show that languages may have distinct types of 'motion' verb within their verbal inventories, and that there may be no simple correspondence between languages over how English translational equivalents will partition.

Interestingly, motion can be coded in many languages in form classes other than the verb. A spectacular example of this is provided by the set of fifteen Arrernte 'associated motion' affixes. These cannot be suffixed to core 'deictic' motion verbs, but can be suffixed to other classes of motion verb, and even to non-motion verbs. The suffixes offer a good range of different meanings. For example, where VERB is the meaning of the action specified in the verb, one suffix will mean 'do VERB while coming', another 'do VERB while going downwards', another 'do VERB hurriedly and go back', and so forth. This particular complex is a central Australian areal feature, but other languages in the sample show that motion coding outside of the verb is by no means restricted to that area. For example, Yélî Dnye has a series of oppositions expressed in portmanteaux morphs, which also code tense, mood, aspect, person and number. A full set of variants of these code something like associated motion, in this case 'Go and VERB' (e.g. 'go and see'). We would need more languages before we felt confident about this, but a tentative suggestion is that there is an implicational scale underlying associated motion semantics:

Go then/to VERB > Go while VERBing > VERB then go

– that is, it seems more common to have motion with purpose, or motion preceding the verbal action, than vice versa.

Mayan languages like Tzeltal have directionals, derived from motion verbs, encoding a full range of path oppositions, like 'go', 'come', 'ascend', etc. As particles following non-motion verbs they indicate that the action was done while going, or was done and then the subject went, etc. Although we do not have such examples in the sample of languages here, there are languages where motion coding is carried entirely by marking the ground elements – for example, Kayardild has verbless motion sentences (Evans 1995: 169), where the allative or ablative cases entail motion (a good case can also be made for these entailments in Arrernte, but the verb cannot be dropped).

Let us now turn to how the trajectory itself is coded. Talmy presumed that one can distinguish between *path* – or abstract direction – and *ground* specification; for example, 'John fell down' specifies a path, 'John fell down into the hole' adds a ground. In many languages, English included, this distinction can be unclear (English syntax is often indeterminate between particle and preposition; see Matthews 1991). In fact the best justification for the distinction comes from languages which use the absolute frame of reference, as in 'He went north', where 'north' is not a ground but a pure path or direction. Still, for most motion coding in most languages, source and goal specification play a crucial role in determining a direction of trajectory. Indeed, in some languages we find a nice kind of alternation between different kinds of marking of ground, one form indicating direction, the other goal. For example, in Arrernte, with one verb class (including 'enter', 'fall', etc.) the allative case encodes direction towards something not necessarily reached, while the dative case encodes goal. (Incidentally, we will treat source and goal as the main grounds for motion, and this is because perlative or *via* notions are more rarely directly expressed in languages. This may well in part follow from many languages having non-durative semantics in their motion verbs.)

The coding of source and goal is cross-linguistically very variable, as the reader will have noticed. On the basis of the languages described above, we can make some tentative typological suggestions. Some languages code sources and goals with zero-marking – that is to say, the relevant noun phrase appears without an adposition or case or other marking (Yélî Dnye is of this kind). Others use a general marker, e.g. a vacuous adposition which does not distinguish between source and goal (Tzeltal and Yukatek are of this kind). Clearly, in these cases the coding is effectively in the verb – a bit like the (uncharacteristic) English *John entered the house*, where the verb encodes motion to the interior of a space, and the noun phrase is bare. Yet other languages can be shown to have verbs of this type, where the semantics encodes the source or goal, but nevertheless redundantly encode source and goal on the NPs (Arrernte is a language of this type). Finally, there are languages where the coding of source vs. goal (or other kinds of ground) is very clearly marked only on the NPs (as in English, Warrwa or Dutch). This gives us a typology as in Figure 14.7. This typology

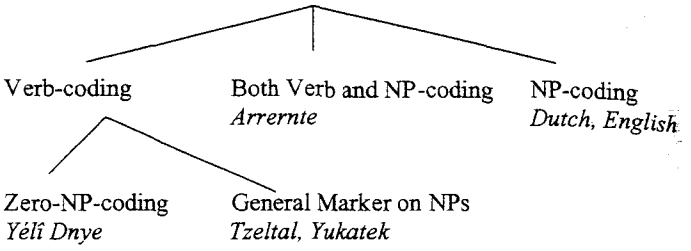

Figure 14.7 Typology of ground-encoding strategies

relates to Talmy's typology of verbal encoding mentioned earlier. First, Talmy (1985) argued that in no language is it characteristic to fuse the ground with the predication in verbal packaging, because the ground phrases are the background to the event description, 'the unvarying component in a situation'. But the languages that code ground specifications in the verb are clear counterexamples to this suggestion.[5] A second point is that verb-framing languages may tend to coincide with ground-encoding in the verb. This is perhaps because Talmy's own distinction between path and ground specification is ultimately a gradation – although path may be encoded in adverbial satellites, path or trajectory is also typically encoded with respect to grounds.

