Final proofs with index 24 February 2010

Language, Cognition and Space

Advances in Cognitive Linguistics

Series editors: Benjamin K. Bergen, Vyvyan Evans & Jörg Zinken

The Cognitive Linguistics Reader

Aspects of Cognitive Ethnolinguistics. Jerzy Bartmiński. Edited by: Jörg Zinken

Forthcoming titles in the series:

Language and Representation. Chris Sinha

Language, Cognition and Space:

The State of the Art and New Directions

Edited by

Vyvyan Evans and Paul Chilton

equinox
LONDON OAKVILLE

Published by

Equinox Publishing Ltd

UK: Unit 6, The Village, 101 Amies St, London, SW11 2JW
USA: DBBC, 28 Main Street, Oakville, CT 06779

www.equinoxpub.com

Language, Cognition and Space
First published 2010

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN 978-1-84553-252-9 (hardback)
ISBN 978-1-84553-501-8 (paperback)

Library of Congress Cataloging-in-Publication Data
Language, cognition, and space : the state of the art and new directions / edited by Vyvyan Evans and Paul Chilton.
p. cm. -- (Advances in cognitive linguistics)
Includes bibliographical references and index.
ISBN 978-1-84553-252-9 (hb) -- ISBN 978-1-84553-501-8 (pb) 1. Cognitive grammar. 2. Space perception. I. Evans, Vyvyan. II. Chilton, Paul A. (Paul Anthony)
P165.L38 2009
415--dc22
2008054445

Typeset by Catchline, Milton Keynes (www.catchline.com)

Contents

List of contributors

Benjamin Bergen (University of California, San Diego)
Johan Blomberg (Lund University),
Laura Carlson (University of Notre Dame)
Daniel Casasanto (MPI for Psycholinguistics, Nijmegen)
Paul Chilton (University of Lancaster)
Caroline David (Université de Montpellier 3)
Banchiamlack Dessalegn (Johns Hopkins University)
Vyvyan Evans (Bangor University)
Michele Feist (University of Louisiana at Lafayette)
Ariel Micah Goldberg (Johns Hopkins University)
Iraide Ibarretxe-Antuñano (Universidad de Zaragoza)
David Kemmerer (Purdue University)
Barbara Landau (Johns Hopkins University)
Yoshihiro Matsunaka (Tokyo Polytechnic University)
Irene Mittelberg (RWTH Aachen University)
Carl Polley (University of Hawai'i, Mānoa)
Stéphanie Pourcel (Bangor University)
Darya Shakhova (Georgetown University)
Kazuko Shinohara (Tokyo University of Agriculture and Technology)
Leonard Talmy (University at Buffalo, State University of New York)
Andrea Tyler (Georgetown University)
Claude Vandeloise (Louisiana State University)
Kathryn Wheeler (University of Hawai'i, Mānoa)
Jordan Zlatev (Lund University)
Joost Zwarts (Utrecht University)
Jörg Zinken (Portsmouth University)

Introduction

Paul Chilton

You have carried out a complex series of non-linguistic spatial tasks when you picked up this book. You were not aware of all of them but they included, perhaps, navigating your way visually to, into and around a book shop or library. Perhaps you used some specifically space-related language: 'Where can I find that book on language and space'?

This book is to a large extent about how these and other unconscious spatial activities – in the mind and by way of the body – relate to language and how language relates to space. This is an area of enquiry that has interested linguists, philosophers, and psychologists for a very long time and for a variety of reasons. In recent years, say over the past fifteen years, cognitively oriented linguists have devoted more and more research effort into trying to understand the space-language relationship. The present book aims to give an overview of some aspects of this effort, its current state, and the directions in which it is heading.

The long view

The concern with language and space can be seen in a historical perspective, one that begins with physical space per se rather than its relationship with language. The study of space emerged among the ancient Babylonians and Greeks and was received by European civilisation in the form of Euclidean geometry. Aristotle added a view of space that saw it as places embedded in containers, rather than relations, ambiguously using the same Greek word, *topos*, for 'place' as well as 'space'. The next conceptual breakthrough was probably the development of analytic geometry by Descartes and projective geometry by Desargues. It was not until the nineteenth century that non-Euclidean geometries were developed – in effect extending the concept of 'space' beyond what could be intuited through everyday perception.

The concept of space posed problems treated in philosophical discourse before space was investigated by scientific methods that were both empirical and also depended on extensions of mathematics. According to Kant (1781/1963), propositions about space are *synthetic a priori*. This means that propositions about space are not analytic, that is, self-defining. Nor are they dependent on sense-experience, or *a posteriori*. Rather, Kant conceives space as intrinsically incorporated in human understanding. The implication is that human minds are innately endowed with a Euclidean perception and conception, or one might say 'construction', of physical reality. While Kant's philosophical formulation leaves many details unaccounted for, it is consonant with modern views that the human mind has evolved in a way that provides a construction of space fitted to human survival. It is important to bear in mind that such a construction may or may not correspond directly to the objective nature of the physical. Indeed, from the beginning of the twentieth century, mathematics and physics have shown that space is not structured in terms of Euclidean space, although Euclidean geometry serves on the local and human scale.

Somewhat later in the twentieth century, neuroscientists began to put together a picture of the way the human brain may construct space at the neuronal level. The neuro-scientist John O'Keefe, who has contributed pioneering work to mammalian spatial cognition, argues that the three-dimensional Euclidean construction is inherent in the human nervous system. He further argues that it is this necessity that leads to the indeterminacies that appear to exist at the quantum level. (O'Keefe 1999: 47–51). In the present volume, Chapter 1 by Vyvyan Evans, as well as Chapter 6 by David Kemmerer, outlines further details of what many other researchers in cognitive science, psychology and neuroscience have discovered about the human embodiment of spatial experience. Among the findings most relevant from the linguistic point of view is the functional distinction between egocentric frameworks of spatial conceptualisation (neurally embodied in the parietal cortex) and allocentric frameworks instantiated in the hippocampus and surrounding structures.

Language enables humans to communicate about many things – to stimulate conceptualisations among members of a group sharing a language. We can assume that communicating about spatial locations and movements is one area that has particular significance in the evolution of language and languages. And possibly it is the most fundamental area. Indeed, it appears that while not all languages have a word for 'space', there is no language that does not have various classes of words that refer to spatial experience. Some of these words have very general referential power, related to the relative position of speaker or addressee – demonstratives like *this* and *that* for example, which are learned very early by children (cf. Haspelmath 2003). It is equally clear that the human brain has neuronal modules specialised for the perception and cognitive processing of various physical phenomena, such as shape, distance, direction, location and locomotion. As already noted, neuroscientists have accumulated and are accumulating evidence of what exactly these systems are. However, what remains very much less clear is the relationship between linguistic expressions for space, in all their variability and similarity across the world's languages, and the various interacting non-linguistic systems of spatial cognition. This relationship is in fact hotly debated among linguists. In order to make headway it is important to be as clear as we can about the precise questions we need to ask about the language-cognition relationship.

How can we think about the relationship between language and space?

The meanings we find coded in linguistic structures relate not to the physicist's culturally developed understanding of space but to the naturally evolved (and possibly limited) way in which the human brain constructs space for purposes of survival. However, stating the issue in this way makes an important assumption that needs to be questioned. To speak of the evolved system of spatial representation in the human brain is to imply that the system is universally the same for all human individuals. And also it implies that *whatever the language they happen to speak* it is the same for all human individuals. It seems natural to think that all humans would negotiate their physical surroundings in the same way but this view is seriously challenged by scholars who emphasize cultural differences and by scholars who take seriously Whorf's well known view that different languages influence or determine conceptualisation, including spatial conceptualisation (Whorf 1939/2000). After the period in which Whorf's claims were widely seen as discredited, 'neo-Whorfian' research has become intellectually respectable (cf. Gumperz and Levinson 1996, Levinson 1996, Levinson 2003 and many other works) and provides the framework for several chapters of the present volume.

An important question in this relativistic perspective concerns the limits of variation in the way languages encode concepts of space. Variation may not be random but follow universal patterns. If this is the case, then we should further want to explain these patterns in terms of properties of the human space-representing apparatus. It remains logically possible that it is physical objective space itself that structures the human experiential apparatus that variably structures linguistic meanings. It might nonetheless be said, following O'Keefe's arguments (O'Keefe 1999), that it is likely that the human apparatus in some sense imposes its way of seeing things on the physical world. Even so, this does not rule out some form of a modified realist view of the relationship between human cognition and an objective physical universe. In any event, we need to pose at least the following questions.

How exactly do languages encode spatial concepts? That is, what spatial meanings does a language enable its users to exchange among one another? It is important to distinguish *language* from *a language*, for languages may in principle differ in the spatial meanings that they encode. A working hypothesis is that languages do indeed differ in this regard and various empirical investigations have been undertaken in an attempt to prove or disprove it. We need to ask whether differences between languages in regard to the way they express spatial relationships are random and unconstrained. It is quite possible, in a purely logical sense, that variation could mean much overlap and small differential features. It cannot be ruled out of course that even small differences in spatial encoding among languages could correspond to significant cognitive distinctions. Two crucial questions for the contributors to this volume are therefore: Do differences in linguistic encoding of spatial concepts affect non-linguistic conceptualisation of space? And, if so, which elements of spatial encoding are involved? There is now a growing body of empirical research aiming to answer this kind of question. The present volume reports on a wide range of empirical investigations that give varying answers.

There is a further dimension of the space-language question that is rather different, and also more speculative, though has been for some time influential in cognitive linguistics. This line of thinking concerns the possibility that spatial conceptualisations provide the basis for *non-spatial* expressions, including abstract ones. The theories associated with such a perspective go much further than, for example, a view of language that simply postulates a separate module of language to deal with a separate module of human spatial cognition (cf. the models of Jackendoff 1993, 2002). According to such theories, spatial cognition motivates the coding of concepts that are not in themselves self-evidently spatial and in many parts of linguistic structure. This perspective is perhaps clearest in the case of lexical meanings. The lexicalisation of spatial meanings to express temporal meanings is the best known case. For example, temporal terms such as *before* and *after* are etymologically spatial terms, as noted by Traugott (1975), while the moving ego and moving time metaphors have been much discussed since at least Lakoff and Johnson (1980). But similar observations for other domains (cf. 'his behaviour *went* from bad to worse') have been noted since Gruber (1965) and discussed in detail by Talmy (2000) under the rubric of 'fictive motion'. But spatial concepts may well be more deeply integrated with linguistic structure, including grammatical constructions. In the 1970s, some linguists went a considerable distance along this road under the banner of 'localism', treating grammatical categories such as tense and aspect as spatially grounded, as well as, for example, causatives, modals, transitivity, instrumental adverbs, possessive and existential constructions (cf. Lyons 1977: 718–724; Anderson 1971). Although the term 'localism' is no longer used, many approaches in cognitive linguistics are consistent with this idea. In some of its manifestations cognitive linguistics is heavily dependent on spatially iconic diagrams for the purpose of describing a very wide range of linguistic phenomena (e.g. Langacker 1987, 1991). Further developments in formalising the spatial basis of language structure are reflected in Chilton, current volume and 2005). It is also worth noting that O'Keefe's pioneering work in mammalian spatial cognition utilises a geometrical framework that he links speculatively with the evolution and structure of human language (O'Keefe 1996 and 2003).

Overview of the present volume

The interest in space and language culminated in the ground-breaking collection of papers by Bloom *et al*. published in 1996. That volume contained an interdisciplinary perspective with contributions from psychologists, cognitive scientists and biologists, as well as linguists. Some of these contributors are also contributors to the present volume. But since the editors' aim, in the present volume, was not only to sketch the state of the art but also to break new ground and explore new directions, there are many important papers from a new generation of researchers working in cognitive linguistics or in areas overlapping with its concerns.

Following the overview of biological mechanisms involved in the non-linguistic perception of spatial relationships in Section I, Section II opens up the questions that cluster around the Whorf hypothesis, questions that also come up in Section III. Barbara Landau and her associates in chapter 2 propose a solution to the problem of whether the structure of a particular language can influence or determine non-linguistic spatial cognition. Spatial language and visual representations of space organise spatial experience in distinct ways, though the two systems overlap. Landau and colleagues propose that this overlap can be considered in terms of two possible mechanisms – *selection* and *enrichment*.

An example of selection is the way in which some languages 'choose' to code direction rather than manner in verbs of motion – a topic that is controversial and much researched, as will be see from later chapters of this book (Section VII). Another example of selection is frames of reference – also an important topic and one that has led to strong Whorfian claims (Levinson 1996, 2003, challenged by Li and Gleitman 2002). An example of enrichment is the use of spatial language to facilitate orientation tasks by combining geometric and non-geometric information. The chapter describes experimental evidence indicating that language can influence non-linguistic cognitions that can lead to erroneous spatial judgements.

For both selection and enrichment, the question is whether language has a permanent impact on non-linguistic cognition. The authors argue that if it has, then such effects are to be found 'in the moment of a task', when it is possible that language is temporarily modulating attention. This is an approach to Whorfian claims that is similar to Slobin's (1996) notion of 'thinking for speaking'.

Chapter 3, by Benjamin Bergen and colleagues also addresses the possibility that the way a particular language encodes spatial expressions may influence the non-linguistic spatial systems, either in the long term or for the task in hand. If a language can have a particular effect, one has to ask what becomes of universalist assumptions about languages. Like the previous chapter, this chapter also takes the view that the variation in respect of spatial expressions lies within limits and that languages overlap. Further, Bergen and colleagues consider the relationship between spatial and abstract meanings. The issue is whether using such expressions as 'the winter is coming' means that speakers are activating spatial concepts or lexicalised non-spatial meanings. In this chapter the authors review a number of experiments, including their own, that strongly indicate two kinds of relationship between language and cognition. The first kind of relationship is one in which linguistic spatial expressions activate the same neuro-circuitry as non-linguistic spatial cognition. These effects seem to apply not only to primary spatial expressions such as prepositions but to spatial components of verbs and nouns. The second kind of relationship is one in which lexicalised expressions metaphorically derived from spatial concepts, such as those for time, also activate spatial cognition, though evidence from brain lesions is reported that may modify this finding. The chapter assesses also in what sense cross-linguistic differences in the space-time metaphor influence non-linguistic temporal thinking. Whatever the linguistic expression, we also want to know to what extent processing language about space uses the same parts of the brain as processing space per se. Here and also in Kemmerer's chapter (Chapter 6), the reported evidence suggests substantial overlap, although the exact nature and extent of this overlap remains for further research.

What methodologies can we use to investigate the relationship between spatial expressions in language and in non-linguistic spatial cognition? What can we learn from these various methodologies? In Section III we present three different approaches to the descriptive analysis of linguistically activated spatial concepts. Scientific models of space, as noted earlier, have depended on mathematical formulations, from Euclid's geometry, Cartesian coordinates, projective geometries, to topology, spacetime and the seemingly strange behaviour of location and time at the quantum level. As was also noted, it may be the case that Euclidean dimensional geometry, coordinates and vectors are intrinsic to the human systems of spatial cognition. However, although language systems also draw on geometrical concepts – as they must if the claims of the last two chapters are right and if it is also true that non-linguistic cognition is fundamentally geometrical, it does not follow that this is *all* there is to linguistic encoding of spatial concepts. Section III opens by addressing some of the issues that confront linguists when they focus on what exactly it is that constitutes spatial meaning in human languages. In general, the picture that is emerging is that linguistic expressions for spatial concepts involve not only concepts that can be described geometrically but also other concepts useful to humans. Michele Feist's chapter (Chapter 4), for instance, argues that the variation and the commonality found among the world's languages with respect to spatial expressions cannot be adequately described by geometrical means alone. Rather there are *three* factors, one of which is geometry, the other two being 'functional' attributes and 'qualitative physics'. The relevance of functional attributes has been widely commented on (by, for example, Talmy, Vandeloise, Tyler and Evans) and this factor implies that language users draw on various cognitive frames in addition to spatial ones when producing and interpreting spatial utterances. Similarly, Feist claims, speakers draw on naïve physics involving force-related cognition such as support and control of movement. By what means can linguists investigate and describe such dimensions? In addition to psycholinguistic experimentation, drawn on extensively in the preceding chapters, Feist highlights the importance of language-typology studies, and reports findings for the semantics of words corresponding to *in* and *on* across 27 languages from 9 language families.

Nonetheless, geometrical approaches remain a fundamental concern. In Chapter 5, Laura Carlson proposes a new methodology for investigating how several factors modify geometrical frames of reference. The term 'frame of reference' has become standard among researchers into human conceptualisation, though with some variation in application (cf. the review in Levinson 2003: 26). Frames of reference are essentially three-dimensional coordinate systems with a scalar (not strictly metric) quality. They involve direction and (non-metric) distance. The coordinate systems vary in the way (or ways) a particular language requires them to be set up. The origin and orientation of the three axes may be located on the self, another person or fixed by the environment (e.g. landscape or earth's magnetic field). The axis system may be geometrically transformed, e.g. rotated, reflected or translated from their origo located at some reference point, typically the speaker. Each of these possibilities constitutes a reference frame. It should be noted that in addition to coordinate geometry, linguists have often invoked, though loosely, the mathematical notion of topological relations, to describe spatial meanings such as those of containment and contact.

So much is taken for granted. Carlson's method, however, shows how spatial terms such as prepositions have meanings that can be defined by 'spatial templates' (another term already in the literature) or spatial regions within the three dimensional coordinate systems. What is crucial here is that Carlson's methodology shows, consistently with some of Feist's points, how such spaces can be influenced by *non*-spatial linguistic meanings. Different kinds of object in different kinds of settings have varying regions projected by a particular preposition. In different cases, what counts as being *in front* of something is shown to vary depending, amongst other things, on what the reference object actually is, the speaker's typical interaction with it, the presence of other objects, and the particular reference frame being used.

Linguistic analysis and psycholinguistic experimentation give only indirect access to what is happening at the level of brain structure. David Kemmerer's chapter (Chapter 6) surveys the work of cognitive neuroscientists, focussing on one aspect of linguistically mediated spatial cognition, location. Like the authors of the preceding chapters, Kemmerer takes into consideration the language-typological questions concerning variation and commonality, as well as the possible Whorfian effect of language on non-linguistic spatial cognition. Kemmerer poses the question 'what types of categorical spatial relations are encoded by language?' Here 'categorical' is the term used by some psycholinguists and cognitive scientists to refer to the class of linguistic terms devoted to spatial relations (e.g. prepositions). Among categorical relations Kemmerer makes important distinctions between three kinds of spatial relation. The first is deictic relations as found in demonstratives, which non-metrically divide space into proximal or distal zones, varying in number from two (most commonly) to four, depending on the language. These systems may be centred on the speaker (again, most commonly), addressee or some geographical feature. The second is topological relations of the kind alluded to earlier. Reviewing the cross-linguistic evidence, Kemmerer tentatively concludes that there is a universal conceptual space with strong 'attractors'. These are kinds of topological relation, e.g. containment, which languages are statistically likely to encode. The third is 'projective' relations – that is, reference frames in the sense outlined above. While acknowledging the cross-linguistic variety and what may seem prodigious cognitive feats associated with the use of an 'absolute' (or geocentric) reference frame, Kemmerer notes the relatively small number of core spatial concepts encoded cross-linguistically.

But what are the neuro-anatomical correlates? The answers come from the field of cognitive neuro-science, a field that draws its evidence from brain-lesion studies, which work by inference, and brain scanning techniques, typically PET (positron emission tomography) and fMRI (functional magnetic resonance imaging), which are capable of providing information of activated brain regions in on-line linguistic and cognitive tasks. The state of the art in this still developing field is summarised by Kemmerer in terms of brain regions likely to be implicated. Many studies indicate that the left inferior parietal lobule is involved – the region into which projects the 'dorsal' or so-called 'where' pathway of the visual system, involved in precise locating of objects for sensorimotor interaction. (The other pathway is the 'ventral' or 'what' pathway, which identifies objects.) Linguists have variously speculated about correlations between these two pathways and aspects of linguistic organisation. See Landau and Jackendoff 1993 and Hurford 2003a and 2003b) Further, Kemmerer reports studies that suggest that categorical space relations are processed in the adjacent regions of the supramarginal gyrus and possibly the angular gyrus. With regard to deictic, topological and projective relations several brain regions in addition to those mentioned may be involved, perhaps the inferotemporal cortex (linked to the ventral 'what' pathway).

The final part of chapter 6 addresses the Whorfian question, drawing evidence from brain-lesion data and from psycholinguistic experiments on normal individuals. One set of findings indicates that linguistic and perceptual-cognitive representations of categorical spatial relations are to some extent distinct. Another set of findings, however, seems to give some plausibility to the claim that linguistic representation of space can, in a certain sense, influence perceptual-cognitive representation of space. This may happen in two ways. The acquisition of a certain language may decrease an infant's sensitivity to certain categorical spatial distinctions (e.g. mirror image differentiation in Tzeltal speakers) or it may lead to the converse, an increase in sensitivity to certain categorical spatial distinctions. An example of the latter is the finding that Tzeltal speakers utilize an absolute (geocentric) frame of reference in non-linguistic cognitive tasks (cf. Levinson 1996, 2003). Kemmerer is cautious, however, in interpreting such findings as indicating permanent influence on spatial cognition, suggesting a similar interpretation to that put forward by Landau and colleagues in Chapter 2 – namely, that particular linguistic encoding has its primary effect in the moment of executing a task. It is premature to close the debate about Whorfian effects: this is an area for future interdisciplinary research.

It is clear that investigating the relationships between linguistic expressions for space and non-linguistic spatial cognition requires an adequate description on the linguistic side. What are the descriptive methods? What theoretical frameworks are required for such methods? And in which new research directions do these frameworks point? Part IV offers three theoretical approaches to spatial representation in language.

In Chapter 7 Claude Vandeloise builds on his previous work to outline a theory of spatial expressions that makes claims in a diachronic perspective as well as a cross-linguistic synchronic perspective. The theoretical starting point is Berlin and Kay's implicational scale for colour terms (modified by MacLaury) and, for spatial terms, the hierarchical classification of adpositions proposed by Levinson and Meira (see Chapter 7 for detailed references). This chapter argues that the Levinson-Meira model has several problematic features and an alternative hierarchy is proposed. To do this Vandeloise uses several of the theoretical ways of categorising spatial expressions that have been introduced in the previous sections of this book. While the Levinson-Meira model includes only topological kinds of expression, excluding projective (frame of reference) expressions, Vandeloise includes projective and dynamic concepts as well as topological ones. The dynamic concepts are of some significance, since they rest on physical notions of force. Vandeloise does, however, like Levinson and Meira exclude 'kinetic' expressions such as English *from* and *to* (which might also be called directional). The most striking result of this reanalysis is the claim that the most abstract spatial distinction is between 'topological' expressions and 'control', the latter involving concepts of force and energy. Sub-divisions of the proposed hierarchy include projective relations under 'location', while 'containment and 'support' come under 'control'. The hierarchy is claimed to apply to languages universally and also to have relevance for the sequential genesis of spatial terms in the history of particular languages.

Is it possible to have a unified model combining both geometric and force dynamic properties? In Chapter 8, Joost Zwarts proposes a way to model forces that uses the same notational and theoretical framework as is used for coordinate geometry. However, he extends this framework by introducing the elementary mathematics of vectors. This approach may seem unusual to linguists who have worked on spatial expressions, but it is a natural complement to the notion of frames of reference. In fact, O'Keefe (1996, 2003) has already proposed a 'vector grammar' for spatial expressions and other researchers, including Zwarts himself, have pursued the idea (see references in Chapter 8 and also in van der Zee, E. & Slack, J. (eds) 2003). Chilton (2005 and Chapter 19 this volume) takes this framework in a more abstract direction. Because vectors are conventionally used in the applied sciences to model forces, as well as locations in axis systems, Zwarts is able to address 'force-dynamic' prepositions such as *against*, which are not in included in the other classifications, as well as the 'control' type and the 'support' type of prepositions. Further, he is able to address spatial verbs that are both directional and 'forceful' like *push* and *pull* and also to accommodate semantic notions such as Agent and Patient. Zwarts's vector-based framework offers a way of representing and combining spatial relations with the general notion of 'force dynamics' that is much invoked by cognitive linguists and sometimes regarded as distinct from geometry-based descriptions.

Vyvyan Evans's approach (Chapter 9) to three English prepositions (*in*, *on* and *at*) maintains the distinction between spatio-geometrical and 'functional' aspects of spatial meaning, where 'functional' covers 'humanly relevant interactions' with objects in particular spatial configurations. It does this within a cognitive-linguistic theoretical framework that proposes an explicit and wide-ranging theory of linguistic meaning, addressing some of the central issues outlined earlier, in particular, the questions concerning the nature of the interface between linguistically-encoded concepts and non-linguistic concepts. This theoretical framework, which Evans calls the Theory of Lexical Concepts and Cognitive Models (LCCM Theory), is a refinement of Tyler and Evans's earlier Principled Polysemy theory and addresses the important question of extended word meanings. The incorporation of a diachronic perspective is a crucial element of this new theory. What is at issue is the problem of accounting for the emergence of non-spatial meanings of originally spatial terms – such as *in love*, *on alert*, and the like.

In regard to the question of the relationship between linguistic meaning and non-linguistic conceptualization, LCCM Theory starts from the claim that language encodes abstracted, schematic or 'skeletal' concepts, referred to as 'lexical concepts', independently of non-linguistic representations stored as in cognitive models, which are richer. There is a further distinction in LCCM Theory, the distinction between closed-class and open-class forms, which have already been given importance in the work of Leonard Talmy (cf. also his chapter, Chapter 13 in the present volume). Both types are pairings of linguistic form and schematic concepts, but the open class forms provide 'access sites' to cognitive models that tend to be more complex, richer and variable, while closed-class forms (including prepositions) are associated with concepts that are relatively more schematic.

Evans's main concern in the present chapter is with prepositions. Evans's claim is that prepositional concepts are made up from a constellation of conventionalised meanings clustered around a central or prototypical 'spatial scene', a concept that includes both a spatio-geometric and a functional element. He further claims that the central concept gives rise to 'parameters' (or is 'parameterized') over time. A parameter is a kind of sub-schema attached to the central or prototypical schema. Furthermore, language only encodes knowledge via parameterization, in Evans's sense of the term. According to the theory, parameterization arises from the extension in use of the functional (rather than the spatial) ingredients of the central concept. Parameters are abstractions over functional concepts resulting from the use of the central concept in particular human contexts. Over time, such uses become associated with particular functional parameters of the central lexical concept, itself associated with a particular language form: whence the phenomenon of polysemy. Thus, for example, the preposition *in* has physical enclosure as its central and earliest meaning, but eventually produces various distinguishable 'state' parameters, for example 'psychosomatic state', as observed in expressions like *in love*. The non-physical meanings of *in* are thus not computed for each expression on line, but are entrenched meanings associated with the linguistic form *in* for speakers of English – which is not to say that new meaning parameters are not so computed, for this is the very mechanism by which polysemy becomes extended diachronically. In sum, 'states are locations', as Conceptual Metaphor Theory showed us, but the LCCM account puts forward a refined and more detailed explanatory and descriptive framework.

Theoretical frameworks such as those just summarised are an integral part of the overall research endeavour. The empirical investigations of individual human languages makes no sense unless the framework pf description is made explicit and this is why much of the literature on language and space has been, and continues to be, taken up with theoretical refinements. But the reverse is also true – theory has to be complemented and integrated with cross linguistic evidence. Sections V and VI of this volume constitute a sample of such evidence. Since it is the descriptive detail that is crucial, we shall summarise only briefly the content of these chapters, leaving the individual chapters to speak for themselves.

In these chapters examples of spatial expressions are examined in four languages: Basque, Russian, Japanese and American Sign Language. So far we have referred in this Introduction mainly to the English encoding of spatial concepts in prepositions. However, the world's languages (including to some extent also English) distribute spatial concepts across various morpho-syntactic categories. In Chapter 10 Iraide Ibarretxe-Antuñano shows how, in Basque, spatial concepts are distributed across case inflections, 'spatial (or locative) nouns', and motion verbs. Ibarretxe-Antuñano investigates the statistical frequencies of choices from these categories made by Basque speakers in responding to pictorial images of the kinds of spatial relationship that have been termed topological. In addition to the findings of these experiments, Ibarretxe-Antuñano notes that two further dimensions may be involved in Basque spatial expressions, namely dynamicity and agentivity. Such concepts may of course also be relevant for research into other languages and for the general theory of linguistic spatial concepts.

Russian is also a language whose encoding of spatial concepts includes case-marking and rather rich polysemy. In Chapter 11, Darya Shakhova and Andrea Tyler take Russian linguistic spatial expressions as a test case for the theory of Principled Polysemy, already discussed in Chapter 9. If a theory of spatial expressions is to be of interest, it needs to make universal claims about human language – that is, it needs to be applicable to individual language systems. In this chapter Shakhova and Tyler claim that Principled Polysemy theory does indeed make descriptive predictions that are borne out when the Russian data for the preposition *za* are examined in detail. This, as we have noted, means that functional concepts play a crucial role in combination with geometrical (in this case projective) concepts. A detailed analysis of the polysemy network of *za* emerges. One particularly interesting finding, relevant for other case-marked languages, concerns the use of Russian instrumental case with verbs like those corresponding to English *follow*.

A similar general approach is adopted in Chapter 12, by Kazuko Shinohara and Yoshihiro Matsunaka, who investigate the concepts associated with three spatial prepositions in Japanese: *mae*, *ushiro*, *saki*. These prepositions have some conceptual similarity with English *in front of*, and so raise questions about their relationship to frames of reference and the geometric transformations of such frames. Again, the aim is not only to enrich our knowledge of the semantics of a particular language but also to test certain theoretical claims. There are two theoretical claims at issue. The first concerns reference frames, which Levinson and others have claimed are encoded linguistically while others (see references in chapter 12 to Svorou, Carlson-Radvansky and Irwin) have asserted that reference frames are not coded linguistically. The second theoretical issue is what Tyler and Evans, following earlier work by Dominiek Sandra, have called the 'polysemy fallacy': the attribution of unnecessarily many separate meanings to a single lexical item when meaning differences can be explained in terms of contextual inference. This kind of proliferation is typical of some of the earlier cognitive linguistics approaches, e.g. Lakoff's multiple image schemas for the preposition *over*. Shinohara and Matsunaka investigate a particular kind of contextual condition on the conceptualisations associated with *mae*, *ushiro*, *saki* – namely, the effect of the motion of the perceiver and the motion of the perceived object. What the authors of Chapter 12 find, on the basis both of linguistic analysis and psycholinguistic experiments, is that the Japanese prepositions have a minimal specification that includes reference frame information and that the contextual conditions in which they are used makes an important contribution to the associated conceptualisations. In general, their findings uphold Levinson's claims and also those of Tyler and Evans.

The accumulation of empirical findings within a coherent theoretical perspective may eventually lead to a deeper understanding both of universals and of variation in the world's languages. The implicit aim is to gain indirect evidence about the human language system itself and about its relationship with the non-linguistic systems of the human brain. We need logically to decide what we mean by 'language'. The prevailing tendency is to focus on spoken language and spoken languages. However, there are two ways (apart from written representations of the spoken) in which language exceeds the boundaries of the spoken, namely by the use of gesture in conventionalised sign languages where the spoken word is absent, and the use of gesture as an integrated accompaniment of spoken languages themselves. Section VII of the volume consists of two chapters that address these two additional aspects in the context of some far-reaching questions concerning the nature of the human language ability.

Leonard Talmy's contribution, Chapter 13, compares the ways in which American Sign Language expresses spatial concepts with the way in which spoken languages (principally American English) do so. This comparative approach is held to have deep theoretical consequences bearing on the long standing issue of the existence and hypothetical nature of a dedicated language module in the human brain. Like Evans, Talmy distinguishes between closed-class language forms and open-class forms, the former including many elements linked to spatial concepts. In signed languages, there is a subsystem, known as 'classifier expressions', which deals specially with the location or motion of objects with respect to one another. Talmy's chapter compares spoken closed-class spatial expressions with signed language classifier expressions. In order to compare these two space-expressing systems in the two languages, Talmy outlines a detailed theoretical framework that aims to provide the 'fundamental space-structuring elements and categories' universally available in spoken languages. Sign language is then compared with this set. Talmy's claim for spoken languages is that all spatial expressions in spoken languages are associated with conceptual schemas made up of combinations of conceptual elements, organised in categories, and pertaining to spatial scenes. Most of these categories appear to be mainly describable in geometric or force dynamic or other physical terms, while non-geometric properties, such as affective state, figure less prominently in this model. Moreover, the notion of 'functional' meaning is not invoked in the way it is in the models of Vandeloise, and Evans, discussed above. Some of the spatial schemas made up of the spatial elements are more basic than others and can be extended by certain regular processes that, to a certain extent, resemble geometric transformations in Talmy's account. How do sign languages compare? What Talmy finds is that the two language modalities, spoken and signed, share a 'core' in the general design principles governing the spatial concepts for which they have forms. However, sign language differs markedly, both qualitatively and quantitatively, and with a high degree of iconicity. The most general claim is that sign language closely parallels the processing of spatial scenes in visual perception. These findings lead to Talmy to propose a new neural model for language. His hypothesis is that the common ground between the two language modalities (including spatial expressions) results from a single neural system that can be considered the fundamental language module of the human brain. The properties of this module are highly abstract and are primarily concerned with categorising and combinatorial principles. What is new is Talmy's emphasis on the linkage between the human language ability and the visual system, a linkage which is crucial in the case of signed languages.

Chapter 14, by Irene Mittelberg, is complementary to the approach of the chapter by Talmy in two respects. First, rather than considering universals, it examines data from particular speech events representing a particular genre of discourse, namely academic lectures on linguistics. Second, the way in which gesture is considered in this chapter concerns the way in which space is used to communicate meaning, rather than the way in which space is represented in communication. Mittelberg's analytic descriptions show that hand shapes and motion patterns recur across speakers. These hand shape configurations, and traced patterns in the air, may constitute a kind of common sense geometry. Many of these patterns also appear to be iconic visual representations of the kinds of image schemas that are postulated in the cognitive linguistic literature. The hand shapes and hand motions, which are finely synchronised with speech, are not all, however, representations of objects perceived as having the shapes represented. In fact, most of them co-occur with spoken lexical items associated with abstract meanings. Even when these abstract items are not recognisably metaphorical (as many abstract concepts often are), they are synchronised with concrete gesture patterns describing a concrete visual geometric shape. For example, in the lecture discourse studied, the concept of 'category' may co-occur with a container-like cupping of the hands; a 'subcategory' may co-occur with the same gesture located relatively lower in the gesture space relative to the body of the speaker. Or a relatively small space between the finger tips may stand for a relatively small object that is being referred to in the accompanying speech. A fundamental iconic principle is clearly at work and is apparently spontaneously applied by both speaker and hearer. The relationship between the geometric shape of the gesture and the spoken abstract concept is not simply visual, but also kinaesthetic. It is important to note that embodied interaction with objects, specifically manipulation, is what is involved. Moreover, metaphoricity as well as iconicity is at work, since the iconically represented container concept and the iconically represented 'sub' relationship are themselves metaphorically related to the abstract concept of subcategory. The analysis of linguistic data alone has in the past been used to show that image schemas are used in the understanding and coding of abstract concepts. The analysis of concrete gesture data can be claimed to reinforce this claim.

The papers summarised so far predominantly concern spatial representations that are static. Of course, in classical physics, nothing is static except in relation to some frame of reference. That is, motion is relative. Now we have already commented on the various attempts to deal with spatial concepts in terms of classical Euclidean geometry. Similar issues arise when linguists turn to the concept of motion. Clearly, languages contain expressions that are associated with the concepts of motion. But what exactly are these concepts? Are they like the classical Newtonian laws? The two chapters in Section VII include alternate, and to some extent competing, attempts to establish the appropriate conceptual framework for the description and investigation of the ways in which different languages encode concepts of movement through space. As Jordan Zlatev and his co-authors point out, in Chapter 15, a well justified descriptive framework is essential if we are to make headway with some of the intriguing questions that have already emerged in work on the linguistic expression of motion.

These questions revolve around three related issues already mentioned as focal in the study of space and language. Are there any universal tendencies in the linguistic encoding of spatial motion? How do languages vary in this respect? And to what extent does linguistic encoding affect, if at all, non-linguistic cognition? From the beginning of the contemporary interest in this area, the linguistic perspective has had typological implications. Some languages (like English and other Germanic languages) encode the manner of motion in a main verb, while the direction of motion is expressed in a 'satellite' (e.g. prepositional phrase; cf. Talmy 2000). Other languages (e.g. Romance languages) appear to prefer the converse. Thus while French has *Jean traverse la rue en courant*, 'John is crossing the street running'), English has *John is running across the street*. Slobin's empirical work (e.g. Slobin 1996) followed up the potential Whorfian implications – does different linguistic encoding of direction and manner imply different ways of non-linguistic cognizing? The chapter by Zlatev and associates takes on a twofold challenge. First, they take issue with Talmy's classification of motion events and propose a new taxonomy. Secondly, they use their new taxonomy as a basis for exploring the Whorfian questions experimentally. It is clear that human conceptual taxonomy is not that of classical mechanics but a rather fine-grained conceptual characterisation humanly relevant to motion that includes such contrasts as self-motion as opposed to caused motion. They find that the binary typology alluded to above is inadequate to capture the range of variation in the syntactic encoding of direction and manner of motion. In a series of experiments the authors investigate possible effects of coding in three languages (French, Swedish and Thai), finding that the results do not support a strong Whorfian account that would claim different languages entail entirely different views of the world. Indeed, they strongly suggest that the similarities between languages in the relevant respect are greater than their differences.

Like Zlatev and colleagues, Stéphanie Pourcel is among the newer generation of cognitive linguists probing and advancing the pioneering work of Talmy and Slobin. In Chapter 16, Pourcel, again like Zlatev, offers a conceptual revision of motion categories as well as experimental explorations with a neo-Whorfian angle. Her logical starting point is the observation that hitherto experimental studies investigating the possible impact of linguistic categories on non-linguistic cognition have tended to rely on motion categories that are drawn from particular languages – a flaw that she considers a manifestation of 'linguacentrism'. The first part of her paper is thus an attempt to provide a theoretical typology of the cognitive domain of motion that is *independent of language* and that can then be used as a basis for empirical work. The second part of the paper then describes experiments designed to test the language-independent validity of the hypothesized motion categories. As far as the typology is concerned, it is compatible to an extent with Zlatev's in its recognition of the considerable complexity of the conceptual domain of motion. Among other features, motion types involve directionality, causality, telicity and force dynamics. The most important part of Pourcel's typology, however, is probably her insistence on the way the 'existential status' of the Figure (the moving or moved entity) constrains the manners of motion associated with it. In this perspective, the conceptualisation of motion is determined by the animacy, agentivity and causal capacity of the entity involved in a motion event. This makes it possible to allow for the effect of conceptualising biological types and physical forces. Pourcel also notes blended types of moving Figure constructed in, for example, fiction and myth. The empirical investigations reported in the second half of the chapter, including earlier experiments conducted with Anetta Kopecka, lend support to Pourcel's major hypothesis that the conceptualisation of the domain of motion is centred around the Figure schema rather than ground, path, manner or causal motivations. The most general claim here is that a figure-based typology of motion is universal, irrespective of the way particular languages syntactically code motion and manner.

The final part of the volume, Section VIII, contains three papers that go beyond the conceptualisations of spatial relations between physical objects. If spatial conceptualisation is somehow fundamental in the human mind, then should we not expect to find it motivating conceptual domains that are not themselves to do with the physical domain? This question has been behind the observations noted by psychologists and linguists at least since the 1970s. (Interestingly, as Daniel Casasanto notes in his chapter, a similar idea is to be found in an 1898 essay by Jules Lafargue, the son-in-law of Karl Marx.) In Chapter 17, Casasanto takes up some of the key questions that we have seen recurring throughout this book. In particular, Casasanto addresses the question of spatial metaphors in language and their relationship to non-linguistic cognition, using spatial metaphors for time as his 'test bed'. Despite the abundance of linguistic observations in the linguistics literature, we still need to know whether people think about time in spatial terms even when they are not using language, thus whether there are purely mental spatial metaphors for time. We need to know if mental metaphors are universal and whether using a particular language to think or speak about time makes us think in ways constrained by the way those languages encode time concepts. We also need to know if any similarities between linguistic and non-linguistic metaphors for time are simply that, similarities, or causally related. In addition to time metaphors, Casasanto extends his empirical investigation of spatial metaphors to the experience of musical pitch. These particular cases serve to help us understand how humans manage to come up with abstract concepts in the first place, an ancient puzzle for the philosophy of mind. If the answer is indeed that it is a capacity for metaphor that facilitates abstract conceptualisation, then this result could be intriguingly consistent with the notion of exaptation in evolutionary biology. According to Casasanto, the experimental evidence points clearly both towards the use of spatial metaphors in thinking about time (but not vice versa) and towards the influence of particular linguistic encoding on non-linguistic spatial metaphors, and this not merely in a 'thinking for speaking' sense.

Chapter 18 also concerns the nature of spatial metaphors for time. As has been seen, many investigators assume that frames of reference, which essentially are Cartesian coordinate systems, provide the basis for the analysis of spatial cognition, whether linguistically expressed or not. Jörg Zinken raises the question of whether frames of reference are therefore also relevant to the description of the abstract conceptualisation of time. He also asks, as Zlatev and Pourcel do for motion, whether the existing typologies for the concept of time are adequate for further cognitive-linguistic investigation. The answer to this last question is that existing assumptions need to take into account the rich anthropological literature concerning cultural variation in the understanding of time, as well as two views of time that are commonly found in the philosophical literature. These two views are, roughly, the experiencer-centred view of time (one event comes to us after another and fades into the past), and the experiencer-independent view, according to which events remain for ever strung out in a sequence. With respect to the question whether frames of reference are relevant, Zinken argues that they are, and shows how the three reference frame types formulated by Levinson (2003) have consequences not just for spatial cognition and linguistic expression but also for spatial metaphors for time. The upshot is a proposal for a new typology of temporal concepts that combines reference frames and the distinction between experiencer-centred and experiencer-independent time concepts. This more detailed framework is required, in Zinken's view, in order to make progress in the empirical exploration of space-time metaphors across the world's languages in relation to non-linguistic cognition. The approach outlined in this chapter is not, however, entirely couched in terms that can be characterised as broadly geometrical. While other contributors to the volume emphasise the 'functional' factor in linguistically encoded conceptualisation of physical space, Zinken focuses on the possible contribution of cultural factors. To what extent is the English tendency to conceptualise immediate and distant future time as 'in front of' the speaker explained by a culture of forward planning and manipulation of events? Such far-ranging questions point to new goals in cognitive-linguistic research.

The final chapter of the volume is a theoretical speculation concerning the possible extension of spatial concepts to the description of more abstract aspects of language structure, including grammatical constructions. In this short chapter I extend the notion of reference frames, a notion that, as we have seen, emerges as fundamental in language-and-space research. I take frames of reference to be Cartesian coordinate systems defining a three dimensional space but I apply them not to physical space (or even a metaphorical target domain such as time in Zinken's paper) but to what I call the abstract 'discourse space' (for more details see Chilton 2005). Of course, since Descartes, n-dimensional spaces have been defined and explored extensively by mathematicians, but the three Euclidean dimensions might be especially significant for humans, as pointed out at the beginning of this Introduction. The model I propose is defined in three axes. These are: discourse distance (essentially Figures are 'closer' to the speaker than Grounds); time (some events in both past and future are 'closer' than others) and epistemic modality (epistemically more certain events are 'close' to the speaker and counterfactual ones are 'remote'). One further ingredient is added, namely simple vectors, which have distance and direction. Consistently with a major component of the account of spatial prepositions, the abstract axis systems can be transformed (cf. the 'projection' of Levinson's 'relative frames'). Using geometrical diagrams for a large part of the argument, I suggest perhaps surprising aspects of viewing certain syntactic and semantic phenomena in terms of geometric transformation. This might appear to be pushing the geometric approach too far, and I certainly do not wish to ignore the 'functional' components that are treated by various authors in the present volume. However, the geometrical description of space provides the essential scaffolding in all accounts of spatial expressions, as we have seen. This is not surprising, if the notion of embodiment is taken seriously. And this is also why my chapter, and indeed this whole volume, begins with a review of the grounding of spatial perception and conception in biological systems.

References

Anderson, J. M. (1971) *The Grammar of Case: Towards a Localist Theory*. Cambridge: Cambridge University Press.

Bloom, P., Peterson, M. A., Nadel, L. and Garrett, M. F. (eds) (1996) *Language and Space*. Cambridge, MA: MIT Press.

Chilton, P. A. (2005) Vectors, viewpoint and viewpoint shift: Toward a discourse space theory. *Annual Review of Cognitive Linguistics*, 3: 78–116.

Gruber, J.S. (1965/1976) *Studies in Lexical Relations*. PhD thesis MIT. Reprinted in *Studies in Lexical Relations*. Amsterdam: North Holland.

Gumperz, J. J. and Levinson, S.C. (eds) (1996) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Haspelmath, M. (2003) The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (ed.) *The New Psychology of Language* (Vol. 2, pp. 211–242). Mahwah, NJ: Lawrence Erlbaum Associates.

Hurford, J. (2003a) The neural basis of predicate-argument structure. *Behavioural and Brain Sciences*, 22(4): 261–283.

Hurford, J. (2003b) Ventral/dorsal, predicate/argument: The transformation from perception to meaning. *Behavioural and Brain Sciences*, 26(3): 301–311.

Jackendoff, R. (1993) *Semantics and Cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Kant, I. (1781/1963) *Critique of Pure Reason.* (Second edition.) Translated by M.K.Smith. London: Macmillan and Co.

Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*. Chicago: Chicago University Press.

Landau, B. and Jackendoff, R. (1993) 'What' and 'where' in spatial language spatial cognition. *Behavioural and Brain Sciences*, 16: 217–238.

Langacker, R. (1987) *Foundations of Cognitive Grammar vol. I*. Stanford: Stanford University Press.

Langacker, R (1991) *Foundations of Cognitive Grammar vol II*. Stanford: Stanford University Press.

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

Levinson, S. (2003) *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Li, P. and Gleitman, L. (2002) Turning the tables: Language and spatial reasoning. *Cognition*, 83(3): 265–294.

Lyons, J. (1977) *Semantics.* (2 vols.) Cambridge: Cambridge University Press.

O'Keefe, J. (1996) The spatial prepositions in English, vector grammar and the cognitive map theory. In P. Bloom et al. (eds) *Language and Space* (pp. 277–316). Cambridge, MA: MIT Press.

O'Keefe, J. (1999) Kant and the sea-horse: An essay in the neurophilosophy of space. In N. Eilan, R. McCarthy and W. Brewer (eds) *Spatial Representation: Problems in Philosophy and Psychology.* Oxford: Oxford University Press.

O'Keefe, J. (2003) Vector grammar, places, and the functional role of the spatial prepositions in English. In E. van der Zee and J. Slack (eds) *Representing Direction in Language and Space*. Oxford: Oxford University Press.

Slobin, D. I. (1996) From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz and S. C. Levinson (eds) *Rethinking Linguistic Relativity.* Cambridge: Cambridge University Press.

Talmy, L. (2000) Fictive motion in language and 'ception'. In *Toward a Cognitive Semantics* (Vol. I pp. 99–176). Cambridge, MA: MIT Press.

Traugott, E. C. (1975) Spatial expressions of tense and temporal sequencing. *Semiotica*, 15: 207–230.

Van der Zee, E. and Slack, J. (eds) (2003) *Representing Direction in Language and Space.* Oxford: Oxford University Press.

Whorf, B. L. (1939/2000) The relation of habitual thought and behavior to language. In J. B. Carroll (ed.) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf* (pp. 134–159). Cambridge, MA: MIT Press.

Zwarts, J. (1997) Vectors as relative positions: A compositional semantics of modified PPs. *Journal of Semantics*, 14: 57–58.

Part I

Perception and space

Final proofs with index 24 February 2010

1 The perceptual basis of spatial representation

Vyvyan Evans

Overview

The human experience of space includes knowledge relating to the size, shape, location and distribution of entities in a stable three-dimensional environment. In this introductory chapter I address the perceptual systems and processes that facilitate this: the sense-perceptory and brain mechanisms that process perceptual information giving rise to spatial experience. I also examine the processes whereby perceptual experience is redescribed into rudimentary representations of space. That is, I examine primitive concepts which form the bedrock of our ability to think, reason and talk about space and, indeed, more abstract realms. Thus, this chapter is concerned primarily with i) the perception of space, and the way in which spatial experience is 'constructed' by virtue of our sense-perceptory systems and brain mechanisms, and ii) how spatial experience is 'redescribed', giving rise to foundational spatial concepts prior to the emergence of language from around one year onwards.

The chapter begins by examining the distinction between spatial representations, exploring the difference between percepts and concepts. I then examine the human perceptual systems which facilitate the detection of sensory stimuli from the external environment. I then look at perceptual theories which attempt to explain how the brain constructs spatial experience from this sensory input. I then turn to the human mapping ability: an innate mechanism that allows us to construct spatial or cognitive 'maps' based on locational information. This ability is essential for wayfaring, which is to say navigating in space. I then examine how percepts are redescribed as the basic spatial primitives, known as image schemas.

1 Introduction: perception vs conception

My main concern in this chapter is to review the way in which space is experienced and constructed by the human *sensory* (or *sense-perceptory*) *systems*, and the brain. I also review the way in which these objects of spatial perception known as *percepts* give rise to rudimentary spatial representations (or *concepts*) known as *image schemas*. Accordingly, at this point I briefly review the distinction between perception (and percepts), and conception (and concepts).

Perception consists of three stages: i) *sensation* ii) *perceptual organisation* and iii) *identification and recognition*. Sensation concerns the way in which external energy, such as light, heat, or (sound) vibrations are converted into the *neural codes* which the brain recognises. Perceptual organisation concerns the way in which this sensory information is organised and formed into a perceptual object, a percept. Identification and recognition relates to the stage in the process whereby past experiences and conceptual knowledge is brought to bear in order to interpret the percept. For instance, a spherical object might be identified and recognised as a football or a coin, or a wheel, or some other object. That is, this stage involves meaning, which is to say understanding the nature, function and significance of the percept. As such, a previously-formed concept is employed in order to identify and categorise the percept.

Table 1. Three stages in perception

Sensation	**external energy stimuli are detected and converted into neural codes**
perceptual organisation	integration of neural codes by the brain to form a percept
identification and recognition	the percept is categorised, which involves matching with stored experiences

The distinction between percepts and concepts relates to distinctions in *representational formats*: how experience is presented at the cognitive level and how it is stored. Percepts constitute coherent representations which derive from sensory experience, and arise from multiple *modalities*. That is, they derive from information which is integrated from a number of different sensory systems, discussed in more detail in the next section. Percepts are typically available to conscious experience. That is, they are the product of *on-line processing*, resulting from a stimulus array perceived in the 'here-and-now'. A consequence of this is that they consist of specific information relating to the specific stimulus array that they are derived from. Thus, they are *episodic* in nature.

Concepts, on the other hand, represent *schematisations*, formed by abstracting away points of differences in order to produce representations which generalise over points of similarity. Thus, the concept CAR, for instance, is a schematisation derived by generalising across many different sorts of specific (episodic) experiences relating to automobiles in order to form a single representation. Of course, this greatly simplifies things, and I emphasise that concepts, while stable schematisations are not static and unchanging. Indeed, they continue to be updated and thus evolve as the human perceiver continues to be exposed to new experiences. A consequence of the schematic nature of concepts is that, unlike percepts, concepts are representations in the sense of re-presentations. That is, they are stored in memory and can be activated during *off-line processing*. That is, they can be recalled in the absence of the percept(s) which may have given rise to them.

A further important point is that while percepts relate primarily to the sensory details of a given entity, concepts include a much greater range of information types, including the nature and function of the entity which is being represented, as well as how it relates to other concepts. Thus, concepts are related to one another in a systematic way, and form a structured knowledge 'inventory', what I will refer to as the human *conceptual system*. Thus, concepts constitute 'theories' concerning a particular entity, and as such bring meaning to bear with respect to any given percept (for discussion see Mandler 2004).

This said, how do percepts and concepts arise? Percepts arise from a process termed *scene analysis* (e.g., Bregman 1990). Scene analysis is the process whereby the perceptual stimulus array is segregated into coherent percepts. This is achieved by both *bottom-up processing* and *top-down processing*.

Bottom-up processing relates to the processing and integration of perceptual 'details' that make up, for instance, object percepts, such as a vase or a ball. I will consider two sorts of perceptual details later in the chapter which are termed *textons* and *geons*. Top-down processing relates to the integration of perceptual information which is guided by global principles. Such principles have been proposed, for instance by Gestalt psychology, an important and influential movement that I will consider in detail below.

Bottom-up and top-down processing cross-cut another important distinction which relates to *primitive segregation* versus *schema-based segregation*. That is, scene analysis proceeds by making use of both innate and learned constraints. Primitive segregation is segregation of the stimulus array based on innate, which is to say, pre-given, primitives. Such primitives, which include, for instance *figure-ground segregation*, discussed below, derive from invariants in the stimulus array which have, through evolutionary processes come to be 'hard-wired' in the human brain. In contrast, schema-based segregation involves scene analysis which employs learned constraints.

Before concluding this section, it is necessary to briefly say something about the relationship between spatial concepts and percepts. In fact, this is an issue I address in greater detail when I present the work of developmental psychologist Jean Mandler later in the chapter. However, for now I note that spatial concepts derive from, in the sense of being 'redescribed' from, perceptual experience. This process, which Mandler refers to as *perceptual meaning analysis*, uses spatial percepts as the basis for the formation of rudimentary spatial concepts: image schemas. I will have more to say about these basic spatial concepts later.

2 Sensory systems

In this section I review the mechanisms that facilitate the processing of energy signals from the environment, the *stimulus array*, and how this information is detected by our sensory systems, and processed. I begin by examining the sensory organs and systems which serve as our windows on our spatial environment.

2.1 The visual system

The crucial organ for the visual system is the eye. The brain and the eye work together to produce vision. Light enters the eye and is changed into nerve signals that travel along the optic nerve to the brain. As light enters the eye it is brought into focus on the rear surface of the eyeball. Light enters at the cornea (see Figure 1), which helps to bend light directing it through the pupil: the small dark circle at the centre of your eye. The amount of light that enters the pupil is controlled by the iris – often coloured brown or blue and encircles the pupil – which expands or contracts making the iris larger of smaller. Behind the pupil is a lens, a spherical body, bringing light waves into focus on the retina, the rear of the eyeball. The retina consists of a thin layer of light receptors known as photoreceptors. There are two kinds of photoreceptors: cones and rods. Cones allow us to see in colour and provide our perception in daylight. Rods facilitate vision under dim conditions and allow only black and white perception. That part of the retina which is most sensitive is called the *macula*, and is responsible for detailed central vision. The part of the macula which produces clearest vision is the *fovea*. It is a tiny area densely packed with cone cells. Accordingly, when we look ahead, light reflected from objects in our 'line of sight' is directed onto our fovea, and objects occupying this area of the macula are perceived by virtue of what is termed *foveal vision*. Objects at the edge of the visual field are perceived less clearly. Vision of this kind is known as peripheral vision.

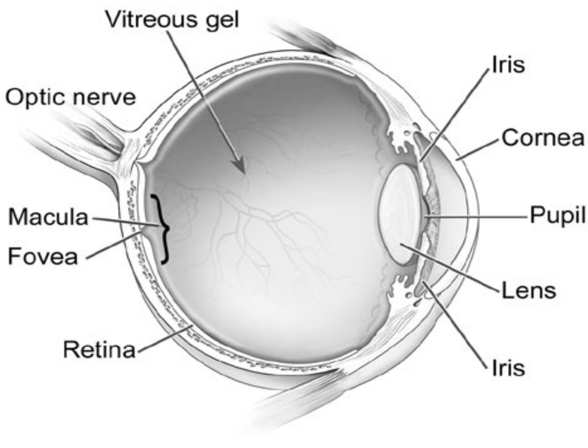


Figure 1. The eye

'What' and 'where' visual systems

The photoreceptor cells on the retina convert light energy into neural information. However, this information from different parts of the retina is carried along two different pathways or 'streams', connecting different parts of the *visual cortex* – that part of the brain responsible for vision – and providing distinct sorts of information. The visual cortex occupies about a third of the (*cerebral*) *cortex*, the outer layer of the *cerebrum* (consisting of four lobes, see Figure 2).

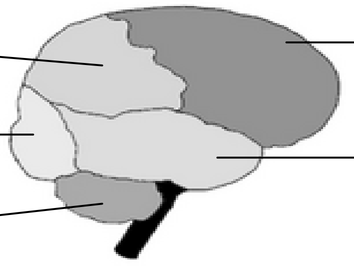


Figure 2. Diagram showing the four lobes of the cerebrum, and the cerebellum (The cebral cortex is the outer layer of the cerebrum. Note: The brain is seen from the right side, the front of the brain, above the eyes, is to the right.)

The visual cortex is divided into approximately thirty interconnected visual areas. The first cortical visual area is known as the primary visual cortex or V1. V1 sends information along two separate pathways or 'streams' through different parts of the visual cortex, giving rise to two separate visual systems each providing different kinds of information (Ungerleider and Mishkin 1982). The primary visual system, known as the *focal system* sends information from the macula along the pathway known as the *ventral stream* (ventral means 'lower'). This system, often referred to as the 'what' system, provides information relating to form recognition and object representation. That is, it allows us to identify and recognise objects, including the recognition of attributes such as colour, for instance.

The second system, known as the *ambient system* sends information from both the macula and more peripheral locations on the retina along a pathway known as the *dorsal stream* (dorsal means 'upper'). This system, also known as the 'where' system, provides information relating to where an object is located in body-centred space, rather than with details of the object itself. Thus, light signals in the eye are transformed by the brain providing two distinct sorts of information relating to 'what' and 'where'.

More recently Milner and Goodale (1995) have demonstrated that the distinction between the two 'streams' does not strictly relate to the type of percept ('what' versus 'where') that visual processing provides, in the way conceived by Ungerleider and Mishkin. Rather, while the ventral stream provides information that allows humans to perceive particular objects ('what'), the dorsal stream provides functional information which facilitates readiness for action in order to interact with objects and other entities in the world. In other words, the ventral stream provides information leading to the conscious understanding of objects and other entities in the physical environment, while the dorsal stream serves to facilitate motor programming.

Important evidence for these two distinct visual systems comes from the phenomenon known as *blindsight*. Some blind individuals appear to be able to localise and orient to objects without actually being able to see them. In other words, some blind people appear to be able to locate objects without knowing what the objects are, that is, without being able to identify the object. This suggests that in such cases while the focal system is damaged, the ambient system, mediated by the dorsal stream allows them to make correct orientation judgments and responses, providing compelling evidence for two distinct kinds of visual information.

Recent work on spatial representation in language suggests that the 'what' and 'where' systems may have linguistic reflexes. For instance, Landau and Jackendoff (1993) argue that spatial relations, as encoded by prepositions, and objects as encoded by count nouns roughly approximate the pre-linguistic representations deriving from the 'where' and 'what' systems respectively. Similarly, Hurford (2003) argues that the 'where' and 'what' systems provide neurological antecedents for predicate-argument structure in language.

2.2 The vestibular system

The *vestibular system*, or *orienting sense* is the sensory system that provides information relating to our sense of balance, and is the dominant system with respect to sensory input about our movement and orientation in space. Together with the cochlea, the auditory organ, discussed below, the vestibular system, is situated in the *vestibulum* in the inner ear (Figure 3).

As our movements in space consist of *rotations* – circular motion, as when we turn around – and *translations* – linear motion, as when we walk along a path (horizontal motion), or climb a ladder (vertical motion or gravity) – the vestibular system comprises two components. The first component consists of semicircular canals which detect rotations. These are interconnected fluid-filled tubes which are located in three planes at right angles to one another. The inner surface of the canals also contain hairs. As the fluid moves in response to rotational movement the hairs detect motion of the fluid and transduce this into neural code. The three distinct canals serve to provide rotational information from three axes.

The second component consists of two fluid-filled sacs, the *utricle* and the *saccule*. These chambers contain *otoliths* – literally 'ear stones' – which are heavier than the fluid in the sacs and respond to linear and vertical motion, including both left-right, forward-back motion and gravity (vertical motion). As before both the utricle and saccule contain hairs which detect movement of the otoliths in response to linear motion. This information is transduced into neural code which is transmitted to the brain for processing.

The vestibular system sends signals primarily to the neural structures that control our eye movements, and to the muscles that keep us upright. One important function of the vestibular system is to coordinate body and head movement with the detection of motion by the visual system. This is referred to as the *vestibulo-ocular reflex (VOR)*, which is necessary for vision to remain clear. This works during head movement by producing an eye movement in the direction opposite to head movement, thus preserving the image on the centre of the visual field. For example, when the head moves to the right, the eyes move to the left, and vice versa. Since slight head movements are present all the time, the VOR is very important for stabilising vision.

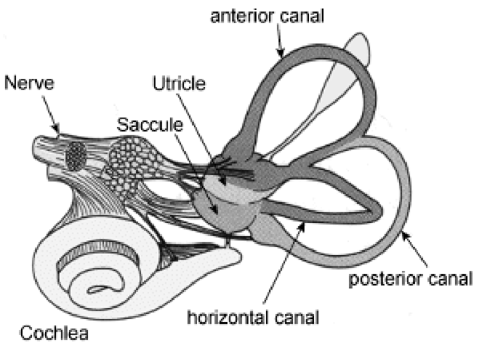


Figure 3. The vestibular system and cochlea

The vestibular system is, in *phylogenetic* (i.e., evolutionary) terms, one of the first systems to have developed. In *ontogenetic* (i.e., developmental) terms it is the first to fully develop, by six months after conception.

2.3 The auditory system

The vestibular system, and the key auditory organ, the cochlea, are closely linked, both occupying the ear bone. It is widely believed that the cochlea evolved from the phylogenetically earlier sensory structures responsible for detecting bodily orientation.

The auditory system works by transforming sensory information first from air to fluid and then to electrical signals that are relayed to the brain. One important function of the ear is to amplify sound vibrations, in preparation for the transformation from air to fluid. The folds of cartilage that comprise the outer ear on the side of the head are called the *pinna* (see Figure 4). The sound waves enter the ear canal, a simple tube which starts to ampify the sound vibrations. At the far end of the ear canal is the eardrum which marks the beginning of the middle ear.

The middle ear includes the *ossicles* – three very small bones shaped like a hammer, an anvil, and a stirrup. The ossicles further amplify the sounds by converting the lower-pressure eardrum sound vibrations into higher-pressure sound vibrations. Higher pressure is necessary because the inner ear contains fluid rather than air. The signal in the inner ear is then converted to neural code which travels up the auditory nerve.

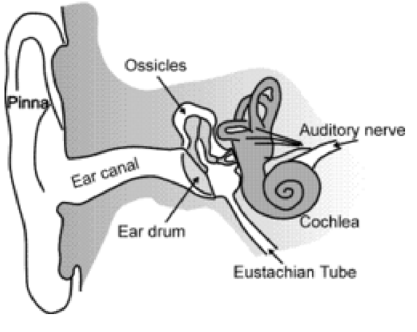


Figure 4. Anatomy of the ear

The auditory nerve takes the neural code to that part of the brainstem known as the cochlear nucleus. From the cochlear nucleus, auditory information is split into two streams, similar to the way in which the visual signal is split into 'where' and 'what' streams. Auditory nerve fibres going to the ventral cochlear nucleus preserve the timing of the auditory signal in the order of milliseconds. Minute differences in the timing of signals received by both ears allow the brain to determine the direction of the sound.

The second, dorsal, stream analyses the quality of sound. It does this by virtue of detecting differences in frequencies and thus allows differentiation of phonemes, such as the distinction between *set* versus *sat*.

2.4 The haptic system

The haptic system includes the combined sensory input from the receptors for touch in the skin and *proprioception receptors* in the body's muscles and joints. Together sense-perception from the haptic system gives rise to perceptual information from a broad range of contact encounters between the body and environment that are sent to, and processed by, a region of the cerebral cortex known as the *somatosensory area*. The haptic system – deriving from *hapsis* which is Greek for 'to grasp' – provides perception of geometric properties including the shape, dimension, and proportions of objects. It also gives rise, through the proprioceptive receptors, to the felt sense of co-ordinated movement, and thus is responsible, in part, for our perception of being distinct from the environment which surrounds us. I review in more detail below the two key components that make up the haptic system, the skin, and proprioception.

The skin

The skin is the largest organ, covering the entire body. It contains specialised nerve endings which can be stimulated in different ways providing different sensations and thus different sorts of sensory information. The sensory effect resulting from stimulation of the skin is known as *cutaneous sensitivity*. There are three main *cutaneous qualities*: pressure (also known as touch), temperature and pain. The somatesensory cortex in the brain represents different skin regions as well as different cutaneous qualities. Thus, the brain is provided with information relating to where on the skin a particular stimulus is being received and what sort of quality is associated with it.

In terms of touch there is an important distinction to be made between *active touch*, and *passive touch*. In active touch, the experiencer actively controls sensory stimulus activation by virtue of picking up an object, for instance. By contrast, passive touch occurs without the reception of the stimulus being controlled by the experiencer, as when an object is placed in contact with the skin. Although the entire surface of the skin responds to touch, the most sensitive receptors are the 'exploratory' parts of the body. These include the fingers and hands, parts of the mouth and the tip of the tongue, as well as the genitalia.

Proprioception

Proprioception – from the Latin *proprius* which means 'one's own' – relates to the sense of body part position and movement. That is, it concerns the posture, location and movement of the arms, legs and other parts of the human skeleton. Another commonly-used term for proprioception is *kinaesthesis* – or kinaesthesia, from the Greek *kineo*, 'to move'. Proprioception is essential for a whole range of coordinated movements. To get a sense of how it functions close your eyes and then touch your nose with a finger tip. Your ability to do this comes from proprioception.

Proprioceptive receptors are known as *mechanoreceptors*. There are two types. The first type provides sensory stimuli for joint information. The second provides information deriving from mechanoreceptors founds in muscles and tendons. The mechanoreceptors for joint information are stimulated by contact between the joint surfaces. This occurs when the angles at which bones are held with respect to one another change, due to movement. The mechanoreceptors in the muscles and tendons respond to changes in the tension of muscle fibres when movement occurs.

3 Spatial perception: how we experience space

In this section I review the perception of objects, form, movement and three-dimensional space. Perhaps unsurprisingly, given the importance of the visual modality for primates in general and humans in particular, much of the work on various aspects of spatial perception has traditionally focused on visual cues. Indeed, visual perception is perhaps the best studied of the sensory systems. Accordingly, in this section I will primarily focus on the role of visual perception in the experience and construction of spatial percepts.

3.1 Texture and object perception

Before objects can be identified visual details must be processed and integrated by the visual system. Variations in visual scenes, in terms of i) light intensity, i.e., adjacent regions of light and dark areas – known as *contrast phenomena* – ii) patterns and iii) colour, form repeated patterns known as *visual texture*. The patterns, for instance, curly versus straight hair, or a tiger's stripes versus a leopard's spots, are often the result of the physical surface properties such as differentially oriented strands, and direction of light and direction of motion.

One important bottom-up theory of visual texture perception is known as *Feature Integration theory*. This theory assumes that there are two major stages involved in the perception of visual texture. The first stage, known as the *preattentive stage*, involves the unconscious processing of visual texture. In a seminal paper, psychologist Bela Julesz (1981) proposed that the preattentive stage serves to process textural primitives, the fundamental components of visual texture. These he labelled *textons*.

Textons are distinct and distinguishable characteristics of any given visual display. For instance, textons include straight lines, line segments, curvature, widths, lengths, intersections of lines, and so on. According to Julesz, the first stage of visual texture perception involves discriminating between the range of textons in a visual display. The second stage in visual texture perception is the *focused attention stage*. This involves conscious processing in order to integrate the textons into complex unitary objects.

Just as textons have been proposed as the primitive elements of visual texture perception, a related bottom-up theory has been proposed to account for object identification. This theory, associated with the work of Biederman (1987) is called *recognition by components*. Biederman's essential insight is that the identification of objects involves the combination of a set of primitive three-dimensional geometric components which he labels *geons*, short for 'geometric icons'. Geons are simple volumes such as cubes, spheres, cylinders, and wedges (see Figure 5). Biederman has proposed 36 geons which can be combined in a range of ways giving rise to complex objects. Biederman argues that object perception crucially relies upon recognising the components which make up an object, the geons. Figure 6 illustrates how a perceived object is comprised of a range of constituent geons. The image on the left corresponds to the perceived object (a desk lamp), and the image on the right to the constituent geons.

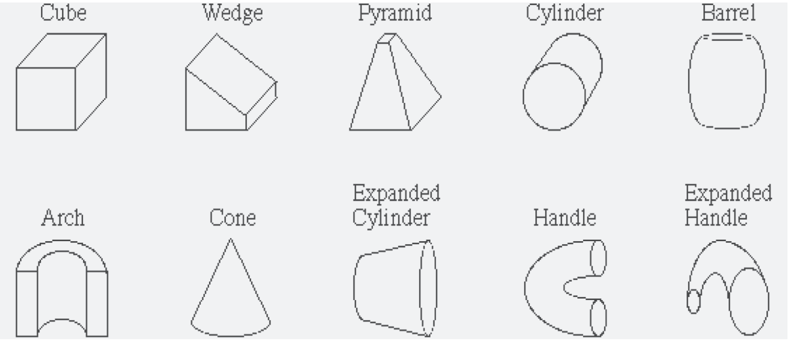


Figure 5. Some examples of geons (After Biederman 1987)

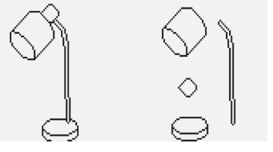


Figure 6. Geons in object perception

3.2 Form perception

In the previous section I briefly looked at primitive elements that have been proposed for textual perception and the identification of objects. However, in addition to identifiable components of images and objects, there are also higher-level processes involved that are essential for the perception of forms and the grouping of objects. Moreover, these appear to be innate. I discuss two sorts of such organising principles below, figure-ground segregation, and the Gestalt grouping principles.

Figure-ground perception

A fundamental way in which we segregate entities in our environment, thereby perceiving distinct objects and surfaces, comes from the our ability to perceive certain aspects of any given spatial scene as 'standing out' from other parts of the scene. This is known as *figure-ground organisation*.

The phenomenon of figure-ground organisation was pointed out by the Danish psychologist Edgar Rubin in 1915. He observed that in visual perception we see parts of a given spatial scene as being made up of well-defined objects, which 'stand out' from the background. That is, we see objects as three-dimensional entities which stand out from the terrain in which they are located. For instance, in Figure 7, the image of the lighthouse, the figure, stands out from the grey horizontal lines, the ground, as a recognisable and distinct image.

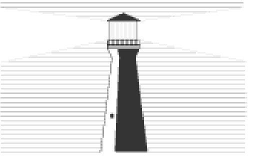


Figure 7. Figure-ground segregation

Rubin proposed a number of perceptual differences between the figure and ground. These are summarised in table 2.

Table 2. Distinctions between figure and ground

Figure	Ground
Appears to be thing-like	appears to be substance-like
a contour appears at edge of figure's shape	relatively formless
appears closer to the viewer, and in front of the ground	appears further away and extends behind the figure
Appears more dominant	less dominant
better remembered	less well remembered
more associations with meaningful shapes	suggests fewer associations with meaningful shapes

In addition, figure-ground perception appears to be innate. For instance, photographs which lack depth, being two-dimensional surfaces, are perceived in three-dimensional terms. That is, the figure-ground organisation associated with photographs is an illusion. A particularly well-known illusion made famous by Rubin is the vase-profile illusion (Figure 8).

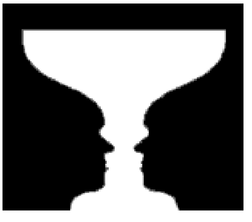


Figure 8. The vase/profile illusion

The vase/profile illusion is an ambiguous figure-ground illusion. This is because it can be perceived either as two black faces looking at each other, on a white background, or as a white vase on a black background. In other words, it undergoes spontaneous *reversal***.** This illusion shows that perception is not solely determined by an image formed on the retina. The spontaneous reversal illustrates the dynamic nature of the perceptual processes. These processes illustrate that how the brain organises its visual environment depends on our innate ability to segregate images on the basis of figure-ground organisation. As this image contains the same percentage of black and white, that part of the image which is assigned the role of figure determines whether a vase or faces are perceived.

Figure-ground organisation appears to be an evolutionary response to our physical environment. Our visual system, for instance, has evolved in order to be able to perceive three-dimensional objects as distinct from the surrounding terrain in which they are embedded. Figure-ground organisation thus constitutes a hard-wired response to this imperative.

Gestalt grouping principles

Gestalt psychology was a movement which emerged in the first decades of the twentieth century. Its primary concern, and those of its three leading proponents, the German psychologists Max Wertheimer, Kurt Koffka and Wolfgang Köhler was to investigate why some elements of the visual field form coherent figures, and others serve as the ground. Gestalt is the German term for 'form', or 'shape' or 'whole configuration'. The Gestalt psychologists proposed a number of innate grouping principles that enable us to perceive forms. Some of these, based on the work of Max Wertheimer (1923) are presented below.

Principle of Proximity (or nearness)

This principle states that the elements in a scene which are closer together will be seen as belonging together in a group. This is illustrated in Figure 9. The consequence of the greater proximity or nearness of the dots on the vertical axis is that we perceive the dots as being organised into columns rather than rows.

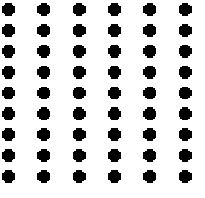


Figure 9. Column of dots

If the scene is altered so that the dots are closer together on the horizontal axis, then we perceive a series of rows, as illustrated in Figure 10.

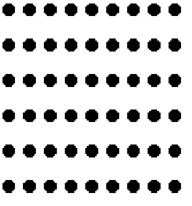


Figure 10. Rows of dots

Principle of Similarity

This principle states that entities which share visual characteristics such as size, shape or colour will be perceived as belonging together in a group. For example, in Figure 11, we perceive columns of shapes (rather than rows). In fact, the shapes are equidistant on both the horizontal and vertical axes. It is due to our innate predisposition to organise based, here, on similarity that similar shapes (squares or circles) are grouped together and, consequently, are perceived as columns.

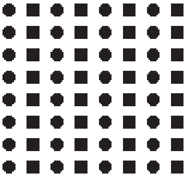


Figure 11. Columns of shapes

Principle of Closure

This principle holds that incomplete figures are often 'completed', even when part of the perceptual information is missing. For instance, in Figure 12 we perceive a circle, even though the 'circle' is incomplete. That is, there is a tendency to close simple figures, by extrapolating from information which is present.

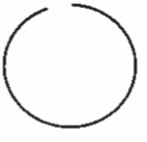


Figure 12. An incomplete figure subject to perceptual closure

A related perceptual process is illustrated by the following. In Figure 13, a white triangle is perceived as being overlaid on three black circles, even though the image could simply represent three incomplete circles. This phenomenon is known as the perception of *subjective* or *apparent contours*. It resembles closure, in so far as there is the appearance of edges across a blank area of the visual field.

Figure 13. Subjective contour: A white triangle

Principle of Good Continuation

This principle states that human perception has a preference for continuous figures. This is illustrated in Figure 14. Here, we perceive two unbroken rectangles, one passing behind another, even though this is not what we actually see. In fact, the shaded rectangle is obscured by the first, so we have no direct evidence that the shaded area represents one continuous rectangle rather than two separate ones.

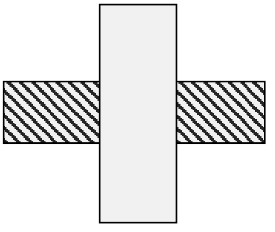


Figure 14. Two rectangles

Principle of Smallness

The Principle of Smallness states that smaller entities tend to be more readily perceived as figures than larger entities. This is illustrated in Figure 15. We are more likely to perceive a black cross than a white cross, because the black shading occupies a smaller proportion of the image.

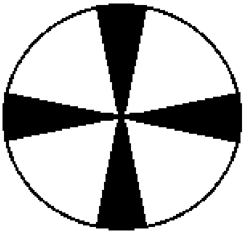


Figure 15. A black cross

Principle of common fate

The final principle I consider here is the Principle of Common Fate. This states that elements that move in the same direction are perceived as being related to one another. For instance, assume that we have two rows of 4 small squares. If the middle two squares from the bottom row begin to move down the page, as depicted by the arrows in Figure 16, they are perceived as belonging together and thus form a separate group from those that remain stationary.

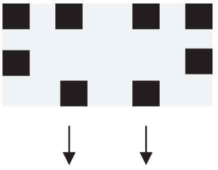


Figure 16. Motion in the same direction

The Gestalt grouping principles I have surveyed conform to the general Gestalt Principle of Good figure, also known as the *Law of Prägnanz***.** This states that we tend to perceive the simplest and most stable of the various perceptual possibilities.

While I have primarily focused in this section on visual perception, it is important to emphasise that the principles I have discussed, both figure-ground and grouping principles, manifest themselves in other modalities. For instance, Kennedy (1983; Kennedy and Domander 1984), present evidence that figure-ground perception, including the analogues to the ambiguous profile/vase illusion occur in the tactile (touch) modality, based on experiments involving raised-line drawings of reversible figures. Similarly, Bregman (1990) has argued that the Gestalt principles apply equally to auditory scene analysis. He makes the point, for instance, that the ability to perceive a series of musical notes as forming a tune is an instance of a gestalt *par excellence*.

3.3 The perception of movement

Our ability to detect movement is essential for the survival of the species. Below I discuss a number of different systems for the detection of motion, and different kinds of motion. I begin with the visual detection of motion.

Two visual systems

Motion detection appears to have evolutionary priority over shape detection (Gregory 1998). Indeed, as observed by Gregory, the evolution of the eye emerged in the first place in order to detect motion. Indeed, only eyes relatively high up the evolutionary scale produce stimulus in the absence of motion. The evolutionary development of vision and the detection of motion are represented in the human eye:

> The edge of our retinas are sensitive only to movement. You can see this by getting someone to wave an object around at the side of your visual field where only the edge of the retina is stimulated. Movement is seen, but it is impossible to identify the object, and there is no colour. When movement stops the object becomes invisible. This is as close as we can come to experiencing primitive vision. The extreme edge of the retina is even more primitive: when it is stimulated by movement we experience nothing; but a reflex is initiated, rotating the eye to bring the moving object into central vision… (Gregory 1998: 98)

The human visual system involves eyes which can move in the head, as when we keep our heads stationary and move our eyes from side to side or up and down. Consequently, our visual system has two distinct ways of detecting motion.

The first involves *image-retina movement*. This involves the eye ball remaining stationary. In this situation the image of moving objects run sequentially across adjacent photoreceptors on the retina. That is, the detection of movement occurs as different photoreceptors are understood by the brain as relating to different locations in space. The second method involves *eye-head movement*. This relates to movement of the eyes in the eye-ball socket when we follow an object in motion. In this situation, an object is not running across different photoreceptors as the eye moves in order to track the object. Rather, information from the eye muscles, which stretch in response to the movement of the eye, is understood by the brain as relating to motion of the tracked object.

Optic flow

In normal vision when we move our eyes, the world remains stable. That is the visual world doesn't spin around. This follows as, during normal eye movements, signals from the image-retina and eye-head systems cancel each other out, such that the world is perceived as stable. While the two visual systems just described relate to the detection of the movement of objects, another source of movement detection comes from the way in which the human experiencer moves about during the world. As we move around the location from which we view our environment changes. The consequence of this is that there is a continuous change in the light stimulus which is projected on the retina. Following the pioneering work of psychologist James Gibson (e.g., 1986), this changing stimulus is known as *optic flow*.

Optic flow relates to a radial pattern which specifies the observer's direction of self-motion and is essential for successful navigation through the environment. As we travel through the world, and as we approach objects, they appear to move towards us, flowing past behind us as we move beyond them. Moreover, different objects at different points in the visual field appear to move towards and past us at different rates. For instance, imagine sitting on a train and travelling through the countryside. Distant objects such as clouds or mountains appear to move so slowly that they are stationary. Closer objects such as trees appear to move more quickly while very close objects appear to whiz by in a blur. This motion, the optic flow pattern, provides important cues as to distance. Moreover, the optic flow varies depending on the relationship between viewing angle and direction of travel. For instance, objects which are dead-ahead and thus centred in the visual field will appear to remain stationary, while objects which are more peripheral in the visual field will appear to move more rapidly. However, because the edges of centred objects will not be in foveal vision, the edges will have optic flow associated with them. Thus, optic flow patterns provide important information about both distance and direction of travel.

Biological motion

The requirement of being able to rapidly detect the motor activities of humans and other organisms is essential for survival. Indeed, under certain lighting conditions, such as at dusk, details relating to the precise nature of the animal in question may not be readily discernable, especially if the animal is distant. Accordingly, humans have evolved an ability to detect what Johansson (1973) terms *biological motion*. Based purely on movement cues, we can quickly distinguish biological from non-biological motion. Moreover, humans can readily distinguish between different types of biological motion based solely on movement cues, for example, running versus jogging versus walking versus jumping, and so on. Each gait represents a gestalt constructed from a sequence of pendulum-like motions, specific to each activity type.

Evidence for this ability comes from the work of the visual psychologist Gunnar Johansson. He videotaped actors in complete darkness. The actors had *point-light displays* (points of light) fixed at ten main body joints which served as the only illumination. This eliminated all non-movement cues, such as the body contours of the actors. Subjects were then asked to identify biological motion and the motor activities engaged in by the actors. Johansson found that in the absence of motion subjects failed to recognise the point light displays as representing a human form. However, with movement the subjects vividly perceived human motion. In other words, subjects related moving lights in order to perceive human movement, and moreover, were able to identify the *pattern of movement*, that is, the kind of movement being engaged in.

3.4 The perception of three-dimensional space

In this section I briefly review how the brain constructs (three dimensional) space, that is, depth, when the retina is a two-dimensional surface. In other words, where does the third dimension come from? I consider below a number of cues that the brain extracts from the visual stimuli in order to construct our experience of (three-dimensional) space.

While depth and distance can be constructed on the basis of a range of visual (and other) stimuli, including auditory cues, and the optic flow patterns described above, an important means of obtaining depth information comes from *binocular cues*. This relates to the spatial stimuli provided by virtue of having two eyes.

The eyes are separated by about 6.5 cm (Gregory 1998). The consequence of this is that each eye sees a different view. As Gregory observes, '[t]his can be seen clearly if each eye is closed alternately. Any near object will appear to shift sideways in relation to more distant objects and to rotate slightly when each eye receives its view.' (Ibid.: 60). The difference between the two retinal images is known as *binocular disparity*, and gives rise to the perception of depth or *stereoscopic vision*. However, stereoscopic vision only applies to objects which are quite near. This follows as binocular disparity reduces the further away an object is. As Gregory notes, '[w]e are effectively one-eyed for objects further than about 100 metres.' (Ibid.: 60). In other words, depth is a consequence of binocular rather than *monocular* (one-eyed) vision.

4 Cognitive maps

In this section I review in more detail the sorts of spatial representations that the brain constructs from the sensory systems and perceptual stimuli described in previous sections. While I have examined distinct sensory systems, in practice perceptual information from a range of modalities is integrated in order to form spatial or cognitive maps. These are complex mental representations which facilitate navigation and moreover, are necessary for the emergence of the concepts of place and location. The concepts of place and location are independent of the entities and objects which occupy specific places or locations. That is, without a cognitive mapping ability which allows us to perceive places and locations independent of the objects which occupy them we would have no means of understanding these concepts. Accordingly, the concepts PLACE and LOCATION are a consequence not of such notions being an inherent aspect of an objective reality, but rather derive from innate cognitive mapping abilities, and particularly our ability to construct spatial maps independently of our egocentric spatial location, as discussed below.

4.1 Egocentric versus allocentric representations

There are two main sorts of spatial cognitive reference frames manifested by humans and many other species. These are *egocentric* representations and *allocentric* representations. In this section I briefly introduce cognitive reference frames of both these sorts.

There is good neurobiological evidence that humans, along with other mammals, maintain multimodal cognitive spatial 'maps' in the parietal cortex (recall Figure 2). The distinguishing feature of egocentric 'maps' is that they represent objects in space with respect to the organism, or part of the organism, such as the organism's hand, body or head. This follows as cognitive 'maps' of this kind represent space in *topographic fashion*. That is, neighbouring areas of neural space represent neighbouring regions of space in the world of the perceiving organism, with respect to the organism which serves as reference point or *deictic centre* for organising the location of the represented objects and regions of space. As regions of space are organised with respect to the organism, spatial maps of this kind are termed *egocentric* representations.

In addition, there is a second kind of spatial representation which is *allocentric* (or other-focused) in nature. These representations, which are more appropriately thought of in terms of maps (for reasons I shall discuss below), integrate information derived from the egocentric spatial representations. Crucially, however, the allocentric mapping ability represents space, and spatial regions independently of the momentary location of the organism. That is, entities and objects, and the locations of objects are related to one another independently of the ego. This system, which is located in the hippocampal region of the brain (O'Keefe and Nadel 1978) represents place, direction and distance information, rather than object details.

4.2 The hippocampus and the human cognitive mapping ability

In now classic work, neurobiologists John O'Keefe and Lynn Nadel (1978) show not only that i) humans have an objective or absolutive spatial framework in which the entities of our experience are located, but also that, ii) this ability is innate, and along with other mammals is associated with the brain region often implicated in motor function: the hippocampus. According to O'Keefe and Nadel, this allocentric mapping system provides 'the basis for an integrated model of the environment. This system underlies the notion of absolute, unitary space, which is a non-centred stationary framework through which the organism and its egocentric spaces move.' (Ibid.: 2). This hippocampal mapping system consists of two major subsystems, a *place system* and a *misplace system*.

The place subsystem is a memory system that allows the organism to represent places in its environment and crucially to relate different locations with respect to each other. That is, the place system allows the organism to represent relationships between different locations without having to physically experience the spatial relations holding between distinct places. In other words, humans, like many other organisms, can compute distances, and other spatial relations between distinct places such as directions, without having to physically experience the spatial relationships in question. Such a cognitive mapping ability is a consequence of the allocentric place subsystem.

The second subsystem to make up the allocentric cognitive mapping ability, the misplace system, facilitates and responds to exploration. That is, it allows new information experienced as a consequence of exploration to be incorporated into the allocentric map of the organism's environment. It thereby allows the organism to relate specific objects and entities to specific locations, and to update the cognitive map held in the place system based on particular inputs (cues) and outputs (responses). Thus, O'Keefe and Nadel demonstrate two things. Firstly, three-dimensional Euclidean space is, in a non-trivial sense, imposed on perceptual experience by the human mind. Secondly, the notion of all-embracing continuous space, 'out there', which 'contains' objects and other entities, as maintained by the misplace system, is in fact a consequence of first being able to represent locations in an allocentric (i.e., a non-egocentric) fashion, as captured by the place subsystem. In other words, our innate ability to form absolutive cognitive maps of our spatial environment is a prerequisite to experiencing objects and the motions they undergo.

4.3 Maps versus routes

In order to illustrate the distinction between egocentric and allocentric spatial mapping abilities, O'Keefe and Nadel provide an analogy which I briefly discuss here. The analogy relates to the geographic distinction between routes versus maps. In geographic terms, a route constitutes a set of instructions which directs attention to particular objects in egocentric space. That is, routes are inflexible, identifying landmarks in order to guide the traveller, and thus do not allow the traveller freedom of choice. Put another way, routes are *guide-post based*. Moreover, routes are goal-oriented, focused on facilitating travel from a specific, pre-specified location to another. In this, routes correspond to egocentric cognitive representations.

In contrast, maps are, in geographic terms, representations of part of space. A map is constituted of places, and the places which the map represents are systematically connected and thus related to each other. Moreover, and crucially, the places captured by the map are not defined in terms of the objects which may occupy a particular location. That is, and unlike routes, maps are not guide-post based. Thus, maps capture space that is held to exist independently of the objects which may be located at particular points in space. Crucially, a map is a flexible representation, which can be used for a range of purposes. In related fashion, this notion of a map is presented as an analogy of the allocentric cognitive mapping ability that many organisms, including humans, possess.

While then map-like representations of the environment are constructed by humans, as well as by other species, it is far from clear, in neurological terms, what the nature of these representations are. Nevertheless, it is by now well established that humans do possess complex information structures which can be used to generate highly-detailed map-like representations, which can be used for a range of behaviours. Indeed, an important finding to have emerged is that place memory has a high information capacity, and can be permanently modified by a single experience. Moreover, experiments reported on by O'Keefe and Nadel reveal that this mapping ability can be used to construct maps in a highly flexible and efficient manner.

Finally, I reiterate that the ability to represent space in an allocentric fashion, i.e., map-like representations, is a trait common to a wide variety of organisms. As O'Keefe and Nadel observe, 'The ability of many animals to find their way back to their nests over large distances would appear to be based on some type of mapping system' (Ibid.: 63). Obvious examples include the migratory and homing behaviour exhibited by many kinds of birds. Indeed, a robust finding from studies on homing pigeons is that they are able to find their way 'home' using novel routes from new release sites. Such abilities would appear to require a cognitive mapping ability.