As mentioned in Chapter 1, an important general dimension of variation concerns the extent to which languages use the same resources in the description of motion vs. stasis. Talmy (1985) has suggested that they universally tend to do so, since static locatives are derivative from or modelled on motion descriptions. Thus (as mentioned in Chapter 1) in English *He went out of the office* is very similar in structure to *He is out of the office*, and the prepositions of motion seem to parallel the prepositions of location (cf. *in~into, on~onto*). But the parallels are often more opaque, as in the systematic distinctions in Tiriyó adpositions between location vs. motion. However, the main point is that some languages make very fundamental distinctions between the two domains. Tzeltal, for example, uses quite different resources in the two domains – for stasis it has a rich set of stative predicates indicating precise figure/ground relationships, and two frames of reference. For motion, it has a special subset of *c.* 20 motion verbs and derived directionals and uses only the absolute frame of reference. Even the coding of the absolute frame of reference differs in the two conditions – 'uphill' is coded with prepositional phrases such as 'at its upness' in the static condition, and is coded in motion verb roots like

[5] Talmy no doubt had in mind that the coding of *specific* grounds in verbs is relatively rare, as in *berth* (or *dock* of boats) or *dive*, but Yélî Dnye might be an exception here with verbs such as 'cross-over-sea' vs. 'cross-over-hill', 'go-home-after-feast' and many of the same ilk.

'ascend' in the motion domain. Similarly, Yélî Dnye has a huge inventory of postpositions used in static descriptions, hardly any of which are used in motion descriptions, because source/goal distinctions are built into the verb roots.

The Frog Story in cross-linguistic perspective

Let us now turn to compare the motion stimulus we have selected for cross-linguistic comparison, namely the Frog Story cliff scene. We chose to illustrate differences in motion coding across languages with this stimulus rather than some of the more detailed elicitation devices (like the ENTER–EXIT film mentioned above) because it gives a good impression of how these detailed differences in formal devices and semantic content work to construct a very different overall picture of a motion event (as the Berman and Slobin 1994 volume illustrates at length). Indeed, comparison shows that this uniform stimulus is coded linguistically in strikingly different ways.

Our focal part of the Frog Story, the cliff scene, portrays what Slobin (1996) calls a 'journey', that is a complex motion event with subpaths. Few tellings of the story mention all the subpaths, but this selection varies systematically with language – it is as if the coding of a visual stimulus into a particular linguistic representation renders some aspects of the event invisible, others prominent, and forces the interpolation of some scenes not visually represented at all. If we put the sample stories from our dozen languages together, and focus on the boy alone, we need to recognize a total of fourteen subevents derived by tellers from the four pictures that make up the cliff scene (see the figures in Chapter 1). The table in Appendix 2 shows which of the tellings in the languages in the sample mention each of these fourteen subevents, with an abbreviated 'propositional' rendering of the boy's journey between his standing on the rock, and his sitting in the water.

The simplified version of that table, Table 14.2, shows which of these subevents maximally recognized across languages are reflected in individual tellings in the dozen languages (we have added English for comparison, choosing one characteristic telling from the collection made by Slobin, by kind permission).[6] The point has been made by Slobin (1994, 1996) that the availability of resources for motion description in a language tends to build a distinctive style, typical for that language's narratives. What stands out from the table is

[6] English Adult from Slobin Sample:20F.
And – and what the boy took to be branches were really – antlers of a deer on which he gets caught – the dog – oblivious to all this looks behind the rock. The deer takes off with – the boy strewed across his antlers – and the dog runs at his feet yelling at him – to – to stop it. Um – they're approaching a cliff – and the deer – stops abruptly – which causes the boy to lose his balance and fall with the dog down into the stream – um – or a little puddle. [4/7]

Table 14.2 *Summary table of the 'cliff scene' from the Frog Story: subevents mentioned (+) in the different languages*

Scene	Jamin-jung	Warrwa	Arrernte	Tzeltal	Yukatek	Tamil	Kilivila	Yélî	Dutch	English
1	+	+	+			+	+	+	+	+
2	+		+	+	+	+	+	+	+	+
3		+	+		+	+	+	+	+	
4			+							
5										+
6						+	+			
7	+				+					+
8	+	+	+	+	+	+	+	+	+	+
9	+		+	+	+	+	+		+	+
10				+						
11									+	
12						+	+			
13		+								
14	+	+	+		+		+	+	+	+

that six subevents (numbered in bold in Table 14.2) are mentioned in at least 70 per cent of the languages – these involve the major stages of the trajectory. But some of the subevents are mentioned by just one or two of the languages, and focussing on these is revealing. For example, subevent 4 is the motion of the boy on the deer past the viewer (Picture 16 shows a side-view of the moving deer). This is mentioned only by Arrernte, which codes 'boy lying while moving past' using the associated motion suffix 'do while moving past' attached to the verb 'lie' – it is hard to escape the conclusion that this subevent is mentioned in Arrernte just because there are special grammatical resources that make it both conceptually prominent and easy to code. Another uniquely mentioned segment is 10, which gives us a snapshot of the boy spread-eagled mid-air. In this case, Tzeltal codes this using its rich set of dispositional predicates, one of which precisely encodes 'lying with limbs outstretched face up' – again the resources of the language seem to make this both natural and efficient to mention. Dutch and English also have unique segments. English mentions the approach of the deer towards the cliff (using the verb *approach* which may be an unusual lexicalization). Only Dutch mentions the source (the cliff edge) in the falling scene – in fact this is the only case in our cross-linguistic sample of 'cliff-scene' descriptions where both source and goal of a motion event are mentioned in the same clause:

gooit	*het jongetje*	*van*	*een klein*	*afgrondje*	*het water*	*in*
throws	the boy	from	a small	cliff	the water	into

As we have seen, many languages (like Yélî Dnye) do not permit the simultaneous mention of source and goal in a single clause, which is partly a function of the type of source/goal coding – where this is coded in the verb, usually only source or goal is subsumed (as in 'enter' vs. 'exit', 'arrive' vs. 'leave'). In addition, as we saw, the semantic construal of motion as a non-durative change of state (as in Yukatek) rather than a translocation (as in Dutch) can also restrict the possibility of having both source and goal encoded (that presuming a durative event). Languages like Yukatek or Yélî Dnye, apparently for rather different reasons, forbid simultaneous mention of both source and goal. Other languages, like Arrernte, permit it, but follow what may be a widespread discourse rule (which we might dub the Preferred Ground Structure) 'mention only one major ground, source or goal, at a time'.

Some languages have strikingly similar distributions of the subevents mentioned, Tamil and Kilivila for example. Both are languages that utilize what one may loosely call verb serialization, that is the concatenating of a string of verbs within a single clause. Thus the first subevent in the sequence, the deer taking the boy away, is rendered in Tamil by a sequence of verb roots, roughly 'raise+hold+run+go', and in Kilivila by 'climb-down+take-away+run'. These resources allow highly compact renderings of many subevents in one clause but require that the grounds for most of these subevents are suppressed and must be inferred.

Another perspective on the differential coding of the event across languages can be had by comparing languages that have very different distributions of subevents that are mentioned, for example Tzeltal and Yélî Dnye. Tzeltal concentrates on the middle subevents, while Yélî Dnye focusses on the initial ones. Why? Tzeltal systematically picks out those scenes where there are complex dispositions in the figure, or the interaction between the figure and the ground, and it codes these in its rich set of dispositional predicates. Thus subevent 2, with the boy on the deer, gets coded as 'boy mounted on deer, wedged between (forked branches)', subevent 8, the falling scene, as 'boy be thrown underhand' where 'throw-underhand' is a single verb, subevent 9 where the fall continues as 'fall_down descending' using the special directional resources of the language, subevent 10 where the boy in mid-air is coded, as already mentioned, as 'boy lying face up limbs outstretched', all of which is likewise packed into a single dispositional root. Two of the scenes are described as static dispositions, from which the motion must be pragmatically inferred, and this is typical of Tzeltal which has unusually rich descriptive resources for static locations. Yélî Dnye has the opposite character – it stresses the action scenes, and allows the resulting states to be inferred – the taking of the boy, the throwing of him, and the landing in the creek. The special portmanteaux tense-aspect-mood-person morphemes, which can alternate to show deixis and associated motion, allow very compact sketches of scenes, complete with a 'camera angle' as it were.

Table 14.3 *Differential loci of encoding for the components of motion description*

	Source	Goal	Manner	Medium, instrument speed	Motion	Path/direction
English, Dutch	PP	PP	V	PP	V	Particles
Arrernte	(Abl NP +V)	Dative NP +V	Subclass of Vs;	Loc NP (medium); associated motion (speed)	V or Associated motion suffix	V
Jaminjung, Warrwa		Allative NP	Coverbs	coverb	V or Coverb	Coverbs
Tzeltal	(*ta NP + V*)	*Ta* NP + V	non-motion V; derivation		V or Directional	Directional
Yukatek	(PP)	PP	non-motion V		V	
Tamil	(Abl NP)	Dative NP	V-serialization		V	V?
Kilivila	(PP)	PP or NP	V-serialization		V	V?
Yélî Dnye	(NP+V)	Unmarked NP + V	V subclass	V/ PP	V or Associated motion proclitic	V
Tiriyó	PP	PP	ideophones, adverbs, nominalizations	Adposition	V	?

Yet another perspective arises if one asks where in the clause specific information is coded. We have discussed many of the details above, but a summary table (Table 14.3) may be helpful here (resources not actually used in the sample stories are in brackets).