5 Primitive spatial concepts

In this section I turn to an examination of spatial concepts and the way in which spatial concepts are derived (or redescribed) from spatial experience. I focus here on the notion of the image schema. Image schemas were first proposed by cognitive linguists (e.g., Johnson 1987, 2007; Lakoff 1987; see Evans and Green 2006 for a review), and represent a rudimentary conceptual building block derived from *embodied experience* (discussed further below). This notion has been subsequently adopted by a range of other cognitive scientists in their work (see papers and references in Hampe 2005). In particular, the notion of the image schema has been developed in the influential work of developmental psychologist Jean Mandler (e.g., 2004) in her work on how conceptual development takes place.

5.1 Embodiment and experience

I begin this brief overview of the image schema by first introducing the role of *embodiment* in the formation of concepts. Due to the nature of our bodies, including our neuro-anatomical architecture, we have a species-specific view of the world. In other words, our construal of 'reality' is mediated, in large measure, by the nature of our embodiment. One obvious example of the way in which embodiment affects the nature of experience relates to biological morphology (i.e., body parts). This, together with the nature of the physical environment with which we interact, determines other aspects of our experience. For instance, while gravity is an objective feature of the world, our experience of gravity is determined by our bodies and by the ecological niche we have adapted to. For instance, hummingbirds – which can flap their wings up to fifty times per second – respond to gravity in a very different way from humans. They are able to rise directly into the air without pushing off from the ground, due to the rapid movement of their wings.

The fact that our experience is embodied – that is, structured in part by the nature of the bodies we have and by our neurological organisation – has consequences for cognition. In other words, the concepts we have access to and the nature of the 'reality' we think and talk about are a function of our embodiment – the phenomenon of *variable embodiment*. That is, we can only talk about what we can perceive and conceive, and the things that we can perceive and conceive derive from embodied experience. From this point of view, the human mind must bear the imprint of embodied experience. This thesis is known as the *thesis of embodied cognition*. This position holds that conceptual structure – the nature of human concepts – is a consequence of the nature of our embodiment and thus is embodied.

5.2 Image schemas

The theoretical construct of the image schema was developed by Mark Johnson in his now classic 1987 book, *The Body in the Mind*. Johnson proposed that one way in which embodied experience manifests itself at the cognitive level is in terms of image schemas. These are rudimentary concepts like CONTACT, CONTAINER and BALANCE, which are meaningful because they derive from and are linked to human *pre-conceptual experience*. This is experience of the world directly mediated and structured by the human body.

The term 'image' in 'image schema' is equivalent to the use of this term in psychology, where *imagistic* experience relates to and derives from our experience of the external world. Another term for this type of experience is sensory experience, because it comes from sensory-perceptual mechanisms that include, but are not restricted to, the visual system.

According to Johnson (1987) there are a number of properties associated with image schemas which I briefly review below.

Image schemas are pre-conceptual in origin

Image schemas such as the CONTAINER schema are directly grounded in embodied experience. This means that they are pre-conceptual in origin. Mandler (2004) argues, discussed further in the next section, that they arise from sensory experiences in the early stages of human development that precede the formation of concepts. However, once the recurrent patterns of sensory information have been extracted and stored as an image schema, sensory experience gives rise to a conceptual representation. This means that image schemas are concepts, but of a special kind: they are the foundations of the conceptual system, because they are the first concepts to emerge in the human mind, and precisely because they relate to sensory-perceptual experience, they are particularly schematic. Johnson argues that image schemas are so fundamental to our way of thinking that we are not consciously aware of them: we take our awareness of what it means to be a physical being in a physical world very much for granted because we acquire this knowledge so early in life, certainly before the emergence of language.

Image schemas form the basis of word senses

Concepts lexicalised by words such as prepositions, for instance, *in*, *into, out, out of* and *out from* are all thought to relate to the CONTAINER schema: an abstract image schematic concept that underlies all these much more specific *senses* – the semantic pole associated with lexical forms (see Tyler and Evans 2003).

The CONTAINER image schema is diagrammed in Figure 17. This image schema consists of the structural elements interior, boundary and exterior: these are the minimum requirements for a CONTAINER (Lakoff 1987). The *landmark* (LM), represented by the circle, consists of two structural elements, the interior – the area within the boundary – and the boundary itself. The exterior is the area outside the landmark, contained within the square. The container is represented as the landmark because the boundary and the exterior together possess sufficient Gestalt properties (e.g., closure and continuity) to make it the figure, while the exterior is the ground (recall my discussion of Gestalt principles above).

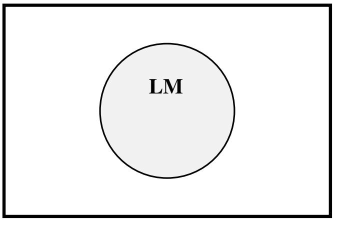


Figure 17. CONTAINER image schema

Although Figure 17 represents the basic CONTAINER schema, there are a number of other image schemas that are related to this schema, which give rise to distinct concepts related to containment. For instance, let's consider one variant of the CONTAINER schema lexicalised by *out*. This image schema is diagrammed in Figure 18 and is illustrated with a linguistic example. The diagram in Figure 18 corresponds to example (1). The *trajector* (TR) *Fred*, which is the entity that undergoes motion, moves from a position inside the LM to occupy a location outside the LM. The terms 'TR' and 'LM' derive from the work of Langacker (e.g., 1987), and relate to the Gestalt notions of figure and ground respectively.

(1) Fred went out of the room

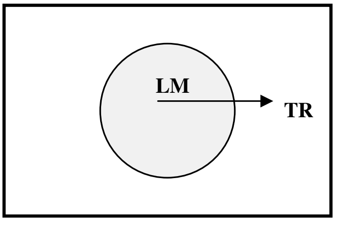


Figure 18: Image schema for OUT

The image schema shown in Figure18 represents a concept that is more specific and detailed than the image schema diagrammed in Figure 17, because it involves motion as well as containment. This shows that image schemas can possess varying degrees of schematicity, where more specific image schemas arise from more fundamental or schematic ones.

Image schemas derive from interaction

As image schemas derive from embodied experience, they derive from the way in which we interact with the world. To illustrate this idea, consider the image schema for FORCE. This image schema arises from our experience of acting upon other entities, or being acted upon by other entities, resulting in the transfer of motion energy. Johnson illustrates the *interactional derivation* of this image schema – how it arises from experience – as follows:

> [F]orce is always experienced through interaction. We become aware of force as it affects us or some object in our perceptual field. When you enter an unfamiliar dark room and bump into the edge of the table, you are experiencing the interactional character of force. When you eat too much the ingested food presses outwards on your taughtly stretched stomach. There is no schema for force that does not involve interaction or potential interaction. (Johnson 1987: 43)

Image schemas are inherently meaningful

As image schemas derive from interaction with the world, they are inherently meaningful. Embodied experience is inherently meaningful in the sense that embodied experiences have predictable consequences. To illustrate, imagine a cup of coffee in your hand. If you move the cup slowly up and down, or from side to side, you expect the coffee to move with it. This is because a consequence of containment, given that it is defined by boundaries, is that it constrains the location of any entity within these boundaries. In other words, the cup exerts force-dynamic control over the coffee. This kind of knowledge, which we take for granted, is acquired as a consequence of our interaction with our physical environment. For example, walking across a room holding a cup of coffee without spilling it actually involves highly sophisticated motor control that we also acquire from experience. This experience gives rise to knowledge structures that enable us to make predictions: if we tip the coffee cup upside-down, the coffee will pour out.

Image schemas are analogue representations

Image schemas are *analogue* representations deriving from experience. The term 'analogue' means image schemas take a form in the conceptual system that mirrors the sensory experience being represented. Because image schemas derive from sensory experience, they are represented as summaries of *perceptual states*, which are recorded in memory. However, what makes them conceptual rather than purely perceptual in nature is that they give rise to concepts that are consciously accessible (Mandler 2004). In other words, image schemas structure (more complex) lexical concepts.

Image schemas can be internally complex

Image schemas are often, perhaps typically, comprised of more complex aspects that can be analysed separately. For example, the CONTAINER schema is a concept that consists of interior, boundary and exterior elements. Another example of a complex image schema is the SOURCE-PATH-GOAL or simply PATH schema. Because a path is a means of moving from one location to another, it consists of a starting point or SOURCE, a destination or GOAL and a series of contiguous locations in between, which relate the source and goal. Like all complex image schemas, the PATH schema constitutes an *experiential gestalt*: it has internal structure, but emerges as a coherent whole.

One consequence of internal complexity is that different components of the PATH schema can be referred to. This is illustrated in example (2), where the relevant linguistic units are bracketed. In each of these examples, different components of the path are profiled by the use of different lexical items.

(2) a. SOURCE
John left [England]

b. GOAL
John travelled [to France]

c. SOURCE-GOAL
John travelled [from England] [to France]

d. PATH-GOAL
John travelled [through the Chunnel] [to France]

e. SOURCE-PATH-GOAL
John travelled [from England] [through the Chunnel] [to France]

Image schemas are not mental images

If you close your eyes and imagine the face of your mother or father, partner or lover, what results is a mental image. Image schemas are not the same as mental images. Mental images are detailed, and result from an effortful and partly conscious cognitive process that involves recalling visual memory. Image schemas are schematic, and therefore more abstract in nature, emerging from ongoing embodied experience. This means that you can't close your eyes and 'think up' an image schema in the same way that you can 'think up' the sight of someone's face or the feeling of a particular object in your hand.

Image schemas are multi-modal

Image schemas derive from experiences across different modalities (different types of sensory experience), and hence are not specific to a particular sense. In other words, image schemas are abstract patterns arising from a range of perceptual experiences, and as such are not available to conscious introspection. For instance, blind people have access to image schemas for CONTAINERS, PATHS, and so on, precisely because the kinds of experiences that give rise to these image schemas rely on a range of sensory-perceptual experiences in addition to vision, including hearing, touch, and our experience of movement and balance.

Image schemas form the basis for abstract thought

Lakoff (1987, 1990, 1993) and Johnson (1987) have argued that rudimentary embodied concepts of this kind provide the conceptual building blocks for more complex concepts, and can be systematically extended to provide more abstract concepts and conceptual domains with structure. According to this view, the reason we can talk about being *in* states like love or trouble (3) is because abstract concepts like LOVE are structured and therefore understood by virtue of the fundamental concept CONTAINER. In this way, image schematic concepts serve to structure more complex concepts and ideas.

(3) a. John is in love.
b. Jane is in trouble.
c. The government is in a deep crisis.

According to Johnson, it is precisely because containers constrain activity that it makes sense to conceptualise POWER and all-encompassing states like LOVE or CRISIS in terms of CONTAINMENT.

5.3 Perceptual meaning analysis

The developmental psychologist Jean Mandler (e.g. 1992, 1996, 2004) has made a number of proposals concerning how image schemas might arise from embodied experience. Starting at an early age infants attend to objects and spatial displays in their environment. Mandler suggests that by attending closely to such spatial experiences, children are able to abstract across similar kinds of experiences, finding meaningful patterns in the process. For instance, the CONTAINER image schema is more than simply a spatio-geometric representation. It is a 'theory' about a particular kind of configuration in which one entity is supported by another entity that contains it. In other words, the CONTAINER schema is meaningful because containers are meaningful in our everyday experience.

Mandler (2004) describes the process of forming image schemas in terms of a redescription of spatial experience via a process she labels *perceptual meaning analysis* (Mandler 2004). This process results from children associating functional consequences with spatial displays. That is, image schemas emerge by virtue of analysing spatial displays of various sorts as relating to the functional consequences with which they are correlated. For example, we saw above that a consequence of coffee being located in a coffee cup is that the coffee moves with the cup. That is, containment has functional consequences in terms of containing, supporting and constraining the location of the entity contained. Thus, the distinction between percepts and concepts such as image schemas is that image schemas encode functional information, that is meaning. As Mandler observes, '[O]ne of the foundations of the conceptualizing capacity is the image schema, in which spatial structure is mapped into conceptual structure' (Mandler 1992: 591). She further suggests that 'Basic, recurrent experiences with the world form the bedrock of the child's semantic architecture, which is already established well before the child begins producing language' (Mandler 1992: 597). In other words, it is experience, meaningful to us by virtue of our embodiment, that forms the basis of many of our most fundamental concepts.

References

Biederman, Irving. (1987) Recognition-by-components: A theory of human image understanding. *Psychological Review* 94: 115–147.

Bregman, Albert. (1990) *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Evans, Vyvyan and Melanie Green. (2006) *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Gibson, James. (1986) *The Ecological Approach to Visual Perception*. Mahwah, NJ: Lawrence Erlbaum.

Milner, David and Melvyn Goodale. (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

Gregory, Richard. (1998) *Eye and Brain.* (Fifth edition.) Oxford University Press.

Hampe, Beate. (ed.) (2005) *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlin: Mouton de Gruyter.

Hurford, James. (2003) The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26: 261–283.

Johansson, Gunnar. (1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* 14: 201–211.

Johnson, Mark. (1987) *The Body in the Mind*. Chicago: University of Chicago Press.

Johnson, Mark. (2007) *The Meaning of the Body*. Chicago: University of Chicago Press.

Julesz, Bela. (1981) Textons: The elements of texture perception and their interactions. *Nature* 290: 91–97.

Kennedy, John. (1983) What can we learn about pictures from the blind? *American Scientist* 71: 19–26.

Kennedy, John and Ramona Domander. (1984) Pictorial foreground-background reversal reduces tactual recognition by the blind. *Journal of Visual Impairment and Blindness* 78: 215–216.

Lakoff, George. (1987) *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, George. (1990) The invariance hypothesis: Is abstract reason based on image schemas? *Cognitive Linguistics* 1(1): 39–74.

Lakoff, George. (1993) The contemporary theory of metaphor. In A. Ortony (ed.) *Metaphor and Thought* 202–251. (Second edition.) Cambridge: Cambridge University Press.

Landau, Barbara and Ray Jackendoff. (1993) 'What' and 'where' in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16: 217–265.

Langacker, Ronald. (1987) *Foundations of Cognitive Grammar.* (Vol. I) Stanford, CA: Stanford University Press.

Mandler, Jean. (1992) How to build a baby II. Conceptual primitives. *Psychological Review* 99: 567–604.

Mandler, Jean. (1996) Preverbal representation and language. In P. Bloom, M. Peterson, L. Nadel and M. Garrett (eds) *Language and Space* 365–384. Cambridge, MA: MIT Press.

Mandler, Jean (2004) *The Foundations of Mind: Origins of Conceptual Thought*. Oxford: Oxford University Press.

O'Keefe, John and Lynn Nadel. (1978) *The Hippocampus as a Cognitive Map.* Oxford: Oxford University Press. [Available on-line at www.cognitivemap.net/]

Tyler, Andrea and Vyvyan Evans. (2003) *The Semantics of English Prepositions: Spatial Scenes, Embodied Experience and Cognition*. Cambridge: Cambridge University Press.

Ungerleider, Leslie and Mortimer Mishkin. (1982) Two cortical visual systems. In D. Ingle, M. Goodale, and R. Mansfield (eds) *Analysis of Visual Behavior* 549–586. Cambridge, MA: MIT Press.

Wertheimer, Max. (1923 [1938]) Untersuchungen zur Lehre von der Gestalt II. *Psycologische Forschung* 4: 301–350. Laws of organization in perceptual forms. In W. Ellis (ed.) (1938) *A Source Book of Gestalt Psychology* 71–88. London: Routledge.

Part II

The interaction between language and spatial cognition

2 Language and space: momentary interactions

Barbara Landau, *Banchiamlack Dessalegn and Ariel Micah Goldberg*

Knowledge of language and space constitute two of our most fundamental ways of knowing the world, and the idea that these systems of knowledge interact is not new. It has inspired work from diverse intellectual circles, including formal approaches to language (Fillmore, 1997; Gruber, 1976; Jackendoff, 1983; Talmy, 1983; Langacker, 1986; Lakoff, 1987); theoretical and empirical studies of language learning (Bowerman, 1973; Brown, 1973; E. Clark, 1973; H. Clark, 1973; Mandler, 1992); studies of the relationship between language and thought (Whorf, 1956; Levinson, 1996; Gleitman and Papafragou, 2005; Munnich, Landau and Dosher, 2001; Hermer and Spelke, 1996), and even theories of the way in which evolution could have built on non-linguistic structures to build a human language (O'Keefe and Nadel, 1978; Hauser, Chomsky and Fitch, 2002). The diversity of interest in the interaction between spatial cognition and language represents the widespread assumption that, at some level, language must map onto our visual and spatial representations of the world. How else would we be able to talk about what we see?

At least two important consequences follow from this mapping. The first is that pre-linguistic representations of space could provide a crucial developmental link for language learning; children might be able to bootstrap their way into the linguistic system by capitalizing on homologous structures in spatial cognition and language. Students of language learning have long assumed that the infant's spatial representations play an important role in allowing him or her to break into the language system, and indeed, there is growing evidence that skeletal spatial representations bear a formal similarity to skeletal linguistic representations (see, e.g. Fisher, 2000; Lakusta, Wagner, O'Hearn and Landau, 2007; Lakusta and Landau, 2005). The second consequence is that language, once acquired, might come to modulate our spatial representations. In this chapter, we focus on the latter effects.

The classical approach to this issue is best known through the views of Benjamin Whorf (1956), who proposed that language shapes thought. Whorf's original observations focused on the coding of time among the Hopi, but was quickly taken up by anthropologists and linguists examining other areas of perception and cognition, notably color (e.g. Berlin and Kay, 1969; Kay and Kempton, 1984). Although experimental studies during the 1960's and 1970's seemed to have settled the Whorfian question of whether language shapes thought (in favor of a resounding 'no'; see Brown, 1976 for review), the same question has recently re-emerged, with a flurry of new research by scientists claiming victory on both sides (see Gentner and Goldin-Meadow, 2003; Majid, Bowerman, Kita, Haun and Levinson, 2004; Munnich et al., 2001; Gleitman and

Papafragou 2005, for recent surveys). Some have strongly argued that the structure of one's spatial lexicon, morphology, and syntax have profound repercussions on non-linguistic representations– i.e. spatial 'thought' (Levinson, 1996; Gentner, 2001). Others have argued quite persuasively that cross-linguistic variation in spatial terminology causes no permanent or substantive effects on one's non-linguistic spatial understanding (Munnich et al., 2001; Malt et al., 2003; Gleitman and Papafragou, 2005; Li and Gleitman, 2002).

The main purpose of our review is to lay out evidence suggesting a new solution to this impasse. In particular, we will suggest that a straightforward 'yes' or 'no' to the question of whether language changes spatial thought is too simplistic. Rather, we will present a different twist to the issue, suggested by some newer developments in thinking about the interaction between language and spatial cognition. Specifically, we will review evidence that language – once acquired – can strongly modulate our non-linguistic spatial representations, but that much of this is done *in the moment* of carrying out a specific task, and does not result in permanent organizational change to spatial representation. The functions of language that we will discuss are in a sense more 'shallow'– more closely related to the immediate on-line time course within which our acquisition, comprehension and production take place. The effects occur in a brief time window, and therefore might be viewed by some as 'mere' temporary mechanisms that operate as we speak and hear sentences, and as we process visual information. However, we will argue that these temporally brief interactions can play a powerful role by engaging language to modulate and enhance what is done by the visual system.

Our chapter is organized as follows. First, we provide a brief review of some of the differences between our representations of language and space, considering some of the ways in which the systems differ from each other. Where the systems overlap, we can ask whether and how language modulates our spatial representations. We focus on two possibilities. One is that language modulates attention because it is inherently *selective:* Languages choose to encode certain spatial properties and not others, directing visual attention accordingly. The second possibility is that language *enriches* visual-spatial representations. The idea here is that language permits us to go beyond what is robustly available to our spatial system of representation, expanding representational power.

Both of these possibilities have been used by researchers to test the Whorfian hypothesis. The cross-linguistic differences in what languages choose to encode and the possibility that language can enrich visual representations have been discussed in the context of permanent changes to our capacity to carry out spatial problems. Despite stronger claims that have been made about the permanent organizing effects of language, we will argue that both of these effects (selectivity and enrichment) take place in a limited time frame. We will review evidence that there are immediate dynamic effects of language on visual-spatial representations, raising the possibility that the powerful effects of language are more time-limited than proponents of the Whorfian hypothesis would suggest.

1 Specialization of language and spatial representation

Several observations suggest that our linguistic and spatial systems are not redundant but are complementary. To start, the primitives in each system are unique, as are their combinatorial rules. Language traffics in primitive symbolic units such as noun and verb, and configurations of these give rise to semantic and syntactic functions such as agent, patient, subject, object. In contrast, the spatial system traffics in primitives such as shapes, objects, locations, landmarks, geometric layouts, angles and directions, all represented within different spatial reference systems. The rules of combination for spatial systems are unique as well. For example, objects are represented as sets of parts that are arranged in terms of hierarchical relationships and layouts are represented in terms of elements and their geometric arrangements (e.g. Marr, 1982; Gallistel, 1990).

The differences in formal properties are accompanied by differences in function. Jackendoff (1987) provides a nice example of the differential power of language and visual-spatial representations. He considers the case of ducks, geese, and swans (see Figure 1). Exemplars of these species clearly differ in some set of geometric properties that are naturally represented by the visual system: Swans have longer necks than geese, and their overall proportions are somewhat different. These differences in overall shape of the two animals, including differences in the length of their necks, are well-captured in visual-spatial representations of objects (e.g. Marr and Nishihara, 1992). But they are not well-captured in the basic lexicons of languages. To the extent that the overall shapes of objects are captured in the lexicon or in morphology, the geometric properties tend to be coarse, such as 'long thin' or 'round' (commonly encoded by classifiers).

Figure 1. Swans and Geese. An example of the differential power of language vs. visual-spatial representations. The visual differences between swans and geese are easily represented by the visual system but are not well-captured by the basic lexicons of languages (Jackendoff, 1987). On the other hand, language (but not visual representations) naturally captures distinctions such as the difference between types and tokens (e.g. 'a swan', 'that swan'), which are not readily captured by visual representations (see text for discussion).

Faces provide another example. Humans are experts at recognizing faces that differ by subtle attributes such as individual features, spacing, and overall arrangement. We easily encode the differences among enormous numbers of faces, and we are experts at recognizing even faces we have seen rarely. Language provides proper names to encode different individuals, but does not encode in any simple and straightforward way the unique organization of different faces. While we can visually recognize a face quite easily, we are remarkably ineffective in verbally communicating what the face looks like.

Finally, some aspects of more complex layouts show that visual-spatial representations can often capture essential relationships that are difficult or impossible to convey efficiently in language. As Morton (2004) points out, although medical students can learn the structure of the human skeletal system given linguistic descriptions (e.g. 'the hip bone is connected to the femur…'), a diagram of the skeletal system is a more natural way to represent all of the spatial relationships at once. In general, global spatial layouts are naturally captured by visual representations but can be quite underspecified when we use language to describe them.

On the other side, many distinctions are uniquely captured in language but are not, in any obvious way, a part of our visual-spatial representations. For example, the very same object may be named as a unique individual ('that swan'), a member of the particular species ('the swan'), or a member of a superordinate class ('that animal'). The distinction between a 'type' and 'token' representation of the same swan, or the different hierarchical levels to which the same swan can belong cannot be differentiated by anything in the visual-spatial representation, but these are clearly distinguished in language. Other distinctions, such as the difference between 'my book' and 'your book' (given two books that look exactly the same) are naturally made by language, whereas they are not distinguishable in visual-spatial representations.

These examples show that language and visual-spatial representations are best suited to conveying distinct sorts of information. But while each system may most naturally convey certain types of information and not others, it is not the case that their functions are exclusive. Language can also encode information about spatial relationships and it is here that we can ask whether and how language modulates our spatial representations. We turn to the two mechanisms of interest: selectivity and enrichment.

1.1 Selectivity of language

Although language encodes many aspects of our spatial representations, it does not encode everything. Selectivity is pervasive in language, and the particular elements that languages select (and the differences over languages) have been central to debates on whether language causes changes in spatial cognition. We consider two examples: Selecting components of motion events, and selecting among reference systems (which are necessarily engaged in order to represent object location).

Across languages, the structure of simple motion events is typically formalized in terms of several major components, including Figure, Ground (or Reference object), Motion, Manner, and Path (see, e.g. Talmy, 1985). In English, motion verbs tend to encode the motion itself plus the manner, for example, run, skip, hop, dance, swim, fly, etc. English also encodes the Path separately, often as a prepositional phrase that includes two elements: the Path function itself (which describes the geometry of the path) and the reference object(s) (in terms of which the path function is defined). In a simple example, the sentence 'Mary ran to the house' includes a Figure (Mary), Motion (+ Manner, ran), the Path-function (to) and its Reference object (house). Paths are further subdivided into TO paths (which focus on the endpoint), FROM paths (which focus on the starting point), and VIA paths (which focus on the intervening segment) (Jackendoff, 1983).

This general pattern of encoding is not universal, however. Although English (sometimes called a 'Manner' language) encodes the path in a prepositional component separate from the verb, other languages ('Path' languages, e.g. Spanish, Greek) tend to encode the path in the main verb, with the manner portion often encoded in a separate phrase or clause. The difference in tendency to encode manner or path in the main verb is a hallmark of a major typological division across languages (Talmy, 1985), and has been used by scientists to examine whether a language's predominant coding tendencies have effects on their speakers' *non-linguistic* coding of events.

A number of theorists have speculated that the path/manner typological distinction could have major ramifications for the way that people represent events. That is, this linguistic difference could affect non-linguistic functions such as memory for visually perceived events. Simply, if a person's language tends to encode the path in its main verb (rather than a separate phrase), then people should represent and remember the path component of the event more robustly than people whose language tends to encode the manner of motion in the main verb. The reverse should hold for the manner component.

Despite the appealing simplicity of this proposal, it has been difficult to find such *non-linguistic* effects of different languages (Gennari et al., 2002; Papafragou et al., 2002; see also Munnich et al., 2001). A much more modest proposal has been advanced by Slobin (1996), who suggests that people's focus on manner vs. path when they observe events might be explained as a consequence of 'thinking for speaking'. That is, individuals may differentially attend to the various aspects of an event strictly in preparation for *talking* about an event– an effect that would not be surprising, since it is a necessary consequence of preparing to linguistically encode manner that one focus on manner (and similarly for path). We return to this issue in Section 2, where we discuss evidence for attentional modulation while people prepare to *describe* an event but not when they are preparing to simply remember it. This evidence is consistent with the idea of temporary, on-line effects, rather than permanent organizational ones.

The second case of selection concerns the use of *reference systems* in encoding spatial relationships. Consider Figure 2. A spatial (geometric) representation of this layout includes the metric locations of each item relative to some reference system. A natural way of representing the locations would be in terms of a reference system centered on the box itself (i.e. with the origin at the center of the box). From this center, we can derive the exact locations of each object relative to the origin, and relative to each other.

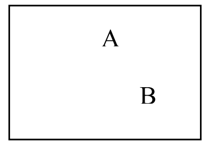


Figure 2. Frames of reference. The locations of A and B can be defined in terms of different reference systems. For example, A and B can both be located relative to a reference system centered on the box itself (i.e. with the origin at the center of the box), in which case, both A and B have specific coordinates in this reference system (or, more coarsely, are 'inside' the box). Or A can be located relative to a reference system centered on B (A is above and left of B), or vice versa (B is below and right of A). Languages can engage several different reference systems (see text for discussion).

But languages do not capture metric relationships in a simple way, a point first elaborated by Talmy (1983). Rather, languages typically have a stock of basic terms that encode what Talmy calls 'topological' relationships – such as whether one object is inside or outside of another, whether it is near or far from the other, etc. Even terms that engage axes (non-topological properties) such as above/below and right/left discount exact metric distance, and rather, encode categories of relationships relative to the axes of a particular reference system.

Like the visual system, languages engage a number of different possible reference systems, with different sets of closed class terms specifying which reference system is being used. The same layout can be mentally represented in terms of different reference systems, thereby changing the spatial interpretation of the layout and hence the particular spatial terms one would choose to describe the inter-object relationships. For example, in Figure 2, we can say that A is near to or to the left of B; or that A is far from or to the right of the box's edge. In the first case, we are representing A in terms of a reference system centered on B; in the second case the reference system is centered on the box's left edge. Terms such as 'right' and 'left' in English are typically used for spatial relationships encoding the two directions along the horizontal axis, with the origin at one's own body (the egocentric system), or at some object or layout in the world (the 'allocentric' system). The same holds for terms above and below, except that they encode the two directions along the vertical axis (again, with the location of the origin defining the reference system). In contrast to these sets of words, terms and phrases such as 'the top/bottom/front/back/side' typically engage an object-centered reference system, that is, a reference system whose origin is centered on an object. Terms 'north/south/east/west' engage a geocentric (earth-centered) reference system, and are usually used (in English) for larger, environmental layouts. Crucially, selection of a particular term assumes that the speaker has selected a particular frame of reference; the hearer will have to adopt the same frame of reference (or be able to translate from the speaker's frame of reference into one the hearer chooses) in order to understand what the speaker has said. The particular word or phrase that the speaker chooses provides information to the hearer about which reference system he has in mind. Additional specification can be made by including the name of the object that serves as the center of the reference frame (e.g. 'above *me*' vs. 'above *the table*').

Although languages generally have the same range of reference systems available for encoding spatial relationships, there are apparently some cross-cultural differences in people's tendency to adopt a given reference frame when they describe locations. Pederson et al. (1998) argued that the speakers of Tzeltal tend to use geocentric frames of reference (or 'absolute' frames in Pederson et al.'s terminology) rather than egocentric or allocentric frames of reference (or 'relative' frames) when describing spatial relationships in relatively small arrays. This is quite different from the general tendency of English speakers: For a tabletop array, English speakers will likely (though not always) use terms 'right/left' (e.g. to my right/left or to the right/left of some object) rather than 'east/west'. In this work, Pederson et al. (1998) claimed that the habitual use of geocentric reference frames by the Tzeltal leads to permanent changes in spatial thought, i.e. non-linguistic problem solving. Pederson's findings and interpretations have been disputed, however, on the basis of logic as well as empirical data (Li and Gleitman, 2002). We will review this dispute and will suggest that the range of empirical findings on reference frame use is best explained within our framework of on-line, temporary modulation – rather than permanent reorganization. We will return to this issue in the next section.

1.2 Enrichment by language

The idea of enrichment is quite different from that of selectivity. Where selectivity emphasizes the power of language in directing attention to a previously available aspect of spatial representation, enrichment suggests that language can add to spatial representations. A recent proposal by Spelke and colleagues (Hermer-Vasquez et al., 1999; Hermer-Vazquez, Moffet and Munkholm, 2001; Spelke et al., 2001) suggests that language can increase the representational power by which we can carry out certain spatial tasks. The particular case they discuss concerns a well-known pattern of error seen in reorientation tasks, in which people are disoriented and then must regain their bearings in space in order to find a hidden object.

The experiments are patterned after research on rats' ability to reorient (Cheng and Gallistel, 1984). In human experiments, a person is typically brought into a small rectangular room (e.g. 4 x 6 feet) that is uniform in color, e.g. with all black walls and ceiling (see Figure 3A). The person is shown an object being hidden in one corner, and is then disoriented (by turning him or her around repeatedly in the center of the room). When asked to search for the object, people ranging from 18 months through adulthood tend to divide their search between the correct corner and the one that is geometrically equivalent, e.g., the long wall to the left of the short wall, as one is facing a corner (see Figure 3). The pattern is quite robust (in these test circumstances), and has been found in species as diverse as rats, chickens, and fish, under some circumstances (see Cheng and Newcombe, 2005, for review). Crucially, even when one of the walls is clearly and uniquely distinguished by color (e.g. one blue, three black walls; see Figure 3B), toddlers and other species do not make use of this information, still producing the geometric error pattern (but see Cheng et al. for some counter examples). The explanation for this pattern, originally proposed by Cheng and Gallistel (1984) is that the reorientation system is modular and encapsulated, operating *only* on geometric layout information (i.e. relative lengths and sense of walls) and not admitting other information that is not relevant to layouts (such as surface color).

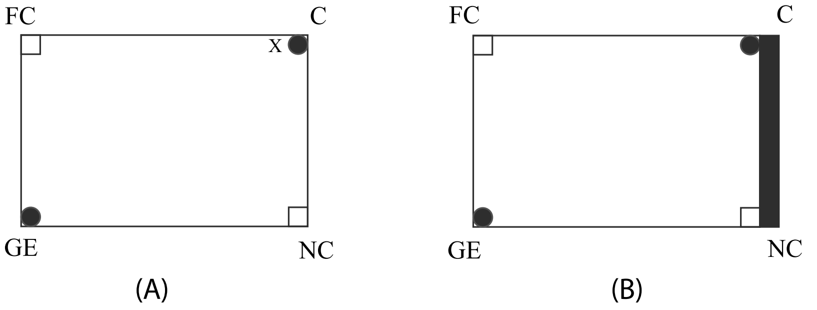


Figure 3. Reorientation by toddlers and adults. (A) When toddlers or human adults are disoriented in an all-black rectangular room, they then reorient themselves and will search for a hidden target equally at the correct location (C) and the geometrically equivalent location (GE). (B) When the room has a single colored wall (shown by thick black bar in 3B), toddlers continue to search at C and GE, whereas older children and adults can use this added information to narrow their search to the correct corner only. Redrawn from Hermer and Spelke (1996).

Hermer and Spelke (1996) found that, unlike toddlers and rats, adults and children from 5 years onward *can* make use of the additional information, and do not commit the geometric error. Rather, they correctly identify the corner where the object was hidden, distinguishing between the geometrically equivalent corners on the basis of wall color. Spelke and colleagues (1996, 1999, 2001) proposed that language plays a crucial role in creating the uniquely human capacity to solve this problem. Specifically, they argue that language has the formal power to combine the outputs of different modules– in this case, the reorientation module (which computes only geometric information) and other systems (such as the one processing surface color of object). If the reorientation module produces two possible solutions, language can then combine the geometric description (e.g. long wall to the left) with non-geometric information (i.e. blue), resulting in the more powerful expression, 'the corner that is *left of the blue wall*'. The larger idea is that language is the only format that allows for combination of properties that are naturally part of separate computational domains.

This hypothesis suggests a powerful role for language: Permanently changing spatial cognition through a new capacity that allows the combination of information from two different systems. In several studies, Hermer, Spelke and colleagues asked whether language could play such a role. One experiment examined adults' ability to solve the reorientation task when carrying out a secondary task, either spatial (shadowing a rhythm) or linguistic (verbal shadowing). If language is the mechanism by which people combine geometric and non-geometric information, then verbal shadowing should cause impairment, and perhaps reversion to the pattern shown by rats and non-verbal humans (i.e. toddlers). This was the pattern found by Hermer-Vasquez et al. (1999). A second study examined the correlation between children's performance in the reorientation task (with blue wall) and their production and comprehension of terms 'left/right' (Hermer-Vazquez, Moffet and Munkholm, 2001). If children's ability to solve the reorientation problem depends on language (specifically, the children's knowledge of 'left' and 'right'), then there should be a strong positive correlation between this linguistic knowledge and performance on the reorientation task. Hermer-Vasquez et al. found that there was a positive correlation between *children's production* accuracy for these terms and their success in the reorientation task. They suggested that this was consistent with the hypothesis that language is crucial to the solution of the task.

There are several problems with this interpretation of the result (see Cheng and Newcombe, 2005, for review). For example, there was no reliable correlation between children's *comprehension* of left/right and success on the reorientation task. This suggests that the role of language may be more limited than was proposed by Hermer-Vasquez et al. In Section 2, we return to this issue, and report evidence that is consistent with a different hypothesis about the role of language: That it may support a *temporary* binding function that permits combination of color and direction.

1.3 Summary

Language and spatial representation are qualitatively different, and hence are functionally specialized for different tasks. Where language and spatial cognition overlap, we can ask whether and how language modulates our spatial understanding. Recently, several different lines of research have proposed a strong role for language in permanently modulating and changing our spatial representations. However, we have hinted that the existing findings might be better understood as temporary (on-line) effects of language in modulating spatial cognition. We now turn to research that illustrates some such temporary effects, and propose that this research casts doubt on stronger Whorfian claims about the effects of language on spatial cognition.

2 How language temporarily modulates visual-spatial representations

2.1 Selectivity and the modulation of attention

In this section, we will discuss two examples showing that language can temporarily modulate attention. These examples concern options for encoding motion events, and options for encoding locations using different frames of reference. Before we turn to these examples, however, it is important to address a key assumption underlying our proposal that language may temporarily influence spatial cognition. For such temporary interactions to occur, language and spatial cognition must be capable of interacting in an online fashion. If they could only influence each other offline (that is, over time, rather than in the moment of computation) temporary interactions would not be possible. To this end, we now briefly present a recent line of research illustrating that language can in fact modulate spatial attention in an on-line fashion, as an integrated part of a basic cognitive task. This research shows that a basic mechanism in visual search can be dynamically modulated by language, and lays the groundwork for our subsequent arguments.

A classical finding in the study of visual search is that the characteristics of search depend on the number of features that must be used to identify the target item (and distinguish it from the other items in the array). If the target is distinguished by only one feature such as color, visual search is very fast and response times do not generally depend on the number of items in the display (e.g., Treisman and Gelade, 1980; Nagy and Sanchez, 1990). For example, when one searches for a vertical bar embedded in an array of horizontal bars, the vertical bar appears to 'pop out' of the display, making search very easy and fast, with little difference due to the number of distracters. However, when the target is distinguished from the distracters by the combination of two or more features (e.g., vertical plus horizontal bars that distinguish a target L from a distracter T) response times typically increase linearly with the size of the display.

Vision scientists have hypothesized that these differences in processing result from the fact that the visual system encodes different visual features such as vertical and horizontal orientation independently of one another. Briefly, search for a single feature is thought to involve parallel pre-attentive mechanisms simultaneously operating over all the stimuli, causing the target item to appear to 'pop out' from the distracters. Search by multiple features (conjunction search) is thought to engage visual attention, bringing it to bear on each object individually to determine whether or not the two features occur together (Treisman and Gelade, 1980; though see Wolfe, 1998 for a review of arguments against this view). Because each object must be individually attended, the reaction time for this kind of search increases linearly with the number of items in the display.

Crucially for our argument, researchers have discovered another important property of visual search: It is capable of using incremental information while searching. Rather than being a rigid procedure that always functions the same way given the visual properties of the stimuli, visual search turns out to be a dynamic process that is capable of making use of information as it becomes available to it. Thus, reaction times for conjunction search are significantly reduced if half of the distracter items are previewed for 1 second before the display is presented (Watson and Humphreys, 1997). The preview allows people to ignore these (distracter) items, narrowing the set of items to which visual attention must be directed (see also Olds et al., 2000).

Recently, Spivey et al. (2001) demonstrated that language can have the same effect on visual search. The general framework for Spivey et al.'s experiments was initiated by Tanenhaus et al. (1995) who showed that people process language incrementally (i.e., they do not wait until the end of a sentence to begin parsing it) and are able to make use of the linguistic information as soon as it becomes available to guide visual processes. Spivey and colleagues showed that this same close time-dependent interaction is engaged during standard visual search tasks. In all of the experiments, subjects were informed of the target item via recorded speech (e.g., 'Is there a red vertical?'), and all of the visual displays were the same over the different conditions. In the 'Auditory First' condition, subjects heard the sentence *before* the visual display was presented and the display appeared immediately following the offset of the audio, as is standard in visual search experiments. In the 'A/V Concurrent' condition, however, the display appeared immediately before the words 'red vertical' were heard. Subjects thus heard the target item and saw the display at the same time (Figure 4).

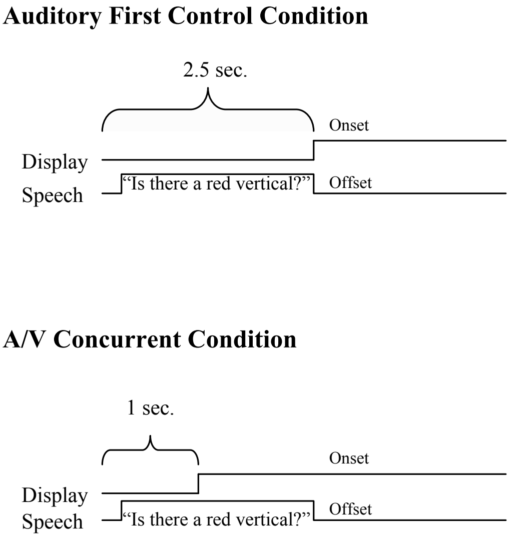


Figure 4. The timeline of presentation. In the 'Auditory First' condition, the visual display was presented following the offset of the audio. In the 'A/V Concurrent' condition, the display was presented as the search target was described auditorily. Adapted from Spivey et al. (2001), Experiment 1.

Spivey et al. replicated the effects of standard visual search tasks in the 'Auditory First' condition (which is the standard method used in the vision science community). Compared to the Auditory First condition, response times in the 'A/V Concurrent' condition increased at a significantly slower rate as the number of distracters increased. Spivey and colleagues hypothesized that subjects were able to make use of the incremental nature of speech to search the display in two phases. Specifically, as soon as subjects processed the word 'red', they were able to narrow their attention to only the red items. This allowed search for the vertical figure to proceed much more quickly when 'vertical' was subsequently heard (Figure 5). This suggests that even basic visual search tasks can be modulated by 'instruction' from language. Crucially, this modulation occurred while the sentence was being processed. Language thus can cause significant modulation of attention in a highly time-bound fashion; language can have an online influence on visual-spatial computations. We now turn to the role of language in directing attention when there are options for encoding motion events and locations (using frames of reference).

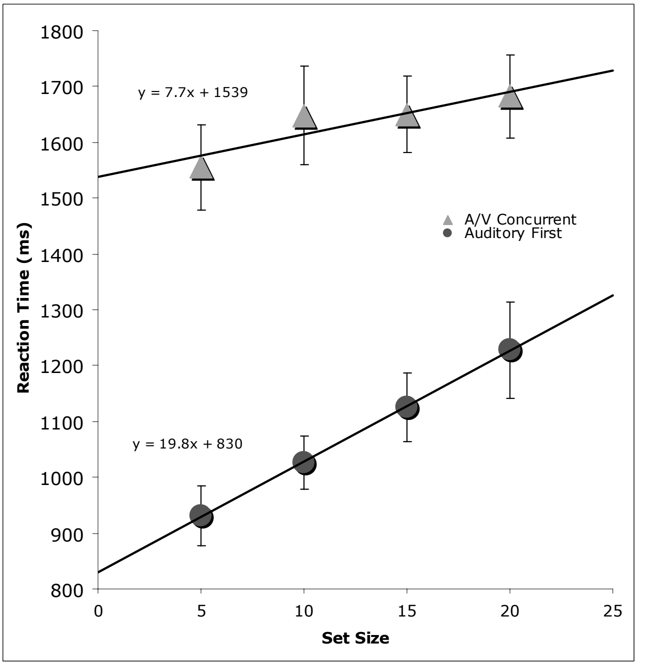


Figure 5. Response time by set size for the 'A/V Concurrent' and 'Auditory First' conditions. The slope in the A/V Concurrent condition is shallower than the Auditory First condition, suggesting that the incremental nature of the auditory stimulus facilitated search. Figure adapted from Spivey et al. (2001), Experiment 1. Data used with permission.

2.1.1 Language directs attention to different components of motion events

a. Manner vs. Path: Attending to different elements

The different typological patterns for motion events described by Talmy (1985) have been used to test the possibility that language causes permanent effects on people's representation of motion events. As discussed earlier, these studies have generally resulted in weak findings, with little convincing evidence that, e.g. speaking a Path language permanently and completely alters our event perception. More importantly, recent findings suggest that the effects are short-lived, and highly task dependent. They suggest that language is a powerful modulator of visual attention *in the moment of a task.*

Papafragou, Trueswell and Hulbert (2006) examined attentional allocation among Greek and English speakers, asking whether the different typological patterns would alter visual search. Greek is predominantly a Path language (in Talmy's typology) whereas English is a Manner language; the prediction is that the former should show heightened allocation of attention to Path and the latter to Manner. Papafragou et al. showed that the different tendencies in Greek vs. English do indeed have consequences for attentional allocation, but these consequences hold principally for the purposes of *linguistic* encoding.

In the study, native speakers of English and of Greek were shown a sequence of brief animated events depicting events that were either bounded (e.g. a person skating around/to a snowman) or unbounded (e.g. a person skating, with no goal in sight). All participants were told that they would view each event, which would be followed by a beep. People in the 'Linguistic' condition were told that they would be asked to verbally describe the events after the beep; people in the 'Non-linguistic' condition were told that they should watch the video and then continue inspecting the event after the beep, because they would be tested afterwards for recognition of the seen events.

Crucially, people in the study were eye-tracked as they carried out the task. Eye movements can be tracked very precisely, and are an excellent way to track people's changing focus of attention throughout a task. Generally, an eye movement to a location is preceded by attention (e.g., Hoffman and Subrahmanian, 1995; Henderson, 1993; see Irwin, 2004, for a review), so the pattern of eye movements provides insight into the allocation of attention. Moreover, eye movements generally occur without conscious awareness and are not usually subject to explicit voluntary modulation, so they reveal how attention is allocated as a consequence of the goals driving different cognitive tasks.

Papafragou et al. reported several notable results. First, people in the 'Linguistic' condition showed language-specific differences in allocation of attention as soon as each event started and as it began to unfold over time: English speakers looked to the region closely related to the manner of motion (e.g. the skates in a skating event) whereas Greek speakers looked to the region closely related to the endpoint of the path (e.g. the snowman in the bounded event). These differences map onto the kind of main verb most likely to be used by English vs. Greek speakers (manner of motion verbs and path verbs, respectively) and suggest that as the speakers were preparing to describe the events linguistically, they attended to those properties of the event that would be most relevant to choosing the proper verb. These cross-linguistic differences occurred only in the bounded events, which are events that would naturally engage different linguistic patterns (see Papafragou et al., for discussion).

In contrast to these clear effects of preparing to describe the events, people in the 'Non-linguistic' condition showed no differences in eye movements as the bounded events unfolded over time. That is, both English and Greek speakers who were inspecting the event for later recall distributed their attention in similar ways. The only differences in this condition appeared after each event had finished, while people continued to inspect the still frames of the finished video in order to prepare for the memory task. Now the English speakers focused attention on the path endpoint (snowman), where as the Greek speakers focused attention on the manner-relevant element (e.g. the skates). Papafragou et al. interpret this as evidence that, in this last phase when they were getting ready for the recall test, people attempted to encode the events verbally. When they did so, they attempted to focus attention on those elements *not* well encoded by the main verbs in their language (e.g. path endpoint for English speakers, manner for Greek speakers), perhaps as a means to clearly remember those 'secondary' elements (i.e. the ones not encoded by their main verbs).

The overall pattern of results shows that the allocation of attention differs depending on the cognitive task, specifically, by the kind of cognitive goal that a person has. When the goal is to tell 'what happened', speakers of different languages will immediately focus attention on the element(s) that is most likely to be encoded in the main elements of their language, in this case, the main verb. When the goal is to merely 'inspect' the event, speakers of different languages simply focus on elements that are highly salient, independent of the language they speak. When the goal is to inspect in order to recall, speakers may attempt to encode the event linguistically, but then they will also focus attention on aspects of the event that are not the highest ranked within their linguistic system (i.e. the aspect not encoded by the main verb). This would seem to be a smart strategy for improving the overall quality of the event representation, perhaps by creating a hybrid representation that includes a compact linguistic representation (i.e. of the main aspect of the event) and a visual-spatial representation that encodes the rest. In Section 2.2, we will report a related example, in which the combination of language and spatial representations provides a powerful combination, perhaps richer than either one or the other alone.

b. TO-Paths vs. FROM-Paths: Reversing natural attentional biases

Although the visual-spatial description of a single path includes the entire geometry (with no particular bias towards the beginning or end), languages allow us to move our 'attentional zoom lens' (Fisher, Hall, Rakowicz and Gleitman, 1994) over any or all conceptual portions. For the same observed event, we can say 'The racehorse ran *from* the starting gate at top speed' or 'The racehorse ran *towards* the finish line' or 'The racehorse ran *past* the viewing stand when he stumbled'. Or we can describe all portions of the event by combining these pieces. Most linguistic analyses put these components on equal footing, without marking any one as more primitive than any other.

But when people describe events, they are not so even-handed. Lakusta and Landau (2005) found that there is a strong tendency for children and adults to encode the goal of spatial motion events in preference to the source. At the same time, we also found that language can provide a powerful means to alter this preference, leading people to focus either on the source or goal, depending on the particular lexical item used in instruction (see also, Fisher et al., 1994).

In our experiments, we asked 3 and 4 year-old children and adults to describe simple videotaped manner of motion events. Each of these events showed a person moving along a specific path from one object to another by a variety of manners, e.g. hopping, walking, running, crawling, etc. After each event was shown, people were asked to tell the experimenter 'what happened'.

In English, events such as these are readily encoded by manner of motion verbs (e.g. hop, walk, run), and these verbs freely and grammatically take either FROM paths, TO paths, both or neither. For example, a person could aptly describe a given event as 'The girl hopped from the mailbox' or 'The girl hopped to the lamppost' or 'The girl hopped from the mailbox to the lamppost' or simply 'The girl hopped'. Although these are all – in principal – equally possible, children and adults showed a strong tendency to explicitly encode the goal path in preference to the source path, saying, e.g. 'The girl hopped to the lamppost' in preference to 'The girl hopped from the mailbox' or 'The girl hopped from the mailbox to the lamppost'. The tendency was slight among adults, who clearly understood that the experimenters wanted a 'nice complete' description, but it was pronounced in young children. This suggests that, in the relatively neutral case of manner of motion events – where the verb does not 'care' whether one or the other, or any path is encoded – people tend to include the goal path but omit the source path.

In follow-up experiments, we asked whether this pattern – which we called the 'Goal bias' – extended to other kinds of events. Here, we built on a linguistic theory developed by Gruber (1976) and extended by Jackendoff (1983) as the Thematic Relations Hypothesis. This hypothesis starts with the observation that there are significant parallels between the way that paths are encoded linguistically in inherently spatial events – such as manner of motion events – and in non-spatial events. For example, the same prepositions 'to' and 'from' (and their relatives) are used freely in the domain of transfer of possession (e.g. verbs give/get, throw/catch, buy/sell). As Jackendoff points out, this kind of transfer is analogous to motion of an object through space from one person to another, and the verbs that encode these transfers show expressions that are parallel to those encoding motion of objects in non-transfer contexts. Thus, we can say 'Mary gave the watch TO Bob' or 'Bob got the watch FROM Mary', focusing on either the path from giver to recipient, or vice versa. The parallel also extends to other domains, for example, change of state (verbs such as turn and grow) and attachment/detachment (verbs such as attach or hook, detach or remove, etc.)

In order to see whether the same source/goal asymmetry applied to these domains, we videotaped events that could readily be encoded using verbs appropriate for transfer of possession, change of state, and attachment/detachment. For example, one set of events showed 'giving/getting', 'throwing/catching' and 'selling/buying'. These events could be encoded with 'goal verbs' (give, throw, sell) or equally well, with 'source verbs' (get, catch, buy)– each of which focuses on a distinctly different 'viewpoint' for the event. A second set of events showed changes of state, e.g., an animal whose ears changed colors and a person whose expression changed from happy to sad. A third set of events showed events in which a person either attached or detached one object to/from another.

When children and adults were asked to freely describe what happened in these events, we found the same goal-source asymmetry (i.e. goal bias), with people choosing 'goal-oriented' verbs (e.g. give, throw, sell) rather than source-oriented verbs (get, catch, buy), and specifying the goal paths (e.g. 'give/throw to X') rather than source paths (e.g. 'got/caught from Y'). Thus the goal bias is very robust and appears to represent a bias for people to construe events in terms of their goal states and/or endpoints rather than their source states and/or starting points (see Lakusta and Landau, 2005 for discussion).

Given this strong bias (especially among children), one might wonder when and how people come to flexibly reorient their attentional lens to choose descriptions in terms of starting points. In a separate experiment, we asked whether we could modulate interpretation of the event (and full description) by providing children with a verb that has a source/starting point bias. Using the same videotaped events, we asked a separate group of children to tell us what happened, but we also told them that we would give them a 'hint'. The hint was the target verb, and it was either a goal-oriented verb (e.g. give) or a source-oriented verb (e.g. get). For example, an event in which an object is transferred from one person to another could equally well be described with the verb 'give' or 'get'. When children were told to describe the event using their hint verb 'give', they all complied, saying e.g. 'The girl gave a present to the boy'. But when they were told to describe the event using their hint verb 'get', they also complied, saying, e.g. 'The boy got a present from the girl' (or 'The boy got a present'). This shows that 3 year-old children were adept at framing their description of the event in terms of the goal or source, even though their spontaneous tendency (found in the previous experiment) was to frame it in terms of the goal-oriented verb (give).

Children's facility in this task shows that reframing the construal of the event in terms of source or goal is relatively easy, if one has the powerful 'hint' from language, i.e. the particular verb that will force the reinterpretation. Although we did not gather eye movement data to monitor the children's changing attentional focus, we expect that, as in Papafragou et al.'s study, eye movements would differ depending on which hint verb is given. The larger point is that modulation of the mental construal of the event– either as a 'source-oriented' or 'goal-oriented' event– can be done quickly, efficiently, and easily through use of different lexical items that focus attention on different possible construals of the very same event. This would appear to be a clear case of the immediate modulation of attention using language as the mental pointer to a new construal.

2.1.2 Language directs attention to different available reference systems

A reference system is a geometric system with an origin and at least two orthogonal axes; locations of an object can be specified in terms of the locations (possibly coordinates) on each axis. The same physical location in space can be represented using many different reference systems, each of which is defined by the location of the origin. Thus, for example, the location of a point (x) in space might be represented relative to a reference system centered on the retina, the head, the torso, the entire body. It can be represented relative to a reference system centered on another object (point x is left of object y), any aspect of the environment (a room, a building, a city), or even larger spaces, e.g. the earth.

The idea of reference systems is crucial to understanding how we (and other mobile species) represent location and how we do this for different purposes– e.g. reaching, pointing, looking, searching, or talking about locations. Because of the importance of reference systems across all of these domains of inquiry, the literature on use of reference frames is vast, ranging from studies of how the visual system programs saccades (e.g. Colby et al., 1999) to how we reach and grasp objects (Milner and Goodale, 2005) to how we deploy attention (Carlson-Radvansky and Logan, 1997) to the acquisition and use of spatial language (Carlson-Radvansky and Irwin, 1993; Landau and Hoffman, 2005; Li and Gleitman, 2002).

What is absolutely clear from the wealth of information available on reference frames is that humans represent location in terms of a variety of reference frames and that they are flexible in engaging these. This flexibility has recently been examined within the attention literature, and a striking fact has emerged: When carrying out tasks requiring that people locate one object relative to another (in order to verify a sentence that describes the location), people activate more than one reference system, and then *select* one over the other (the latter of which is inhibited; Carlson-Radvansky and Logan, 1997; Carlson-Radvansky and Jiang, 1998).

The evidence in these studies shows that the engagement of any particular reference frame is subject to the goals of the perceiver, but that multiple reference frames are likely to be available for any given task (see also Gallistel, 2002). Still, the mechanisms by which people select reference frames are not well understood. This makes it surprising that the literature on language and thought has recently been dominated by a strong Whorfian hypothesis: Levinson (2003) has proposed that the reference system most frequently used in one's native language will lead to permanent organizational changes in one's *non-linguistic* spatial cognition.

This hypothesis was spurred by findings from Pederson et al. (1998), who investigated the *linguistic* choices of reference frames among speakers of a variety of languages including Tzeltal, Mopan, Longgu, Dutch and Japanese. The task involved a Director and a Matcher, who were seated next to each other with a screen placed between them. They viewed a set of photos and as the Director described a picture, the Matcher was supposed to select the corresponding one in his own array. Individual pictures were set up so that a person could use one of several possible frames of reference to describe the picture. The question was whether speakers of different languages would use different frames of reference.

For example, one picture displayed a man and a tree; the picture could equally well be described using ego-centered frame of reference (e.g. 'the man is on the *left*', relative to the viewer), a geocentric one (or what Levinson and others have called 'absolute', e.g. 'the man is to the *north* of the tree'), or an object-centered frame of reference[1] (e.g. 'the man is facing the tree'). All of these frames were used to some extent, but language groups differed in their tendencies. Speakers of some languages mainly used one of the frames of reference (e.g. object-centered in Mopan) while speakers of other language groups used a combination of egocentric and object-centered (Dutch, Japanese) or object-centered and geocentric (Tzeltal).

Given this variation, Pederson et al. then tested the hypothesis that 'users of different language systems (in a given context) should correspondingly vary in their choice of nonlinguistic spatial problem-solving strategies (in analogous contexts)' (p. 574). In a new task, subjects viewed three animals placed facing a particular direction on a table. They were asked to 'remember the objects just as they are' (see Figure 6A). Subjects were then rotated 180 degrees and after a 30 second delay, walked over to a new table. At this table, they were given the same animals and were asked to arrange the animals in the same sequence they had just seen. Notice that the task is ambiguous – given the setup on the stimulus table, subjects could reproduce the pattern on the recall table using an egocentric (relative) frame of reference or a geocentric (absolute) frame of reference (Figure 6B).

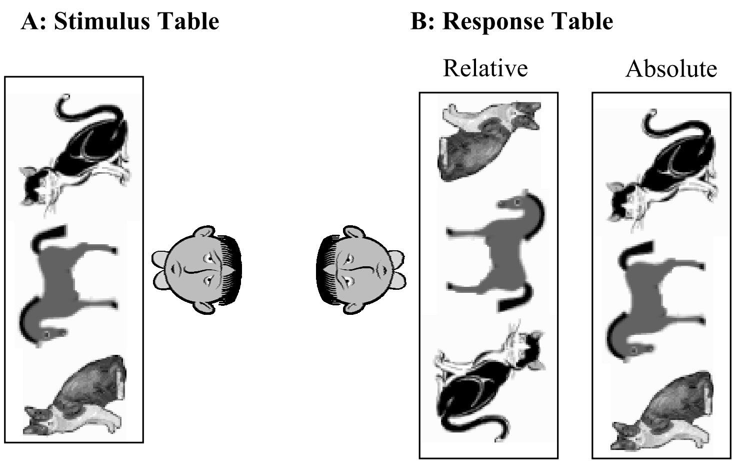


Figure 6. The Animals in a Row task. (A) Subjects are shown the arrangement of animals on the Stimulus Table. They are then rotated 180 degrees to the Response Table (B), and are asked to recreate the arrangement of animals. People may replicate the pattern using an egocentric frame of reference ('relative') or a geocentric frame of reference ('absolute'), depending on the conditions of test and the person's native culture (adapted from Pederson et al., 1998 and Li and Gleitman, 2002).

Pederson et al. reported that people from language groups in which geocentric system is often used gave geocentric (absolute) responses, while other language groups (Dutch and Japanese) gave egocentric (relative) responses. Thus, taken together the results '… indicate that the frame of reference identified in the linguistic elicitation task correlates well with the conceptual frame of reference used in this recall task.' (pg. 580). These correlations were taken to suggest that '…we must represent our spatial memories in a manner specific to the socially normal means of expression.' (pg. 586). That is, the language one uses determines the choice of frame of reference one uses in nonlinguistic representations.

Li and Gleitman (2002) argued on both empirical and theoretical grounds that Pederson et al.'s findings did not reflect the effects of language on spatial thought, but rather, could be viewed as the reverse causal chain: A group's choice of a particular frame of reference for encoding in language could easily be the result of culture, terrain, or other variables such as environmental factors; and this *non-linguistic* choice could then bias the tendency of language to use one or the other frame of reference. Li and Gleitman tested this possibility by manipulating the conditions under which monolingual speakers of English solve the rotation problem. If such speakers (who do not naturally tend to describe the small scale layout using geocentric terms) also tend to change their choice of reference frame (without accompanying changes in their primary language, obviously), then it would follow that the choice of reference frames is not caused by language.

Li and Gleitman began by administering Pederson et al.'s linguistic and nonlinguistic tasks to native English speakers. In the linguistic task, with a Director describing photos to a Matcher, they found that English speakers overwhelmingly used terms that engage an egocentric frame of reference, i.e. left and right relative to their body. In the nonlinguistic task (reconstructing the order of toys they had seen), choice of reference frame was strongly affected by the surrounding spatial context. When minimal landmark cues were present (i.e. the experiment was conducted in a lab room with window blinds down) subjects responded using an egocentric frame of reference, e.g. 'to my left' before and after rotation. This pattern was consistent with that of Dutch and Japanese speakers in Pederson et al.'s study. But when there were salient landmarks (e.g. the experiment was conducted outside or in a lab room with the blinds up), people varied quite a bit, responding either on the basis of an egocentric or geocentric frame of reference. Moreover, people changed their responses as significant landmarks were introduced: If a small landmark was placed on both tables such that it was in the same geocentric location (e.g. north), then people responded using a geocentric frame of reference. If the landmark was placed in the same egocentric location (e.g. on the left end of the pre-rotation table but the right end of the post-rotation table), then people adopted an egocentric frame of reference.

These findings show that, contrary to Pederson et al.'s claim, the language one speaks has no permanent organizing effect on one's choice of reference frame for a *non-linguistic* task. Rather, as shown by many experiments on spatial cognition on humans and non-humans, there is great flexibility in which reference system will be adopted for what task (for review, see Landau 2002; Gallistel, 2002). Moreover, as Carlson-Radvansky and colleagues have shown, multiple reference frames are likely to be activated in parallel (Carlson-Radvansky and Logan, 1997). It is the selection of one, and the inhibition of others, that causes a particular response (Carlson-Radvansky and Jiang, 1998). This selection occurs in a limited time frame (in less than a second, in Carlson's studies), and is unlikely to persist.

So why are there tendencies for speakers of one language to choose one reference system rather than another in Pederson's studies? We propose that their findings can be easily explained as the tendency to solve the 'non-linguistic' problem using language– in which case, of course, one's linguistic coding would automatically select one or another reference frame. If one uses the dominant coding of one's native language, it is not at all surprising that one would then recreate the test array in accordance with that linguistic coding. What people would have effectively done in this case is to activate all reference systems, choose one (if linguistically coding the location, choosing the dominant coding of their language) and then recreate the array using that coding. Li and Gleitman's results show that people's selection of reference frames can be easily changed depending on external task conditions, which surely interact with the perceiver/actor's goal.

The bottom line is that people can freely choose to represent a particular location within many different reference systems; language may be a mechanism that ramps up attention to certain reference systems over others, without forcing any permanent change in the *availability* of multiple reference frames. And language surely has a powerful function in providing the means by which a speaker informs the hearer which of the multiple reference system he or she has in mind. Without such power, it would be hard to imagine how any of us could ever understand directions.

2.2 Enrichment: language binds together elements, modulating and enhancing our visual representations

Our second focus in this chapter is on the role of language in enriching spatial representations. Earlier, we discussed a hypothesis put forth by Spelke and colleagues, suggesting that language provides the computational power to combine outputs of different systems that are otherwise modular. The case that they offered concerned reorientation by humans, and they proposed that the solution to this task might require that language be used to combine geometry and color. We now discuss a related case from our own research which shows that such combinations might be the product of on-line, temporary computations in which language and spatial representation enrich each other.

The general question we have been investigating is the extent to which language can play a role in enriching visual representations. To do this, we have examined a case involving a well-known problem in the visual system. The visual system is thought to process different visual features independently of each other (at some level) and abundant research has shown that the visual system sometimes fails to bind together visual features that co-occur in a single object or multiple objects. A classic example was reported by Treisman et al. (e.g., 1982): If people are very briefly presented with a display containing a red O adjacent to a green L, they will often mistakenly report that they have seen either a red L or a green O. This phenomenon is called illusory conjunction and is thought to reflect difficulty in binding together the two kinds of features (color and shape). Theories of visual attention have suggested that binding requires active allocation of focused attention at the location of the target object, with the location serving as the 'glue' that binds the features together. Although the bulk of work has been carried out with normal adults, there are also reports of brain-damaged individuals who experience both attentional difficulties and a pronounced occurrence of illusory conjunctions. This combination suggests that attention is necessary for the process of binding together individual visual features (Arguin, Cavanagh and Joanette, 1994).

We took the case of failure to bind as a possible arena within which to test the effects of language. Previous findings had shown that young children (around 6 years of age) might have binding problems when the features are color and location. Hoffman, Landau and Pagani (2003) showed children geometric blocks that were split in half either horizontally, vertically, or diagonally and were filled with two different colors in each half (see Figure 7A). Children were shown a target block, and were then asked to match it to one of a set of 8 blocks below the target. Children tended to choose the correct split (e.g. if the target was a vertical split, they chose a block with a vertical split), but they also tended to err in assignment of color: They might choose a vertically split block with red on the left/green on the right, or the mirror image. Even a very short (1 second) delay between viewing the target and selecting the match, there were significant errors.

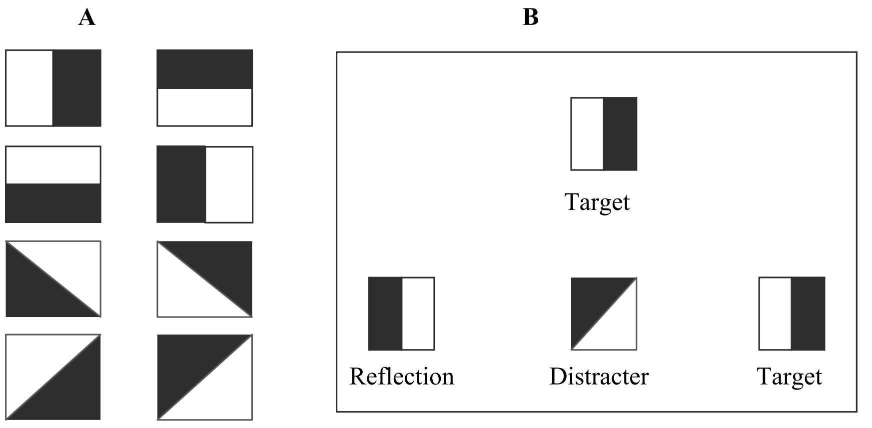


Figure 7. Geometric blocks split in half horizontally, vertically, or diagonally. (A) The eight blocks used by Hoffman et al. (2003). When subjects about 6 years old were given a block split, e.g., vertically (black right, white left) and were asked to match the target after a 1 second delay, subjects confused the target with its reflection (black left, white right). (B) Dessalegn and Landau (2008) used a task where a target block is shown followed by three test items after a 1 second delay. Subjects confused the target and its reflection in all conditions except when the position of one of the colors was labeled with a directional phrase: 'The black is on the right' (see text for discussion of the conditions). Note that in the text and actual experiments, blocks split by red and green were used. The figure displays black and white splits for ease of printing.

Dessalegn and Landau (2005; 2008) asked if such apparent failures to bind visual features could be modulated – or even overcome – by language. In a first experiment, we used a task similar to that used by Hoffman et al. (2003). Four year-old children were shown target squares that were split vertically, horizontally, or diagonally by color (see Figure 7B). One square was presented on each trial, in the top center of a computer screen. Children were instructed to look at the target very carefully, so they could find exactly the same one out of a set of alternatives. When they had finished inspecting the target, the experimenter clicked the mouse, and the target disappeared; after a 1 second delay, three options appeared on the bottom of the screen. Options included the target replica, its reflection (i.e. colors and locations switched), and a distracter square having a different split from the target (see Figure 7B). The key question was whether children would be able to retain the exact identity of the target without confusing it with its reflection. Doing so would require that each color (e.g. red, green) be bound with its location (e.g. left, right, top, bottom).

Results showed that children performed above chance, selecting the correct target block on about 60% of the trials. However, 88% of the errors were target reflection confusions, e.g. selecting a red-left/green-right block instead of a red-right/green-left block. This pattern of performance suggests that the visual-spatial representation of the target did not include a stable representation in which colors were bound to their respective locations.

In a second experiment, Dessalegn and Landau asked whether such failures in forming a stable representation could be overcome by using linguistic instructions that explicitly provide the location of each color. The same targets were presented using the same method, except that the child was told, e.g. 'See this? The red is on the left' (or right, top, bottom). Children's performance improved by about 20%, with accurate performance now hovering around 80%.

In some ways, it may seem trivial that the linguistic instruction 'The red is on the left' helped children to select the correct target after a one second delay. One interpretation is that language can enhance attention (as shown by the phenomena discussed in the previous sections), drawing the child's attention to the figures and enhancing the process of binding color and location. Several other experiments explored the mechanisms that might accomplish this. One experiment tested whether simply labeling the target with a novel whole object label would do the trick. Some have argued that labeling objects has a powerful attentional pull for children, resulting in heightened attention to certain properties over others (Smith et al., 1996; Smith et al., 2002). However, when the targets were labeled as whole objects (e.g. 'See this? This is a dax.'), children's performance dropped back to the level observed with no specific linguistic instruction (i.e. around 60%). The same pattern of performance occurred when we substituted neutral spatial terms for the directional ones (e.g. 'See this? The red is *touching* the green.'). Moreover, other attempts to directly manipulate the child's attention failed. For example, the child was shown the target and told 'Let's see where the red part is', after which, the red section of the block flashed briefly on and off for several seconds. This did not result in better performance. In another condition, the red section grew and shrunk, which presumably should have drawn the child's attention. But it did not result in better performance. In yet another condition, the child was asked to point to the red part, but this did not help.

So how did the 'left/right' sentences help? The most obvious interpretation of this entire pattern of results is that the children had a full and accurate long-term representation of the terms top, bottom, left, and right, and that they used this knowledge to distinguish between the target and the reflection. That is, they stored the linguistic representation (e.g., 'the red is on the left') and were able to make use of it in the matching task without using their visual-spatial representation of the target at all. We had anticipated this possibility, and had therefore carried out a production and comprehension task after the main task, testing children's long term knowledge of the terms top, bottom, left, and right.

Children had very accurate representations for terms *top* and *bottom*. For left and right, however, they were near chance at distinguishing the two directions. When asked to place a dot 'to the left of' a square, they correctly placed the dot along the horizontal axis, but often erred on the direction, showing that they did not know which end of the horizontal axis was left and which was right. Crucially, there was no significant correlation between accuracy in the production and comprehension tasks and accuracy on the main matching task, suggesting that they were not using their long-term understanding of left/right to carry out the matching task.

Dessalegn and Landau (2008) interpreted this pattern to suggest that children were using language in this task as follows: When they heard the sentence 'The red is on the left', the children temporarily represented the term 'left' accurately, by just noting the location (i.e. direction) of the red part relative to the green. This temporary representation could then be held over the span of the one second delay, and brought to bear on the task of choosing the target. That is, when the test items appeared, children could use their temporary representation to select the test-item that had the red on the left side of the object, successfully distinguishing it from its mirror image. But ten minutes later, when given the production and comprehension tasks, this representation was gone, resulting in failure to distinguish between left and right.

As a whole, these findings suggest that language did not have its powerful effect because of stable, long-term representations of words like left/right. Instead, the findings point to a powerful but temporary enhancement of the representation of the target's direction, which was used on-line, for the purposes of matching, but which rapidly disappeared. This enhancement augmented the visual-spatial representation of the target in the context of the task, working to bind together the color and location in the moment of test.

3.0 Summary and conclusions

As we noted in the beginning of this chapter, the idea that language and visual-spatial cognition interact is not new. What remains unclear, however, is exactly how these two quite different systems of representation affect each other– whether the effects are temporary or permanent, task-dependent or quite general, and the degree to which the interactions confer greater increased representational power to human cognition. In this chapter, we have proposed two specific mechanisms of interaction– selectivity and enrichment. *Selectivity* occurs because language is inherently selective, encoding certain distinctions and not others; and because language can serve as a mental pointer, indicating which of many possible representations we have in mind. Surprisingly, these effects occur incrementally, as we speak and hear, providing a continually changing pointer to our different mental construal of the world. *Enrichment* occurs because language has the representational power to robustly encode certain properties that are only encoded in fragile form in the visual-spatial system. We have provided examples of how each of these mechanisms operates in a time-bound fashion, as people carry out particular cognitive tasks. The evidence that language can play a time-bound role in modulating spatial cognition makes us question whether any effects of language on spatial cognition can be considered permanent, as envisioned by strong versions of the Whorfian hypothesis. But giving up a strong version of Whorf's hypothesis does not mean relinquishing the idea that language is a powerful modulator of human thinking. Indeed, the real power of language may be precisely in its time-bound effects, which ultimately permit humans the flexibility to communicate to others, on a moment to moment basis, the rich variety of mental construals of the world.

References

Arguin, M., Cavanagh, P. and Joanette, Y. (1994) Visual feature integration with an attention deficit. *Brain and Cognition* 24(1): 44–56.

Berlin, B. and Kay, P. (1969) *Basic color terms, their universality and evolution.* Berkeley: University of California Press.

Bowerman, M. (1973) Structural relationships in children's utterances: Syntactic or semantic? In T. E. Moore (ed.) *Cognitive development and the acquisition of language*. New York: Academic Press.

Brown, R. (1973) Development of the first language in the human species. *American Psychologist* 28(2): 97–106.

Brown, R. (1976) Reference in memorial tribute to Eric Lenneberg. *Cognition* 4(2): 125–183.

Carlson-Radvansky, L. A. and Irwin, D. E. (1993) Frames of reference in vision and language: Where is above? *Cognition* 46(3): 223–244.

Carlson-Radvansky, L. A. and Jiang, Y. (1998) Inhibition accompanies reference-frame selection. *Psychological Science* 9(5): 386–391.

Carlson-Radvansky, L. A. and Logan, G. D. (1997) The influence of reference frame selection on spatial template construction. *Journal of Memory and Language* 37(3): 411–437.

Cheng, K. and Gallistel, C. R. (1984) Testing the geometric power a spatial representation. In H. L. Roitblat, H. S. Terrace and T. G. Bever (eds) *Animal cognition* 409–423. Hillsdale, NJ: Erlbaum.

Cheng, K. and Newcombe, N. S. (2005) Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review* 12(1): 1–23.

Clark, E. (1973) What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (ed.) *Cognitive development and the acquisition of language* 65–110. New York: Academic Press.

Clark, H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive development and the acquisition of language* 27–64. New York: Academic Press.

Colby, C. and Goldberg, M. (1999) Space and attention in parietal cortex. *Annual Review of Neuroscience* 22: 319–349.

Dessalegn, B. and Landau, B. (2005) *Relational language binds visual-spatial representations.* Paper presented at the 4th Biennial meeting of Cognitive Development Society, San Diego, CA.

Dessalegn, B. and Landau, B. (2008) More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science* 19(2): 189–195.

Fillmore, C. J. (1997) *Lectures on deixis*. Stanford, CA: CSLI Publications.

Fisher, C., Hall, D. G., Rakowitz, S. and Gleitman, L. (1994) When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. In L. Gleitman and B. Landau (eds) *Acquisition of the lexicon* 333–375. Cambridge, MA: MIT Press.

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

Gallistel, C. R. (2002) Language and spatial frames of reference in mind and brain. *Trends in Cognitive Sciences* 6(8): 321–322.

Gennari, S. P., Sloman, S. A., Malt, B. C. and Fitch, W. T. (2002) Motion events in language and cognition. *Cognition* 83(1): 49–79.

Gentner, D. (2001) Spatial metaphors in temporal reasoning. In M. Gattis (ed.) *Spatial schemas and abstract thought* 203–222. Cambridge, MA: MIT Press.

Gentner, D. and Goldin-Meadow, S. (2003) *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Gleitman, L. and Papafragou, A. (2005) Language and thought. In K. J. Holyoak and R. G. Morrison (eds) *The Cambridge handbook of thinking and reasoning* 633-661. Cambridge: Cambridge University Press.

Gruber, J. S. (1976) *Lexical structures in syntax and semantics*. New York: North-Holland Publishing Company.

Hauser, M. D., Chomsky, N. and Fitch, W. T. (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598): 1569–1579.

Henderson, J. M. (1993) Visual attention and saccadic eye movements. In G. d'Ydewalle and J. Van Rensbergen (eds) *Perception and cognition: Advances in eye-movement research* 37–50. Amsterdam: North-Holland.

Hermer-Vazquez, L., Moffet, A. and Munkholm, P. (2001) Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition* 79: 263–299.

Hermer-Vazquez, L., Spelke, E. S. and Katsnelson, A. S. (1999) Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology* 39(1): 3–36.

Hermer, L. and Spelke, E. (1996) Modularity and development: The case of spatial reorientation. *Cognition* 61(3): 195–232.

Hoffman, J., Landau, B. and Pagani, B. (2003) Spatial breakdown in spatial construction: Evidence from eye fixations in children with Williams syndrome. *Cognitive Psychology* 46(3): 260–301.

Hoffman, J. and Subramaniam, B. (1995) The role of visual attention in saccadic eye movements. *Perception & Psychophysics* 57(6): 787–795.

Irwin, D. E. (2004) Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson and F. Ferrreira (eds) *The interface of language, vision, and action: Eye movements and the visual world* 105–133. New York: Psychology Press.

Jackendoff, R. (1983) *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1987) On Beyond Zebra: The relation of linguistic and visual information. *Cognition* 26(2): 89–114.

Kay, P. and Kempton, W. (1984) What is the Sapir-Whorf hypothesis? *American Anthropologist* 86(1): 65–79.

Lakoff, G. (1987) *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakusta, L. and Landau, B. (2005) Starting at the end: The importance of goals in spatial language. *Cognition* 96(1): 1–33.

Lakusta, L., Wagner, L., O'Hearn, K. and Landau, B. (2007) Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development* 3(3): 179–197.

Landau, B. (2002) Spatial cognition. In V. Ramachandran (ed.) *Encyclopedia of the human brain* 395–418. (Vol. 4) San Diego: Academic Press.

Landau, B. and Hoffman, J. E. (2005) Parallels between spatial cognition and spatial language: Evidence from Williams syndrome. *Journal of Memory and Language* 53(2): 163–185.

Langacker, R. (1986) An introduction to cognitive grammar. *Cognitive Science* 10(1): 1–40.

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

Levinson, S. C. (2003) Language and mind: Let's get the issues straight! In D. Gentner and S. Goldin-Meadow (eds) *Language in mind: Advances in the study of language and thought* 25–46. Cambridge, MA: MIT Press.

Li, P. and Gleitman, L. (2002) Turning the tables: Language and spatial reasoning. *Cognition* 83(3): 265–294.

Malt, B., Sloman, S. and Gennari, S. (2003) Universality and language specificity in object naming. *Journal of Memory and Language* 49(1): 20–42.

Majid, A., Bowerman, M., Kita, S., Haun, D. and Levinson, S. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3): 108–114.

Marr, D. (1982) *Vision*. San Francisco: Freeman.

Marr, D. and Nishihara, H. (1992) Visual information processing: Artificial intelligence and the sensorium of sight. *Frontiers in cognitive neuroscience* 165–186. Cambridge, MA: MIT Press.

Mandler, J. (1992) How to build a baby: II. Conceptual primitives. *Psychological Review* 99(4): 587–604.

Milner, A. and Goodale, M. (2005) *The visual brain in action*. New York: Oxford University Press.

Morton, J. (2004) *Understanding developmental disorders: A causal modeling approach*. Oxford: Blackwell Publishing.

Munnich, E., Landau, B. and Dosher, B. (2001) Spatial language and spatial representation: A cross-linguistic comparison. *Cognition* 81(3): 171–207.

Nagy, A. L. and Sanchez, R. R. (1990) Critical color differences determined with a visual search task. *Journal of the Optical Society of America, A, Optics, Image & Science* 7(7): 1209–1217.

O'Keefe, J. and Nadel, L. (1978) *The hippocampus as a cognitive map*. Oxford: Clarendon.

Olds, E. S., Cowan, W. B. and Jolicoeur, P. (2000) The time-course of pop-out search. *Vision Research* 40(8): 891–912.

Papafragou, A., Massey, C. and Gleitman, L. (2002) Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84(2): 189–219.

Papafragou, A., Trueswell, J. and Hulbert, J. (2006) *Mapping event perception onto language: Evidence from eye movements*. Paper presented at the 80th Annual LSA Meeting, Albuquerque.

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

Slobin, D. I. (1996) From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz and S. C. Levinson (eds) *Rethinking linguistic relativity* 70–96. Cambridge: Cambridge University Press.

Smith, L., Jones, S. and Landau, B. (1996) Naming in young children: A dumb attentional mechanism? *Cognition* 60: 143–171.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L. and Samuelson, L. (2002) Object name learning provides on-the-job training for attention. *Psychological Science* 13(1): 13–19.

Spelke, E. S. and Tsivkin, S. (2001) Language and number: A bilingual training study. *Cognition* 78(1): 45–88.

Spivey, M. J., Tyler, M. J., Eberhard, K. M. and Tanenhaus, M. K. (2001) Linguistically mediated visual search. *Psychological Science* 12(4): 282–286.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. and Sedivy, J. C. (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268: 1632–1634.

Talmy, L. (1983) How language structures space. In H. Pick and L. Acredolo (eds) *Spatial orientation: Theory, research, and application* 225–282. New York: Plenum Press.

Talmy, L. (1985) Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (ed.) *Language typology and syntactic description* 57–149. (Vol. III) Cambridge: Cambridge University Press.

Treisman, A. and Gelade, G. (1980) A feature integration theory of attention. *Cognitive Psychology* 12: 97–136.

Treisman, A. and Schmidt, H. (1982) Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14(1): 107–141.

Watson, D. G. and Humphreys, G. W. (1997) Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review* 104: 90–122.

Whorf, B. L. (1998) *Language, thought, and reality*. Cambridge, MA: MIT Press.

Wolfe, J. M. (1998) Visual search. In H. Pashler (ed.) *Attention* 13–74. Hove, UK: Psychology Press/Erlbaum.

Notes

1 This case is more complex than stated. A simple object-centered frame of reference (origin on the tree) would not uniquely capture the fact that the man is *facing* the tree. This requires an additional coordination with a frame of reference centered on the man, including specification of front, back, etc. For purposes of exposition, we are staying close to Pederson's own analysis.

Final proofs with index 24 February 2010

3 Language and inner space

Benjamin Bergen, Carl Polley and Kathryn Wheeler

1 Introduction: space in language and cognition

Much of language, including spatial prepositions and verbs of motion, is dedicated to describing the physical space that our bodies operate in. But it isn't just that humans occupy space; in an Escherian twist, our conceptual systems also contain internal representations of the world around them. These internal spatial representations become activated when we reason about spatial events and relationships in the world, when we recall such spatial events and relationships, and, the focus of the current paper, when we understand language about space and related domains. The convergent findings we survey in the following pages are nothing short of remarkable; they show that, in order to understand language about space, human language users internally reconstruct the spatial world and experience it through dynamic mental simulations. In short, understanding language about space is cognitively akin to perceiving space.

For many years, spatial language has been a primary domain of typological and cognitive linguistic research. All languages appear to dedicate resources to describing spatial relationships (Majid et al. 2004), but they differ in exactly how they express them. Different languages group spatial configurations in different ways (Bowerman and Pederson 1992). For instance, English clusters together vertical and horizontal attachment as *on* (a piece of paper can be *on* a wall or *on* a desk), while German distinguishes between these relations using *an* and *auf*. Even more striking, across languages, different frames of reference are used for spatial location descriptions. Languages like English prefer a relative frame of reference for small objects and short distances (*the pen is <u>to the left of</u> the lamp*), while languages like Guugu Yimithirr prefer an absolute frame of reference (something like *the pen is <u>North of</u> the lamp*) in such cases (Majid et al. 2004).

Despite cross-linguistic variation, however, systematic typological research has suggested at least two interesting ways in which spatial language is similar across languages. The first is that in all languages, words describing space, including closed-class sets of function morphemes such as spatial adpositions, appear to be organized in terms of conceptually constrained semantic primitives. Sometimes termed *image schemas* (Johnson 1987, Lakoff 1987), these conceptual primitives capture only schematic information, such as contact or containment relations between objects, but not the absolute size, color, or position of such objects or their containers (Talmy 2000). For instance, English *in* encodes the topological relation between an object and its container but not their Euclidean properties.

A second apparent cross-linguistic spatial universal is that many classes of abstract concepts are described spatially (Lakoff and Johnson 1980, Lakoff 1993). For instance, *in* doesn't only have a spatial meaning; it can also be used to relate an entity to a state, as in

We're in trouble, or *This research project is in its final throes*. Such use of concrete spatial language for abstract domains is analyzed in the literature as being driven by *conceptual metaphor* (Lakoff and Johnson 1980). Across languages, spatial terms come over time to acquire abstract meanings (Sweetser 1990), which results synchronically in words that are polysemous, with both spatial and abstract meanings. It has been suggested that this relation between abstract and concrete goes beyond mere language, such that understanding of abstract concepts is grounded in experiences with their concrete spatial counterparts, such as states being understood as containers (Lakoff 1993).

Until recently, there was limited evidence for the psychological reality of the spatial primitives that purportedly underlie the semantics of spatial and abstract language. The evidence that abstract concepts are understood in terms of concrete spatial ones was predominantly linguistic.

During the last decade, however, a number of lines of research have converged upon a set of common findings, suggesting that understanding language about space, as well as other abstract domains that are figuratively described in terms of space, results in the activation of the same cognitive mechanisms that are responsible for perceiving actual space. This is an instance of a larger movement within cognitive science variably called *simulation semantics* (Bergen 2007), *motor resonance* (Zwaan and Taylor 2006), or *perceptual simulation* (Barsalou 1999). Evidence collected in support of these views indicates that deep language understanding is accomplished by the internal creation or recreation of experiences of the world, which are triggered by descriptions encoded in language. These recreated mental experiences – known as mental imagery or mental simulation – make use of the same neurocognitive resources (the motor or perceptual systems, for instance) that are typically used for acting on or perceiving aspects of the world. This paper surveys evidence showing that spatial language does indeed engage those neurocognitive systems.

2 Spatial language

All normal humans, like most other animals, are endowed with brain systems that serve the functions of perceiving and acting in space. In humans, these systems are highly specialized, but at the same time they are so complex that the study of their localization and mechanics remains in its infancy. One thing we do know is that a variety of higher cognitive processes recruit systems dedicated to spatial cognition in order to bootstrap off the existing functionality of these systems.

Foremost among these parasitic cognitive capacities are memory and imagery. Behavioral and brain imaging evidence over the past century convergently indicates that recalling or imagining aspects of space involves activating a set of neural circuits that overlap with those used to perceive or act in space. One of the earliest demonstrations of this reuse of spatial circuits for imagery employed early image projection technology. Perky (1910) asked one set of subjects to imagine seeing an object (such as a banana or a leaf) while they were looking at a blank screen, while the other group was just asked to look at the screen. At the same time, unbeknownst to them, an actual image of the same object was projected on the screen, starting below the threshold for conscious perception, but with progressively greater and greater definiteness. Perky found that subjects who were imagining a banana or a leaf failed to recognize that there was actually a real, projected image, even at levels where the projected image was perfectly perceptible to those subjects who were not performing simultaneous imagery. The interference between imagining and perceiving shown in this early study and scores of subsequent experiments demonstrates that the system for perceiving objects is also used for imagining objects.

More recently, work on the Perky effect has shown that interference can also arise from shared location of a real and imagined object. Craver-Lemley and Arterberry (2001) presented subjects with visual stimuli in the upper or lower half of their visual field, while they were (i) imagining objects in the same region where the visual stimulus appeared, (ii) imagining objects in a different region, or (iii) performing no imagery at all. They were asked to say whether they saw the visual image or not, and were significantly less accurate at doing so when they were imagining an object (of whatever sort) in the same region than when they were performing imagery in another region or performing no imagery.

Behavioral measures like the Perky effect show that systems dedicated to spatial cognition are recruited during mental imagery. These same behavioral tools have also been used to investigate the processing of language about space; in order to test the hypothesis that understanding language about space, like performing imagery and recalling aspects of space, makes use of perceptual and motor systems.

A first such line of research investigated whether language denoting action along the horizontal versus the vertical axis produced Perky effects, similar to those described above. When concrete spatial language denotes motion, that motion is often likely to occur along a particular axis. For example, actions like *springing* and *shoving* typically involve vertical or horizontal motion, respectively. Simulation-based theories of language understanding predict that the processing of such language involving motion will automatically drive the activation of dynamic and analog mental simulations, capturing the embodied experience of motion. Preliminary work by Richardson et al. (2001) showed that naïve subjects systematically associated gradient axes of motion with action verbs like *spring* and *shove*. This finding was interpreted as suggesting that language understanders can access image schemas implicating direction of spatial movement during semantic judgment tasks. A subsequent experiment took these same verbs and placed them in the context of a Perky-like task. Subjects first listened to sentences denoting horizontal (1a) or vertical (1b) motion:

(1) a. The miner pushes the cart. (horizontal)
    b. The plane bombs the city. (vertical)

Following this, they then saw a shape – either a circle or a square – flash in either the vertical or the horizontal axis of a computer screen and were asked to press a button as soon as possible to indicate whether the shape was a circle or a square. Naturally, subjects were kept unaware of the experimenters' hypothesis: that the previously presented sentence would interfere with processing the shapes if it denoted motion in the same axis. What Richardson et al. found in this spatial language-induced replication of the Perky effect was that, as predicted, reaction times to the shapes were longer when the implied sentential axis matched that of the picture presentation axis.