14.4 Frames of reference

In the introductory chapter to this volume, we introduced a restricted typology of just three types of frame of reference to be found in language, absolute, relative and intrinsic. Frames of reference are coordinate systems whose function it is to designate angles or directions in which a figure can be found with respect to a ground, where the two are separated in space (in contrast where they are contiguous, the topological system comes into play). As we explained, the absolute frame depends on the antecedent fixing by community consensus of arbitrary bearings, the relative frame depends on mapping the bodily coordinates of the viewer onto the scene, and the intrinsic frame relies on designating facets of a ground object.

As mentioned in Chapter 1, languages sometimes use only a subset of these three kinds of frames of reference. The language sample described in this volume nicely illustrates the kinds of variation to be found. A full summary can be found in Appendix 3, which includes notes of any trace of any of the systems, but to obtain a simplified picture, let us extract the major frames of reference that are used in ordinary discourse without recourse to prosthetic devices like maps and compasses. Note that we then set aside fragments of systems which do not offer a full polar system of coordinates, or full systems which may be coded but not used in daily parlance. We then have the distribution of frames of reference across languages represented in Table 14.4: each of these frames of reference may be instantiated in different ways. Take, for example, the absolute frame of reference, which requires consensus in the community about named, fixed directions. Such directions may be fully abstract notions, as in the fixed bearings used in Warrwa and Arrernte. Systems of this kind, which are not based on compass points or any one meteorological or landscape feature, are triumphs of human abstract reasoning. After all, it matters not at all what directions are fixed and named, only that members of a community can consistently find and name them. Such abstract systems are probably anchored on a complex set of cues, involving a solar compass (i.e. the abstraction of an ideal axis orthogonal to an average of solstitial variation on the east–west axis) supplemented with wind directions, dune axes and such like (see Levinson 1997a, 2003). Other systems are more directly linked to ecological cues, but these can be of quite different kinds. Jaminjung in this volume illustrates a type (also common in Alaska) based on the main river-drainage system – this gives us upstream/downstream and across directions. Again, though, this drainage system merely provides a

Table 14.4 *Distribution of main frames of reference across the sample languages*

Absolute, relative and intrinsic:	Tamil, Yukatek, Tiriyó, Ewe, Kilivila(?)
Absolute and intrinsic:	Warrwa, Arrernte, Jaminjung, Yélî Dnye, Tzeltal
Relative and intrinsic:	Japanese, Dutch

basis for a fully abstract system of directions – most of the river beds are dry for much of the year, and of course wander in all directions, but nevertheless the directions are fixed. Another reasonably common kind of system is illustrated by Tzeltal, which draws the inspiration for its abstract directional system from the major geographical tilt of the country that forms the territory of the Tenejapans who speak the dialect described here. The Tenejapans identify an uphill–downhill axis, with an orthogonal across axis, which can be shown to have fixed compass directions even though every valley of course has its own meandering path in this rugged mountain land. Another kind of system is illustrated by Yélî Dnye, which also has an up–down system, in this case based primarily on the prevailing eastern winds – one sails and poles up against the wind. As a further axis, they use a mountainwards vs. seawards axis. These people live on a relatively small island, with a high central mountain range, so as one goes around the island the two axes change their angle with respect to each other, thus demonstrating that there is nothing essential in orthogonals to the human geometric imagination.

Peoples like these who use absolute directions as their main way of designating figure–ground relations for things separated in space (at least in contexts where deictic systems and pointing will not alone do the job) can be shown to have acquired a distinct cognitive style, involving a constant sense of direction and a conceptual coding scheme based on fixed directions in memory and reasoning (Pederson et al. 1998, Levinson 2003).

Let us turn now to the intrinsic frame of reference. This frame of reference is the only one that may be – at least in rudimentary form, with topological antecedents – universal.[7] Developmental psychologists have long noted that it is the first frame of reference to be used systematically in language and reasoning by children (Piaget and Inhelder 1956, Johnston and Slobin 1979, Tanz 1982). One reason is perhaps because the intrinsic system is based on a simple binary relation between figure and ground (unlike the other two frames which involve ternary relations between figure, ground and viewer or fixed bearing). Another

[7] Guugu Yimithirr is a language where the intrinsic system plays arguably little or no systematic role in spatial description, but even here perhaps rudiments of such a system can be found (see Levinson 1997a). See also Warrwa in this volume.

reason for its fundamental nature is that it is closely linked to topology, where the geometry of the ground object is also relevant – knowing the parts of an object is a precondition to using intrinsic systems. Indeed one can think of the relation between topological part systems and intrinsic systems as forming a cline: for example, between the topological concept where a figure is designated as in contiguity with the back part of the ground, and the corresponding intrinsic projective system, where the figure is conceived of as in the region projected from the back part of the ground (Heine 1997: 44–5). Because of the close relation to topology, it is not always easy to decide whether one is dealing with a topological spatial description using part-names, as in (the gloss) 'The ladder is AT the back part of the house', or an intrinsic description projecting a region of part-names, as in 'The ladder is WITHIN the region at the back of the house'. The description of Warrwa in this volume raises issues of this kind and allows one at least to question the universality of the intrinsic frame of reference.

Given the close relation to topology, the binary simplicity of the relations and the early development of intrinsic notions in childhood, one might expect all intrinsic systems to conform to a single, simple pattern. But this is not the case. As mentioned in Chapter 1, there are different ways to assign parts or facets to objects. The English (and Dutch) systems are remarkably complex. At their heart there is a simple, gravitationally oriented armature, giving us 'top', 'bottom', and 'sides' for any object. But then to obtain named sides – essential for projecting regions on the horizontal plane – we need to take into account a wide range of factors: does the object characteristically move (if so, the side going forward is the 'front', as with trucks), do we primarily use one side of it (if so, that is the 'front', as with clocks), do we enter the object from a particular side (as in the 'front' of buildings), do we mould it to our frame (as in the 'front' of a jacket or a chair)? Despite these varied functional factors, children of two can already line up different objects so that their 'fronts' all face in one direction! In contrast, the system described in the Tzeltal chapter has none of these features. The main body-part system does not use a vertically oriented armature – there is nothing universal about 'top', 'bottom' and 'sides'. Instead, for inanimate objects, the whole system is driven by the internal axial structure of the object. Thus a stone lying down with a flat surface on the ground will have its 'face' upside down, with its 'head' and 'butt' determined by the shapes at the end of its longest axis (see Levinson 1994). Neither vertical orientation nor function play a role in part assignment, which can be shown to be almost entirely a matter of internal geometry.

Finally, we turn to consider variants of the relative frame of reference. Fully developed relative systems have clearly evolved out of intrinsic systems, especially to deal with cases where the ground object lacks unique intrinsic sides (as with a ball, tree, or box). Hence the implication: if a language has a relative system, it has an intrinsic one, usually with shared lexemes. Relative systems

involve the speaker's coordinates (his or her own front/back/left/right), but in addition a secondary coordinate system mapped from the speaker's coordinate system onto the ground object. This is the source of considerable complexity. This secondary system may be a rotation, translation or reflection of the primary coordinates, or even a mixture of these. Thus when I say *in front of the tree* in English, I mean 'between me and the tree' – the tree has acquired a front by mapping my coordinates onto the tree under reflection: *front* and *back* are reversed (as if the tree was someone facing me), but *left* and *right* have stayed constant (imagine writing my coordinates *front/back/left/right* on a transparent sheet and turning it over away from me and now overlaying the assignments onto the tree). In actual fact, the correct analysis of English relative usage is arguably even more complex than this (see Levinson 2003: 86–8), but this will do as a first approximation.

But there are other possible ways to assign the secondary coordinates. Instead of reflecting them onto the ground object, one could rotate them onto it (instead of flipping the transparent sheet, we now rotate it 180°). Now we will have 'front' and 'back' as in English, but 'left' and 'right' will be reversed: 'The ball is to the left of the tree' would mean what in English we express as *The ball is to the right of the tree*, because the tree is like a person with its own front (facing us), and its own left (at our right). Such systems have been reported from, for example, some dialects of mainland Tamil. A third possibility is simpler: the speaker's coordinates are simply translated (shifted across without rotation or reflection) to the ground. Now the tree has a 'front' on the far side away from us, and we are looking at its 'back': 'left' and 'right' remain as in English. Such systems have been reported from Hausa and many other languages. Finally, languages can borrow from these different possibilities and assign some terms in one manner, and others in another, or they may even use terms ambiguously. Many languages make no use at all of relative systems (as in Jaminjung in Chapter 3). Quite often languages with intrinsic terms allow relative uses only in a few marginal cases (as in Tzeltal in Chapter 7), or a bit more systematically just where the ground object lacks intrinsic facets (as in Yélî Dnye in Chapter 5). At the other extreme are languages where relative systems are central, as in Japanese and Dutch (Chapters 12 and 13).

Thus, just as in the intrinsic and absolute frames of reference, there are many distinct variants of the relative system, and varying degrees to which terms which have an intrinsic origin may also have gained relative uses (see Levinson 2003: 84–9 for details). It is important then to bear in mind that when we talk about just three frames of reference in language, we are talking at an abstract level about *types* of coordinate system, not about how these are instantiated in particular languages, which can be quite diverse.