Reversing the order of presentation, Lindsay (2003) performed a replication in which he showed subjects an object moving along the horizontal or vertical axis before presenting a sentence involving movement. Reading times once again showed a significant interaction between the direction of the perceived movement of the object and that of the implied movement in the sentence.

Naturally, the question arises how detailed these activated spatial representations are. The visual field can in principle be divided more finely than just into axes, and indeed it would be surprising if internal representations of described events did not distinguish at least between left and right, let alone up and down. To investigate the level of detail displayed by language-driven spatial representations, Bergen et al. (2007) presented subjects with sentences describing actions that would by default occur in either the upper (2a) or lower (2b) part of the visual field:

(2) a. The mule climbed. (upward movement)
b. The pipe dropped. (downward movement)

Bergen et al. measured how long it would take subjects – after hearing these sentences – to categorize shapes that appeared in the upper or lower parts of a computer screen. They found the same Perky-like interference effect described above, with visual objects presented in the same part of the visual field being processed more slowly.

In order to investigate what sentence parts drive spatial imagery, Bergen et al. performed an additional manipulation, in which language stimuli had spatial connotations only by dint of their subject nouns and not their verbs (3).

(3) a. The rainbow faded. (up-related noun)
b. The ground shook. (down-related noun)

The results of this study came back the same as those of the previous one; when verbs are neutral for spatial orientation, a sentence with just an up- or down-connoted noun can drive location-specific spatial imagery. While a full discussion of the ramifications of these finding are beyond the scope of the current paper, they suggest that spatial meanings, perhaps in the form of conceptual primitives like image schemas, appear to be at work not only in function words, as has often been reported, but also in content words like nouns and verbs.

Spatial representations of actions are not limited to static relationships like location along an axis or in a quadrant but can also include direction of movement within the visual spatial field. Zwaan et al. (2004) had language understanders perform a task similar to those described above, but with moving target object stimuli. Subjects first heard a sentence denoting motion that, if perceived, would move away from (4a) or towards (4b) the body of the listener.

(4) a. The shortstop hurled the softball at you. (motion towards the body)
b. You hurled the softball at the shortstop. (motion away from the body)

Subjects then saw two slightly differently sized images of the same object in quick succession, which subtly yielded the appearance of motion away from or towards the subject. They were asked to decide if the two objects were the same or different. Just as in previous studies, subjects' response times to decide if the objects were the same or not was affected by the direction of the sentence they had just heard. But interestingly, in this study, subjects were faster to say the objects were the same if the implied movement was in the same direction as that of the sentence they had heard. The reasons why studies like this one yield compatibility effects, whereas other work has shown interference effects, is still hotly debated (Lindsay 2003, Kaschak et al. 2005, Bergen 2007).

One further finding of note pertains to fictive motion, that is, the use of language about motion through space to describe static scenes (5).

(5) a. The road *runs* across the desert.
b. The road *winds* through a rocky ravine.

In a series of experiments, Matlock (2004) has shown that even fictive motion language yields dynamic spatial imagery. Subjects reading fictive motion sentences like those in (5) take significantly less time when the sentence describes fictive motion across short distances, on smooth terrain or by fast means of travel, as contrasted with descriptions of long, impeded or slow fictive travel. This effect suggests that language understanders build up a mental image of the path of described motion through space as they process spatial language.

The behavioral experiments described above support the idea that, when understanding language involving spatial information, people activate spatial simulation and imagery in a dynamic, analog and modal fashion. In the next section, we examine evidence for the use of these same systems during the processing of spatial language used to describe abstract concepts.

3 Metaphorical language

Metaphorical language uses space as a source domain for a number of basic conceptual target domains. Chief among these are quantity (*tax rates are rising again*), quality (*their newest film is top-notch*) and time (*let's move the meeting forward an hour*). The linguistic facts are unambiguous: spatial language can progressively acquire new conventionalized non-spatial meanings, and it can also be used in novel ways to describe non-spatial scenarios with figurative expressions (*the price of corn has cannonballed*). Nonetheless, this evidence from language leaves open the question of whether, when processing a word with a spatial meaning (like *rising*) to describe a non-spatial event (like an increase in price), language users actually engage their systems for spatial cognition in the same way that the behavioral evidence above suggests they do for literal spatial language understanding. The empirical results we discuss in this section suggest that the processing of abstract target domains, such as time, does indeed involve activation of spatial systems.

There are a number of different ways in which time is linguistically and conceptually cast in terms of space, and any given language is likely to employ a combination of these. Time is commonly viewed as a landscape across which the speaker moves, often labeled the Time Passing is Motion over a Landscape metaphor (6a). Alternatively, it may be viewed as a row of objects that move in relation to a stationary speaker (6b), as in the Time Passing is Motion of an Object metaphor (Lakoff 1993, Boroditsky 2000, see also McTaggart 1908).

(6) a. We're coming up quickly on Easter. (Time Passing is Motion over a Landscape)
b. Easter flew by. (Time Passing is Motion of an Object)

English employs both of these metaphorical construals of time and, in some cases, ordinary expressions can be ambiguous as to the underlying spatial metaphor. For instance, when told that *Wednesday's meeting has been moved forward two days*, the metaphorical forward motion may be interpreted in terms of the motion over a landscape metaphor, in which case forward is defined in terms of the experiencer's direction of motion – into the future so that moving the meeting forward makes it later. It can alternatively be interpreted in terms of the motion of an object metaphor, in which case a line of times move along in a queue with the earliest times first, making forward motion temporally earlier.

Do these two ways of interpreting language about time in terms of space also rely on thinking about space, using neurocognitive resources dedicated to spatial reasoning? Logically, if reasoning about time depends on spatial structures, then inducing language understanders to think either about themselves moving through space, or contrarily about objects moving through space, should lead them to interpret ambiguous temporal language according to the primed spatial schema. In a series of innovative studies, this is precisely what Boroditsky and colleagues have shown. Boroditsky and Ramscar (2002) demonstrated that, when standing at the end of a line or waiting for someone to arrive, a speaker is more likely to adopt the Time Passing is Motion of an Object metaphor when interpreting ambiguous descriptions of temporal events. In contrast, when first entering a vehicle or preparing to disembark during the course of a long journey, a speaker is more likely to employ the Time Passing is Motion over a Landscape metaphor (Boroditsky and Ramscar 2002). In other words, interpreting language about time seems to depend upon contextually modulated activation of spatial knowledge.

While all languages cast time as space, the dimensions of space involved in metaphorical time language can vary across languages. Chinese, for example, employs not only the front-back axis to describe past and future events, as in English, but also the vertical dimension. The Chinese character *shang* ('up') is used in compound words that refer to past events, while *xia* ('down') denotes future events (Yu 1998). The psychological reality of this up/down metaphorical mapping is supported by experiments showing that native Chinese speakers do indeed conceive of time as abstractly laid out along the vertical axis (Boroditsky 2001). In these studies, subjects were primed by pictures to think about vertical or horizontal motion of objects and then asked to answer a temporal question (*Is April before May?*). Even when performing the task in English, native speakers of Mandarin showed better performance on the time question when they had just performed a vertical spatial reasoning task, as compared with native English speakers, whose correct time responses were faster following a horizontal spatial reasoning task.

Another way that languages vary in their use of spatial terms for time is the orientation of the speaker within the Time Passing is Motion over a Landscape metaphor. English and many other languages describe future events as being in front of a speaker (*We're just coming up on midterms now*) and past events as behind the speaker (*I'm so glad we're past the rainy season*). However, in Aymara, a language spoken by indigenous people of Bolivia, Peru and Chile, as well as other languages, time is conceived of and described as though the future were behind and the past ahead. Aymara speakers say *quipa pacha* (literally 'behind time') to refer to the future and *nayra pacha* (literally 'sight time' or 'front time') to refer to the past. At first blush, this arrangment is jarring to speakers of languages like English that use the reverse orientation. But it is quite well motivated. The past is known, thus seeable, thus in front, while the future is unknown and, as such, still hidden or unseen and behind. The backwards-motion-through-time perspective that underlies metaphorical Aymara expressions can also be seen in the gestures that monolingual Aymara speakers use when referring to temporal events. When describing events occurring in the past, they gesture only toward the front, but when referring to the future they gesture exclusively toward the back (Núñez and Sweetser 2006).

Languages can also differ in the number of dimensions they use to measure time. English and Indonesian, among many others, commonly describe the duration of an event with linear spatial descriptors: *a <u>short</u> wait*. In contrast, Greek and Spanish speakers tend to describe event durations in terms of volume rather than distance, with expressions equivalent to *it took <u>much</u> time*. To what extent, though, do these language differences result in differences in cognitive processing of space and time independently of language? Casasanto et al. (2004) addressed this question through a series of psychophysical experiments.

In their first experiment, Casasanto et al. requested native English, Indonesian, Greek and Spanish speakers to state the most natural phrases in their languages describing a large period or a long period of time. As predicted, English and Indonesian speakers used expressions corresponding to *long time*, while Greek and Spanish responses predominantly described *much time*. To determine whether there were relativistic effects of these metaphors on speakers' cognition, Casasanto et al. presented English and Indonesian speakers (who tend to quantify time linearly) with a video of a growing line and asked them to estimate the period of time for which it was presented on a screen. As predicted, the length of the line interfered with subjects' judgments of temporal length: the longer the line was spatially, the more time subjects thought it had remained on the screen. However, the reverse was not found: duration of display did not affect subjects' judgments of spatial length. Showing moving images of an abstract 'container' gradually filling also interfered with English and Indonesian speakers' temporal judgments, but the dynamic container displays did so to a much lesser extent than the linear motion displays. In contrast, the temporal reasoning of Greek and Spanish speakers was modulated to a greater degree by the filling-container animation than by the growing-line animation (Casasanto et al., 2004, Casasanto and Boroditsky 2003). In other words, cross-linguistic differences in mappings from space to time correlated with non-linguistic differences in the extent to which speakers' temporal judgments were influenced by their spatial perception.

We started this section with the well-worn observation that time can be described using spatial terms, in language after language, even among those with no historical genetic relationship. Cross-linguistic variations in the directions that are mapped (front/back in English and up/down in Chinese), in the orientation of a speaker in relation to temporal events (future-facing English speakers versus past-facing Aymara speakers), and in the image schemas appropriated for temporal terms (*long* periods of time for English and Indonesian speakers versus *much* duration of time for Greek and Spanish speakers) correlate with cross-linguistic differences in the behavior of speakers of those languages in reasoning and psychophysical tasks. All of this goes to show that metaphorical language about time is grounded in spatial processes, and additionally that the ways in which a language construes time in terms of space modulate its speakers' conceptual representations of time.

4 The spatial brain

If processing space and processing language about space really do use a shared biological substrate, then this should be corroborated by imaging studies of the living brain. The brain exhibits a good deal of localization according to function, with certain regions, like the well-known Broca's and Wernicke's areas of the left hemisphere, often selectively active during language behavior. Other areas, such as the visual cortex and the so-called parietal *where* pathway, are active predominantly in the right hemisphere during the processing of spatial scenes. Despite evidence that language and space are localized in discrete neuroanatomical regions, however, recent neurophysiological research indicates that there is overlap between the structures associated with attending to spatial relations and processing language about such relations (Kemmerer 2006). This result corroborates the behavioral evidence described in section 2 above.

The parietal cortex houses neural regions involved in attending to and processing spatial relationships. These same areas become active during retrieval of words identifying spatial relationships. Using positron emission tomography (PET), Damasio et al. (2001) imaged the brains of subjects performing naming and spatial relation judgments. In the first, they were presented with static pictures involving two objects in a spatial relation (*X is on Y*) and were asked to name the item (*X*), and in the second, they had to name the spatial relationship between them (*on*). Results showed that the regions dedicated to perceiving spatial relationships – left parietal and frontal cortices, and in particular the supramarginal gyrus (SMG) – were also active during spatial language processing. In other words, the neural structures necessary for perceiving and understanding spatial relationships appear to get selectively activated for retrieval of language describing spatial relations. Emmorey et al. (2002) also showed left SMG activation in bilingual speakers of English and American Sign Language (ASL) in a naming task similar to Damasio et al. (2001). The critical region was more active when the subjects were naming spatial relations than when they named objects, implicating this area in semantic processing.

The significance of the SMG in spatial language processing has also been subsequently reinforced by a series of studies by Kemmerer (2005) with subjects who had incurred damage to the SMG. Kemmerer asked subjects to perform a task that used prepositions in a spatial and then temporal test (*at* the store; *at* nine o'clock), and subjects were required to fill in the appropriate preposition in given contexts. Subjects with damage to the left SMG performed well on the temporal test but poorly on the spatial test, while the subject with no damage to this region was able to competently produce spatial prepositions. These results not only underscore the importance of the SMG – a region dedicated to spatial perception – in processing spatial language, but also raise important considerations for studying the neurological basis for metaphorical uses of spatial terms. The results showed a double dissociation between temporal and spatial prepositions, suggesting that independent neural substrates can process the same prepositions when used spatially versus temporally.

These studies show that language users employ neural structures initially devoted to concrete spatial reasoning when processing and producing language about spatial relations. This overlap of neural regions is critical to a theory of spatial language processing that is grounded in spatial cognition. At the same time, the evidence that spatial and temporal uses of prepositions like *at* do not require identical neural substrates implies that metaphorical language using space as a source domain is not processed in an identical fashion as are literal, spatial uses. This is hardly surprising, since it could not in principle be the case that spatial and temporal uses of prepositions for example have precisely the same neural substrates. Without different neural patterns, there would be no behavioral or subjective differences between these different uses. However, it does open up important questions about the relation between spatial and metaphorical uses of spatial language. How is the apparent use of spatial cognitive mechanisms, evidenced by the behavioral studies cited above, realized in neural terms? What effect might the conventionality of the metaphorical sense of a word have on the likelihood that its use will engage the spatial cognition system? Regardless of the answers, which we can only hope that future work will give hints to, the neural basis of spatial language processing appears to overlap significantly with that of spatial cognition.

5 Computational models of language and spatial cognition

Due to the complexity of human linguistic cognition, computational models of language learning and use are also valuable tools for testing the relative viability of competing views. The most successful models of spatial language bootstrap linguistic behavior off of representations of the spatial cognition system. Models of the human semantic potential for learning static (e.g. *at*) and dynamic (e.g. *into*) spatial language (Regier 1996) and the evolution and acquisition of language for spatial perspectives (Steels et al. In Prep) have succeeded when spatial language is learned through computational mechanisms responsible for aspects of spatial cognition. The success of these models supports the view that the human language faculty is not a discrete component within the mind, but rather the product of many interconnected units, including, in the case of spatial language learning and use, some dedicated principally to spatial cognition.

In developing a computational model of the acquisition of spatial terms like *in*, *out*, *into*, and *out of*, Regier (1996) drew inspiration from the architecture of the neural cortex, with discrete computational subunits for (i) creation and comparison of perceptual maps; (ii) orientation and directional awareness on the basis of perceptual maps; (iii) motion detection and analysis; and (iv) association of signals from the above three subunits with an array of locative prepositions. The first three structures within this architecture process stimuli in a manner globally similar to that of the human visual cortex, while the fourth serves as an interface between perceptual representations and the lexicon.

Regier's model incorporates several explicit constraints, or principles, that guide the classification of spatial relationships according to sets of primitive features. For example, an 'endpoint configuration constraint' allows the model to recognize the static perceptual feature of inclusion with a series of images showing movement of a trajectory from the exterior to the interior of a landmark, which can then be associated with an appropriate word such as *into*. This endpoint configuration constraint mirrors findings of behavioral studies in developmental psychology indicating that children categorize events more often on the basis of their results than by event-interior relationships (Behrend 1989, 1990, Smiley and Huttenlocher 1994) and provides a computational mechanism for linguistic processing according to a Source-Path-Goal image schema (Lakoff 1987).

Using this architecture, Regier's model can learn the correct classifications of spatial percepts according to sets of spatial terms from English, German, Japanese, Mixtec or Russian. Since each of these languages groups spatial features differently in its encoding of spatial relationships, Regier's model supports the idea that spatial language learning and use is grounded in a primitive and universal set of computational mechanisms arrayed in the human perceptual cognitive system.

This model, however insightful, is based on the implausible assumption that spatial language describes scenes viewed from an invariant perspective. Of course, in the real world, this is rarely true: two speech participants tossing a ball back and forth will have dramatically different views of the scene, where what is on the right for one will be on the left for the other, and what is close for one will be far for the other. Successful communication about spatial scenes thus requires language that responds to these differences in perspective. One key tool that languages provide for this end is the use of perspective encoding, namely, language indicating the perspective from which a particular description holds. In English, possessive nouns can identify the orientation from which an object description holds (*my right*, *your right*, *John's right*). On the basis of the evidence described in the preceding sections, we might hypothesize that learning and using language describing different perspectives relies on language users engaging those components of the spatial cognition system responsible for adopting alternative spatial perspectives. In other words, to calculate whether the ball is on John's right, a language user might have to perform a mental rotation of the scene as they see it (Shepard and Metzler 1971) to envision what it would look like from John's perspective.

Steels et al. (In Prep) conducted a series of simulations where robots were programmed to describe for each other scenes in which objects moved in one direction or another, but where speakers did not necessarily share the same perspective as their interlocutors. The aim was to determine whether endowing these communicating agents with the ability to mentally rotate a perceived scene would facilitate their communicative success. The agents for this study were a community of autonomous robots endowed with various processing modules including (i) a real-time image processing system; (ii) probabilistic modeling of a three-dimensionally perceived visual world; (iii) active vision systems to track moving objects; (iv) motor control and obstacle avoidance mechanisms; and (v) behavioral mechanisms for exploration of the immediate environment, establishment of joint attention and communication of motion events.

During a given trial of the experiment, two robots explored a room until finding a colored ball and establishing joint attention to it. The human experimenter would then move the ball and, after perceiving this movement, the robot agents verbally described the movement. In each trial, if the communication was successful (i.e., if the first robot was able to describe the movement in terms that the second robot could recognize as matching the perceived event), the cognitive and linguistic mechanisms used for the communication task were positively reinforced. If the communication was unsuccessful, however, the cognitive and linguistic mechanisms involved were incrementally inhibited for future trials. Over the course of a large number of trials (usually on the order of several thousand), the population collaboratively evolved linguistic terms for movement events.

By manipulating the robots' cognitive mechanisms, Steels et al. (In Prep) discovered that the agents were much more successful at evolving adequate spatial language when endowed with a cognitive mechanism allowing them to adopt their interlocutor's perspective before formulating or interpreting an utterance, in comparison to those that were not endowed with such an ability. Moreover, when allowed to invent new words to indicate the perspective from which a spatial description held, they consistently did so. This study thus suggests that perspective reversal through mental rotation of an egocentric view to simulate that of the hearer allows for more efficient development of language. In light of this finding, it comes as no surprise that human languages universally encode perspective in spatial language (*<u>my</u> left*, *in front <u>of you</u>*).

These experiments assume that certain neurocognitive resources used for human language, such as the mechanisms required to perform mental rotation or to calculate orientation and direction, ontogenetically precede linguistic capacities. If these computational models are any indication, human language learning and use seems to draw from other existing specialized cognitive systems. For the autonomous agents used in the last study, language development and use were most successful when mechanisms for movement recognition and mental rotation systems were accessible for recruitment. Just like the behavioral and imaging evidence above, the computational models described here indicate a critical role for cognitive mechanisms dedicated to space in the use of spatial language.

6 Conclusions

The behavioral, neural, and computational evidence surveyed above suggests that language users recruit cognitive systems dedicated to spatial cognition when processing language with spatial content, and that such systems also appear to be used for processing metaphorical spatial language about abstract domains. These findings coalesce with those of related work on language processing, which similarly shows that language about action is understood through the activation of motor circuitry (Glenberg and Kaschak 2002, Pulvermueller et al. 2001), and that language about visually perceptible scenes is understood through the activation of visual imagery of the implied shape and orientation of described objects (Stanfield and Zwaan 2001, Zwaan et al. 2002). At the same time, evidence suggesting that spatial language used for metaphorical purposes recruits spatial cognition is in line with other work showing that metaphorical language using other source domains like containment, searching, and possession are also activated when used to metaphorically structure abstract domains like emotional states (Tseng et al. 2005, Sato et al. 2006).

The picture of the human capacity for language that emerges from these convergent results is one where linguistic capacities recruit and make consistent use of existing cognitive systems – in the case of the studies described in this paper, systems dedicated to spatial cognition. The use of spatial language, whether literal or metaphorical, appears to involve a process whereby the spatial configurations described by language trigger the internal reconstruction by the language user of experiences that are akin to actually perceiving motion through or relations in space. In processing language about space, an understander recreates the described spatial experience. This mechanism may serve to explain large expanses of what it means to deeply understand spatial language.

References

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660.

Behrend, D. (1989) Default values in verb frames: Cognitive biases for learning verb meanings. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

Behrend, D. (1990) The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development* 61: 681–696.

Bergen, B. (2007) Experimental methods for simulation semantics. In Gonzalez-Marquez, M., Mittelberg, I., Coulson, S. and Spivey, M. J. (eds) *Methods in cognitive linguistics*. Amsterdam: John Benjamins.

Bergen, B., Lindsay, S., Matlock, T. and S. Narayanan. (2007) Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science* 31: 733–764.

Boroditsky, L. (2000) Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* 75(1): 1–28.

Boroditsky, L. (2001) Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology* 43(1): 1–22.

Boroditsky, L. and Ramscar, M. (2002) The roles of body and mind in abstract thought. *Psychological Science* 13(2): 185–188.

Bowerman, M. and Pederson, E. (1992) Crosslinguistic perspectives on topological spatial relationships. Paper presented at the 87th annual meeting of the American Anthropological Association, San Francisco, CA.

Casasanto, D. and Boroditsky, L. (2003) Do we think about time in terms of space? In *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., Fotokopoulu, O., Pita, R. and Gil, D. (2004) How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek and Spanish. In *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Lawrence Erlbaum.

Craver-Lemley, C. and Arterberry, M. (2001) Imagery-induced interference on a visual detection task. *Spatial Vision* 14: 101–119.

Damasio, H., Grabowski, T. J., Tranel, D. Ponto, L. L. B., Hichwa, R. D. and Damasio, A. R. (2001) Neural correlates of naming actions and naming spatial relations. *NeuroImage* 13: 1053–1064.

Emmorey, K., Damasio, H., McCullough, S., Grabowski, T., Ponto, L. L. B., Hichwa, R. D. and Bellugi, U. (2002) Neural systems underlying spatial language in American Sign Language. *NeuroImage* 17: 812–824.

Glenberg, A. M. and Kaschak, M. P. (2002) Grounding language in action. *Psychonomic Bulletin & Review* 9: 558–565.

Johnson, M. (1987) *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: The University of Chicago Press.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. and Zwaan, R. A. (2005) Perception of motion affects language processing. *Cognition* 94(3): B79–B89.

Kemmerer, D., (2005) The spatial and temporal meanings of English prepositions can be independently impaired. *Neuropsychologia*, 43, 797–806.

Kemmerer, D. (2006) The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia* 44: 1607–1621.

Lakoff, G. (1987) *Women, fire and dangerous things*. Chicago: The University of Chicago Press.

Lakoff, G. (1993) The contemporary theory of metaphor. In Ortony, A. (ed.) *Metaphor and thought* 202–251. Cambridge: Cambridge University Press.

Lakoff, G. and Johnson, M. (1980) *Metaphors we live by.* Chicago: University of Chicago.

Lindsay, S. (2003) *Visual priming of language comprehension.* Unpublished University of Sussex Master's Thesis.

Majid, A., Bowerman, M., Kita, S., Haun, D. and Levinson, S. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3): 108–114.

Matlock, T. (2004) Fictive motion as cognitive simulation. *Memory & Cognition* 32: 1389–1400.

McTaggart, J. (1908) The unreality of time. *Mind* 17: 457–474.

Núñez, R. and Sweetser, E. (2006) With the future behind them: Convergent evidence From Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30(3): 401–450.

Perky, C. W. (1910) An experimental study of imagination. *American Journal of Psychology* 21: 422–452.

Pulvermüller, F., Haerle, M. and Hummel, F. (2001) Walking or talking?: Behavioral and neurophysiological correlates of action verb processing. *Brain and Language* 78: 143–168.

Regier, T. (1996) *The human semantic potential*. Cambridge, MA: The MIT Press.

Richardson, D., Spivey, M., Edelman, S. and Naples, A. (2001) Language is spatial: Experimental evidence for image schemas of concrete and abstract spatial representations of verbs. *Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society* 873–878. Mawhah, NJ: Erlbaum.

Sato, M., Schafer A. and Bergen, B. (2006) Effects of picture perception on the expression of abstract concepts in sentence production. Paper presented at the 19th annual CUNY Conference on Human Sentence Processing, New York, NY.

Shepard, R. N. and Metzler, J. (1971) Mental rotation of three-dimensional objects. *Science* 171: 701–703.

Smiley, P. and Huttenlocher, J. (1994) Conceptual development and the child's early words for events, objects and persons. In Tomasello, M. and Edward, W. (eds) *Beyond names for things: Young children's acquisition of verbs*. Hillsdale, NJ: Lawrence Erlbaum.

Steels, L. and Loetzsch, M. (2009) Perspective alignment in spatial language. In Coventry, K. R., Tenbrink, T. and Bateman, J. A. (eds) *Spatial language and dialogue*. Oxford: Oxford University Press.

Stanfield, R. A. and Zwaan, R. A. (2001) The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12: 153–156.

Sweetser, E. (1990) *From etymology to pragmatics*. Cambridge: Cambridge University Press.

Talmy, L. (2000) *Toward a cognitive semantics*. (2 volumes) Cambridge, MA: MIT Press.

Tseng, M., Hu, Y., Han, W. and Bergen, B. (2005) 'Searching for happiness' or 'Full of joy'? Source domain activation matters. In *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Yu, N. (1998) *The contemporary theory of metaphor: A perspective from Chinese*. Amsterdam: John Benjamins.

Zwaan, R. A., Madden, C. J., Yaxley, R. H. and Aveyard, M. E. (2004) Moving words: Dynamic mental representations in language comprehension. *Cognitive Science* 28: 611–619.

Zwaan, R. A., Stanfield, R. A. and Yaxley, R. H. (2002) Do language comprehenders routinely represent the shapes of objects? *Psychological Science* 13: 168–171.

Zwaan, R. A. and Taylor, L. J. (2006) Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General* 135(1): 1–11.

Part III

Typological, psycholinguistic and neurolinguistic approaches to spatial representation

4 Inside *in* and *on*: typological and psycholinguistic perspectives

Michele I. Feist

1 Introduction

Spatial language offers us many windows onto the landscape of human spatial cognition. But how can we best understand the insights offered by spatial language? What do we pay attention to when we talk about space? Researchers investigating these questions have suggested a variety of factors, often individually. How then to make sense of this complex landscape?

In this chapter, I will sketch the view through two windows onto the landscape of spatial cognition: one being that of a semantic typologist; the other, that of a psycholinguist. The evidence gathered by looking through these two windows will suggest that despite surface differences in how we talk about space, all humans are attuned to the same three abstract families of factors – geometric, functional, and qualitative physical – which together influence the ways in which we talk about relations in space. I will examine each of these families of factors in turn, along with limitations on proposed meanings based on a single type of factor.

The importance of geometry to the meanings of spatial relational terms has long been noted (Bennett, 1975; Feist, 2000; Feist and Gentner, 2003; Herskovits, 1986; Landau, 1996; Lindkvist, 1950; Miller and Johnson-Laird, 1976; Talmy, 1983; Tyler and Evans, 2003). Geometry includes information such as the relative vertical and horizontal positions of the Figure and Ground,[1] their proximity to one another (with inclusion being the closest possibility and contact the next closest), their shapes, and their relative sizes. Such information forms the basis of many proposed meanings of topological spatial prepositions, exemplified by the following two researchers' proposed meanings for *in*:

(1) A[locative[interior of B]]
(Bennett, 1975, p. 71)

(2) inclusion of a geometric construct in a one-, two-, or three-dimensional geometric construct
(Herskovits, 1986, p. 48)

Consistent with geometric approaches to spatial meaning, it has been found that simply changing the geometric relations in a spatial scene can shift speakers' intuitions regarding the most appropriate preposition to describe the scene (Coventry and Prat-Sala, 2001; Coventry, Prat-Sala and Richards, 2001; Feist, 2000, 2002; Feist and Gentner, 1998, 2003). For example, Coventry and Prat-Sala (2001) showed participants piles of objects placed in containers. They varied the heights of the piles, placing the Figure at the very top, then asked participants to rate the appropriateness of *in*, *over*, and *above* to the resultant scenes. They found that this manipulation resulted in higher ratings for *in* when the piles were low, and for *over* and *above* when the piles were high.

Although they are intuitively appealing, there are a variety of problems with representations of the semantics of spatial relational terms based solely on geometry. First, and most importantly, there are many static spatial uses that cannot be accounted for by a purely geometric meaning. A simple example will suffice. Consider the two proposed meanings for *in* cited above. In both cases, *in* is described as applicable to situations in which the Figure is located at the interior of, or included in, the Ground, as in the case of the pear in the bowl in Figure 1a. However, many spatial terms used to describe situations of full inclusion, like English *in*, can also be used for partial inclusion (Figure 1b; cf. Levinson, Meira and the Language and Cognition Group, 2003) or, in some cases, situations in which the Figure is not geometrically included in the Ground at all (Figure 1c). It is difficult for a geometric approach to account for such uses.

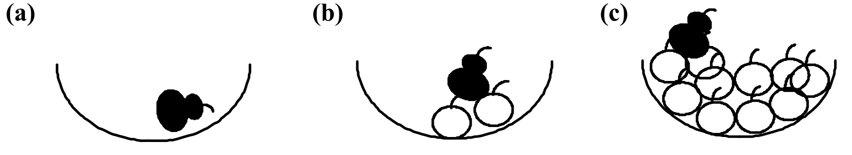


Figure 1. Three pears in three bowls

A second problem faced by geometric accounts of spatial relational meaning is the existence of multiple possible descriptions for a single scene, as demonstrated in example (3). Although one can argue that there are distinct shades of meaning, or conceptualizations (Tyler and Evans, 2003), corresponding to the two sentences, the fact remains that there is but one geometric relation being described. In addition to failing to motivate alternate conceptualizations, purely geometric approaches are unable to provide a principled means of explaining why a speaker might choose one over the other for a particular situation.

(3) (a) The players are *on* the field.
(b) The players are *in* the field.

More recently, researchers have begun to argue that the meanings of spatial relational terms rely crucially on functional attributes of spatial scenes (Coventry, Carmichael and Garrod, 1994; Coventry and Garrod, 2004; Coventry and Prat-Sala, 2001; Feist, 2000, 2005b; Feist and Gentner, 2003; Vandeloise, 1991, 1994), as in the proposed meanings in (4) and (5). Functional attributes include knowledge about the normal uses (if any) of the objects (particularly the Ground), with special attention to the purpose for which they were created (Coventry et al., 1994; Feist, 2000, 2002; Feist and Gentner, 1998, 2003; Vandeloise, 1991), knowledge about whether or not the Figure and Ground normally interact (Coventry and Prat-Sala, 2001), and knowledge of the manner in which they are interacting in the current scene.

(4) D/H: *a est* [=is] *dans/hors de b* if the landmark and the target are/are no longer the first and second elements in the container/contained relation.
(Vandeloise, 1991, p. 222)

(5) in: functional containment – *in* is appropriate if the [G]round is conceived of as fulfilling its containment function.
(Coventry et al., 1994)

Consistent with such analyses, Coventry and his colleagues (Coventry et al., 1994) found that the typical function of the Ground object influenced participants' judgments about the applicability of spatial relational terms: solid objects were judged more *in* bowls, which typically hold solids, than jugs, which more typically hold liquids. Similarly, Feist (2000; Feist and Gentner, 1998, 2003; see below) found that participants were reliably more likely to use *in* than *on* if a pictured Ground was labeled as a *bowl* rather than a *plate*, despite the fact that all participants saw the same picture.

The functional approach provides a superior explanation for the range of pictures in Figure 1, as the bowl in each case is fulfilling its usual function as a container, motivating the use of *in*. The approach meets up with problems, however, when the Ground object does not have a normal function (as, for example, in the case of natural kinds), or when it is filling a qualitative physical role different from its normal function (see below). In such situations, it is unclear how a functional approach might predict speakers' uses of spatial relational terms.

Finally, it has been suggested that the meanings of spatial relational terms are influenced by the qualitative physics of the spatial scene per se (Bowerman and Choi, 2001; Bowerman and Pederson, 1992, 1996; Feist, 2000, 2005a, 2005b; Feist and Gentner, 2003; Forbus, 1983, 1984; Talmy, 1988; Vandeloise, 2003). Although considerably less attention has been paid to the independent role of qualitative physical attributes (such attributes, in fact, do not form the basis for any proposed spatial prepositional meanings), these may prove to be equal to geometry and function in their importance. By qualitative physics, I am referring to information about the physics of the configuration, including the presence or absence of a support relation and the ability of one entity to control the movement of itself or another (cf. Coventry and Garrod's 2004 discussion of location control). Often, qualitative physical aspects of the scene result from functional features, as when a canonical container fulfills its typical function by constraining the location of another entity. However, this is not always the case. As a case in point, the typical function of an umbrella is to protect the bearer from falling rain. In the scene in Figure 2, however, the umbrella is constraining the location of the apple, motivating the appropriate use of *in*. As this example shows, it is important to carefully separate qualitative physical and functional features, despite their normal co-occurrence.

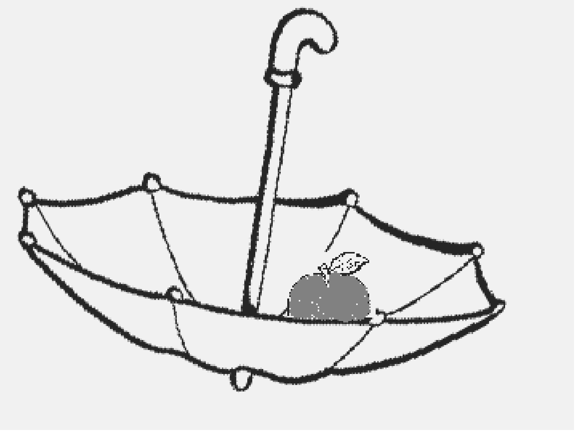


Figure 2. An apple in an umbrella

Although much theoretical work has suggested important roles for geometry, function, and qualitative physics in the semantics of spatial relational terms, there remain large gaps in our knowledge. First, most proposed meanings of spatial relational terms, such as those cited above, have their basis in a single feature, noting other aspects only as they support the prominent feature (as, for example, geometric inclusion is a characteristic of the functional containment relation (Vandeloise, 1991)). Such a view of spatial meaning, however, leaves many uses of spatial relational terms – even static spatial uses – unexplained (Feist, 2000), as outlined above. Second, the majority of the work to date has considered a single language (most commonly English). Yet because linguistic typology helps to separate out the motivated and explainable from the arbitrary (Croft, 1999), a deep understanding of the semantics of spatial terms may benefit from a wider crosslinguistic perspective. Third, while the roles of geometry, function, and qualitative physics have been suggested, their importance awaits detailed empirical verification (although there have been some efforts in this area, as noted above). To address these gaps, I will describe two studies. The first, a crosslinguistic survey, addresses the question of which, if any, of the identified factors recur in the spatial vocabularies of a variety of languages. The second, a psycholinguistic experiment, addresses the question of whether small changes in the geometric, functional, and qualitative physical attributes of a depicted spatial relationship will lead to concomitant changes in speakers' use of English spatial prepositions, thus providing empirical evidence for the importance of these factors to English prepositional meaning. As such, I will be presenting two complementary views onto the landscape of factors that combine to make up spatial relational meanings – one typological and one psycholinguistic. What we seek are the organizing principles around which spatial vocabularies are built.

2 A view through the window of typology

If there is any domain where we might expect universals, it is surely space, due in part to the universality of our early experience with space (Clark, 1973). It is perhaps this assumption that has led researchers to examine the semantics of spatial terms largely in single languages, as the simple topological notions into which spatial terms have been decomposed (Bennett, 1975; Herskovits, 1986; Miller and Johnson-Laird, 1976) are largely considered universal, with neurocognitive correlates (Landau and Jackendoff, 1993). In contrast to this intuition, however, the variation in spatial descriptions that has been uncovered in crosslinguistic studies is astonishing (Bowerman and Choi, 2001; Bowerman and Pederson, 1992, 1996; Brown, 1994; Feist, 2000, 2004, 2008; Gentner and Bowerman, 1996, 2009; Levinson et al., 2003; Majid, Bowerman, Kita, Haun and Levinson, 2004; Pederson, Danziger, Wilkins, Levinson, Kita and Senft, 1998; Sinha and Thorseng, 1995; Sinha, Thorseng, Hayashi and Plunkett, 1994). Careful examination of the extensional range of spatial terms in multiple languages further suggests that the very dimensions of variation may differ across languages, as in the oft-cited difference between English and Korean spatial terms (Bowerman and Choi, 2001). A simple example will illustrate this difference. Imagine two scenes: a cassette in its case, and an apple in a bowl. In English, the two scenes would be described using the same word, as both are instances of inclusion. In Korean, however, it would be inappropriate to describe them alike, as one (the cassette in its case) is an instance of tight fit, while the other (the apple in the bowl) is an instance of loose fit. In describing these two scenes, the dimensions of contrast that are important in English and Korean are in fact quite different (but see Vandeloise, 2003, this volume, for an alternate view of this distinction).

Does this mean that the sets of attributes of spatial scenes that languages encode in their spatial relational vocabularies are incommensurable? Perhaps not. Consider again the English-Korean distinction. English *in* communicates inclusion, which is both geometric, and (due to our three-dimensional gravitational world) physical. Korean, on the other hand, distinguishes tight and loose fit – a qualitative physical (Vandeloise, 2003, this volume) and geometric distinction. Thus, despite surface differences in the ways in which words map to scenes, there are similarities at the abstract level of attention to geometry and qualitative physics. This explanation echoes the findings of Levinson and his colleagues (2003), who suggested that there may be universal 'attractors', or abstract relations which languages will tend to recognize. This is also in line with Croft and Poole's (2008) suggestion that what is universal across languages may be the constraints on variation, rather than the specifics of how languages work (see also Feist, 2008).

In addition to uncovering abstract similarities in the semantics of spatial relational terms – and verifying them across a wide range of languages – there is yet another reason to examine the typology of spatial semantics. By including more languages in a sample, we increase the chances that potentially important factors will be identified, as in the identification of tight vs. loose fit as a result of studying Korean. In addition to shedding light on human spatial cognition in their own right, some of these factors may prove relevant even in languages where they were previously discounted. As a case in point, attributes of the Figure object have largely been considered unimportant to the uses of English spatial prepositions (Landau and Stecker, 1990; Talmy, 1983). Looking across languages, this is by no means a universal fact about spatial relational terms. For instance, in Mayan languages such as Tzeltal, the nature of the Figure seems to carry particular importance in the selection of a spatial relational term to describe a scene (Brown, 1994; Levinson, 1996). Upon reexamination of the role of the Figure in the use of the English prepositions *in* and *on,* Feist (2000; Feist and Gentner 2003; see below) found a small but reliable effect, suggesting that the role of the Figure had been mistakenly discounted in previous accounts of English spatial meanings.

Although the field of semantic typology is still in its infancy, seminal work has already laid the foundations for important advances in our understanding of the ways in which languages categorize spatial relations (Bowerman and Choi, 2001; Bowerman and Pederson, 1992, 1996; Feist, 2008; Levinson et al., 2003). I will here describe one further contribution to this growing area (for complete details of this study, see Feist 2000, 2004, 2008), based on the pioneering work of Bowerman and Pederson (1992; 1996).

Bowerman and Pederson elicited descriptions of a range of topological spatial relations from speakers of thirty-four languages (see also Levinson et al., 2003), using a single set of pictures to elicit the terms from all of the languages in a uniform manner. Their findings illustrated a number of facts about the extensions of spatial terms across a range of languages. First, none of the languages in their sample used separate terms for each of the relations exemplified by pictures in their set. Rather, spatial terms in each of the languages grouped together multiple spatial relations for the purpose of communication. This finding is important, as it validates the study of the extensions of spatial relational terms as a means of examining those factors of spatial scenes that humans deem important. By examining the ways in which the elicited spatial terms grouped the pictures in their set, Bowerman and Pederson were able to infer the kinds of semantic distinctions that tend to appear in spatial language. They found that, along with prodigious cross-linguistic variation, there was a striking commonality. The pictures in their set could be arranged in a semantic map (Haspelmath, 2003), or 'similarity gradient' (Bowerman and Choi, 2001), over which the range of application of each of the elicited terms could be mapped. Further, in keeping with Croft's Semantic Map Connectivity Hypothesis (Croft, 2001, 2003; Croft and Poole, 2008), Bowerman and Pederson found that none of the terms which they had elicited grouped together discontinuous portions of their similarity gradient. This systematicity suggests that significant variation co-exists with deep commonality.

By presenting a single set of pictures to speakers of a wide variety of languages, Bowerman and Pederson were able to directly compare the extensions of the languages' spatial terms. Inspired by this, my study borrows Bowerman and Pederson's methodology in order to elicit a data set from which the crosslinguistic importance of particular attributes to the semantics of spatial relational terms may be inferred. If geometry, function, and qualitative physics are important structuring elements for human spatial cognition, we can expect to see their influence in the spatial terms of a variety of unrelated languages.

Twenty-nine simple line drawings, each depicting two objects in a topological spatial relation, were used in this study. In each picture, one object (the Figure) was colored in yellow; the second object (the Ground) was left in black and white. Twenty-seven of the drawings were borrowed from Melissa Bowerman and Eric Pederson's Topological Picture Series (Bowerman and Pederson, 1992, 1996; Gentner and Bowerman, 1996, 2009; Levinson et al., 2003), one of the remaining two was modified from the Topological Picture Series, and the final one was borrowed from an example in Coventry (1998). Participants were asked to describe the locations of the yellow objects with respect to the other objects in the most natural manner. Twenty-seven speakers volunteered to describe the picture series, providing terms from sixteen languages and nine language families. The languages are listed, along with their genetic affiliations[2] and the number of speakers participating, in Table 1.

Table 1. Languages surveyed in the crosslinguistic study

Language	Language Family	Number of speakers in sample
Polish	Indo-European, Slavic, West, Lechitic	3
Russian	Indo-European, Slavic, East	2
Croatian	Indo-European, Slavic, South, Western	1
German	Indo-European, Germanic, West, Continental, High	3
Swedish	Indo-European, Germanic, North, East Scandinavian	1
Italian	Indo-European, Italic, Romance, Italo-Western, Italo-Romance	1
French	Indo-European, Italic, Romance, Italo-Western, Western, Gallo-Romance, North	2
Hindi	Indo-European, Indo-Iranian, Indo-Aryan, Central zone, Western Hindi, Hindustani	2
Hebrew	Afro-Asiatic, Semitic, Central, South, Canaanite	3
Hungarian	Uralic, Finno-Ugric, Ugric, Hungarian	2
Cantonese	Sino-Tibetan, Chinese	1
Telegu	Dravidian, South-Central, Telugu	1
Turkish	Altaic, Turkic, Southern, Turkish	1
Tagalog	Austronesian, Malayo-Polynesian, Western Malayo-Polynesian, Meso Philippine, Central Philippine, Tagalog	2
Japanese	Japanese, Japanese	1
Korean	Language Isolate[3]	1

In order to understand the ways in which a small set of attributes may influence the use of spatial relational terms across the language sample, the pictures were first analyzed separately from the elicited terms. Afterwards, the analysis of the pictures was combined with an examination of the extensional maps of each of the elicited terms in order to isolate attributes which may influence the uses of the terms.

First, each of the pictures was coded for whether it matched each of a small set of geometric, functional, and qualitative physical attributes. The set of attributes was chosen largely from characterizations of spatial terms in the literature. The geometric attributes examined were:

a difference in vertical position – important to terms such as *above, below, over,* and *under* (O'Keefe, 1996; Tyler and Evans, 2003)

contact – important to terms such as *on* (Cienki, 1989; Herskovits, 1986; Miller and Johnson-Laird, 1976)

inclusion[4] – important to terms such as *in* (Cienki, 1989; Herskovits, 1986; Miller and Johnson-Laird, 1976; Tyler and Evans, 2003)

relative size – not cited in the literature, but chosen because a larger Ground might facilitate other attributes, such as *inclusion* (above) and *support* (below).

One functional attribute – the presence of a functional relation based on the Ground's typical function (Coventry et al., 1994; Vandeloise, 1991, 1994) – was examined. To make this concrete, coffee and a coffee cup are functionally related, as the typical function of a cup is to contain a volume of liquid. As such, a functional relation would be coded as present for a picture of coffee in a coffee cup. On the other hand, a cloud and a mountain are not functionally related, and a functional relation would be coded as absent for a picture of a cloud over a mountain.

Finally, the following three qualitative physical attributes were examined:

support – important to terms such as *on* (Bowerman and Pederson, 1992, 1996; Herskovits, 1986; Miller and Johnson-Laird, 1976)

control by Ground – important to terms such as *in* (Coventry et al., 1994; Coventry and Garrod, 2004)

animacy – important to terms such as *in* (Feist, 2000; Feist and Gentner, 2003)

Next, the range of application of each of the terms was examined as follows. For each term, all of the pictures described by the term were grouped together for further analysis. Each of the groups was then examined in order to isolate the attribute or attributes that were common to the pictures in the group, based on the picture codings just described.

Four of the coded attributes emerged as unifying factors in this analysis: *a difference in vertical position, contact, support,* and *inclusion*. The influences of these attributes, individually and in combination with one anther, are exemplified by the representative terms in Table 2. For each of the terms listed in Table 2, a plus under an attribute indicates that the attribute is present in all of the pictures described by the term; a minus indicates that the attribute is absent from all pictures described by the term.

Table 2. Representative terms

Term	Figure higher than Ground	Contact	Ground supports Figure	Inclusion
[sɔŋ] (Cantonese)	+			
taas (Tagalog)	+			
[nad] (Russian)	+	-		
[məʔal] (Hebrew)	+	-		
sotto (Italian)	-	+		
sous (French)	-	+		
na (Polish)		+	+	
på (Swedish)		+	+	
auf (German)		+		
an (German)		+		
u (Croatian)				+
[lɑ], [lopəla] (Telegu)				+
iqinde (Turkish)				+
[tʃə] (Cantonese)				-

As further evidence of the unifying nature of these attributes, they together served to categorize fifty-six of the sixty-three collected terms into the following seven classes of meaning.[5]

Figure higher than Ground
Figure higher than Ground, no contact
Figure lower than Ground, with contact
Ground supports Figure with contact
Contact
Inclusion of Figure by Ground
Absence of inclusion of Figure by Ground

These four attributes together provide evidence for the importance of geometry, function, and qualitative physics to the meanings of spatial terms across a variety of languages. The first two, *a difference in vertical position* and *contact*, provide information about the geometry of the Figure-Ground relationship. The third, *support*, provides qualitative physical information about the Figure-Ground interaction and the forces acting between the objects. Finally, *inclusion* provides information about geometry, function, and qualitative physics. In a three-dimensional world, geometric inclusion of one entity by another entails that the location of the included entity (the Figure) is constrained by the including entity (the Ground): in order to be geometrically included, the Figure must be located at the interior of the Ground. As such, the geometric attribute *inclusion* validates inferences about the presence of the qualitative physical attribute *location control*.[6] Similarly, as control of the location of an object is a common human goal, many artifacts have been created to fulfill this function, with the result that if the Ground is an artifact, *inclusion* of the Figure likely results from the fact that the Ground was created for this purpose (as was the case for the pictures in the current study).

The view through the window of semantic typology shows a landscape in which significant variation coexists with abstract similarities. Although spatial relations are grouped differently by the languages studied, attributes from all three families – geometric, functional, and qualitative physical – recurred in the meanings of the collected spatial terms. However, while typological studies such as the one presented here may suggest factors that are important to the meanings of spatial relational terms, controlled experimental investigation is necessary in order to test the roles of the factors in speakers' decisions to use specific terms. It is to this issue that we will now turn.

3 A view through the window of psycholinguistics

The view through the window of typology provided support for the theoretical import of geometry, function, and qualitative physics to the meanings of spatial relational terms. In language after language, it was found that geometric, functional, and qualitative physical properties united the disparate sets of scenes that could be described by a single term. Yet to be sure that these attributes truly influence speakers' choice of a word to describe a situation, we must seek corroborating evidence. Will a change in one of these factors lead to a concomitant change in a speaker's likelihood to employ a given term?

In order to closely examine the influence of any given attribute, it is desirable to hold as many other attributes constant as possible. This problem is nontrivial, as many of the attributes of spatial scenes that participate in spatial relational meaning co-occur in the real world. For example, support (a qualitative physical attribute) seldom occurs without contact (a geometric attribute) in everyday interactions (Feist, 2000). Similarly, as discussed above, many artifacts are created for the purpose of constraining the location of other objects, thus combining geometric, functional, and qualitative physical attributes in relations resulting from their normal use. In an attempt to tease apart a small set of attributes of scenes that influence the use of the English spatial prepositions *in* and *on*, Feist (2000; Feist and Gentner, 1998; 2003) adapted a method developed by Labov (1973) to study complex interacting factors in the use of English nouns. The details of the experimental study are reported in Feist (2000; see also Feist and Gentner, 2003). I present here an outline of the main experiment along with reasonably complete results.

In his classic study of naming patterns for cuplike artifacts, Labov (1973) systematically varied the functional context (e.g., *holding coffee* or *holding flowers*) and the relative height and width of a set of similarly shaped objects, which he then asked participants to name. He found that the variation in these factors led to changes in the nouns adults chose to name the objects. Similarly, I created a set of spatial scenes which were systematically varied with respect to geometric, functional, and qualitative physical factors in order to closely examine their influences on the use of the English prepositions *in* and *on*. The extent to which the differences in the pictures correlate with the changing rate of use of these English spatial prepositions is taken as indicative of the roles of these factors in the meanings of the prepositions.

In approaches to the meanings of *in* and *on* based on geometry, it is apparent that, while *in* requires that there be an interior of the Ground at which the Figure may be located, *on* requires merely a surface with which the Figure must be in contact (Bennett, 1975; Herskovits, 1986; Miller and Johnson-Laird, 1976). Consider a Figure in contact with the upper surface of a Ground. By manipulating the concavity of the Ground, without further change in the position of either object, it is possible to shift the relative applicability of the prepositions *in* and *on* (Figure 3). The influence of geometry was thus examined via changes in the curvature of the Ground. If geometry influences preposition choice, greater curvature (and concomitantly deeper concavity) of the Ground should correspond to a higher proportion of *in* responses.

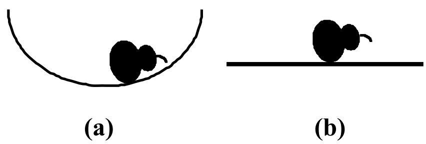


Figure 3. Two scenes differing only with respect to the concavity of the Ground

To vary the perceived function of the Ground, we took advantage of Labov's (1973) finding that the choice of a noun to label an object is influenced by the functional context within which the object is presented. Thus, we introduced the inanimate Ground with one of five labels, each communicating different information about the typical function of the Ground. The labels chosen were *dish, plate, bowl, rock,* and *slab.* If function influences preposition choice, we should see the greatest use of *in* when the inanimate Ground is labeled as a *bowl,* which is a prototypical container. Use of *in* should be lower for *plate,* which typically names objects that function as a supporting surface, and intermediate for *dish,* which is a superordinate term for *plate* and *bowl.* Finally, use of *in* should be low for *rock*, which is an afunctional solid, and for *slab*, which is an afunctional surface.

As information about qualitative physics is difficult to directly manipulate in static scenes while holding geometry constant, we indirectly manipulated qualitative physical properties by varying the animacy of the Figure and the Ground. An animate Figure, by virtue of its ability to enter and exit a configuration under its own power, may be conceived of as being less under the control of the Ground than would be an inanimate Figure. Conversely, an animate Ground is able to exert volitional control over the location of the Figure, while an inanimate Ground is not. If indirect effects of animacy on qualitative physical attributes related to location control influence preposition use, we might expect to see the greatest use of *in* for those situations that are physically most stable – situations where the Ground is animate and situations where the Figure is not. Similarly, we might expect to see the least use of *in* for those situations which are least stable – situations in which the Figure is animate and situations where the Ground is not. Thus, we should see greater use of *in* when the Ground is animate than when it is not. Likewise, we should see that the use of *in* is more prevalent when the Figure is inanimate than when it is animate.

In all, there were a total of twelve pictures. The set included two Figure objects – one animate and one inanimate. The Figures were each placed with respect to two Ground objects – one animate and one inanimate – and the Grounds were depicted at three levels of concavity, with the concavity of the two Grounds being equal at each level. The complete design is sketched in Figure 4.

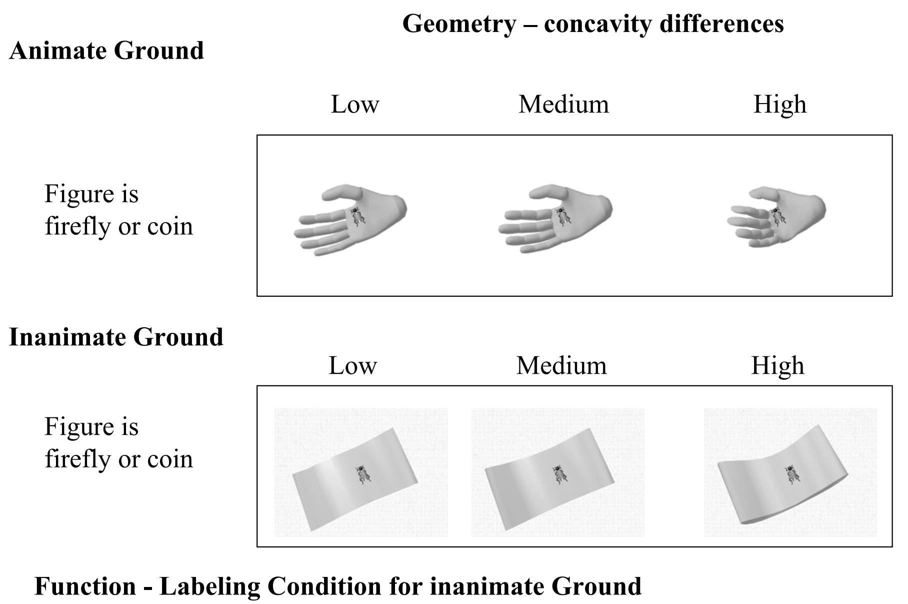


Figure 4. Design of the psycholinguistic study

The twelve pictures were presented individually on a computer screen in random order, and participants were given answer sheets with sentences of the following form:

> The *Figure* is IN/ON the *Ground*.

Figure was replaced with the noun referring to the pictured Figure (*firefly* or *coin*); likewise *Ground* was replaced with *hand* when the pictured Ground was the animate, and with the noun corresponding to the participant's labeling condition (*dish, plate, bowl, rock,* or *slab*) when the inanimate Ground was shown. The participant's task was to circle *in* or *on* to make each sentence describe the corresponding picture on the computer screen.

As predicted, participants' choices between *in* and *on* were found to be influenced by geometric, functional, and qualitative physical factors, as confirmed by a 2 (Ground: hand or inanimate) x 2 (Figure: firefly or coin) x 3 (concavity) x 5 (labeling condition) repeated measures analysis of variance. I will discuss each of these factors in turn.

That geometry plays a role in the meanings of *in* and *on* can be seen from the effect of changing the concavity of the Ground. As the concavity of the Ground increased, so did the use of *in*, with the average proportion of *in* responses for scenes depicting low concavity at .38, the average proportion for scenes depicting medium concavity at .45, and the average proportion for scenes depicting high concavity at .54, $F(2,172) = 28.34$, $p < .0001$ (Figure 5).

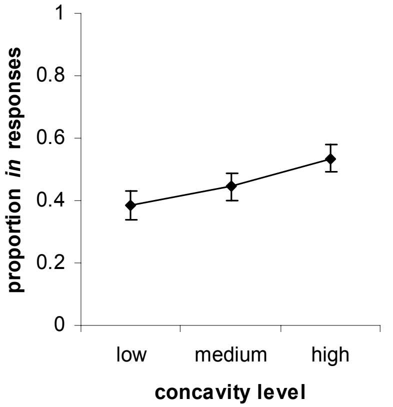


Figure 5. Effect of concavity, averaged across both Figures, both Grounds, and all five labeling conditions

That functional information plays a role in the meanings of *in* and *on* can be seen from the effect of varying the label provided for the inanimate Ground ($F(4,86) = 10.77$, $p < .0001$). As expected from the fact that the label was only changed for the inanimate Ground, there was also an interaction between the labeling condition and the animacy of the Ground ($F(4,86) = 5.43$, $p = .001$) (Figure 6). When the inanimate Ground was labeled as a *bowl*, a label normally applied to prototypical containers, the use of *in* was most prevalent (mean proportion *in* responses = .65). When the inanimate was labeled with *plate*, a noun normally used to label a functional surface, the proportion *in* responses was much lower (mean proportion *in* responses = .09). When the superordinate term *dish* was used, the proportion *in* responses was in between (mean proportion *in* responses = .50). Finally, the use of *in* was quite rare when the Ground was presented along with a label which suggested that it was not a functional artifact (mean proportion *in* responses for *rock* = .07; mean proportion *in* responses for *slab* = .08).

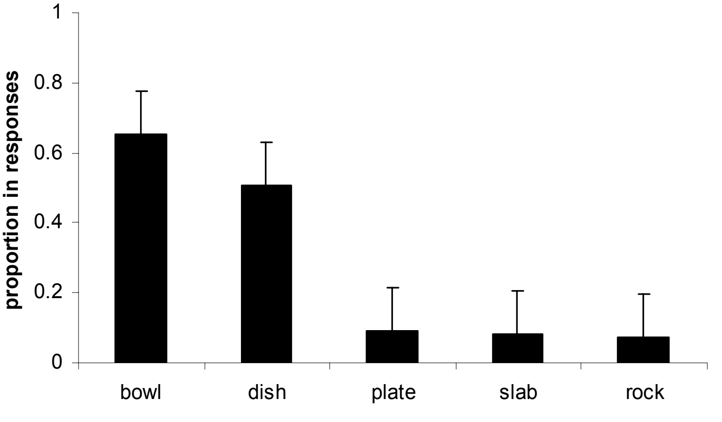


Figure 6. Effect of labeling condition for the inanimate Ground, averaged across all three concavities and both Figures

The influence of qualitative physics on the meanings of *in* and *on* can be inferred from the effects of the animacy of the Ground and the animacy of the Figure. When the depicted Ground was a hand, which is able to exert volitional control over another entity, the use of *in* was more prevalent than when the depicted Ground was inanimate (mean proportion *in* responses, hand as Ground = .63; mean proportion *in* responses, inanimate Ground = .28, $F(1,86) = 65.59$, $p < .0001$). Further, I found an interaction between the animacy of the Ground and its concavity whereby the increase in the proportion *in* responses as concavity increased was sharper for the hand than for the inanimate Ground ($F(2,172) = 5.50$, $p = .005$) (Figure 7). This difference makes sense in qualitative physical terms: because it can continue to close, a hand may be thought of as having more control over the location of its contents as it becomes more concave (more closed), while an inanimate object's degree of control, like its ability to continue closing, would remain constant across concavities.

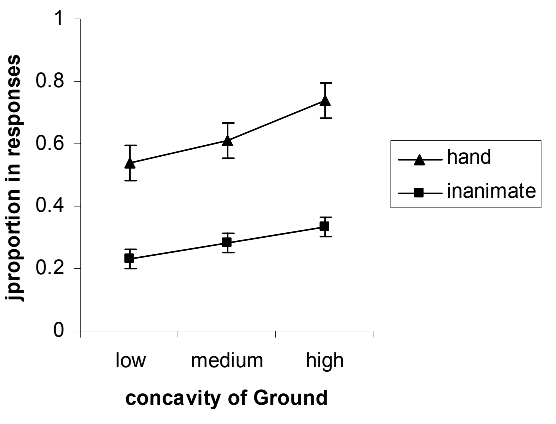


Figure 7. Interaction of animacy of the Ground and concavity, whereby the increase in *in* responses with increased concavity is sharper for the hand than for the inanimate Ground

In support of this explanation of the effect of the animacy of the Ground, and consistent with the predictions, when the depicted Figure was animate (a firefly), and thereby able to exert control over its own location, the use of *in* was *less* prevalent than when the depicted Figure was inanimate (mean proportion *in* responses, firefly as Figure = .43; mean proportion *in* responses, coin as Figure = .49, $F(1,86) = 9.69$, $p < .005$). Further, the influence of the animacy of the Figure interacted with the influence of functional information about the Ground: the extent to which *firefly* received a lower proportion *in* responses than did *coin* was greatest when the label for the inanimate Ground suggested a containment function (*bowl* and *dish*), $F(4,86) = 2.73$, $p < .05$ (Figure 8). The function of a container is, at its most basic, to fulfill the qualitative physical role of constraining the location of another object. This function can best be fulfilled if the object is more constrainable. As such, qualitative physics and function reinforce one another in scenes depicting an inanimate Figure and a Ground labeled as a container, hence raising the applicability of *in*.

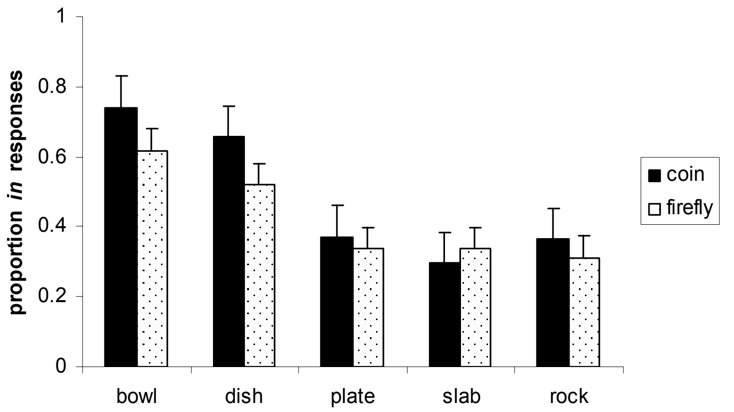


Figure 8. Interaction of labeling condition and animacy of the Figure, whereby the difference between responses to the coin and the firefly appear predominantly when the Ground is labeled as a functional container

Taken together, this set of results demonstrates that geometric, functional, and qualitative physical properties all influence speakers' uses of the English spatial prepositions *in* and *on*. Furthermore, although each exerts an independent influence on English prepositional usage, these three families of factors are not completely independent. Rather, they influence one another in complex ways, often providing reinforcing information that can raise the applicability of a preposition to a scene. Thus, the view through the window of psycholinguistics echoes the view through the window of typology, providing evidence that those factors which recur in the uses of spatial terms across languages also individually influence speakers' choices in a controlled communicative environment.

4 Conclusions

Multiple times each day, speakers choose from among a relatively small set of spatial relational terms (Landau and Jackendoff, 1993) to describe one of infinitely many possible spatial configurations between two objects in the environment. Their decisions are quick and sure, reflecting the automaticity of spatial relational terms. What attributes of spatial configurations must speakers attend to in order to fluently use the set of spatial relational terms available in their language?

While the semantics of spatial relational terms has received extensive attention, the picture of spatial relational meaning that emerges from an examination of theoretical treatments of spatial semantics is difficult to interpret. First, most characterizations of the meanings of spatial relational terms rely on a single type of feature. As a result, many common uses of spatial relational terms are left unexplained by the proposed meaning. Further, there is disagreement about whether geometric or functional features are criterial for spatial relational meaning. Second, the majority of the studies to date have involved single languages. Although these studies have catalogued the uses of the terms in the language under consideration, they are unable to provide a sense of spatial language more generally. Such a sense can only be gotten by considering the spatial vocabularies of many languages. It is precisely this sense of spatial language more generally that may provide the insights necessary to arrive at a descriptively adequate account of the meanings of individual spatial relational terms. Third, while theoretical treatments of spatial relational terms have proposed hypotheses about the factors that participate in the meanings of the terms, very few controlled experimental tests of the hypotheses have appeared.

In recent years, all three of these open issues have begun to be addressed, leading to a clearing picture of the factors participating in the semantics of spatial relational terms. With regard to the first issue, meanings incorporating more than one type of factor have been proposed (Coventry and Garrod, 2004; Feist, 2000; Herskovits, 1986), expanding the range of uses that can easily be accounted for within the proposed meaning. On the second count, researchers have begun to examine the spatial relational terms of multiple languages within a single project (Bowerman and Choi, 2001; Bowerman and Pederson, 1992, 1996; Feist, 2000, 2004, 2008; Levinson et al., 2003), concomitantly expanding the range of distinctions of which they are aware. Finally, with regard to the third open issue, researchers have begun to test the validity of the proposed factors in controlled psycholinguistic experiments (Carlson, Regier, Lopez and Corrigan, 2006; Coventry et al., 1994; Coventry and Prat-Sala, 2001; Coventry et al., 2001; Feist, 2000, 2002, 2005b; Feist and Gentner, 1998, 2003), allowing them to verify the role that each one plays in the meanings of individual spatial relational terms.

In this chapter, I have provided an overview of two studies designed to address the second and third of the identified gaps in our understanding of the semantics of space. In doing so, these studies provide valuable data which can be used to further efforts to address the first gap.

The first of the studies discussed compared the extensional ranges of sixty-three spatial relational terms collected from sixteen languages, representing data from nine language families. In order to be made maximally comparable, the terms were elicited by having all of the participants describe the same set of simple line drawings. The results showed that four attributes of spatial scenes, *a difference in vertical position, contact, support,* and *inclusion*, together provided unifying explanations for the individual extensional ranges of the fifty-six specific spatial terms collected (those encoding relatively detailed information about the Figure's location; see Feist (2004, 2008)). At a more abstract level, these four attributes impart information about geometric, functional, and qualitative physical aspects of the spatial scenes, providing evidence that these three families of factors influence the uses of spatial relational terms across a range of languages.

The second of the studies discussed in this chapter examined English speakers' uses of the prepositions *in* and *on* to describe a small set of scenes designed to vary along geometric, functional, and qualitative physical parameters. The results suggest roles for all three kinds of factors in the meanings of these two prepositions. The influence of geometry was demonstrated by the rise in *in* responses as concavity of the Ground increased (Figure 5). The influence of function was demonstrated by the observed effect of labeling condition: the use of *in* was most prevalent when the noun labeling the inanimate Ground typically names a container (*bowl*), with concomitantly low rates of use when the noun labeling the Ground typically names a functional surface (*plate*) or a nonfunctional entity (*rock* or *slab*) (Figure 6). Finally, the influence of qualitative physics was indirectly demonstrated via the effects of animacy of the Figure and Ground: use of *in* was most prevalent when the Ground was animate, enabling it to exert control over the location of the Figure, and when the Figure was inanimate, preventing it from exerting control over its own location.

Taken together, these two studies sketch two complementary views onto the landscape of human spatial cognition. The first view, that of the semantic typologist, considers both the unity and diversity of spatial language in order to arrive at a comprehensive picture of the set of factors involved in spatial relational meaning. The second view, that of the psycholinguist, considers the separable effects of a complex set of interacting factors on the uses of spatial relational terms. Both views suggest roles for three families of attributes of spatial scenes: geometric, functional, and qualitative physical. In combination, these three types of attributes can form the basis for a new representation of spatial relational meaning which, with one eye on typology and one on psycholinguistics, may better account for the uses of spatial relational terms than any one type of factor alone.

Notes

1 Following Talmy (1983), I will be referring to the located object, alternately called the trajector, or TR (Langacker, 1987), as *Figure*, and the reference object, alternately called the landmark, or LM, as *Ground*.

2 Data on genetic affiliations from Ethnologue, produced by the Summer Institute of Linguistics: http://www.ethnologue.com.

3 There is a difference of opinion among scholars as to whether or not Korean is related to Japanese. Further, Korean is possibly distantly related to Altaic.

4 Although *inclusion* is listed here as a geometric attribute, its presence bears on both functional and qualitative physical inferences, as will be discussed below.

5 The remaining seven terms fall into an eighth class, general spatial terms, which do not encode any specific attribute values. For details, see Feist (2000, 2004, 2008).

6 Note that this is not the case for the other geometric attributes. For example, although *contact* tends to co-occur with *support* across a variety of situations, the two attributes can easily be dissociated (e.g., in the case of two boxes side-by-side on the floor – they are in contact, but neither supports the other (Feist, 2000)).

References

Bennett, D. C. (1975) *Spatial and Temporal Uses of English Prepositions*. London: Longman.

Bowerman, M. and Choi, S. (2001) Shaping meanings for language: Universal and language specific in the acquisition of spatial semantic categories. In M.

Bowerman and S. C. Levinson (eds) *Language Acquisition and Conceptual Deveolpment*. Cambridge, UK: Cambridge University Press.

Bowerman, M. and Pederson, E. (1992) *Cross-linguistic perspectives on topological spatial relationships.* Paper presented at the 91st Annual Meeting of the American Anthrolopogical Association, San Francisco, CA.

Bowerman, M. and Pederson, E. (1996) *Cross-linguistic perspectives on topological spatial relationships*. Unpublished manuscript.

Brown, P. (1994) The INs and ONs of Tzeltal locative expressions: The semantics of static descriptions of location. *Linguistics* 32: 743–790.

Carlson, L. A., Regier, T., Lopez, B. and Corrigan, B. (2006) Attention unites form and function in spatial language. *Spatial Cognition and Computation* 6: 295–308.

Cienki, A. J. (1989) *Spatial Cognition and the Semantics of Prepositions in English, Polish, and Russian*. Munich, Germany: Verlag Otto Sagner.

Clark, H. H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive Development and the Acquisition of Language* 27–63. New York: Academic Press.

Coventry, K. R. (1998) Spatial prepositions, functional relations, and lexical specification. In P. Olivier and K.-P. Gapp (eds) *The Representation and Processing of Spatial Expressions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Coventry, K. R., Carmichael, R. and Garrod, S. C. (1994) Spatial prepositions, object-specific function, and task requirements. *Journal of Semantics* 11: 289–309.

Coventry, K. R. and Garrod, S. C. (2004) *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions*. London: Psychology Press.

Coventry, K. R. and Prat-Sala, M. (2001) Object-specific function, geometry, and the comprehension of in and on. *European Journal of Cognitive Psychology* 13: 509–528.

Coventry, K. R., Prat-Sala, M. and Richards, L. (2001) The interplay between geometry and function in the comprehension of *over, under, above,* and *below*. *Journal of Memory and Language* 44: 376–398.

Croft, W. (1999) Some contributions of typology to cognitive linguistics, and vice versa. In T. Janssen and G. Redeker (eds) *Foundations and Scope of Cognitive Linguistics* 61–94. Berlin: Mouton de Gruyter.

Croft, W. (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, W. (2003) *Typology and Universals.* (2nd ed.) Cambridge: Cambridge University Press.

Croft, W. and Poole, K. T. (2008) Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics* 34: 1–37.

Feist, M. I. (2000) *On in and on: An investigation into the linguistic encoding of spatial scenes.* Northwestern University, Evanston, IL.

Feist, M. I. (2002) Geometry, function, and the use of *in* and *on*. Paper presented at the Sixth Conference on Conceptual Structure, Discourse, and Language, Houston, TX.

Feist, M. I. (2004) Talking about space: A cross-linguistic perspective. In K. D. Forbus, D. Gentner and T. Regier (eds) *Proceedings of the Twenty-sixth Annual*

Meeting of the Cognitive Science Society 375–380. Mahwah, NJ: Lawrence Erlbaum Associates.

Feist, M. I. (2005a) In support of *in* and *on*. Paper presented at *New Directions in Cognitive Linguistics*. Brighton, UK.

Feist, M. I. (2005b) *Semantics and pragmatics of a general spatial term: The case of Indonesian di.* Paper presented at the 2005 Annual Meeting of the Linguistic Society of America, Oakland, CA.

Feist, M.I. (2008) Space between languages. *Cognitive Science* 32: 1177–1199.

Feist, M. I. and Gentner, D. (1998) *On plates, bowls, and dishes: Factors in the use of English IN and ON.* In M. A. Gernsbacker and S. J. Derry (eds) *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* 345–349. Mahwah, NJ: Lawrence Erlbaum Associates.

Feist, M. I. and Gentner, D. (2003) Factors involved in the use of in and on. In R. Alterman and D. Kirsh (eds) *Proceedings of the Twenty-fifth Annual Meeting of the Cognitive Science Society* 390–395. Mahwah, NJ: Lawrence Erlbaum Associates.

Forbus, K. D. (1983) Qualitative reasoning about space and motion. In D. Gentner and A. L. Stevens (eds) *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum.

Forbus, K. D. (1984) Qualitative process theory. *Journal of Artificial Intelligence* 24: 85–168.

Gentner, D. and Bowerman, M. (1996) *Crosslinguistic differences in the lexicalization of spatial relations and effects on acquisition.* Paper presented at the Seventh International Congress for the Study of Child Language, Istanbul, Turkey.

Gentner, D. and Bowerman, M. (2009) Why some spatial semantic categories are harder to learn than others: The Typological Prevalence hypothesis. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, S. Özçaliskan and K. Nakamura (eds) *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* 465–480. New York, NY: Lawrence Erlbaum Associates.

Haspelmath, M. (2003) The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (ed.) *The New Psychology of Language Volume* 2 211–242. Mahwah, NJ: Lawrence Erlbaum Associates.

Herskovits, A. (1986) *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Cambridge, UK: Cambridge University Press.

Labov, W. (1973) The boundaries of words and their meanings. In C-J. N. Bailey and R. W. Shuy (eds) *New ways of analyzing variation in English*. Washington, DC: Georgetown University Press.

Landau, B. (1996) Multiple geometric representations of objects in languages and language learners. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and Space* 317–363. Cambridge, MA: MIT Press.

Landau, B. and Jackendoff, R. (1993) 'What' and 'where' in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16: 217–265.

Landau, B. and Stecker, D. S. (1990) Objects and places: Geometric and syntactic representations in early lexical learning. *Cognitive Development* 5: 287–312.

Langacker, R. W. (1987) *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford: Stanford University Press.

Levinson, S. C. (1996) Relativity in spatial conception and description. In J. J. Gumperz and S. C. Levinson (eds) *Rethinking linguistic relativity* 177–202. Cambridge: Cambridge University Press.

Levinson, S. C., Meira, S. and The Language and Cognition Group. (2003) 'Natural concepts' in the spatial topological domain – adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79: 485–516.

Lindkvist, K.-G. (1950) *Studies on the Local Sense of the Prespositions in, at, on, and to in Modern English*. Lund, Sweden: C.W.K. Gleerup.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M. and Levinson, S. C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Science* 8: 108–114.

Miller, G. A. and Johnson-Laird, P. N. (1976) *Language and Perception*. Cambridge, MA: Belknap Press.

O'Keefe, J. (1996) The spatial preposition in English, vector grammar, and the cognitive map theory. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and Space* 277–316. Cambridge, MA: MIT Press.

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

Sinha, C. and Thorseng, L. A. (1995) A coding system for spatial relational reference. *Cognitive Linguistics* 6: 261–309.

Sinha, C., Thorseng, L. A., Hayashi, M. and Plunkett, K. (1994) Comparative spatial semantics and language acquisition: Evidence from Danish, English, and Japanese. *Journal of Semantics* 11: 253–287.

Talmy, L. (1983) How language structures space. In H. Pick and L. Acredolo (eds) *Spatial Orientation: Theory, Research, and Application*. New York: Plenum Press.

Talmy, L. (1988) Force dynamics in language and cognition. *Cognitive Science* 12: 49–100.

Tyler, A. and Evans, V. (2003) *The Semantics of English Prepositions*. Cambridge: Cambridge University Press.

Vandeloise, C. (1991) *Spatial Prepositions: A Case Study from French* (A. R. K. Bosch, Trans.) Chicago: University of Chicago Press.

Vandeloise, C. (1994) Methodology and analyses of the preposition in. *Cognitive Linguistics* 5: 157–184.

Vandeloise, C. (2003) Containment, support, and linguistic relativity. In H. Cuyckens, R. Dirven and J. Taylor (eds) *Cognitive Approaches to Lexical Linguistics* 393–425. Berlin and New York: Mouton de Gruyter.

Vandeloise, C. (this volume) Genesis of spatial terms. In V. Evans and P. Chilton (eds) *Language, Cognition and Space: The State of the Art and New Directions*. London: Equinox.

5 Parsing space around objects

Laura Carlson

1.0 Introduction

Imagine you are planning to get together with a friend after a class, and that you arrange to meet in front of the lecture hall. After class, you momentarily position yourself at the front door, but because there are a lot of students milling in and out of the building, you walkway halfway down the steps and sit down. While you wait, you become very warm in the sun, and move down past the steps to sit on a bench off to the side under a tree to await your friend. Although you are now quite a distance from the building, you feel confident that your friend will find you. The research project described in the current chapter investigates how all of these locations (at the door of the building, halfway down the steps, and on the bench to the side of the building) are understood to be '*in front of*' the lecture hall. The chapter begins with an overview of the cognitive processes and representations that assist in defining projective spatial terms such as *front*. This is followed by a brief summary of the previous attempts at studying the regions denoted by such terms. A new methodology is described that addresses some limitations with these previous approaches. The utility of the new methodology is established by demonstrating that various factors known to affect the interpretation of spatial terms also impact the size and shape of the region denoted by *front*. These factors include: the identity of the objects, the functional characteristics of the objects, the presence of additional objects in the scene, and the reference frame used to define the spatial term.

2.0 Projective spatial terms

One way to describe the location of a target object is by spatially relating it to an object whose location is known, as in 'My friend is at the front of the lecture hall'. In this utterance, 'my friend' is designated the located object (also known variously as figure, locatum or trajector); finding this object is the presumed goal of the utterance. The 'lecture hall' is designated the reference object (or variously, relatum, ground, or landmark), an object whose position is presumed known or easily found (see Miller and Johnson-Laird, 1976; Levelt, 1996; Talmy, 1983; Tyler and Evans, 2003). Understanding of this statement involves not only linking the objects in the world to the referents in the sentence, but also mapping the spatial relational term (i.e., *front*) to the appropriate region of space around the reference object. Terms such as *front* belong to the class of projective spatial terms that convey direction information; these contrast with terms such as 'near' that belong to the class of proximal terms that convey distance information (for an organizational chart, see Coventry and Garrod, 2004).

Logan and Sadler (1996) present a computational framework that specifies the processes and representations involved in mapping spatial terms such as *front* onto regions of space around the reference object. For the current chapter, two representations and their constituent processes are of interest: reference frames and spatial templates. For projective terms, a reference frame consists of a set of three axes that assign directions onto space. Specifically, the reference frame is imposed on the reference object, and its axes extend outward defining the space beyond. One set of axes corresponds to the vertical dimension, and its endpoints delineate space *above* and *below* the reference object. Two sets of axes correspond to the horizontal dimensions, with one set delineating space '*in front of*' and '*in back of*' the reference object, and the other set delineating space '*to the right of*' and '*to the left of*' the reference object. (Garnham, 1989; Levelt, 1984; Levinson, 1996; Miller and Johnson-Laird, 1976). Figure 1, Panel A shows a reference frame imposed on a reference object (a chair) with its axes defining the space around the reference object.

Reference frames are flexible representations that have a set of parameters that define their use in a given context (Logan and Sadler, 1996; for a summary of evidence, see Carlson, 2004). These parameters include the origin, orientation, direction and distance. The *origin* parameter defines where on the reference object the reference frame is imposed. For example, in Figure 1, Panel A, the reference frame is imposed at the center of the chair. More generally, the origin may be defined on the basis of the reference object's geometry and/or on the basis of its functional properties (Carlson-Radvansky, Covey and Lattanzi, 1999). The *orientation* parameter determines how to orient the vertical and horizontal axes. For example, in Figure 1, Panel A, the vertical axis is aligned with the top/bottom of the chair and with the top/bottom of the picture plane. The *direction* parameter determines the endpoints of these axes (e.g., the above and below endpoints of the vertical axis). In Figure 1, Panel A, the endpoint of the vertical axis closest to the top of the chair is *above*; the endpoint closest to the bottom of the chair is *below*. The orientation and direction parameters can be defined on the basis of various sources of information, such as the environment, the object, or interlocutor (speaker or addressee) (Logan and Sadler, 1996), with the source typically taken to define the type of reference frame in use (i.e., absolute, intrinsic or relative, respectively; see Levinson, 1996). For example, in Figure 1, Panel B, the chair is rotated 90 degrees to the right. It is now possible to define two different axes as vertical – one corresponding to the top/bottom of the picture plane (absolute) and one corresponding to the top/bottom of the chair (intrinsic). Finally, the *distance* parameter indicates the spatial extent of the region, and is defined at least in part by properties of the objects and the spatial term relating them. For example, Carlson and Covey (2005) asked participants to imagine sentences such as 'The squirrel is in front of the flowers'. The main task was to provide an estimate for how far apart the objects were in their image. Distance estimates varied systematically as a function of the spatial term relating the objects, both for proximal terms such as 'near' and 'far' (as might be expected by the semantics of these terms) and for projective terms such as *front* and *back*, with *front* estimates consistently smaller than *back* estimates.

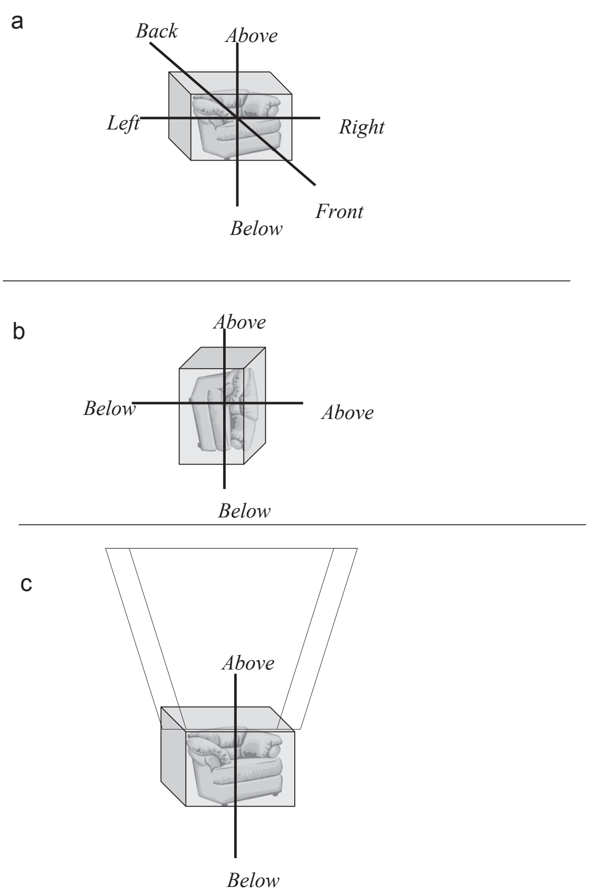


Figure 1. Panel A. A reference frame imposed on a reference object (chair), with its axes defining the space around the reference object. Panel B. With the reference object rotated, *above* and *below* can either be assigned to the vertical axis consistent with the top/bottom of the picture plane or the horizontal axis consistent with the top/bottom of the chair. Panel C. An illustration of a possible spatial template representing the region *above* with respect to the chair. Note that the size and shape of this region are speculative.

A second representation that has been subsequently considered an additional parameter of a reference frame (Carlson-Radvansky and Logan, 1997) is a *spatial template*. Spatial templates can be thought of as an illustration of the spatial region around a reference object for a given spatial term. For example, Figure 1, Panel C shows a possible spatial template for *above* for the chair in Figure 1, Panel A, that extends upward and outward from the topside of the chair. Note, however, that the size and shape of this region are speculative. One of the critical goals of the current chapter is to explore the factors that help determine the shape and spatial extent of such regions.

Spatial templates reflect two important assumptions of the mapping of the term onto space around the reference object: first, that the use of a spatial term does not correspond to a single location in space but rather encompasses a region. Theoretically, this idea is consistent with Miller and Johnson-Laird (1976) conceptualization of the location of an object as containing an area of space immediately surrounding it, referred to as its penumbra or region of interaction (Morrow and Clark, 1988). Two objects are said to be in a spatial relation with each other when these areas overlap (see also Langacker, 1993, 2002). Second, a spatial term does not apply equally across this region (Logan and Sadler, 1996; Hayward and Tarr, 1995); rather, some locations are preferred over others. Theoretically, this idea is consistent with Hayward and Tarr's (1995) treatment of the meaning of spatial terms as category representations with prototypes and graded membership (see also the idea of a preferred subspace, Herskovits, 1986 and a proto-scene, Tyler and Evans, 2003).

3.0 Previous empirical approaches to understanding spatial regions

There have been theoretical treatments of spatial regions (e.g., Herskovits, 1986; Miller and Johnson-Laird, 1976; Vandeloise, 1991). For example, Herskovits (1986) suggests that there are ideal meanings of spatial prepositions based on geometric descriptions that are then applied in a given context via particular use types that may serve to modify these underlying descriptions (for similar ideas, see Coventry and Garrod, 2004; Tyler and Evans, 2003). As applied to the particular regions associated with projective terms, she discusses how context, the presence of additional objects, and characteristics of the objects may serve to identify preferred sub-spaces. However, there has been relatively little systematic empirical work supporting these ideas that explicitly addresses how the size and shape of the regions are defined, and whether they can be modified. In this section, I'll briefly describe two empirical approaches adopted in the previous literature that examine spatial regions, and discuss some of their limitations as applied to the questions of interest. The new methodology presented in Section 4.0 will integrate components from each of these approaches.

3.1 Spatial templates

One approach to examining spatial regions has been to ask participants to rate the acceptability of a spatial term as describing placements of located objects at various locations around the reference object (Hayward and Tarr, 1995; Logan and Sadler, 1996). For example, Logan and Sadler (1996) presented participants with a located object (the letter 'O') that was placed across trials within various locations in an invisible 7 X 7 grid whose center contained a reference object (the letter 'X'). The task was to rate how well a given spatial term described the relation between the two objects. Figure 2, Panel A shows a sample trial. The endpoints of the scale were labeled as 1 = bad and 9 = good; intermediate values were permitted. The ratings were then plotted as a function of the placement of the located object within the grid, as shown in Figure 2, Panel B. The divot in the plot at cell 4,4 corresponds to the location of the reference object. Logan and Sadler (1996) referred to this plot as a spatial template.

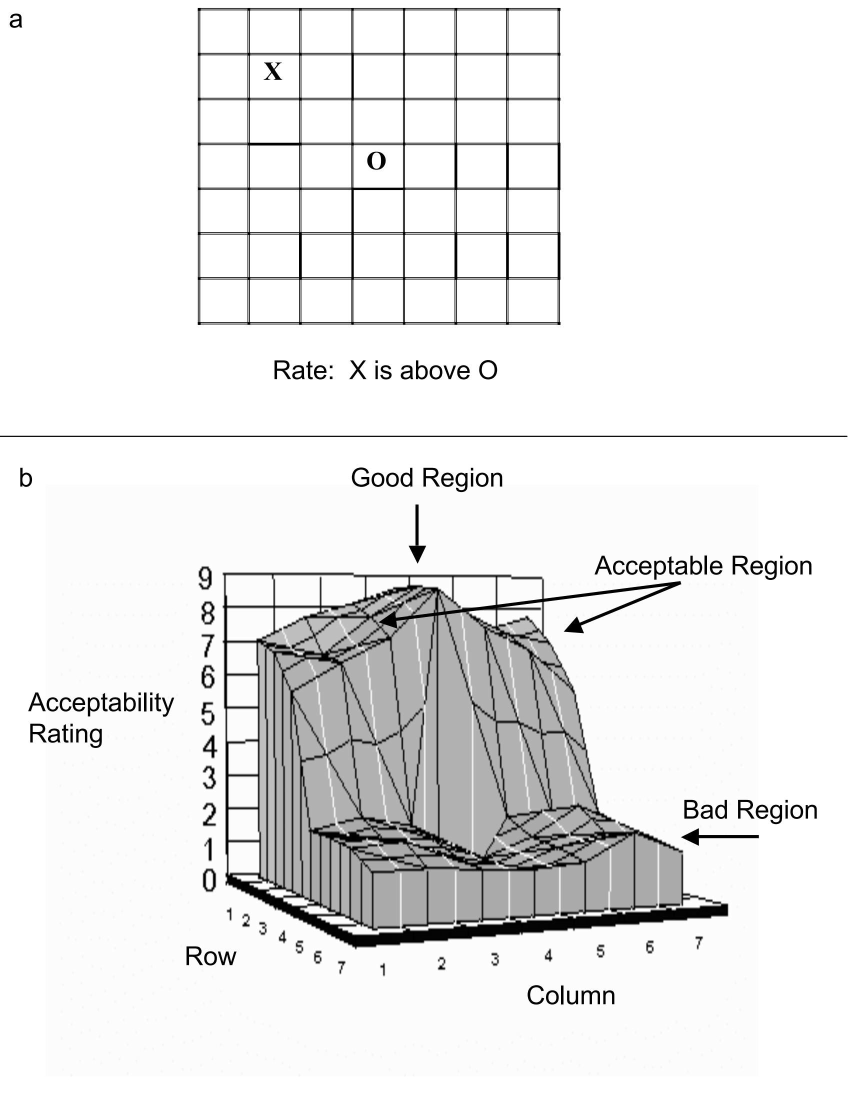


Figure 2. Panel A. A sample trial from Logan and Sadler (1996) in which participants are asked to rate how well the sentence 'X is above O' describes the relative locations of the X and O in the display. The 7X7 grid defines the possible placements of the X across trials; the grid was not visible to participants. Panel B. The spatial template for *above* from data from Logan and Sadler (1996). See text for descriptions of the good, acceptable and bad regions.

Within the spatial template three distinct regions were identified. The peak of the template comprised the 'good' region, and corresponded to placements along the axis. The regions flanking the peak showed a drop in acceptability, and comprised the 'acceptable' regions. Finally, the flat area at the opposite endpoint with uniformly low acceptability ratings comprised the 'bad' region. Logan and Sadler (1996) derived spatial templates for many different terms, and showed that the overall shapes and sizes for projectives such as *above* and *left* were remarkably similar, differing only in orientation. Hayward and Tarr (1995), Carlson-Radvansky and Logan (1997), and Regier and Carlson (2001) obtained spatial templates using a variety of reference objects, and showed global similarity in their size and shape, suggesting a common underlying mechanism that defined the space independently of the objects. For example, Regier and Carlson (2001) presented the attention vector sum model (AVS) in which the shape of the spatial template was defined by the joint components of attention and a vector sum representation of direction. Carlson, Regier, Lopez and Corrigan (2006) further extended AVS to incorporate the functional characteristics of the objects, thus making it sensitive to the objects being related.

Although this approach offers a means of visualizing the region associated with a given spatial term, it has several possible limitations. First, it is not clear that a rating of acceptability will translate into actual use in a more naturalistic task. While one might infer that ratings at the endpoints may indicate whether a speaker would use the term (presumably *yes* for placements in the good region, and *no* for placements in the bad region), it is not clear whether intermediate ratings (as in the acceptable region) would necessarily translate into selection, nor whether such selection would be constant within the acceptable region. Second, the spatial extent of the template is largely defined by the experimenter before the study during construction of the grid that contains the placements of the located object. This is potentially problematic, as the grid may not directly encompass the boundaries of a given region. For example, in the Logan and Sadler (1996) plot for *above* (Figure 2, Panel B), there is no drop-off in acceptability as a function of distance within the good region, suggesting that all placements within this region are acceptable. However, Regier and Carlson (2001) found that ratings did vary within the good region when a reference object with spatial extent was probed at multiple locations using different distances. Moreover, Carlson and Van Deman (2004) showed faster response times to placements of the located object in the good region that were closer to the reference object than those that were farther away from the reference object, indicating a potential effect of distance. Thus, it is not clear that the edge of the spatial template as constructed by the experimenter will necessarily reflect the boundary of the region. A final limitation is that spatial templates have been collected using a 2D projection of space; however, most of our everyday use of these projective terms involves mapping them onto 3D space. It is not clear whether such 2D projections will necessarily generalize to the 3D case.

3.2 Regions around oneself

A second approach to examining the spatial regions defined by projective spatial terms has been to ask participants to explicitly divide the space around their bodies. Franklin, Henkel and Zangas (1995) had participants stand in the center of a circular room constructed by hanging a curtain from the ceiling, thereby eliminating any visible sides. Their task was to indicate the boundaries of the regions corresponding to *front*, *back*, *left* and *right* around themselves. To do this, they held a pointer, and were asked to move the pointer as far to one side as possible while staying within a given region. For example, to identify the left edge of the front region, participants were told to move the pointer to the left as far as possible but so that the pointer still indicated *front*. The pointer extended outside the curtain where a protractor was printed on the floor; this enabled the experimenter to record the angle corresponding to the pointer, and to identify the size of the various regions. For example, for the left edge of the front region, if a participant was facing 60 degrees and placed the pointer at 110 degrees, this would indicate that the front region extended 50 degrees to the left from their midline. Across trials, each participant indicated each boundary of each region. This enabled Franklin et al. (1995) to determine the sizes of each region; a schematic is shown in Figure 3 with the viewer standing in the center, facing the top o the page. Franklin et al. (1995) found that the space around oneself was not divided into equally spaced 90 degree regions. Rather, the front region was the largest (124 degrees), followed by the back region (110 degrees) and then left and right regions (91 and 92, respectively). The regions for *front* and *back* did not differ significantly from each other, but both were significantly larger than the regions for *left* and *right*, which also did not differ from each other.

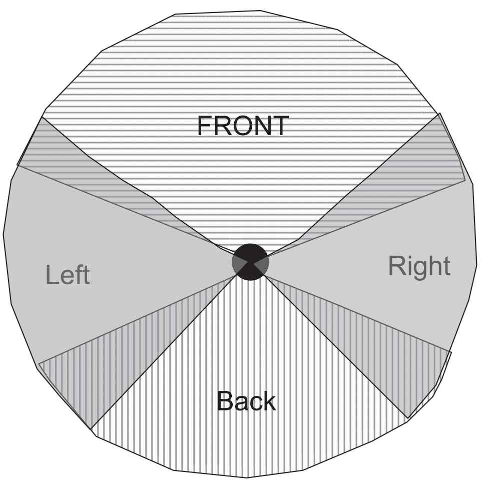


Figure 3. A schematic of the *front*, *back*, *left* and *right* regions obtained by Franklin, Henkel and Zangas (1995).

This approach offers a direct means of assessing the boundaries of the various regions. Interestingly, the overlap between the front and left/right regions indicates that there areas in which multiple spatial terms may be possible, consistent with an interpretation of these spatial terms as having fuzzy boundaries within a categorical representation (Hayward and Tarr, 1995; for other work on overlapping regions, see Carlson-Radvansky and Logan, 1997; Herskovits, 1986). Another benefit of this approach is that it was conducted in 3D space. However, there are also several limitations. First, it does not enable one to assess the spatial extent of the regions. Participants did not indicate a distance associated with the regions, but only the borders of the regions. Indeed, a constant distance was imposed by the walls of the room, and there was an implicit assumption that the regions would extend at a minimum to the walls at all angles (see Figure 3). Yet, the spatial template data from Logan and Sadler (1996) suggest that this may not be correct. Specifically, the plot in Figure 2, Panel B, suggests that ratings drop as a function of angle and distance, as one moves from the good to the acceptable region, and one moves within the acceptable region. Moreover, the identity of the objects may impact the extent of these regions as well. Miller and Johnson-Laird (1976) suggest that objects evoke distance norms that represent typical values derived from interactions with other objects. Morrow and Clark (1988) refer to these areas as zones of interaction. Carlson and Covey (2005) showed that the distances inferred between two objects that were spatially related depended not only on the spatial term used to describe the relation (e.g., *front* versus *back* as discussed in Section 2.0), but also on the size and shapes of the objects. For example, the distance estimates inferred for descriptions relating large objects (i.e., 'The St. Bernard is in front of the tree.') were consistently larger than the distance estimates inferred for descriptions relating small objects (i.e., 'The squirrel is in front of the flower.'). This finding suggests that the size of the regions observed by Franklin et al (1995) may be specific to the space surrounding the particular object that they investigated (the participants themselves), and may not generalize to other objects.

4.0 A new methodology for determining size and shape of the spatial region

4.1 Combining components from previous approaches

In addition to the limitations specific to each of the two approaches described in section 3.0, more generally, there has been no systematic attempt to examine how spatial regions may vary as a function of factors asserted to be relevant (Herskovits, 1986; Langacker, 1987; Miller and Johnson-Laird, 1976; Tyler and Evans, 2003; Vandeloise, 1991), and known to impact other aspects of spatial term interpretation. These factors include the identity of the objects (Carlson-Radvansky, et al, 1999), its functional characteristics (Carlson et al., 2006; Carlson-Radvansky et al., 1999; Coventry and Garrod, 2004), the presence of additional objects (Carlson and Logan, 2001; Carlson and Hill, submitted) and the type of reference frame used to define the spatial term (Carlson-Radvansky and Irwin, 1993; Carlson-Radvansky and Radvansky, 1996). This section presents a new methodology to assess the role that these factors may play in defining these spatial regions. The validity of the new methodology can be evaluated by determining whether it is sensitive to manipulations of these established factors.

The new methodology builds on features of the two previous approaches described in Section 3. Specifically, from the spatial template approach, we adopt the idea of probing at multiple locations, drawn from the best, acceptable and bad regions. However, rather than collect acceptability ratings, we asked participants to directly indicate placement that corresponds to the best (prototypic) use of a given spatial term at various angles from the reference object. In this manner we obtain a direct measure of this distance at multiple locations. From the Franklin et al. (1995) paradigm we adopt the idea of directly assessing the boundaries, asking participants to indicate the farthest points at which a spatial term still applies. Combining the best and farthest measures enables us to obtain a fairly direct means of representing the spatial extent of the regions. In addition, we also asked participants to indicate whether alternative terms also apply at given locations, as a way of getting at the fuzziness of the boundaries and the degree of overlap with other spatial terms. We developed the methodology in a number of studies by focusing on the size and shape of *front,* given that this spatial term is considered privileged relative to the other horizontal relations (i.e., *back*, *left* and *right* (Clark, 1973; Fillmore, 1971; Franklin, Henkel and Zengas, 1995; Garnham, 1989).

4.2 The specific methodology and data

To measure the spatial region corresponding to *front*, we placed a small dollhouse cabinet that served as a reference object in the center of a 102 cm X 82 cm uniform white foam board. The white board was placed on a large conference table, and the participant was seated at one end of the table facing the cabinet. The set up is shown in Figure 4. Eleven lines were identified that radiated out from the reference object, numbered from left to right, counterclockwise, as shown in Figure 5. These 11 lines can be defined in terms of angular deviation (0 – 90 degrees, in either direction; i.e., unsigned) from the front of the cabinet, and categorized with respect to the regions (good, acceptable, bad) within which they fall on the *front* spatial template of Logan and Sadler (1996). Specifically, Lines 5–7 were directly in front of the cabinet, at 0 degrees, located in the good region, with lines 5 and 7 at the edges of the region, and line 6 at its center. Lines 2 and 8 were each at 22.5 degrees, lines 3 and 9 were each at 45 degrees and lines 4 and 10 were each at 67.5 degrees from the front of the cabinet; these lines all fell into the acceptable region. Finally, lines 1 and 11 were each at 90 degrees from the front of the cabinet, located in the bad region. The lines were not visible to participants; marks on the edges of the white foam board indicated the endpoints of the lines for the experimenter.

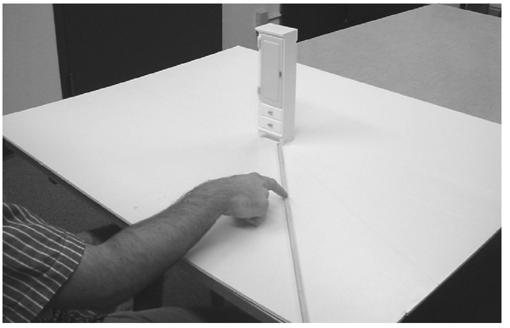


Figure 4. Experimental set-up in which the participant is seated in front of a white board containing the reference object, and indicates distance judgements associated with the spatial term by pointing to a location on the dowel. The dowel is marked in centimeters on the side not visible to participants.

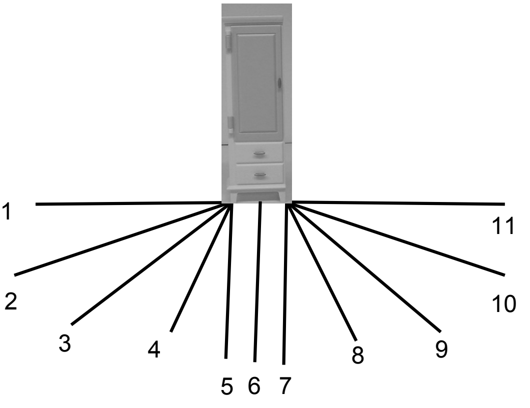


Figure 5. The cabinet is the reference object; measures were collected along each of lines 1 – 11 in a random order. Lines 1 and 11 are within the bad region; Lines 2–4 and 8–10 are within the acceptable region; Lines 5–7 are within the good region.

On each trial, the experimenter placed a square dowel rod in position along one of the 11 lines. The lines were not visible to the participant; the experimenter lined up the rod with one end at the cabinet and the other end at a mark on the edge of the white foam board (not visible to the participant) that indicated the appropriate angle. One side of the dowel was marked in centimeters, starting with 0 at the cabinet, and ending at 67 cm, just past the edge of the white board. For each placement of the dowel, participants were asked to make three judgments pertaining to their definition of *front* with respect to the cabinet. First, participants indicated by pointing to a location on the dowel the distance that corresponded to the best use of *front*. The experimenter read this value on the dowel, and recorded it. This measure was intended to define the ideal or prototypical distance (Hayward and Tarr, 1995). Second, participants indicated along the dowel the farthest distance for which *front* would still be deemed acceptable. The experimenter read this value on the dowel, and recorded it. This measure was intended to define the outer extent of the front region. Third, at this farthest distance, participants were asked to indicate whether they would prefer to use an alternative spatial term to describe an object at this location rather than *front*. If there was a preference for an alternative term, the participants reported the term, and were instructed to move back along the dowel toward the cabinet and indicate the point at which *front* became preferred to this alternative. This measure was intended to take into account the fact that some locations could be defined with respect to competing spatial terms. As such, the farthest distance at which *front* may be used may not be the same as the farthest distance at which *front* is preferred relative to these other terms. That is, just because a participant may indicate that *front* may be acceptable, this does not necessarily mean that a participant would use the term *front* to describe an object at that location. Each participant provided these three measures for each of the 11 lines, with the sequence of lines randomly determined for each participant.

The best and farthest data can be summarized by averaging the values per line across participants and then plotting these means on Lines 1–11. Connecting the means reveals the spatial regions defined by the best and farthest distances, as shown in Figure 6. With respect to the competing term measure, the data are interesting but complicated, given that not all participants supplied competing terms, or indicated a new preferred *front* distance. For the sake of the current chapter, we will focus on the best and farthest measures only. In the studies that we describe in the next section, we were interested in how these regions changed as a function of various manipulations that are known to impact the interpretation of projective spatial relations.

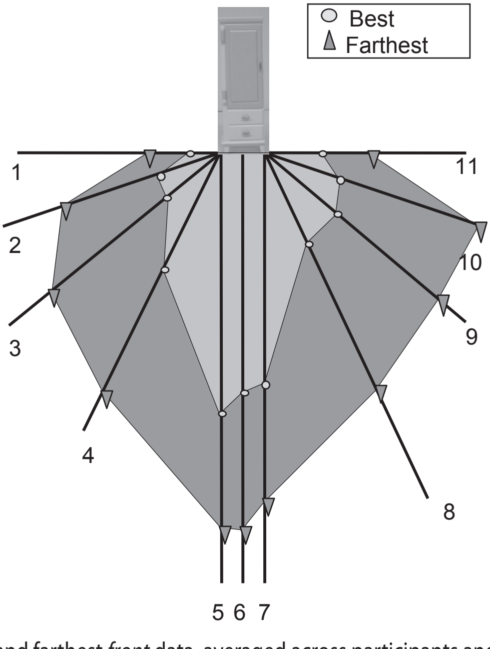


Figure 6. Sample best and farthest *front* data, averaged across participants and connected to form the corresponding regions.

5.0 Factors that impact the size and shape of the spatial region

Using the basic methodology described in Section 4, across experiments we examined the impact of a diverse set of factors on the size and shape of the *front* region, including the identity of the objects, the functional characteristics of the objects, the presence of additional objects in the display, and the type of reference frame that defined the spatial term.

5.1 Identity of the reference object

Past theoretical and empirical work (Tyler and Evans, 2003; Herskovits, 1986; Langacker, 1987; Miller and Johnson-Laird, 1976; Vandeloise, 1991) has shown that the identity of the reference object has a significant impact on the way in which a spatial term is applied to space around it. Such influence has been observed with respect to where the reference frame is imposed on the reference object (Carlson-Radvansky et al, 1999), with respect to the distance inferred between the objects (Carlson and Covey, 2005; Morrow and Clark, 1988); and with respect to the locations that are deemed acceptable around the object (for extensive review of object effects, see Coventry and Garrod, 2004). At the outset of this project we were interested in understanding how the same physical space (the white board) may be interpreted differently as a function of the reference object being used. Previous work has shown that the absolute size of the reference object makes a difference in the distances associated with spatial terms, with shorter distances associated with smaller objects (Carlson and Covey, 2005; Morrow and Clark, 1988). We were interested in whether conceptual size would have a similar impact. To assess this, we contrasted a scaled-down model object (dollhouse cabinet) with an actual size object (lotion bottle), closely equating the physical sizes of the objects.[1] These objects are shown in Figure 7. If participants scale the whiteboard space with respect to the

Figure 7. Dollhouse cabinet and lotion bottle matched in actual size and shape but mismatched in conceptual size.

conceptual size of the object, then we should observe differences in the best and farther measures for the *front* region; however, if participants define the *front* region with respect to physical size, there should be no differences in these measures.

Figure 8 shows the best regions for the cabinet (Panel A) and the bottle (Panel B), with the two regions superimposed and flipped vertically to enable an assessment of how well the regions overlap (Panel C). Figure 9 shows the data in an alternate form, plotting the best distances for each object as a function of line in panel A, and the farthest distances for each object as a function of line in panel B. We excluded lines 1, 2, 10, 11 from the plots because most responses for these lines corresponded to a 0 distance, reflecting the fact that *front* would not be used for positions on these line. These plots render differences among the contrasting conditions easiest to see; accordingly, the data in the remaining sections will be presented and discussed in this manner. There was no difference in the best measure as a function of object (cabinet or lotion bottle). However, in the farthest measure, a significant difference occurred within the good region (lines 5,6,7), with *front* extending farther for the lotion bottle than the cabinet. This suggests that participants may have been scaling the size of the regions to the conceptual size of the objects, with the actual size object having a larger zone of interaction than the model sized object.

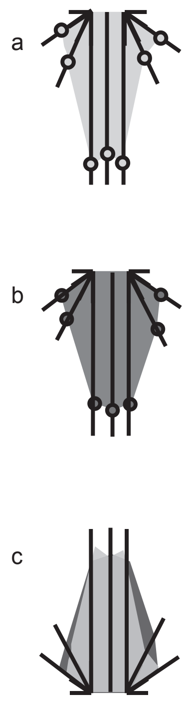


Figure 8. Panel A. Plot of best *front* for the cabinet. Panel B. Plot of best *front* for lotion bottle. Panel C. Superimposed plots for cabinet and lotion bottle; differences in size and shape are of interest.

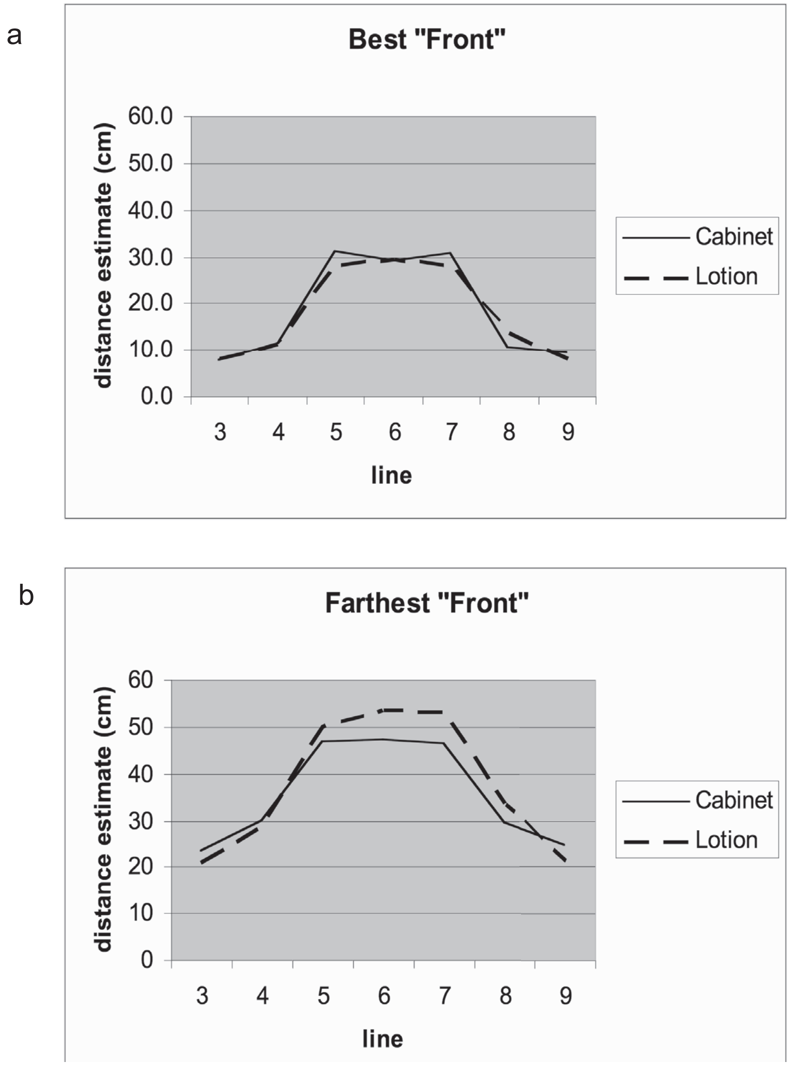


Figure 9. Panel A. Plot of best *front* as a function of line for cabinet and lotion bottle. Panel B. Plot for farthest *front* as a function of line for cabinet and lotion bottle.

5.2 Functional parts of the reference object

Previous research has suggested that not only the identity of the reference object but also the manner in which the reference and located objects interact may impact the way in which spatial terms are mapped onto space. For example, Carlson-Radvansky et al. (1999) demonstrated that the best placement for a located object was influenced by the location of a prominent functional part of the object. Carlson and Kenny (2006) further showed that this influence depended not only on the association between the objects, but also upon their ability to interact. To see whether functional effects would be observed in this new methodology, we contrasted best and farthest measures for *front* for two versions of the dollhouse cabinet that differed only with respect to the side at which the door opened. The two cabinets are shown in Figure 10, Panel A. If the way in which one might interact with the object impacts the way in which the *front* region is defined, then one would expect the best measure to be influenced by the side at which the door opens. Specifically, the good region should not extend as far on the opening side, as one needs to be closer to that side in order to interact with (e.g., reach into) the cabinet. In contrast, if one investigated the back region with respect to the cabinet, one would not expect such a difference due to side, because one typically doesn't interact with the back of the cabinet. Figure 10, Panel B shows the backs of the cabinets – note that the door handles can be seen; in addition the cabinets were presented with the doors slightly ajar. Thus, information about the doors was available for defining the back regions. However, the prediction was that this would not influence the size or shape of the regions. Finally, the predicted effect on the *front* region was expected to be limited to the best measure; the farthest measure reflects the putative boundary of front, and would be presumably beyond the area of interaction with the object.

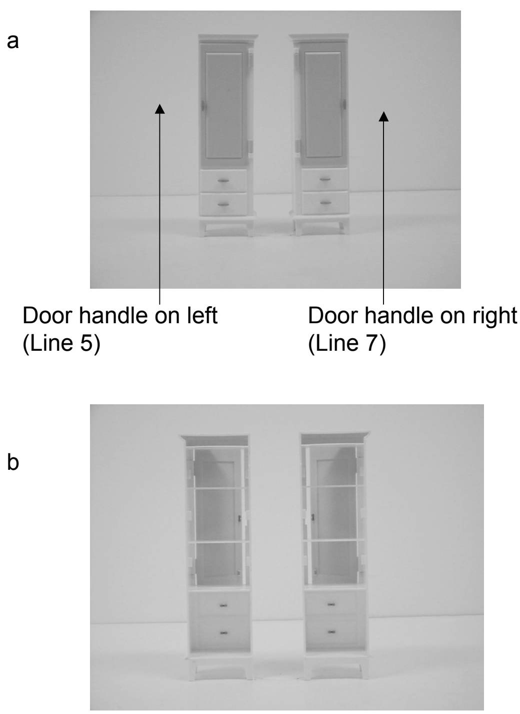
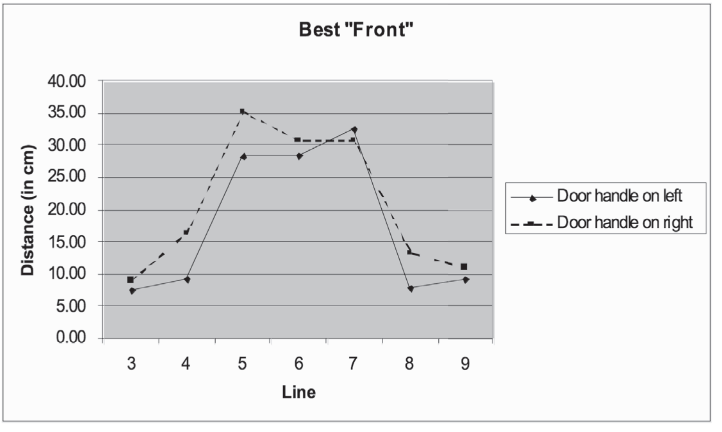


Figure 10. Panel A. Two dollhouse cabinets. The one on left has door handle on the left (from the reader's perspective), aligned with line 5. The one on the right has the door handle on the right (from the reader's perspective), aligned with line 7. Panel B. The corresponding backs of the dollhouse cabinets; backs were removed and doors were slightly ajar so that participants could see how the doors would open.

Figure 11, Panels A and B show the best and farthest data for *front*. The best data clearly show an interaction with the door of the cabinet. When the door opens on the right (from the reader's perspective) (dotted line), the location of the best *front* along the line close to the handle (line 7) was closer to the cabinet than for line close to the opposite edge of the cabinet (line 5); however, when the door opened on the left (from the reader's perspective) (solid line), the location of the best *front* was closer to the cabinet by the door (line 5) than on the opposite side (line 7). This asymmetry in the best *front* region is consistent with the way in which one might interact with the object. Moreover, no such asymmetry was observed for the farthest *front* measure (Panel B), nor for either the best or farther *back* measures, plotted in Panels A and B of Figure 12, respectively. These data replicate the influence of functional parts on defining spatial regions around a reference object within the new methodology.

a



b

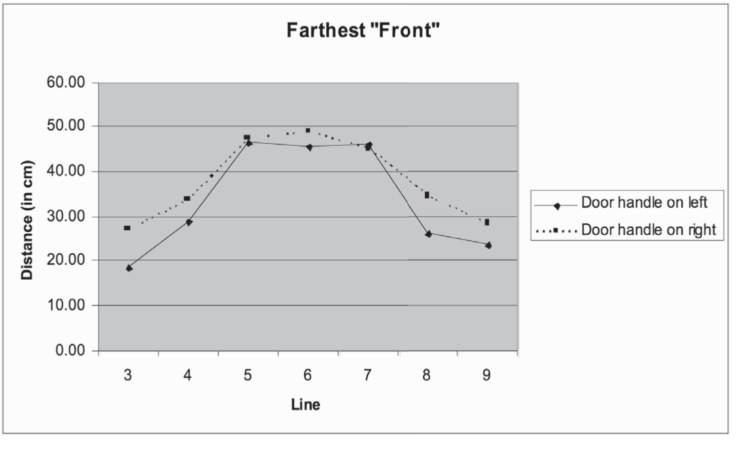


Figure 11. Plot of best *front* as a function of line for the two dollhouse cabinets. Panel B. Plot for farthest *front* as a function of line for the two dollhouse cabinets.

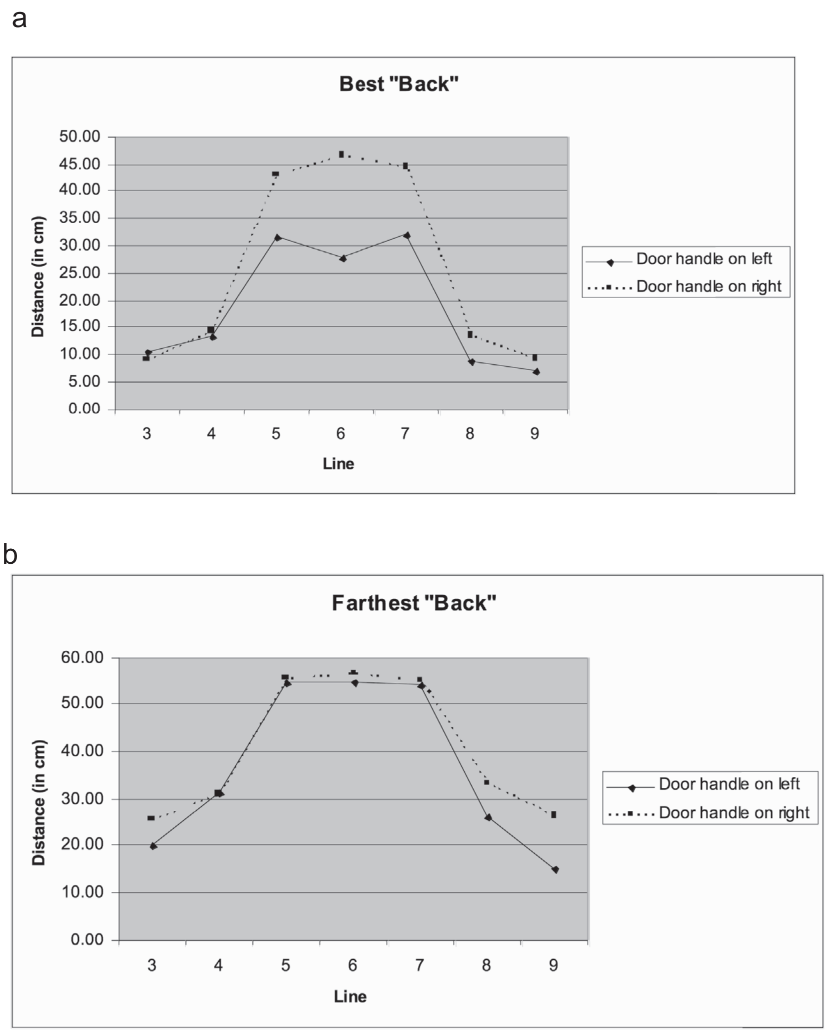


Figure 12. Plot of best *back* as a function of line for the two dollhouse cabinets. Panel B. Plot for farthest *back* as a function of line for the two dollhouse cabinets.

5.3 The addition of a located object

In the studies described in sections 5.1 and 5.2, a reference object was placed in the middle of the display board and participants indicated the best and farthest distances associated with *front* along the dowel using their finger. However, most often spatial descriptions include two objects, the located object and the reference object. In this study, we asked participants to indicate the best and farthest *front* by placing a located object at the desired distance along the dowel. Previous research has shown that the identity of the located object and the manner in which it interacts with the reference object has a significant impact on the way in which spatial terms are applied to space around the reference object (Carlson and Kenny, 2006; Carlson-Radvansky et al, 1999). Therefore, we contrasted two located objects: a doll and a dog that were both scaled relative to the dollhouse cabinet that was used as a reference object (i.e., these were sold together as a playset). The objects are shown next to the cabinet in Figure 13, Panel A. Figure 13, Panel B shows a sample participant placing these objects during the task. Given Carlson and Covey's (2005) results that the distance associated with a spatial term depended upon the size of the objects, we expected the best and farthest measures for *front* to be smaller for the dog (a smaller object) than for the doll.

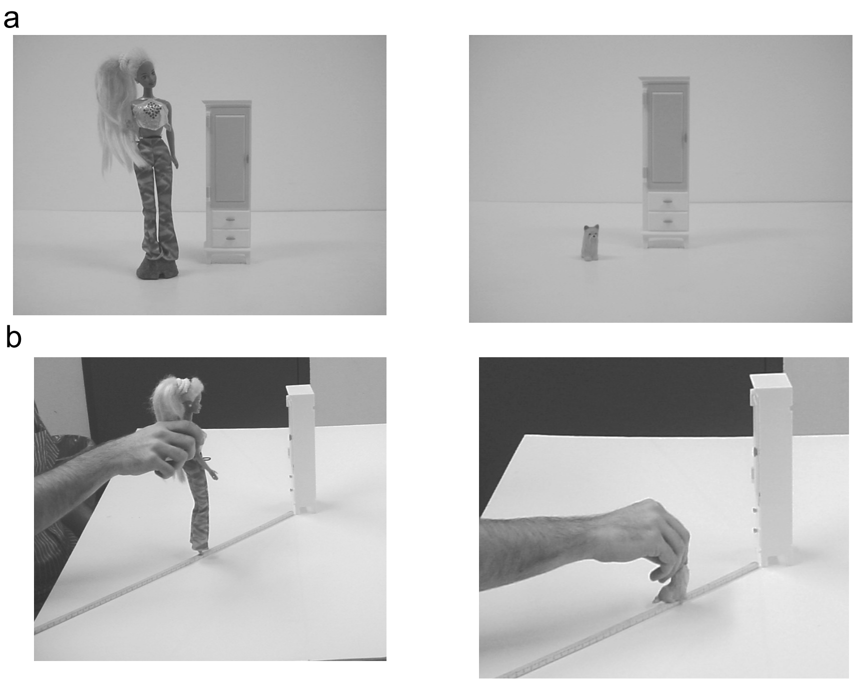
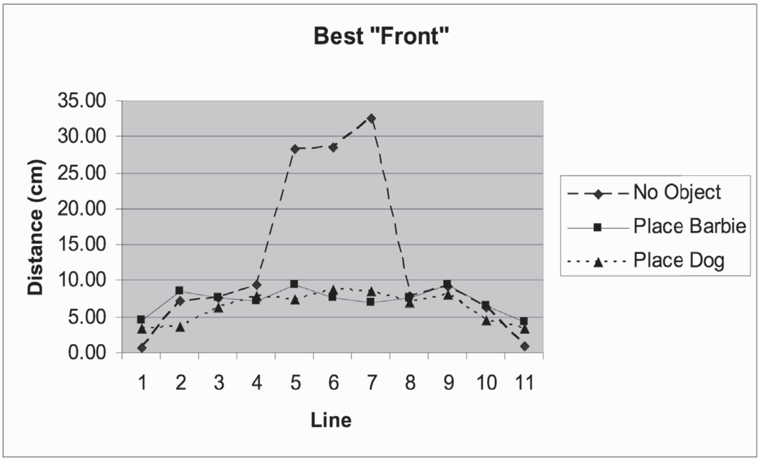


Figure 13. Panel A. Doll as located object next to cabinet as reference object (on left) and dog as located object next to cabinet as reference object (on right). Panel B. Placement of doll (on left) and dog (on right) during the experimental task.

Figure 14, Panels A and B show the data for the best and farthest measures, respectively. For the best measure, two effects are readily apparent. First, adding a located object compresses the best *front*, relative to the condition in which no object was placed and participants indicated the distance with their fingers. Second, averaging across lines, there was a small but reliable difference, such that the best *front* for the dog was closer to the cabinet than the best *front* for the doll. For the farthest measure, there were also differences due to placing an object. First, relative to not placing an object, the distances associated with placements of the dog were much smaller, but of the same general shape. This stands in contrast to the data for the lines in the good region with the best measure (contrast lines 5–7, Panels A and B) in which distances for placing the dog were much reduced. Second, the differences between placing the doll and not placing an object depended upon line. In the good region (lines 5–7), there was not much difference; in contrast, in the acceptable and bad regions, the distances dropped off steeply when there was no object to place but remained relatively large when placing the doll. Third, the distances for the dog were uniformly shorter across lines than the distances for the doll. In summary, across both measures there were systematic effects of adding a locating object, with its characteristics (identity or size) impacting the way in which the term *front* was applied to space around the reference object.

a



b

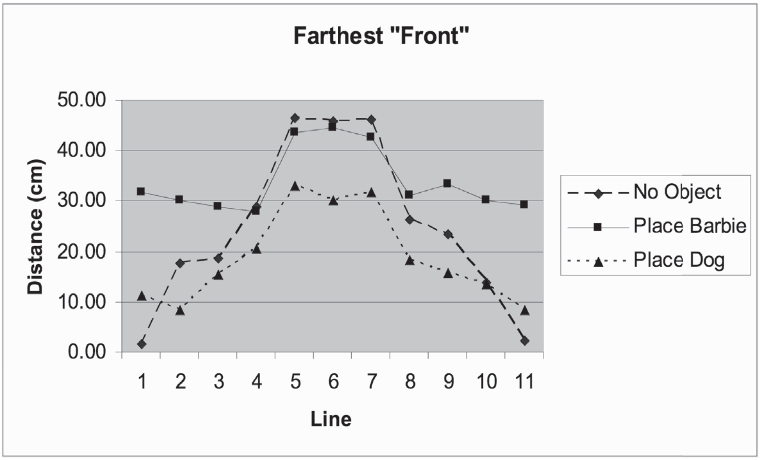


Figure 14. Panel A. Plot of best front as a function of line for conditions where no object was placed (finger indicated location), doll was placed and dog was placed. Panel B. Plot of farthest front as a function of line for conditions where no object was placed (finger indicated location), doll was placed and dog was placed.

5.4 Reference frame used to define front

The source of information used to define the axes of a reference frame, and thereby set the orientation and direction parameters, is typically used to define the type of reference frame. For example, within Levinson's (1996) typology, an absolute reference frame uses features of the environment that are invariant to the viewer or reference objects to assign directions; the intrinsic reference frame uses the predefined sides of an object to assign directions; and the relative reference frame uses the perspective of a viewer or other object in the scene to assign directions to space around the reference object (see Levinson, 1996, for other notable differences among the types of reference frames). We were interested in whether the front region defined with respect to an intrinsic reference frame based on the cabinet would be of a different size or shape than the front region when defined with respect to a relative reference frame based on the participant. This is interesting because when the participant faces the cabinet to perform the task (see Figure 4), their front regions overlap in physical space. Thus, any observed differences would be due to the way in which the particular reference frames were imposed on the space, rather than to the space itself.

Figure 15 shows the data for best and farthest *front* measures for space around a cabinet in Panels A and B, respectively. There is a small but consistent effect of a larger 'best' region when the space was defined with respect to the intrinsic frame based on the cabinet than with respect to the relative frame based on the participant's front. A similar trend was observed when comparing the intrinsic and relative reference frames with the lotion bottle. These effects may be due to the fact that greater attention is paid to the object when *front* is defined by the intrinsic frame than relative frame, thereby emphasizing its zone of interaction. No such effect was observed in the farthest measure.

6.0 Conclusion

In this chapter I have presented a new methodology for examining how regions associated with the projective term *front* are defined. Several factors that have been previously shown to impact the interpretation of such spatial terms within alternative approaches were examined, and initial findings suggest that effects of these factors can be observed within this methodology as well. The idea that these regions may change shape and size as a function of characteristics of the objects being related is consistent with a current dominant theme in spatial language research that incorporates influences of the objects, the context, and goals into one's interpretation of a spatial description (Coventry and Garrod, 2004; Tyler and Evans, 2003; more generally, see Zwaan, 2004).

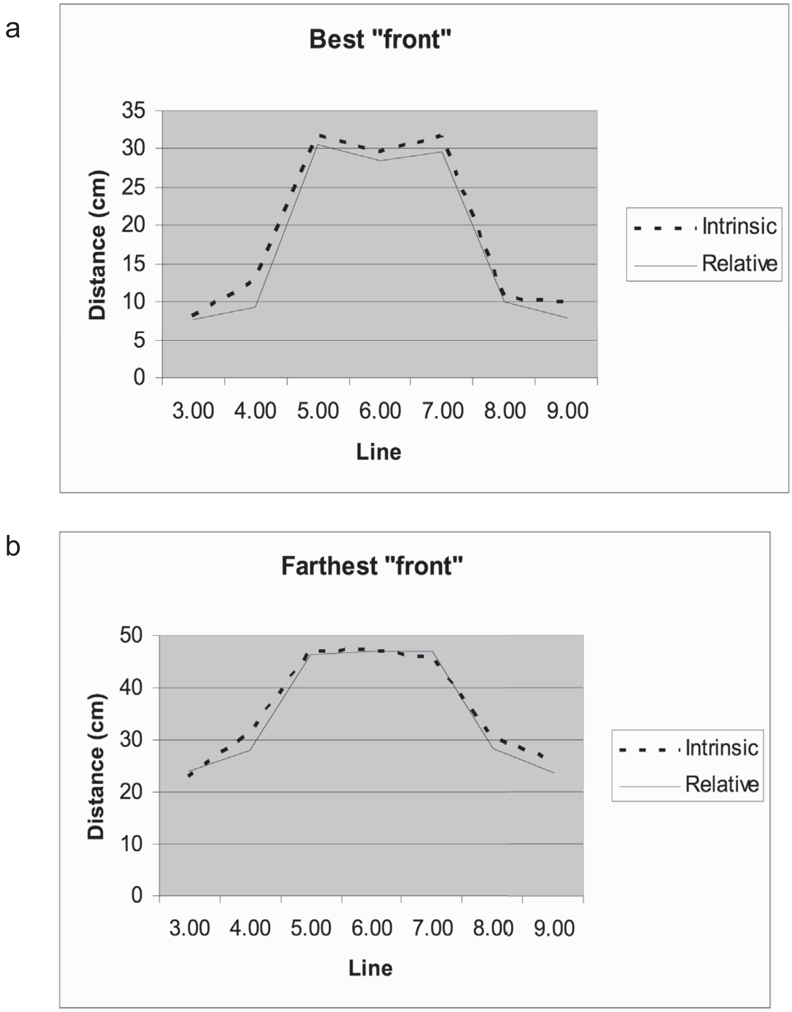


Figure 15. Plot of best *front* as a function of line when using the intrinsic and relative reference frames to define the spatial term. Panel B. Plot for farthest *front* as a function of line when using the intrinsic and relative reference frames to define the spatial term. The cabinet was the reference object.

Acknowledgements

I thank Meghan Murray, Christina Shreiner, and Padraig Carolan for assistance in data collection, and Aaron Ashley and Patrick Hill for discussion. Address all correspondence to Laura Carlson, Department of Psychology, 118-D Haggar Hall, University of Notre Dame, Notre Dame, IN 46556; lcarlson@nd.edu. Portions of this research were presented at the 2004 meeting of the Association for Psychological Science, Chicago, IL.

Note

1. An additional constraint also directed our selection of these objects: specifically, that the objects had functional parts that could be moved from one side to the other (the door of the cabinet; the nozzle on the lotion bottle; see Section 5.2).

References

Carlson, L. A. (2003) Using spatial language. In B. H. Ross (ed.) *The psychology of learning and motivation* 127–161. (Vol. 43) New York: Academic Press.

Carlson, L. A. and Covey, E. S. (2005) How far is near? Inferring distance from spatial descriptions. *Language and Cognitive Processes* 20: 617–632.

Carlson, L. A. and Hill, P. L. (2008) Processing the presence, placement and properties of a distractor during spatial language tasks. *Memory & Cognition* 36: 240–255.

Carlson, L. A. and Kenny, R. (2006) Interpreting spatial terms involves simulating interactions. *Psychonomic Bulletin & Review* 13: 682–688.

Carlson, L. A. and Logan, G. D. (2001) Using spatial relations to select an object. *Memory & Cognition* 29: 883–892.

Carlson, L. A., Regier, T, Lopez, W. and Corrigan, B. (2006) Attention unites form and function in spatial language. *Spatial Cognition and Computation* 6: 295–308.

Carlson, L. A. and van Deman, S. R. (2004) The space in spatial language. *Journal of Memory and Language* 51: 418–436.

Carlson-Radvansky, L. A., Covey, E. S. and Lattanzi, K. L. (1999) 'What' effects on 'where': Functional influences on spatial relations. *Psychological Science* 10: 516–521.

Carlson-Radvansky, L. A. and Irwin, D. E. (1993) Frames of reference in vision and language: Where is above? *Cognition* 46: 223–244.

Carlson-Radvansky, L. A. and Logan, G. D. (1997) The influence of reference frame selection on spatial template construction. *Journal of Memory and Language* 37: 411–437.

Carlson-Radvansky, L. A. and Radvansky, G. A. (1996) The influence of functional relations on spatial term selection. *Psychological Science* 7: 56–60.

Clark, H. H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive development and the acquisition of language*. New York: Academic Press.

Coventry, K. R. and Garrod, S. C. (2004) *Saying, seeing and acting: The psychological semantics of spatial prepositions*. New York: Psychology Press.

Fillmore, C. J. (1971) *Santa Cruz lectures on deixis*. Bloomington, IN: Indiana University Linguistics Club.

Franklin, N., Henkel, L. A. and Zengas, T. (1995) Parsing surrounding space into regions. *Memory & Cognition* 23: 397–407.

Garnham, A. (1989) A unified theory of the meaning of some spatial relational terms. *Cognition* 31: 45–60.

Herskovits, A. (1986) *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge: Cambridge University Press.

Hayward, W. G. and Tarr, M. J. (1995) Spatial language and spatial representation. *Cognition* 55: 39–84.

Langacker, R. W. (1987) *Foundations of cognitive grammar*. Stanford: Stanford University Press.

Langacker, R. W. (1993) Grammatical traces of some 'invisible' semantic constructs. *Language Sciences* 15: 323–355.

Langacker, R. W. (2002) A study in unified diversity: English and mixtec locatives. In N. J. Enfield (ed.) *Ethnosyntax: Explorations in grammar and culture*. Oxford: Oxford University Press.

Levelt, W. J. M. (1984) Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind and J. J. Koenderink (eds) *Limits in perception* 323–358. Utrecht: VNU Science Press.

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

Levinson, S. (1996) Frames of reference and Molyneux's questions: Cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel and M. Garret (eds) *Language and space* 109–169. Cambridge, MA: MIT Press.

Logan, G. D. and Sadler, D. D. (1996) A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel and M. Garret (eds) *Language and space* 493–529. Cambridge, MA: MIT Press.

Miller, G. A. and Johnson-Laird, P. N. (1976) *Language and perception*. Cambridge, MA: MIT Press.

Morrow, D. G. and Clark, H. H. (1988) Interpreting words in spatial descriptions. *Language and Cognitive Processes* 3: 275–291.

Regier, T. and Carlson, L. A. (2001) Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General* 130: 273–298.

Talmy, L. (1983) How language structures space. In H. L. Pick and L. P. Acredolo (eds) *Spatial orientation: Theory, research and application* 225–282. New York: Plenum Press.

Tyler, A. and Evans, V. (2003) *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.

Vandeloise, C. (1991) *Spatial prepositions: A case study from French*. Chicago: Chicago University Press.

Zwaan, R.A. (2004) The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (ed.) *The psychology of learning and motivation* 35–62. (Vol. 44) New York: Academic Press.

Final proofs with index 24 February 2010

6 A neuroscientific perspective on the linguistic encoding of categorical spatial relations[1]

David Kemmerer

Slus, a Mayan speaker of the language Tzeltal, says to her husband, facing an unfamiliar contraption, 'Is the hot water in the uphill tap?' It is night, and we have just arrived at an alien hotel in a distant, unfamiliar city out of the hills. What does she mean? She means, it turns out, 'Is the hot water in the tap that would lie in the uphill (southerly) direction if I were at home?' Levinson (2003, p. 4)

1 Introduction

The semantic domain of space arguably consists of three subdomains – shape, motion, and location. Most readers of the current volume are probably aware that all three of these subdomains have been intensively investigated by cognitive linguists during the past few decades. However, many readers may not realize that in recent years cognitive neuroscientists have begun to use the tools of their trade – especially the lesion method and hemodynamic methods – to illuminate the brain structures that underlie each subdomain. The greatest progress has been made in understanding the neural correlates of the subdomain of shape, but a substantial amount has also been learned about the anatomical bases of the subdomains of motion and location (for reviews see Kemmerer, 2006, in press, forthcoming). This chapter focuses on the subdomain of location and attempts to integrate new findings from linguistics and neuroscience.

At the very outset, it is important to note that much of the neuroscientific work on the meanings of locative morphemes has been partly motivated by an interest in Kosslyn's (1987) hypothesis that the human brain contains separate systems for computing two types of spatial relations – coordinate and categorical (for reviews see Jager and Postma, 2003; Laeng et al., 2003; Postma and Laeng, 2006). Representations of coordinate spatial relations involve precise metric specifications of distance, orientation, and size; they are useful for the efficient visuomotor control of object-directed actions such as grasping a cup; and they may be processed predominantly in the right hemisphere. In contrast, representations of categorical spatial relations involve groupings of locations that are treated as equivalence classes; they serve a variety of perceptual functions, such as registering the rough positions of objects in both egocentric and allocentric frames of reference; and they may be processed predominantly in the left hemisphere. It has often been observed that categorical spatial relations are usually referred to linguistically by words like English prepositions, many of which specify binary oppositions – e.g., *on/off, in/out, left/right, above/below.* For instance, Laeng et al. (2003, p. 308) state that 'all natural languages seem to have a special class in their grammar (i.e., prepositions) devoted to the expression of categorical spatial relations'. As I demonstrate below, however, prepositions are not the only relevant grammatical category, and the range of categorical spatial relations that are linguistically encoded goes well beyond the meanings of English prepositions.

The chapter is organized as follows. In section 2 I summarize recent research on the kinds of categorical spatial relations that are encoded in the 6000+ languages of the world and that are also, *ipso facto,* implemented in the brains of the speakers. Emphasis is placed on crosslinguistic similarities and differences involving deictic relations, topological relations, and projective relations, the last of which are organized around three distinct frames of reference – intrinsic, relative, and absolute. During the past few decades, a voluminous literature on the meanings of locative morphemes has emerged, including several new approaches such as the Functional Geometry framework (e.g., Coventry and Garrod, 2004; Carlson and Van Der Zee, 2005) and the Principled Polysemy model (e.g., Tyler and Evans, 2003). However, I will draw mostly on recent typological research, especially studies conducted by the Language and Cognition Group at the Max Planck Institute for Psycholinguistics (e.g., Levinson, 2003; Levinson and Wilkins, 2006). Next, section 3 reviews what is currently known about the neuroanatomical correlates of linguistically encoded categorical spatial relations, with special focus on the left supramarginal and angular gyri. In addition, suggestions are offered for how crosslinguistic data can help guide future research in this area of inquiry. Finally, section 4 explores the interface between language and other mental systems, specifically by summarizing studies which suggest that although linguistic and perceptual/cognitive representations of space are at least partially distinct, language nevertheless has the power to bring about not only modifications of perceptual sensitivities but also adjustments of cognitive styles.

2 What types of categorical spatial relations are linguistically encoded?

Very few languages have a word for 'space' in the abstract sense employed by philosophers and scientists such as Newton, Leibniz, Kant, and Einstein. However, current evidence suggests that all languages have *Where*-questions (Ulltan, 1978) that tend to elicit answers in which the figure object (F) – i.e., the thing to be located – is described as being within a search domain defined by some kind of categorical spatial relation to a ground object (G) – i.e., a thing that serves as a point of reference (Talmy, 1983). Several classes of categorical spatial relations are encoded to different degrees in different languages, and although they interact in complex ways, each one usually constitutes a fairly independent semantic field that is 'carved up' by a specialized set of lexical items and grammatical constructions (Levinson and Wilkins, 2006).

2.1 Deictic relations

Deixis involves the many ways in which the interpretation of utterances depends on aspects of the speech event (Fillmore, 1997). In the present context, the most relevant deictic expressions are demonstratives – e.g., *here* vs. *there, this* vs. *that* (Burenhult, 2008; Diessel, 1999, 2005, 2006; Dixon, 2003; Dunn et al., forthcoming). These words specify the location of F directly in relation to the location of the speech participants, instead of in relation to some G outside the speech situation. The proper functional characterization of demonstratives requires close attention to details of social interaction (Enfield, 2003; Hanks, 2005). However, I will not discuss these complex social parameters here, since the main focus is on how demonstratives are often used to divide the radial egocentric space surrounding the speaker (or addressee) into categorically discrete zones. Crucially, demonstratives do not encode metrically precise degrees of remoteness from the deictic center, but rather have abstract meanings that are pragmatically modulated by either the discourse context or the referential scenario, thereby allowing speakers to flexibly expand or contract the zones so as to express an unlimited range of distance contrasts – e.g., *here in this room* vs. *here in this galaxy*.

In a sample of 234 languages from diverse families and geographical regions, Diessel (2005) found that the kind of demonstrative system manifested in English, with a binary proximal/distal contrast, is actually the most frequent, showing up in 127 (54%) of the languages. However, this is the minimal type of system, and other languages exhibit systems of greater complexity. For example, some languages include the addressee as a possible deictic center. Such person-oriented systems come in several varieties. One type, exemplified by Pangasinan (Western Austronesian, Philippines),[2] has a three-way contrast between 'near speaker', 'near addressee', and 'far from both speaker and addressee', while another type, exemplified by Quileute (Chimakuan, Washington State), has a four-way contrast between 'near speaker', 'near addressee', 'near both speaker and addressee', and 'far from both speaker and addressee'. These person-oriented systems resemble the English two-term system insofar as they specify just two zones – proximal and distal. The key difference is that person-oriented systems require the speaker to perform more elaborate spatial calculations which take into account not only his or her own egocentric frame of reference, but also that of the addressee. Perhaps for this reason, person-oriented systems are relatively rare. A more common way to increase the complexity of a demonstrative system is to partition the dimension of distance into more fine-grained zones. Eighty-eight (38%) of the languages in Diessel's sample follow this strategy by distinguishing between three zones – proximal, medial, and distal. Spanish and Yimas (Sepik-Ramu, Papua New Guinea) have systems like this. A very small proportion of languages (less than 4% in Diessel's sample) go one step further by distinguishing between four zones – proximal, medial, distal, and very distal. Tlingit (Na Dane, Yukon) is the most often cited example. There are even reports of languages with demonstrative systems that encode five distance contrasts (Anderson and Keenan, 1985), but Diessel supports Fillmore (1997), who maintains that systems with more than four terms invariably combine other semantic parameters.

These other semantic parameters include visibility, elevation, and geography. A striking example of how local geographic features can be incorporated into the semantics of demonstrative systems comes from the Himalayan language Limbu (Kiranti, Nepal), which has the following terms: *ma:dha:mbi* means 'on the slope of the mountain ridge across the valley from where the speaker is situated', *kona:dha:mbi* means 'on the same slope of the mountain ridge as the speaker', and *khatna:dha:mbi* means either 'on the back side of the mountain ridge on which the speaker is situated' or 'on the far side of the mountain ridge across the valley from which the speaker is situated' (van Driem, 2001). Even more remarkable is Cora (Uto-Aztecan, Mexico), which encodes in multimorphemic words the distance of F relative to the speaker (proximal vs. medial vs. distal), the location of F relative to the speaker's line of sight (inside vs. outside), and the location of F relative to a mountain slope (foot vs. face vs. top) – e.g., *mah* means roughly 'away up there to the side in the face of the slope' (Casad and Langacker, 1985).

2.2 Topological relations

According to the loose, non-mathematical sense of 'topology' employed in research on spatial semantics, topological relations involve various types of allocentric contiguity between F and G, such as the notions of penetration and containment encoded by the English prepositions *through* and *in,* respectively. In an influential article building on a rich tradition of previous work, Landau and Jackendoff (1993) point out that the spatial concepts found in English prepositions are extremely coarse – in other words, very abstract, schematic, and categorical – since they place few geometric constraints on F and G. They also argue that these sorts of concepts are likely to be crosslinguistically universal. For example, based on the observation that English prepositions are insensitive to the specific shapes of F and G, they state that no language should have a locative element like the hypothetical *sprough,* which means 'reaching from end to end of a cigar-shaped object', as in *The rug extended sprough the airplane.* Similarly, given that English prepositions do not discriminate between the subregions of Gs that are containers, they propose that no language will manifest a locative element like the hypothetical *plin,* which means 'contact with the inner surface of a container', as in *Bill sprayed paint plin the tank.*

This orthodox view has been challenged by studies that have revealed considerable diversity in the kinds of topological relations that are lexicalized in various languages. To begin with the blackest fly in the ointment, Levinson (2003: 63, 72) notes that the putative non-existence of an expression like *sprough* is directly contradicted by Karuk (Hokan, Northwestern California), which has a suffix *-vara* meaning 'in through a tubular space'. Similarly, expressions of the *plin* type, which specify subregions of G, have been attested in Makah (Wakashan, Washington State), which has suffixes encoding locations such as 'at the rear of a house', 'at the base of an upright object', and 'at the head of a canoe' (Davidson, 1999). Equally if not more threatening to Landau and Jackendoff's theory is Tzeltal (Mayan, Southeastern Mexico), which describes topological relations with a large but, importantly, closed class of so-called dispositional adjectives that specify quite detailed, yet still essentially categorical, distinctions involving the location of F relative to G (Brown, 1994). When combined with the single, all-purpose relational marker *ta,* these words extensively cross-classify spatial arrays that would be described in English by using semantically more general prepositions like *in* and *on* (Table 1). Thus, if asked 'Where are the tortillas?' an English speaker might reply simply 'On the table', a statement that semantically reduces the tortillas to a mere point or shapeless blob; however, a Tzeltal speaker would probably select one of several terms that encode geometric information about the appearance of the tortillas, such as *latzal* (if they are stacked) or *pakal* (if they are folded).

Table 1. Examples of Tzeltal dispositional adjectives encoding topological relations that would normally be described in English as *in* or *on*. In each case, *ta* is a general-purpose marker meaning 'be located'. (Data reproduced from Brown, 1994.)

A. Ways of conveying 'in' relationships involving containment.

Form	Meaning	Eliciting F and G
t'umul ta	be located, by having been immersed in liquid in a container	apple, water in bucket
tik'il ta	be located, by having been inserted into a container with a narrow opening	bull, corral
xijil ta	be located, of long-thin object, by having been inserted carefully into a container	pencils, cup
xojol ta	be located, by having been inserted singly into a close-fitting container	coffee bag, pot
tz'apal ta	be located, by having been inserted at its end into supporting medium	stick, ground
lapal ta	be located, of long-thin-sharp object, by having been inserted through a flexible object	safety pin, cloth

B. Ways of conveying 'on' relationships involving contact with, and support by, a horizontal surface.

Form	Meaning	Eliciting F and G
pachal ta	be located, of a wide-mouthed container canonically 'sitting'	bowl, table
waxal ta	be located, of a tall oblong-shaped container or solid object canonically 'standing'	bottle, table
pakal ta	be located, of a blob with a distinguishably flat surface lying 'face' down	dough, table
lechel ta	be located, of a wide flat object lying flat	frying pan, table
chepel ta	be located, of a full (bulging) bag supported underneath	netbag, table
cholol ta	be located, of multiple objects arranged in a row	beans, table

Although languages differ greatly in the kinds of topological relations they encode, there are underlying patterns. In a recent study, nine unrelated languages[3] were investigated by comparing native speaker responses to a standardized set of 71 pictures showing a wide range of topological relations (Levinson and Meira, 2003). Results indicated that crosslinguistically the labels for pictures were not randomly distributed but instead tended to cluster, suggesting that the topological domain forms a coherent similarity space with a number of strong 'attractors', i.e., taxonomically basic-level categories that are statistically likely to be recognized by languages – in particular, notions such as containment, attachment, superadjacency, subadjacency, and proximity. Several generalizations about the organization of this abstract similarity space emerged from the study. First, each core concept has a prototype structure. For example, at the center of the cluster of containment pictures were scenes in which F is enclosed within G (e.g., a dog in a cage); scenes involving partial two-dimensional containment on a planar surface (e.g., a dog in a yard) were more peripheral, implying that English is somewhat unusual in using *in* for such topological relations. Second, the core concepts are arranged as neighbors along gradients in the similarity space, making some conflations of categories more natural than others. For instance, English *on* embraces both superadjacency (e.g., a cup on a table) and attachment (e.g., a picture on a wall), Berber *di* embraces both attachment (e.g., a picture on a wall) and containment (e.g., an apple in a bowl), and Spanish *en* embraces all three categories; however, there should not be, and do not as yet appear to be, any languages with a spatial morpheme that applies to superadjacency and containment while excluding attachment, since the latter concept is intermediate between the other two along the relevant gradient of the abstract similarity space. Third, each core concept can be further fractionated, leading to more fine-grained categories of topological relations. For example, the cluster of pictures for superadjacency included scenes both with and without contact (e.g., a cup on a table, and a lamp above a table), suggesting that languages are likely to use the same morpheme for these kinds of relations – a tendency that seems somewhat surprising from the perspective of English, since *on* and *above/over* divide the superadjacency category into separate subcategories distinguished by the presence or absence of contact between F and G. Levinson and Meira also report many intriguing cases of category fractionation in other languages, such as the exotic Tiriyó morpheme *awee,* glossed 'astraddle', which applies to the subset of attachment pictures in which F is suspended from a point on G and hangs down on either side of it (e.g., a coat on a hook, an earring dangling from a person's ear, a pendant on a chain, clothes drying on a line, a balloon on a stick, and a tablecloth on a table). Further analyses of crosslinguistic similiarities and differences in the subdomain of topological relations can be found in the detailed case studies compiled by Levinson and Wilkins (2006).

2.3 Projective relations

Projective relations involve locating F within a search domain that radiates out some distance from G along a specified angle or line. This class of categorical spatial relations breaks down into several subclasses, each of which exhibits substantial, but not unconstrained, crosslinguistic variation. The following summary is based mainly on Levinson's (2003) analysis. According to Levinson (2003, p. 76; see also Levinson and Wilkins, 2006), languages use, to varying degrees, three frames of reference for encoding (primarily) horizontal projective relations: 'the intrinsic system, which projects out a search domain from a named facet of a landmark object; the relative system, which imports the observer's bodily axes and maps them onto the ground object thus deriving named angles; and the absolute system, which uses a fixed set of bearings or a conceptual 'slope' to define a direction from a ground object'.

2.3.1 The intrinsic frame of reference

The first locative strategy has two steps: the speaker identifies a salient part or facet of G – e.g., the 'front' – and then extracts from the designated component an angle which extends outward a certain distance, thereby defining a search domain within which F can be found – e.g., *The ball is in front of the house.* In English this system operates mainly by imposing on G a six-sided, box-like 'armature' that yields a front, back, top, bottom, and two lateral (i.e., left and right) sides as the major intrinsic parts. Functional criteria are often used to identify, for instance, the 'front' of G based on factors like the typical direction of the perceptual apparatus (for animate entities), the typical direction of motion (for vehicles), or the typical direction of encounter (for houses, TVs, etc.). Some objects resist this decompositional approach because they appear to lack intrinsic asymmetries – e.g., English speakers do not construe trees and mountains as having fronts and backs. But judgments of this nature vary across languages – e.g., in Chamus (Nilo-Saharan, Kenya) the front of a tree is the side it leans toward, or, if it is vertical, the side with the biggest branch or the most branches, and in Kikuyu (Nilo-Saharan, Kenya) the front of a mountain is the side opposite its steepest side (Heine, 1997, p. 13).

It is crosslinguistically common for locative terms employing the intrinsic frame of reference to derive historically from body-part terms (Svorou, 1994; Heine, 1997; for recent crosslinguistic work on body part terms, see Majid, Enfield, and van Staden, 2006, as well as the critique by Wierzbicka, 2007). This can be seen in the English example used above – *The ball is in front of the house* – and in a number of fixed English expressions like *the face of a cliff, the mouth of a cave, the eye of a hurricane, the nose of an airplane, the head of a nail, the neck of a guitar, the arm/leg of a chair,* etc. In many languages, however, the body-part-based intrinsic system is quite complex, requiring regular linguistically driven visual analysis of the axial geometry as well as the major and minor protrusions of inanimate objects so that the relative appropriateness of different body-part terms can be computed instantly on the basis of these inherent properties, i.e., independent of the object's orientation or the speaker's viewpoint. Perhaps the best-studied language of this type is Tzeltal (Levinson, 1994), in which even a G as seemingly nondescript as a stone may be assigned a 'face', a 'nose', an 'ear', a 'back', a 'belly', or any of about fifteen other quasi-metaphorical body parts in order to specify that F is located within a search domain projected from one of these facets – e.g., an *s-jol* 'head' is a protrusion that can be found at one end of the major axis of G and that has a gently curved, circular outline with only minor concavities on either side.[4]

2.3.2 The relative frame of reference

To describe spatial arrays in which F is at some remove from G but G is classified as 'unfeatured' by the intrinsic system of the given language, the front/back and left/right axes of the observer's body can be introduced to provide a frame of reference for structuring the scenario. This increases the complexity of the spatial relations from binary (F and G) to ternary (F, G, and the observer). Thus, whereas *The ball is in front of the house* specifies a binary relation in which F is located with respect to an intrinsic facet of G, *The ball is in front of the pole* specifies a ternary relation in which F is located with respect to a non-intrinsic facet of G that can only be identified by taking into account the observer's perspective.

The type of relative system found in English involves imposing on G the mirror reflection of the observer's bodily axes (Figure 1A). A mirror flips the front/back axis but not the left/right axis of the object it reflects. To designate F as being *in front of* or *in back of* G, the observer's front/back axis is mapped onto G under 180º rotation, so that *The ball is in front of the pole* means 'From this viewpoint, the ball is in a search domain projected from the side of the pole that 'faces' me'. To designate F as being *left* or *right* of G, directions are projected laterally from G along angles that correspond to the observer's left/right axis. Besides the English system, there are two other logical possibilities for organizing the relative frame of reference on the horizontal plane, and both are utilized by other languages (Levinson, 2003: 84–89). One strategy, exemplified by some dialects of Tamil (Dravidian, India), involves mapping the observer's bodily axes onto G under complete 180º rotation, generating not only front/back reversal but also left/right reversal, so that *The ball is in front of the pole* has the same meaning as it does in English, but *The ball is to the left of the pole* means that the ball is located in the region that English speakers would consider 'to the right' (Figure 1B). The other strategy, exemplified by Hausa (Chadic, Nigeria), involves mapping the observer's bodily axes onto G without any rotation whatsoever, so that *The ball is in front of the pole* means that the ball is located in the region that English speakers would consider 'in back of', but *The ball is to the left of the pole* means the same thing as it does in English (Figure 1C).

A. Reflection (e.g., English)	B. Rotation (e.g., Tamil)	C. Translation (e.g., Hausa)
B	B	F
L **Ground** R	R **Ground** L	L **Ground** R
F	F	B
⇑	⇑	⇑
F	F	F
L **Observer** R	L **Observer** R	L **Observer** R
B	B	B

Figure 1. Crosslinguistic variation in the organization of the relative frame of reference on the horizontal plane. In the illustration of each type of system, the observer is shown at the bottom and the ground object at the top, with the observer's line of sight indicated by an arrow. Abbreviations: F = front, B = back, L = left, R = right.

2.3.3 The absolute frame of reference

The third type of angular specification on the horizontal plane involves an absolute frame of reference that provides a set of fixed bearings or cardinal directions, similar to north, south, east, and west. These bearings define 'an infinite sequence of parallel lines – a conceptual 'slope' – across the environment' (Levinson, 2003, p. 90). To indicate the location of F with respect to G, one projects an angle from G to F, assesses the orientation of this angle in relation to the grid of cardinal directions, and selects the appropriate term – e.g., something like *The ball is north of the pole.* Absolute systems are fundamentally geocentric, and languages often base terms for cardinal directions on stable environmental features like mountain slopes, river drainages, and prevailing wind patterns. For example, returning yet again to Tzeltal, it has an absolute system that is anchored in the mountain incline of the local landscape, giving rise to three directional terms: *ajk'ol* 'uphill' (roughly south), *alan* 'downhill' (roughly north), and *jejch* 'across' (either east or west) (Brown and Levinson, 1993, forthcoming). It is important to note (since this issue has been previously misunderstood – see the debate between Li and Gleitman, 2002, and Levinson et al., 2002) that although the terminology of absolute systems derives from environmental landmarks, such systems are fully abstracted, and in order to use them spontaneously and accurately, speakers must constantly monitor their spatial orientation by running a kind of mental compass. This is a remarkable neurocognitive capacity, as revealed by the anecdote about the Tzeltal speaker, Slus, in the epigraph of this chapter. Another vital point is that unlike the English *north/south/east/west* system, which has extremely limited use, the absolute systems under discussion are regularly employed to describe spatial arrays at every level of scale, from inches to miles. This is clearly shown in Levinson's (2003, p. 114) description of Guugu Yimithirr (Pama-Nyungan, Australia), which uses exclusively the absolute frame of reference for characterizing horizontal projective relations:

> In GY, in order to describe someone as standing in front of the tree, one says something equivalent (as approximate) to 'George is just north of the tree', or, to tell someone to take the next left turn, 'go north', or, to ask someone to move over a bit, 'move a bit east', or, to instruct a carpenter to make a door jamb vertical, 'move it a little north', or, to tell someone where you left your tobacco, 'I left it on the southern edge of the western table in your house', or, to ask someone to turn off the camping gas stove, 'turn the knob west', and so on. So thoroughgoing is the use of cardinal directions in GY that just as we think of a picture as containing virtual space, so that we describe an elephant as behind a tree in a children's book (based on apparent occlusion), so GY speakers think about it as an oriented virtual space: if I am looking at the book facing north, then the elephant is north of the tree, and if I want you to skip ahead in the book I will ask you to go further east (because the pages would then be flipped from east to west).

2.3.4 The vertical dimension

Finally, with regard to the linguistic encoding of projective relations along the vertical dimension, the three frames of reference – intrinsic, relative, and absolute – usually coincide and yield the same answer to the question 'Where is F in relation to G?' (Levinson, 2003, p. 75). For example, consider a scene in which a fly hovers above a bottle. F is 'above' G according to all three criteria: it is located within the search domain that radiates from the top of the bottle (intrinsic frame); it is higher than the bottle in the observer's visual field (relative frame); and it is higher than the bottle along the vertical axis defined by gravity (absolute frame). However, as a number of experiments have shown (e.g., Friederici and Levelt, 1990; Carlson-Radvansky and Irwin, 1993; Carlson, 1999), the three frames of reference can be manipulated independently of each other (e.g., by rotating either G or the observer, or, more radically, by shifting the entire array to a zero gravity environment) to create special situations in which they yield conflicting answers to the *Where*-question. Also, as noted earlier, although English clearly distinguishes *above/over* from *on* according to whether F contacts G, this may be the result of splitting into two subcategories the crosslinguistically more common (and perhaps conceptually more basic) category of superadjacency, which is neutral with respect to contact and is directly encoded in languages like Japanese and Arrernte (Pama-Nyungan, Australia). This is one of several ways in which the vertical dimension interacts with topology. Another manifestation of this interaction is that *over* and *under* are not synonymous with *above* and *below,* respectively, because the former prepositions have a topological component that makes them more suitable than the latter for describing spatial arrays that involve an encompassment relation – e.g., it is more felicitous to say that a penny is *under* than *below* an inverted cup on a table (Coventry et al., 2001).

2.4 Summary

Two major generalizations emerge from this review of the kinds of categorical spatial relations that are encoded in languages. First, there is a huge amount of crosslinguistic variation regarding the specific concepts that are lexicalized, suggesting that every language has its own unique spatial ontology with idiosyncratic notions ranging from Limbu's *ma:dha:mbi* ('F is on the slope of the mountain ridge across the valley from where the speaker is situated') to Tiriyó's *awee* ('F is astraddle G') to Tzeltal's *ajk'ol* ('F is uphillwards, i.e., roughly south, of G'). Second, despite this tremendous diversity, a number of patterns can be identified that lend coherence to each of the semantic fields comprising the overall conceptual domain. For instance, in the field of deictic relations, over 50% of languages appear to have demonstrative systems that specify a binary proximal/distal contrast; in the field of topological relations, a relatively small number of core concepts tend to recur across languages and hence constitute statistical attractors for lexicalization; and in the field of projective relations, languages typically have complex sets of expressions that instantiate up to three frames of reference – intrinsic, relative, and absolute.

3 What are the neuroanatomical correlates of linguistically encoded categorical spatial relations?

Very little research in cognitive neuroscience has explored which brain structures subserve the rich variety of categorical spatial relations that are lexicalized in languages around the world. Nevertheless, all of the studies that have addressed this issue suggest that the left inferior parietal lobule (IPL) is an especially important cortical region. It is well-established that the visual system consists of two major subsystems – the so-called 'what' pathway that projects from the occipital lobe to the ventral temporal lobe and processes complex information about shape, color, and texture that is necessary for conscious object perception and recognition; and the so-called 'where' pathway that projects from the occipital lobe to the parietal lobe and processes complex information about space that is necessary for efficient sensorimotor interaction with objects (for a review see Milner and Goodale, 2006). In an influential article, Landau and Jackendoff (1993) used this distinction as the basis for speculating that the meanings of English locative prepositions are represented in the left IPL. The studies summarized below not only corroborate this proposal but allow it to be made more precise by suggesting that the critical neuroanatomical structures are the supramarginal gyrus and, perhaps to a lesser extent, the angular gyrus.

3.1 Studies implicating the left inferior parietal lobe

3.1.1 Supramarginal gyrus

Damasio et al. (2001) report a positron emission tomography (PET) study in which English speakers viewed drawings of static spatial relations between objects (e.g., a cup on a table) and performed two tasks: naming F, and naming the spatial relation between F and G with an appropriate preposition. When the condition of naming objects was subtracted from that of naming spatial relations, the largest and strongest area of activation was in the left supramarginal gyrus (SMG). The authors do not indicate which prepositions were targeted for production, but it appears that a mixture of topological and projective prepositions were included, which suggests that the SMG activation reflects semantic processing of both types. More recently, a functional magnetic resonance imaging (fMRI) study also found significant SMG activation during a task requiring semantic processing (within the relative frame of reference) of the Dutch equivalents of the terms *left* and *right* (Noordzij et al., 2008).

Additional evidence comes from a neuropsychological study conducted by Tranel and Kemmerer (2004; see also Kemmerer and Tranel, 2000, 2003, and Kemmerer, 2005). They administered a set of tests that require production, comprehension, and semantic analysis of 12 English prepositions (encoding topological relations as well as several kinds of projective relations) to 78 brain-damaged subjects with lesions distributed throughout the left and right cerebral hemispheres, and then compared the lesion sites of the subjects who were impaired on the tests with the lesion sites of those who were unimpaired. Poor performance was linked specifically with damage in the left SMG and the left frontal operculum. The involvement of the left SMG strengthens the hypothesis that this region plays an essential role in representing the spatial meanings of English prepositions. The investigators did not, however, conduct separate analyses to determine whether the different semantic classes of prepositions dissociate from each other behaviorally and neuroanatomically. As for the involvement of the left frontal operculum, it may reflect either or both of two functions: phonological encoding, possibly in Brodmann area 44 (e.g., Amunts et al., 2004), and semantic working memory, possibly in Brodmann areas 45 and/or 47 (e.g., Devlin et al., 2003; Thompson-Schill et al., 1999; Wagner et al., 2001).

To my knowledge, no other studies of spoken languages have identified a strong association between the left SMG and morphemes that denote categorical spatial relations;[5] however, further evidence for precisely this association comes from two functional neuroimaging studies of locative classifier constructions in American Sign Language (ASL; Emmorey et al., 2002) and British Sign Language (BSL; MacSweeney et al., 2002). Locative classifier constructions are complex coding devices that exploit the three-dimensional medium of signing space in the following ways: the relative positions of the hands in front of the body correspond schematically and iconically to the relative positions of F and G in the physical world, and the shape of each hand indicates the general class to which each object belongs (Emmorey, 2003). For example, to express the equivalent of *The bike is near the house,* the referential handshapes for house and bike are articulated sequentially (G preceding F), and then the classifier for vehicles (thumb, middle, and index fingers extended) is placed directly adjacent to the classifer for large bulky objects (five fingers spread and curved) to indicate topographically that F is 'near' G. To investigate the neural substrates of this unique form of spatial description, Emmorey et al. (2002) conducted a PET study in which deaf native ASL signers viewed the same kinds of drawings of spatial relations that were used in Damasio et al.'s (2001) PET study, and performed two tasks: naming F, and naming the spatial relation between F and G with an appropriate locative classifier construction. Relative to naming objects, naming spatial relations engaged the left SMG; moreover, the centroid of activation was similar to that found for English speakers in Damasio et al.'s (2001) study, suggesting that it reflects semantic processing. In another study, MacSweeney et al. (2002) used fMRI to investigate the neural systems underlying comprehension of BSL sentences containing locative classifier constructions. Compared to sentences without such constructions, activation was observed in the same sector of the left SMG as in Emmorey et al.'s (2002) study, providing additional support for the hypothesis that this cortical area contributes to the semantic processing of linguistically encoded categorical spatial relations.

3.1.2 Angular gyrus

Neuroimaging and neuropsychological studies suggest that the left angular gyrus (AG) is also involved in the linguistic representation of categorical spatial relations, but perhaps to a more limited degree than the left SMG. Baciu et al. (1999) report an fMRI study in which significantly stronger left than right AG activation was observed while subjects judged whether a dot was presented above or below a bar. This task has a core linguistic component because, as the instructions clearly indicate, the two categories that must be discriminated are directly encoded by the projective prepositions *above* and *below*. This particular spatial contrast may seem natural and intuitive to English speakers, but it is by no means crosslinguistically universal, since some languages do not have morphemes that distinguish 'above' from 'on' or 'below' from 'in' (Levinson, 2003, p. 73; Levinson and Meira, 2003, p. 507). Hence the left AG activation may reflect, in part, the essentially lexicosemantic process of classifying the location of the dot as falling within one of two projected search domains – 'above' or 'below' the line – that are both familiar categories in the spatial ontology of the subject's native language.[6]

Another linguistically encoded categorical spatial contrast that has been linked, albeit loosely, with the left AG is the distinction between *left* and *right* within the intrinsic frame of reference. Lesions centered in the left AG sometimes produce Gerstmann syndrome, which comprises the following four symptoms: left/right confusion, finger agnosia, agraphia, and acalculia (Gerstmann, 1957; Mayer et al., 1999; Mazzoni et al., 1990; Morris et al., 1984; Roeltgen et al., 1983; Varney, 1984). The symptom of left/right confusion is usually manifested as difficulty pointing to left and right body parts on command. However, the relevance of this particular deficit to the issue of the neural correlates of the meanings of *left* and *right* is limited in two ways. First, knowledge of the actual meanings of the terms is not always disrupted; instead, what seems to be impaired are certain cognitive operations that are necessary to <u>apply</u> the meanings appropriately in certain situations, such as the ability to mentally rotate visual images of the body in space (Bonda et al., 1995; Mayer et al., 1999; Zacks et al., 1999). For example, Mayer et al.'s (1999) subject performed well (15/16 correct) when asked to point with either hand to designated left and right parts of his own body, but performed poorly (11/16 correct) when asked to point with a specified hand to designated left and right parts of a line drawing of a human body that was facing him and hence had a 180° reversal of left and right sides relative to his own body. Second, studies of Gerstmann syndrome are generally restricted to the use of *left* and *right* to refer to the intrinsic sides of the human body under various conditions; they do not pursue the inquiry further by systematically assessing whether the subject understands how the terms are also used – in English but not in all languages (like Guugu Yimithirr and Tzeltal) – to specify regions of space that are (a) projected outward from the intrinsic sides of the body (e.g., *The ball is on your left*), (b) projected outward from the intrinsic sides of inanimate objects (e.g., *The ball is on the car's left-hand side*), and (c) projected outward from the sides of unfaceted objects by importing the speaker's own left/right bodily axis as a frame of reference (e.g., *The ball is to the left of the pole*).

3.2 Further neuroanatomical questions raised by linguistic typology

The studies reviewed above suggest that the left IPL is a key cortical region for representing the meanings of locative expressions. But when these studies are considered in the context of the preceding typological survey of the kinds of categorical spatial relations that are encoded crosslinguistically, it immediately becomes clear that the research done so far is merely spadework, and that most of this rich neurocognitive terrain remains to be mined. For example, at this point it is not known whether the three major classes of categorical spatial relations – deictic, topological, and projective – are subserved by separate neural networks within the left IPL. Many questions can also be raised about the specific neural organization of each class, as suggested below.

3.2.1 Deictic relations

I am not aware of any studies that have explored the neural correlates of demonstratives that specify egocentrically-anchored deictic spatial relations, although in a previous paper (Kemmerer, 1999) I pointed out that this topic is interesting in light of the mounting evidence for separate circuits representing, on the one hand, near or peripersonal space which extends roughly to the perimeter of arm's reach, and on the other hand, far or extrapersonal space which extends outward from that fuzzy boundary (for a review see Berti and Rizzolatti, 2002; see also Longo and Lourenco, 2006, and Makin et al., 2007). The representational division of near and far sectors of space may derive from computational differences in the forms of sensorimotor control that are typical for each sector – i.e., primarily visually guided manual activity in the near sector, and primarily visual search and object scanning or 'parsing' in the far sector. It is tempting to speculate that this fundamental division is causally relevant to the fact that the majority of languages worldwide have demonstrative systems that encode a binary proximal/distal contrast.[7] It is also important to bear in mind, however, that demonstratives are not restricted to quantitative spatial distinctions such as within vs. beyond arm's reach; instead, objective distances are semantic variables that are assigned values on-the-fly by pragmatic factors, thereby allowing speakers to expand or contract the referential range of demonstratives as needed – e.g., as noted by Levinson (1983, p. 80), the statement *Place it here* 'may have quite different implications of precision if said to a crane operator or a fellow surgeon'. In addition, some languages have demonstrative systems that carve the radial space surrounding the speaker into three or, as in the unusual case of Tlingit, even four concentric zones, thereby violating the two-way perceptual distinction. Perhaps the abstract meanings of demonstratives are subserved by the left IPL, just like the other types of linguistically encoded categorical spatial relations described above. But for demonstrative systems that incorporate geographical information, such as the Limbu and Cora systems involving mountain slopes, the semantic structures may recruit not only the dorsal 'where' pathway extending into the left IPL, but also the ventral 'what' pathway extending into the inferotemporal cortex (Milner and Goodale, 2006).

3.2.2 Topological relations

Further research on the neural correlates of linguistically encoded topological relations could benefit greatly by utilizing carefully designed stimuli that take into account theoretically important semantic dimensions, like the standardized set of 71 pictures that Levinson and Meira (2003) employed in their crosslinguistic comparison (see also Levinson and Wilkins, 2006). By conducting high-resolution functional neuroimaging studies with such materials, it may be possible to test the hypothesis that the conceptual similarity space discovered by Levinson and Meira (2003) – a similarity space organized in terms of notions such as containment, attachment, superadjacency, subadjacency, and proximity – is neuroanatomically implemented in the form of a topographically structured cortical map in the left IPL, most likely the SMG. Within this map, the representational dimensions of the conceptual space might be captured, albeit in a warped manner, by the physical distribution of cortical columns (Kohonen and Hari, 1999; Simmons and Barsalou, 2003; Graziano and Aflalo, 2007; Kriegeskorte et al., 2008). This is, however, an admittedly bold conjecture.

Another hypothesis is that the inferotemporal cortex contributes to representing the detailed geometric features of objects that Tzeltal incorporates into the meanings of dispositional adjectives. Besides encoding various forms of allocentric contiguity between F and G, such as containment or surface contact and support, many dispositional adjectives also indicate, in a manner much more specific than Indo-European languages, the shape or configuration of F relative to G (see Table 1). These terms are semantically somewhat similar to the classifiers that are prevalent in sign languages, and Emmorey et al. (2002) report that in their PET study the production of locative classifier constructions engaged not only the SMG but also the left posterior inferotemporal region – a finding which supports the view that the same region might contribute to the geometric component of the meanings of Tzeltal dispositional adjectives.

3.2.3 Projective relations

Projective relations may constitute the subdomain of spatial representation with the greatest potential for interdisciplinary cross-talk between linguistic typology and cognitive neuroscience, because research on the central issue of frames of reference is highly developed in both areas of inquiry (for the best linguistic overview, see Levinson, 2003; for excellent neuroscientific overviews, see Hillis, 2006; Previc, 1998; Robertson, 2004). The direction of influence can certainly go both ways, but here I restrict the discussion to a small sample of the many ways in which recent findings from linguistic typology can generate intriguing questions about the neural substrates of linguistically encoded categorical spatial relations involving intrinsic, relative, and absolute frames of reference.

An important discovery in linguistic typology is that terms for projective relations involving the intrinsic frame of reference often derive historically from body-part terms. Moreover, in some languages the application of such terms to the facets of inanimate objects, for the purpose of anchoring a search domain within which F can be located, usually requires a complex visuospatial analysis of axial and contour features – e.g., in Tzeltal an *s-ni* 'nose' is a pointed extremity or an extremity having a sharp convexity, and an *x-chikin* 'ear' is a flattened protrusion. What is the neural basis of terms like these? One hypothesis that warrants investigation (Kemmerer and Tranel, 2008) is that the meanings of such terms depend on shape-sensitive regions of the posterior lateral/

inferior temporal cortex that receive input from the recently discovered 'extrastriate body area' (EBA), which appears to be especially important for the visual categorization of human body parts (e.g., Peelen and Downing, 2007).

The fMRI study by Noordzij et al. (2008) implicates the left SMG in the use of *left* and *right* to designate projective relations involving the relative frame of reference. Would the same type of activation be observed when Tamil speakers perform the same task? As noted earlier, Tamil employs a strategy of rotation rather than reflection, so that a sentence like *The triangle is to the left of the circle* means that the triangle is located within a search domain that English speakers would consider 'to the right' (see Figure 1B).

Perhaps the best example of how linguistic typology can inspire future research on the neural representation of categorical spatial relations involves the systems of cardinal direction terms analogous to *north/south/east/west* that speakers of languages like Tzeltal and Guugu Yimithirr use habitually to specify the angular location of F relative to G according to an absolute frame of reference. Such linguistic behavior requires a mental compass that constantly computes one's orientation within a conventional framework of fixed bearings. Many nonhuman species have evolutionarily specialized sensory devices that enable them to use absolute coordinates for navigation – e.g., some species of migratory birds have light-absorbing molecules in their retinae that are sensitive to the magnetic field of the earth and that may enable the birds to see this information as patterns of color or light intensity (Ritz et al., 2004); sea turtles have the biological equivalent of a magnetically based global positioning system that allows them to pinpoint their location relative to geographically large target areas (Luschi et al., 2007); and locusts perceive polarization patterns in the blue sky and use them as cues for spatial orientation (Heize and Homberg, 2007). But for people in 'absolute' communities the mental compass that generates their superb sense of direction – a sense comparable in accuracy to that of homing pigeons (Levinson, 2003, p. 232) – is presumably not genetically programmed but may instead be a 'knock-on' effect of the intensive training in orientation tracking that comes with speaking a language that regularly employs cardinal direction terms to describe spatial arrays at every level of scale (Levinson, 2003, p. 278; see also Haun et al., 2006a, 2006b). It is reasonable to suppose that relevant brain areas include parietal as well as hippocampal and entorhinal structures that have been implicated in both constructing landmark-based cognitive maps of the environment and monitoring one's movement through them (e.g., Ekstrom et al., 2003; Hartley et al., 2003; Janzen and van Turennout, 2004; Hafting et al., 2005; Leutgeb et al., 2007; Spiers and Maguire, 2006). However, because the use of the mental compass does not require input from visually perceived landmarks (as illustrated in the epigraph of this paper), other neural systems must also be recruited, presumably to carry out the computations that underlie dead-reckoning – that is, keeping track of distances traveled along each angular heading. Identifying these systems is clearly an exciting direction for future research (for important new clues, see Sargolini et al., 2006; Heyman, 2006; Jeffrey and Burgess, 2006).

3.3 Summary

Research on the neuroanatomical substrates of linguistically encoded categorical spatial relations has only recently begun, but the studies conducted so far consistently point to the left IPL as an essential region. Taking into account the view from linguistic typology, which provides not only a well-developed theoretical framework but also detailed semantic analyses of the variety of spatial coding systems manifested in the languages of the world, can lead to many new questions regarding the neural basis of this rich conceptual domain.

4 How do linguistically encoded categorical spatial relations interact with perception and cognition?

The final topic of discussion involves the interaction between linguistic and perceptual/cognitive representations of categorical spatial relations. Very little is currently known about the nature of this interaction, but the existing data suggest that it is quite complicated (Papafragou, 2007). Two main points are elaborated below. First, several studies with both normal and brain-damaged populations indicate that the kinds of categorical spatial distinctions that are encoded for non-linguistic perceptual/cognitive purposes are to some extent separate from the diverse spatial categorization systems of languages around the world. Second, a number of other studies suggest that the unique spatial ontology of one's language nevertheless has the power to influence one's perceptual/cognitive representations of categorical spatial relations by both *decreasing* sensitivity to distinctions that are not captured by one's language and *increasing* sensitivity to distinctions that are. The fact that these two sets of findings are not easy to reconcile is a clear sign that we are still far from understanding the intricacies of the interaction between linguistic and non-linguistic representations of space.

4.1 Linguistic and perceptual/cognitive representations of categorical spatial relations are to some extent distinct

Although English distinguishes *on* from *above/over,* many other languages – perhaps even the majority (Levinson and Meira, 2003) – have morphemes that encode the general notion of superadjacency, which is neutral with respect to whether F contacts G. Korean is one such language. To investigate whether this form of crosslinguistic variation influences non-linguistic spatial memory, Munnich et al. (2001) asked native speakers of English and Korean to perform two tasks with the same stimuli, which consisted of spatial arrays showing a ball in any of 72 locations superadjacent to a table. In the naming task, subjects completed the sentence 'The ball is ___ the table' (or the equivalent sentence in Korean). In the memory task, they viewed an array for 500 ms, and then after a 500 ms delay they saw another array which they judged as being either the same as or different from the initial one. In the naming task the English speakers consistently employed the lexical contrast between *on* and *above/over,* whereas the Korean speakers rarely mentioned the contact/noncontact distinction. In the memory task, however, the two subject groups had almost identical patterns of accuracy for all 72 locations, including an advantage for locations aligned with the surface of the table. This study therefore suggests that non-linguistic spatial memory is not constrained by whether the contact/noncontact distinction is linguistically encoded on a regular basis throughout one's life. In other words, even though Korean does not force speakers to fractionate the category of superadjacency according to the presence or absence of contact between F and G, this spatial distinction is nevertheless perceptually salient enough to influence the operation of recognition memory in Korean speakers.