Let us now consider the details about how such systems are used. Since in all the languages, with the possible exception of Warrwa, more than one frame

of reference is available, a first issue is in what kind of context which frame of reference is deployed. The best way to get a firm understanding of this is to look at our comparative task, the Men and Tree Game. As described in Chapter 1, the game involves distinguishing *inter alia* between a set of photos with position and orientation of a tree and a man systematically alternated (now the man is to the left of the tree, now to the right, now he's facing us, now he's looking away, etc.) – each of the numbered photographs is reproduced as a line drawing (Figure 1.3) in Chapter 1.

From the details in the chapters, we can make quite a lot of comparative observations. Consider first three pairs of players of the Men and Tree Game in Dutch. Table 14.5 gives the full propositional coding. We have found it essential to distinguish what we call 'standing' (positional) from 'facing' (orientational) information – in these scenes that amounts to making a distinction about the placement of the man versus the tree in the left/right plane (standing information), versus the direction the man is facing (towards the viewer, away from the viewer, to the left, or to the right). If we look at the three pairs of Dutch players, we get a very consistent picture, with a systematic distribution of standing and facing information:

		Standing information	*Facing information*
Dutch	Pair 1	Relative	Intrinsic
	Pair 2	Relative	Intrinsic
	Pair 3	Relative	Intrinsic*

* (One proposition combined relative and intrinsic for both standing and facing information)

Here we find the coding of facing information done in the intrinsic frame of reference. This is probably because, since a part or orientation of the figure has to be described anyway, this invokes the frame of reference for which this is a precondition. But standing information is consistently given in relative terms. Of course, this all seems natural to us – the larger framework invokes the fixing of a point of observation, and a placing within the visual field, while the detail within the scene invokes a scene-internal, intrinsic description. So the reader will not be surprised to find that Japanese – a language with a very similar frame-of-reference inventory – behaves in exactly the same way: standing information is systematically in the relative frame, facing information in the intrinsic frame.

But now consider three languages, Tzeltal, Arrernte and Yélî Dnye, that do not have full, or much used, relative systems. Instead, each has available both absolute and intrinsic frames of reference as fully coded systems. We can code the linguistic details for the description of each photo as in Table 14.6. By way of background, recollect that Tzeltal has an absolute system in which what is roughly south is coded as 'uphill', north as 'downhill', with an orthogonal 'across' in both directions; Arrernte has a fully abstract cardinal direction system which we

Table 14.5 *Propositions used to distinguish Photos 2.3, 2.4 and 2.5 in the Men and Tree Game in two relative + intrinsic languages*

Dutch	Standing/Facing	Photo 2.3	Photo 2.4	Photo 2.5
Pair 1	Standing	Man stands left of the tree	Man at left next to the tree	Man at right next to the tree
	Facing	Man looks towards the tree	He looks off away from the tree	He looks towards the tree
Pair 2	Standing	Man is standing on the left Tree is standing on the right (you see the stick towards you)	Man is standing to the left of tree	Man standing to the right of tree (with stick to the back)
	Facing	Man is looking at the tree	Man is looking left Man with the back to tree	Man is looking at the tree
Pair 3	Standing	—	Man is standing to the left of tree	Man standing on right near tree
	Facing	Man looking left towards tree	He isn't looking at the tree His back towards the tree	To look . . . with his face towards the tree
Japanese				
Pair 1	Standing	Man is at left side of tree	Man standing at left side of tree	Man is at the right side of tree
	Facing	Man is looking at tree	He is orienting his back to tree	Man is looking at the tree

can gloss directly as 'north', 'south', etc., and Yélî Dnye utilizes two axes, not necessarily orthogonal, although approximately so in the location of the recording – one glosses 'up' (east) vs. 'down' (west), and the other 'hillwards' (here south) vs. 'seawards' (here north).

Now a number of important observations can be read off the table. First, in all three communities the absolute frame of reference can be used in this task to make discriminations in what we can call 'tabletop space'. To make this clear: these are linguistic systems in which micro-locations, centimetres apart, may be distinguished in terms of coordinates like north and south. Second, in none of the communities were any discriminations drawn from the relative frame of reference utilized in this task – thus no terms meaning to the visual left or right, or terms meaning in front or behind an object facing the viewer. Arrernte does have 'front', 'back' terms of this relative kind, and so does Yélî Dnye, but they are not often used, and apparently are not appropriate or reliable means of communication in this context. Since all three languages also make available an intrinsic system, it is interesting to see how, within the same tabletop space, the two systems are deployed. We find the following pattern:

	Standing information	*Facing information*
Tzeltal	Absolute and Intrinsic	Absolute
Yélî Dnye	Absolute and Intrinsic	Absolute
Arrernte	Absolute and Intrinsic	Absolute

Thus orientation (facing) is systematically coded in absolute-only terms, although this has to be combined with a body part (e.g. front) or an action (moving, looking) to fix an orientation. This is efficient – no ground needs to be specified to fix an absolute orientation: the man can simply be said to be looking north. But relative position may be coded intrinsically or in absolute terms, or most often both. Here since the relative location of figure and ground is involved, both an intrinsic and an absolute coding are efficient, and both are in fact employed. What is interesting is that we do not find what we might (on an analogy from the European languages) have thought likely – namely that specifying the larger framework will invoke the larger orientational system (here absolute), while specifying the details of the scene will invoke the intrinsic system. The intrinsic systems used in Dutch (and other languages with primary relative frames of reference) and in these three absolute languages are at least in part similar in kind – but it does not follow that the sharing of a frame of reference entails that it will be used for the same functional distinctions.