Neuropsychological data also support the view that linguistic and perceptual/cognitive representations of categorical spatial relations are at least partially separate. As noted earlier, Tranel and Kemmerer (2004) found maximal lesion overlap in the left SMG for a group of brain-damaged subjects who had pervasive and severe defects in the knowledge of the meanings of English prepositions. In a follow-up experiment with these subjects, non-linguistic visuospatial processing was assessed by administering a battery of standardized neuropsychological tests, including three subtests from the Wechsler Adult Intelligence Scale-III (Matrix Reasoning, Block Design, and Object Assembly), the Benton Facial Recognition Test, the Benton Judgment of Line Orientation Test, the Hooper Visual Organization Test, the Complex Figure Test (copy), and the Benton Three-Dimensional Block Construction Test (Benton and Tranel, 1993; Tranel, 1996). Overall, the subjects performed extremely well on the various tests. Although two of the tests – the Benton Facial Recognition Test and the Benton Judgment of Line Orientation Test – emphasize sensitivity to coordinate spatial relations, the remaining tests arguably require an appreciation of categorical spatial relations.[8] Moreover, Kemmerer and Tranel (2000) describe a subject with a large right-hemisphere lesion affecting frontoparietal and temporal regions who manifested a dissociation that was the opposite of the kind manifested by the brain-damaged subjects in Tranel and Kemmerer's (2004) study – namely, intact knowledge of the meanings of English prepositions but impaired nonlinguistic visuospatial processing of coordinate as well as categorical spatial relations. Taken together, these findings constitute evidence for what Jager and Postma (2003, p. 513) call 'a tripartition between perceptual coordinate spatial codes, perceptual categorical spatial codes, and verbal categorical spatial codes'.

Additional neuropsychological evidence for this 'tripartition' comes from Laeng (1994), who evaluated the performance of 60 brain-damaged subjects, 30 with unilateral left-hemisphere (LH) lesions and 30 with unilateral right-hemisphere (RH) lesions, on the following tasks. First, subjects were shown a drawing of two objects bearing a certain spatial relation to each other (e.g., a large cat to the left of a small cat), and after a short delay they were shown another drawing and were asked to make a same/different judgment (analogous to the recognition memory task in Munnich et al.'s 2001 study); half of the drawings were different, and the change was along either the categorical dimension (e.g., a large cat to the right of a small cat) or the coordinate dimension (e.g., a large cat to the left of a small cat, but a different distance away). Second, once again subjects were shown a drawing of a spatial relation, but this time after a short delay they were shown two other drawings and were asked to decide which was more similar to the initial one; alterations were either categorical or coordinate. On both tasks, LH-damaged subjects had greater difficulty detecting categorical than coordinate changes, and RH-damaged subjects exhibited the opposite pattern. Importantly, Laeng (1994) also found that the LH-damaged subjects' scores on several aphasia tests – including the Token Test, which has commands that incorporate prepositions (e.g., 'Point to the square to the left of the blue circle') – did not correlate with their scores on the nonlinguistic spatial tests, supporting the view that linguistic and perceptual representations of categorical spatial relations are to some extent distinct. Further research is clearly necessary, however, to explore the nature of this distinction in greater detail.

4.2 Linguistic representations of space can influence perceptual/cognitive representations of space

The studies reviewed above suggest that the kinds of categorical spatial distinctions that are encoded for nonlinguistic purposes are at least partially separate from the spatial ontology of one's language. However, a number of recent studies suggest that language can nevertheless influence perceptual/cognitive representations of space by modulating sensitivity to certain distinctions. These studies support the 'Whorfian hypothesis' that language modifies thought – a hypothesis that was widely embraced during the 1950s and 1960s, fell into disrepute during the 1970s and 1980s, and was resurrected in the mid-1990s because of theoretical and methodological advances that have continued to develop (e.g., Gentner and Goldin-Meadow, 2003; Gilbert et al., 2006).

4.2.1 Language can decrease sensitivity to certain categorical spatial distinctions

As summarized above, Munnich et al. (2001) found that even though Korean does not lexicalize the contact/noncontact distinction, speakers are still sensitive to it for nonlinguistic purposes such as recognition memory. However, there is also evidence that in some cases sensitivity to a particular categorical spatial distinction is present in infancy but then gradually diminishes during an early stage of language acquisition because the distinction is not captured by the target language being learned. This type of scenario is illustrated by a study that focused on the following contrast between English and Korean strategies for describing actions involving topological relations of containment (McDonough et al., 2003). The English expression *put in* specifies that F ends up occupying an interior region of G, but is neutral with respect to whether F fits tightly or loosely within G. In Korean, on the other hand, the notion of containment is subdivided into two different categories: *kkita* designates the creation of a tight-fitting relation between F and G (e.g., putting a cassette in a case), and *nehta* designates the creation of a loose-fitting relation between F and G (e.g., putting an apple in a bowl). Using a preferential looking paradigm as an indirect measure perceptual categorization, McDonough et al. (2003) found that infants as young as 9 months of age, from both English- and Korean-speaking environments, can discriminate between tight and loose containment events (see also Hespos and Spelke, 2004). This kind of spatial sensitivity is clearly useful for infants growing up in Korean-speaking environments, but it is ultimately less valuable for infants growing up in English-speaking environments, and in fact when adult speakers of each language were given the same preferential looking task, the Korean speakers exhibited sensitivity to the tight/loose distinction, but the English speakers did not. In another experiment that evaluated the adult speakers' recognition of the distinction more explicitly, subjects observed the enactment of three tight containment events and one loose containment event, and then answered the question 'Which is the odd one?' Significantly more Korean- than English-speaking adults based their choice on degree of fit (80% vs. 37%).

The investigators interpret their findings as evidence that when language-specific spatial categories are being learned, the perceptual judgments that are necessary to use them efficiently become increasingly rapid and automatic. Thus Korean speakers implicitly monitor the tightness of fit of containment relations because the grammatical system of their language regularly forces them to encode distinctions along this parameter. However, spatial sensitivities that are not needed in order to use the local language may fade – e.g., English speakers can safely ignore the tight/loose contrast most of the time. As McDonough et al. (2003) point out, the loss of sensitivity to the tight/loose contrast is remarkably similar to another dramatic instance of perceptual tuning that takes place during early language development, namely the loss of phonetic contrasts that are not phonemic in the target language (Kuhl, 2004; Kuhl et al., 1992; Werker and Tees, 1984).

Linguistically induced downgrading of spatial acuity is also illustrated by Levinson's (2003: 152–4) report that Tzeltal speakers have difficulty distinguishing mirror stimuli – e.g., **b** vs. **d.** This intriguing perceptual deficiency may derive from the fact that Tzeltal makes no use of the egocentrically anchored relative frame of reference, relying instead on two other locative strategies (apart from dispositional adjectives for describing topological relations): body-part terms based on the intrinsic frame of reference for describing closely contiguous projective relations, and directional terms based on the absolute frame of reference for describing more distant projective relations. Because the mirror stimuli employed in the experiment were multicomponent line figures, they were most likely processed according to the intrinsic system, which is essentially orientation-free and viewer-independent. The Tzeltal speakers' difficulty in detecting the difference between unreflected and reflected images cannot be attributed to low education, since educationally matched speakers of Totonac, a Mayan language that does use the relative frame of reference, performed the task much like Dutch speakers. From the perspective of cognitive neuroscience, it is striking that the behavior of the Tzeltal speakers resembles that of brain-damaged English and Italian speakers who have selectively impaired mirror-stimulus discrimination (Davidoff and Warrington, 2001; Priftis et al., 2003; Turnbull and McCarthy, 1996; McCloskey, 2009). A direction for future research would be to carefully compare the presumably linguistically induced decrement in discriminating mirror stimuli exhibited by Tzeltal speakers with the clearly neurologically induced form of mirror stimulus agnosia exhibited by brain-damaged subjects.

4.2.2 Language can increase sensitivity to certain categorical spatial distinctions

There are also reasons to believe that language can cause speakers to become more attuned to particularly subtle or non-obvious types of spatial relationships. According to Bowerman and Choi (2003, p. 417), 'In cases like this, an important stimulant to comparison can be hearing the same word. As the child encounters successive uses of the word, she 'tries' (although this process is presumably rarely if ever conscious) to align the referent situations and work out what they have in common. Sometimes … there is no existing concept that does the job, and the child has to construct a new one to account for the distribution of the word'. An excellent example is the Tiriyó morpheme *awee* which refers to situations in which F is suspended from a point on G and hangs down on either side of it, hence treating as equivalent such superficially diverse spatial arrays as a necklace around a person's neck, a tablecloth draped over a table, and a clothespin dangling from a line. It is not known whether infants are sensitive to this highly language-specific spatial category, but it seems likely that they are not and that they must therefore gradually construct the concept through multiple exposures to *awee* when acquiring Tiriyó. Another good example is the Chamus strategy of treating the intrinsic front of a tree as either the side it leans toward or, in case it is perfectly vertical, the side with the biggest branch or the most branches. It seems safe to assume that these are features of trees that English speakers do not usually register, although Chamus speakers must attend to them in order to use the grammatical system of the language appropriately. In this manner language can be said to provide 'on-the-job training for attention' (Smith et al., 2002). As Majid (2002) observes, it is useful to think of this form of linguistically driven perceptual tuning as similar to the novice-to-expert shift in categorization abilities that is known to engender more refined representations for the target domain (Palmeri et al., 2004).

The most systematic and intensive investigation of linguistic influences on cognitive representations of space has been conducted by Stephen Levinson and his colleagues (Levinson, 2003; Majid et al., 2004; Pederson et al., 1998). Although this area of inquiry is controversial (see the debate between Li and Gleitman, 2002, and Levinson et al., 2002), several experiments suggest that there may be deep cognitive consequences of speaking a language that employs predominantly either the relative frame of reference, like English and Dutch, or the absolute frame of reference, like Guugu Yimithirr and Tzeltal, for describing projective relations. The central experimental method involves a rotation paradigm which makes it possible to identify the frame of reference that subjects use to carry out various types of nonlinguistic cognitive tasks, such as memory tasks that probe both recognition and recall, maze task that require tracking motion and path direction, and reasoning tasks that evaluate transitive inference. To take a straightforward example, subjects are first seated at a table on which three toy animals are lined up headed leftward, or south, and then they are rotated 180º and seated at a different table where they must arrange an identical set of toy animals so that they are just as before, with an emphasis on remembering the linear order. If subjects orient the animals in a leftward direction, they are invoking an egocentric frame of reference, but if they orient the animals in a rightward (i.e., southerly) direction, they are invoking an absolute frame of reference. When performing this as well as all other nonlinguistic cognitive tasks involving the rotation paradigm, subjects overwhelmingly follow the coding pattern of their language. Such results have been obtained with speakers of a variety of 'relative' languages – e.g., English, Dutch, Japanese, and Yukatek (Mayan, Mexico) – and 'absolute' languages – e.g., Guugu Yimithirr, Tzeltal, Arrernte, Hai//om (Khoisan, Namibia), Longgu (Austronesian, Solomon Islands), Balinese (Austronesian, Indonesia), and Belhare (Sino-Tibetan, Nepal).

Levinson (2003: 290–291) argues that these effects are due to the fact that relative and absolute frames are incommensurable – e.g., from the proposition 'The knife is to the right of the fork' one cannot derive the proposition 'The knife is to the south of the fork', or vice versa:

> Once a language has opted for one of these frames of reference and not the other, all the systems that support language, from memory, to reasoning, to gesture, have to provide information in the same frame of reference. If I remember an array as 'The knife is to the right of the fork' but live in a community where no left/right terminology or computation is part of everyday life, I simply will not be able to describe it. For my memory will have failed to support the local description system, in, say, terms of north and south. The use of language thus forces other systems to come into line in such a way that semantic parameters in the public language are supported by internal systems keeping track of all experience coded in the same parameters.

Despite Levinson's assertions, this area of research remains quite contentious. Nevertheless, there is sufficient data to motivate questions regarding the neural substrates of linguistically driven cognitive restructuring. For example, although neuropsychological studies have shown that some brain-damaged subjects with impaired knowledge of the meanings of English prepositions can still accomplish tasks requiring nonlinguistic processing of categorical spatial relations (Kemmerer and Tranel, 2000; Tranel and Kemmerer, 2004), it is unknown how such subjects would perform on the various kinds of nonlinguistic 'space games' that Levinson and his colleagues have developed around the rotation paradigm. How would an English-speaking brain-damaged subject with severely disrupted knowledge of *left* and *right* perform on the 'animals in a row' task described above? And what would the results reveal about the interface between linguistic and cognitive representations of space? These questions, and many others that involve integrating linguistic typology and cognitive neuroscience, await future research.

4.3 Summary

Experimental studies with normal as well as brain-damaged subjects suggest that the meanings of locative expressions are language-specific semantic structures that are activated primarily when a person packages his or her conceptualizations of space in a manner that can easily be communicated in words – a process that Slobin (1996) calls 'thinking for speaking'. These linguistic representations are at least partially distinct from the perceptual/cognitive representations used in many visuospatial and visuomotor tasks such as recognizing, drawing, and constructing complex spatial arrays. At the same time, however, recent findings from the neo-Whorfian movement suggest that the unique way in which one's language structures space has implications for other mental systems, bringing about not only modifications of perceptual sensitivities but also adjustments of cognitive styles. I consider it a safe bet that most if not all of the readers of this article are usually *oblivious* to their orientation with respect to north, south, east, and west; however, there are a great many cultures in which people can *instantly* indicate cardinal directions like these. Such profound cognitive differences may be largely due to linguistic differences. The correct theory of the interface between linguistically and nonlinguistically encoded categorical spatial relations remains a topic of future research, and I submit that the most progress will be made through interdisciplinary efforts that include the mutually informing perspectives of linguistic typology and cognitive neuroscience.

5 Conclusion

People worldwide talk on a daily basis about the three-dimensional spatial world that surrounds them, and most of the time the content of their discourse concerns broadly defined spatial categories – e.g., *The book is right here in front of me* – as opposed to metrically exact notions – e.g., *The book is* 53 *centimeters from my chest.* This is obviously a natural state of affairs. After all, coordinate details usually cannot be consciously quantified with precision, and even if they could be, taking them into account in ordinary linguistic communication would require an astronomical number of locative morphemes to express all of the possible spatial relations – a situation that would be computationally unmanageable and pragmatically otiose, not to mention utterly impossible in languages that completely lack the lexical and grammatical resources for complex counting (Gordon, 2004). But even though almost all spatial discourse is restricted to schematic, coarse-grained distinctions, there is nevertheless a vast range of coding possibilities, and languages vary tremendously in how they carve up this multidimensional conceptual domain, while still conforming to certain overarching tendencies. In this chapter I have attempted to describe this rich field of semantic diversity from the perspective of cognitive neuroscience. I have also suggested ways in which typological data can help guide future research on how linguistically encoded categorical spatial relations are implemented in the brain, and on how they interact with perceptual and cognitive representations of space. The upshot of the chapter is captured by the following question: Would research on the neural substrates of spatial representation be substantially different if the dominant language in the world were, say, Tzeltal instead of English?

Notes

1 Portions of this chapter are reproduced from Kemmerer (2006).

2 Here and in what follows, the family and geographical area of languages that may not be familiar to the reader are provided in parentheses.

3 Basque (Isolate, Europe), Dutch (Indo-European, Europe), Ewe (Niger-Congo, West Africa), Lao (Tai-Kadai, Southeast Asia), Lavukaleve (Isolate, Solomon Islands), Tiriyó (Cariban, South America), Trumai (Isolate, South America), Yélî Dnye (Isolate, Papua New Guinea), Yukatek (Mayan, Mexico).

4 The search domain is always restricted, however, to the region directly adjacent to the designated part of G, because when F is separated from G by a larger distance (even a few inches, in most cases), a different set of locative terms is applied – specifically, terms for cardinal directions, as described below in the subsection on the absolute frame of reference.

5 Since this chapter went to press, the following new studies have come to my attention: Wu et al. (2007), Chatterjee (2008), Amorapanth et al. (in press).

6 In a related study, Carlson et al. (2002) measured event-related brain potentials (ERPs) while subjects judged the appropriateness of *above* for describing the location of a dot relative to a watering can in a series of pictures in which the intrinsic and absolute frames of reference were systematically manipulated.

7 See Coventry et al. (2008) for a recent psycholinguistic study that investigated this topic. See also Bonfiglioli et al. (2009) for another perspective.

8 An important caveat, however, is that none of the tests was specifically designed to distinguish between impaired processing of categorical spatial relations and impaired processing of coordinate spatial relations.

References

Amorapanth, P., Widick, P. and Chatterjee, A. (in press). The neural basis for spatial relations. *Journal of Cognitive Neuroscience*.

Amunts, K., Weiss, P.H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J.M., Marshall, J.M., Shah, N.J., Fink, G.R. and Zilles, K. (2004) Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space – the roles of Bodmann areas 44 and 45. *Neuroimage* 22: 42–56.

Anderson, S. and Keenan, E. (1985) Deixis. In T. Shopen (ed.) *Language typology and syntactic description Vol.* 3, *Grammatical categories and the lexicon* 259–307. Cambridge, UK: Cambridge University Press.

Baciu, M., Koenig, O., Vernier, M.-P., Bedoin, N., Rubin, C. and Segebarth, C. (1999) Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *NeuroReport* 10: 1373–1378.

Benton, A.L. and Tranel, D. (1993) Visuoperceptual, visuospatial, and visuoconstructive disorders. In K.M. Heilman and E. Valenstein (eds) *Clinical neuropsychology* 165–213. (Third edition.) New York: Oxford University Press.

Berti, A. and Rizzolatti, G. (2002) Coding near and far space. In H.-O. Karnath, D. Milner and G. Vallar (eds) *The cognitive and neural bases of spatial neglect* 119–30. Oxford, UK: Oxford University Press.

Bonda, E., Petrides, M., Frey, S. and Evans, A. (1995) Neural correlates of mental transformations of the body-in-space. *Proceedings of the National Academy of Sciences USA* 92: 11180–11184.

Bonfiglioni, C., Finocchiaro, C., Gesierich, B., Rositani, F. and Vescovi, M. (2009) A kinematic approach to the conceptual representations of *this* and *that*. *Cognition* 111: 270–274.

Bowerman, M. and Choi, S. (2003) Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner and S. Goldin-Meadow (eds) *Language in mind* 387–428. Cambridge, MA: MIT Press.

Brown, P. (1994) The INs and ONs of Tzeltal locative expressions. *Linguistics* 32: 743–790.

Brown, P. and Levinson, S.C. (1993) 'Uphill' and 'downhill' in Tzeltal. *Journal of Linguistic Anthropology* 3: 46–74.

Brown, P. and Levinson, S.C. (forthcoming) *Tilted worlds.* Cambridge: Cambridge University Press.

Burenhult, N. (2008) Spatial coordinate systems in demonstrative meaning. *Linguistic Typology* 12: 99–142.

Carlson, L. and Van Der Zee, E. (eds) (2005) *Functional features in language and space*: *Insights from perception, categorization, and development.* Oxford, UK: Oxford University Press.

Carlson, L., West, R., Taylor, H.A. and Herndon, R.W. (2002) Neural correlates of spatial term use. *Journal of Experimental Psychology*: *Human Perception and Performance* 28: 1391–1408.

Carlson-Radvansky, L.A. and Irwin, D.A. (1993) Frames of reference in vision and language: Where is above? *Cognition* 46: 223–244.

Casad, E. and Langacker, R.W. (1985) 'Inside and 'outside' in Cora grammar. *International Journal of American Linguistics* 51: 247–281.

Chatterjee, A. (2008) The neural organization of spatial thought and language. *Seminars in Speech and Language* 29: 226–238.

Coventry, K. R., Prat-Sala, M. and Richards, L. (2001) The interplay between geometry and function in the comprehension of *over, under, above,* and *below. Journal of Memory and Language* 44: 376–398.

Coventry, K.R. and Garrod, S.C. (2004) *Saying, seeing, and acting*: *The psychological semantics of spatial prepositions.* New York: Psychology Press.

Coventry, K.R., Valdés, B., Castillo, A. and Guijarro-Fuentes, P. (2008) Language within your reach: Near-far perceptual space and spatial demonstratives. *Cognition* 108: 889–895.

Damasio, H., Grabowski, T.J., Tranel, D., Ponto, L.L.B., Hichwa, R.D. and Damasio, A.R. (2001) Neural correlates of naming actions and of naming spatial relations. *NeuroImage* 13: 1053–1064.

Davidoff, J. and Warrington, E.K. (2001) A particular difficulty in discriminating between mirror images. *Neuropsychologia* 39: 1022–1036.

Davidson, M. (1999) Southern Wakashan locative suffixes: A challenge to proposed universals of closed-class spatial meaning. Paper presented at the Annual Meeting of the International Cognitive Linguistics Association, Stockholm, Sweden.

Devlin, J.T., Matthews, P.M. and Rushworth, M.F.S. (2003) Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience* 15: 71–84.

Diessel, H. (1999) *Demonstratives.* Amsterdam: John Benjamins.

Diessel, H. (2005) Distance contrasts in demonstratives. In M. Haspelmath, M.S. Dryer, D. Gil and B. Comrie (eds) *The world atlas of language structures* 170–173. Oxford: Oxford University Press.

Diessel, H. (2006) Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17: 463–490.

Dixon, R.M.W. (2003) Demonstratives: A cross-linguistic typology. *Studies in Language* 27: 62–112.

Dunn, M., Meira, S. and Wilkins, D. (eds) (forthcoming) *Demonstratives in cross-linguistic perspective.* Cambridge: Cambridge University Press.

Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T.A., Isham, E.A., Newman, E.L. and Fried, I. (2003) Cellular networks underlying human spatial navigation. *Nature* 425: 184–187.

Emmorey, K. (2003) *Perspectives on classifier constructions in sign languages.* Mahwah, NJ: Lawrence Erlbaum Associates.

Emmorey, K., Damasio, H., McCullough, S., Grabowski, T., Ponto, L.L.B., Hichwa, R.D. and Bellugi, U. (2002) Neural systems underlying spatial language in American Sign Language. *NeuroImage* 17: 812–824.

Enfield, N.J. (2003) Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis. *Language* 79: 82–117.

Fillmore, C.F. (1997) *Lectures on deixis.* Stanford: CSLI Publications.

Friederici, A. and Levelt, W.J.M. (1990) Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception and Psychophysics* 47: 253–266.

Gentner, D. and Goldin-Meadow, S. (eds) (2003) *Language in mind.* Cambridge, MA: MIT Press.

Gerstmann, J. (1957) Some notes on the Gerstmann syndrome. *Neurology* 7: 866–869.

Gilbert, A.L., Regier, T., Kay, P. and Ivry, R.B. (2006) Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences* 103: 489–904.

Gordon, P. (2004) Numerical cognition without words: Evidence from Amazonia. *Science* 306: 496–499.

Graziano, M.S.A. and Aflalo, T.N. (2007) Rethinking cortical organization: Moving away from discrete areas arranged in hierarchies. *The Neuroscientist* 13: 138–147.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B. and Moser, E.I. (2005) Microstructure of a spatial map in the enthorinal cortex. *Nature* 436: 801–806.

Hanks, W.F. (2005) Explorations in the deictic field. *Current Anthropology* 46: 191–220.

Hartley, T., Maguire, E.A., Spiers, H.J. and Burgess, N. (2003) The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron* 37: 877–888.

Haun, D.B.M., Call, J., Janzen, G. and Levinson, S.C. (2006a) Evolutionary psychology of spatial representations in the Hominidae. *Current Biology* 16: 1736–1740.

Haun, D.B.M., Rapold, C.J., Call, J., Janzen, G. and Levinson, S.C. (2006b) Cognitive cladistics and cultural override in Hominid spatial cognition. *Proceedings of the National Academy of Sciences* 103: 17568–17573.

Heine, B. (1997) *Cognitive foundations of grammar.* Oxford, UK: Oxford University Press.

Hespos, S.J. and Spelke, E.S. (2004) Conceptual precursors to language. *Nature* 430: 453–456.

Heyman, K. (2006) The map in the brain: Grid cells may help us navigate. *Science* 312: 680–681.

Hillis, A.E. (2006) Neurobiology of unilateral spatial neglect. *The Neuroscientist* 12: 153–163.

Jager, G. and Postma, A. (2003) On the hemispheric specialization for categorical and coordinate spatial relations: A review of the current evidence. *Neuropsychologia* 41: 504–515.

Janzen, G. and van Turennout, M. (2004) Selective neural representation of objects relevant for navigation. *Nature Neuroscience* 7: 673–677.

Jeffery, K.J. and Burgess, N. (2006) A metric for the cognitive map: Found at last? *Trends in Cognitive Sciences* 10: 1–3.

Kalaska, J.F., Cisek, P., Gosselin-Kessiby, N. (2003) Mechanisms of selection and guidance of reaching movements in the parietal lobe. In A.M. Siegel, R.A. Andersen, H.-J. Freund and D.D. Spencer (eds) *The parietal lobes* 97–120. Philadelphia, PA: Lippincott Williams & Wilkins.

Kemmerer, D. (1999) 'Near' and 'far' in language and perception. *Cognition* 73: 35–63.

Kemmerer, D. (2005) The spatial and temporal meanings of English prepositions can be independently impaired. *Neuropsychologia* 43: 797–806.

Kemmerer, D. (2006) The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia* 44: 1607–1621.

Kemmerer, D. (in press) How words capture visual experience: The perspective from cognitive neuroscience. In B. Malt and P. Wolff (eds) *Words and the world*: *How words capture human experience.* Oxford, UK: Oxford University Press.

Kemmerer, D. (forthcoming) Visual and motor features of the meanings of action verbs: A cognitive neuroscience perspective. In R. de Almeida and C. Manouilidou (eds) *Verb concepts*: *Cognitive science perspectives on verb representation and processing*. Oxford, UK: Oxford University Press.

Kemmerer, D. and Tranel, D. (2000) A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology* 17: 393–414.

Kemmerer, D. and Tranel, D. (2003) A double dissociation between the meanings of action verbs and locative prepositions. *Neurocase* 9: 421–435.

Kemmerer, D. and Tranel, D. (2008) Searching for the elusive neural substrates of body part terms: A neuropsychological study. *Cognitive Neuropsychology* 25: 601–629.

Kosslyn, S. (1987) Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review* 94: 148–175.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K. and Bandettini, P.A. (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60: 1126–1141.

Kuhl, P.K. (2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5: 831–843.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N. and Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606–608.

Laeng, B. (1994) Lateralization of categorical and coordinate spatial functions: A study of unilateral stroke patients. *Journal of Cognitive Neuroscience* 6: 189–203.

Laeng, B., Chabris, C.F. and Kosslyn, S.M. (2003) Asymmetries in encoding spatial relations. In K. Hugdahl and R.J. Davidson (eds) *The asymmetrical brain* 303–339. Cambridge, MA: MIT Press.

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

Leutgeb, J.K., Leutgeb, S.L., Moser, M.-B. and Moser, E.I. (2007) Pattern separation in the dendate gyrus and CA3 of the hippocampus. *Science* 315: 961–966.

Levinson, S.C. (1983) *Pragmatics.* Cambridge, UK: Cambridge University Press.

Levinson, S.C. (1994) Vision, shape, and linguistic description: Tzeltal body-part terminology and object description. *Linguistics* 32: 791–856.

Levinson, S.C. (2003) *Space in language and cognition.* Cambridge, UK: Cambridge University Press.

Levinson, S.C., Kita, S., Haun, D.B.M. and Rasch, B.H. (2002) Returning the tables: Language affects spatial reasoning. *Cognition* 84: 155–188.

Levinson, S.C. and Meira, S. (2003) 'Natural concepts' in the spatial topological domain – adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79: 485–516.

Levinson, S.C. and Wilkins, D. (eds) (2006) *Grammars of space.* Cambridge, UK: Cambridge University Press.

Li, P. and Gleitman, L. (2002) Turning the tables: Language and spatial reasoning. *Cognition* 83: 265–94.

Longo, M.R. and Lourenco, S.F. (2006) On the nature of near space: Effects of tool use and the transition to far space. *Neuropsychologia* 44: 977–981.

Luschi, P., Benhamoou, S., Girard, C., Ciccione, S., Roos, D., Sudre, J. and Benvenuti, S. (2007) Marine turtles use geomagnetic cues during open-sea homing. *Current Biology* 17: 126-133.

Majid, A. (2002) Frames of reference and language concepts. *Trends in Cognitive Sciences* 6: 503–504.

Majid, A., Bowerman, M., Kita, S., Haun, D.B.M. and Levinson, S.C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8: 108–114.

Majid, A., Enfield, N. and van Staden, M. (eds) (2006) *Parts of the body*: *Crosslinguistic categorization.* Special issue of *Language Sciences* 28: 137–360.

MacSweeney, M., Woll, B., Campbell, R., Calvert, G.A., McGuire, P.K., David, A.S., Simmons, A. and Brammer, M.J. (2002) Neural correlates of British Sign Language comprehension: Spatial processing demands of topographic language. *Journal of Cognitive Neuroscience* 14: 1064–1075.

Makin, T.R., Holmes, N.P. and Zohary, E. (2007) Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience* 27: 731–740.

Mayer, E., Martory, M.-D., Pegna, A.J., Landis, T., Delavelle, J. and Annoni, J.-M. (1999) A pure case of Gerstmann syndrome with a subangular lesion. *Brain* 122: 1107–1120.

Mazzoni, M., Pardossi, L., Cantini, R., Giorgetti, V. and Arena, R. (1990) Gerstmann syndrome: A case report. *Cortex* 26: 459–467.

McCloskey, M. (2009) *Visual reflections*: *A perceptual deficit and its implications.* Oxford, UK: Oxford University Press.

McDonough, L., Choi, S. and Mandler, J.M. (2003) Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology* 46: 229–259.

Milner, A.D. and Goodale, M.A. (2006) *The visual brain in action.* (Second edition.) Oxford, UK: Oxford University Press.

Morris, H.H., Luders, H., Lesser, R.P., Dinner, D.S. and Hahn, H. (1984) Transient neuropsychological abnormalities (including Gerstmann's syndrome) during cortical stimulation. *Archives of Neurology* 34: 877–883.

Munnich, E., Landau, B. and Dosher, B.A. (2001) Spatial language and spatial representation: A crosslinguistic comparison. *Cognition* 81: 171–207.

Noordzij, M.L., Neggers, S.F.W., Postma, A. and Ramsey, N. (2008) Neural correlates of locative prepositions. *Neuropsychologia* 46: 1576–1580.

Palmeri, T.J., Wong, A.C.N. and Gauthier, I. (2004) Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences* 8: 378–386.

Papafragou, A. (2007) Space and the language-cognition interface. In P. Carruthers, S. Laurence and S. Stich (eds) *The innate mind*: *Foundations and the future.* Oxford, UK: Oxford University Press.

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

Peelen, M.V. and Downing, P.E. (2007) The neural basis of visual body perception. *Nature Reviews Neuroscience* 44: 1515–1518.

Postma, A. and Laeng, B. (2006) New insights in categorical and coordinate processing of spatial relations. *Neuropsychologia* 8: 636–648.

Previc, F.H. (1998) The neuropsychology of 3-D space. *Psychological Bulletin* 124: 123–164.

Priftis, K., Rusconi, E., Umilta, C. and Zorzi, M. (2003) Pure agnosia for mirror stimuli after right inferior parietal lesion. *Brain* 126: 908–919.

Ritz, T., Thalu, P., Phillips, J.B., Wiltschko, R. and Wiltschko, W. (2004) Resonance effects indicate a radical-pair mechanism for avian magnetic compass. *Nature* 429: 177–180.

Robertson, L. (2004) *Space, objects, minds, and brains.* New York: Psychology Press.

Roeltgen, D.P., Sevush, S. and Heilman, K.M. (1983) Pure Gerstmann's syndrome from a focal lesion. *Archives of Neurology* 40: 46–47.

Sakata, H. (2003) The role of the parietal cortex in grasping. In A.M. Siegel, R.A. Andersen, H.-J. Freund and D.D. Spencer (eds) *The parietal lobes* 121–140. Philadelphia, PA: Lippincott Williams & Wilkins.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.B. and Moser, E.I. (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312: 758–762.

Slobin, D.I. (1996) From 'thought and language' to 'thinking for speaking'. In J.J. Gumperz and S.C. Levinson (eds) *Rethinking linguistic relativity* 70–96. Cambridge, UK: Cambridge University Press.

Smith, L.B., Jones, S.S., Landau, B., Gershkoff-Stowe, L. and Samuelson, L. (2002) Object name learning provides on-the-job training for attention. *Psychological Science* 13: 13–19.

Spiers, H.J. and Maguire, E.A. (2006) Thoughts, behaviour, and brain dynamics during navigation in the real world. *NeuroImage* 31: 1826–1840.

Svorou, S. (1994) *The grammar of space.* Amsterdam: John Benjamins.

Talmy, L. (1983) How language structures space. In H. Pick and L. Acredolo (eds) *Spatial orientation*: *Theory, research, and application* 225–282. New York: Plenum Press.

Thompson-Schill, S.L., D'Esposito, M. and Kan, I.P. (1999) Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron* 23: 513–522.

Tranel, D. (1996) The Iowa-Benton school of neuropsychological assessment. In I. Grant and K.M. Adams (eds) *Neuropsychological assessment of neuropsychiatric disorders* 81–101. (Second edition.) New York: Oxford University Press.

Tranel, D. and Kemmerer, D. (2004) Neuroanatomical correlates of locative prepositions. *Cognitive Neuropsychology* 21: 719–749.

Turnbull, O.H., Beschin, N. and Della Sala, S. (1997) Agnosia for object orientation: Implications for theories of object recognition. *Neuropsychologia* 35: 153–163.

Turnbull, O.H. and McCarthy, R.A. (1996) Failure to discriminate between mirror-image objects: A case of viewpoint independent object recognition? *Neurocase* 2: 63–72.

Tyler, A. and Evans, V. (2003) *The semantics of English prepositions*: *Spatial scenes, embodied experience and cognition.* Cambridge: Cambridge University Press.

Ulltan, R. (1978) Some general characteristics of interrogative systems. In J. Greenberg (ed.) *Universals of human language Vol. IV* 211–248. Stanford, CA: Stanford University Press.

Van Driem, G. (1987) *A grammar of Limbu.* Berlin: Walter de Gruyter.

Varney, N.R. (1984) Gerstmann syndrome without aphasia. *Brain and Cognition* 3: 1–9.

Wagner, A.D., Pare-Blagoev, E.J., Clark, J. and Poldrack, R.A. (2001) Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31: 329–338.

Werker, J.F. and Tees, R.C. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63.

Wierzbicka, A. (2007) Bodies and their parts: An NSM approach to semantic typology. *Language Sciences* 29: 14–65.

Wu, D.H., Waller, S. and Chatterjee, A. (2007) The functional neuroanatomy of thematic role and locative relational knowledge. *Journal of Cognitive Neuroscience* 19: 1542–1555.

Zacks, J., Rypma, B., Gabrieli, J.D., Tversky, B. and Glover, G.H. (1999) Imagined transformations of human bodies: An fMRI investigation. *Neuropsychologia* 37: 1029–1040.

Final proofs with index 24 February 2010

Part IV

Theoretical approaches to spatial representation in language

7 Genesis of spatial terms

Claude Vandeloise

A parallelism is often established between the production of a language by a culture (phylogeny) and its reproduction by children (ontogeny). The basic spatial words in different languages will be used in this article in order to investigate the similarities and the discrepancies between the two processes. Concerning the creation of spatial words, section 1 establishes a contrast between what I call *external lexical formation*, in which a word is associated to an extra-linguistic concept, and *internal lexical formation*, that proceeds by *division* or *union* of established lexical categories. In section 2, I will discuss a hierarchy in the formation of spatial terms in languages of the world (Levinson and Meira 2003) inspired by an implicational scale for the creation of basic color terms proposed by Berlin and Kay (1968). MacLaury (1993) motivates this development by creating a hierarchy involving a process of internal lexical formation by division. I will compare these hypotheses to another hierarchy proposed in Vandeloise (2003, 2005). This hierarchy establishes a basic contrast between the relation of *localization*, conveyed in French by the preposition *à,* and the dynamic relations of *control*, expressed by *in* and *on*.

Three modes of lexicon development are investigated in section 3. Whereas the creation of basic color terms may go from the most general to the most specific, as illustrated by MacLaury, the creation of words often evolves in the reverse direction, from the application of a word to very specific situations to its extension to more general uses. In contrast to the former mode of creation that operates by division, the latter mode of internal lexical formation proceeds by union. If external lexical creation anchors a word in the middle of a hierarchy of concepts, both processes can occur to create supercategories and subcategories. In contrast to the linguistic community that builds its language from scratch, infants pick their first spatial words inside a complete and well-structured language. In section 4, I will attempt to explain how the different levels of abstraction of spatial words can influence the acquisition of spatial words.

1 External and internal lexical formation

According to one of the main dogmas of structuralism, the meanings of words emerge negatively, from their differences with other words in the language. These differential meanings are called *values* (Saussure 1916). I will come back to them when I speak of internal lexical formation. This conception of meaning, however, poses an obvious logical problem once one considers the production of language and the first words created in the lexicon. This problem has not been urgent as long as language creation was considered a taboo subject, unworthy of linguists' attention. Once this interdiction is transgressed, though, one must admit that, according to the differential hypothesis, the first words can only be created by pairs (*x, y*), with *x* determining the value of *y* and vice versa. This may make sense for pairs like *here* and *there* or *yes* and *no*. But if one admits that among the first words appear also terms for actions (like *eat*) or names for persons (like *Peter*), complementary words designating any action that is not eating, or any human who is not Peter, are more difficult to conceive. The meaning of these words cannot emerge from differences in a system but may only be explained by the extra-linguistic stimulations that make these terms convenient to ensure the good functioning of the society in which they emerge. This is what I call external lexical formation. It occurs when the members of a society share a common interest in an aspect of their environment or of their social life; when they are able to recognize this aspect in a sufficiently similar way; and when they associate a term to this aspect of their lives.

The existence of external lexical formation, mainly based on *similarities* between the occurrences in the world designated by a word, does not preclude a very important role for internal lexical formation. In this case, a term is applied to aspects of environment or social life because *differences* with aspects of the world designated by the available words in the lexicon begin to appear pertinent for the ease of communication. Linguists have been much more interested in internal lexical formation. Its functioning is much better documented than the development of external lexical formation. The domain of colors is a perfect field to observe this type of formation. The work of Berlin and Kay (1968), devoted to basic color terms and to their hierarchical appearance in the development of language, will be essential for this article. This book was equally an important source of inspiration for the typology proposed by Levinson and Meira (2003) that will be discussed in the next section. According to these authors' interpretation of Berlin and Kay's implication scale, color terms appear in the following order in the languages of the world:

White +Black→ Red→ Green or Yellow→ Yellow or Green→ Blue→ Brown→ Purple
Pink
Orange
Grey

This means that a language that possesses a color term on the right of the scale necessarily includes all others to the left of this term.

The formation of basic color terms in this implicational scale cannot be explained by internal lexical formation only. At the beginning of the scale, an internal lexical formation of *white* and *black* might be justified by the contrast between *day* and *night* (Wierzbicka, 1990). In this case, however, it would not be a genuine color contrast. Taken together, *white* and *black* might be opposed to a word meaning *colorful* but certainly not to *red* alone as proposed in the above scale. At the end of the implication scale, *brown* is also very unlikely to be created from a category including *brown* and *blue* by internal lexical formation. Some amount of external lexical formation, then, must be involved in the creation of the first and the last 'basic color terms'.

The creation of basic color terms has been carefully observed by MacLaury (1993), who compares the evolution of two Mayan languages, Tzeltal of Tenejapa and Tzotzil of Novenchuuc, from a system of two color terms to a system of six color terms. He proposes an interpretation of Berlin and Kay's implicational scale that is more compatible with internal lexical formation (see chart 1).

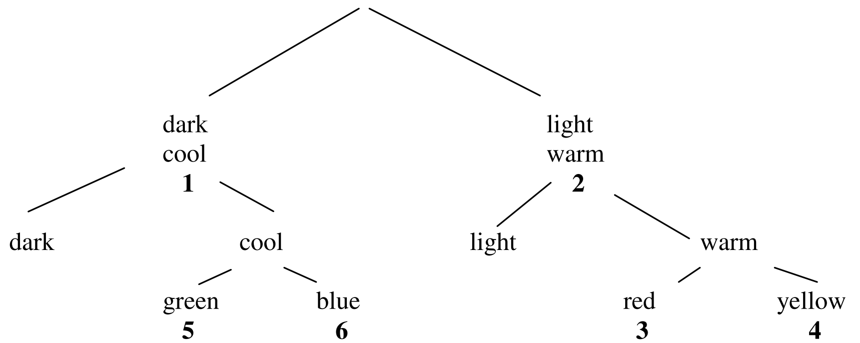


Chart 1. MacLaury's implicational scale

In chart 1, the places of *black* and *white*, the two first terms of the implicational scale, are occupied by the category of dark or cool colors and by the category of light or warm colors, respectively. *Red* and *yellow*, third and fourth in the implicational scale, are the result of the split of warm colors whereas *green* and *blue*, fifth and sixth in the implicational scale, are the result of the split of cool colors.[1] What is the destiny of *dark* and *light* at the second level remains an open question: do these categories remain without linguistic representation, or are they conveyed by words equivalent to *black* and *white*?

The category of cool colors that gathers green and blue in chart 1 is often called 'grue'. The split of the category of cool colors into *green* and *blue* provides an exemplary case of internal lexical formation. At the beginning, suppose that *green* is indifferently used for the green or blue tokens of the 'grue' category. Inside this general category appears a new word *blue* that is used by innovative speakers for blue objects only. At a first stage, *green* can still be used for blue objects, in such a way that *blue* may be considered as a hyponym of *green*. With evolution, and the disappearance of more conservative speakers who prefer *green* to *blue*, a second stage appears at which *green* can no longer be applied to blue objects and restricts itself to 'grue' objects that are not blue, i.e. to green objects. In this way, at the third and final stage, the connection between *green* and *blue* is severed: *green* applies only to green objects and *blue* to blue objects.

The case of the 'grue' category is a perfect example of internal lexical creation, because the similarities between green and blue makes plausible the existence of a category including the two colors. In contrast, at the origin of the implicational scale of

Berlin and Kay, it is difficult to imagine a category gathering black and white from which these terms emerge. *White* and *black*, then, are examples of external lexical formation and *green* and *blue* are examples of internal lexical formation, as illustrated in Figure 1. External and internal lexical formation will prove useful in section 2 for the comparison of English spatial words *in* and *on* with Spanish *en* and *sobre*.

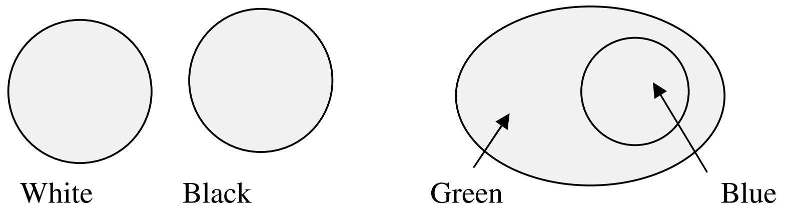


Figure 1. Internal vs. external lexical creation

2 Creation of spatial words

Berlin and Kay (1968) provide an implicational scale according to which basic color terms develop in the languages of the world. This may provide a first insight in lexical formation. In this article, I will be concerned with the creation of spatial terms. I will present the analysis of Levinson and Meira (2003) before arguing for an alternative solution based on preceding articles (Vandeloise 2003, 2005). Like Berlin and Kay (1968), Levinson and Meira use a sample of genetically unrelated languages. Informants in each language were asked to ascribe an adposition in their language to a booklet of 71 line-drawings known under the name of 'topological relations picture series'. As in the case of the attribution of basic color terms, the choices tended to cluster and were not randomly distributed as they would be if there were no crosslinguistic generalizations. The five main clusters are labeled IN, NEAR/UNDER, ON/OVER, ATTACHMENT and ON-TOP. On the basis of these data, Levinson and Meira propose the following implicational scale for spatial terms.

AT < IN < ON < OVER< ON TOP < ATTACHED < INSIDE < SPIKED
UNDER NEAR HANGING
DISTRIBUTED OVER

They elaborate this implicational scale to show how different languages can develop spatial terms in different ways. I modify the presentation of their analysis (Figure 18, p. 512), in order to make the comparison with my solution easier.

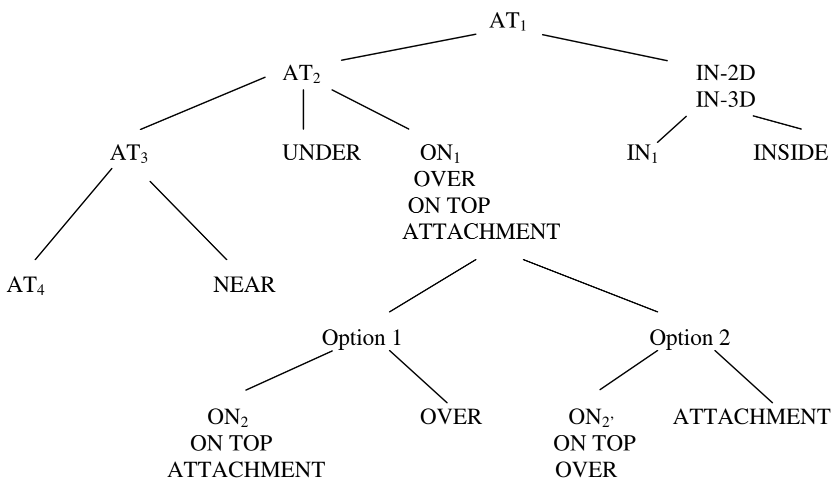


Chart 2. Levinson and Meira's implicational scale

AT$_1$ is a unique spatial notion that covers all the spatial relations and corresponds to the adposition *ta* in a language like Tzeltal or *di* in Indonesian (Feist 2004). AT$_2$, AT$_3$ and AT$_4$ are more and more specific notions. Thus, AT$_2$ covers all spatial relationships with the exception of those conveyed by IN-2D and IN-3D and AT$_4$ is a residue that excludes all the preceding notions, including NEAR. These processes correspond to internal lexical formation by division. Vertically aligned notions, such as IN-3D and IN-2D, as well as ON$_1$, OVER, ON-TOP and ATTACHMENT, correspond to composite notions. Levinson and Meira split IN-2D and IN-3D because they attribute two foci to the category IN, one focus specifying containment in a three-dimensional container and the other inclusion in a two-dimensional plane. I will come back to this decision later in this section. Indices attached to ON$_1$ and ON$_2$ do not appear in Levinson and Meira's chart but they will make the exposition easier. At the last level of specification of chart 2, two options are offered for the decomposition of the complex concept ON$_1$, OVER, ON TOP and ATTACHMENT.

Levinson and Meira use capitals AT, IN and so forth to represent the 'central meanings of the relevant sort' (footnote 2, p. 486) associated to the basic topological notions conveyed by *at, in* and so forth.[2] ATTACHMENT is an exception to this convention. The authors are obliged to use this notional term instead of a preposition because English has no specific adposition to convey attachment. The use of capitals may be a handy way of introducing the prototypes of spatial relations but it raises some questions. First, why is AT chosen to represent the most general category in chart 2? The preposition *at* appears nowhere in the data or in the article and it is certainly not a good example of an inclusive spatial preposition. As we will see later, the Old English preposition *æt* might fit this role better. Second, there are discrepancies between the data coming from the experiments summarized in the map proposed by Levinson and Meira (p. 505) and chart 2. Indeed, IN represents a coherent cluster in the map but it splits in IN-2D and IN-3D in the chart. As a matter of fact, this is so because all the members in the IN-cluster correspond to IN-3D. Therefore, one may doubt whether IN-2D represents a 'central meaning of the relevant sort'. The reason why IN-3D and IN-2D are grouped in the chart is not because they are notionally related but because an identical preposition is assigned to them in English. On the other hand, NEAR/UNDER, as well as ON/OVER, corresponds to one cluster in the map of data but NEAR and UNDER, as well as ON and OVER, are disjointed in the chart. In my alternative, instead of AT, IN (with or without contact), I will use explicit notions like LOC(ALIZATION), CONTROL and so forth. For each notion, I will provide an example of a preposition attached to this notion in a language of the world.

Levinson and Meira are uniquely concerned with basic topological relationships between the located *target* and the *landmark* that locates it. For example, among the clusters in the map of data, ON/OVER is characterized by superadjacency (with or without contact), UNDER by subadjacency (with or without contact) and NEAR by proximity. Contiguity and coincidence are further topological notions present in the analysis. In contrast, the cluster IN corresponds to full containment (p. 508). Containment is certainly not a topological notion but the distinction between containment and the topological notion of inclusion is often blurred in the literature and many scholars appear to use them indifferently. In chart 2, Levinson and Meira make a distinction between IN-3D (a notion close to CONTAINMENT) and IN-2D (close to INCLUSION). In this way, they introduce a further topological notion in their analysis. The authors mention that a reviewer of their article 'questions to what extent 'attachment' (and indeed other notions like 'containment' and 'support') are really spatial as opposed to mechanical in conception' (footnote 9, p. 487). The authors admit that some doubt is in order. The alternative proposed in chart 3 reinforces the contrast between topological basic categories and puts the role of force and energy to the forefront. Therefore, the first dichotomy established in chart 3 distinguishes between LOC (topology[3]) and CONTROL (dynamics). The discrepancies in the linguistic representation of these notions in the languages of the world create a problem for the typology proposed in chart 2. Indeed, the Spanish preposition *en* conveys both ON and IN notions. But in chart 2, ON and IN do not have a common direct hyperonym. If these notions belong to different branches of the structure, one may wonder why so many languages, like Spanish and Modern Greek, have a common adposition to designate both notions. The common status of IN and ON as opposed to AT constitutes the main discrepancy between my analysis and that of Levinson's and Meira's. Another famous example concerning control is Korean in which two verbs correspond to the English preposition *in*: the verb *kkita* that conveys tight fit, as opposed to loose containment conveyed by the verb *nehta*.[4] Chart 3 makes this distinction possible.

Levinson and Meira exclude very pervasive spatial relations such as projective notions IN FRONT and BEHIND from their inquiry because 'projective concepts belong to a different conceptual subdomain, where coordinate systems or frames of reference are necessary' (p. 488). It is true that projective prepositions do not correspond to topological notions, but control prepositions like *in* and *on* do not either.

Projective spatial prepositions are basic spatial prepositions that will be introduced in chart 3. On the other hand, like Levinson and Meira, I will limit myself to static spatial relations. However, chart 4 describing the evolution of the Old English preposition *œt* will show how kinetic prepositions such as *from* and *to* might be incorporated in the genesis of spatial terms. UNDER and OVER will also be excluded from basic spatial categories in chart 3 for reasons I will give below. These two categories have a peculiar status in the analysis of Levinson and Meira. First, UNDER is the only basic category to appear simultaneously with another category (ON_1) in the partition of AT_2. No explanation is provided for this exception.[5] Second, OVER appears first as a component of the composite concept ON_1/ON TOP/OVER/ATTACHMENT. Depending on the languages, this composite category can split in two different ways. The former option, with OVER excluded, appears in English whereas Levinson and Meira attribute the latter option to Yucatec and Ewe. In their chart, they put ON_1, OVER, ON-TOP and ATTTACHMENT at the same level. The relationship between ON_1 and ON-TOP is unclear. The only definition provided for the latter notion is 'location above the eye-level' (p. 512), a definition contradicted by the utilization of ON-TOP for a picture representing a table covered by a tablecloth.[6] As a matter of fact, except for this example, the pictures to which ON-TOP is ascribed correspond to the prototypical uses of *on* in English. Therefore, instead of being at the same level as ON-TOP, ATTACHMENT and OVER, ON_1 might be considered as a more general notion, including these three categories. One may furthermore cast in doubt the usefulness of the category ON-TOP. At any rate, in contrast to OVER and ATTACHMENT, ON-TOP never dissociates itself from ON_1 at the last level of specification.

In chart 3, the equivalent of AT_1 is called RELATION IN SPACE. These relations imply *accessibility* in space between two *material entities*; between a material and a *spatial entity*; or between two spatial entities. A spatial entity may be a place occupied by a material entity or a portion of space that material entities might occupy. Linguistic communities attribute names to *geographic* spatial entities. In the case of material entities, accessibility is guaranteed by *contact* or *proximity.*[7] When a spatial landmark is involved in the relationship, there is *coincidence* or proximity of the target with the landmark. This coincidence is often partial since the landmark is usually larger than the target. Coincidence between two material entities is impossible. In many respects, proximity appears to be the most general ingredient of a relation in space. As a matter of fact, if contact and coincidence are considered as limit cases of proximity, this notion might be chosen to characterize relations in space at the most general level. Therefore, the occurrence in chart 2 of NEAR (representing proximity) at the same level as the most specific notion AT_4 is surprising. How can one and the same concept appear together at the most general and at the most specific level? Because it is associated to the primitive notion of proximity, *near to* appears deceptively as a basic expression. However, far from being basic, the syntax of *close to* or *near (to)* in English or of *près de* in French demonstrate that these locutions, though related to the primitive concept of proximity, are complex notions. *Near* in English may be an adjective as well as a preposition and there are discussion in French (Gunnarson 1986) about whether *près* should be treated as an adverb, an adjective or a preposition. In the genesis I propose in chart 3, projective relationships in the vertical axis and in the horizontal plane will appear instead of NEAR. As a matter of fact, *near (to)* might be considered as a late hyperonym for all the prepositions involving proximity in the horizontal plane.

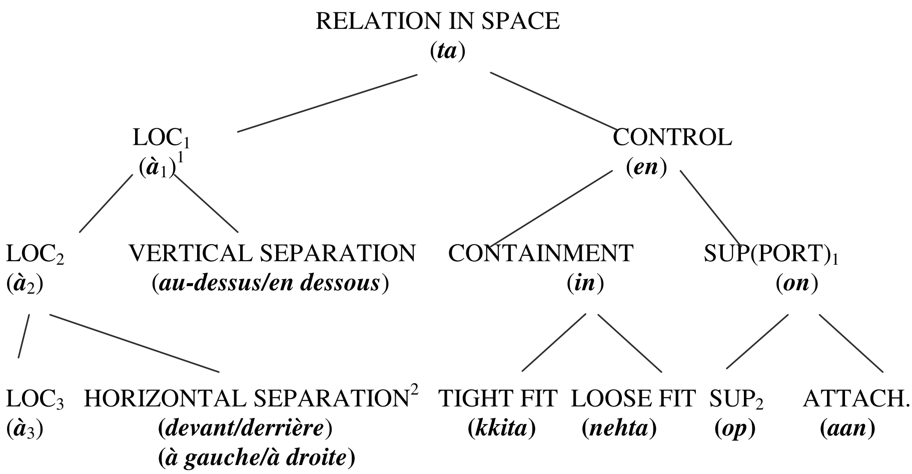
The most important division in chart 3 separates CONTROL (that implies an exchange of energy between the landmark and the target) from a residue of spatial relations, called LOC_1, that do not involve such an exchange of energy.[8] By this division, LOC_1 is deprived of all the relations in space involving contact between two material entities, since – if one forgets magnets and radiations – contact is a necessary condition for control. LOC_2, then, is left with the spatial relationships involving at least one spatial entity on the one hand; and with the relationships between material entities that are not in contact on the other hand. Thus, LOC_2 means that the target (partially) coincides with a spatial landmark; or that it is close to a spatial or material landmark. At the corresponding stage of the development, chart 2 subtracts NEAR from AT_3. If PROXIMITY were similarly subtracted from LOC_2 in chart 3, LOC_3 would be restricted to the spatial relationships containing at least one spatial entity and implying coincidence of the target and the landmark. Reasons to avoid the presence of PROXIMITY at this stage have been evoked in the preceding paragraph. Categories split because a subset of their members attracts more attention than the others or because there is a need for explicitness. In conformity with this principle, IN-3D in chart 2 or CONTROL in chart 3 are relations in space more constrained and prominent than AT_2 or LOC_1, respectively. But why should relationships of proximity be more prominent than relationships of coincidence? Quite the contrary! Coincidence with the landmark locates the target more precisely than proximity which needs specification. This specification is, I believe, the role of projective prepositions. Actually, in the vertical axis, proximity may be too strong a word since *the sun is above the earth* does not involve proximity between the sun and the earth. Separation between the target and the landmark may be sufficient for the use of projective prepositions. I will make a distinction between separation in the vertical axis (VERTICAL SEPARATION) and separation in the horizontal plane (HORIZONTAL SEPARATION). Whereas VERTICAL SEPARATION admits material landmarks (*the lamp is over the table*) as well as spatial landmarks (*the airplane is above Paris*), HORIZONTAL SEPARATION prefers material landmarks and is used with difficulty with spatial landmarks: *?The car is to the left of Paris.* Also, in French, *au-dessus* and *en dessous* maintain a connection with coincidence since, in contrast to *plus haut* and *plus bas,* these prepositions require coincidence of the vertical projection of the target with the landmark. For these reasons, in chart 3, I will first subtract VERTICAL SEPARATION from LOC_2 and then, subtract HORIZONTAL SEPARATION from LOC_3. In internal lexical division, the emergence of a new term makes the use of the old term obsolete. This is true in the vertical axis in the French examples below:

(1) La lampe est au-dessus/*à/?près de la table

(2) La chaise est devant/ *à/ près de la table

(3) Le chat est à gauche/*à/ près de la table

Près de, in contrast, is compatible with the projective horizontal prepositions.



[1] Besides static location, the French preposition *à* can also introduce the goal of the target like the preposition *to*.
[2] Separation along the frontal direction and along the lateral direction might need to be treated separately.

Chart 3. A hierarchy of concepts

Chart 3 can be understood as a hierarchy of concepts going from the most abstract level to the most concrete levels. Languages like Tzeltal (Brown 1994) have only one adposition at the first level – *ta* – that introduces any spatial relations in space and leaves to verbs and nouns the elaboration of these relations. At this abstract level, the only spatial term opposes relations in space to the other grammatical functions, marked by cases like nominative, accusative and so forth. All the notions below LOC_1 are illustrated by prepositions in the same language. I have chosen French because *à* is more clearly related to localization (Vandeloise 1991) than *at* or *in*. Whenever a more specific preposition is added, the extension of the most general preposition diminishes. The process going from $à_1$ to $à_3$ proceeds by subtraction like the evolution from AT_1 to AT_4 in the analysis of Levinson and Meira. LOC_1 splits in LOC_2 and VERTICAL SEPARATION; and LOC_2 in LOC_3 and HORIZONTAL SEPARATION.

The nature of the development going from LOC_1 to the more specific levels in the chart is different from the development of CONTROL. Whereas the former notion evolves by division, the latter develops by specification. In contrast to the development of LOC_1, the development of CONTROL in chart 3 is exemplified by prepositions in different languages. General control is conveyed by the Spanish preposition *en*. The prepositions *in* and *on* in English, and the Korean verbs *kkita* and *nehta* as well as the Dutch prepositions *op* and *aan*, correspond to more and more specific types of control. Whereas in support, conveyed by *on*, the bearer controls the burden in the vertical direction only, containment, conveyed by *in*, requires control in more than one direction. *Kkita* and *nehta* mark tight fit and loose containment of the target in the landmark respectively[9] while *op* and *aan* convey direct support and indirect support, respectively. I propose this hierarchy in Vandeloise (2003) to show that the relativity in the description of space illustrated by Spanish, English, Korean and Dutch is less dramatic than claimed by Bowerman (1996). Spanish *en*, English *in* and Korean *kkita* convey control at different levels of specificity.

In the development of localization, the evolution occurs mainly by internal lexical formation. In the case of control, some languages might overlook the most abstract notions and immediately establish a connection at the level of more specific concepts. Thus, whereas in the development of localization, a specific preposition (like *au-dessus*) reduces the scope of a more general preposition (like *à*), in the case of control, there is no evidence that a general preposition of control existed at an earlier stage of French, even though *dans* and *sur* specify the Spanish preposition *en*. The more specific prepositions *dans* and *sur*, covering approximately the scope of the Spanish preposition *en*, might have appeared simultaneously, or at least independently, by external lexical formation. This is the reason why the examples illustrating the development of control prepositions in chart 3 are taken from different languages. This does not mean that the expression of control never developed in the same way as the expression of localization. I would like to remain neutral on this point.

From AT_1 to AT_4, the extension of the preposition of general localization shrinks each time a more specific preposition appears. The history of languages might support this type of development. Notably, the evolution of the preposition *œt* in Old English illustrates the mode of production by internal formation. Besides the meaning of the present preposition *at*, *œt* could convey (1) the origin of movement; (2) proximity to a living being; and (3) the goal of a movement. It progressively lost these meanings to the profit of the prepositions *with* or *by, from* and *to.* The first shift occurred around 1500, the second in the sixteenth century and the concurrence with *to* lasted until Early Modern English (Lindkvist 1978). This evolution is represented in chart 4.

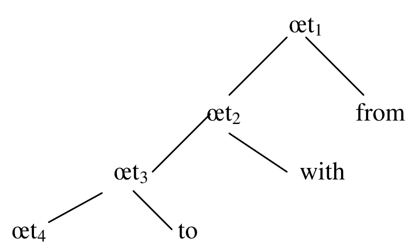


Chart 4. Evolution of AT

Œt$_2$ has all the meanings of *œt*$_1$ with the exception of the origin; *œt*$_3$ has all the meanings of *œt*$_2$ but its landmark cannot be a living being; and *œt*$_4$[10] has all the meanings of *œt*$_3$ with the exception of the goal. In contrast to *œt,* the French preposition *à* does not leave the introduction of the goal of the target to another preposition.

A comparison between the case of *en* and *sobre* in Spanish and between *in* and *on* in English might reveal two different modes of lexical creation. In Spanish, one may use the preposition *en* for an object placed on a table, but if the object is placed on a chest of drawers, s*obre* must be used instead of *en*. Indeed, using the latter preposition would imply a reference to the interior of the drawers, as a preferred option. To avoid the confusion with the objects contained by the drawers, *sobre* must be chosen.

(4) El libro está en la mesa

(5) El libro está sobre la cómoda
El libro está en la cómoda (inside a drawer)[11]

Therefore a need for clarification pushes Spanish to work with two prepositions like English. What is different between the two languages is not that Spanish has only one preposition to describe CONTAINMENT and SUPPORT, but that English does not allow one to use *in* when *on* is adequate. The diagrams in Figure 2 illustrate the distribution of the prepositions *in* and *on* in English and of the prepositions *en* and *sobre* in Spanish.

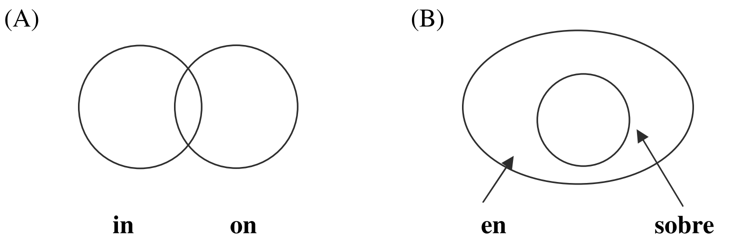


Figure 2. Distribution of prepositions in English and Spanish

To be accurate, the schema describing *in* and *on* requires an intersection since, in some cases, speakers hesitate between these two prepositions to describe the same situation. Whereas schema (A) is compatible with external lexical formation, in which *in* and *on* are directly attached to support and containment, schema (B) is a case of internal lexical formation at its first stage. This means that *sobre* has not reached the stage in which it would prevent *en* being chosen when the conditions for the use of *sobre* are met. This preposition is preferred to *en* only in the cases in which an ambiguity must be avoided.

In the case of *in* and *inside*, morphology shows that the formation of *in* is likely to precede the formation of *inside* since it is much easier to imagine the addition of –*side* to *in* than to build *in* from *inside* by truncation. Whereas *in* can be used for the interior of closed containers (sentence 7a), for the interior of open containers (sentence 8a), for the material of containers (sentence 9a) and for masses (sentence 10a), *inside* can only be used in sentences (7b) and (9b):

(7) a. The jewels are in the box
b. The jewels are inside the box

(8) a. The wine is in the glass
b. *The wine is inside the glass

(9) a. The termites are in (the wood of) the cupboard
b. The termites are inside (the wood of) the cupboard

(10) a. The fish is in the water
b. *The fish is inside the water

Therefore, *inside* may be considered as a hyponym of *in*. I am not aware of examples in which *inside* must be used instead of *in* in order to avoid ambiguity, as was the case for *sobre* and *en* in Spanish. As long as this need is not felt by the speakers, a split of *in* and *inside* similar to the split of the category 'grue' in *green* and *blue* has not occurred. Hyponymy, then, does not necessarily lead to separation. Therefore, IN_1 does not correspond to IN-2D and IN-3D except those cases in which INSIDE can be used, as it is suggested in chart 2. I conclude this section with other discrepancies between chart 2 and the analysis proposed in chart 3.

According to the analysis of Levinson and Meira, IN-3D (containment) and IN-2D (inclusion in a plane) appear simultaneously at the second level of abstraction. The first notion implies the control of the target by the landmark whereas the second notion localizes the target in a two-dimensional landmark. The interaction occurs between two material entities in the former case while the landmark is a spatial entity in the latter case. In chart 3, IN-3D corresponds to CONTAINMENT. The notion corresponding to IN-2D should be in the LOC part of the chart and, indeed, an important function of LOC_3, conveyed by *à* in French is to locate a material entity (Jean) or a spatial entity (Montmartre) in a spatial entity (Paris):

(11) Jean est à Paris

(12) Montmartre est à Paris

As a matter of fact, the contrast between a material landmark and a spatial landmark might determine the difference between IN-2D and IN-3D better than the contrast between two-dimensional and three-dimensional. Indeed, the dimensionality of a spatial entity is a matter of conceptualization and the two-dimensional wood in sentence (13) looks rather three-dimensional in sentence (14):

(13) The rabbits play in the wood

(14) The birds fly in the wood

Interestingly, English uses *in* to translate (11) and French uses *en* – coming from the Latin preposition *in* – in front of feminine country names as well as in front of masculine country names beginning with a vowel:

(15) John is in Paris

(16) Jean est en France

Further hesitation between *at* and *in* to locate a target in geographic entities appears in the development of English. Indeed, whereas *in* was used for this function in Old English, *œt* introduces countries and large areas in Middle English and survives in Early Modern English to disappear in the nineteenth century (Lindkvist 1978). In order to explain these variations, I would like to claim that chart 3 captures only the prototypical values of the basic spatial prepositions. According to my analysis of *dans* ('in'), the first function of this preposition is the representation of the relationship CONTAINER/CONTENT (Vandeloise 1994, 2005). It accounts for the initial value of *in* in chart 3. From this initial value, *in* develops different meanings that can be more or less close to the prototypical value of other basic spatial preposition in the chart (Vandeloise 1995). Thus, what Levinson and Meira call IN-2D might be a later development of IN-3D. Whereas the landmark in IN-3D is a material entity with boundaries that allow physical control of the target, the landmark of IN-2D is a spatial entity. Spatial entities may have determinate boundaries – think of countries! – but they are virtual rather than material. Therefore, like *à* in French, IN-2D orientates itself toward localization and may compete with AT_2. Even with a spatial landmark, however, the French preposition *dans* keeps the memory of its first function. Compare sentences (17) and (18):

(17) ?Hans est dans Paris

(18) Les soldats sont dans Paris

Whereas sentence (17) looks odd, the use of *dans* in sentence (18) is perfect because the idea of a conflict evoked by the soldiers makes control more salient.[12]

Two notions introduced in the analysis of Levinson and Meira – UNDER and OVER – do not appear in chart 3. Numerous studies have been dedicated to *over* (Lakoff 1987, Brugmann 1988, Dewell 1994, Tyler and Evans 2001, Deane 2005). In contrast to Tyler and Evans, Dewell (1994) treats this preposition as a path preposition. This would be a sufficient condition to ignore *over* in a chart devoted to static spatial prepositions. One may also doubt whether this preposition belongs to basic prepositions since, besides English, Levinson and Meira do not mention another language with the category OVER. Brugman (1988) and Lakoff (1987) associate *over* to *above* and *across*. In fact, the two pictures illustrating OVER in Levinson and Meira's data might as well be described by *above*. However, in the analysis of Levinson and Meira, the link of *over* with *above* is ignored and OVER is considered as a notion that confines the scope of ON_2 in languages like English, in the same way as ATTACHMENT does for languages like Dutch.

Like OVER, UNDER has a particular status in chart 2 since it is introduced simultaneously with ON_1. As with IN, UNDER can convey control between two material entities when there is contact (sentence 19), or localize a target relative to the landmark (sentence 20):

(19) The red book is under the yellow book

(20) The shoes are under the table

Under looks like a converse of *on* in sentence (19) since this sentence implies that the yellow book is *on* the red book.[13] However, as illustrated by sentence (20), the converse relation between *on* and *under* is not as complete as the converse relation between the projective prepositions *in front* and *in back*. It would be easy to integrate *under* in chart 3. Its first meaning might be introduced below SUPPORT, in the same way as the prepositions *au-dessus* and *en dessous* are introduced below VERTICAL SEPARATION.

CONTROL

CONTAINMENT (*in*) SUPPORT (*on/under*)

Chart 5. Incorporating *under*

The meaning of *under* in sentence (20) might then be considered as an extension of its meaning in sentence (19) (Vandeloise 1991, chapter 12), just as IN-2D may be an extension of IN-3D. Compared to chart 2, this alternative presents the advantage of justifying the simultaneous introduction of *on* and *under* by their common relationship to SUPPORT. However, this might suggest too strong a connection between *on* and *under* and I will ignore the notion UNDER in chart 3.

3 Three modes of development

The implicational scale of Berlin and Kay for the basic terms of colors describes the order of appearance of these terms in the formation of languages. With the assumption that languages evolve from little sets of words to their complete lexicon, one may assume that a language with a system of seven basic color words is more evolved in this domain than a language with a system of five words. When the new terms occur through internal lexical formation by division, the development of languages can only go from the top to the bottom, i.e. from the most general terms to the most specific ones. If only internal lexical formation were involved in the creation of spatial terms, the same conclusions might be drawn for the typology of Levinson and Meira in chart 2 and for the conceptual hierarchy proposed in chart 3. This means that Korean would be a development of English, itself a development of Spanish. But then, *in* and *on* in English should derive from a word conveying the same situations as *en* in Spanish, just as *green* and *blue* are created by the split of the category 'grue'. And *kkita* and *nehta* in Korean would be created by internal lexical formation from a word with a larger distribution corresponding to English *in*.[14] If we do not have evidence in the history of English and Korean for such a development, this may simply mean that the formation of spatial terms is not parallel to the formation of basic color terms, and that there are different modes of genesis of spatial terms. Indeed, besides internal lexical formation, external lexical formation plays a role in their creation. In this case, *in* and *on* in English, as well as *kkita* and *nehta* in Korean, do not have to be the result of the split of a larger category.

They may have been created separately because the speakers of these languages attach a communicative virtue to the categories represented by these words.

With external lexical formation, the first spatial terms can appear at any level of generality in the hierarchy proposed in chart 3. Whereas Spanish might attach *en* directly to control, English may attach *in* to containment and Korean can associate immediately *kkita* to tight fit. The process of internal lexical formation proposed by MacLaury proceeds by division: a larger category 'grue' is replaced by two more specific categories designated by *green* and *blue*. This type of formation, therefore, can only go from the top to the bottom of the hierarchy. But, if some languages create words at a high degree of specificity by external lexical formation, there may be a different type of internal lexical formation going from the bottom of the hierarchy to the top. Besides *internal lexical formation by division*, then, there might be a mode of *internal lexical formation by union*. This mode of formation is internal because it relies on the existence of two more specific words. In contrast to internal lexical formation by division, however, internal lexical formation by union goes toward the top of the hierarchy. It can begin from the bottom of the hierarchy, with the most specific terms, or in the middle with intermediary notions. With these three modes of lexical formation, the developments illustrated in chart 6 are logically possible in languages.

(A) 1→2→3

(B) 1←2→3

(C) 3→2→1

Chart 6. A hierarchy of formation of spatial terms

In the case of schema (A), the creation of spatial terms begins with RELATIONS IN SPACE, a concept that gathers LOC and CONTROL. Schema (B), beginning in the middle of the hierarchy, is very reminiscent of the relationship between basic categories, supercategories and subcategories proposed by Rosch (1973). One goes from basic categories to supercategories by abstraction and to subcategories by specification. Schema (C) goes from the most concrete concepts to the most abstract.

Which of schemas (A), (B) and (C) is dominant in the creation of language? Schema (A) is illustrated by the development of basic color terms proposed by MacLaury in chart 1, in which the number of words increases from the top to the bottom of the hierarchy. The implicational scale of Levinson and Meira suggests a similar development for spatial terms. One may also surmise that languages have fewer words at their beginning than when they are fully developed. This parallelism pleads in favor of schema (A). Other arguments, however, show that schema (C) has a dominant role in the creation of languages. Indeed, Lévy-Bruhl (1922) claims that 'primitive' thought is characterized both by its concreteness and the absence of general concepts. For example, many Amerindian languages do not have a general term for walking but they have many more specific verbs that specify the direction, the trajectory or the manner of walking. According to Merleau-Ponty (1945), Maoris have 3000 terms for colors, not because they distinguish numerous colors, but because they do not recognize the same color when it belongs to different objects and use different words for it. Concrete specific concepts, then, might be at the origin of many words. As far as schema (B) is concerned, numerous experiments in cognitive psychology by Rosch and her colleagues (1975) demonstrate the preponderance of basic categories over subcategories and supercategories. If one may recognize basic categories in the middle of chart 3, this might plead for schema (B). CONTAINMENT and SUPPORT, then, should be more prototypical notions than CONTROL and TIGHT FIT or ATTACHMENT. Experiments by Choi et al. (1999) cast some doubt about the predominance of CONTAINMENT over TIGHT FIT. Indeed, English infants demonstrate more interest in the latter relation than in the former. If TIGHT FIT were universally dominant, Spanish children should begin their journey in language by limiting the use of *en* to the most specific contexts before enlarging its distribution to CONTROL. As far as attachment is concerned, Levinson and Meira found that many languages consider it a central topological notion. An explanation for this predominance might be that, with the exception of fruits attached to trees, ATTACHMENT is mainly an artificial way of stabilizing the target. This is in contrast to CONTAINMENT and SUPPORT that occur frequently in the nature. ATTACHMENT, then, would contrast with all the natural spatial relationships.

4 Language acquisition

At the beginning of the nineteenth century, 'primitive' thought was often compared to the thought of children (Lévy-Bruhl 1922). In this way, ontogeny, the acquisition of one language by one child, would reproduce phylogeny, the creation of a language by a civilization. However, there is an obvious difference between language creation and its recreation by the child since, in contrast to the community that must begin a language from scratch, the child is immediately confronted with a completely developed language. Furthermore, whereas the creation of a language requires production, children first learn a language through understanding and reproduction. In this section, I will first attempt to understand how the acquisition of a language without the help of a pre-linguistic conceptual system could occur. This eventuality appears very unlikely. Therefore, in the second part of this section, I will evaluate the incidence of schemas (A), (B) and (C) and of the pre-linguistic concepts in chart 3 on the acquisition of different languages.

An extreme form of determinism claims that no structured thought can exist without language. Therefore, only language can help to learn language. However, the first use of a word *W* by a child must be triggered by a situation in the world to which he associates *W*. Since, by hypothesis, the concept corresponding to the word *W* does not exist before *W* is acquired, its association with the situation must be referential or indexical. At the time of anchorage, the knowledge of the word is, of course, very tentative. Language can help to develop the full knowledge of the word in two ways. First, when the word *W* is used for a new situation in which the child would not have used it, he knows that his language establishes a connection between this situation and the other occasions on which he uses *W*. In contrast, when a different word is used for a situation in which he would have used *W*, he realizes that his language is sensitive to a difference that justifies the choice of another word. The use of *W* will be under-extended as long as the child does not know all the relevant similarities and overextended as long as he does not know all the relevant differences.

The strength of linguistic determinism depends a great deal on the nature of the connections language reveals to the child. Indeed, if they are based on similarities and differences recognizable in the extra-linguistic situations, language does not so much create the concept associated to the word as it guides the child through an array of differences and similarities available in the world. In this case, at each stage of the development of his acquisition of a word, the child has expectations that correspond to his knowledge of the word. Does the final stage corresponding to the complete acquisition of the word – if there is such a thing – have a special linguistic flavor that singles it out from the preliminary stages? I would rather guess that there is a continuum going from the anchorage situation to the final stage of knowledge. In this case, there may not be a clear-cut distinction between the established 'linguistic' concept and its elaboration.

If determinism is rejected and pre-linguistic concepts[15] are admitted, how can the acquisition of words expressing containment in languages like Spanish, English and Korean help us to understand what they are? Since these concepts are pre-linguistic, they are independent of language and can be shared by the infants speaking each language.[16] For example, a Spanish child could be receptive to the notion of TIGHT FIT (associated to *kkita* in Korean) and a Korean child could be sensitive to CONTROL (associated to *en* in Spanish). In this way, there might be a common set of pre-linguistic concepts shared by all the children in the world. On the other hand, children might have different pre-linguistic conceptual systems, even among children learning the same language. For example, there might be concrete-minded Spanish boys ready to anchor *en* to TIGHT FIT whereas other boys, more abstract-minded, would associate it directly to CONTROL and others, in the middle, would associate *en* to CONTAINMENT. As a result, these children should use different schemas in order to reach a complete knowledge of *en*: concrete-minded boys should use schema (C), going from the concrete to the abstract, whereas abstract-minded boys would get an almost immediate knowledge of the distribution of the word. In this way, schemas (A), (B) and (C) constitute the most economical ways of learning Spanish, English, and Korean respectively, since the concept corresponding to the level of abstraction chosen by these languages would be acquired directly. I do not have empirical data answering this question. They would be very helpful to choose between the existence of a common universal set of pre-linguistic concepts on the one hand, and the existence of individual variations in the acquisition of spatial terms on the other hand. Spanish infants under-extending *en* to TIGHT FIT or to CONTAINMENT, for example, would provide strong evidence for these pre-linguistic concepts since these underextensions cannot be justified by their language. The same thing would be true for English infants limiting the use of *in* to TIGHT FIT.

Schema (C), proceeding from the most specific to the most abstract concepts, might be built entirely conceptually, without the help of language, by the child who recognizes the commonalities between TIGHT FIT and LOOSE FIT, and afterwards between CONTAINMENT and SUPPORT. In this way, a child going through this process of generalization would have the three pre-linguistic concepts at his disposal before he begins to acquire his language. It is very easy, however, to see how language can contribute to the building of these concepts. Indeed, suppose that an English child under-extends the meaning of the preposition *in* and restricts its use to the representation of TIGHT FIT. He will quickly realize that adults are also using the same word for LOOSE FIT. Therefore, he will be inclined to look for similarities that he might otherwise have overlooked. In this case, one might say that language is a necessary condition, if not a sufficient one, for the constitution of concepts. A Spanish child who would under-extend the meaning of the Spanish preposition *en* and associate it with TIGHT FIT or CONTAINMENT would also receive plenty of warnings from adult language until he extends the use of the preposition *en* to CONTROL, which embraces the whole extension of the preposition in adult language. Korean children, in contrast, will not find in their language any incentive to extend TIGHT FIT to CONTAINMENT or to CONTROL. A Spanish child who underextends *en* will correct himself more easily than a Korean child who overextends *kkita* since the former will receive positive evidence (each time he hears *en* used in circumstances he was not using it), whereas the Korean child will only receive negative data (when adults correct him if he uses *kkita* inappropriately).

5 Conclusion

The genesis of basic colors (Berlin and Kay 1968, MacLaury 1993) provides hints to better understand the genesis of spatial terms. Two modes of internal lexical formation inside the language system (by division and by union) have been opposed to external lexical formation that attaches words directly to extra-linguistic notions of utmost importance in the linguistic community.

Before presenting my views on the genesis of spatial terms, I have discussed the analysis of Levinson and Meira (2003). They exclude projective prepositions from their investigation because, according to the authors, these prepositions belong to a different subsystem. The development of spatial terms begins with an all-encompassing adposition AT covering all the relationships in space. In chart 2, the system enriches itself through internal lexical formation by division. The new notions introduced are mainly topological basic categories like ON/OVER (superadjacency with or without contact), UNDER (subadjacency with or without contact), NEAR (proximity). IN-3D (containment) and IN-2D (inclusion in a surface) are also notions proposed in the analysis, even though I believe that containment is a dynamic notion rather than a topological one. According to my proposition, the dichotomy between CONTROL (a general dynamic notion) and LOC (a general topological notion of localization) constitutes the first step in the genesis of spatial terms. As illustrated by the preposition of Old English *æt,* this part of the system evolves essentially by internal lexical formation by division. In contrast to Levinson and Meira, I have introduced the projective notions. As far as the dynamic spatial system is concerned, different levels of specification may be observed in different languages. For example, the Spanish preposition *en* represents a general notion of CONTROL whereas the English prepositions *in* and *on* convey more specific notions of CONTAINMENT and SUPPORT. No historical data show that this enrichment occurs by internal lexical formation by division, which means that IN and ON might occur by external lexical formation. In this case, the comparison of the different levels of abstraction cannot be done inside one and the same language but requires a comparison between different languages.

As far as color terms are concerned, one may consider that languages with more specific terms are a development of languages with more general terms according to schema (A) in section 3. If language creation was proceeding by internal lexical formation only, one might draw the same conclusion for spatial terms related to containment in chart 3. But external lexical formation may attach a word directly to different levels of abstraction. Such is the case for natural kinds like dogs and birds. Nouns in basic categories are considered more prototypical than nouns for supercategories and subcategories and are acquired first. The creation of these words conforms to schema (B): supercategories and subcategories develop from basic categories by abstraction and specification respectively. Finally, according to Lévy-Bruhl, human thought at its beginning evolves from the concrete to the abstract, according to schema (C). This schema would give precedence to the most specific basic terms.

In the last section of this article, I investigate how the acquisition of language might help to provide clues about the development of spatial terms. How do children adjust to the level of abstraction of control terms in the language they are learning: general like *en* in Spanish, intermediary like *in* and *on* in English or specific like *kkita* in Korean and *aan* in Dutch? Any discrepancies between child and adult language, as well as the adjustments children are making to reach a complete command of spatial control terms, may be helpful to understand the genesis of language. Three extreme – and much caricatured – avenues may be proposed. First, the *universal view*: before speaking, all the children in the world first pay attention to the same concepts and, afterwards, adjust to their language through schemas (A), (B) or (C). Second, the *relativist view*: after a period of passive understanding, children are immediately tuned to the level of abstraction that characterizes their language. And finally, the *individualistic view*: even in a single language, different children make different hypotheses and reach the command of control terms by different ways. It might be useful to keep the three possibilities in mind when we analyze any data that might be relevant for the genesis of language.

Notes

1 According to MacLaury, the category of warm colors splits before the category of cool colors and *red* appears in third position because the perceptual difference between red and yellow is more conspicuous than the contrast between green and blue.

2 'Relevant sort' might only have a specific sense if there was a consensus about the central meanings of these prepositions, which is far from being be the case.

3 *Topology* here has not a mathematical meaning but refers to static common sense relationships in space, such as neighborhood and inclusion, as used in Piaget and Inhelder (1956).

4 Levinson and Meira do not need to be concerned by verbs since they explicitly limit their analysis to adpositions.

5 Maybe the authors consider that *on* appears simultaneously with its converse *under*. However, in language acquisition, *under* is understood much later than *on* (Rohlfing 2003).

6 If the tablecloth covers the table entirely, its situation would be described in English by *the tablecloth is over the table* rather than by *the tablecloth is on the table*.

7 For some spatial relationships, like the situation described by *the sun is above the earth* or *the airplane is over the house*, proximity of the two material entities is not a necessary condition. In these particular cases, however, accessibility may be obtained by the rays in the case of sun or by bombs (or landing) in the case of the airplane.

8 Adpositions marking control may help to locate the target but they do it only indirectly. A sentence like *the wine is in the glass* is used to indicate that the wine is available for drinking – as opposed to *the wine on the floor*. French children are well aware that the preposition *dans* conveys localization only indirectly when their answer to the question 'Where is the King?' is: Dans sa chemise ('In his shirt').

9 *Kkita* might also be considered as a specification of *on* when it represents a relation of tight fit between the target and a horizontal landmark. However, these situations are extremely rare since, except for magnetic objects, the pressure exerted by the target on its support is not stronger than its weight. Two horizontal pieces of Lego fitting together are an example of horizontal tight fit. However, *in* might be used in this case, in contrast to *on*, preferred if one piece is simply put on the other, without adjustment.

10 In Early Modern English, *oet* had acquired the modern form *at*.

11 These sentences are adapted from Fortis (2004). Thanks to Ignasi Navarro-Ferrando for comments on these examples.

12 As noted by an anonymous reader, the control here is exerted by the target (*the soldiers*) rather than by the landmark (*Paris*).

13 In French, the phonetic similarity between *sur* and *sous* reinforces the parallelism between the two spatial relations they convey.

14 Even though this hypothesis looks similar to the hypothesis concerning the common origin of *in* and *on* in English, there is an important difference since *in* and *on* are acquired approximately at the same time by children whereas *kkita* appears to be learned earlier than *nehta* in Korean. These two words, then, do not have the same status in acquisition.

15 Tye (2000: 176) speaks of 'perceptual concepts' that are 'a matter of having a stored memory representation that has been acquired through the use of sense organs and available for retrieval, thereby enabling a range of discriminations to take place '

16 Society can introduce differences in the set of pre-linguistic concepts independently of language. This is the case for societies that have no containers or societies that have only round symmetrical objects.

References

Berlin, B. and Kay, P. (1968) *Basic Color Terms.* Berkeley: The University of California Press.

Bowerman, M. (1996) The origins of children's spatial semantic categories. In J. Gumperz and S. Levinson (eds) *Rethinking Linguistic Relativity* 145–176. Cambridge: Cambridge University Press.

Brown. P. (1994) The INs and ONs of Tzeltal locative expressions: The description of stative descriptions of location. *Linguistics* 32 (4/5): 743–790.

Brugman, C. (1988) *The Story of* over: *Polysemy, Semantics and the Structure of the Lexicon.* New York: Garland Press.

Choi, S., Mc Donough, L., Bowerman, M. and Mandler, J. (1999) Early sensitivity to language specific spatial categories in English and Korean. *Cognitive Development* 14: 241–268.

Deane, P. (2005) Multimodal spatial representation: On the semantic unity of over. In Beate Hampe (ed.) *From Perception to Meaning: Image Schema in Cognitive Linguistics* 235–285. Berlin: Mouton De Gruyter.

Dewell, R. (1994) *Over* again: Image-schema transformation in semantic analysis. *Cognitive Linguistics* 5: 351–380.

Feist, M. (2004) Talking about space: A cross-linguistic perspective. In K.D. Forbus, D. Gentner and T. Regier (eds) *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Fortis, J-M. (2004) L'espace en linguistique cognitive: Problèmes en suspens. *Histoire Epistémologie Langage* XXVI(1): 43–88.

Gunnarson, K.A. (1986) *Loin de X, près de X et parallèlement à X.* Syntagmes prépositionnels, adjectivaux ou adverbiaux? *Le français moderne* 54(2): 1–23.

Lakoff, G. (1987) *Women, Fire and Dangerous Things.* Chicago: The University of Chicago Press.

Levinson, S. and Meira, S. (2003) Natural concepts in the spatial topological domain – adpositional meanings. *Language* 79(3): 485–516.

Lévy-Bruhl, L. (1922) *La mentalité primitive.* Paris: Presses Universitaires de France.

Lindkvist, K. (1978) *AT* vs. *ON, IN, BY*. Stockolm: Almqvist and Wiksell International.

MacLaury, R. (1993) Social and cognitive motivation of change: Measuring variability in color semantics. *Language* 69(3).

Merleau-Ponty (1945) *Phénoménologie de la perception.* Paris: Gallimard.

Piaget, J. and Inhelder, B. (1956) *The Child's Conception of Space.* London: Routledge and Kegan Paul.

Rohlfing, K. (2003) *UNDERstanding. How Infants Acquire the Meaning of* Under *and other Spatial Relational Terms.* Bielefield University: PhD dissertation.

Rosch, E. (1973) Natural categories. *Cognitive Psychology* 4: 328–350.

Rosch, E. and Mervis, C. (1975) Family resemblances: Study in the internal structure of categories. *Cognitive Psychology* 93: 10–20.

Saussure, F. (1916) *Cours de linguistique générale.* Paris: Payot.

Tye, M. (2000) *Consciousness, Color and Content.* Cambridge, MA: The MIT Press.

Tyler, A. and Evans, V. (2001) Reconsidering prepositional polysemy networks: The case of *over*. *Language* 77: 724–65.

Vandeloise, C. (1991) *Spatial Prepositions: A Case Study in French.* Chicago: The University of Chicago Press.

Vandeloise, C. (1994) Methodology and analyses of the preposition. *Cognitive Linguistics* 5(5): 157–185.

Vandeloise, C. (1995) De la matière à l'espace. *Cahiers de grammaire* 20: 123–145.

Vandeloise, C. (2003) Containment, support and linguistic relativity. In H. Cuyckens, R. Dirven and R. Taylor (eds) *Cognitive Approaches to Lexical Semantics* 393–425. Berlin: Mouton De Gruyter.

Vandeloise, C. (2005) Force and function in the acquisition of the preposition *in*. In L. Carlson and E. van der Zee (eds) *Functional Features in Language and Space* 219–231. Oxford: Oxford University Press.

Wierzbicka, A. (1990) The meaning of colour terms: Semantics, culture and cognition. *Cognitive Linguistics* 1: 11–36.

8 Forceful prepositions[1]

Joost Zwarts

Introduction

As the title suggests, the focus of this paper is on prepositions with a force-dynamic aspect, as in the following example sentence:

(1) Alex ran into a tree

This sentence does not mean that Alex ended up inside a tree, but that she came forcefully into contact with the tree. There is a spatial component in this sentence (Alex followed a path that ended where the tree was), but there is also a force-dynamic component (the tree blocked her movement).

This use of *into* brings together two conceptual domains that are fundamental in the semantics of natural language: the spatial domain and the force-dynamic domain, each of which comes with its own intricate system of concepts and relations. The spatial domain is primarily concerned with location, movement and direction, the force-dynamic domain with causation, control and interaction. The basic thematic roles of the spatial domain are Figure and Ground (Talmy 1983), Theme, Goal, and Source (Gruber 1976), and Trajector and Landmark (Langacker 1987), while the force-dynamic domain has Agent and Patient (Jackendoff 1987, Dowty 1991) or Agonist and Antagonist (Talmy 1985).

The interaction of these two domains in the verbal domain has been relatively well-studied (for example in Jackendoff 1987, Croft 1991, 2009, and others), but this is different with the prepositions, that seem to be the spatial words *par excellence*. However, there is a growing awareness in the study of prepositions and spatial language that force-dynamic notions do play an important role (Vandeloise 1991, Bowerman and Choi 2001, Coventry and Garrod 2004, Carlson and Van der Zee 2005). It is becoming clear that geometric notions alone do not suffice to capture the meaning of even very basic prepositions like *in* and *on*, let alone an obviously force-dynamic preposition like *against*. However, what is not yet clear is how the role of force-dynamics can be transparently and adequately modeled in representations of the meaning of prepositions. This paper makes some specific proposals about how to do this.

In section 1 I will single out a few important phenomena that concern prepositions, most of which are well-known from the literature, that require reference to forces in one way or another. I will then argue in section 2 that the general semantic mechanics that underlies reference to forces can best be captured in terms of *vectors* (O'Keefe 1996, Zwarts 1997, Zwarts and Winter 2000). These force vectors will allow an interface between the force-dynamic part and the geometric part of the semantics of prepositions along lines worked out in section 3. In a concluding section 4 I will sketch the potentials of the model in understanding cross-linguistic variation in the domain of containment and support.

1 Forced beyond geometry

In order to illustrate the need for force-dynamics in the semantics of prepositions, this section will briefly discuss some relevant aspects of the interpretation of the English prepositions *against* and *in* and *on* and the Dutch prepositions *op* and *aan*.

Against

Against is the clearest example of a preposition that is not purely geometric:

(2) Alex bumped against the wall

Dictionaries characterize the meaning of *against* in such terms as 'collision', 'impact', and 'support'. It typically combines with verbs like *crash*, *lean*, *push*, *bang*, and *rest,* verbs that all involve forces, either dynamically (3a) or statically (3b):

(3) (a) There was a loud bang against the door
(b) The rifle rested against the tree

Against is a relation that always implies physical contact between the Figure and the Ground. This contact is usually lateral, i.e. from the side, involving horizontal force exertion. We can see this clearly when we contrast *against* with *on*:

(4) (a) Alex leaned against the table
(b) Alex leaned on the table

(4a) refers to a horizontal force, requiring contact with the side of the table, but (4b) to a downward force, involving the tabletop. Notice finally that the result of the force is left unspecified when *against* is used:

(5) (a) Alex pushed against the car
(b) Alex pushed the car (to the garage)

Sentence (5a) does not tell us what the 'reaction' of the car is, whether it moves as a result of the pushing or stays put. It simply leaves the result of Alex' force open. Notice the contrast with (5b) in this respect, a construction that allows directional PPs like *to the garage*, apparently because the transitive use of *push* implies that pushing results in motion of the direct object.

In and *on*

Although probably two of the most common prepositions in English, *in* and *on* have also proved to be the most difficult ones to define in geometric terms (Herkovits 1986). Intuitively, the geometric condition for *in* is 'inclusion' and the geometric for *on* 'contiguity'. But, as Vandeloise (1991) and Coventry and Garrod (2004) have argued, these conditions are not always necessary for the proper use of *in* and *on*, respectively, and they are not always sufficient either. Here are two well-known examples:

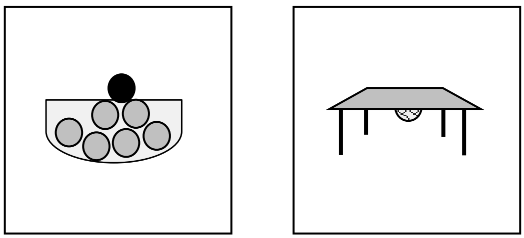


Figure 1 *Figure 1a* *Figure 1b*

In Figure 1a the black marble is not included in the (interior of the) bowl, but we still would describe this situation with sentence (6a) below. In Figure 1b there is contiguity of the ball with the table, but nevertheless, the description in (6b) is not felicitous.

(6) (a) The black marble is in the bowl
    (b) The ball is on the table

So, there are relations without inclusion that we call *in*, as in (6a), and there are relations with contiguity that we don't call *on*, (6b). These observations have led Vandeloise (1991) and Coventry and Garrod (2004) to propose that force-dynamic conditions are needed instead of, or in addition to, the geometric conditions of containment and contiguity. Even though the black marble in Figure 1a is not included in the bowl, its position is in some sense controlled by the bowl through the grey marbles. There is a force-dynamic relation of *containment*. The position of the ball underneath the table in Figure 1b is not controlled by the table in the way that would be necessary for *on* to be apply, namely by *support*, a force relation that requires the ball to be on the opposite, upper side of the tabletop.

Aan and *op*

For the third example of the role of force-dynamics, we need to turn to Dutch. The English preposition *on* corresponds to two distinct Dutch words, *op* and *aan* (Bowerman and Choi 2001, Beliën 2002):

(7)	(a)	a cup on the table	een kopje op de tafel	'support'
	(b)	a bandaid on a leg	een pleister op een been	'adhesion'
	(c)	a picture on the wall	een schilderij aan de muur	'attachment'
	(d)	a handle on the door	een handvat aan de deur	'attachment'
	(e)	a leaf on a twig	een blaadje aan een tak	'attachment'

As Bowerman and Choi (2001) show, Dutch uses *aan* for spatial relations that involve *attachment* (7c-e) while *op* is used for relations of *support* (7a) and *adhesion* (7b). So the distinction between *aan* and *op* is again not purely geometric, but also force-dynamic, given that relations of attachment, support and adhesion presuppose that the related objects exert forces on each other.

Extended location

Herskovits (1986) noted that the applicability of *on* can be extended in an interesting way, crucially involving force-dynamics again. The English sentence (8a) and its Dutch translation (8a') describe the situation in Figure 2a below, even though the cup is really standing on a book that is lying on the table.

(8) (a) The cup is standing on the table
(a') Het kopje staat op de tafel
(b) De lamp hangt aan het plafond
The lamp hangs on the ceiling
'The lamp is hanging from the ceiling'

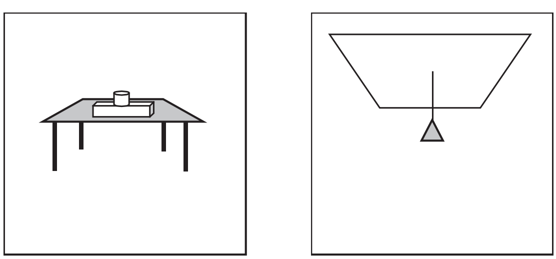


Figure 2 *Figure 2a* *Figure 2b*

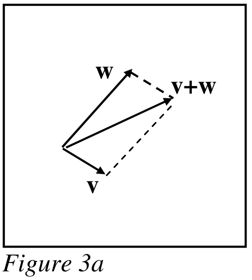
In the same way, Figure 2b fits the Dutch description in (8b), although the lamp is not directly hanging from the ceiling, but connected to it by a cable. What we see then is that these prepositions usually require direct contact ('contiguity' or 'attachment') between Figure and Ground, but it is possible to make the relation indirect, by the intervention of a third object.

So, we have seen a range of force-related notions that seem to play a role in the semantics of spatial prepositions: 'impact', 'control', 'containment', 'support', 'adhesion', 'attachment'. What is the force-dynamic system behind these notions? And, given that there is interaction with purely spatial concepts (like vertical and horizontal direction, inclusion and contiguity), how does this force-dynamic system interface with the spatial system? In other words: what is the geometry of forces?

2 A geometry of forces

Vectors

Since the notion of vector is going to play an important role in our analysis of the geometry of forces, we will start with a brief and informal overview of some core concepts. Vectors are a powerful tool to analyze geometrical concepts. Essentially, a vector **v** is a directed line segment, an arrow, as illustrated in the diagrams in Figure 3. There are different ways to represent a vector in linear algebra, but for our purposes it is sufficient to understand it at this basic level. Free vectors have a length and a direction only, located vectors have a particular starting point. The zero vector has no length and no direction, but it can have a location.



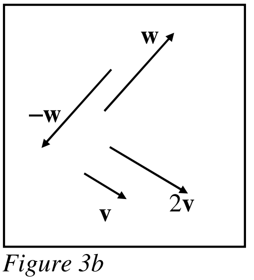


Figure 3 *Figure 3a* *Figure 3b*

In the algebra of vectors two vectors **v** and **w** can be added up to form the vector sum **v**+**w**. Figure 3a illustrates how this vector sum forms the diagonal of the parallelogram of which **v** and **w** are the sides. Scalar multiplication is another operation, in which a vector **v** is multiplied by a real number *s*, to form the scalar multiple *s***v**, which is *s* times as long as **v** (see Figure 3c). Each non-zero vector **v** has an inverse -**v** of the same length, but pointing in the opposite direction. With this background, we can now take a closer look at vectors in the force-dynamic domain.

Force vectors

The literature about force-dynamics is not extensive, but it would still go too far for the purposes of this paper to give here even a short overview of what has been written in Talmy (1985) and works inspired by it, like Johnson (1987), Langacker (1991), Croft (1991, 2009), Jackendoff (1993), Wolff and Zettergren (2002). I will restrict myself to extracting from the literature some useful ingredients for a rudimentary model of forces.

(i) The first ingredient is that forces have vector properties. Even though this is not made explicit by all the authors mentioned above, forces have two parameters: they have a *magnitude* (they can be smaller or bigger) and they have a *direction* (they point in a particular, spatial, direction). These two parameters define a vector. The third parameter, less relevant here, is the *location* of the force, i.e. the physical point where the force is exerted.

(ii) Usually, a force is exerted by one object, the Agent, on another object, the Patient. The Agent is what Talmy (1985) calls the Antagonist, the Patient is the Agonist. Talmy's terms have not found general currency and I will therefore use the more common terms Agent and Patient here, even though this occasionally leads to somewhat awkward results, as we will see at the end of the next section.

(iii) The Patient may also have its own force vector. This vector represents the inherent tendency of the Patient to move in a particular direction. The tendency of material objects to go downwards, because of gravitation, is an example of such an inherent force vector (even though, strictly speaking, the earth is the Agent here).

(iv) Because of the interaction between the forces of Agent and Patient, there is a resultant vector that determines the result of this interaction. This resultant vector is simply the sum of the Agent's and the Patient's vector (according to the parallelogram rule) and this sum can be zero, when the forces of Agent and Patient are equal but opposite.

All of these ingredients can be illustrated with a concrete example, based on the experiment and analysis of Wolff and Zettergren (2002). Consider the example:

(9) The fan prevented the boat from hitting the cone

In their experiment, subjects were asked to judge whether sentences like these applied to short and simple animations in which different kinds of objects were seen to exert forces on a moving boat. Wolff and Zettergren found that the conditions for using causative verbs like *prevent* could be analyzed in terms of the vector force interaction of the objects involved. A situation that falls under sentence (9) might look as follows:

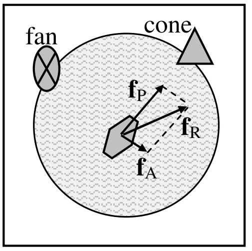


Figure 4

In this picture, $\mathbf{f}_A$ is the force of the Agent, the fan, blowing against the Patient, the boat. The boat has its own force tendency $\mathbf{f}_P$, that is directed towards the cone. In this example, the Patient's force vector is determined by the engine and the rudder of the boat. When we add up the two vectors we get the resultant vector $\mathbf{f}_R = \mathbf{f}_A + \mathbf{f}_P$ that tells us where the boat is heading, as a result of the combination of the two forces. All of this is simple high-school physics, but it allows Wolff and Zettergren to isolate the directional parameters that determine how people actually apply causative verbs to dynamic scenes: the directions of $\mathbf{f}_A$, $\mathbf{f}_P$ and $\mathbf{f}_A$+$\mathbf{f}_P$ with respect to a target T.

In the model of Wolff and Zettergren, the relative magnitudes of these force vectors are essential for understanding how people label particular situations. A stronger force vector $\mathbf{f}_A$ results in a stronger sum $\mathbf{f}_A$+$\mathbf{f}_P$, which will then bring the Patient far enough away from the target to judge the situation as an instance of *prevent*. Notice that the *absolute* lengths of the force vectors in the spatial diagrams have no direct linguistic significance. Multiplying all the force vectors in a situation by the same scalar would represent the same force-dynamic concept. What matters for the understanding of verbs like *prevent* are ultimately the relative magnitudes and absolute directions of the three vectors.

For *prevent* to be applied to a force-dynamic situation, it is necessary that $\mathbf{f}_P$ is directed towards the target T, while $\mathbf{f}_A$ and $\mathbf{f}_A$+$\mathbf{f}_P$ are not. The verbs *cause* and *enable* are different, in that the result $\mathbf{f}_A$+$\mathbf{f}_P$ is directed towards the target. *Enable* requires that the vectors of both Patient and Agent point towards the Target, with *cause* they are opposite. See Wolff and Zettergren (2002) for further explanation and evidence concerning this vector-based force-dynamics of causative verbs. I will turn now to a class of verbs that refer to forces in a more direct and more spatial way.

Forceful verbs

The first two verbs that I would like to consider are *push* and *pull*. Obviously, these two verbs are opposites, more specifically directional opposites (Cruse 1986):

(10) (a) Alex pushed the pram
(b) Alex pulled the pram

But what is it exactly about their meanings that makes them opposite? It is not the directions of motion that are opposite, because Alex can push or pull the pram without the pram actually moving. In this respect, *push* and *pull* are different from opposite motion verbs like *enter* and *leave* or *come* and *go* that have opposite spatial trajectories. The opposition of *push* and *pull* is also different from the opposition between *cause* and *prevent* seen in the following examples:

(11) (a) The fan caused the boat to hit the cone
(b) The fan prevented the boat from hitting the cone

where the results are in opposition (hitting the cone vs. not hitting the cone). The opposition between *push* and *pull* lies purely in the opposite directions of the force vectors involved, relative to the Agent. With *push* the force vector is pointing away from the Agent, with *pull* it is pointing in the direction of the Agent. This is schematically indicated in the following two figures:

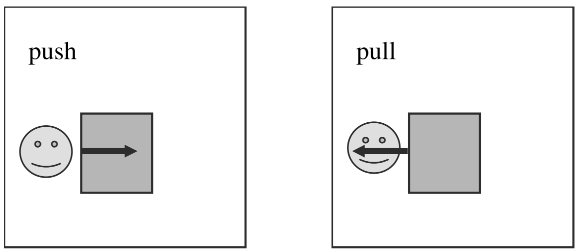


Figure 5 *Figure 5a* *Figure 5b*

The vector is located at that point of the Patient where the Agent exerts its force and its length represents the magnitude of the force. If there are no other forces interacting with the pushing or pulling force, the Patient will move in the direction of the force vector, so either away from the Agent in Figure 5a or towards the Agent in Figure 5b. The force relation between Agent and Patient is closely related to a purely locative relation between them. With pushing, the Agent is *behind* the Patient, with pulling it is *in front of* the Patient. We can already see here how force-dynamic and spatial notions interface in a way that is crucially based on direction and that requires forces to have spatial direction.

What is the role of the *length* of the force vectors in Figure 5? As I said above, the particular scale with which we represent force vectors in spatial diagrams is arbitrary. However, the magnitude of forces does play a role, in two ways. First, verbs like *push* and *pull* can be modified by an adverb like *hard*, which suggests that the length of a force vector has linguistic relevance, although in a non-quantitative way, of course. Second, on a more conceptual level, we could imagine that there are two people pulling *equally hard* on opposite sides. In that case, we need to compare the magnitudes of forces to conceptualize and describe this situation as one of balance.

Because I am mainly interested here in the *directions* of the force vectors, relative to Agent and Patient, and not so much in their location and length, I will use a simpler and much more schematic graphical representation, that abstracts away from the other two parameters:

(12) push: Agent --> Patient
pull: Agent <-- Patient

The arrows in (12) represent the spatial directions of the force vector, either pointing from Agent to Patient, or from Patient to Agent.

Let me make a bit more precise how this could be represented in a formal vector model. Let us assume that the *spatial* relation between Agent and Patient is represented by a spatial vector $\mathbf{v}_{PA}$ pointing from the Patient to the Agent (connecting their centers of gravity, for instance). This vector $\mathbf{v}_{PA}$ then gives us the spatial frame with respect to which we can represent a *force* vector $\mathbf{f}_A$, as indicated in Figure 6a and 6b. What *push* and *pull* express, is how $\mathbf{f}_A$ is aligned with respect to vector $\mathbf{v}_{PA}$. $\mathbf{v}_{PA}$ and $\mathbf{f}_A$ are opposite for *push*, they point in the same direction for *pull*. This is what (12) intends to represent in an informal way.

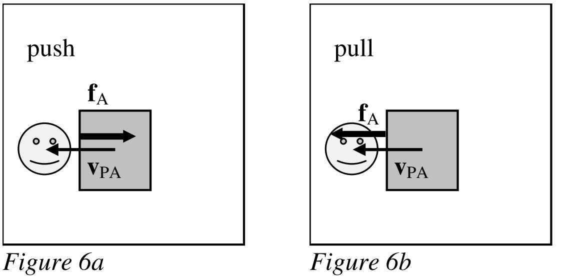


Figure 6 *Figure 6a* *Figure 6b*

Another pair of opposite force verbs is *squeeze* and *stretch*, that are very close to *push* and *pull*. *Squeeze* can be defined as 'press from opposite sides', while *stretch* is 'pull in opposite directions:

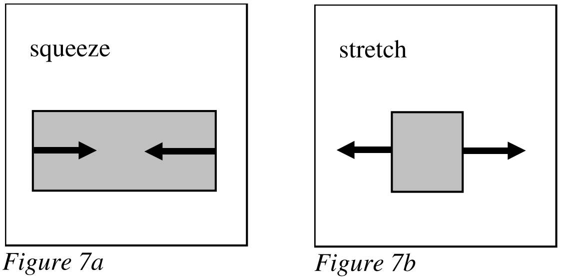


Figure 7 *Figure 7a* *Figure 7b*

Again, there is a close relation with basic spatial notions: the forces of *squeeze* have an *inward* direction with respect to the Patient and the forces of *stretch* an *outward* direction. If there is a resulting change, it is a change of shape or volume, a shrinking or expanding. Here also, I will use a more schematic representation of the force-dynamic relation between Agent and Patient:

(13) squeeze: Agent –> Patient <– Agent
stretch: Agent <– Patient –> Agent

The third and last pair of forceful verbs to be discussed here is *lean* – *hang*. Both verbs can refer to a downward force exerted by the subject, as in the following examples:

(14) (a) Alex was leaning on the table with his elbows
(b) There was a light bulb hanging from the ceiling

The distinction lies in the relative position of the Agent and the Patient. In (14a) the Agent (Alex, or rather, his elbows) is above the Patient (the table), in (14b) the Agent (the light bulb) is below the Patient (the ceiling). In one sense, leaning and hanging are a bit like pushing and pulling. Leaning is like pushing from above and hanging is like pulling from below. But there are two important differences. The first difference is that the forces don't come from within the Agent, but are the result of gravitation. Alex does not have to *do* something to the table when he is leaning on it. The second difference is that the force exerted by the Agent is counterbalanced by an equal but opposite force of the Patient (indicated by the grey arrow), creating a static situation of balance, as illustrated in the following two figures:

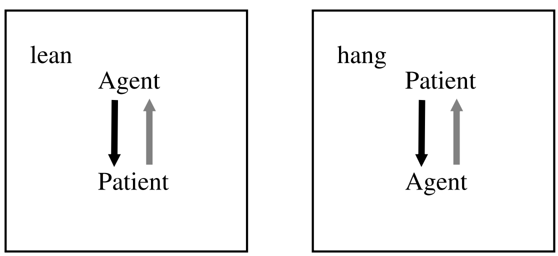


Figure 8 *Figure 8a* *Figure 8b*

We can see the configuration of Figure 8a as a representation of *support*: the Patient is supporting the Agent. Figure 8b, on the other hand, captures an important aspect of the notion of *attachment*: the Agent is attached to the Patient.

It is in this situation that the use of the terms Agent and Patient becomes somewhat awkward. From the perspective of the theory of thematic roles, we would not usually call the subject of *lean* or *hang* an Agent and the table or the ceiling a Patient, because we cannot say that the subject is doing something to the object of the prepositions. Talmy's term Agonist and Antagonist are not appropriate either. I will therefore use a slightly different representation for situations of leaning and hanging, respectively:

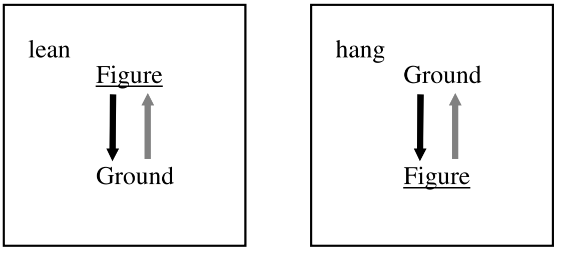


Figure 9 *Figure 9a* *Figure 9b*

The objects are labeled Figure and Ground here. The underlining of Figure indicates that it is this participant that exerts the primary downward force to which the Ground is reacting. What is not made explicit in the representation is that gravitation is responsible for the Figure's force.

Two kinds of arrows

The arrows in the representation that I proposed in the previous section should not be confused with the arrows that are found in Langacker (1991) and Croft (1991, 2009). There an arrow is used to indicate the direction in which energy is transmitted from one object to another. The direction of the arrow is non-spatial and non-vectorial, and it is always pointing from the Agent to the Patient, or from a more agentive to a less agentive participant in a situation, e.g. from an Agent to an Instrument. In fact, notions like Agent, Patient and Instrument can be more or less defined from their position in a chain of causal relations:

(15) X --> Y --> Z

In such a chain, X will be the Agent, Y the Instrument and Z the Patient. In other words, the arrow is *thematic*, representing the roles that objects play in a force *relation*.

In the representation that is used here, and also in Johnson (1987) and Wolff and Zettergren (2002), the arrow represents the spatial direction of the force with respect to given objects and dimensions. (15) then means that there is a force working away from X and towards Y and a force working away from Y towards Z. It does not specify the origins of these forces: this is where we need to label objects as Agent or Patient, or underline them to indicate their force-dynamic primacy.

Both representations are justified, but for different reasons and for different purposes. The first kind of arrow is useful for representing the thematic side of causal relations, particularly for analyzing aspectual and argument structure, as argued for in Croft's work. The second kind of arrow is needed for the spatial side of causal relations and is indispensable for understanding verbs with a directional component, as we saw in this section, but also for force-dynamic prepositions, as we will see in the next section.

3 Prepositional forces

Verbs and prepositions in Dutch

Verbal forces and prepositional forces interact. One area where we can see this clearly is in some relevant verb preposition patterns in Dutch (Beliën 2002). While *trekken* 'pull' is used with *aan*, as shown in (16a) and (16b), the opposites *duwen* 'push' or *drukken* 'press' are used with *op* or *tegen*, (16a') and (16b'):

(16) (a) aan de wagen trekken
on the car pull
'pull the car'

(a') tegen de wagen duwen
against the car push
'push the car'

(b) aan de bel trekken
on the bell pull
'pull the bell'

(b') op de bel drukken
on the bell press
'press the bell'

The choice between *op* and *tegen* is subtle, depending on the direction and the granularity of the force. While (16a') is used for a horizontal force exertion, (17a) below is used for a force that comes from above. *Op* in (16b') is the normal preposition to use when a bell is pressed with a finger, but *tegen* is found, as in (17b) when something bigger exerts a force on the bell, in a non-canonical way:

(17) (a) op de wagen duwen
on the car push
'push on the car'

(b) tegen de bel drukken
against the bell press
'press against the bell'

Hangen 'hang' and *leunen* 'lean' also correlate with particular prepositions:

(18) (a) aan de wagen hangen
on the car hang
'hang on the car'

(a') op/tegen de wagen leunen
on/against the car lean
'lean on/against the car'

(b) aan de bel hangen
on the bell hang
'hang on the bell'

(b') op/tegen de bel leunen
on/against the bell hang
'lean against the bell'

Hangen clearly goes with *aan*, (18a) and (18b), while *leunen* goes with *op* and *tegen*, (18a') and (18b'). However, *hangen* is also possible with *op* and *tegen*. Notice the contrasts in the following examples:

(19) (a) Het gordijn hangt aan het plafond
The curtain hangs on the ceiling
'The curtain is hanging from the ceiling'

(b) Het gordijn hangt op de grond
The curtain hangs on the ground
'The curtain is hanging on the ground'

(c) Het gordijn hangt tegen het raam
The curtain hangs against the window
'The curtain is hanging against the window'

The curtain is suspended from the ceiling, and *aan* is used in (19a) to describe this relation. However, to describe the situation in which the curtain touches the ground at the lower end *op* is used in (19b) and its contact with the window in the vertical direction is indicated by *tegen* in (19c). In the remainder of this paper I will ignore the use in (19b) and (19c).

In order to make sense of the patterns of (16) and (18), the prepositions *aan*, *tegen* and *op* need to involve a force relation between the Figure and the Ground. The basic idea is that *aan* 'on' is like pulling and hanging: a relation in which the Figure is at the same time a kind of Agent, exerting a force on the Ground that is directed towards itself. I will represent this as follows:

(20) aan: Figure <-- Ground

What characterizes *aan* is that the force vector is pointing from the Ground towards the Figure. The Figure is underlined to indicate the division of agentivity in this relation: it is the Figure that has an intrinsic tendency to move. *Tegen* 'against' and *op* 'on' are the opposite of *aan*, in the sense that the force points away from the Figure towards the Ground:

(21) tegen: Figure --> Ground
op: Figure --> Ground

In this respect, *tegen* and *op* are like pushing and leaning. The directional nature of forces allows us to capture the distinction between *aan* on the one hand from *op* and *tegen* on the other hand, but it also explains why prepositions cooccur with push and pull verbs in the way they do.

Interestingly, the directional nature of forces has a direct reflex in English in the the use of the directional preposition *from* with the verb *hang*:

(22) The lamp was hanging from the ceiling

The *from* that usually designates a *path of motion* away from the Ground is used here for a force vector pointing away from the Ground. It is difficult to account for this use if we don't allow forces to have spatial directions.

More properties of support and attachment

What we have seen in the previous section is just the basic core of the force-dynamics of contact prepositions like *on* in English and and *op* and *aan* in Dutch. There are a number of other observations to make about these prepositions.

The first effect is the contact effect: the Figure and the Ground are in contact or spatially contiguous. But note that this is not a spatial condition that is separate from their force-dynamic properties. As we noted already with forceful verbs, spatial contact is necessary for force-dynamic interaction. The Figure and Ground have to touch to allow the configurations in (20) and (21) to obtain in the first place. So, the force-dynamic and spatial components of the relations expressed by prepositions like *on* and *against* are closely tied together.

The second effect, which we already described in section 1, is the chaining effect, a way of extending the contiguity between two objects. The force interaction between objects does not need to be direct, but it can be mediated by a third object. In our schematic representation, we can represent this for *op* and *aan* (support and attachment) as follows:

(23) op: Figure --> X --> Ground (support)
aan: Figure <-- X <-- Ground (attachment)

With *op* the Figure has a pushing relation with X and X with the Ground, with *aan* the Figure is pulling X and X is pulling the Ground. The X can only fulfil its role if it is literally between Figure and Ground, so if it is also a spatial intermediary, which is also what we see in the situations from section 1, repeated here:

(24) (a) Het kopje staat op de tafel
The cup stands on the table
'The cup is standing on the table'

(b) De lamp hangt aan het plafond
The lamps hangs on the ceiling
'The lamp is hanging from the ceiling'

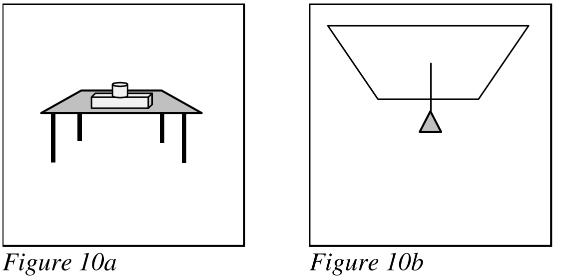


Figure 10 *Figure 10a* *Figure 10b*

The book in Figure 10a is between the table and the cup, just as the cable in Figure 10b is between the ceiling and the lamp.

There is a third effect that usually occurs with *aan* and *op*. We can call it a default effect, because it concerns the prototypical use of these prepositions. Again, we need to refer to the spatial direction of forces to account for this effect. Unless otherwise specified by the context or the sentence, we assume that *aan* (attachment) applies in a situation in which the force vector is downward, because of gravitation. *Aan* is not just 'pulling', it is downward pulling, i.e. 'hanging'. This is especially the case if the sentence does not have an explicit Agent. Also with *op* (support) the default is downward, as a result of the gravitational pull. So, this is what we get in prototypical situations:

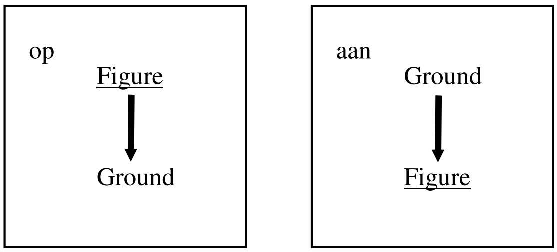


Figure 11 *Figure 11a* *Figure 11b*

This also implies that in the prototypical case, *aan* (attachment) implies 'under', while *op* (support) implies 'above'. These spatial relations follow again from the force-dynamic specifications. However, it is not difficult to find situations in which the force relations hold in a different direction, especially with *on*/*op*:

(25) (a) The fly is sitting on the wall
(a′) De vlieg zit op de muur

(b) The fly is sitting on the ceiling
(b′) De vlieg zit op het plafond

Finally, with *aan* and *op*, we get what we might call stative effects: situations in which the force that the Figure exerts on the Ground is counterbalanced by an equal but oppositely directed force exerted by the Ground. We normally interpret the sentences in (18) as referring to situations of stasis, similar to what we saw with *lean* and *hang*:

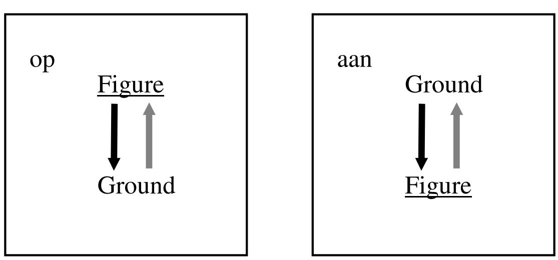


Figure 12 *Figure 12a* *Figure 12b*

In general, in such a situation of stasis, the force vectors $\mathbf{f}_{\text{Figure}}$ and $\mathbf{f}_{\text{Ground}}$ are opposite and of equal length, i.e. $\mathbf{f}_{\text{Ground}} = -\mathbf{f}_{\text{Figure}}$, or, in other words: $\mathbf{f}_{\text{Ground}} + \mathbf{f}_{\text{Figure}} = \mathbf{0}$, where $\mathbf{0}$ is the zero force vector.[2]

Stasis is not necessary, however. There can also be situations with *on*, *op* and *aan* in which the counterforce is non-existent or such that no balance results:

(26) (a) Alex trok aan de wagen
Alex pulled on the car
'Alex pulled the car'

(b') Alex drukte op de bel
Alex pressed on the bell
'Alex pressed the bell'

These Dutch sentences don't specify what happens with the car and the bell. This depends on particulars of the situation and properties of these objects.

Containment as a force-relation

We have finally come now to the most common and at the same time most complicated preposition of Dutch and English: *in*. Vandeloise (1991) and others have argued that the semantics of this preposition should be understood in terms of containment. Given what we know now about force vectors, how can we capture containment in these terms, such that the phenomena in section 1 are accounted for?

The idea is to take our inspiration again from what we see with verbs. We take *in* to share important force-dynamic characteristics with *squeeze*. We have proposed in section 2 to treat *squeeze* as a configuration in which there is concavity of forces: the Patient is between (parts of) the Agent and the Agent's forces are pointing towards the Patient. I propose to represent *in* in a similar way, but since we are talking about prepositional relations, I will use Figure and Ground:

(27) in: <u>Ground</u> –> Figure <– <u>Ground</u>

This is like a minimal configuration, which says that the Ground exerts forces on the Figure from at least two opposite sides. Of course, the Ground might enclose the Figure on all sides (and maybe this is even true for typical containment), but for the time being I will assume that containment minimally requires what we see in (27). Notice that a kind of spatial inclusion follows from this force-dynamic configuration. The forces of the Ground can only come from different sides if the Ground somehow spatially includes the Figure. Just as with *on*, *aan* and *op*, we see an intimate connection between forces and locations, made possible by the way force vectors are embedded in space.

Obviously, there are important differences between *squeeze* and *in*. The verb *squeeze* involves active and dynamic exertion of forces from at least two opposite sides, involving close contact, typically by an animate Agent. The preposition *in* involves a passive and stative configuration of forces, not necessarily involving contact, typically by an inanimate Ground. I believe that many of these differences correlate with the fact that *squeeze* is a verb, while *in* is a preposition. Nevertheless, the two words both take part in an abstract force-dynamic schema.

The configuration in (27) gives a basic condition for containment. What we see are only two *parts* of the Ground, on opposite sides of the Figure. In a sense, (27) gives us a one-dimensional cross-section of a two-dimensional situation in which the Ground is a ring around the Figure or of a three-dimensional situation in which the Ground is all around the Figure.

Even though (27) is very rudimentary, it does give us a way to capture what goes on in the following two situations, both describable by *the black marble in the bowl*:

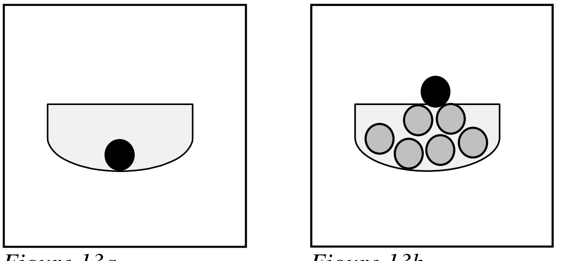


Figure 13 *Figure 13a* *Figure 13b*

Figure 13a is the simple situation in which there are two forces of the Ground pointing from opposite sides towards the Figure, as in (27). However, in Figure 13b, there is *chaining*:

(28) in:

Figure
↓
<u>Ground</u> –> X <– <u>Ground</u>

There are force vectors pointing from the Ground to X, but there is also a force vector connecting X to the Figure. The force-dynamics of containment by the Ground is transmitted here through an object X to the Figure. What is interesting here is that the chaining is not homogeneous. The force-dynamic relation between the black marble in Figure 13b (the Figure in (28)) and the other marbles (X in (28)) is not itself a relation of containment, but rather one of *support*, it seems. But because the grey marbles contained in the bowl support the black marble this marble is also indirectly contained in the bowl.

This is only one simple example and it is not clear what will happen with this primitive model of prepositional force-dynamics when we confront it with the diversity of uses of topological prepositions like *in* and *on*. Nevertheless, as semanticists we should go beyond simple descriptive labels like 'containment' and 'support' and look for the system behind these relations. Modeling such a system will allow us to generate testable hypotheses about the role that containment and support play in the semantics of prepositions. We would predict, for example, that *in* can also be used in a situation where Ground and Figure are related through attachment:

(29) in: Ground –> X <– Ground
↓
Figure

Whether this is the case remains to be seen, but it illustrates an important point. With a general unanalyzed notion of containment it remains unclear what is possible and impossible. A model of force-dynamic relations in which we can manipulate parameters is more adequate from a semantic point of view.

4 Conclusion

In this paper I have shown that verbs and prepositions are based on the same 'geometry' of forces. The notion of geometry can be taken quite literally, because forces are represented as vectors with a direction in space. This is essential for providing the interface between probably the two most basic components of natural language semantics: force-dynamics and space.

One avenue to explore is how this model can help us to model the typological results of Bowerman and Choi (2001:485). They show that there is a universal hierarchy of topological static relations that ranges from a typical instance of support (cup on table) to containment (apple in bowl), with less typical relations in between:

(30)

cup on table	bandaid on leg	picture on wall	handle on door	apple on twig	apple in bowl
<––––––––––––––––––– ON –––––––––––––––––––––>					<–IN–>
<––––––––OP––––––><–––––––––––AAN––––––––––––>					<–IN–>
<–––––––––––––––––––––––––EN–––––––––––––––––––––––>					

Languages carve up this scale in different ways, but terms always correspond to *continuous* regions. If a language uses a term X for two situations then it also uses it for every situation in between. This is illustrated in (30) for English, Dutch and Spanish, respectively. This continuity property is strongly related to the property of *convexity* that Gärdenfors (2000) proposed as a constraint on regions in conceptual spaces, but also to the notion of *connectivity* in the semantic map approach in typology (Haspelmath 2003).

My point here concerns not so much the nature of this general property, but rather the way we could use the force-dynamic schemas proposed here to give us more insight into the conceptual space underlying prepositions like *in* and *on* in various languages, in other words, in the conceptual space of containment and support. If we compare our representations for Dutch *op*, *aan* and *in*, we can see the beginnings of a way to model the hierarchy in (30).

(31)	op:	Figure	-->	Ground		
	aan:	Figure	<--	Ground		
	in:	Ground	->	Figure	<-	Ground

There are different parameters here that can be manipulated: whether the Figure or the Ground is the agentive participant, whether the force vector is directed towards the Ground or towards the Figure and additionally, whether the force vector is typical downward (as with *op*) or not. Another parameter is whether the force is simplex (with *op* and *aan*) or complex (with *in*). In this way, we might hope to get a scale in which (the prototypes of) *op* and *in* are maximally distinct with a gradient in between, in which the parameters change from *op* to *in*:

(32)		*op*		*in*
	Force source:	Figure	....	Ground
	Force orientation:	Ground	....	Figure
	Force direction:	Down	....	Not down
	Force complexity:	Simplex	....	Complex

In this way the analysis of forceful prepositions proposed here is not only relevant for English and Dutch, but for all languages across the world that refer to notions of containment, support and attachment.

If the approach of this paper is on the right track, then it also sheds an interesting light on two common and influential ideas in the literature on topological prepositions like *in* and *on*, which go back to work of Herskovits (1986) and Vandeloise (1991). One idea is that the semantics of *in* and *on* is based on a particular type of geometry, namely the *topological* one, in which basic relations between spatial regions play a role (as opposed to the axis-based semantics of *projective* prepositions like *above* and *behind*). *In* corresponds to 'inclusion' while *on* corresponds to 'contiguity' or 'connectedness'. Vandeloise came with an alternative, non-geometric approach based on functional or force-dynamic notions like 'containment' and 'support'. The results of this paper suggest, however, that geometry vs. function may be a false dichotomy. Spatial geometry and force-dynamics are not mutually exclusive, but they are both based on a more fundamental notion of vector, which makes it possible to take a more unified approach towards these conceptual domains.

Notes

1 This paper was presented at the ICLC 9 in Seoul, July 22, 2005. The research for this paper was financially supported by a grant from the Netherlands Organization for Scientific Research NWO to the PIONIER project 'Case Cross-Linguistically' (number 220–70–003), which is gratefully acknowledged. The comments of an anonymous reviewer have been very helpful for me in revising the paper.

2 This zero *force* vector should be kept distinct with the zero *spatial* vector that might potentially be used to represent the purely spatial relation of *contact* between Figure and Ground. However, as argued in Zwarts and Winter (2000), there are several reasons to analyze the spatial contact relation of *on* in terms of non-zero vectors.

References

Beliën, M. (2002) Force dynamics in static prepositions: Dutch *aan*, *op*, and *tegen*. In H. Cuyckens and G. Radden (eds) *Perspectives on Prepositions* 195–209. Tübingen: Niemeyer.

Bowerman, M. and Choi, S. (2001) Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman and S.C. Levinson (eds) *Language Acquisition and Conceptual Development* 475–511. Cambridge: Cambridge University Press.

Carlson, L. and van der Zee, E. (eds) (2005) *Functional Features in Language and Space: Insights from Perception, Categorization, and Development*. Oxford: Oxford University Press.

Coventry, K.R. and Garrod, S. (2004) *Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions.* Psychology Press, Hove.

Croft, W. (1991) *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

Croft, W. (2009) Aspectual and causal structure in event representations. In V. Gathercole (ed.) *Routes to Language Development: In Honor of Melissa Bowerman* 139–166. Mahwah: New Jersey: Lawrence Erlbaum Associates.

Cruse, D.A. (1986) *Lexical Semantics*. Cambridge: Cambridge University Press.

Dowty, D. (1991) Thematic proto-roles and argument selection. *Language* 67(3): 547–619.

Gärdenfors, P. (2000) *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.

Gruber, J.S. (1976) *Lexical Structures in Syntax and Semantics*. Amsterdam: North-Holland.

Haspelmath, M. (2003) The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (ed.) *The New Psychology of Language* 211–243. (Vol. 2) New York: Erlbaum.

Herskovits, A. (1986) *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Cambridge: Cambridge University Press.

Jackendoff, R. (1987) The status of thematic relations in linguistic theory. *Linguistic Inquiry* 17: 369–411.

Jackendoff, R. (1993) The combinatorial structure of thought: The family of causative concepts. In E. Reuland and W. Abraham (eds) *Knowledge and Language, Volume II: Lexical and Conceptual Structure* 31–49. Dordrecht: Kluwer Academic Press.

Johnson, M. (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.

Langacker, R.W. (1987) *Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, R.W. (1991) *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.

O'Keefe, J. (1996) The spatial prepositions in English, vector grammar, and the cognitive map theory. In P. Bloom et al. (eds) *Language and Space* 277–316. Cambridge, Mass: MIT Press.

Talmy, L. (1983) How language structures space. In H. Pick and L. Acredolo (eds) *Spatial Orientation: Theory, Research, and Application* 225–282. New York: Plenum Press.

Talmy, L. (1985) Force dynamics in language and thought. In *Papers from the Twenty-first Regional Meeting of the Chicago Linguistic Society* 293–337. Chicago: University of Chicago.

Vandeloise, C. (1991) *Spatial Prepositions: A Case Study from French*. Chicago: University of Chicago Press.

Wolff, P. and Zettergren, M. (2002) A vector model of causal meaning. In W. D. Gray and C.D. Schunn (eds) *Proceedings of the 24th Annual Conference of the Cognitive Science Society* 944–949. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Zwarts, J. (1997) Vectors as relative positions: A compositional semantics of modified PPs. *Journal of Semantics* 14: 57–86.

Zwarts, J. and Winter, Y. (2000) Vector space semantics: A model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information* 9: 169–211.

9 From the spatial to the non-spatial: the 'state' lexical concepts of *in*, *on* and *at*

Vyvyan Evans

1 Introduction

This paper is concerned with modelling the lexical representation of spatial relations, particularly as encoded by English prepositions, and examining how these spatial relations give rise to non-spatial meanings. In previous work Andrea Tyler and I (Evans and Tyler 2004a, 2004b; Tyler and Evans 2001, 2003) modelled the extensive polysemy exhibited by prepositions, and sought to provide a principled framework for characterising their distinct sense-units. We also sought to establish boundaries between senses as they inhere in semantic memory. In so doing, we attempted to account for this polysemy in a motivated way, as an outcome of situated language use, the nature of human socio-physical experience and the relevant cognitive mechanisms and processes.

Nevertheless, the framework of *Principled Polysemy* we developed was not primarily concerned with modelling the complexity of the spatio-geometric and functional semantic properties, and the extremely complex functional knowledge that prepositional sense-units assist in conveying. This follows as it was primarily concerned with addressing perceived methodological weaknesses in early work in cognitive lexical semantics, as exemplified by the work of Brugman and Lakoff (Brugman [1981] 1988; Brugman and Lakoff 1988; Lakoff 1987). In particular, it is becoming clear that Tyler and I, in our work on Principled Polysemy, may, in fact, have underestimated the functional complexity that 'spatial' prepositional sense-units encode.

Accordingly, the goal of this paper is to present a more recent theory of lexical representation which builds on and refines the framework of Principled Polysemy. This approach, I argue, better accounts for some of the complexities I will be describing with respect to the sorts of knowledge structures that prepositions encode, as evidenced in language use. Following Evans (2004a 2004b; see also Evans 2006, 2009), this theory employs two central constructs: the *lexical concept*, and the *cognitive model*. In brief a lexical concept is a relatively complex sense-unit which is conventionally associated with a specific form. Moreover, certain kinds of lexical concepts afford access to large-scale multi-modal knowledge structures: cognitive models. Cognitive models constitute relatively stable, non-linguistic knowledge structures, which are subject to ongoing modification as we continue to interact in the world and in communicative settings. Moreover, cognitive models provide the complex informational characterisation lexical concepts invoke in meaning construction processes. As the constructs of the lexical concept and the cognitive model are of central importance, the theory of lexical representation to be presented is termed the theory of *Lexical Concepts and Cognitive Models*, or LCCM Theory for short. The theoretical discussion presented later in the paper is based on more detailed explications of LCCM Theory (Evans 2006, 2009).

The main analytical focus of the paper is the so-called 'state' senses of English prepositions, as associated with prepositions such as *in*, *at*, and *on*. While these sense-units presumably derive from, and are certainly related to 'spatial' senses encoded by the same forms, they are not, in and of themselves, primarily spatial in nature. Representative examples are provided below.

(1)	We are in love/shock/pain	'state' sense
	cf. We are in a room	'spatial' sense
(2)	We are at war/variance/one/dagger's drawn/loggerheads	'state' sense
	cf. We are at the bus stop	'spatial' sense
(3)	We are on alert/best behaviour/look-out/the run	'state' sense
	cf. We are on the bus	'spatial' sense

In these examples, *in*, *at* and *on* mediate a relation between human experiencer(s) and a particular state. While some of these expressions, for instance, to be 'at daggers drawn' are clearly idiomatic, the contention of cognitive lexical semantics is that while such expressions may be highly conventionalised, and the source of the idiom may not be accessible to contemporary language users, the fact that *at* is employed is, diachronically at least, motivated (see Evans and Green 2006: chapter 10, for a review; see also Evans, Bergen and Zinken 2007).

If the perspective offered by cognitive semantics is correct, namely that the use of *in*, *at* and *on* to encode a 'state' meaning is motivated, deriving from historically earlier, and synchronically, perhaps, more primary 'spatial' senses, then there are a number of issues which await explanation. Firstly, how do we account for the derivation of non-spatial, what we might dub 'abstract' senses from historically earlier spatial senses? One solution to this problem has been to posit underlying conceptual metaphors as the solution (Lakoff and Johnson 1999). That is, due to the conceptual metaphor, *qua* sub-symbolic knowledge structure, of the sort glossed as STATES ARE LOCATIONS, states of the type captured in (1) to (3) inclusive are conceptualised as locations. On the metaphor account, the existence of an independently motivated conceptual metaphor licenses the development of new polysemous senses associated with *in*, *at* and *on*.

Despite the intuitive appeal of the conceptual metaphor account, this cannot be the whole story. After all, each of the 'state' senses associated with the prepositions evident in (1)-(3) exhibit distinct patterns in terms of the semantic arguments with which they collocate. Put another way, the 'state' senses associated with the different prepositional forms: *in*, *on* and *at*, are not equivalent. For instance, the 'state' sense associated with *in* relates to semantic arguments which have to do with emotional or psychological 'force' such as being 'in love', 'in pain' and so on. In contrast, the semantic arguments associated with *at* have to do, not with emotional force but, rather, with mutual (or interpersonal) relations, such as being 'at war'. Meanwhile, *on* relates to semantic arguments that have to do with time-restricted activities and actions which involve being currently active in some sense. These include being 'on alert', 'on duty', and so forth. That is, the semantic arguments associated with each of the 'state' senses for these prepositions is of a quite different kind. This suggests that the 'state' meanings conventionally associated with each of these prepositional forms is also of a distinct kind. While this does not preclude a conceptual metaphor account as part of the story, positing a unified metaphoric account for examples of the kind provided in (1) to (3) does not, in itself, adequately account for the linguistic facts.

The challenge, then, for a theory of lexical representation, which assumes that the 'state' sense-units are motivated and related, is to account for the fact that i) each of these prepositions exhibits a conventional 'state' lexical concept, and ii) that each of the 'state' lexical concepts diverges. Put another way, we must account for the differential motivation that gives rise to the similar, yet distinct, 'state' lexical concepts associated with each of these prepositions. Thus, the 'state' lexical concepts present an intriguing challenge which, I shall argue, existing theories of lexical representation, notably the theory of Principled Polysemy, cannot, at present provide an account for. For this reason, we require a more sophisticated account of lexical representation.

I will employ linguistic data associated with these 'state' lexical concepts in order to provide a reasonably detailed illustration of how LCCM Theory accounts for the functional complexity of the semantics involved. I argue that LCCM Theory facilitates i) a revealing descriptive analysis of the 'state' lexical concepts of these prepositions, including the way in which these sense-units are in fact distinct from one another; and ii) a revealing account of the spatio-geometric and functional knowledge that the core 'spatial' lexical concepts associated with *in*, *at* and *on* encode; and finally, in view of this, iii) a revealing account of how each of the 'state' lexical concepts involved is motivated by, and related to, the core 'spatial' lexical concepts associated with each preposition.

A further reason for selecting the 'state' lexical concepts as a case study is as follows. While there is now a voluminous literature on spatial semantics, especially within cognitive lexical semantics, this work has primarily been concerned with examining the range of distinct sense-units associated with a given preposition, including a now impressive body of research which has focused on principles for determining sense-boundaries, including psycholinguistic and corpus-based approaches (e.g., Sandra and Rice 1995 and Gries 2005 and the references therein). However, hitherto, there has been, in relative terms, comparatively little research on the non-spatial lexical concepts associated with prepositional forms, and how they are related to one another and derived from spatial lexical concepts. This lack of research makes an examination of the 'state' lexical concepts of different prepositions an issue worthy of attention.

There are two claims that I make, and which the findings presented serve to substantiate. Firstly, 'new' lexical concepts derive from already extant lexical concepts by virtue of inferential processes, relating to situated instances of language use. Hopper and Traugott (1993) refer to such a mechanism as *pragmatic strengthening*: an inferential process whereby a new semantic representation is abstracted from an extant semantic representation in what has been referred to as a *bridging context* (N. Evans and Wilkins 2000). A bridging context is a context of use in which the new lexical concept emerges as a situated inference (or an 'invited inference', Traugott and Dasher 2004). A polysemous relationship thereby holds between the extant and the derived lexical concept. I argue that the polysemous lexical concepts associated with the prepositional forms to be examined arise due to new *parameters* being encoded, giving rise to distinct lexical concepts. These parameters arise due to the functional consequences of spatio-geometric properties in situated language use, about which I shall have more to say below.

The second claim is as follows. The 'state' lexical concepts for each prepositional form are distinct, as revealed by an examination of their *lexical profiles*: the semantic and grammatical selectional tendencies exhibited. Moreover, each form has a number of conventional 'state' lexical concepts associated with it, which are different from one another. Put another way, there are clear differences in terms of 'state' lexical concepts both across and within the prepositions I address here.

2 The functional nature of the spatial semantics of prepositions

The point of departure for this study relates to the functional nature of the semantics associated with spatial relations as lexicalised by prepositions. Recent work in the framework of cognitive semantics (e.g., Herskovits 1986, 1988; Vandeloise 1991, 1994) has shown that the received or traditional view is descriptively inadequate in terms of accounting for how the core, prototypical or ideal 'spatial' sense-units associated with prepositions are actually used. The received view, which following Herskovits I refer to as the *simple relations model*, holds that the prototypical sense-unit associated with a given preposition straightforwardly encodes purely spatio-geometric properties, i.e., 'simple' relations.

My purpose in this section is to make the case for a functional characterisation of the 'spatial' lexical concept associated with a given preposition. By 'functional' I mean the following. To understand how language users employ the core 'spatial' lexical concept of a preposition we must also allow for non-spatial parameters which form part of the linguistic content encoded by the lexical concept. The use of the term 'functional' is motivated by the observation that such non-spatial parameters are a functional consequence of humanly relevant interactions with the spatio-geometric properties in question. Moreover, the way 'spatial' lexical concepts are ordinarily employed by language users would appear to require such a functional understanding if 'spatial' lexical concepts are to be correctly interpreted in context.

Providing a functional account is of further importance as the derived lexical concepts which result from sense-extensions (such as the 'state' lexical concepts of *in*, *on* and *at*), cannot be adequately accounted for without first recognising that in addition to spatio-geometric parameters, the core 'spatial' lexical concept associated with a prepositional form also includes functional information. That is, if we assume that the derived lexical concepts are motivated by the prototypical lexical concept, as is the case in cognitive lexical semantics, then we must assume a relatively complex (albeit schematic) body of 'functional' knowledge, if we are to account for the derivation of extended lexical concepts.

In this section, therefore, I briefly review some of the arguments made by Herskovits, and Vandeloise for thinking that functional information also constitutes part of the linguistic content associated with 'spatial' lexical concepts for prepositions (see also Coventry and Garrod 2004; Deane 2005, and Feist This volume for a related perspective).

I begin with Herskovits. In her work she observes that the received view has assumed that the 'basic' function of the spatial sense-units associated with prepositional forms is to encode purely spatial relations. On this view, the semantic contribution of any given spatial use of a preposition relates to spatio-geometric properties, typically designating a relation involving notions such as dimensions, axes or proximity (e.g., Bennett 1975; Miller and Johnson-Laird 1976 for representative examples).

This general approach, particularly as has been evident in formal and computational accounts of prepositions, as noted above, Herskovits (e.g., 1988) refers to as the simple relations model. Yet, as Herskovits shows in detail, the simple relations model is descriptively inadequate. That is, the 'simple' spatial relations posited are unable to account for the range of spatial representations that prepositions ordinarily designate. Some of the descriptive shortcomings of the simple relations model relate to phenomena such as the following.

Firstly, the same preposition often appears to include quite distinct geometric descriptions:

(4) a. the water in the vase
    b. the crack in the vase

The example in (4a) relates to an entity: *the water*, the trajector (TR), 'contained' by the landmark (LM), *the vase*. That is, it relates to the volumetric interior of the LM. In contrast, in (4b) the semantic contribution of *in* concerns a relation between a 'negative' region, namely a lack of substance, *a crack*, which is not part of the volumetric interior of the vase, but rather forms part of the landmark-boundary, namely the physical structure of the vase. Put another way, *in* relates to quite distinct spatio-geometric relations in these examples. This is problematic for the simple relations model which assumes that a given preposition encodes a single spatio-geometric relation,

Secondly, the spatial relations encoded by prepositions often appear to diverge from straightforward 'simple' relations. For instance, the following expression:

(5) the dictionary on the table

can be used unproblematically to refer to a dictionary placed on top of another book which is 'on' the table. That is, the dictionary is not actually 'on' the table, but rather 'on' the book which is in direct contact with, and therefore 'on', the table.

Thirdly, there often appears to be what Herskovits refers to as 'added constraints' which apply to prepositions. For instance, in examples of the following kind:

(6) a. the man at the desk
b. the schoolboy at the bus-stop

the relation implied is more specific than 'simple' spatio-geometric relations. That is, the example in (6a) implies, and is understood to mean, that not only is the TR in question, *the man*, in close proximity to his desk, but he is also working at his desk (or at least in a position to do so). Similarly, in (6b), in addition to the co-locational relation, this expression implies that the schoolboy is 'waiting' at the bus-stop, presumably for a bus. In other words, part of the meaning of these utterances is functional in nature. The schoolboy is co-located with the bus-stop *in order to* catch a bus. Implications such as these are not explained by the simple relations model. In fact, we seldom employ prepositions simply to describe a purely spatio-geometric relationship.

Fourthly, there are often unexplained *context dependencies* associated with prepositions which the simple relations model fails to account for. In an example such as the following:

(7) Max is at the crèche

this utterance appears only to work when both speaker and addressee are not also present at the crèche. In the case when the speaker and addressee are located at the crèche, the following would be more likely:

(8) Max is (somewhere) in the crèche

Finally, there are a number of other restrictions which appear to relate to discursive salience and/or relevance. Again, these are not accounted for by the simple relations model. For instance, in a scenario such as that represented by Figure 1, in which there is an apple located beneath an upturned bowl, the following expression is semantically anomalous:

(9) #the apple in the bowl

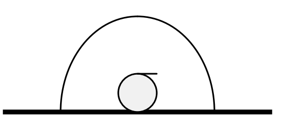


Figure 1. The apple beneath the bowl

Herskovits argues that in view of the failure of the simple relations approach a modified view of the lexical representation for spatial prepositions is required.

A related perspective has been presented by Vandeloise in his work. Vandeloise (1991, 1994) argues compellingly that any account of spatial semantics that leaves out the functional nature of prepositional lexical concepts fails to properly account for how they are actually employed. That is, spatio-geometric relations have functional consequences, consequences which arise from how we interact with objects and entities in our physical environment, and in our daily lives. To illustrate, take the mundane example of a cup of coffee. Imagine holding it in your hand. If you move the cup slowly up and down, or from side to side, the coffee moves along with the cup. This follows as the cup is a container with a bottom and sides and thus constrains the location of any entity within these boundaries. Tyler and I (2003) referred to this property of bounded landmarks as 'location with surety'.

The force-dynamic properties associated with a cup as a container also show up in linguistic content, as illustrated by the semantic contribution of the preposition *in*. Consider the diagram in Figure 2 drawn from the work of Vandeloise (1994).

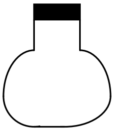


Figure 2. A bottle or a light bulb? (adapted from Vandeloise 1994)

Vandeloise observes that the image depicted in Figure 2 could either represent a bottle or a light bulb. As example (10) shows, we can use the preposition *in* to describe the relation between *the light bulb* (TR) and *the socket* (LM).

(10) The bulb is in the socket

In contrast however, we cannot use *in* to describe the relation between a bottle and its cap, as illustrated by (11).

(11) #The bottle is in the cap

Vandeloise points out that the spatial relation holding between the TR and LM in each of these utterances is identical, and yet while (10) is a perfectly acceptable sentence (11) is semantically odd. Vandeloise suggests that it is not the spatial relation holding between the TR and LM that accounts for the acceptability or otherwise of *in*. He argues that the relevant factor is one of force-dynamics: '[W]hile the socket exerts a force on the bulb and determines its position, the opposite occurs with the cap and the bottle' (Vandeloise, 1994: 173). In other words, not only is the position and the successful function of the bulb contingent on being *in* (contained by) the socket, but the socket also prevents the bulb from succumbing to the force of gravity and falling to the ground. In contrast, the position and successful functioning of the bottle is not contingent on being *in* the cap. This suggests that our knowledge of the functional consequences associated with located containment affects the contextual acceptability of a preposition such as *in*.

3 Principled polysemy revisited

Having begun to consider the functional nature of the spatial semantics of prepositions, I now reconsider the model of Principled Polysemy as an account of spatial semantics. In developing this model Tyler and I (e.g., Tyler and Evans 2001, 2003) sought to model the nature of the lexical representations associated with spatial particles such as prepositions. In so doing we were concerned with two sorts of issues. Firstly, we were concerned with accurately describing the nature and range of the distinct (albeit related) lexical concepts (what we referred to as 'senses') associated with lexical categories such as prepositions. That is, we were concerned with providing a constrained (i.e., principled) methodology for establishing sense-units and thus sense-boundaries.

Secondly, we were concerned with accounting for how sense-units (lexical concepts in present terms) arise. We posited that the lexical concepts which populate a semantic network for a given preposition are diachronically related, and the derivation of 'new' lexical concepts (i.e., sense-extension) is motivated (see Evans and Tyler 2004a in particular). Both these issues required detailed analysis of the lexical representations associated with the various lexical concepts for a given preposition. Moreover, this in turn entailed examination of spatio-geometric, and non-spatio-geometric, aspects of prepositional lexical concepts.

For instance, while an important part of the semantic representation for *over* in (12) has to do with the spatio-geometric relationship holding between the TR and the LM, in (13) an important part of the lexical representation relates to non-spatio-geometric aspects, i.e., occlusion.

(12) The picture is over the sofa

(13) The veil is over her face

In (12) the semantic contribution of *over* relates to an 'above' relation, which concerns the spatio-geometric relationship in a 3-dimensional region holding between the TR and LM. In (13), while part of the linguistic content of *over* must also encode spatio-geometric information – as occlusion is a consequence of a physical relationship holding between artefacts and the vantage point of a perceiver from which the artefacts are viewed – nevertheless, the semantic contribution of *over* is more saliently identifiable as the functional notion of 'occlusion'. Examples such as this, in which *over* is not interpreted as providing a semantic contribution relating to 'above' but rather 'occlusion', provide good evidence that the occurrence of *over* in (13) is *sanctioned* by a distinct lexical concept: we are dealing with a lexical concept which is distinct vis-à-vis the 'above' lexical concept which sanctions the use of *over* in (12).

In our analyses, Tyler and I made the point that functional lexical concepts such as what we referred to as the Covering Sense of *over* (i.e., the [OCCLUSION] lexical concept in present terms), obtain because spatial experience is inherently meaningful for humans. That is, as human beings we interact with objects around us in our spatial environment (see Johnson 1987; 2007). Particular spatial relations, as manifested by the linguistic content encoded by prepositional lexical concepts, have functional consequences. These functional consequences we described as arising from *experiential correlations*, an idea we borrowed, and adapted from the work of Grady (1997).

For instance, a consequence of the spatio-geometric property associated with *over* in examples such as (12), i.e., an 'above' relation, is that in certain contexts, occlusion occurs. To illustrate consider (14):

(14) The tablecloth is over the table

In this example, the use of *over* is sanctioned by a lexical concept that encodes a spatio-geometric relation in which the TR is in an 'above' relation with respect to the LM. However, a functional consequence of how we interact with TRs such as tablecloths, and LMs such as tables, and given the dimensions of tablecloths, such that they often have a greater extension than tables, is that by virtue of being over (i.e., above), the tablecloth thereby occludes the table. Thus, we argued that due to such contexts of use, *over* can, by virtue of the process of reanalysis termed pragmatic strengthening (as briefly introduced above), lead to the 'occlusion' reading becoming 'detached' from the context in which it grounded, and reanalysed as a distinct sense-unit of *over* in its own right.

A related idea that was important in the Principled Polysemy framework was the notion of a *functional element*, an idea inspired by the work of Vandeloise (e.g., 1991, 1994) in his functional approach to spatial semantics. This notion related to the central or core sense in a semantic polysemy network. Such lexical concepts we termed *proto-scenes*. The proto-scene for *over*, what we termed the Above Sense as exemplified in (12), constitutes an abstraction over spatio-geometric properties associated with the range of spatial scenes in which a given preposition, such as *over*, is used.

However, as already noted, a large part, perhaps the majority, of uses of the proto-scene of a given prepositional form relate to usages which are not purely or even wholly spatio-geometric in nature (see Vandeloise 1991, 1994 and especially Herskovits 1986, 1988 as described above). Thus, Tyler and I argued that functional information forms part of the semantic representation of any given proto-scene (see Evans and Tyler 2004a; Tyler and Evans 2003 for details).

In sum, Principled Polysemy posits two kinds of lexical concept which populate a prepositional polysemy network. The first kind, the proto-scene, is primarily spatio-geometric in nature. Moreover, the proto-scene corresponds – for most of the prepositions we surveyed – to the historically earliest lexical concept associated with a given prepositional form (Tyler and Evans 2003). Nevertheless, proto-scenes include a functional element, reflecting the way in which proto-scenes are ordinarily used. That is, language users typically employ proto-scenes in ways which draw upon the functional consequence of interacting with spatial scenes of certain kinds in humanly relevant ways. Thus, linguistic knowledge associated with proto-scenes appears to involve more than simply knowing the particular spatio-geometric properties encoded by a particular form.

The second sort of lexical concept – the remainder of the senses in a prepositional polysemy network – we hypothesised as being motivated by, and ultimately derived from, the proto-scene. This said, we observed that the derivation is often complex and indirect (see Tyler and Evans 2003 for detailed discussion). These derived lexical concepts we referred to as *sense-extensions*. These 'new' lexical concepts, we argued, were derived by virtue of the process of reanalysis (pragmatic strengthening) due to experiential correlations of the sort described above for the development of the Occluding Sense from the Proto-scene (i.e., the Above Sense).

One issue which Tyler and I largely side-stepped, in the version of Principled Polysemy which appeared as Tyler and Evans (2003), concerned how best to account for 'common' lexical concepts of different prepositions, such as the 'state' lexical concepts for *in*, *on* and *at*, illustrated above in the examples in (1) to (3). The difficulty here is that as the 'state' lexical concepts associated with *in*, *at* and *on*, for instance, are all identified by a common label, this might be construed as suggesting that there is common semantic representation. Yet, the 'state' lexical concepts appear, on the contrary, to be distinct sense-units as evidenced by the distinct semantic arguments with which they each collocate: their lexical profiles, in present terms. What is required is a theory of lexical representation which has methodological tools for distinguishing between ostensibly 'similar' lexical concepts associated with different forms.

A further difficulty is that it is unclear, in Principled Polysemy, what the nature of the functional relationship is holding between the lexical representation associated with the proto-scene, and the diverse 'functional' lexical representations associated with the range of derived senses we posited. That is, while Principled Polysemy posited a single functional element associated with each proto-scene, it is not clear how this would motivate the functional complexity apparent in the plethora of functionally diverse extended senses, posited for each prepositional form.

Thus, while an important construct, there is good reason, therefore, to think that the notion of a functional element associated with the proto-scene, as presented in Evans and Tyler (2004b) and Tyler and Evans (2003) actually underestimates the functional complexity that must be readily available to language users, as encoded by the range and various combination of parameters associated with the distinct 'state' lexical concepts across and within prepositions.

Ultimately, the difficulty for the Principled Polysemy framework is that while it attempted to provide a detailed account of lexical representation, because of its primary concern with detailing a rigorous methodology for establishing distinct sense-units, it failed to work out the implications of the functional nature of spatial semantics for lexical representation.[1]

4 The Lexical Concepts and Cognitive Models (LCCM) approach to lexical representation

In recent work (Evans 2006, 2009), I have begun to develop an approach to lexical representation which is consistent with the context-dependent nature of the meanings associated with words. Indeed, part of the focus of this particular research programme is to develop an account of how lexical representations give rise to situated meaning construction, and thus to provide a cognitively-realistic approach to meaning construction. While the issue of situated meaning construction is less relevant to the analysis of how best to represent the 'state' lexical concepts in the present paper, and won't be addressed further, Evans (2006) constitutes an attempt to model lexical representation that is relevant for any lexical class, including prepositions.

The starting point for the LCCM Theory account is the premise that linguistic knowledge is usage-based. That is, I assume that the organisation of our language system is intimately related to, and derives directly from, how language is actually used (Croft 2000; Langacker 2000; Tomasello 2003). Through processes of abstraction and schematisation (Langacker 2000), based on pattern-recognition and intention-reading abilities (Tomasello 2003), language users derive linguistic units. These are relatively well-entrenched mental routines consisting of conventional pairings of form and semantic representation. The semantic representations conventionally associated with a given unit of form, I refer to, as already noted, as a lexical concept.

While lexical concepts are mental representations, they underspecify the range of situated meanings associated with a given form in an individual utterance. Thus, I make a fundamental distinction between lexical concept as a mental unit, and its context-dependent realisation in an utterance. This is akin to the distinction in Phonological Theory between the abstract notion of a phoneme and the actual unit of realised context-dependent sound, the allophone. My claim is that there is an essential distinction between lexical representation and meaning. While meaning is a property of the utterance, lexical representations are the mental abstractions which we infer must be stored as part of the language user's knowledge of language, in order to produce the range of novel uses associated with situated instances of a particular word such as a preposition. The meaning associated with an utterance I refer to as a *conception*. Thus, conceptions are a function of language use.

There are a number of important properties associated with lexical concepts. I briefly review some of the most relevant here (for detailed discussion see Evans 2009). Firstly, and as noted above, linguistic units, as I use the term, are conventional pairings of form and meaning. From this it follows that lexical concepts are *form-specific*. Secondly, as mentioned above, although lexical concepts are form-specific, a single form can be conventionally associated with a potentially large number of distinct lexical concepts which are related to degrees as attested by the phenomenon of polysemy.[2] That is, forms are not *lexical concept-specific*. A consequence of this is that the lexical concepts which share the same form can be modelled in terms of a *semantic network* (see Evans and Green 2006: chapter 10 for discussion).

Thirdly, the definitional property of any given lexical concept is that it has a *lexical profile*, its unique 'biometric' identifier. A lexical profile is an extension of criteria presented in Evans (2004a), and akin to the notion of an 'ID tag' (Atkins 1987) and 'behavioural profile' (Gries 2005). While a lexical concept associated with a particular form can be provided with a semantic gloss, as in the case of lexical concepts associated with *over*, an example of which I glossed as [ABOVE] or the lexical concepts associated with *in*, *at* and *on* to be examined later which I preliminarily gloss as [STATE], whether a particular usage of a form relates to one lexical concept rather than another is a matter of examining the 'selectional tendencies' (the lexical profile) associated with a given usage. While any given usage of a lexical concept will have its own unique collocational pattern, general patterns can be established, and form part of the conventional knowledge associated with a particular lexical concept.

Two sorts of information form a lexical concept's lexical profile. The first relates to semantic selectional tendencies. In Evans (2004a) this was referred to as the Concept Elaboration Criterion. The second relates to formal or grammatical selectional tendencies. In Evans (2004a) I referred to this as the Grammatical Criterion. Gries (2005) has advocated the way in which corpus methodologies can be used to examine the lexical profile associated with a specific lexical concept. For instance, each of the 'state' lexical concepts associated with *in*, *at* and *on* have distinct lexical profiles. In the remainder of this chapter I primarily rely on semantic selectional tendencies for adducing distinct lexical concepts.

To provide a preliminary illustration of the construct of the lexical profile, I briefly consider two lexical concepts, both of which I provisionally gloss as [STATE] – although I revise this gloss later in the chapter – and which are conventionally encoded by the English prepositional forms *in* and *on*. These are evidenced by the following examples:

(15) a. John is in trouble/danger
b. Jane is in love/awe
c. Fred is in shock
d. Jake is in a critical condition

(16) a. The guard is on duty
b. The blouse is on sale
c. The security forces are on red alert

While both *in* and *on* have 'state' lexical concepts conventionally associated with them, the lexical profile for each is distinct. For instance, the [STATE] lexical concept associated with *on* selects semantic arguments which relate to states which normally hold for a limited period of time, and which contrast with salient (normative) states in which the reverse holds. For instance, being 'on duty' contrasts with being off-duty, the normal state of affairs. Equally, being 'on sale' is, in temporal terms, limited. Sales only occur for limited periods of time at specific seasonal periods during the year (e.g., a winter sale). Similarly, being 'on red alert' contrasts with the normal state of affairs in which a lesser security status holds. Further, the states in question can be construed as volitional, in the sense that to be 'on duty/sale/red alert' requires a volitional agent who decides that a particular state will hold and takes the requisite steps in order to bring such a state of affairs about.

In contrast, the semantic arguments selected for by the [STATE] lexical concept for *in* relate to states which do not necessarily hold for a limited period of time, and do not obviously contrast with a 'normal' state of affairs. Moreover, while states encoded by *on* are in some sense volitional, states associated with *in* are, in some sense, non-volitional. That is, we do not usually actively choose to be in love, shock or a critical condition, nor can we, by a conscious act of will, normally bring such states about. That is, these states are those we are affected, constrained and influenced by, rather than those which are actively (in the sense of consciously) chosen.

The fourth and final property of lexical concepts that I review here concerns the position that they have bipartite organisation. That is, lexical concepts encode *linguistic content* and facilitate access to *conceptual content*. Linguistic content represents the form that conceptual structure takes for direct encoding in language, and constitutes what might be thought of as a 'bundle' of distinct knowledge types. There are a large number of different properties encoded by linguistic content which serve to provide a schematic or skeletal representation, which can be encoded in language (for a review see Evans 2009: chapter 6). The one which is relevant for the present study relates to the notion of *parameterisation*.

One way in which knowledge, in general terms, can be represented is in terms of richly inflected nuances that serve to reflect the complexity of experience. An alternative way is to 'compress' such fine distinctions into two, three or more, much broader, and hence, more general distinctions. These I refer to as *parameters*. Linguistic content serves to encode content by adopting the latter strategy, which is to say, to employ parameterisation. Parameters are hence part of the 'bundle' of information that a lexical concept encodes.

To illustrate this notion, consider the complex range of expressions that a language user might employ, in English, in order to 'locate' themselves with respect to time, thereby facilitating time-reference. Any one of the following could conceivably be employed, depending upon context: *today, January, 2008, the day after yesterday, the day before tomorrow, this moment, now, this second, this minute, this hour, today, this week, this month, this quarter, this year, this half century, this century, this period, the 8th day of the month, this era, this millennium*, and so on. A potentially unlimited set of finer and finer distinctions can be made (e.g., *1 second ago, 2 seconds ago, 1 hour 4 minutes and 3 second ago, 2 days ago*, etc.), reflecting any manner of temporal distinction we might care to make.

In contrast, paramaterisation functions by dividing all the possible distinctions relating to a given category, such as time-reference, into a small set of divisions: parameters. Such parameters might distinguish between the past, for instance, and the non-past. Indeed, this is the basis for the tense system in English, as illustrated by the following:

(17) a. He kick**ed** the ball Past
b. He kick**s** the ball Non-past

English encodes just two parameters that relate to Time-reference: Past versus Non-past, as exhibited by the examples in (17), and thus manifests a binary distinction. Some languages, such as French, have three parameters: Past, Present and Future. Some languages have more than three parameters, distinguishing additionally remote past from recent past, for instance. The language with the most parameters for linguistically encoding time-reference is an African language: Bamileke-Dschang with eleven. Crucially, parameters are encoded by specific lexical concepts, and thus form part of the knowledge 'bundle' that constitutes a lexical concept. For instance, the parameter 'past' is encoded by the lexical concept associated with the –*ed* form in (17a). However, other lexical concepts also include the parameter 'past' such as the lexical concepts associated with the following forms: *sang*, *lost*, *went*, etc.

I argue, then, that a key feature of linguistic (as opposed) to conceptual content is that it only encodes knowledge in parametric fashion. This is not to say that conceptual content does not parameterise knowledge. Indeed, parameterisation is simply a highly reductive form of abstraction: it serves to abstract across the complexity exhibited by a particular category. The point, however, is that the parameters encoded by linguistic content serves to 'strip away' most of the differences apparent in the original perceptual experience, thereby reducing it to a highly limited number of parameters.

In addition to encoding linguistic content, a subset of lexical concepts – those conventionally associated with open-class forms (see Evans 2009 for discussion of this), serve as *access sites* to conceptual content. Conceptual content relates to non-linguistic information to which lexical concepts potentially afford access. The potential body of non-linguistic knowledge, what I also refer to as a lexical concept's *semantic potential*, is modelled in terms of a set of cognitive models. I refer to the body of cognitive models, and their relationships as accessed by a given lexical concept, as the *cognitive model profile*. A design feature of language is that it involves a bifurcation of lexical concept types: those which are relatively more schematic in nature, such as those associated with prepositional forms, the subject of the present study, and those which are relatively richer in nature. As I am dealing with lexical concepts associated with closed-class forms in this study, namely prepositions, I will have little more to say about cognitive models in the remainder of this chapter.

5 Two factors in accounting for 'state' lexical concepts: lexical profiles and parameters

In the Principled Polysemy framework the prototypical (i.e., spatial) sense with respect to which a semantic network is structured is a proto-scene. As we saw earlier, proto-scenes have a single functional element associated with them. In LCCM Theory in contrast, lexical representations, and thus proto-scenes, are representationally more complex than this, especially with respect to their functional properties. In this section I briefly reconceptualise the nature of the core lexical concept associated with a prepositional polysemy network in the light of LCCM Theory.

The prototypical semantic representation associated with a preposition, like the other lexical concepts in the prepositional polysemy network, is a lexical concept. As we saw in the previous section, lexical concepts have bipartite organisation: they facilitate access to conceptual content and encode linguistic content. As prepositional lexical concepts are associated with prepositions: closed-class forms, they constitute closed-class lexical concepts. As such, while they encode linguistic content they do not serve as access sites to conceptual content.

There are two aspects of linguistic content that will be relevant for the discussion of the polysemy exhibited by the range of 'state' lexical concepts in this study. The first concerns the lexical profile exhibited by lexical concepts, as manifested by distinct collocational patterns in language use. As we saw earlier in the chapter, two sorts of information form a lexical concept's lexical profile: semantic selectional tendencies, and formal or grammatical selectional tendencies. In this study I employ distinctions in the semantic arguments which, I hypothesise, collocate with distinct 'state' lexical concepts to uncover distinctions in lexical concepts both within and between prepositions.

The second aspect of linguistic content that will be relevant relates to parameterisation. One characteristic that serves to distinguish between lexical concepts, both across prepositions and within a single preposition, relates to the parameters encoded. For instance, the prototypical 'spatial' lexical concept associated with *in*, which I gloss as [enclosure], encodes the parameter Containment, as evidenced by the example in (18). In contrast, the [emotion] lexical concept – one of the 'state' lexical concepts associated with *in* – encodes the parameter Psycho-somatic State, as evidenced in (19), but not the Containment parameter.

(18) The kitten is in the box Parameter: Containment

(19) John is in love Parameter: Psycho-somatic state

That is, the [enclosure] lexical concept in (18) encodes a schematic dimension abstracted from sensory-motor experience in which a TR is contained by the LM. Notice that the relation encoded is highly schematic in nature; it says nothing about whether there is contact or not between the TR and LM as in (20), nor as to whether the TR represents part of the LM or not as in (21):

(20) a. The fly is in the jar (i.e., flying around)
b. The fly is in the jar (i.e., stationary on one interior surface)

(21) There's a crack in the vase

Indeed, the precise spatio-geometric nature of the TR, LM and their relationship is a function of the TR and LM and their possible forms of interaction, rather than the abstract parameter encoded by the [enclosure] lexical concept associated with the prepositional form *in*. This information derives from the semantic potential accessed via the open-class lexical concepts, as mediated by compositional processes (see Evans 2009 for details).

In contrast, the [EMOTION] lexical concept associated with *in* encodes the parameter Psycho-somatic state. This information is highly schematic in nature. That is, the parameter encoded does not determine which sorts of psycho-somatic states can collocate with this lexical concept. This is a function of the lexical profile: knowledge relating to the semantic selectional tendencies associated with this lexical concept, and hence the range of psycho-somatic states which can co-occur with the [EMOTION] lexical concept. Hence, while the parameters encoded by a lexical concept determine the possible range of semantic arguments that can co-occur, the lexical profile, which relates to stored knowledge based on usage-patterns, provides information relating to the range of permissible states which can co-occur with this lexical concept.

6 Functional consequences of parameters

I now consider how the 'state' lexical concepts arise from historically earlier spatial lexical concepts, giving rise to the phenomenon of polysemy. Put another way, polysemy is a consequence of new, or derived lexical concepts emerging, thereby exhibiting a semantic relationship with a synchronically present – albeit diachronically antecedent – lexical concept.

Based on arguments developed in Tyler and Evans (2001, 2003) I argue that the spatio-geometric knowledge, encoded, in present terms, as abstract parameters by the 'spatial' lexical concepts associated with prepositional forms gives rise to the development of non-spatial lexical concepts. In other words, 'state' lexical concepts emerge by virtue of parameters such as that of Psycho-somatic state arising as a functional consequence of spatio-geometric properties, in particular usage contexts. Hence, the emergence of derived lexical concepts is a consequence of the functional consequences of spatio-geometric parameters in a specific context of use. Such contexts of use Tyler and I (2001, 2003) referred to as *spatial scenes*.

For instance, there are a large number of distinct sorts of spatial scenes that involve the prototypical spatial lexical concept: [ENCLOSURE], associated with *in*, and which hence encode the parameter Containment. This follows as different *bounded landmarks* – a landmark which exhibits the structural properties interior, boundary and exterior – have different functions, are employed for different ends and are viewed from different vantage points. For instance, while a playpen, prison cell and a coffee cup all restrict the containee to a specific location, they do so in service of different objectives, respectively: safety, punishment and consumption. Hence, without understanding the functional consequence of being located 'in' a bounded landmark such as a prison (cell), the question in (22) would be uninterpretable:

(22) What are you in for?

After all, *in,* here, does not relate directly to a given spatial relation, but rather to the specific sets of knowledge systems relating to the 'containment' function of prison in a particular society. Thus, in (22), being 'in' relates not purely to containment, a functional consequence of the [enclosure] lexical concept, but rather, and in addition, to punishment, a functional consequence of being contained in enclosures (i.e., bounded landmarks) of a certain kind, i.e., prisons, which occupy a certain position, and fulfil a specified role in the socio-cultural and legal institutions of a particular society.

Now consider a different sort of functional consequence associated with the [enclosure] lexical concept for *in*. One consequence of certain sorts of bounded landmarks is their utility in providing security. This is evident in the scenario involving a very small child in a playpen for instance. But it is also true of bounded landmarks such as safes used to safeguard valuable commodities such as money or jewels. Indeed, a functional consequence of bounded landmarks of this sort is that the contents are occluded. This of course assumes that the vantage point from which the bounded landmark is viewed is exterior with respect to the volumetric interior of the bounded landmark in question, here the safe. Thus, 'containment' or 'location with surety' is a functional consequence of the spatial relation (i.e., the lexical concept) conventionally associated with *in*, i.e., of [enclosure].

The point is, then, that when *in* is employed in any given utterance, the conception which derives will almost certainly always relate to a functional consequence attendant on a specific sort of spatial scene, involving a containment relation, but will do so in service of objectives and consequences specific to the sort of spatial scene in question. Put another way, bounded landmarks are of many different kinds, a consequence of the many different ways in which we interact with, and the complex range of functions to which we put, bounded landmarks.

In terms of the phenomenon of polysemy, which is to say the emergence of derived lexical concepts, it is precisely functional consequences of this sort which give rise to new parameters. Such new parameters become conventionally associated with a lexical form, and hence contribute to the formation of a new lexical concept. The occlusion afforded by certain kinds of bounded landmarks, such as a jeweller's safe, is a consequence of placing valuables in a landmark that serves to protect the commodity in question. Typically, such landmarks are made of materials that serve to occlude the contents, a consequence – rather than the objective – of employing the types of materials used for constructing the safe. This functional consequence has become abstracted from such spatial scenes to give rise to a distinct parameter. This forms part of the linguistic content encoded by a distinct lexical concept. Evidence for this comes from examples of the following sort:

(23) The sun is in

This utterance relates to lack of visibility of the sun, rather than the sun, the TR, being enclosed by a bounded LM of some sort. That is, the functional consequence of certain sorts of containment relations has given rise to a distinct lexical concept which has a Lack of Visibility parameter encoded as part of its linguistic content.

7 Lexical concepts for *in*

In this section I present an LCCM analysis of the 'state' lexical concepts associated with *in*. That is, I argue that there is more than one distinct 'state' lexical concept conventionally associated with the prepositional form *in*. I also show how these 'state' lexical concepts relate to and are motivated by the functional consequences attendant upon the range of spatial scenes which involve usages of *in* sanctioned by the [ENCLOSURE] lexical concept.

7.1 'Spatial' lexical concepts for *in*

As noted above, the central 'spatial' lexical concept associated with *in* I gloss as [ENCLOSURE]. This lexical concept encodes the parameter Containment. This parameter constitutes an abstraction across the spatio-geometric properties associated with bounded landmarks, such as a box, as lexicalised by the example in (18). The key spatio-geometric components associated with a LM such as a box is that it has the structural elements interior, boundary and exterior (see Tyler and Evans 2003: chapter 7 for detailed discussion). There are a diverse range of complex conceptualisations across which the parameter Containment is abstracted. This includes, at the very least, experiences relating to a TR: the entity enclosed, and a bounded landmark which serves to enclose the TR. Bounded landmarks themselves consist of many types even in everyday experience. For instance, a bounded landmark includes an interior, which further subsumes an interior surface, and the volumetric interior bounded by the interior surface. It also subsumes a boundary, which can be rigid, as in a metal safe, or non-rigid, as in a plastic carrier bag. The boundary also has other physical characteristics such as permeability and degrees of opacity. Finally, the bounded landmark has, by definition, an exterior: that region which constitutes the inverse of the volumetric interior. Accordingly, part of the exterior includes the exterior surface.

As observed earlier, due to our interaction with enclosures, *in* is associated with a number of functional consequences. That is, there are a number of identifiably distinct sorts of *functional categories* associated with spatial scenes involving enclosure. These include Location with Surety, Occlusion and Affecting conditions. Bounded landmarks that are specialised for providing a Location with Surety function are known as 'containers'. These can provide a support function by virtue of containing (i.e., holding and restricting) the location of the TR. This was illustrated with the discussion of the light bulb in the socket example earlier. Alternatively, containers can restrict access (and escape), as in the case of prisons, and safes. The second functional category mentioned relates to Occlusion. A consequence of certain bounded landmarks, due to the opacity of the material which forms the boundary, is that the figure located on the volumetric interior is occluded, and hence hidden from view. The third functional category, that of Affecting conditions, relates to the fact that an enclosure provides a delimited environment which thereby affects the TR located on the volumetric interior. For instance, a prisoner held in solitary confinement in a windowless sound-proofed room is thereby subjected to a particular sensory environment, which is a direct consequence of the nature of the bounded landmark in which s/he is located.

I suggest that it is these functional categories, which arise from the spatio-geometric property of Enclosure, that serve to become abstracted as distinct parameters. Put another way, abstracting across different sorts of sense-perceptory experiences, namely the spatio-geometric properties associated with enclosures, gives rise to an Enclosure parameter. Abstracting across re-occurring functional consequences of the spatio-geometric properties associated with enclosure gives rise to further parameters notably Location with Surety, Occlusion and Affecting Conditions. These parameters, which arise from spatial scenes involving enclosure, are diagrammed in Figure 3.

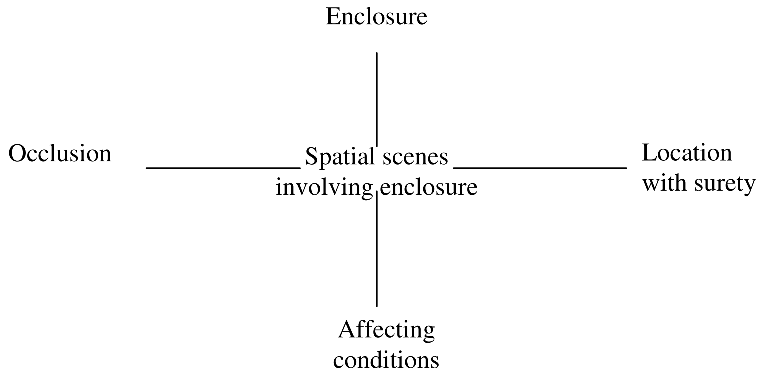


Figure 3. Parameters deriving from spatial scenes involving enclosure

I suggest that the emergence of the parameters: Location with Surety, Occlusion and Affecting Conditions, associated with the linguistic content encoded by *in*, can, under certain conditions, give rise to new 'state' lexical concepts. While the parameter Enclosure, entails, under certain conditions, all of the other parameters illustrated in Figure 3, the other parameters do not necessarily entail the Enclosure parameter. For this reason, as I shall argue, the Enclosure parameter can be seen to be primary; the other parameters arise from spatial scenes in which Enclosure is a key attribute.

The means whereby new lexical concepts arise is due to a disjunction between the various parameters. I illustrate this with the examples below which reveal the disjunction between the Enclosure and Location with Surety parameters.