If we now return to the large table, which includes many details about the contexts of use of each frame of reference, we can add a number of further generalizations. First, here are some generalizations about the relative deployment of the absolute frame of reference in languages with such systems. We can say that the absolute frame is more likely to be used:

Table 14.6 *Propositions used to distinguish Photos 2.3, 2.4 and 2.5 in the Men and Tree Game in three absolute+intrinsic languages*

Language	Standing/Facing	Photo 2.3	Photo 2.4	Photo 2.5
Tzeltal	Standing	Man standing uphillwards Tree is downhillwards	Tree is at his back	Man is downhillwards Tree standing uphillwards
	Facing	Man looking downhillwards	He's looking uphillwards	He is looking uphillwards
Arrernte	Standing	He is standing in the east Tree in region west from him	He's standing on the east-side He puts the tree at his back Tree is behind from him	Man is standing on the west-side Tree is in eastern region
	Facing	He's standing facing westwards	He is standing facing eastwards	He is facing towards the tree
Yélî Dnye (pair Y-L)	Standing	—	Tree at man's back	Tree standing seawards
	Facing	Man facing hillwards Man facing shrub (deduce standing relations)	Man facing seawards	Man approaching tree

1. To provide facing (orientation) rather than standing (placement) information (as we have seen exemplified in the Men and Tree data; see also notes on Kilivila).
2. To describe motion rather than static location (see notes on Jaminjung, Arrernte, Yukatek).
3. To describe static figure–ground as the separation between them increases (see notes on Tzeltal, Arrernte).
4. To describe large-scale space rather than 'tabletop' space (see notes on Jaminjung, Yukatek, Japanese).

Each of these generalizations can be stated as implicational scales: if a language uses the absolute frame for standing information, it will use it for facing information; if it uses it for static description, it will use it for motion; if it uses it when figure and ground are close to one another, it will use it when they are distant; if it uses it for small-scale spatial discrimination, it will use it for large-scale description. Incidentally, one might expect there to be implications from the use of the absolute frame in small-scale space, to its utility in describing the position or parts of the body – but this is an independent variable. Warrwa uses the absolute frame at all scales, but one cannot talk about 'my eastern leg' in Warrwa as one can in Arrernte.

Similar observations are no doubt possible for languages where the relative frame of reference is predominant, but in our sample we have only Japanese and Dutch. Still, a number of languages show small rudiments of a relative frame of reference, and this is very telling from the point of view of cognitive universals – it suggests that the relative frame of reference is always incipiently available, if not fully used in many languages. For example, Tzeltal is a language where the relative frame of reference does not exist as a systematic system, but there are marginal interpretations of terms which are relative in character. If we look at these kinds of cases it is clear that 'front'/'back' terms with relative interpretations are present even where there is no corresponding left/right axis linguistically coded. This suggests a corresponding implicational generalization: if a language has relative 'left', 'right' expressions (as in 'The man is to the left of the tree') then it certainly has relative 'front', 'back' ones (as in 'The man is in front of the tree'). However, Yukatek is a counterexample: it does have clear relative uses of 'left' and 'right' terms, more restricted relative use of 'back' and no relative use of 'front'. This is in line with the suggestion in Levinson (1996b) that left/right oppositions are often different in kind from front/back ones – the former may have to do with place in visual field, the latter with occlusion. The implicational generalization may then be restricted to:

relative 'back' > relative 'front'
(intrinsic 'back' terms are more likely to generalize to relative 'back' terms than 'front' terms – evidence from Yukatek, Tzeltal).

What the observations in this section show is that in this fundamental area of spatial language and cognition, which psychologists have imagined to be conceptually uniform across the species, we find once again significant variation at almost every level. First, although there are only three global frames of reference, not all languages utilize them all. Second, the way in which each selected frame of reference is conceptually constructed can vary in a fundamental way – in the way in which fixed bearings are abstracted, the way in which designated sides are assigned to objects, or the way in which body-axes are mapped to spatial scenes. Third, where more than one frame of reference is deployed, the contextual conditions under which one is used rather than another can be quite various, and it doesn't follow that because a language has, for example, an intrinsic system, that it will employ it for the same purposes that another does. Nevertheless, behind all this cultural, cognitive and linguistic variation there are underlying universals and uniformities. First, all coordinate systems are polar, and only three major classes exist, with different logical and rotational properties (see also Levinson 2003 for geometric primitives shared across the systems). Second, there are constraints on the selection from this set – a relative frame of reference, for instance, implies the use of an intrinsic one. Third, there are many detailed implicational tendencies about the usage of such systems, of the kind illustrated by 'If a language uses an absolute system for the description of stasis, then it certainly uses the same system for the description of motion, but not necessarily vice versa', and the many other examples given above. Once again, then, the picture that emerges is of considerable variation under abstract universal constraints.