To do so, consider examples of the following kind:

(24) The toy is in the box

(25) a. The bulb is in the socket
b. The flower is in the vase
c. The umbrella is in his hand

The example in (24) is, I suggest, a consequence of the two parameters: Enclosure and Location with Surety. That is, by virtue of being located in the interior portion of the bounded landmark, the TR is thereby enclosed. Moreover, by virtue of being enclosed, the TR is located with surety: if the box is moved, so also, is the TR – the toy – as a direct consequence. This is what it means to say that Location with Surety is entailed by Enclosure.

Evidence for thinking that the Location with Surety and Enclosure parameters are, nevertheless, distinct units of knowledge encoded as part of a lexical concept's linguistic content comes from spatial scenes involving partial enclosure. In the examples in (25), the TR is only partially enclosed by the bounded landmark: only the base of a bulb is enclosed by the socket as illustrated in Figure 2, above, only the stem, and not the whole flower, is enclosed by the vase (see Figure 4); and only the umbrella handle is enclosed by the hand (see Figure 5). Indeed, the reason that the form *in* can relate to spatial scenes involving partial, as well as full, enclosure is due to the parameter of Location with Surety. It is precisely because the bounded LM that partially encloses the TR serves to provide location with surety that the form *in* is sanctioned in these instances.

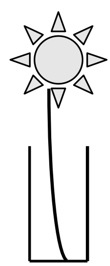


Figure 4. The flower is in the vase

Figure 5. The umbrella is in his hand

On the basis of the examples in (24) and (25), there is no reason, however, to be convinced that Enclosure and Location with Surety constitute distinct parameters, and hence distinct knowledge units encoded as part of the linguistic content associated with the [ENCLOSURE] lexical concept.

However, the example in (26) illustrates a crucial disjunction between the two. While the TR, the bottle, is partially enclosed by the bounded LM, *the cap*, in exactly the same way as the relationship between the bulb and the socket, this use of *in* in (26) is semantically anomalous, as indicated by the hash sign. In the spatial scene described by this example, the bottle is not located with surety by virtue of being partially enclosed by the cap. That is, the bottle's location is not determined by being partially enclosed by the cap – although access to its contents are. Hence, in a situation where partial enclosure applies, but location with surety does not, the [enclosure] lexical concept associated with *in* cannot be applied. This reveals that in the absence of the Location with Surety parameter, *in* cannot be applied to spatial scenes involving only partial enclosure.

(26) #The bottle is in the cap

The examples thus far considered reveal that the Enclosure parameter entails Location with Surety. Hence, in spatial scenes in which there is no location with surety, yet there is (partial) enclosure, as in the spatial scene to which (26) refers, the use of the [enclosure] lexical concept cannot apply, as shown by the semantic unacceptability of (26).

We must next examine whether the Location with Surety parameter can be employed independently of the Enclosure parameter. If so, then we can posit that there is a distinct lexical concept, which we can gloss as [location with surety], a lexical concept which encodes the Location with Surety parameter as part of its linguistic content but does not also feature the Enclosure parameter. Evidence for such a state of affairs is provided by the following example, which relates to the spatial scene depicted in Figure 6.

(27) The pear is in the basket

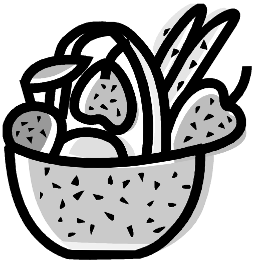


Figure 6. The pear is in the basket

In this example, the pear (in the centre of the image) is not enclosed by the basket, as it is supported by other fruit; although the supporting fruit are enclosed by the basket. Yet, the form *in* can be applied to this spatial scene, as is evident in (27). I argue that this is due to a [location with surety] lexical concept which sanctions this particular usage. While the [enclosure] lexical concept apparent in (24) and (25) encodes the Enclosure and Location with Surety parameters, the [location with surety] lexical concept encodes the Location with Surety parameter but not the Enclosure parameter as part of its linguistic content. This difference in linguistic content between the two lexical concepts explains the difference in linguistic behaviour in the examples just considered. The [enclosure] lexical concept requires full enclosure, or, partial enclosure *plus* location with surety. However, in (27) neither full nor partial enclosure is apparent, yet *in* is sanctioned. This follows as the independent, but semantically related (and hence polysemous) [LOCATION WITH SURETY] lexical concept sanctions this use, I suggest. Thus, we see that there are, plausibly, at least two 'spatial' lexical concepts associated with *in*, [ENCLOSURE] and [LOCATION WITH SURETY], which encode different configurations of parameters, and hence, subtly distinct linguistic content.

7.2 'State' lexical concepts for *in*

I now turn to the 'state' lexical concepts, in order to see how these arise from the spatial lexical concepts. Consider the following examples involving *in*.

(28) a. The cow is in milk
b. The girl is in love
c. John is in trouble/debt
d. He's in banking [i.e., works in the banking industry]

While each relates to a 'state' of some kind, these examples in fact relate to slightly different 'states': those that have a physical cause, as in (28a) – the state of being 'in milk', which is a consequence of the physical production of milk – those that have a psychological or emotional cause, as in (28b) – the state is a consequence of a subjective state, which may (or may not) have physical, i.e., observable, manifestations – those that have a social/inter-personal cause, as in (28c) – resulting from social/interpersonal interactions which result in an externally-maintained state – and those that are a result of a habitual professional activity, as in (28d). Put another way, each of these 'states' take distinct semantic arguments, relating a particular entity to quite different sorts of states. In essence, I argue that these examples are sanctioned by four distinct 'state' lexical concepts for *in*. These distinct 'state' lexical concepts, as we shall see below, I hypothesise to emerge from the functional category Affecting Conditions, which arises from spatial scenes involving enclosure. I spell out the distinctions between the 'state' lexical concepts for *in*, below, with additional examples.

Physiological state (resulting in a 'product')
(29) a. The cow is in milk
b. The cow is in calf
c. The woman is in labour

Psycho-somatic state (i.e., subjective/internal state)
(30) a. John is in shock/pain (over the break-up of the relationship)
b. John is in love (with himself/the girl)

Socio-interpersonal state (i.e., externally-maintained state)
(31) a. The girl is in trouble (with the authorities)
b. John is in debt (to the tune of £1000/to the authorities)

Professional state (i.e., professional activity habitually engaged in)

(32) a. He is in banking

b. She is in insurance

The fact that *in* collocates with semantic arguments of the distinct kinds illustrated in (29–32), relating to physiological, psycho-somatic, socio-interpersonal and professional conditions or properties suggests that we are dealing with four distinct lexical concepts. This follows as LCCM Theory claims that each distinct lexical concept has a unique lexical profile.

In addition to evidence based on semantic selectional tendencies, the position that there must be a number of distinct 'state' lexical concepts associated with *in*, along the lines illustrated by the distinct examples in (29) to (32) inclusive can also be demonstrated by virtue of ambiguities associated with an utterance of the following kind:

(33) She's in milk

The utterance in (33) could potentially be interpreted as relating to a woman who is nursing a baby, and thus lactating, or as relating to a woman who works in the dairy industry. That is, given an appropriate extra-linguistic context, an example such as this can be interpreted in at least two ways. The potential for divergent interpretations is a consequence, in part, of our knowledge that *in* has a number of distinct lexical concepts associated with it: what is relevant for this example is the distinction between a [physiological state] lexical concept and a [professional state] lexical concept. Moreover, ambiguities can be generated even when a relatively well entrenched example is employed. For instance, even examples of the following kind:

(34) She is in labour

(35) He is in love

can be interpreted in alternate ways. For instance, (34) could be interpreted as relating to childbirth or to a professional activity, e.g., the trade union movement. Similarly, (35) could be interpreted as relating to an emotional state or a professional activity, e.g., marriage guidance counselling. The former reading is only possible by virtue of assuming something akin to an [psycho-somatic state] lexical concept which is distinct from a [professional state] lexical concept. That is, both lexical concepts must exist if 'love' can be interpreted in these ways in this example.

7.3 Derivation of the 'state' lexical concepts

I now consider how the 'state' lexical concepts for *in* exemplified in (29) to (32) inclusive may have been extended from the prototypical [enclosure] lexical concept. I observed above that in previous work with Andrea Tyler, Tyler and I argued that polysemy derives from regular processes of semantic change, in which situated implicatures associated with a particular context can become reanalysed as distinct semantic components, in present terms, lexical concepts, which are associated with the relevant preposition (Hopper and Traugott 1993; Traugott and Dasher 2004; cf. Levinson 2000). That is, Tyler and I argued for a usage-based approach to language change, a position adopted by LCCM Theory.

In terms of an LCCM account of the emergence of closed-class lexical concepts such as the 'state' lexical concepts for *in*, the trajectory is as follows. Situated implicatures arise in bridging contexts, as briefly discussed above. These are contexts in which a usage sanctioned by the relevant 'spatial' lexical concept, such as the [ENCLOSURE] lexical concept, also gives rise to a situated implicature, such as an affecting condition. If the form is repeatedly used in such bridging contexts, the situated implicature may give rise to the formation of a parameter: a highly abstract unit of knowledge, specialised for being encoded as part of the linguistic content associated with a lexical concept, as discussed earlier. I argue below that bridging contexts, involving the functional category of Affecting Conditions, give rise to the formation of a number of related but distinct 'state' parameters, and hence lexical concepts.

In order to trace the development of the functional category Affecting Conditions, we need to consider spatial scenes that might provide appropriate bridging contexts. To illustrate, consider the following expressions:

(36) a. in the dust
b. in the sand
c. in the snow

While dust, sand and snow are physical entities which can 'enclose' they cannot, normally fulfil the functions provided by, for instance, containers. That is, they do not typically serve to locate with surety, exceptional circumstances such as quicksand and avalanches excepted. For instance, dust, sand and snow, by virtue of enclosing, do not normally have the structural attributes that allow an entity to be supported and thus transported (cf. a bucket), nor do they normally restrict access in the way a prison cell does, for instance.

Nevertheless, these examples exhibit some of the spatio-geometric properties associated with the [ENCLOSURE] lexical concept. This is a consequence of the properties associated with these 'bounded' landmarks: they provide an affecting condition, an environmental influence which affects our behaviour. For instance, they determine the kinds of apparel we wear, and how we behave when we are exposed to the dust/sand/snow, and so on. While examples such as sand, snow and dust can be construed as enclosures with boundaries, there are other related examples of what we might refer to as Prevailing Conditions which are much less clear-cut in terms of the nature of the boundaries involved:

(37) a. the flag in the storm
b. the flag in the wind

I suggest that these instances of *in* are sanctioned by virtue of there existing a distinct parameter Affecting conditions, which forms part of the linguistic content encoded by a distinct [PREVAILING CONDITIONS] lexical concept. Clearly a storm and wind are much less prototypically enclosures, and more saliently provide prevailing conditions which thereby constitute an environment which affects us. As such, spatial scenes involving more prototypical enclosures have given rise to the functional category Affecting Conditions, which has led to the formation of a distinct Affecting Conditions parameter in semantic memory. The existence of a distinct [PREVAILING CONDITIONS] lexical concept, as evidenced by examples in (37) provides suggestive evidence that such a distinct Affecting Conditions parameter must exist, and has come to form the core a distinct [AFFECTING CONDITIONS] lexical concept.

I argue that there are a number of distinct 'state' lexical concepts associated with *in* that encode the parameter of Affecting Conditions, rather than Enclosure – those evidenced in (29)-(32). Indeed, these lexical concepts are what I have referred to as 'state' lexical concepts, as the states invoked all provide, in some sense, affecting conditions. Moreover, all these 'state' lexical concepts are relatively, and to degrees, far removed from the physical notion of enclosure from which they most likely originally evolved. In essence, once an Affecting Conditions parameter becomes conventionalised, it can be applied to distinct kinds of affecting conditions, even those that are non-spatial in nature, such as states. This leads to the development of new lexical concepts, with correspondingly new lexical profiles.

The first such 'state' lexical concept relates to the physical condition of an organism which thus provides an affecting condition. Such physical conditions include good/ill health, pregnancy, and any salient physical aspect of the organism's condition which affects and thus impacts on the organism's functioning. This lexical concept I gloss as [PHYSIOLOGICAL STATE]. In addition to environmental and physical conditions, affecting conditions can be caused by psycho-somatic states, such as grief, happiness and sadness which are internal in nature. This 'state' gives rise to a [PSYCHO-SOMATIC STATE] lexical concept associated with *in*. In addition, social interactions which give rise to social or interpersonal relationships lead to conditions which may affect the individual. Such extrinsic or socially-induced affecting conditions might include debts, or other sorts of difficult situations which impose conditions on the behaviour of an individual. This set of affecting conditions gives rise, I suggest, to what I gloss as the [SOCIO-INTERPERSONAL STATE] lexical concept associated with *in*. Finally, one's habitual professional activity provides an affecting condition by virtue of the physical and social interactions that are attendant upon such activities. This provides an affecting condition giving rise to a lexical concept glossed as [PROFESSIONAL STATE] associated with *in*. These are illustrated in Figure 7.

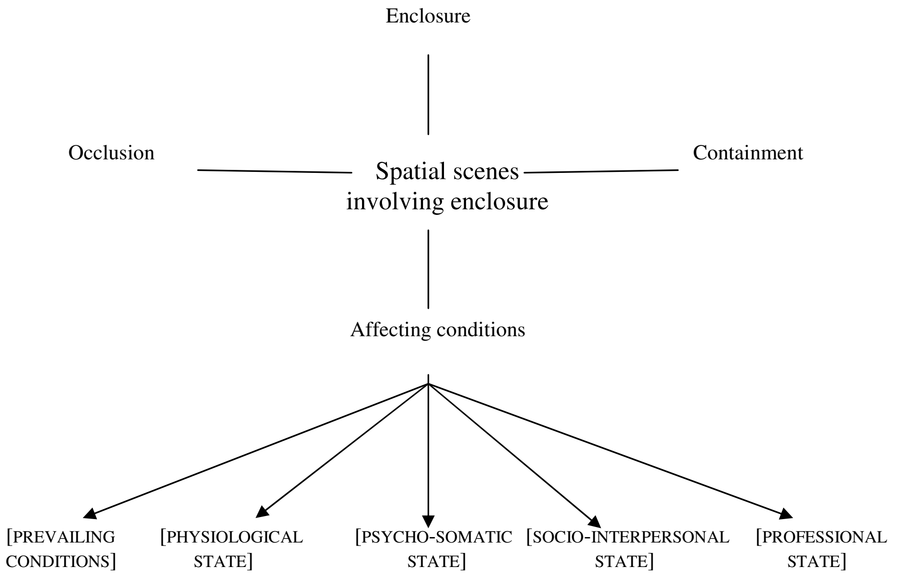


Figure 7. Parameters and their relationship with 'state' lexical concepts for *in*

8 Lexical concepts for *on*

In this section I deal, somewhat more briefly, with *on*.

8.1 The prototypical lexical concept for *on*: [CONTACT]

The spatial relation designated by *on* involves the relation of contact or proximity to the surface of a LM, and so the functional consequence of being supported or upheld by it. I gloss the prototypical 'spatial' lexical concept conventionally associated with *on* as [CONTACT]. This serves to encode the geometric parameter Contact and functional parameter Support as part of its linguistic content. This lexical concept licenses an example of the following sort:

(38) the apple on the table

Note that evidence that the parameters Contact and Support are both encoded by the lexical concept [CONTACT] comes from the fact that *on* can only felicitously be employed to describe spatial scenes in which both parameters are apparent. For instance, if an apple is held against a wall by someone, the utterance in (39) is semantically anomalous. However, if the apple is affixed to the wall, for instance by glue, then (39) is entirely appropriate.

(39) the apple on the wall

That is, while the apple is in contact with the wall in both scenarios, in the first scenario it is the person, rather than the wall, that affords support, while it is the wall (and the glue, which employs the wall as a means of affixing the apple) in the second. Hence, the example in (39) applies when there is both physical contact between the TR and the LM, and when the latter has a role in supporting the former.

Indeed, there are a number of distinct 'support' lexical concepts associated with *on* which privilege the Support parameter, at the expense of the Contact parameter, as illustrated by the following examples:

Body part which provides support

(40) a. on one's feet/knees/legs/back
b. on tiptoe
c. on all fours

In the examples in (40), the use of *on* relates to that part of the body which provides support, rather than being concerned with contact.

Means of conveyance

(41) a. on foot/horseback
b. on the bus

With respect to the example in (41b), it is worth pointing out, as Herskovits (1988) does, that if children were playing on a stationary bus, for instance, that had been abandoned, then it would not be appropriate to say: *on the bus*, but rather *in* would be more natural. This supports the view that the [means of conveyance] lexical concept is a distinct 'support' lexical concept encoded by *on*.

Supporting pivot

(42) The Earth turns on its axis

Drug dependency/continuance

(43) a. Are you on heroin?
b. She's on the pill

Psychological support

(44) You can count/rely on my vote

Rational/epistemic support

(45) on account of/on purpose

8.2 The [ACTIVE STATE] lexical concept for *on*

There is just one 'state' lexical concept for *on*, which I gloss as [ACTIVE STATE]. This lexical concept derives not from the functional category of Support. Rather, it pertains to a functional category concerning 'functionality' or 'activity'. That is, in many spatial scenes, a consequence of contact is that the TR, as it comes into contact with a particular surface, becomes functional. This category I refer to as Functional Actioning. Removing contact precludes functional actioning. Such forms of contact, for instance, invoke scenarios involving physical transmission, such as the very salient one of electricity. Many times a day we plug-in or switch 'on' electrical appliances. It is by facilitating contact between the appliance and the electrical circuit that an appliance is rendered functional. A 'switch' provides a means of facilitating this contact, which is why we employ the term 'switch on' in English. In other words, I suggest that the [ACTIVE STATE] lexical concept associated with *on* encodes a Functional Actioning parameter as part of its linguistic content. It is this which makes it distinctive from the 'spatial' lexical concepts of *on* discussed in the previous examples.

The [ACTIVE STATE] lexical concept associated with *on* relates to adjectives or nouns of action which involve a particular state which can be construed as 'active' or 'functional', as contrasted with a, perhaps, normative scenario in which the state does not hold. In other words, states described by instances of *on* sanctioned by this lexical concept are often temporally circumscribed and thus endure for a prescribed or limited period of time. In this, the states referred to are quite distinct from those that *in* serves to describe. Here, the notion of being 'affected', apparent with *in*, is almost entirely absent. Consider some examples:

(46) a. on fire
b. on live (i.e., a sports game)
c. on tap (i.e., beer is available)
d. on sleep (as in an alarm clock on a particular mode)
e. on pause (as in a DVD player)
f. on sale
g. on loan
h. on alert
i. on best behaviour
j. on look-out
k. on the move
l. on the wane
m. on the run

Figure 8 depicts the parameter that underpins this lexical concept.

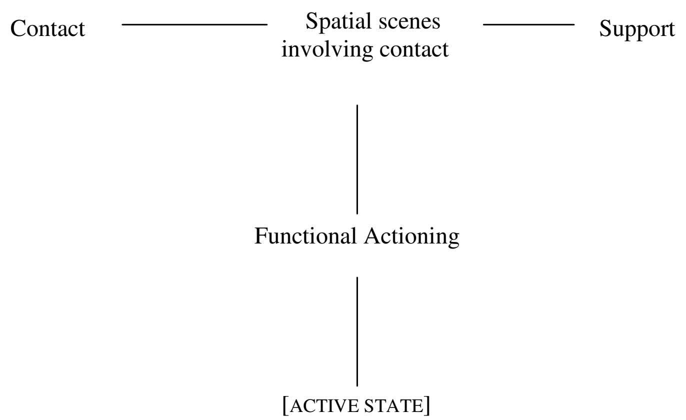


Figure 8. Parameters and their relationship with 'state' lexical concept for *on*

9 The state senses of *at*

This section briefly examines the 'state' lexical concepts of *at*.

9.1 The prototypical lexical concept for *at*: [co-location]

The lexical concept which licenses spatial uses of *at* affords the most general expression of localisation in space in English, expressing the relation between a TR and a point of space that it is contiguous or proximal with. This lexical concept I gloss as [co-location]. Consequently, it is one of the most polysemous of all English prepositions. Indeed, this lexical concept for *at* forms a *contrast set* (Tyler and Evans 2003) with the 'place' identifying lexical concepts associated with other prepositions. The [co-location] lexical concept encodes the Co-location parameter, designating a highly abstract spatial relation between a TR and a place, when the relation is not more precisely expressed by 'spatial' lexical concepts associated with the following prepositional forms: *near, by, on, in, over, under*, all of which, at times, can be encoded by *at*.

Perhaps the most salient functional category associated with *at* constitutes what I will refer to as that of Practical Association. That is, a functional consequence of being co-located with a particular LM is that the TR has some practical association with the reference object. This is evidenced in the examples in (6) discussed earlier (e.g., *at the desk/bus-stop*), and is particularly evident with examples such as the following:

(47) a. at school
b. at sea

In these examples, the activity associated with the school buildings or being out on the sea is extremely salient.

9.2 The 'state' lexical concepts for *at*

There are three distinct lexical concepts associated with *at* that might be described as relating to 'states'. These are illustrated below:

State (or condition) of existence

(48) at rest/peace/ease/liberty
(e.g., *He stood at ease*, or *He is at peace* [=dead])

States relating to mutual relations

(49) at war/variance/strife/one/dagger's drawn/loggerheads
(e.g., *The EU is at war with the US over the imposition of steel tariffs*)

States relating to external circumstances

(50) at peril/risk/hazard/expense/an advantage/a disadvantage
(e.g., *The company is at risk of going under*)

The 'state' lexical concepts for *at* appear to be motivated by the functional consequence of close-proximity between two point-like entities giving rise to the formation of a parameter: Practical Association.

In the case of the [STATE OF EXISTENCE] lexical concept, the practical association resulting from the co-location is the state of existence which holds. That is, there is a practical association which holds between a given entity and its state of existence.

The second lexical concept I gloss as [STATE OF MUTUAL RELATIONS], as evidenced by (49). This lexical concept arises due to a salient practical association resulting from co-location of two entities involving mutual relations. For instance, while warfare often involves combatants who must be proximal to one another, the state of being 'at war' need not, as evidenced by the so-called 'phoney war' which held during 1939 when the United Kingdom, France and Germany were officially 'at war', and yet no troops engaged. Thus, the use of *at* to designate a state of mutual relations, independent of spatio-geometric co-location, is due to the parameter of Practical Association being invoked as part of the linguistic content encoded by this lexical concept. Put another way, this lexical concept encodes a state of a particular kind, rather than the 'spatial' notion of proximity.

Finally, states pertaining to external circumstances may relate to evaluations concerning circumstances associated with mutual relations. This is instantiated by the lexical concept which I gloss as [STATES OF EXTERNAL CIRCUMSTANCES], as evidenced by the examples in (50). The relationship between the parameter of Practical Association and the 'state' lexical concepts is diagrammed in Figure 9.

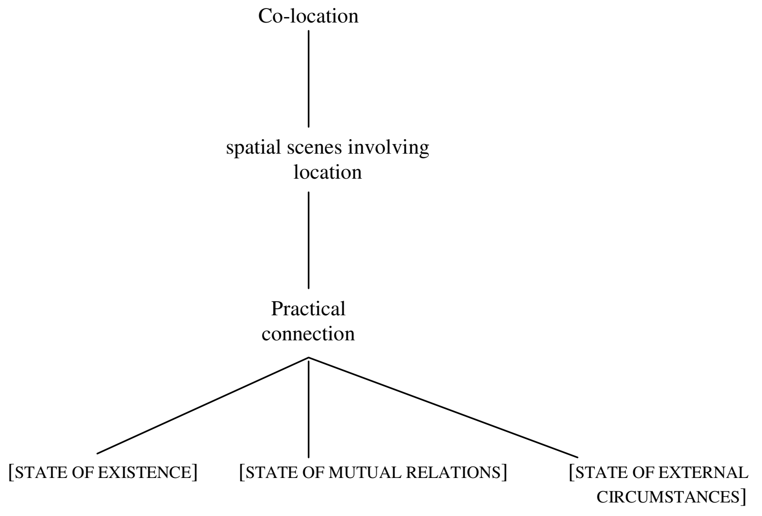


Figure 9. Parameters and their relationship with 'state' lexical concepts for *at*

10 Conclusion: *in* vs. *in* vs. *at*

Having presented an analysis of i) distinct 'state' lexical concepts for *in*, *on* and *at*, and ii) how these encode distinct parameters which relate to functional categories arising from spatial scenes involving spatio-geometric relationships, I now return to one of the observations with which I began this study. I observed that each of the 'state' lexical concepts associated with *in*, *on* and *at*, as exemplified in (1)-(3), is minimally distinct in that it is associated with distinct semantic arguments. Consequently the lexical concepts exemplified in these examples relate to states of distinct kinds. The analysis presented here has supported this initial assessment, and elaborated on it in three ways.

Firstly, the perspective offered here, particularly with respect to the construct of the lexical concept, allows us to establish in a reasonably precise way the nature of the distinction between the 'state' lexical concepts associated with *in*, *on* and *at*. That is, given that lexical concepts are form-specific and moreover have distinct lexical profiles – for instance they collocate with distinct kinds of semantic arguments – we are able to establish that the 'state' lexical concepts (within and between) prepositional forms are distinct.

Secondly, by taking seriously the functional nature of spatial relations, and the formation of parameters: highly abstract knowledge structures specialised for being directly encoded 'in' language, this allows us to understand the sorts of functional motivations, and thus distinctions, between the 'state' lexical concepts among different forms.

Thirdly, prepositions, particularly *in* and *at* have more than one so-called 'state' lexical concept associated with them. We have seen that the prototypical spatial lexical concept associated with a given preposition is associated with a number of parameters, including parameters derived from what I referred to as functional cognitive categories.

Providing an LCCM analysis gives us a way of establishing the sorts of distinctions that exist between the 'state' lexical concepts associated with the same form. That is, we have a means of understanding how these lexical concepts are distinct (based on a distinction in parameters encoded) despite their conceptual similarity. We also have a means of empirically verifying hypotheses as to distinctions in the underlying lexical concepts which are assumed to sanction instances of use. This followed due to an examination of semantic selectional tendencies, which relate to the theoretical construct of the lexical profile in LCCM Theory: distinct lexical concepts are held to have a unique lexical profile which forms part of the knowledge encoded by a given lexical concept.

Notes

1 This said, the framework developed in Tyler and Evans (2001, 2003) and Evans and Tyler (2004a, 200b) remains important. Principled Polysemy, as articulated in those publications, was and remains an important theoretical development in terms of what it brings to descriptive accounts of spatial semantics. In particular, it sought, for good reason, to address the sorts of methodological criticisms that had been levelled at the early pioneering work of Brugman and Lakoff (Brugman [1981] 1988; Brugman and Lakoff 1988; Lakoff 1987) in developing cognitive lexical semantics. While it doubtless requires modification (see Evans 2004a), it nevertheless provides a relatively robust set of methodologically constrained, and above all principled 'decision principles' (in Sandra's 1998 terms) for identifying and distinguishing between senses-units, and for, a principled means of modelling lexical polysemy with respect to spatial relations. While important developments in the use of psycholinguistic testing (see Sandra and Rice 1998; Cuyckens et al. 1997) and corpus-based techniques (see Gries 2005) have added to the arsenal of cognitive lexical semanticists in this regard, empirical methods will always require a theoretical framework in order to motivate the sorts of questions that can be asked and to provide a lens for interpreting results, even though this may mean modifying the theoretical framework. Indeed, this perspective is in fact compatible with the desire to have more empirical methods in cognitive lexical semantics.

2 See Evans (2005) and Tyler and Evans (2001, 2003) for detailed discussion of polysemy.

References

Atkins, Beryl T. (1987) Semantic ID tags: Corpus evidence for dictionary senses. *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary* 17–36.

Bennett, David (1975) *Spatial and temporal uses of English prepositions*. London: Longman.

Brugman, Claudia ([1981] 1988) *The Story of 'over': Polysemy, semantics and the structure of the lexicon.* New York: Garland.

Brugman, Claudia and George Lakoff (1988) Cognitive topology and lexical networks. In S. Small, G. Cottrell and M. Tannenhaus (eds) *Lexical ambiguity resolution* 477–507. San Mateo, CA: Morgan Kaufman.

Coventry, Kenny and Simon Garrod (2004) *Saying, seeing and acting. The psychological semantics of spatial prepositions*. Hove: Psychology Press.

Croft, William (2000) *Explaining language change: An evolutionary approach*. London: Longman.

Cuyckens, Hubert, Dominiek Sandra and Sally Rice (1997) Towards an empirical lexical semantics. In B. Smieja and M. Tasch (eds) *Human Contact Through Language and Linguistics* 35–54. Frankfurt: Peter Lang.

Deane, Paul (2005) Multimodal spatial representation: On the semantic unity of *over*. In B. Hampe (ed.) *From perception to meaning: Image schemas in cognitive linguistics* 235–284. Berlin: Mouton de Gruyter.

Evans, Nicholas and David Wilkins (2000) In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3): 546–592.

Evans, Vyvyan (2004a) *The structure of time: Language, meaning and temporal cognition*. Amsterdam: John Benjamins.

Evans, Vyvyan (2004b) How we conceptualise time. *Essays in arts and sciences* 33(2): 13–33.

Evans, Vyvyan (2005) The meaning of *time*: Polysemy, the lexicon and conceptual structure. *Journal of Linguistics* 41(1): 44–75.

Evans, Vyvyan (2006) Lexical concepts, cognitive models and meaning-construction. *Cognitive Linguistics* 17(4): 491–534.

Evans, Vyvyan (2009) *How words mean*. Oxford: Oxford University Press.

Evans, Vyvyan, Benjamin K. Bergen and Jörg Zinken (2007) The cognitive linguistics enterprise: An overview. In V. Evans, B. Bergen and J. Zinken (eds) *The cognitive linguistics reader*. London: Equinox Publishing Co.

Evans, Vyvyan and Melanie Green (2006) *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Evans, Vyvyan and Andrea Tyler (2004a) Rethinking English 'prepositions of movement': The case of *to* and *through*. In H. Cuyckens, W. De Mulder and T. Mortelmans (eds) Adpositions of movement (*Belgian Journal of Linguistics* 18) Amsterdam: John Benjamins.

Evans, Vyvyan and Andrea Tyler (2004b) Spatial experience, lexical structure and motivation: The case of *in*. In G. Radden and K. Panther (eds) *Studies in linguistic motivation* 157–192. Berlin: Mouton de Gruyter.

Feist, Michele (This volume) Inside *in* and *on*: Typological and psycholinguistic perspectives.

Grady, Joseph (1997) *Foundations of meaning: Primary metaphors and primary scenes*. Doctoral Thesis, Linguistics dept., UC, Berkeley.

Gries, Stefan Th. (2005) The many senses of run. In S. Gries and A. Stefanowitsch (eds) *Corpora in cognitive linguistics.* Berlin: Mouton de Gruyter.

Herskovits, Annette (1986) *Language and spatial cognition*. Cambridge: Cambridge University Press.

Herskovits, Annette (1988) Spatial expressions and the plasticity of meaning. In B. Rudzka-Ostyn (ed.) *Topics in cognitive linguistics*. Amsterdam: John Benjamins, 271–298.

Hopper, Paul and Elizabeth Closs Traugott (1993) *Grammaticalization*. Cambridge: Cambridge University Press.

Johnson, Mark (1987) *The body in the mind: The bodily basis of meaning, imagination and reason*. Chicago: University of Chicago Press.
Johnson, Mark (2007) *The meaning of the body: Aesthetics of human understanding.* Chicago: University of Chicago Press.
Lakoff, George (1987) *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
Lakoff, George and Mark Johnson (1999) *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York, NY: Basic Books.
Langacker, Ronald (2000) A dynamic usage-basked model. In M. Barlow and S. Kemmer (eds) *Usage-based models of language* 1–64. Stanford, CA: CSLI Publications.
Levinson, Stephen (2000) *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
Miller, George and Philip Johnson-Laird (1976) *Language and perception*. Harvard: Harvard University Press.
Sandra, Dominiek (1998) What linguists can and can't tell you about the human mind: A reply to Croft. *Cognitive Linguistics* 9(4): 361–478.
Sandra, Dominiek and Sally Rice (1995) Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's? *Cognitive Linguistics* 6(1): 89–130.
Tomasello, Michael (2003) C*onstructing a language: A usage-based theory of language acquisition*. Harvard: Harvard University Press.
Traugott, Elizabeth Closs and Richard Dasher (2002) *Regularity in semantic change*. Cambridge: Cambridge University Press.
Tyler, Andrea and Vyvyan Evans (2001) Reconsidering prepositional polysemy networks: The case of *over. Language* 77(4): 724–765.
Tyler, Andrea and Vyvyan Evans (2003) *The semantics of English prepositions: Spatial scenes, and the polysemy of English prepositions*. Cambridge: Cambridge University Press.
Vandeloise, Claude (1991) *Spatial prepositions: A case study from French.* (Translated by Anna R. K. Bosch.) Chicago: University of Chicago Press.
Vandeloise, Claude (1994) Methodology and analyses of the preposition *in*. *Cognitive Linguistics* 5(2): 157–184.

Part V

Spatial representation in specific languages

10 Static topological relations in Basque

Iraide Ibarretxe-Antuñano

1 Space in language and cognition

Space is one of the most studied areas not only from the point of view of linguistic description, that is, the description of the linguistic devices that languages have to express and describe space, but also from the perspective of cognition, how space is understood and computed in our brains. In recent years, a major focus of analysis in the domain of space has been the relationship between language and cognition (Landau and Jakendoff, 1993; Levinson, 2003; Majid et al., 2004).

In this paper, I will offer a description of the linguistic means available in Basque for the lexicalisation of space as well as their usage, that is, which of these devices is used more often by Basque speakers when they talk about space. More concretely, I will deal with those devices used for the description of topological relations, what Levinson et al. (2003: 486) call *basic locative constructions*, i.e. answers to 'where' questions. Based on empirical data elicited by means of the Topological Relation Picture Series, stimuli developed at the Max Planck Institute for Psycholinguistics (Pederson, Wilkins, and Bowerman 1993), I will concentrate on the type of spatial information that is crucial for Basque speakers in the description of space, and also on certain features that seem to influence how Basque speakers conceptualise space.

2 Topological relational markers in Basque

One of the main difficulties that a linguist faces when s/he starts to analyse how space is described in a given language is the wide range of semantic and morpho-syntactic elements and mechanisms that are more or less directly involved in its codification. Space is expressed not only by means of nouns, verbs, adverbs, adpositions, cases… but also by combinations of these elements, that is to say, spatial description is not always localised on one single lexical item but distributed alongside several words (Sinha and Kuteva, 1995). For example, the Basque ablative means 'through' only if it is accompanied by a transversable ground (e.g. door), otherwise it means 'source'. The situation becomes even more complicated if one tries to describe languages where some elements do not exactly fit into traditional linguistic categories such as the 'category of associated-motion' in Arrernte (Koch, 1984; Wilkins, 1991, 2004), and almost impossible when some elements are classified differently depending on the linguistic framework one works in as is the case for the distinction between case and postposition in structuralism, generativism, and functionalism (cf. Agud, 1980 and Blake, 2001).

In order to avoid this terminology problem, I adopt the cover term *topological relational marker* proposed by Levinson et al. (2003: 486) and use it for all the 'various form classes involved in coding topological relations'. In this section, I analyse some of the most commonly used topological relational markers in Basque: spatial cases, spatial nouns, and motion verbs.

Spational cases in Basque

There are five spatial cases in Basque (see Ibarretxe-Antuñano, 2001, 2004a, for a detailed account):

- The *locative case* (*-n*) is one of the most productive cases in the Basque case system. Its basic meaning is 'location' in space ('in', 'on', 'at'), as in *kale-eta-n* [street-pl-loc][1] 'in the streets', *mahai-an* [table-loc.sg] 'on the table', *etxe-an* [house-loc.sg] 'at home'. Sometimes it also expresses motion ('into'), as in *geltoki-an sartu* [station-loc.sg enter] 'to go into the station'.

- The *ablative case* (*-ti(k)*) is usually defined as expressing the 'source of motion'. For example, *etxe-tik* [house-abl.sg] means 'from the house'. In specific contexts, the ablative can also convey the meaning 'through' as in *leiho-tik* [window-abl.sg] 'through the window'.

- The *allative case* (*-ra(t)*) expresses the 'goal of motion' in the domain of space, as in *etxe-ra* [house-all.sg] 'to the house'.

- The *goal***,** *destinative* or *terminative allative* (*-raino*) conveys the meaning 'up to' in the spatial domain, as in *etxe-raino* [house-ter.sg] 'up to the house'. It indicates a telic motion event, that is, the trajector reaches its final destination.

- The *directional allative* (*-rantz, -runtz, -rontz*) indicates the notion of 'towards' in the spatial domain as in *etxe-rantz* [house-dir.sg] 'towards home'. This spatial case profiles the directionality of the motion event. The trajector moves towards a specific destination but it is not specified whether the trajector reaches or not the place towards which it moves.

Within the Basque case system, spatial cases form a special group, not only because they share a common reference to space, but also because they behave morphologically differently from the other Basque cases. Their main properties are the following: (i) they are of direct relevance to the distinction between animate and inanimate head nouns, (ii) they lack the article *-a* in the definite singular form, and (iii) they have the infix *–(e) ta* in non-singular inanimate NPs.

Apart from the five spatial cases, two other cases have been used for spatial description in our elicited data. The adnominal marker, also called locative genitive (*-ko*) as in *Euskal Herri-ko-a* [basque country-adn.sg-abs.sg] 'from the Basque Country', and the dative (*-i*) as in *Lakioa kandela-ri lotuta dago* [ribbon.abs.sg candle-dat.sg tie.pple is.3sg] 'the ribbon is tied to the candle'. The dative argument when used with verbs of motion such as *etorri* 'come', *joan* 'go', and *ibili* 'walk' usually refers to a goal as in *na-tor-ki-zu* [1sg.abs-come.stem-dative-2sg.dat] 'I am coming to you'.

Spatial nouns

Basque has more than thirty or more spatial nouns[2] that specify even more finely spatial relations between figure and ground. Table 1 summarises some of the most widely used. All these spatial nouns follow the same structure: the ground they belong to usually takes the genitive case, which can be dropped, and, to a lesser extent, the absolutive, ablative, dative or instrumental cases, and, the spatial noun in turn takes any of the five spatial cases. For example, *mahai-aren gain-etik* [table-gen top-abl] 'from the top of the table' or *mendi-tik zehar* [mountain-abl through] 'through / over the mountain'. In our elicited data, almost all spatial nouns are inflected in the locative case.

Table 1. Spatial nouns in Basque

Case for Ground	Spatial noun	Meaning	Example
(Gen)	*Aitzin*	'front' (eastern)	*Eliza(ren) aitzin-era* 'to the front of the church'
(Gen)	*Albo*	'side'	*Zuhaitz(aren) albo-an* 'next to the tree'
(Gen)	*Aldamen*	'side'	*Ikastola(ren) aldamen-erantz* 'towards near the school'
(Gen)	*Alde*	'side'	*Etxe(aren) alde-tik* 'from near the house'
(Gen)	*Arte*	'space between, among'	*Arrok(en) arte-tik* 'from between the rocks'
Abl	*At*	'outside'	*Etxe-tik at* 'outside from the house'
Abl, Instr	*Ate*	'door'	*Eliza-tik ate-an* 'outside the church'
(Gen)	*Atze*	'back'	*Etxe(aren) atze-tik* 'from the back of the house'
(Gen)	*Aurre*	'front'	*Eliza(ren) aurre-an* 'in front of the church'
(Gen)	*Azpi*	'bottom, lower part'	*Mahai(aren) azpian* 'under the table'
(Gen)	*Barne*	'interior, inside'	*Etxe(aren) barne-tik* 'from inside the house'
(Gen), Loc	*Barren*	'interior, inside; bottom, lower part'	*Eliza(ren) barren-ean* inside the house' *Oihan-ean barren-a* 'through the forest'
(Gen)	*Barru*	'interior, inside'	*Etxe(aren) barru-ra* 'to the interior of the house'
(Gen), Loc, Abl, Ø	*Behe*	'bottom, ground, lower part'	*Mendi-an behe-ra* 'to the lower part along the mountain'
(Gen), All	*Buru*	'centre' 'extremity'	*Kale(aren) buru-an* 'at the end of the street'
(Gen)	*Erdi*	'middle, centre'	*Eliza(ren) erdi-tik* 'from the middle of the church'
(Gen)	*Gain*	'top, upper part'	*Mahai(aren) gain-era* 'to the top of the table'
Loc	*Gaindi*	'through'	*Mendi-an gaindi* 'through the mountain'
(Gen)	*Gibel*	'back' (eastern)	*Etxe(aren) gibel-etik* 'from the back of the house'

Case for Ground	Spatial noun	Meaning	Example
(Gen), Loc, Abl, Ø	*Goi*	'top'	*Etxe-tik go-ra* 'from the house to the top'
(Gen)	*Inguru*	'vicinity'	*Eliza(ren) inguru-an* 'in the vicinity of the church'
Abl, Instr	*Kanpo*	'outside, exterior'	*Etxe-tik kanpo* 'outside the house'
(Gen), Dat	*Kontra*	'against'	*Horma-ri kontra* 'against the wall'
Abl	*Landa*	'field'	*Hiri-tik landa* 'outside the city'
(Gen)	*Ondo*	'side'	*Ikastola(ren) ondo-raino* 'up to near the school'
(Gen)	*Oste*	'back'	*Eliza(ren) oste-an* 'at the back of the church'
(Gen)	*Pare*	'opposite side'	*Etxe(aren) pare-an* 'across from the house'
(Gen)	*Pe*	'lower part, below, under'	*Zuhaitz bat(en) pe-an* 'below the tree'
(Gen)	*Saihets*	'side'	*Ama(ren) saihets-ean* 'next to the mother'
Loc, Abl	*Zehar*	'through, across'	*Mendi-an zehar* 'through the mountain'

Motion verbs

In Basque, there are more than 2000 different types of motion verbs (Ibarretxe-Antuñano, in prep.). A possible explanation for this rich repertoire, which contradicts the general prediction about the number of this type of verbs in verb-framed languages (Slobin, 1996), can be found in the various morphological strategies that Basque employs to create its motion verb lexicon (Ibarretxe-Antuñano, in press). Let us briefly explain some of these strategies:

> Simple verbs. These can be classified into four classes according to their perfective participle[3] (Hualde and Ortiz de Urbina, 2003: ch. 3.5; Trask, 1997: 103): (i) verbs in –*i* such as *iritsi* 'arrive', (ii) verbs in –*tu* such as *jarraitu* 'follow'. This is the only suffix that can be used in borrowings from other languages such as *ailegatu* 'arrive' and *bueltatu* 'return', (iii) verbs in –*n* such as *joan* 'go', and (iv) verbs with no suffix such as *igo* 'ascend'.
>
> Derived verbs. There are two categories: (i) verbs formed from nouns and adjectives plus the past participle suffix in –*tu* (-*du* after a nasal or lateral) as in *zuzen-du* [straight-suf] 'head, set off', and (ii) verbs formed from (spatial) nouns inflected in the allative singular plus the past participle suffix –*tu* as in *lurre-ra-tu* [ground-all-suf] 'go down', *aurre-ra-tu* [front-all-suf] 'go forward'. In these cases the verb always means 'go/bring to (spatial) noun' (Hualde, 2003a: 347).
>
> Compound verbs.[4] There are two categories: (i) uninflected spatial noun plus the verb *egin* 'make, do' as in *alde egin* [side make] 'leave', and (ii) inflected spatial noun, usually in the allative or directional allative, plus the verb *egin* 'make, do' as in *eskuma-rantz egin* [right-dir.all make] 'go right', *behe-ra egin* [below-all make] 'go down'.

The possibility of using these strategies for conveying motion verbs implies that the lexicon is very rich. For instance, if we wanted to say 'go out' in Basque, the lexicon would give us the opportunity to choose among four different possibilities: *atera*, *irten*, *kanpo-ra egin*, and *kanpo-ra-tu*, plus a construction with the spatial noun (*kanpo* 'outside') with the allative and the verb *joan* 'go', i.e. *kanpo-ra joan*. This means that for the same motion description we can use a wide variety of choices: compound verbs, derived verbs, pairs of synonyms which belong to different verb classes, such as *iritsi* and *heldu* 'arrive', even loans from other surrounding languages such as *ailegatu* and *arribatu* 'arrive' from Spanish *llegar* and French *arriver* respectively.

In the specific subcase of locative verbs, Basque is one of those languages that only offers a *small set* of locative/posture verbs (Ameka and Levinson, 2007). The static verb par excellence is *egon* 'static be'. In standard Basque, there is a distinction between *egon* 'static be' as in *etxe-an dago* [house-loc is.3sg] 's/he at home' and *izan* 'existential be' as in *hizkuntzalari-a da* [linguist-abs is.3sg], but in eastern dialects there is only one basic copular verb, *izan*, which covers both functions (*etxe-an / hizkuntzalari-a da*). Basque also has a set of posture verbs such as *eseri* 'sit', *zutitu, jagi* 'stand', *zintzilikatu*, *eskegi* 'hang', *etzan* 'lie down', *makurtu* 'crouch'… These verbs are mostly used for the description of an active change of posture as in *eseri da* [sit-perf aux.2sg] 's/he sits down' and *Mikel-ek soka zuhaitz bat-etik zintzilikatu du* [Mikel-erg rope.abs tree one-abl hang.perf aux.3sg.abs.3sg.erg] 'Michael hung the rope from the tree'. For the description of static postures, these verbs are always used in the participle form (*-ta, -rik*) with the verb *egon* as in *eseri-ta dago* [sit-pple is.3sg] 'he's sitting' and *soka zuhaitz bat-etik zitzilikatu-ta dago* [rope.abs tree one-abl hang.pple is.3sg].

3 Topological relational markers and their usage

Data elicitation

The elicitation tool I used in this study is called the *Topological Relation Picture Series* and was developed at the Max Planck Institute for Psycholinguistics in Nijmegen (Pederson, Wilkins, Bowerman, 1993; see also Bowerman, 1996). This tool is a booklet that consists of seventy-one line-drawings that depict different topological spatial relations equivalent to English prepositions *in, on, at, under, near, in the middle of*…and such like. Each drawing has a figure coloured yellow and marked with an arrow, and a ground object in black and white as shown in Figure 1.

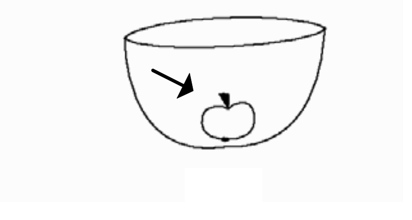


Figure 1. Example of a drawing from the Topological Relation Picture Series

The procedure is very simple. The researcher asks the informant to answer a question of the form: 'Where is the [Figure] (with respect to the [Ground])?'. Informants provide descriptions of the drawings, including the spatial relational marker that would most naturally be used to describe the relation depicted. If informants provide more than one answer these are also noted down as second or third choices. Twenty-six Basque native speakers participated in this study.

Which topological relational markers does Basque use more frequently?

The topological relational marker par excellence in Basque is the locative case *–n*. The locative is usually found with the verb *egon* 'static be' (1), with other spatial nouns (2): and with participles (3).

(1)	*Sagarra*	*ontzi-an*	*dago* [2][5]	
	apple.abs	bowl-loc	is.3sg	
(2)	*Sagarra*	*ontzi-aren*	*barru-an*	*dago* [2]
	apple.abs	bowl-gen	inside-loc	is.3sg
(3)	*Tirita*	*hanka-n*	*jarri-ta*	*dago* [35]
	bandaid.abs	leg-loc	put.pple	is.3sg

The Basque locative is a good example of what Feist (2004, 2008) calls a 'General spatial term'. Similarly to Turkish *–da*, Ewe *le* and Indonesian *–di*, the Basque locative occurs in all spatial descriptions, with or without more specific terms, and with no specific information about the location of the figure. The locative only lexicalises the semantic component LOCATION, i.e. position in space, leaving for the other elements of the spatial description (semantic content of figure, ground, verbs…) the details about the exact spatial configuration. For example, in *boat on water* [11] the locative only tells us that the figure *boat* is located in the area of interaction of the ground *water*. It is only thanks to the characteristics of boats and water and their canonical relationship that the inference that the boat is on the water and not inside the water arises.

As far as its usage is concerned, the locative (alone, with no other spatial noun) is the most widely used topological relational marker in our data. Informants employ the locative case as their first choice in 38 drawings (54%) and as their second choice in 26 drawings (37%). There are only seven drawings (10%) where the locative case is not used on its own, although it does appear in the spatial noun.

Spatial nouns are also used quite frequently for the description of these drawings; more concretely, they are the first choice in 29 drawings (40%) and appear as second and third choice in virtually all the pictures. In our data, spatial nouns fulfil two different functions. On the one hand, they cover spatial scenes that cannot be or are not usually described by the locative case as in *cloud over mountain* [36]. On the other, they give more specific information about the spatial scene than the locative case. Here, several subcases can be distinguished. Spatial nouns are used when the informant wants to be more specific about the scene s/he is describing. For example, in drawings such as *book on self* [8] and *hat on head* [5], informants tend to use the spatial noun *gain-ean* [top-loc] 'on top of', even though the locative could be enough for the description of such spatial configurations. The canonical positions of these figures – books and hats – with respect to their grounds do not really allow for other type of interpretation than 'be on top'. Spatial nouns are also used when the locative is too uninformative; that is to say, when the locative offers only very general and ambiguous spatial information that gives rise to different spatial configurations. This is the case of drawings like *owl in tree hole* [67] and *lamp over table* [13], where informants prefer to mention as their first choices spatial nouns such as *barru-an* [inside-loc] and *gain-ean* [top-loc] respectively. Finally, they are also used for the description of a common spatial type scene. These are cases where the spatial relationship between figure and ground is prototypical, that is, they are located in a way that is expected by the informant due to the characteristics of both elements. Drawings such as *cup on table* [1], *cat under table* [31] are good examples of this subcase. Here, informants consistently choose a spatial noun like *gain-ean* [top-loc] and *azpi-an* [below-loc] instead of the locative alone.

With respect to the utilisation of motion verbs, the most widely used verb is *egon* 'static be'. The semantic information of this locative verb is very poor and general, and thanks to the tolerance of Basque for verb omissions (Ibarretxe-Antuñano, 2004b, in press), the verb *egon* is often elided. As mentioned in the previous section, Basque also has several posture verbs, but these are mainly used as adverbial participles accompanying the verb *egon*. Adverbial participles such as *itsatsi-ta, inke-ta,* 'stuck', *eskegi-ta, zintzil-ik* 'hung', *eutsi-rik* 'clung', *ipini-ta, jarri-ta* 'put', and *lotu-ta* 'tied, joined' are basically used to specify the information in the locative (or ablative) and/or spatial noun. For instance, in drawings like *coat on hook* [9] and *clothes on line* [37], the majority of informants – 22 (84,6%) in the former and 21 (80,7%) in the latter – mention 'hung' in their descriptions. It is interesting to point out that in a couple of drawings, informants seem to interpret and describe the scenes dynamically by means of motion verbs such as *ingururatu* 'go around' in *fence around the house* [15], and *zeharkatu* 'cross' in *arrow through apple* [30]. These dynamic descriptions give rise to a very important question: Can location be conceptualised both as dynamic and static? I will come back to this issue later in the discussion.

4 Spatial information

In this section, I will discuss the type of spatial information that Basque topological relational markers pay attention to. Authors such as Herskovits (1986), Jackendoff (1993), Miller and Johnson-Laird (1976), Vandeloise (1991) have proposed different labels for spatial components. Here I will follow the terms commonly used in other work related to the Topological Relational Picture Series produced by the Language and Cognition group at the MPI (see Bowerman, 1996; Bowerman and Choi, 2001; Levinson et al., 2003). Spatial components are written in small caps. For each spatial component, I will mention the topological relational markers (TRM) related to them and the drawings that better exemplify these spatial configurations. The complete distribution of these relational markers and the pictures – extensional maps – can be found in the Appendix. Here Venn diagrams are used to show how Basque topological relational markers group certain scenes together. Six are the semantic components covered by Basque TRMs in our elicited data:

LOCATION

The locative case is the prototypical vehicle for the expression of LOCATION. As said in the previous section, the locative case by itself does not offer any further specification about the spatial configuration. This is inferred from the figure and ground, which can be one-, two-, or three-dimensional. It can be used in all sorts of situations with or without CONTAINMENT, with or without CONTACT-AND-SUPPORT as shown in the description of drawings such as *rabbit in cage* [54], *boat on water* [11], and *ribbon around candle* [4].

VERTICALITY

This spatial component is present in the spatial nouns *gain-ean* [top-loc] and *azpi-an* [bottom-loc]. CONTACT-AND-SUPPORT information is not relevant for these spatial nouns since they are used both in contact-and-support situations as in the drawing *cup on table* [1] and in non-contact-and-support cases as in *lamp over table* [13].

CONTAINMENT (ENCLOSURE)

Barru-an [inside/interior-loc] is the spatial noun that pays attention to this spatial component. The distinction between PARTIAL and TOTAL INCLUSION is not relevant for this spatial noun as its appearance in the drawings *box in bag* [14] and *apple in bowl* [2] demonstrates. A three-dimensional CONTAINMENT is not necessary for *barruan*; this spatial noun appears both in a two-dimensional scene such as *dog in basket* [47] and in a three-dimensional scene such as *apple in bowl* [2]. Although both *barruan* and *barnean* refer to the interior of a given ground, the conceptualisation of interior does not seem to be exactly the same. *Barru* is usually applied to grounds that happen to have an interior as part of their intrinsic configuration as in the case of *dog in basket* [47]. *Barne*, on the other hand, only occurs in situations where the ground itself is the one that delimits or creates an interior as in the drawing *house in fence* [60].

HORIZONTALITY

This spatial element is present in the spatial nouns *atze-an* [back-loc] and *aurre-an* [front-loc] as represented in drawings such as *tree in front of church* [49] and *boy behind the chair* [64]. It is important to point out that these spatial nouns not only involve a topological concept – HORIZONTALITY – but also projective meanings, i.e. they specify an angle in relation to a ground and project a search-domain for the referent from that ground. *Atzean* and *aurrean* fit into the intrinsic frame of reference (FOR). This coordinate system is object-centred and its coordinates are determined by the conceptual characteristics of the ground object: its shape, canonical orientation and so on (see Levinson, 2003: ch.2, for a complete discussion of frames of reference).

DISTANCE

This is represented in our data by spatial nouns like *ondo-an* [side-loc], *albo-an* [side-loc], and *alde bat-ean* [side one-loc]. All of them denote PROXIMITY as in the drawing *boy next to fire* [38]. This spatial element can be complemented with relative frame of reference information. The relative coordinate system 'presupposes a viewpoint, and a figure and ground distinct from the viewpoint […] and utilizes coordinates fixed on the viewpoint to assign directions to figure and ground' (Levinson, 2003: 43). For example, informants use *eskubi-tara* [right-all] in the drawing *dog near kennel* [6] and *ezkerr-etara* [left-all] in *boy next to fire* (38). Although more research is needed, the elicited data suggest that *aldean* is preferred in cases where the figure is not just located near the ground but also constitutes an integral part of it as in drawings like *tree on hillside* [17] and *strap on purse* [66], whereas *ondoan* and *alboan* do not necessarily entail this identity connection.

POINT-TO-POINT ATTACHMENT

This spatial information is lexicalised by means of the ablative case together with a posture verb as adverbial participle as in *–tik zintzilik / eskegita* [-abl hang.pple], and *-tik eutsirik* [-abl hold.pple]. Drawings like *apple on twig* [27] and *lamp on ceiling* [63] are good examples of this spatial configuration.

5 Conceptualisation of space

In this section, I would like to touch on two issues that seem to play a role in how Basque conceptualises space: the opposition between dynamicity and stativity in spatial configurations and the role of agentivity.

All the drawings used as an elicitation tool in this study depict static topological scenes; that is, they are static representations of a figure located at some position in relation to a given ground. No information is given about the procedure (movement, change of position…) followed by the figure in order to reach that location or the agent that caused that state of affairs. However, it seems that informants consistently conceptualise some of these scenes as dynamic and some others with an implicit agent. Speakers have other options to describe these scenes statically, but they choose to include this type of information. At this stage, I can only say that these two elements are present in the conceptualisation of space in Basque but I would like to argue that dynamicity and agentivity are two intrinsic components of space in Basque. Further research will tell us whether these results are true or just a coincidence. Let us briefly look at some examples.

Dynamicity vs. stativity

If all these scenes are static, it is only natural to expect static descriptions of these drawings. As discussed in previous sections, the typical formula in Basque is: (spatial noun) + locative case + (adverbial participial of posture verb) + *egon* 'static be', where elements in brackets can be omitted. This formula is employed in the description of most of the drawings, but there are some scenes where speakers used more dynamic descriptions.

This dynamicity is often obtained by the use of a motion verb instead of the static verb. A good example is the picture *arrow through apple* [30], where speakers (39,3%) tend to use a motion verb such as *zeharkatu* and *gurutzatu* 'cross' instead of the static description *sagarran sartuta dago* [apple.loc enter.pple is.3sg]. The choice of a specific spatial case seems to be relevant too. In the drawing *insects on wall* [52], some informants prefer the ablative case, which implies more dynamicity, over the locative. A possible explanation is that these speakers distinguish between things that are statically on the wall and cannot move, and things like the insects that happen to be on the wall – attached to the wall – but that can change places.

Spatial nouns are also a good example for the distinction between dynamicity and stativity because some seem to be intrinsically dynamic such as *zehar* 'through, across' and *barren* 'interior, inside; lower part', and some others intrinsically static such as *gain* 'top, upper part'. As mentioned above, this is an area that requires further research but I would like to bring forward a case that seems to support this distinction between static and dynamic spatial nouns: the pair *barru* and *barren*. Both spatial nouns refer to 'interior, inside', but I would like to argue that their conceptualisation is different on the basis of dynamicity. Compare these two examples.

(4) *Kaia* *barru-an*
port inside-loc

(5) *Kaia* *barren-ean*
port inside-loc

Whereas *barru-an* in (4) only entails location at that port, *barren-ean* means not only that the boat is located at the port but also that the boat moved there. The result is the same, the boat is at the port, but their conceptualisation is different as represented in Figure 2.

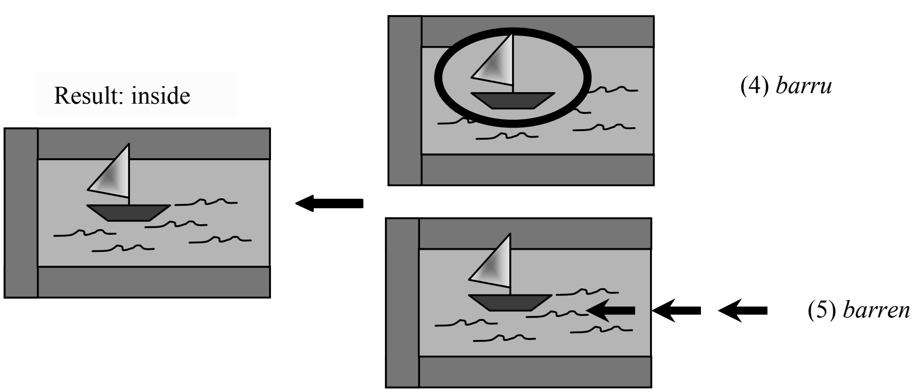


Figure 2. Stativity and dynamicity in *barru* and *barren*

In other words, if *barru-an* is used, the final state – location at the port – is profiled. It is assumed that the boat was at point *a* at time *x* – although this is information is neither present nor relevant for *barru-an* – but what *barru-an* tells us is that the boat is at the port at this time. If *barren-ean* is used, there is a double profiling: the final state – location at the port – and the process followed to reach that final destination – the boat moves from point *a* to the port.

In the elicited data, there is a very illustrative example that can shed some more light on the dynamic conceptualisation of *barren*. One of the informants uttered the following sentence for the description of *insects on wall* [52]:

(6)	*Xomorrok*	*paret-etik*	*barren-a*	*dabiltz*
	bug.abs.pl	wall.abl.pl	inside-all	move.3.pl

In (6), this informant uses the spatial noun *barren* with the allative case, which is usually translated as 'through'. Here, the idea is that the insects are located on the wall, got there because they moved from some other place, and are still moving or creeping along.

Agentivity

All the pictures in the data reproduce static topological relations where the figure is located in relation to some ground. In most cases, informants describe these pictures without mentioning the agent that causes that figure to be located at that position. Although one can assume that in drawing [5] *hat on head*, for example, the man put on his hat or that in drawing [31] *cat under table* the cat himself moved to that position, this information is neither represented in the drawings nor relevant for the speakers, and therefore, it is only natural that informants omit this type of information. There is, however, a group of drawings where informants do mention an implicit agent. For example, in picture [20] *balloon on stick*, more than half of the speakers (55%) use the dative instead of the locative as illustrated in (7) and (8).

(7)	*Globoa*	*makila-ri*	*lotuta*	*dago*
	balloon.abs	stick-dat	tie.pple	is.3sg

(8)	*Globoa*	*makilaren*	*punta-n*	*dago*
	balloon.abs	stick-gen	tip-loc	is.3sg

In these cases, the dative implicitly shows that somebody had to tie the balloon to the stick, whereas the locative only tells us that the balloon is located there. The use of the dative in some descriptions can be explained if we take into account the relationship between figure and ground. Although more research is needed, I would like to argue that the dative is used in those cases where the topological relationship is not natural, either because the figure is not in a typical position or because the figure cannot be located at the ground on its own without the help of an external agent. This would help us to understand why *apple on tree* [27] is described with the locative (*zuhaitz-ean* [tree-loc]) or the ablative (*zuhaitz-etik zintzilik* [tree-abl hang.pple]) and *balloon on stick* [20], *ribbon around candle* [4], and *clothespin on line* [33] with the dative (*makila-ri lotuta* [stick-dat tie.pple], *kandela-ri lotuta* [candle-dat tie.pple], *soka-ri lotuta* [rope-dat tie.pple]).

6 Conclusions

In this paper, I have offered an overview of the main linguistic means used in Basque for the lexicalization of space. Three are the main types of topological relational markers: spatial cases, spatial nouns, and motion verbs. Based on elicited data from the Topological Relation Picture Series, I have shown that the structure that Basque speakers mostly prefer for the description of static topological relations is the locative case and the static verb *egon* 'be'. Therefore, Basque can be classified as one of those languages with a general spatial term and a small set of locative/posture verbs – mainly used as adverbial participles. Speakers opt for spatial nouns in cases when the locative case cannot be utilised for the description of a spatial configuration or when it is semantically too uninformative and vague. To a lesser extent, informants also use the dative case and path motion verbs. This usage is restricted to very specific spatial scenes where an agent is implicitly needed and where the location is conceptualised as dynamic. With regard to the spatial information provided by these topological relational markers, I have found six main spatial components: LOCATION, VERTICALITY, CONTAINMENT, HORIZONTALITY (with intrinsic FOR information), DISTANCE-PROXIMITY (with relative FOR information), POINT-TO-POINT ATTACHMENT.

There is still a lot of work to do in order to get to grips on the conceptualisation of space in Basque. In the last part of this paper, I have pointed out that elements such as dynamicity and agentivity are to be taken into account since they seem to be present in the linguistic characterisations of some of the space scenes used as stimuli. The spatial noun *barren-ean* has the same meaning as *barru-an* 'inside', but its conceptualisation is different, *barren-ean* profiles not only final location at some point but also the dynamic procedure followed by the figure in reaching that destination. Another area that deserves more attention is the contrastive study among topological relational markers with a similar function. Spatial nouns like *ondo*, *alde*, and *albo* are all used for 'near' but we need to know whether they can be employed in the same contexts or whether there are any subtle differences among them. Another interesting issue for further research is the study of hierarchical relations between spatial nouns. Levinson et al (2003: 489) have found that languages with large sets of spatial adpositions are bound to be arranged in taxonomic trees where 'subordinate terms are more specific: they have, if one likes, additional features missing from their superordinate or more general terms'. Basque is one of those languages with numerous spatial resources, and hyponymic chains like *locative* 'location' → *barru* 'any interior, inside part' → *barne* 'interior, inside part delimited by ground' are worth investigating.

Notes

1 List of abbreviated morphemes: abl Ablative; abs Absolutive; adn Adnominal; all Allative; ben Benefactive; dat Dative; det Determiner; dir Directional Allative; erg Ergative; gen Genitive; ger Gerund; ter Terminative Allative; ind Indicative; indf Indefinite; inst Instrumental; loc Locative; perf Perfective; pl Plural; pple Participle; sg Singular. Morphemes are separated with a hyphen only in those cases when they are relevant for the discussion, otherwise they will be written together.

2 There is a great deal of discussion about the categorial status of these space elements in Basque grammars. I use the general theory-free term of *spatial noun* to cover what other authors call locative nouns (de Rijk, 1990), postpositions (Euskaltzaindia, 1991; Hualde y Ortiz de Urbina, 2003), adverbs (Bostak Bat, 1996) and locative cases (Laka, n.d.).

3 The citation form of Basque verbs in most dialects is the perfective participle. Basque is a language in which the majority of finite verb forms are largely analytical – also called periphrastic in the Basque grammar tradition – and as such, verb expressions are formed by a participle, which carries the lexical meaning and aspect information, and an auxiliary, which contains information about tense, mood and argument structure (cf. Hualde and Ortiz de Urbina, 2003: chapter 3.5, for a more detailed description of verbs in Basque).

4 Syntactically these are called 'complex predicates' in contrast with simple verbs and derived verbs which are 'simplex predicates' (Etxepare, 2003).

5 Examples taken from the data will always show the number of the corresponding drawing in square brackets.

References

Agud, A. (1980) *Historia y Teoría de los Casos*. Madrid: Gredos.

Ameka, F. K. and Levinson, S. C. (2007) Introduction: The typology and semantics of locative predicates. Posturals, positionals, and other beasts. *Linguistics* 45(5–6): 847–871.

Blake, B.J. (2001) *Case*. Cambridge: Cambridge University Press.

Bostak Bat. (1996) *Diccionario Hiru Mila Hiztegia*. Bilbao: Elkar.

Bowerman, M. (1996) Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, Peterson, M., Nadel, L. and Garrett, M. (eds) *Language and Space* 385–436. Cambridge, MA: MIT Press.

Bowerman, M. and Choi, S. (2001) Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman and Levinson, S. (eds) *Language Acquisition and Conceptual Development* 475–511. Cambridge: Cambridge University Press.

Etxepare, R. (2003) Valency and argument structure in the Basque verb. In J. I. Hualde and Ortiz de Urbina, J. (eds) *A Grammar of Basque* 363–425. Amsterdam and Philadelphia: John Benjamins.

Euskaltzaindia. (1991) *Euskal Gramatika: Lehen Urratsak I-II*. Iruñea: Euskaltzaindia.

Feist, M.I. (2004) Talking about space: A cross-linguistic perspective. In K.D. Forbus, Getner D. and Regier, T. (eds) *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* 375–380. Mahwah, NJ: Lawrence Erlbaum Associates.

Feist, M.I. (2008) Space between languages. *Cognitive Science* 32(7): 1177–1199.

Herskovits, A. (1986) *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*. Cambridge: Cambridge University Press.

Hualde, J.I. (2003) Derivation. In J. I. Hualde and Ortiz de Urbina, J. (eds) *A Grammar of Basque* 328–351. Amsterdam and Philadelphia: John Benjamins.

Hualde, J.I. and Ortiz de Urbina, J. (2003) *A Grammar of Basque.* Amsterdam and Philadelphia: John Benjamins.

Ibarretxe-Antuñano, I. (2001) An overview of Basque locational cases: Old descriptions, new approaches. *International Computer Science Institute Technical Report No.* 01–006. University of California, Berkeley.

Ibarretxe-Antuñano, I. (2004a) Polysemy in Basque locational cases. *Belgian Journal of Linguistics* 18: 271–298.

Ibarretxe-Antuñano, I. (2004b) Dicotomías frente a continuos en la lexicalización de los eventos de movimiento. *Revista de Lingüística Española* 34(2): 481–510.

Ibarretxe-Antuñano, I. (in press) Basque: Going beyond verb-framed typology. *Linguistic Typology*.

Ibarretxe-Antuñano, I. (In preparation) Towards a typological classification of motion verbs in Basque: Their structure and meaning. Unpublished Manuscript. Universidad de Zaragoza.

Jackendoff, J. (1993) *Semantics and Cognition*. Cambridge, MA: MIT Press.

Koch, H. (1984) The category of 'associated motion' in Kaytej. *Language in Central Australia* 1: 23–34.

Laka, I. (n.d.) *A Brief Grammar of Basque*. Retrieved on June 2005 from http://www.ehu.es/grammar/index.htm.

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

Levinson, S.C. (2003) *Space in Language and Cognition. Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.

Levinson, S.C., Meira, S. and Language and Cognition Group (MPI) (2003) 'Natural concepts' in the spatial topological domain – Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79(3): 486–516.

Majid, A., Bowerman, M., Kita, S., Haun, D.b.M. and Levinson, S.C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* II(3): 108–114.

Miller, G. and Johnson-Laird, P.N. (1979) *Language and Perception*. Cambridge, MA: Harvard University Press.

Pederson, E., Wilkins, D. and Bowerman, M. (1993) Static topological relations. *MPI Fieldmanual*.

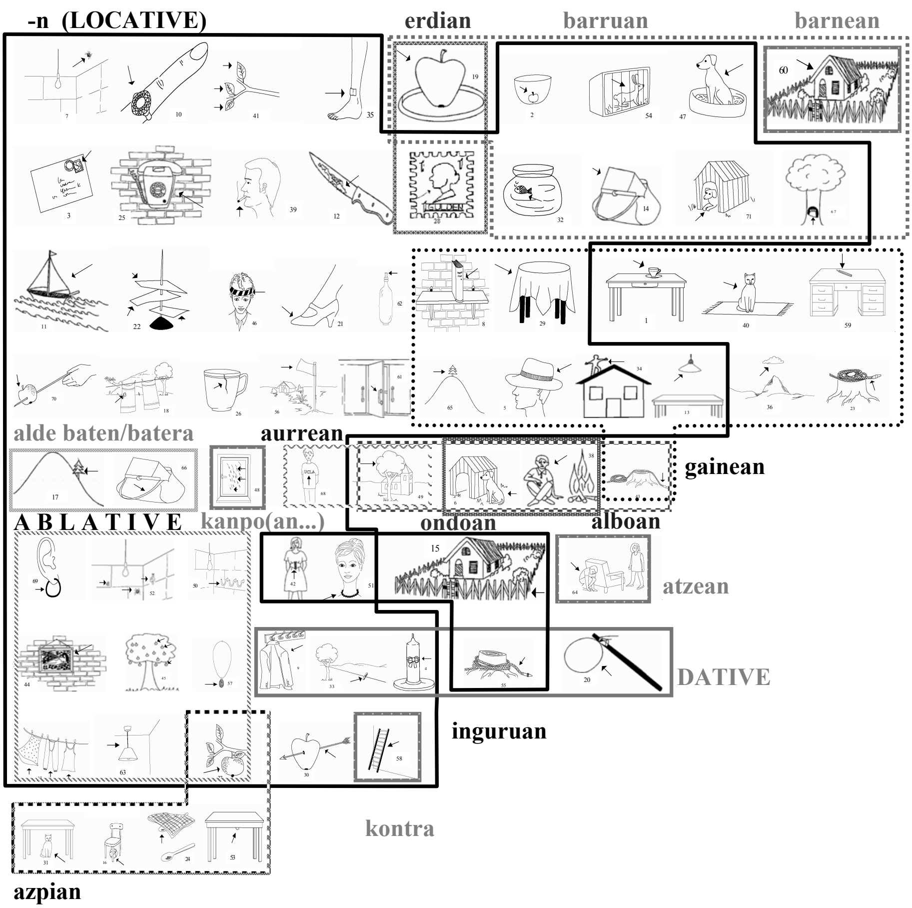
Pederson, E., Danzinger,E., Wilkins, D., Levinson, S.C., Kita, S. and Senft, G. (1998) Semantic typology and spatial conceptualisation. *Language* 74(3): 557–589.

Rijk, R.P.G.de (1999) Location nouns in Standard Basque. *Anuario del Seminario de Filología Vasca 'Julio de Urquijo'* XXVI(1): 3–20.

Sinha, C. and Kuteva, T. (1995) Distributed spatial semantics. *Nordic Journal of Linguistics* 18: 167–199.

Slobin, D. I. (1996) Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani and Thompson, S.A. (eds) *Grammatical Constructions. Their Form and Meaning* 195–219. Oxford: Clarendon Press.
Trask, R.L. (1997) *The History of Basque*. London: Routledge.
Vandeloise, C. (1991) *Spatial Prepositions: A Case Study from French*. Chicago: Chicago University Press.
Wilkins, D. (1991) The semantics, pragmatics and diachronic development of 'associated motion' in Mparntwe Arrernte. *Buffalo Papers in Linguistics* 1: 207–257.
Wilkins, D. (2004) The verbalization of motion events in Arrernte. In S. Strömqvist and Verhoeven, L. (eds) *Relating Events in Narrative: Typological and Contextual Perspectives* 143–157. Hillsdale, NJ: Lawrence Erlbaum.

Appendix: Basque topological relational markers and their distribution



11 Taking the Principled Polysemy Model of spatial particles beyond English: the case of Russian *za*

Darya Shakhova and Andrea Tyler

1 Introduction

Motivated by the goals of providing a replicable methodology and a theoretically grounded model of the polysemy networks of English spatial particles, Tyler and Evans (e.g. 2001, 2003) developed a model of semantic extension, termed the Principled Polysemy model. Relying primarily on established principles of language processing, such as embodied experience, taking multiple perspectives on a scene, and pragmatic inferencing, the model offered both a method for determining the central sense of a preposition and a more comprehensive accounting of the meaning extension mechanisms involved in the polysemy networks of English spatial particles.

The model emphasizes universal properties of human cognition, such as knowledge of real world force dynamics, leading Tyler and Evans (2003) to hypothesize that the model is likely to be applicable to many languages. However, Tyler and Evans' analyses have been based almost solely on English prepositions and the hypothesis concerning universal application is yet to be tested. A primary purpose of this paper is to begin to test the universality of the Principled Polysemy model by applying it to one of the most highly polysemous prepositions in Russian, *za*. A second purpose is to investigate how the model might be flexibly augmented when applied to a language whose system of spatial referencing includes a complex system of case marking, which is lacking in English.

Tyler and Evans (2001, 2003) noted that, in spite of many years of research, scholars still disagreed about the appropriate representation of the central meaning of prepositions, how to distinguish independent senses, and how to systematically account for the many extended meanings associated with each preposition. Their model aimed at providing a replicable methodology for determining the central sense and distinguishing independent, extended senses.

Under their analysis, the many meanings associated with a single spatial particle are represented as a complex category organized around a central sense, or the proto-scene. The meaning of a spatial particle is thus represented as a motivated, systematic polysemy network. Tyler and Evans (2003) posited that the central sense of a preposition represents a particular spatial relation between a focus element, the trajector (TR), a background or locating element, the landmark (LM), and a functional element (Vandeloise, 1991); the functional element represents the humanly meaningful consequences of the TR and LM being in a particular spatial configuration and is understood to be an integral part of the central sense.

The Principled Polysemy Model accounts for semantic extension through application of several established cognitive and linguistic principles such as: taking multiple perspectives on a scene, including highlighting subparts of a spatial scene; experiential correlation; knowledge of force dynamics; attention to the role of context in establishing an interpretation; and distributed meaning, i.e. appropriately attributing the overall interpretation of a preposition (or any lexical item) within an utterance to the meaning provided by all the elements of the utterance (including the rich social-cognitive context in which the utterance occurs).

Tyler and Evans provided a systematic account of many of the most commonly occurring prepositions in English. However, languages vary widely in how they represent spatial relations and in the patterns of polysemy associated with those spatial elements. For instance, Japanese employs a combination of particles and special nouns to indicate spatial relations. Languages such as Finnish and German employ a combination of spatial particles and case marking. In order to test the universality of the Principled Polysemy model, analyses of additional languages, with different grammatical mechanisms for representing spatial relations, is required. The purpose of the present work is to begin to test the universality of the Principled Polysemy model by applying the model to the Russian spatial particle, *za*.[1] *Za* is a particularly good test of the model because it is highly polysemous**.** In addition, the analysis of *za* entails considering how the Principled Polysemy model applies to a language whose system of spatial referencing includes a complex system of case marking, which is lacking in English. In this area of the analysis, we have been greatly influenced by Janda's (1993, 2000) work on Russian case.

2 The present study

2.1 The data

All the proposed senses are supported by attested uses of *za.* Our examples come from the following sources: (1) Ozhegov's Russian Language Dictionary (1984); (2) online dictionaries: Dal' Dictionary (1998), Dictionary 'Obschaĭa Leksika'; (3) grammar books: Poltoratsky and Zarechnak (1961), Pulkina and Zakhava-Nekrasova (2000); (4) on-line corpora: online corpus of the works of Dostoevsky, the National Corpus of the Russian Language. Approximately 1500 sentences in which *za* occurred were translated by Shakhova, a native speaker of Russian and English.

2.2 The problem

The Russian preposition *za* is associated with a broad range of meanings, e.g. standard Russian-English dictionaries list as many as 21 different senses including the following English prepositions: *behind, over, outside, beyond, after, for, at, by,* and *near*. To begin to get a sense of the challenge of representing the meaning of *za*, consider a small sample of its uses and their English translations:

(1) Behind:

Za	domom	275a	budet	vystroen	novyĭ	supermarket.
[**Za**	building-INST	275a	will be	built	new	supermarket-NOM]

The new supermarket will be built ***behind*** *Building 275a.*

(2) Over:

On	zhivët	**za**	rekoĭ,	na	ĭugo-vostoke.
[He-NOM	lives	**za**	river-INST,	at	south-east-PREP]

He lives ***over*** *the river, in the south-east.*

(3) For:

Fuks	zaplatil	**za**	svoë	osvobozhdenie	15000	dollarov.
[Fuks-NOM	paid	**za**	his	release-ACC	15000	dollars-GEN]

Fuks paid 15,000 dollars ***for*** *his release.*

(4) At:

Sima	poslushno	sela	**za**	pianino.
[Sima-NOM	obediently	sat down	**za**	piano-ACC]

Sima obediently sat down ***at*** *the piano.*

In addition to having a wide range of senses, *za* interacts with case in a complex way. Some senses occur with both Instrumental and Accusative Case, others occur with only one or the other.

To date, no unified analysis has been offered for the many, seemingly unrelated meanings associated with *za* nor its interaction with case. The grammar books do offer a few general rules, but for the most part, the meanings are considered as an arbitrary list. Dictionaries often attempt to represent the many meanings of *za* by defining a particular use in terms of another Russian preposition. The definitions are often misleading in that they usually fail to provide key elements of contextualized interpretation of *za*. *Ozhegov's Russian Language Dictionary*, one of the most widely used reference volumes for Russian language, defines *za* as *okolo,* 'near' or *vokrug* 'around' as in:

(5) *sidet'* ***za*** *stolom = sidet' okolo/vokrug stola*
'to sit **near/around** the table' (Ozhegov 1984)

However, while *za* denotes a spatial scene where the TR is positioned in close proximity to the LM (here the table) and therefore conforms to the notion 'near', in this context *za* also evokes an additional understanding that the TR is **purposefully** sitting proximal to the table. The same is true of a spatial scene involving multiple TRs positioned 'around' the table. Thus the full interpretation of *sidet'* ***za*** *stolom* includes the notion that the TR is sitting in order to use the table and prompts for an implicature of the TR's legs being under the table. Such a fine-grained configuration is not captured by the preposition *okolo* 'near' or *vokrug* 'around'. The analysis presented here argues that applying the Principled Polysemy model, in conjunction with Janda's analysis of Russian case, allows us to represent the range of meanings exhibited by *za,* including fine-grained interpretations illustrated above, as a systematic, motivated polysemy network.

3 The analysis

3.1 Establishing the proto-scene

Key to the successful analysis of any polysemy network is establishing the appropriate central sense, or in Tyler and Evans' terminology, the appropriate proto-scene, from which all other senses are held to derive. Tyler and Evans (2003) suggest several steps for establishing the proto-scene. Here we consider only three:

1) Examining the spatial configuration of LM and TR in multiple sentences in which the spatial particle is used.

2) Examining sentences that use contrasting spatial particles (members of a contrast set). Tyler and Evans established that a contrast set can involve aspects of the spatial configuration of the TR and LM, as in the contrast between English *over* versus *under*. Or a contrast set can involve the variations in the functional element, as in English *over*, whose functional element involves proximity or mutual influence between the TR and LM, versus *above*, whose functional element involves distance or lack of mutual influence between the TR and LM.

3) Frequency in the polysemy network. By this Tyler and Evans mean that, in the case competing analyses of the central sense, the majority of independent senses should be derivable from the proto-scene.

3.2 Case marking

Before examining the proto-scene, a few words about Russian case are in order. The Russian case system is rather complex and scholars have offered many analyses of just what the various cases mean. We have found Janda's (1993, 2000) analysis most convincing. Using a Cognitive Linguistic approach she carried an extensive analysis of Instrumental and Accusative case, the two cases with which *za* occurs. Under her analysis Instrumental primarily denotes stable physical configurations or 'setting'. Thus, it tends to contribute a sense of a static scene. In contrast, Accusative generally occurs in situations depicting motion. Janda's unique contribution here is the discovery that the most prototypical meaning of the Accusative involves a destination (which may involve an extended dimensionality) or the endpoint of an action.

3.3 The proto-scene

The proto-scene posited for the central sense of *za* involves a spatial configuration in which the LM is horizontally oriented away from the TR. In other words, the LM is conceptualized as having asymmetrical front and back, and the TR is conceptualized as being in back of the LM. In this respect, *za* is similar to the English preposition *behind*, as the following examples illustrate:

(6) The oriented LM:

(a) Rĭadom so mnoĭu sidel Vanĭa, a **za** nim Marusĭa.
[Next to I-INST was sitting Vanĭa-NOM, and **za** him-INST Marusĭa-NOM]
Vania was sitting next to me with Marusia (TR) behind him (LM).

(b) **Za** domom 275a budet vystroen novyĭ supermarket.
[**Za** building-INST 275a will be built new supermarket-NOM]
The new supermarket (TR) will be built ***behind*** *Building 275a (LM).*

In each of these examples, the LM is clearly inherently oriented and the TR's position is understood as being at the back of or 'behind' the LM. A common property of human perception is to also assign orientation to objects based on gravity, resemblance to human beings, behavior, or function. The next example serves as an illustration:

(c) Kogda prishël Tom, malysh sprĭatalsĭa **za** zerkalo.
[When came Tom-NOM kid-NOM hid **za** mirror–ACC]
When Tom came, the kid hid ***behind*** *the mirror.*

Here, the LM, the mirror, is functionally oriented, with the reflecting surface interpreted as the front.

While the landmark's orientation is salient in the proto-scene, the trajector's orientation is not specified. Consider the following examples:

(7) (a) Gde-to **za** nimi vozvyshalsĭa Ėl'brus.
[Somewhere **za** they-INST was rising Elbrus-NOM]
Elbrus was rising somewhere ***behind*** *them.*
(TR is unoriented.)

(b) Na urokakh, Kolĭa obychno sidit **za** Sëmoĭ.
[At classes-PREP Kolĭa-NOM usually sits **za** Sëma–INST]
In classes, Kolĭa sits ***behind*** *Sëma*.
(The TR, can be variably oriented, e.g., Kolĭa can be turned either toward or away from Sëma.)

These two examples show that the particular orientation of the TR is inconsequential for use of *za***.** However, were we to switch the orientation of the LM, a new preposition, like '*pered* 'in front of' would be required.

One of the criteria posited by Tyler and Evans (2003) for deriving the proto-scene of a spatial particle is examining how this particle interacts with members of a contrast set. A contrast set is a minimal pair of spatial particles that have complementary functions in dividing up the conceptual space along a particular dimension. For example, *up* and *down* divide the space along the vertical dimension, while *in front of* and *behind* divide up space along the horizontal dimension. Tyler and Evans hypothesize that the meaning component used to differentiate members of a contrast set is likely to be key to establishing the primary sense. In terms of spatial configuration, *za* most clearly contrasts with *pered* 'in front of' as in:

(8) (a) **Za** mnoĭ stoĭal stol.
[Za I-INST stood table-NOM]
*The table stood **behind** me.*

(b) **Peredo** mnoĭ stoĭal stol.
[Peredo me-INST stood table-NOM]
*The table stood **in front of** me.*

Unlike *za*, *pered* is associated with a relatively limited cluster of senses. Its most basic spatial meaning is a configuration in which a horizontally oriented landmark is 'facing' the trajector. Switching between *za* and *pered* in Russian results in switching the orientation of the landmark from facing the trajector (*pered*) to facing away from the trajector (*za*).

Za also forms a minimal pair with *pozadi* in terms of distance along the horizontal axis. *Za* indicates a proximal relationship while *pozadi* indicates a distal one (analogous to the distinction between English *over* versus *above* which divide spatial relations on the vertical axis).

(9) (a) On stoĭal **za** mnoĭ i sheptal mne v ukho.
[He-NOM stood **za** I-INST and whispered me-DAT in ear-ACC]
*He stood **behind** me and whispered in my ear.*

(b) **?**On stoĭal **pozadi** menĭa i sheptal mne v ukho.
[He-NOM stood **pozadi** I-INST and whispered I-DAT in ear-ACC]
*He stood **behind** me and whispered in my ear.*

(c) Magda sela obedat' **za** stol.
[Magda-NOM sat to dine **za** table-ACC]
*Magda sat down **at** the table to have dinner.*

(d) **?**Magda sela obedat' **pozadi** stola.
[Magda-NOM sat to dine **pozadi** table-ACC]
*Magda sat down **at** the table to have dinner.*

In these pairs the version containing ***pozadi*** strikes native speakers as odd because the contexts require proximity between the TR and the LM.

Considering all this evidence, we conclude that *za*'s central scene involves an oriented LM (represented in the diagrams by a 'nose'; in diagram 1, the 'nose' is on the LM and is pointing away from the TR) and a neutral (i.e. non-oriented) TR which is positioned at the back of and proximal to the oriented LM:

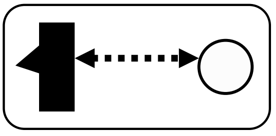


Diagram 1: The proto-scene for *za*
Oriented LM facing away from a proximal TR; vantage point is off-stage.

The functional element arising from this configuration is one of proximity or mutual influence or potential interaction between the TR and LM, analogous to the functional element posited for English *over*. Positing this functional element accounts for the notion of the purposefulness of the TR being positioned proximal to the LM, hence the interpretation that if a person is located *za* the table, she is positioned such that she can purposefully interact with the table.

3.4 The extended network

Tyler and Evans (2003) posited a basic principle for determining whether an extended sense exists independently. That principle was an independent, extended sense had to involve either a different spatial configuration than represented by the proto-scene or an independent non-spatial meaning. In either case, the extended meaning could not be inferred from the proto-scene as it occurred in the context of the utterance. For instance, Tyler and Evans posited an independent on-the-other-side sense for English *over* on the basis of sentences such as:

(10) *The boathouse is over the river from Rosslyn, under the Key Bridge* (linguist list, Oct. 2005).

They noted that unless the listener knew that *over* had an independent on-the-other-side sense, it would be impossible to appropriately interpret this sentence. In contrast, in a sentence such as:

(11) *The plane flew over the desert.*

The sense of motion derives from our understanding of the verb *fly,* which denotes motion, and our knowledge of planes. Contra several analyses, e.g. Dewell (1994), Tyler and Evans argued that with such sentences any sense of motion is derived from contextual inferences and should not be taken as evidence of a separate sense which contains the information +motion.

This analysis largely derives from the principle of distributed meaning. Following Sandra and Rice (1995), Tyler and Evans (2003) noted that many cognitive analyses of prepositions had fallen into what they termed the 'polysemy fallacy'. Essentially, the polysemy fallacy resulted in positing overly many senses. The polysemy fallacy arises from assigning aspects of the overall interpretation of an utterance, including regularly occurring implicatures, as part of the meaning of the preposition. Such analyses fail to appropriately determine which aspects of the interpretation are contributed by the various elements in the utterance. Similar to the flying plane example, Tyler and Evans argued that in an utterance such as, *The cat jumped over the wall,* the interpretation of the TR being in motion and following an arc-shaped trajectory comes from an integration of the central meaning of *over* (a TR located higher than, but proximal to a LM), the meaning of the verb *jump,* and application of our basic understanding of force dynamics, rather than a special sense of *over* that includes a trajectory and a TR in motion.

The principle of distributed meaning is particularly important in accurately analyzing Russian prepositions, as Russian prepositions combine with nouns (in LM position) which occur in various cases. Case marking is in part governed by the preposition itself. In the case of *za,* which combines with both Accusative and Instrumental case, the case marking appears to contribute to the exact interpretation of the scene being prompted for, for instance, whether the TR in the scene being depicted is interpreted as being static or involving motion. Consider the following sentences in which the case on the LM noun varies:

(12) On zhivët **za** rekoĭ, na ĭugo-vostoke.
[He-NOM lives **za** river-INST, at south-east-PREP]
He lives ***over (on the other side of)*** *the river, in the south-east.*

In this sentence, the LM carries Instrumental case and the interpretation involves a static scene. Our understanding of the scene is that there is no TR in motion and no trajectory involved. Use of Accusative case is unacceptable.