14.5 Conclusions

In this chapter, we have tried to extract some of the major underlying patterns arising from a comparison of the coding of spatial distinctions in a dozen languages. The findings are likely to come as a surprise to the reader. The literature, and our own preconceptions, have led us to expect a dominant pattern, in which topological distinctions are coded in rich adpositions, motion in verbs and particles (following the satellite-framing patterns in Germanic), and frames-of-reference issues are exhausted by noting ambiguities between 'intrinsic' and 'deictic' *front* and *back*. In fact, this pattern is in every feature a minority profile in our sample, where the majority of languages do not code topological distinctions in adpositions, do not have satellite-framed motion systems, and do not have an intrinsic plus relative frame-of-reference inventory.

The first impression is one of overwhelming diversity, and apparently endless mismatches between any two languages in both the formal coding of distinctions, and the semantical basis for them. Closer examination, made possible by the use of the same stimulus materials across languages, shows, however,

that there are significant constraints on the diversity. The constraints appear to be of essentially two types. First, there seem to be underlying dimensions of universal relevance for the structuring of spatial sub-domains. Sometimes, as in frames of reference, these amount to a finite set of very abstract types from which languages select. In other cases, as in the topological sub-domain, we seem to have a shared space of possibilities structured on half a dozen universal parameters. And in the motion sub-domain, we have seen that even the semantic construal of the brute fact of motion has to be deconstructed into its underlying components – from these underlying components, distinct types of construal are built up which are reflected in different languages. The second major set of constraints are implicational, rules of the sort 'if a language has a relative frame of reference, it also has an intrinsic one'. These generalizations may have rather different sources, the one just mentioned, for example, probably having a source in diachronic generalization of intrinsic parts to relative interpretations, in order to cope with grounds that lack inherently distinguishable sides.

As far as we know, this book contains the first careful cross-linguistic comparison of this kind outside well-defined, more restricted domains like colour, kinship or ethnobotany. In the absence of this information, many theorists have assumed a strong universal structuring of the spatial domain. Our intuitions about the way space is conceptually structured seem so strong, and children learn at least some of these concepts so easily and early, that we have been led to assume that notions like ON (superadjacent with direct horizontal support) or LEFT as in 'to the left of the tree' are universal primitives in language and cognition. Moreover, it seems to us natural and perhaps therefore unavoidable that motion coding should be in verbs and static locations coded in adpositions. But the picture that comes out of this comparative exercise is altogether more varied and complex. There really is no room at all for the Fodorean view that universal concepts are macro-packages, unanalysable wholes, which now has such a following in linguistics (see Lyons 1995) or psycholinguistics (see Levelt, Roelofs and Meyer 1999). The evidence points to much more abstract underlying parameters as the common root of human conceptualization – old-fashioned componential analysis seems a necessary mode of analysis in comparative semantics, even if for processing purposes speakers treat complex semantic macro-molecules as chunks.

The implications are that the child language learner is a constructivist – he or she is not just mapping local forms onto pre-existing innate concepts but building those concepts as he or she learns the language. Those constructive processes are channelled by universal structuring in the different sub-domains of space, but constructing meanings for spatial words, morphemes and constructions constitutes a significant intellectual achievement – and indeed we know that spatial language is not fully mastered until late childhood (see, e.g., Berman and Slobin, 1994, Brown and Levinson 2000). The task is harder because neither

the meanings nor the forms are antecedently given – the child must construct both domain and range and then the mappings between them. Philosophers (like Quine) and psychologists (like Gleitman) have thought this task impossible and have argued from this to the innate structuring of concepts. But children are better detectives than presumed, and they abduct their way into the system, on the assumption (it seems) that languages have consistent patterns of meaning and coding within them – one solved clue can reveal the patterning of the whole subsystem (see Levinson 2000b). It is this that perhaps explains the divergent styles of motion representation, best reflected above in the Frog Story tellings.

The generalizations we have made in this chapter over the dozen languages should be taken merely as an example of the kinds of comparative observation that can be extracted from controlled comparison. Readers alerted to the potential will be able to find many other points of comparison and contrast between the languages described in the chapters. The point we would like to emphasize here is that such comparison is made possible only by carefully designed elicitation tools. In this domain, as in most others, these tools are in their infancy, and we hope that readers will be inspired to develop such methods further, and through doing so, help to construct a field that today still hardly exists, namely semantic typology. This field has enormous implications for all the disciplines that study language, cognition and their interaction.