(13) Potom ĭa uekhal **za** granitsu.
[Then I-NOM left **za** border-ACC**]**
Then, I moved abroad.

In contrast, in example (13), the LM is marked with Accusative case. The interpretation is that the TR is in motion and therefore a trajectory is involved. Use of Instrumental case is unacceptable. In both these examples, case alone does not establish the scene as static or involving motion and destination. Certainly the verbs and our background knowledge contribute to the interpretation, but case is consistent with these meanings.

Under certain analyses, e.g. Lakoff (1987), these two instances of *za* would constitute two separate senses, on the grounds that one involves motion while the other does not. Under such an analysis, the assumption is that movement (or non-movement) is part of the meaning of a particular sense of the preposition. Following the principle of distributed meaning, we posit that these two instances of *za* constitute one independent sense which means 'beyond or on-the-other side'. The exact interpretation of +/- movement is not part of the semantics of the preposition, but rather provided by case marking and other elements of the sentence (such as the verb *uehkal* 'left', which prompts for motion, and *zhivët* 'lives', which prompts for a static scene). Thus, in our analysis of the polysemy network for *za,* we do not posit separate senses based on case even though there are differences in the exact understanding of the spatial scene being depicted when Accusative occurs versus Instrumental.

As mentioned in the introduction, the Principled Polysemy model also identified a set of cognitive mechanisms by which the central sense of a preposition could be extended to create independent, distinct senses. These mechanisms are all independently established in the literature. They include (1) multiple ways of viewing a scene (Langacker, 1987); (2) knowledge of real world force dynamics (Talmy, 2000); (3) making pragmatic inferences based on the linguistic prompts and background knowledge (Grice, 1975; Wilson and Anderson, 1986). Tyler and Evans (2003) argued that in most cases an extended sense could be traced back to an utterance in which the proto-scene (or an established sense derived from the proto-scene), in conjunction with the context, created a novel interpretation of the preposition. After repeated uses, such a contextualized use of a preposition could be established as an independent sense in the network. Once the meaning was established in the network, the context which originally gave rise to the new sense would no longer be needed in order for the speakers to interpret the preposition. The many meanings associated with a preposition were thus represented as a motivated polysemy network.

Our analysis of the 1500 naturally occurring examples of *za* revealed that, in addition to the proto-scene, five independent senses occur with both Instrumental and Accusative. We term these senses the Shared Network. The following sentences illustrate each of these senses:

(14) Behind-Deictic orientation:

(a)

Misha	sprĭatalsĭa	**za**	kustikom	naprotiv	kamysheĭ	i
[Misha-NOM	hid	**za**	(small) bush-INST	opposite	rushes-GEN	and

stal zhdat'.
waited]
*Misha hid **behind** a small bush opposite the rushes and waited.*

(b)

Alekseĭ	s siloĭ	shvyrnul	v	gruzoviki	limonku-ACC	i
[Alekseĭ-NOM	forcefully	hurled	at	trucks-ACC	grenade-ACC	and

pryglnul	**za**	kuchu	khvorosta.
leaped	**za**	pile-ACC	brushwood-GEN]

*Alekseĭ forcefully hurled a grenade at the trucks and leaped **behind** the pile of brushwood.*

(15) Functional:

(a) Kak i prezhde, vechera oni korotali **za** chteniem
[As before, evenings-ACC they-NOM whiled away **za** reading-INST

vslukh.
aloud]
*As before, they whiled away their evenings **at** reading aloud.*

(b) On na sekundu pripodnĭal golovu i snova prinĭalsĭa
[He-NOM for second-ACC raised head-ACC and again began/set to

za chtenie.
za reading-ACC]
He raised his head for a second and then went back to reading.

(16) Beyond/on-the-other-side 'over':

(a) Ĭa odin provozhal bol'nuĭu starukhu **za** reku.
[I-NOM alone accompanied ailing old woman-ACC **za** river-ACC]
*I alone accompanied the ailing old woman **to the other side** of the river.*

(b) Tam, **za** rekoĭ… uzhe zagoralis' pervye zvëzdy.
[There, **za** river-INST… already were lighting up first stars-NOM]
*There, **over** the river, first stars were already lighting up.*

(17) Focus of attention:[2]

(a) Mama volnuetsĭa **za** devochku.
[Mom-NOM is worried **za** girl-ACC]
*Mom is worried **for** the girl.*

(b) Moĭa mat' obeshchala pereekhat' k nam i smotret' **za**
[My mother-NOM promised to move in with we-DAT and look **za**
det'mi.
children-INST]
*My mother promised to move in with us and look **after** the children.*

(18) Purpose:

(a) Stol' ozhestochënnaĭa bo'rba **za** golosa razvernulas' eshchë i v
[Such fierce battle-NOM **za** votes-ACC unfolded also in

predvidenii nizkoĭ ĭavki.
anticipation-PREP low turn-out-GEN]
*Such fierce battle **for** votes unfolded in anticipation of low turn-out.*

(b) My zabespokoilis' i poslali **za** doktorom Brusesom.
[We-NOM started to worry and sent **za** doctor-INST Bruses-INST]
*We started to worry and sent **for** doctor Bruses.*

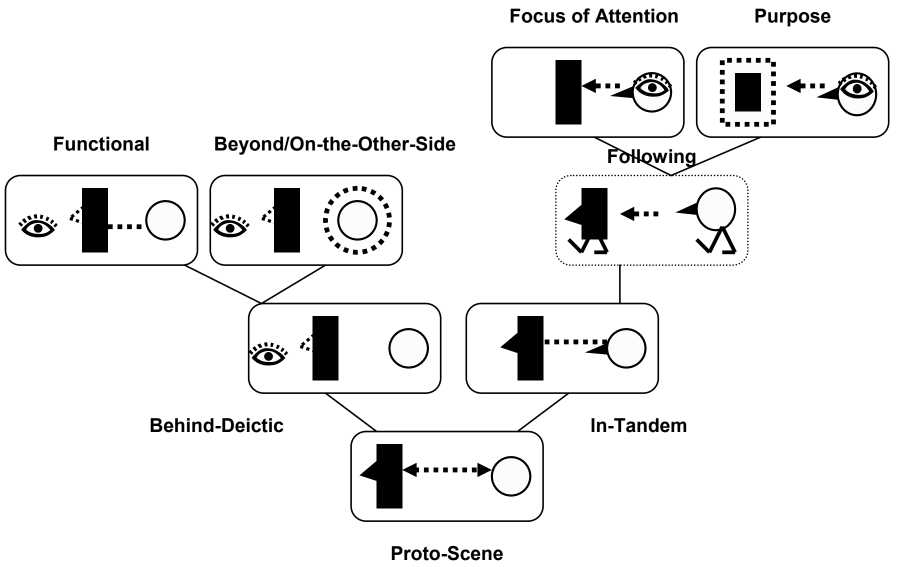


Diagram 2: The Shared Network

We found a number of additional senses that occur in conjunction with one of the cases but not the other. In these instances, the data indicate that distinct senses arise as a result of an extended sense of *za* as it combines with particular aspects of meaning associated with either Instrumental or Accusative case.

3.5 Accusative network

Sentences illustrating the senses which occur only with Accusative case:

(19) Exchange:

Fuks	zaplatil	**za**	svoë	osvobozhdenie	15000 dollarov.
[Fuks-NOM	paid	**za**	his	release-ACC	15000 dollars-GEN]

*Fuks paid 15000 dollars **for** his release.*

(20) Substitution:

Luchshe	gluptsa	priniat'	**za**	umnogo,	chem	umnogo	**za**	gluptsa.
[Better	fool-GEN	to take	**za**	sage-ACC,	then	sage-GEN	**za**	fool-ACC]

*Better to take a fool **for** a sage, than a sage **for** a fool.*

(21) Contact:

Militsioner	vzial	starushku	**za**	ruku	i	perevël	cherez
[Militiaman-NOM	took	old lady-ACC	**za**	hand-ACC	and	walked	across

dorogu.
road-ACC]

*The militiaman took the old lady **by** the hand and walked her across the road.*

(22) Support:

Gosduma	progolosovala	**za**	prinïatie	zakonoproėkta.
[State Duma-NOM	voted	**za**	approval-ACC	bill-GEN]

*The State Duma voted **for** approval of the bill.*

(23) Cause:

Neozhidanno	polkovnik	razozlilsïa	na	sebïa	**za**	svoïu
[Unexpectedly	colonel-NOM	became angry	at	himself-ACC	**za**	his

sentimental'nost'.
sentimentality-ACC]
*Unexpectedly, the colonel became angry with himself **for** his sentimentality.*

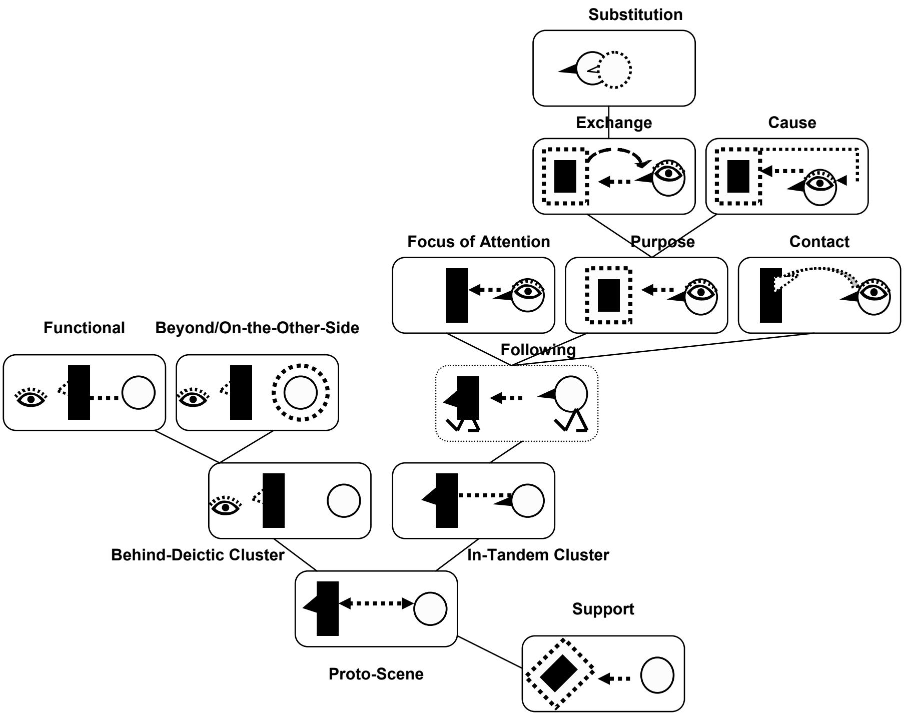


Diagram 3 The Accusative Network

3.6 Instrumental Network

We also found many examples of senses which occur only with Instrumental case:

(24) Covering:

Za	shumom	voln	nichego	nel'zïa	bylo rasslyshat'.
[**Za**	sound-INST	waves-GEN	nothing	impossible	was to hear]

*It was impossible to hear anything **behind** the sound of waves.*

(25) Obstacle:

Za neimeniem sredstv izdatel'stvo 'Prosveshchenie'
[**Za** lack-INST resources-GEN publisher-NOM 'Prosveshchenie'-NOM

prekratilo izdanie zamechatel'noĭ knigi.
stopped publication-ACC wonderful book-GEN]
For *lack of financial resources, the publishing company 'Prosveshchenie' stopped its publication of a wonderful book.*

(26) Following:

Za mnoĭ, moĭ chitatel', i tol'ko **za** mnoĭ, i ĭa
[**Za** me-INS, my reader-NOM, and only **za** me-INS, and I-NOM

pokazhu tebe takuĭu lĭubov'!
will show you-DAT such love-ACC]
Follow *me, my dear reader,* ***follow*** *me, and I will show you such love!*

(27) Sequence:

Odin **za** drugim na ėkrane krutĭatsĭa fil'my pro
[One **za** another-INST on screen-PREP play films-NOM about

zalozhnikov.
hostages-PREP]
One ***after*** *another, films about hostages play on the screen.*

(28) Possession:

Brat'ĭa dali **za** nevestoĭ ogromnoe pridanoe…
[Brothers-NOM gave **za** bride-INST large dowry]
The brothers gave the bride a large dowry…

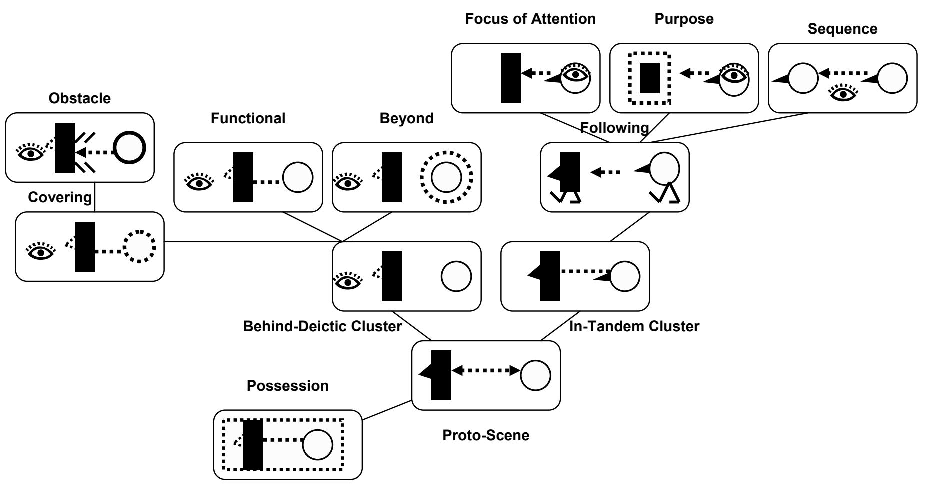


Diagram 4: The Instrumental Network

3.7 Motivation for specific senses

Because of space limitations, we cannot provide a detailed justification for all the spatial scenes in the network; thus, we will look at only five. However, we note that our analysis revealed striking similarities between the spatial scenes for senses of *za* often translated by English prepositions such as *over* and the scenes posited for either the proto-scenes or extended meanings of these English prepositions. We will illustrate this in our discussion of the 'on-the-other-side' sense of *za*. There are also notable similarities between the 'covering' sense (often expressed by 'over' as in *She placed her hands over her eyes);* the 'purpose', 'cause', and 'exchange' senses (often translated as 'for'), the 'in-tandem' sense (often translated as 'after' or 'behind'), and the focus of attention sense (often translated as *at).*

Key to a systematic analysis of the majority of the senses involved in the network is recognizing two major clusters of senses, the 'Behind-Deictic Center' cluster and the 'In-tandem' cluster. A cluster involves a key extension from the proto-scene, which in turn forms the basis for other extended senses. Tyler and Evans also found clusters of senses within the polysemy networks of English prepositions.

The first main extension we will consider is the Behind-Deictic Center cluster, as represented by the following diagram:

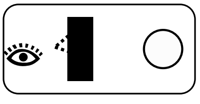


Diagram 5: Behind-Deictic Center

Note that the spatial scene represented in this diagram looks very like the proto-scene except that the vantage point has shifted from off-stage to on-stage (as represented by the eye to the left of the LM). The Deictic Orientation scene represents a natural extension of the proto-scene in that humans in their everyday lived experience are constantly viewing the same spatial configurations from varying vantage points. Indeed, Clark (1973) has argued that the human perceptual system crucially relies on humans constantly shifting perspective. Langacker (1987, 1991) has identified changing perspectives on spatial scenes as a key cognitive process that has multiple manifestations in language.

An important consequence of this particular shift in vantage point from off-stage to on-stage and in front of the LM is a shift in how orientation of the LM is established. This is essentially a shift from intrinsic orientation in the proto-scene that emanates from the nature of the LM itself to a deictic orientation imposed by the 'vantage point'. Although the LM is still understood as being oriented, the mechanism by which orientation is assigned is different. Determining the front and back of a LM by means of deictic orientation has the effect of de-emphasizing whatever intrinsic orientation that might be inherent in the LM. This de-emphasis is represented in the diagram by the 'nose' of the LM appearing in broken lines rather than solid lines. The following is an example from our corpus illustrating the 'Behind-Deictic' sense:

(29)	**Za**	zaborom	roslo	derevo.
	[**Za**	fence-INST	was growing	tree-NOM]

The tree was growing ***behind*** *the fence.*

Note that this sentence could be uttered by someone standing inside the area enclosed by the fence, looking out, or standing outside the enclosed area, looking in. This ambiguity of the scene being depicted clearly shows that the LM, the fence, does not have an intrinsic orientation, in the sense that a human is intrinsically oriented with a front and back; nor does the fence have an intrinsic functional orientation as a mirror or house would have. What is understood as 'in front of' the fence or 'behind' the fence depends on the viewer's perspective. In other words, in the scene described here, the orientation of the LM is assigned by the viewer's vantage point. (For additional discussion of similar shifts in perspective, see Zinken, this volume).

Now we turn to an examination of one of the extended senses within the 'Behind-Deictic Orientation' cluster – the 'Beyond' sense.

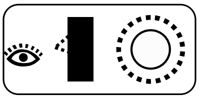


Diagram 6 The Beyond sense

As with the other extensions, the change in this spatial scene and the one it is linked to, that is the Deictic Center sense, is incremental. Here the change involves a shift in interpretation of the TR. In both the proto-scene and the Deictic Center scene, the TR is neutral, that is, it is not highlighted or given particular salience. In the spatial scene associated with the Beyond sense, the interpretation of the TR and its location in relation to the LM are highlighted. Such highlighting represents a shift in perspective (Langacker, 1987) or a shift in the conceptualization of the scene. In the Beyond scene the viewer is particularly focused on the location of the TR that is highlighted. Highlighted status is represented in the diagram by the dotted line ringing the TR. The following sentences illustrate this use of *za*.

(30)	On	zhivёt	**za**	rekoĭ,	na	ĭugo-vostoke.
	[He-NOM	lives	**za**	river-INST,	at	south-east-PREP]

He lives ***over*** *the river, in the south-east (beyond/on the other side of).*

In this sentence, the LM is the river. The speaker is standing on one side of the river and focusing on the location of the TR ('he' is a metonymy for 'his' house) on the opposite side of the river. There also seems to be a shift in conceptualization of the LM. In the proto-scene the LM serves a neutral locating function, as in:

(31)	**Za**	mal'chikami	vozvyshalsĭa	staryĭ	dub.
	[**Za**	boys-INST	was towering	old	oak-tree-NOM]

Behind *the boys towered an old oak tree.*

In the Beyond scene, the LM is conceptualized as a barrier or boundary between the on-stage viewer and the TR. The TR is conceptualized as being at a distance from the viewer, i.e. on the other side of the barrier.

In English, either *over* or *on-the-other-side* can be used to appropriately convey the interpretation of sentence 30. This scene is strikingly like the spatial scene for the on-the-other-side sense of *over* in English, with an on-stage perspective point and a LM conceptualized as a barrier or boundary between the on-stage viewer and the TR as in the sentence,

(32) *Arlington is over the river from Georgetown.*

Notice that in sentence (30) the LM, river, is marked with Instrumental case. Following Janda's analysis, we hypothesize that Instrumental indicates a static scene in which no trajectory occurs.

Now consider a second sentence illustrating the Beyond sense:

(33) Potom ĭa uekhal **za** granitsu.
[Then I-NOM left **za** border-ACC]
Then, I moved abroad. (Then, I left to the other side of/ over/beyond the border.)

Here we infer that the TR, *ĭa*/I, was on one side of the border and crossed it, ending up on the other side of the border. We also understand that the purpose of the movement was to reach a destination. Note that this movement/destination interpretation is prompted for by the verb and Accusative case on *granitsu*, the LM.

3.8 The In-tandem cluster

The second main extension revealed by the data is what we term the 'In-tandem' sense, as represented by diagram 7:

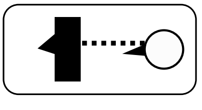


Diagram 7 The In-tandem scene

Note again, that this spatial scene looks very similar to the proto-scene. The only difference is that the TR is oriented, as well as the LM. This is represented by the 'nose' on the TR, which is fully darkened to represent the importance of the TR being oriented.

In our real world, lived experiences, we often observe two entities, often animate, aligned in what Hill (1978) calls an 'in-tandem' configuration. In such a configuration, both the LM and the TR are interpreted as being oriented in the same direction, one behind the other. An unavoidable consequence of both entities being oriented is the potential introduction of a second vantage point, that of the TR, if the TR is animate. This is an important distinction from the proto-scene. The following sentence exemplifies the In-tandem configuration in which both the LM and TR are inherently oriented and an on-stage perspective point residing with the TR exists:

(34) Na kontserte mne bylo plokho vidno, potomu chto ǐa sidela
[At concert-PREP I-DAT was poorly visible because I-NOM was sitting

za vysokim muzhchinoǐ.
za tall man-INST]
*At the concert, I couldn't see much, because I was sitting **behind** a tall man.*

Here the speaker, the TR, clearly has a perspective point. However, not all instances of this sense require the TR to have a perspective point.

(35) Prǐamo **za** muzhchinoǐ sidel bolshoǐ plǐushevyǐ mishka.
[Directly **za** man-INST was sitting big teddy bear-NOM]
*Sitting directly **behind** the man, was a large teddy bear.*

Here both the TR and LM are oriented, but the TR has no vantage point.

(36) …drug **za** drugom, v odin rǐad, opustiv golovy, shli
[one **za** other-INST, in one line-ACC, heads-ACC down, walked

desǐatki sobak.
dozens-NOM dogs-GEN]
*Heads down, dozens of dogs walked in one line, one **behind** the other.*

Here both the TR and LM are oriented, but the vantage point is off-stage.

Recall we noted Tyler and Evans (2003) argued that if a particular interpretation was derivable from integrating the proto-scene, the other elements in the sentence and our knowledge of the world, then it should not be considered an independent sense in the polysemy network. What that analysis overlooked is that there can be instances of a spatial scene that have particular properties, which, although they could be inferred from context, occur so frequently that they may become entrenched in memory and thus become part of the polysemy network. We suggest that the In-tandem sense is such a case. Consider the sentence: *I couldn't see because I was sitting behind a tall man*. With our basic knowledge of the human body and the proto-scene for *za* we can infer that both the LM and TR are oriented and that the LM and TR are aligned such that the TR is posterior to the LM. Thus the interpretation of the scene could be derived through on-line processing and inferencing. But it is important to note that this configuration does add two components to the basic proto-scene, namely that the TR must be oriented and the LM and TR must both be aligned such that they are facing in the same direction. As we saw earlier, these are not requirements for the proto-scene.

Hill (1978) noted that in-tandem alignment is ubiquitous and hypothesized that it has an important cognitive status. Moreover, the Russian data indicate that the In-tandem scene forms the basis for a wide range of *za*'s extended meanings. It seems questionable to argue that there is no independent, entrenched representation of the In-tandem sense in the polysemy network if it forms the basis of many established meanings. Thus we posit the In-tandem sense as an independent sense in the network and designate it as the foundational scene in the In-tandem cluster.

3.9 The Following sense

If the two in-tandem entities are in motion, the one we experience as first encountered is interpreted as **leading**, the entity we experience as second encountered is interpreted as **following** the first.

(37)	Muzhchina	shël	**za**	kolonnoĭ	demonstrantov.
	[Man-NOM	was walking	**za**	line–INST	demonstrators-GEN]
	The man was ***following*** *a line of demonstrators.*				

Interpreting the LM and TR as being in a leading/following relationship assigns a degree of intentionality to the spatial-physical relationship. This sentence might be translated at some 'literal' level as 'The man was walking behind a line of demonstrators' but this translation loses the native speaker's interpretation that the man didn't just happen to be taking a stroll and inadvertently ended up taking the same route as the demonstrators. Rather native speakers interpret this sentence to mean that the man was purposefully walking behind or following the demonstrators. We represent this notion of intentionality on the part of the TR by the eye in the TR's head, which emphasizes the animacy and viewpoint of the TR. Another distinguishing characteristic of the Following sense is that the LM and TR **must** be in motion. We designate this by representing both the LM and TR as walking.

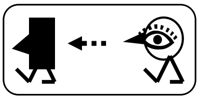


Diagram 8 The Following sense

3.10 The Following sense and case

Let's return to sentence 37. Note that in this sentence *kolonnoĭ* 'line' (lit. 'column') occurs in the Instrumental case. At first blush, this may seem strange as it is often the case that if the TR is in motion, Accusative case is used (as in examples 6c and 13). However, in our rather extensive corpus, we found no instances of the Following sense occurring with Accusative case. We hypothesize that this is so because while in this scene the TR may be close to the LM and moving along with the LM, the TR doesn't actually reach the LM. Moreover, the LM seems not to be conceptualized as a destination. For example, there is nothing in sentence 37 that implies the man is trying to catch up with the demonstrators. We hypothesize that in the Following scenes, even though the LM and TR are in motion, there is no change in the TR's position vis-a-vis the LM. In other words, this is a stable scene, in which the LM is not conceptualized as a destination.

This analysis is consistent with Janda's analysis of the central meanings of Accusative and Instrumental case. By Janda's analysis, the prototypical meaning of Instrumental is scene setting, while the prototypical meaning of Accusative case indicates a destination or end point of motion, not that the noun marked with Accusative is in motion. According to Janda, this interpretation holds for the meaning of Accusative case with all Russian prepositions as well as case marking of the direct objects (which are analyzed as representing the end point of the motion, or in Langacker's terms, the energy sink). Thus, since the LM is in Following sense does not represent a destination, the LM is marked with Instrumental case. A similar analysis applies to the In-tandem sense.

3.11 The Purpose sense

Even though the In-tandem sense does not co-occur with the Accusative, a number of senses that derive from the In-tandem scene can co-occur with the Accusative. By our analysis this is possible because once a scene is entrenched in memory it is subject to re-analysis. This includes being viewed from different perspectives, which can potentially give rise to re-interpretations, new implicatures and eventually semantic extensions. For instance, the stable LM-TR relationship prompted for by the Following sense can be reinterpreted such that the LM is understood as a goal. As a goal, the LM is privileged or highlighted within the scene. This shift in conceptualization reflects the common experience of humans being in a following situation and having the additional desire to reach the LM. Thus, the LM is no longer simply a neutral locater for the TR. Consider the English sentence *The hunter followed the fox.* An inference that fits our schema for fox hunting includes the notion that capture of the fox (which necessarily entails physically reaching the fox) is likely the hunter's goal. Because we understand from our lived experience that following often includes the goal of reaching the LM, a Purpose sense has been extended from the Following sense.

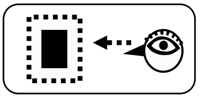


Diagram 9: The Purpose sense

The diagram representing the spatial scene we posit for the Purpose sense is similar to the Following scene. There are two main differences. The LM is reconceptualized as a goal to be reached by the TR. This privileging is indicated by dotted lines surrounding the LM. Second, the LM is no longer explicitly oriented. Key elements are an intentional TR (who has a particular vantage point on the scene), the interaction between the TR and LM (a continuation of the functional element in the proto-scene) and the LM (re)conceptualized as a goal. Importantly, once the LM is reconceptualized as a goal and the Purpose sense becomes entrenched in the network, the exact qualities of the original scene, in this case the orientation of the LM, can drop away. Again, in our lived experience we often intentionally move towards entities that we conceptualize as goals that are not necessarily inherently oriented. The Purpose sense co-occurring with the Accusative is exemplified below:

(38) Shkoly	sorevnovalis'	**za**	luchshuĭu	uspevaemost'.
[Schools-NOM	competed	**za**	best	results-ACC]

*Schools competed **for** best results.*

Here it seems clear that the schools' purpose in competing is to obtain the best results.

The data from our corpus show that the interaction of case with the Purpose sense is rather complex. The Purpose sense can co-occur with either the Instrumental or the Accusative. However, unlike examples we saw in the Deictic Center cluster, whether or not the TR is in motion does not explain the distribution of case. In the following sentence, the sentence-level TRs are in motion, but 'milk' appears in the Instrumental:

(39) My	s	mamoĭ	poshli	v	magazin	**za** molokom.
[We-NOM	with	mom-INST	went	to	store-ACC	**za** milk-INST]

*Mom and I went to the store **for** milk.*

We hypothesize the choice of case in the Purpose sense has largely to do with the underlying schema prompted for by the verb. The verb 'go' prompts for a Path schema with a TR moving along the Path. In sentence 39, we find an overtly articulated physical destination, 'the store', which is marked with the Accusative case, as we would expect from Janda's analysis. The TR following *za* articulates the reason for moving to the particular physical location. The physical destination does not need to be overtly articulated in order for the concept of a destination to be available:

(40) Mama	poshla	**za**	molokom.
[Mom-NOM	went	**za**	milk-INST]

*Mom went **to get** milk.*

Here the physical destination is not overtly articulated, but a physical destination can be easily deduced from our background knowledge about how and where we obtain milk. Importantly, the notion of physical destination is coherent with the semantics of the verb 'to go'. 'To go' is a member of a class of verbs that can be called 'directional', which prompt for scenes that involve destinations. Following Tyler and Evans' (2003) account, destination is part of the general Path schema. They argue that Path is 'a consequence of an endpoint or goal being related to a starting point or locational source by virtue of a series of contiguous points. That is, the concept of path requires a particular spatial goal…' (Tyler and Evans, 2003:218) This understanding of Path would predict that 'directional verbs' will co-occur with endpoints marked by Accusative.

Now re-consider example 38, reproduced below for ease of argumentation:

(38)	Shkoly	sorevnovalis'	**za**	luchshuĭu	uspevaemost'.
	[Schools-NOM	competed	**za**	best	results-ACC]

*Schools competed **for** best results.*

This example employs a 'non-directional' verb, 'compete', that does not prompt for a scene involving the prototypical Path schema and a particular spatial destination. Although we clearly understand that the competition took place somewhere, a path with a particular physical destination does not seem to be part of the scene. Rather the reason for or purpose of the activity seems to also be the desired endpoint of the activity. In other words, non-directional verbs seem to prompt for a different schema in which purpose involves some of the attributes of destination or endpoint. With directional verbs, destination and purpose are clearly distinguished.

This distinction is illustrated in the following minimal pairs:

(41)	Ehat'	na	Olimpiĭskie	Igry	**za**	zolotoĭ	medal'ĭu.
	Go	to	Olympic	Games-ACC	**za**	gold	medal-INST]

*Go to the Olympic Games **for** the gold medal.*

Here our understanding of the scene involves a Path schema with France as the physical destination. The purpose for going to the Olympic Games is to obtain the gold medal, which is marked with Instrumental case.

(42)	Borot'sĭa	na	Olimpiĭskikh	Igrakh	**za**	zolotuĭu	medal'.
	[Fight	at	Olympic	Games-PREP	**za**	gold	medal-ACC]

*Fight **for** the gold medal at the Olympic Games*

Here our understanding of the scene does not involve a path with a physical destination. The purpose for competing at the Olympic Games is to win the gold medal; winning the gold medal is the desired endpoint of the activity. These examples demonstrate that if a Path schema is prompted for, as in sentences 39 and 41, the physical destination is coded by Accusative case and the purpose is coded by Instrumental case. In contrast, if a Path schema (with a spatial destination) is not prompted for, as in sentences 38 and 42, the purpose is interpreted as the desired endpoint of the activity and is coded by Accusative case.

A consequence of this entrenched distinction in interpretation of destination or endpoint with directional verbs versus non-directional verbs is that a purpose phrase that is marked with Instrumental case tends to be physical or concrete, while a purpose marked with Accusative tends to be more abstract. This explains the oddity of the following examples:

(43) (a)

Poĭti	za	molokom
[Go	for	milk-INST]

(b)

?*Poyti	za	obrazovaniem
[?*Go	for	education-INST]

In sentence 43a, we see the familiar directional verb 'go' plus purpose in which 'milk' (the purpose for going) is marked with Instrumental. However, in sentence 43b, even though the directional verb 'go' plus the Purpose sense co-occurs with the expected Instrumental case, native speakers judge the sentence as odd. We believe the sense of oddness stems from the non-physical TR, 'education', being marked with Instrumental case. We find a similar pattern in sentence 44.

(44) (a)

Borot'sĭa	za	obrazovanie
[Fight	for	education-ACC]

(b)

?*Borot'sĭa	za	moloko
[?*Fight	for	milk-ACC]

In sentence 44a, the non-directional verb 'fight' co-occurs with a Purpose phrase whose TR is non-physical and marked with the expected Accusative case. In contrast, in 44b, even though 'fight' plus the Purpose sense occurs with the typical Accusative case, the sentence sounds odd to native speakers. We hypothesis that this oddity arises from the physical TR, 'milk', being marked with the Accusative case. (We do note that certain contexts can be created in which these questionable sentences sound less odd).

To conclude our discussion of case and the Purpose sense, we can represent the case distribution for the Purpose sense of *za* with the following patterns:

(45) Instrumental:

[N + destination verb + (destination -ACC) + **za** (purpose) + N (physical)-**INST**]
Zvonarëva + returned + (to Memphis) + **za** + victory-**INST**
Zvonarëva returned to Memphis ***for*** *victory.*

Zvonarëva	vernulas'	v	Memfis	**za**	zolotoĭ	medal'ĭu
[Zvonarëva-NOM	returned	in	Memphis-ACC	**za**	gold	medal-INST]

(46) Accusative:

N + non-destination verb + (location) + **za** (purpose) + N (non-physical)-**ACC**]
Zvonarëva + borolas + (v Memfise) + **za** + victory-**ACC**
Zvonarëva fought ***for*** *victory in Memphis.*

Zvonarëva	borolas'	v	Memfise	**za**	zolotuyu	medal'.
[Zvonarëva-NOM	fought	in	Memphis-PREP	**za**	gold	medal-ACC]

4 Conclusion

In this chapter we have demonstrated that Tyler and Evans' (2003) Principled Polysemy model can be successfully extended to languages other than English. By applying the basic principles laid out in the Principled Polysemy model, we have been able to provide a systematic, motivated analysis for the highly polysemous Russian preposition, *za*. The analysis revealed that while the proto-scene for *za* appears to bear similarities to that of English *behind*, many of the spatial scenes associated with *za's* extended senses are quite different. Indeed, a number of the extended senses associated with *za* that are standardly translated into various English prepositions, such as *over* and *for,* represent spatial scenes that are very similar to the spatial scenes associated with the extended senses of these English prepositions. For instance, the spatial scene associated with the on-the-other-side sense of *za,* one of the extended senses in the Deictic Center cluster, is very similar to the spatial scene Tyler and Evans posit for the on-the-other-side sense of *over,* one of the extended senses from the ABC trajectory cluster. This is one of the contexts in which *za* is regularly translated as 'over'. This is consistent with Tyler and Evans' predictions.

Moreover, the analysis has shed light on some puzzling aspects of the distribution of Instrumental and Accusative case in Russian. It has been common to associate Accusative case with motion and Instrumental case with lack of motion. However, Janda has recently offered a more refined analysis in which Instrumental case is represented as prototypically linked with scene setting and Accusative with destination. Drawing on Janda's insights, we argued that the Following sense, in which *za* always co-occurs with Instrumental case even though the participants are in motion, does not entail the TR reaching the LM and therefore does not involve the notion of destination in its interpretation.

The distribution of case with the Purpose sense also challenges the simple association of Accusative with motion and Instrumental with lack of motion. Analysis of our data revealed that with directional verbs, such as 'go', which clearly involve a Path schema containing a beginning and a particular spatial destination, Accusative is used to mark the spatial destination; the TR in the Purpose phrase is consistently marked with Instrumental case and seems to be quite separate from the spatial destination. In contrast, with non-directional verbs, such as 'compete' and 'fight', which do not evoke a prototypical Path schema with a spatial destination, the TR in the Purpose clause is marked with Accusative. We hypothesize that with these non-directional verbs there is a conceptual coalescence of purpose and the end of the action. This analysis provides support for Tyler and Evans' (2003) distinction between Path and the trajectory followed by the TR in a specific motion.

Finally, the analysis of *za*'s polysemy network forced us to reconsider the strong claim made by Tyler and Evans that if a non-proto-scenic interpretation is derivable from context, then it should not be considered an independent sense. We argued that certain spatial scenes, although derivable from the proto-scene and context, may occur so frequently that they are entrenched in memory. Once they are entrenched in memory, they are free for re-analysis and can form the basis for further extended meanings. We believe that *za*'s In-Tandem sense represents such a case.

Notes

1 *Za* is both a preposition and a verbal prefix. In this paper we only address its uses as a preposition. We believe that the verbal prefix meanings associated with *za* are related to its prepositional meanings, but that analysis goes beyond the scope of this paper.

2 This sense is part of what we called the 'in tandem' cluster, following Hill (1978). Hill noted that a salient, frequently occurring orientation for humans involves two individuals facing the same direction and lined up one behind the other. He termed this spatio-physical arrangement 'in tandem'.

References

Clark, H. (1973) Space, time, semantics and the child. In T. E. Moore (ed.) *Cognitive Development and the Acquisition of Language* 27–64. New York: Academic Press.

Dal', V. I. (1998) *Tolkoviy Slovar Zhivogo Velikorusskogo Yazika*. Moscow: Tsitadel. (Original published in 1881.) Retrieved on 10 February 2003 from http://www.rusword.org/rus/dal_in.php

Dewell, R. (1994) *Over* again: Image-schema transformations in semantic analysis. *Cognitive Linguistics* 5(4): 351–380.

Dostoevsky's online corpus: *Slovar' Konkordans Publicistiki Dostoevskogo*. Retrieved on 20 December 2002 from http://mmedia3.soros.karelia.ru:8080/~dost_voc/

Grice, H. P. (1975) Logic and conversation. In P. Cole and J. Morgan (eds) *Syntax and Semantics* 3: *Speech Acts* 41–58. New York: Academic Press.

Hill, C.A. (1978) Linguistic representation of spatial and temporal orientation. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society* 524–538. Berkeley, CA: UC Berkeley Press.

Janda, L. A. (1993) *A Geography of Case Semantics: The Czech Dative and the Russian Instrumental.* Berlin/New York: Mouton de Gruyter.

Janda, L. A. (2000) A cognitive model of the Russian accusative case. In R. K. Potapova, V.D. Solov'ev and V.N. Poljakov (eds) *Trudy meždunarodnoj konferencii Kognitivnoe modelirovanie* 4(Part I): 20–43. Moscow: MISIS.

Lakoff, G. (1987) *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Langacker, R. (1987) *Foundations of Cognitive Grammar Vol. I: Theoretical Prerequisites.* Stanford, CA: Stanford University Press.

Langacker, R. (1991) *Foundations of Cognitive Grammar* V*ol. II: Descriptive Application.* Stanford, CA: Stanford University Press.

National Corpus of the Russian Language. Retrieved in February-March 2006 from http://ruscorpora.ru

Ozhegov, S.I., Shvedova, N.Y. (1984) *Slovar' Russkogo Yazika*. Moscow: Russkiy Yazik.

Poltoratsky, M. and Zarechnak, M. (1961) *Russkiy Yazik: Pervaya Kniga* 2. Milwaukee: The Bruce Publishing Company.

Pulkina, I.M. and Zakhava-Nekrasova, E.B. (2000) *Russkiy Yazik: Prakticheskaya Grammatika S Uprazhneniyami. Uchebnik (dlia govoryashih na angliyskom yazike)*. (Eighth edition.) New York, Moscow: Russkiy Yazik.

Sandra, D. and Rice, B. (1995) Network analysis of prepositional meaning: Mirroring whose mind – the linguist's or the language user's? *Cognitive Linguistics* 6(1): 89–130.

Slovar': *Obschaya Leksika Russkogo Yazyka*. Retrieved on 5 March 2003 from http://lingvo.yandex.ru/cgi-bin/lingvo.pl?text=%E7%E0

Talmy, L. (2000) *Toward a Cognitive Semantics Vol. I: Concept Structuring Systems.* The MIT Press.

Talmy, L. (2000) *Toward a Cognitive Semantics Vol. II: Typology and Process in Concept Structuring.* The MIT Press.

Tyler, A. and Evans, V. (2001) Reconsidering prepositional polysemy networks: The case of 'over'. *Language* 77(4): 724–765.

Tyler, A. and Evans, V. (2003) *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition.* Cambridge: Cambridge University Press.

Vandeloise, C. (1991) L'espace en francais. Paris: Editions du Seuil. English translation by A. Bosch: *Spatial Prepositions*. Chicago: University of Chicago Press.

Wilson, P. and Anderson, R. (1986) What they don't know will hurt them: The role of prior knowledge in comprehension. In J. Orasunu (ed.) *Reading Comprehension: From Research to Practice* 31–48. Hillsdale, NJ: Lawrence Erlbaum Associates.

Zinken, Jörg (this volume) Temporal frames of reference. In V. Evans and P. Chilton (eds) *Language, Cognition and Space: The State of the Art and New Directions*. London: Equinox.

12 Frames of reference, effects of motion, and lexical meanings of Japanese front/back terms

Kazuko Shinohara and Yoshihiro Matsunaka

1 Introduction*

Spatial cognition is often said to play the central and fundamental role in our thinking. Spatial concepts and how they are expressed have been discussed for many years in a wide variety of disciplines including philosophy, physics, cognitive science, and anthropology. Of course, these topics have also attracted linguists, who have long noticed that spatial concepts are related to a vast range of linguistic phenomena. For example, words that express spatial relations are among the most basic elements in language, as instantiated by adpositions, conjunctions, and so forth. These functional elements in language are often the result of the process known as grammaticalization. Moreover, spatial concepts also serve as the source domains of widespread metaphors, such as metaphors of time, state, emotion, and life. It is not a simple task, however, to clarify how spatial cognition and lexemes denoting space are related. Naturally, while innumerable previous studies have addressed this issue, there remain a lot of unsolved problems. For example, one such problem involves two contradicting positions, one that claims that frames of reference are an extra-linguistic matter (Levinson 2003), and the other that rejects the possibility that frames of reference may apply at the linguistic level (Svorou 1994; Carlson-Radvansky and Irwin 1993).

In addition to the difficulty in specifying the relationship between spatial cognition and lexical meanings, there is an issue of how much specification concerning space should be attributed to spatial lexemes. Some researchers describe lexical meanings of space in terms of rich and detailed information. For example, Lakoff (1987) presents a highly rich description of the image-schematic networks of 'over'. Other researchers like Tyler and Evans (2003) avoid assigning such specified information to each lexeme but attribute much of spatial relations associated with spatial expressions to contextual information and encyclopedic knowledge.

In this study, we aim to consider these issues by examining the uses and meanings of three Japanese spatial lexemes *mae* (front), *ushiro* (back), and *saki* (front/ahead). Our analysis and empirical data will support Levinson's position and Tyler and Evans's position stated above. In the remainder of this paper, we first review relevant previous studies and present our goal in Section 2. Then we examine unmarked uses of these lexemes in Section 3. In Section 4, our experiment on these lexemes is reported. Finally, Section 5 concludes this study.

2 Previous studies and issues

Two lines of research are critical for the present study. One is Levinson's (1996, 2003) framework of spatial frames of reference, and the other is the theory of lexical meanings by Tyler and Evans (2003) and Evans (2004). We will first introduce these theories and ideas, and then present the issues we are addressing.

2.1 Frames of reference

Many researchers (Clark 1973; Talmy 1983, 2000; Vandeloise 1991; Svorou 1994; Levelt 1996; Levinson 1996, 2003, among others) describe the notion of frames of reference as playing one of the most fundamental roles in the study of spatial cognition and its linguistic expression. Though this notion has been defined and classified in various ways in different disciplines, the shared view seems to be that cognition of the spatial relationships of objects involves at least the following three elements.

(i) a referent, trajector, or a figure (the object to be located)
(ii) a relatum, landmark, or a ground (the object relative to which the referent is located)
(iii) a perspective system or a frame of reference (the system that determines the relation of a referent to a relatum)

Levelt (1996: 78) uses the terms 'referent', 'relatum' and 'perspective system', which correspond to 'figure', 'ground' and 'frame of reference' respectively in Talmy's (1978, 1983) and Levinson's (1996) terminology. Langacker (1987) calls the first two elements 'trajector' (TR) and 'landmark' (LM). In the main part of this paper, the terms 'figure', 'ground', and 'frame of reference' will be consistently used to avoid confusion. In the expression 'X is in front of Y', for example, X is the figure, Y is the ground, and the frame that determines the spatial relation of X to Y is the frame of reference.

Scholars differ in their ideas on what kinds of frames of reference are necessary and sufficient. Some posit two types of frames of reference, e.g., egocentric frame of reference versus allocentric frame of reference in developmental psychology, or deictic frame of reference versus intrinsic frame of reference in linguistics. Others posit three types of frames of reference, e.g., viewer-centered frame of reference versus object-centered frame of reference versus environment-centered frame of reference in psycholinguistics. (These classifications are reviewed by Levinson 2003: 26.) Among these different subdivisions of frames of reference, this paper follows Levinson's three-way classification of linguistic frames of reference: the intrinsic, the relative, and the absolute frames of reference. Figure 1 illustrates the three frames of reference.

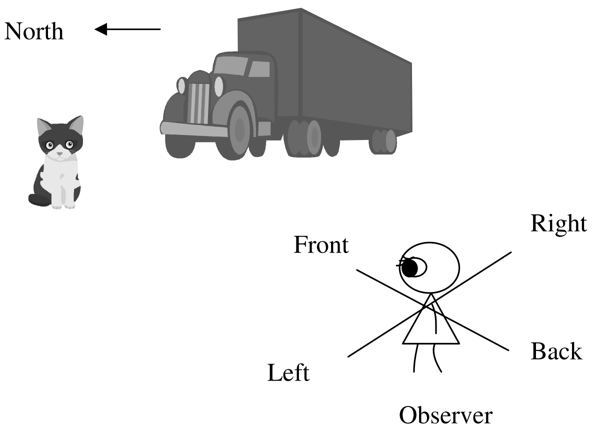


Figure 1. Frames of reference

In the intrinsic frame of reference, the cat (figure) is said to be 'in front of the truck (ground)'. This relation is based on the coordinate system (or particularly the front/back axis) determined by the intrinsic properties of the truck including its functional aspects. For example, the side of the truck that faces the default direction of motion may be regarded as the 'front' of the truck. The other three directions, 'back', 'right' and 'left' are derived from 'front'. In this frame, the figure object (cat) is located in the direction of 'front' of the ground object (truck) which is determined in this way. Note that the intrinsic front/back asymmetry of the figure object (cat) does not determine the orientation of the ground object in this case. Though the cat is looking in a different direction from that of the truck, the cat is still referred to as being 'in front of' the truck when the intrinsic frame of reference is employed.

In the relative frame of reference, the cat is described as being 'to the left of the truck'. This relation comes from the observer's viewpoint. That is, the observer's coordinate system (front, back, right, and left) provides the ground object with these directions. The truck thus obtains its front, back, right, and left based on the observer's viewpoint, which makes it possible to say 'The cat is to the left of the truck'. The observer's left is regarded as the left side of the truck in this case, though the same side of the truck may be the 'front' side in the intrinsic frame of reference.

In the absolute frame of reference, the coordinate system of the truck is determined by the configuration of the outside world, which is non-relative or non-intrinsic. The expression 'The cat is north of the truck' is based on the earth's magnetic field and cardinal orientations derived from it, which are determined independent of the observer's viewpoint or the intrinsic orientation of the truck.

In the intrinsic frame, the ground object itself provides the coordinate system. In the relative and absolute frames, the coordinate system of some other object is projected onto the ground object. In case of the relative frame, the observer is the source of the coordinate system of the ground object, while in the absolute frame, the earth is the source of the coordinate system of the ground object. Talmy (2000) classifies the latter two cases together and states that both of them have Secondary Reference Objects while the intrinsic frame has a Primary Reference Object. In the present study, we do not refer to the distinction between Primary and Secondary Reference Objects, since we only deal with the relative frame of reference.

Among these three frames of reference, the relative frame has three subtypes of projecting the coordinate system on the viewpoint of the observer onto the ground object. They are reflection analysis, translation analysis, and rotation analysis (Levinson 2003: 86–88). Figure 2 illustrates them.

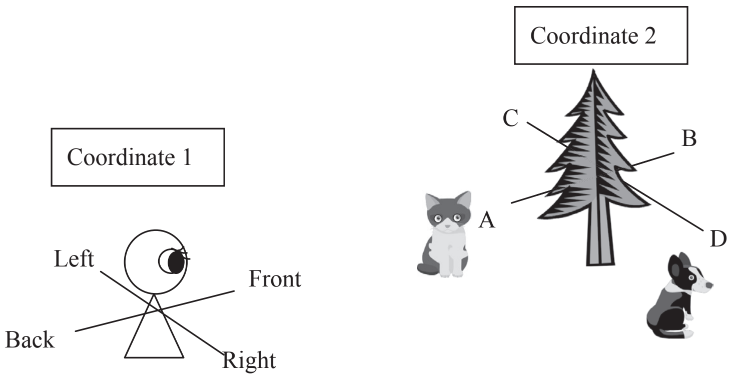


Figure 2. Subtypes of the relative frame of reference

In reflection analysis, the polar A of Coordinate 2 in Figure 2 is said to be front of the tree; B is back of the tree; C is left of the tree; D is right of the tree. The coordinate system on the viewpoint is projected onto the ground (the tree) in such a way that the front/bask axis is reversed but the right/left axis is not reversed. 'The cat is in front of the tree' means that the cat is between the observer and the tree. In this analysis, the dog is said to be 'to the right of the tree'. In other researchers' works, reflection analysis uses different terms. Clark (1973) uses the term 'canonical encounter', Hill (1978) adopts 'mirror-image' strategy, and Moore (2000) calls this 'ego-opposed' strategy. English, Japanese, and perhaps many other languages have this type of projection system.

The second subtype of the relative frame is the translation analysis. In this case, A is said to be back of the tree; B is front of the tree; C is left of the tree; D is right of the tree. In this subtype, the coordinate system on the viewpoint is translated without any reversal or rotation. If this frame is employed, the 'front' side of the tree is the farther side of the tree from the observer, and the space between the observer and the tree is regarded as 'behind' the tree. Thus, speakers who take this frame will describe the scene in Figure 2 as 'The cat is behind the tree'. As for the right/left axis, this frame results in the same expressions as reflection analysis. Hill (1978) calls this frame 'in-tandem' strategy, and Moore (2000) adopts the term 'ego-aligned' strategy because the ground object that has no intrinsic front/back axis is construed as if it stood looking in the same direction as the observer. Hausa, an African language of the Chad family, is said to have this type of frame of reference as the dominant, unmarked frame for FRONT/BACK terms (Hill 1975, 1978, 1982, Levinson 2003).

The third subtype, the rotation analysis, is a rare case in world languages. In this frame, A is front of the tree; B is back of the tree; C is right of the tree; D is left of the tree. The coordinate system on the viewpoint is projected onto the ground object after being rotated 180 degrees. Thus, the right/left axis and the front/back axis are both reversed. In Figure 2, the cat between the observer and the tree is said to be 'in front of the tree' and the dog is said to be 'to the left of the tree, while in English the dog may be to the right of the tree. It is said that Tamil, a Dravidian language, uses this system (Levinson 2003: 88).

2.2 Frames of reference in Japanese

In Japanese, it is observed that all three frames of reference, the absolute, the intrinsic, and the relative frames of reference, are available. However, the absolute frame of reference is relatively limited in use, except in some cases in rural dialects (Inoue 1998, 2002, 2005, Kataoka 2003). The absolute frame of reference is mostly used to refer to distal geographical places. For example, when describing maps or directing ways to places in the distance, the absolute frame of reference tends to be used. To describe things in more proximal space, the intrinsic and the relative frames are likely to be used by most Japanese, especially city-dwellers. Things that we see around us and manipulate or use in daily life are mostly described in the intrinsic or the relative frame of reference. It is not that we cannot describe proximal spatial relations in terms of the absolute frame of reference, but this is limited to certain dialects, and it requires a considerable amount of effort for speakers of standard Japanese to do this. This was confirmed by the authors' in-class research. It took students more time and induced more mistakes to use the absolute frame to describe the relative position of pens, books, and other objects on a desk, than to use the relative frame to do the same thing. (Inoue (2005) reports, however, that there are dialectal variations concerning the use of these reference frames: in some rural regions in Japan the absolute reference frame tends to be used for proximal spatial relations or even body parts such as teeth.)

As for the three subtypes within the relative frame of reference, most of the relevant previous studies report that the reflection analysis is dominant in adult Japanese (Inoue 1998; Imai and Ishizaki 1999; Imai, Nakanishi, Miyashita, Kidachi and Ishizaki 1999; Odate, Shinohara and Matsunaka 2003; Shinohara, Matsunaka and Odate 2003; Shinohara, Odate and Matsunaka 2003; Yoshida 2003; Shinohara and Matsunaka 2004; Shinohara, Kojima and Matsunaka 2004a, b; Matsunaka and Shinohara 2005, etc). Some have also shown that children tend to use translation analysis for *mae* (front) and *ushiro* (back) more than adults do (Odate et al. 2003; Yoshida 2003). All of these points hold in English 'front' and 'back' as well (Clark 1973, Harris and Strommen 1972, Hill 1978, 1982, Levinson 1996, 2003, etc.).

2.3 Lexical specification of frames of reference

Several issues have been raised concerning frames of reference in linguistics. One of them is at which level they are coded. Levinson states this question as follows.

> In psycholinguistic discussions about frames of reference, there seems to be some unclarity, or sometimes overt disagreement, about at which level – perceptual, conceptual or linguistic – such frames of reference apply. … [W]e need to distinguish in discussions of frames of reference between at least three levels, perceptual, conceptual and linguistic, and we need to consider the possibility that we may utilize distinct frames of reference at each level. (Levinson, 2003: 33–34)

Some researchers deny the possibility that frames of reference may apply at the linguistic level. Svorou (1994: 23) states that 'typically RFs are not coded linguistically in spatial expression' (RF stands for 'reference frame'). Carlson-Radvansky and Irwin, in their studies on the spatial term 'above', find that frames of reference are not linguistically coded (1993: 242). However, Levinson argues against these views.

> [I]n most languages there are many subtle details of the use of expressions that generally mark which frame of reference they are being used with – thus *at the truck's front* or *in the front of the truck* can only have an intrinsic reading, not a relative one – so this cannot be treated as an extralinguistic matter. (Levinson, 2003: 108)

In addition to this Levinson's analysis, previous studies on Japanese spatial lexemes like *mae* (front), *saki* (front/ahead), *temae* (front), and *ushiro* (back) provide more evidence of linguistically coded frames of reference. Imai et al. (1999), through an experimental study of *mae* and *ushiro*, show that 97% of decisions concerning the front/back axis of objects without an intrinsic axis are based on reflection analysis. Matsunaka and Shinohara (2004, 2005) state that *mae, saki* and *temae* exhibit some restriction in the choice of frames of reference, which renders it difficult to shift freely to other frames of reference or subtypes. These restrictions cannot be explained if we assume that frames of reference reside only in pre-linguistic cognition or perception, independent of linguistic coding. We must assume, instead, that lexical items can at least in some cases determine which frames of reference they can relate themselves to. Moreover, Shinohara et al. (2004a) examine two Japanese spatial terms (*mae* and *saki*) denoting the frontal concept, and show that the unmarked usage of *mae* is based on the reflection analysis while that of *saki* is based on the translation analysis. This, they conclude, indicates that at least some information about frames of reference is included in each of these words. Thus, contra Svorou's and Carlson-Radvansky and Irwin's view that spatial frames of reference cannot be settled at the linguistic level, evidence presented by Levinson and previous studies on Japanese spatial lexemes show that spatial frames of reference are linguistically-coded and included in lexical meanings at least to some extent.

2.4 Meanings of spatial lexemes

As we have seen, Levinson's view that spatial frames of reference are at least to some extent coded at the linguistic level seems adequate for Japanese spatial lexemes such as *mae*, *saki*, and *temae*. However, it does not necessarily lead to the idea that the senses of spatial lexemes should be as rich as they can be. Contrary to this, scholars like Tyler and Evans (2003: 17–18) argue that semantic properties of lexemes should be described in a simple manner, avoiding excessively specified information. They suggest that meanings that can be obtained by elaborating lexical meanings using contexts should be excluded from semantic description of the lexeme. For example, Tyler and Evans's (2003) description of English spatial lexeme 'over' is simpler than Lakoff's (1987) well-known analysis of 'over'. They state their idea as follows.

> In essence, by attempting to build too much redundancy into the lexical representation, Lakoff's model vastly inflates the number of proposed distinct meanings associated with a spatial particle such as 'over'. An implicit consequence of this representation is that discourse and sentential context, which is utilized in the conceptual processes of inferencing and meaning construction, is reduced in importance, as much of the information arising from inferencing and meaning construction is actually built into the lexical representation. (Ibid.: 42)

Thus, they argue that the general inference system elaborates the meanings of lexemes, for example, 'over', in terms of the context in which the expressions are used. In their analysis, the detailed shape of the landmark, verticality of the landmark, multiplex nature of the trajector(s), coverage, contact between trajector and landmark, etc. are not included in the semantic network of the lexeme 'over' but these kinds of information are claimed to be obtained through elaboration.

The present study deals with slightly different aspects of meaning shifts of spatial lexemes, but we take a similar standpoint as Tyler and Evans' in that we intend to claim that the meaning shifts we are looking at are induced by contextual information and thus they should not be included in the senses of each spatial lexeme.

2.5 Issues and the goal of this study

In this study, we will consider the issue of spatial frames of reference specified by spatial lexemes, and the issue of semantic shifts induced by context. We will demonstrate that three Japanese front/back terms, *mae, ushiro*, and *saki,* exhibit interesting tendencies that seem to support Levinson's position that frames of reference can be settled at linguistic level, and the position that lexically determined meanings of these terms, especially for relative frames of reference (Levinson 2003), may be rather simple, but perceptual spatial context can affect the uses of these terms to produce varying construals of spatial relations of objects that have no front/back axis. We support this argument by demonstrating the following two points: (1) specifications of frames of reference are included in the lexical meanings of *mae*, *ushiro* and *saki*; (2) the effects of motion on the uses of these terms indicate that their meanings are not so rich as to include concrete, specific spatial regions or positions. We present evidence for the former argument from our previous studies, and for the latter claim, we employ an experimental method to demonstrate that different conditions of motion can add to the basic, unmarked orientation of the ground object that these terms prompt. Visual-perceptual contexts such as the motion of the observer or of the objects can add extra information about orientation of the ground object whereby interaction between lexical senses and contextual information can take place. Such context can induce shifts in what these lexemes mean in each case. However, we regard these meaning shifts as not included in the lexical senses. Thus, we argue that these spatial lexemes are rich enough to prompt the unmarked reference frame to refer to, but that they are simple enough in meaning so as to not designate concrete spatial regions or positions.

3 Basic meanings and frames of reference of *mae*, *ushiro*, and *saki*

In this section we describe basic, unmarked usage of the Japanese spatial terms *mae* (front), *ushiro* (back), and *saki* (front/ahead). (All of these three terms are related to the FRONT/BACK concepts, rather than other spatial axes or directions like RIGHT/LEFT, NORTH/SOUTH, UP/DOWN, etc.) Then we will show that they have different specifications of frames of reference. This is especially clear when an object that has neither intrinsic directions nor an asymmetrical shape, such as a block, a cylinder, or a ball, is the ground object.

Mae and *ushiro* originally derived from bodily meanings. *Mae* is related to the word *me* (eye), and *ushiro* is related to the word *shiri* (hip or buttock). They are unmarked words for FRONT and BACK in Japanese. *Saki* basically means a tip or a sharp point of a stick-like object (e.g., the *saki* of a pencil is its pointed end).

The critical question for these unmarked uses of the three spatial terms is which of the frames of reference is employed for each term. In Section 2, we described three different frames of reference, the intrinsic, the relative, and the absolute frames. In fact, *mae*, *ushiro*, and *saki* can all be used for the intrinsic frame and the relative frame, but not for the absolute frame. When the ground object has an intrinsic front/back axis, like a truck, a car, a house, etc., *mae* can mean their frontal part, side, or region, and *ushiro* can mean their back part, side, or region, based on their intrinsic front/back axis. If the ground object has a gradually narrowing shape and a sharp tip, then *saki* can mean the tip itself, or the direction of that tip, or the region in that direction. These uses are based on the intrinsic frame of reference.

However, if the ground object has no such intrinsic axes or directions, then these terms are interpreted based on the relative frame. That is, the viewer's front/back axis is projected onto the ground object. In this case, one of the three subtypes (reflection, translation or rotation) of the relative frame of reference (see Figure 2) is employed.

Figure 3 shows the unmarked uses of these words in the relative frame. Both the figure object (ball) and the ground object (block) lack an intrinsic front/back axis. That is, the orientation of the ground object cannot be determined by its shape or function. Moreover, the situation described in Figure 3 is assumed to be static, that is, neither of the objects are moving. Even in such cases, the lexemes *mae*, *ushiro*, and *saki* can be used to designate front/back relation. These uses are understood and shared by the speakers of standard Japanese. If you say, in Japanese, a sentence that means 'There is a ball in *mae* of the block', the ball is normally in area A in Figure 3. If you say 'There is a ball in *ushiro* of the block, the ball is in area B. If you say 'There is a ball in *saki* of the block', the ball is in area B.

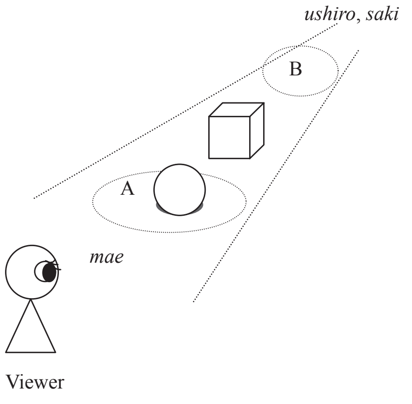


Figure 3. Unmarked uses of *mae*, *ushiro* and *saki*

The unmarked uses of *mae* and *ushiro* are based on reflection analysis: the nearer side to the viewer from the ground object is referred to as *mae* (front), and the farther side is referred to as *ushiro* (back). The unmarked use of *saki* is based on translation analysis: the farther side of the ground object is referred to as *saki* (front). Figure 4 and 5 illustrates these frames for *mae*, *ushiro*, and *saki*.

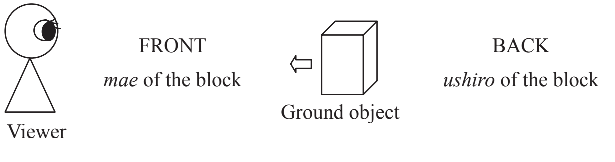


Figure 4. Reflection Frame

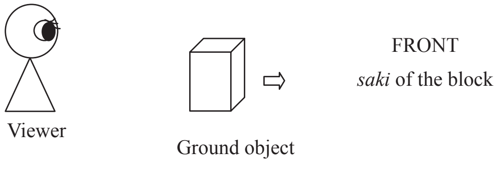


Figure 5. Translation Frame

As shown above, the basic, unmarked uses of *mae*, *ushiro*, and *saki* are crucially related to certain frames of reference. Previous studies show that 80% to 97% of the uses of these words are based on these unmarked patterns (Imai et al. 1999; Imai and Ishizaki 1999; Odate et al. 2003; Shinohara, Matsunaka and Odate 2003; Shinohara, Kojima and Matsunaka 2004a). This indicates that these spatial lexemes have, as part of their lexical properties, at least some specification of the unmarked frame of reference to be referred to (Matsunaka and Shinohara 2004, 2005). Without such specification, it would be impossible for most native speakers to share the same judgment about the spatial usage of these terms. It is even more obvious when we consider the fact that both *mae* and *saki* mean FRONT but the regions they actually refer to are the opposite. Their difference seems to reside in the different frames of reference they are associated with. We cannot say that in Japanese the frontal direction is generally based on the reflection frame, since *saki* is a strong counterexample to this claim. Each of these Japanese spatial lexemes seems to have, as its meaning, specification of the frame of reference to be referred to in unmarked cases. This is, as we have stated in the foregoing, inconsonant with Svorou's (1994) and Carlson-Radvansky and Irwin's (1993) argument.

What we have argued in this section concerns the Japanese FRONT/BACK terms *mae*, *ushiro*, and *saki*, and therefore, it cannot be directly applied to similar spatial terms in other languages. However, some implication for semantics of spatial terms in language may be derived from our analysis. Tyler and Evans (2003) describe meanings of the English phrase 'in front of' as follows.

> As we have just seen in our discussion of 'in front of', the Priority Sense and the proto-scene involve essentially the same relationship between the TR and LM. In both senses, the LM is oriented towards the TR. (Ibid.:164)

In their analysis, 'orientation' is treated as an essential part of the meaning of 'in front of'. This argument seems quite convincing. We would like, however, to point out that their analysis of the meanings of 'in front of' does not include the case where both the figure and the ground objects are symmetrical in terms of the front/back axis (that is, lack an intrinsic front/back axis derived from physical shape, default direction of motion, or functional properties like accessibility). The only example they give for symmetrical objects is the case where two bottles without intrinsic front/back axis are moving in line on a conveyor-belt. The concept of FRONT, however, is obtainable even in cases without such motion. As we have seen in this section, static objects that have no intrinsic front/back axis, such as a block and a ball, can be construed in terms of FRONT relation. By including this instance, Tyler and Evans's analysis would be more exhaustive. It should be emphasized that this does not mean that their analysis is wrong. Rather, the FRONT/BACK conceptualization of objects that have no intrinsic axis, no motion, nor privileged accessibility may occupy a peripheral position in the radial category structure of frontal terms. We will discuss this later.

In this section, we have examined basic, unmarked uses of *mae*, *ushiro*, and *saki* in the relative frame, and have shown that the first two lexemes select the reflection frame as the unmarked frame to refer to, but *saki* selects the translation frame as its unmarked frame of reference when the objects have no intrinsic front/back axis. It has been made clear that each lexeme has its own specification of frame of reference for unmarked uses.

4 The effect of motion on the meanings of *mae*, *ushiro*, and *saki*

In the previous section, we have shown that Japanese spatial lexemes *mae*, *ushiro*, and *saki* designate, at least to some extent, specification of the frames of reference they are associated with. In this section, we will further show that these spatial lexemes can sometimes shift their frames of reference when motion is involved in perceptual context. We do this by way of experimentation. By using three-dimensional computer graphic images to create a sense of virtual reality, we made the viewer feel as if she were moving toward the objects on the screen or as if the objects were drawing nearer to the viewer. In the following, we will describe the method of our experiment, show the data we have obtained, and then discuss what the results mean.

4.1 Method of experiment**

In our experiment, the stimuli consisted of twelve pairs of three-dimensional computer graphic images and sentences. In each image, two objects (a green ball and a red block, both of which have no intrinsic front/back or right/left axes) were located along the viewer's frontal axis, so that the viewer sees these two objects aligned with her orientation. Figure 6 shows the spatial configuration of the objects shown on the screen. The numerals in Figure 6 represent the ratios of the distance between the objects and the viewer. Figure 7 is a rough image of what the participants see just before the objects start moving. Two conditions for motion were set: the viewer-in-motion condition and the objects-in-motion condition. (Under the viewer-in-motion condition, the viewer feels as if she were moving toward the two objects in the screen. Under the objects-in-motion condition, the viewer feels as if the two objects on the screen were moving toward her.)

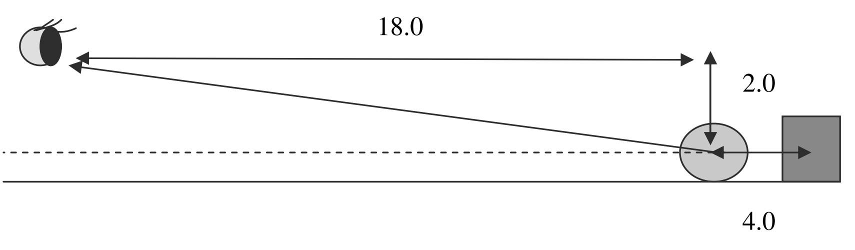


Figure 6. Design of the stimuli

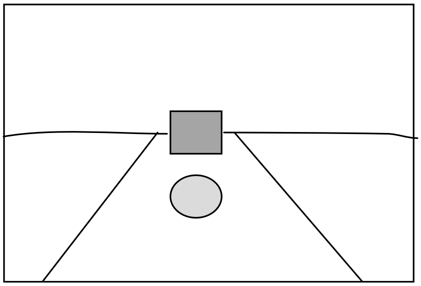


Figure 7. Computer graphic image

Each graphic image was accompanied by a Japanese sentence describing the spatial relation of the objects, which has one of the following two sentence structures.

(a) *midori-no kyuu-wa akai rippoutai-no mae [ushiro / saki]-ni aru*.
green-Gen. ball-Nom. red block-Gen. front [back / front]-Loc. be
'The green ball is in front [back / ahead] of the red block.'

(b) *akai rippoutai-wa midori-no kyuu-no mae [ushiro / saki]-ni aru*.
red block-Nom. green-Gen. ball-Gen. front [back / front]-Loc. be
'The red block is in front [back / ahead] of the green ball.'

In each sentence, one of the three Japanese spatial terms *mae*, *ushiro*, or *saki* was used, so that the sentences were all complete ones. Thus, we prepared twelve stimuli in total. Two motion conditions {viewer-in-motion, objects-in-motion}, two sentence patterns {the green ball as the subject, the red block as the subject}, three spatial lexemes {*mae*, *ushiro*, *saki*} were fully crossed.

Thirty native speakers of Japanese participated in this experiment. They were instructed to (1) read the Japanese sentence shown on the screen, (2) press the key when they understood the sentence (then a three-dimensional computer graphic image appeared on the screen), (3) look at the motion image, and (4) rate how the sentence matched the image in a 4-point scale (where -2=complete mismatch, +2=complete match), by pressing a key on the keyboard. When the subject pressed a key for rating, the next sentence appeared on the screen and the process (1) to (4) was repeated. The twelve stimuli were presented in a random order for each subject.

4.2 Results

The data obtained were categorized into two classes, positive responses (+2 and +1) and negative responses (-2 and -1). The numbers of responses for each category were counted and the total numbers for each condition were statistically analyzed using the Chi-square test.

First, we considered the use of *mae* (front). The arrangement of the two objects, as shown in Figure 7, matches the sentence 'the ball is in *mae* of the block' in the unmarked, normal interpretation, but does not match the sentence 'the block is in *mae* of the ball' (See Section 3). Hence, it is expected that, for the sentence 'the block is in *mae* of the ball', negative responses will dominate. We examined this case, that is, the case where negative responses are expected to dominate. If motion conditions did not affect how this spatial arrangement is perceived and expressed using *mae*, this expectation (the dominance of negative responses) would be satisfied equally under the viewer-in-motion condition and the object-in-motion condition. However, as shown in Figure 8, the viewer-in-motion condition received significantly more positive responses than the object-in-motion condition (*Chi-square*(1)=7.72, $p< .001$). That is, the farther side of the ground object, which is not normally thought of as *mae* (front) of the ground object, was judged as *mae* of the ground object more frequently under the viewer-in-motion condition than under the object-in-motion condition.

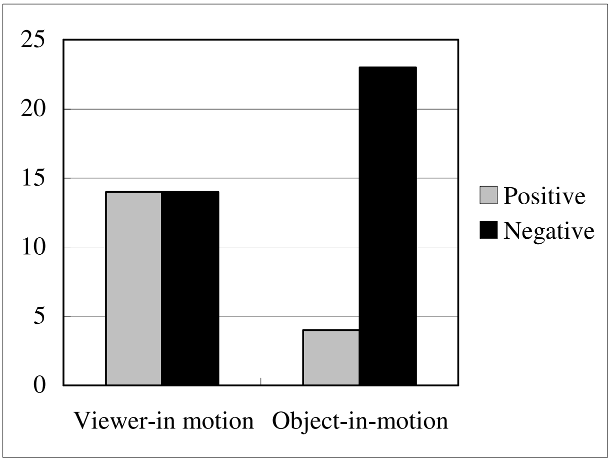


Figure 8. The block is in *mae* of the ball.

Next, we examined the use of *ushiro* (back). In Figure 7, the block is normally said to be in *ushiro* of the ball but the ball is not said to be in *ushiro* of the block in static situation. Consequently, it is expected that the sentence 'the ball is in *ushiro* of the block' will receive dominantly negative responses. If motion conditions did not affect how this spatial arrangement is perceived and expressed using *ushiro*, this expectation (dominance of negative responses) would be satisfied equally for the viewer-in-motion condition and the object-in-motion condition. However, as Figure 9 shows, the viewer-in-motion condition received significantly more positive responses than the object-in-motion condition (*Chi-square*(1)=6.66, $p< .01$). That is, the nearer side, which is not normally regarded as *ushiro* (back) of the ground object, was judged as *ushiro* of the ground object more frequently under the viewer-in-motion condition than under the object-in-motion condition.

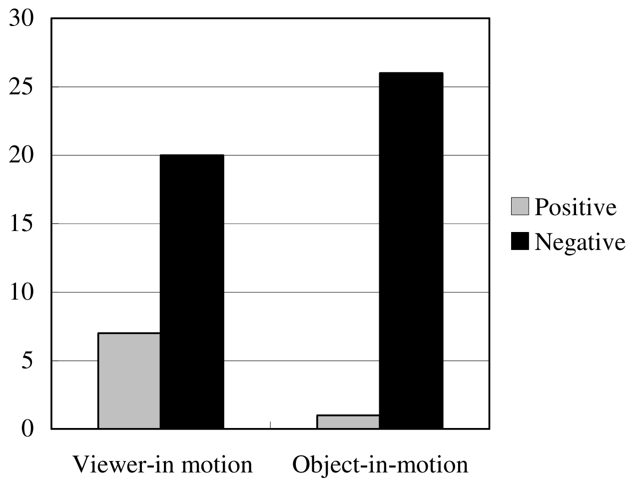


Figure 9. The ball is in *ushiro* of the block

In the third test, *saki* exhibits quite different results. In Figure 7, the block is normally said to be in *saki* (front/ahead) of the ball but the ball is not in *saki* of the block in static situation. Therefore, it is expected that the sentence 'the ball is in *saki* of the block' will receive dominantly negative responses. If motion condition did not affect how this spatial arrangement is perceived and expressed using *saki*, this expectation (dominance of negative responses) would be satisfied equally for the viewer-in-motion condition and the object-in-motion condition. However, the object-in-motion condition received significantly more positive responses than the viewer-in-motion condition as Figure 10 shows ($Chi\text{-}square(1)=6.98$, $p< .01$). That is, the nearer side of the ground object was judged as *saki* (front/ahead) of the ground object more frequently under the object-in-motion condition than under the viewer-in-motion condition.

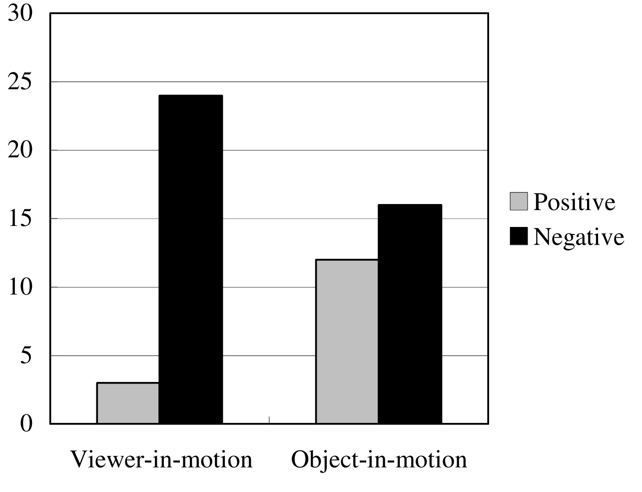


Figure 10. The ball is in *saki* of the block

Thus, we obtained different effects of motion for the three Japanese spatial lexemes. Table 1 summarizes these results.

Table 1. Summary of the results

	Motion condition that received more positive responses for non-standard expressions
mae	Viewer-in-motion
ushiro	Viewer-in-motion
saki	Object-in-motion

4.3 Discussion

As shown in the previous section, the two motion conditions, the viewer-in-motion condition and the object-in-motion condition, affect the subjects' responses. The viewer-in-motion condition, compared with the object-in-motion condition, induced greater positive judgments for non-standard, unusual uses of *mae* and *ushiro*, while the object-in-motion condition, compared with the viewer-in-motion condition, induced greater positive responses for those of *saki* (see Table 1). Why do these motion conditions affect the responses in such different ways?

A likely answer to this question may be that the direction of motion perceived in visual context is projected onto the ground object. For example, if the viewer is moving toward the objects, the direction of this frontward motion may be projected onto the ground object. An illustration of this is given in Figure 11. The viewer is moving in the forward direction, toward the ground object (block). This is indicated by the black arrow. As already explained in Section 3, the unmarked, dominant construal of an object's orientation is based on the reflection frame when the term *mae* (front) or *ushiro* (back) is used. Hence, the *mae* side of the block is normally the nearer side to the viewer (indicated by the bright arrow in Figure 11; this is the unmarked orientation of *mae*) and *ushiro* is the opposite side. Onto this block, the direction of the viewer's frontward motion is projected (indicated by the broken arrow and the dark arrow on the right side of the block). Thus, the block obtains the same direction as the viewer's motion. This projection, it seems, does not override the unmarked orientation of the object. Since our data indicate that half of the subjects responded positively to the sentence that designates the nearer side of the block as *mae* of the block (see Figure 8), the unmarked construal of *mae* of the ground object is not totally cancelled by the projection of the direction of viewer's motion onto the ground object. Rather, it suggests that the ground object obtains the projected direction in addition to the unmarked one.

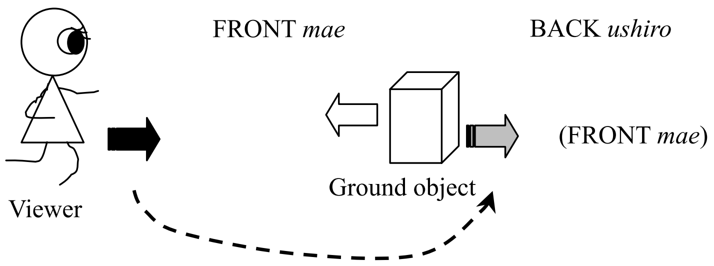


Figure 11. The effect of viewer's motion on the ground object

The above case is based on the viewer-in-motion condition. How about, then, the object-in-motion condition? The effect of the object-in-motion condition is quite different. Figure 12 illustrates the different effect. In this case, there is no viewer's motion, but instead, the ground object moves toward the viewer. This motion defines the moving object's front side (indicated by the black arrow in Figure 12). As stated above, the unmarked, dominant construal of the ground object's orientation is defined by the reflection frame when the term *mae* or *ushiro* is used (indicated by the bright arrow). Hence, in this case, the frontal direction given by the block's motion toward the viewer coincides with the frontal direction given by the reflection frame. Thus, the motion of the block does not add a new direction but just reinforces the unmarked frontal orientation of the block. This explains why the farther side of the block was judged not as *mae* but as *ushiro* of the block by most of the subjects under the object-in-motion condition (Figure 8, 9).

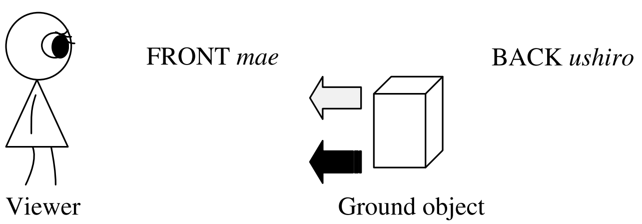


Figure 12. The effect of object's motion on the ground object

Thus, our data for *mae* and *ushiro* can be explained in terms of the effect of motion on the orientation of the ground object. The viewer-in-motion condition affects positively for non-standard expressions of *mae* and *ushiro* because the viewer's motion can give the opposite direction to the ground object.

The same mechanism seems to work for *saki*, but the apparent effects look different because the unmarked orientation defined by *saki* is based on a different frame, not the reflection but the translation frame. As the bright arrow in Figure 13 illustrates, *saki* of the block is the farther side of the block from the viewer's point of view. This is because the unmarked orientation of the object is based on the translation frame in the case of *saki*. The black arrow in Figure 13 indicates the direction of the viewer's motion that is projected onto the block. The block, then, obtains the same direction as the viewer's, as indicated by the dark arrow. As the Figure 13 shows, this projected direction and the unmarked construal of *saki* of the block coincide. Thus, the viewer's motion does not add a new direction but just reinforces the unmarked construal of *saki*. This explains why most subjects responded negatively to the sentence that designates the nearer side of the block as *saki* under the viewer-in-motion condition.

Under the object-in-motion condition, however, the direction of motion of the object and the direction designated by the unmarked use of *saki* oppose each other as illustrated in Figure 14. (The bright arrow indicates the unmarked orientation of the ground object in the case of *saki*.) Since a moving object obtains a front axis defined by the direction of motion, the ground object, the block, obtains the direction toward the viewer as its front. (The black arrow indicates this direction.) Thus, both the farther side and the nearer side of the block can be the *saki* (front) of the block. This explains why about half of the subjects responded positively to the nearer side of the block being called *saki* of the block (Figure 10).

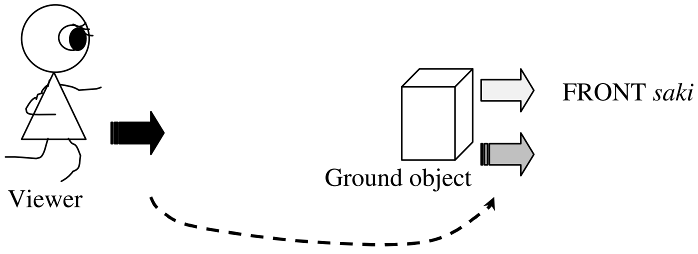


Figure 13. The effect of viewer's motion on the ground object

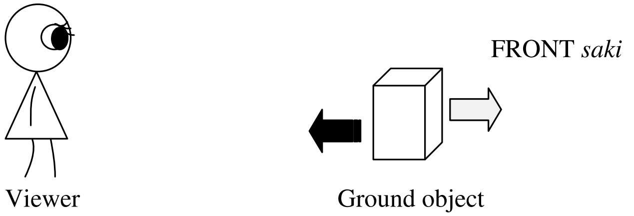


Figure 14. The effect of object's motion on the ground object

In this way, the result of our experiment can be explained if we assume that (1) a moving object obtains frontal direction defined by the direction of that motion, (2) the direction of the viewer's motion can be projected onto the objects being observed, and (3) *mae*, *ushiro*, and *saki* lexically designate certain frames of reference that determine the unmarked frontal orientation of the ground object (as discussed in Section 3).

4.4 Implication for lexical meanings

In Section 3, we have demonstrated that the three Japanese spatial lexemes *mae*, *ushiro*, and *saki*, include specification of frames of reference (either the reflection frame or the translation frame within the relative frame of reference) for the cases where the ground object has no intrinsic front/back axis. The question is, then, whether these lexemes are rich enough to include specific, concrete spatial regions or positions as their senses. The results of our experiment indicate that what these lexemes denote may not be such concrete information about spatial regions or positions but may be only the basic setting of front/back orientation of the ground object.

Support for this argument comes from the effect of motion on the interpretation of the spatial relations of the ball and the block. As discussed in the previous section, the results of our experiment can be explained if we assume that the motion of the viewer or of the objects can affect the orientation of the reference object ((1) and (2) in Section 4.3.), and that each of the three lexemes designates a certain frame of reference ((3) in Section 4.3). We also assume that frames of reference determine the orientation (frontal direction) of the ground object.

If, however, these lexemes denote specific, concrete regions or positions relative to the ground object, our results cannot be explained in such a simple manner. This is because the concepts of REGION and POSITION may be of a quite different kind than MOTION. Though it seems reasonable to assume that the concept of MOTION includes the conceptual element of DIRECTION, and thus it seems natural that motion can affect direction, it is difficult to explain why direction can change the specification of REGION or POSITION. In short, interaction between two directions is far more intelligible than interaction between DIRECTION and REGION or POSITION. Evidence also comes from the present authors' previous studies (Shinohara, Kojima and Matsunaka 2004a, b). In these previous studies we carried out a similar kind of experiment but compared the viewer-in-motion condition and the static condition. We obtained the results that the viewer-in-motion condition works adversely for the standard, unmarked use of *mae* (front). When a ball and a block were placed as in Figure 15, the ball was dominantly (about 91%) judged as being *mae* of the block, but the viewer-in-motion condition made this judgment significantly lower. That is, significantly greater numbers of negative responses were obtained for the sentence 'The ball is in *mae* of the block' under the viewer-in-motion condition.

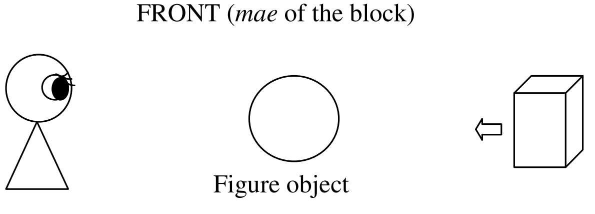


Figure 15. Arrangement of objects in Shinohara et al. (2004a, b)

If *mae* denotes the region that is nearer to the viewer from the ground object, i.e. between the viewer and the ground object, then it is quite difficult to explain why this region is less likely to be called *mae* of the block when the motion of the viewer comes into perceptual context. By contrast, if we assume that the meaning of *mae* defines the orientation (frontal direction in this case) of the block, it becomes understandable that some other motion that orients the object in the opposite direction can have an adverse effect. This explains the increase of negative responses to the ball being called *mae* of the block under the viewer-in-motion condition. The same phenomenon (increase of negative responses to the unmarked, normal uses) is seen for *ushiro* and *saki* as well (Shinohara, Kojima and Matsunaka 2004a, b).

To sum up, we suggest the following three points: (1) what frames of reference do is to determine the (frontal) orientation of the ground object; (2) the spatial terms under examination 'prompt' (Evans 2004: 54) certain spatial frames of reference rather than denote specific regions or positions; (3) the effect of motion included in perceptual context is not a part of the lexical meanings of these terms, but it is a kind of contextual elaboration.

The second and third points especially concern Tyler and Evans's (2003) and Evans's (2004) argument about lexical meanings. As we have shown, it is not reasonable to attribute contextual effects to lexical meanings. We cannot specify all the different motion conditions that can affect the uses of spatial lexemes. Nor can we describe all the specific effects of motion as parts of the lexical properties because, as we have demonstrated, such effects can only be described as a tendency in certain perceptual contexts. The effect of motion is not truth-conditional: it only provides different possibilities of spatial construal with each expression. Each lexeme can be elaborated and has various possibilities of interpretation when contexts permit. Thus, we support the position taken by Tyler and Evans.

What strikes us is that the visual-perceptual context like the one we examined in this study does not constitute the linguistic context that Tyler and Evans (2003) treat as the material of lexical elaboration. Still, as we have seen, such context can affect the interpretation of sentences so strongly that the totally opposite spatial direction can be referred to by the same spatial lexeme. Such an effect might reside not at the linguistic level but in a deeper cognitive level, and it may be that spatial lexemes such as the ones we examined in this study have radial structures that include prototypes as the core senses and gradually diffusing peripheral members, of which the latter may be more susceptible to such cognitive-level influence of perceptual contexts. Tyler and Evans's analysis seems to concern the core senses, and we expect that our findings can add to their theory.

5 Conclusion

In this paper we have examined the meanings of three Japanese FRONT/BACK terms: *mae*, *ushiro*, and *saki*. After reviewing previous studies in Section 2, we have shown in Section 3 that these lexemes have different specifications of frames of reference. Each of the lexemes has, as part of its lexical properties, at least some information about the unmarked frame of reference to refer to. Thus we support Levinson's (2003) position that there exists a certain degree of lexical specification of frames of reference, rather than Svorou's (1994) and Carlson-Radvansky and Irwin's (1993) position that rejects lexically specified frames of reference.

In Section 4, we have demonstrated that these spatial lexemes, when used for ground objects that have no intrinsic directions or prominent axes, do not designate specific, concrete spatial regions or positions. If the perceptual context includes the motion of the viewer or of the objects, the direction of that motion can be added onto the ground object, and thus, interpretation of the orientation of the ground object can vary. We claim that these effects and the consequent interpretations about spatial relations are not included in the lexical meanings of these words, but are a kind of contextual elaboration (in a broad sense). This position is consonant with what Tyler and Evans (2003) and Evans (2004) have suggested, i.e. the claim that lexical meanings should not include contextual elaboration but rather be as narrow as can be.

In conclusion, we suggest that the three Japanese spatial lexemes, *mae, ushiro*, and *saki*, can be made semantically narrow by eliminating the conceptual properties of REGION and POSITION, but that they must have at least a certain specification of unmarked frame of reference. Effects of motion can be observed and these lexemes can have various uses in actual perceptual contexts, but this may be the consequence of contextual elaboration in a broad sense. Thus, we have combined Levinson's (2003) position concerning linguistic specification of frames of reference and Tyler and Evans's (2003) position that lexical meanings and contextual elaboration should be distinguished.

Notes

* This study is supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Grant No. 16500159. We would like to thank the participants at the 9th International Cognitive Linguistics Conference for helpful comments and encouragement, as well as an anonymous reviewer for very helpful comments. All remaining shortcomings are ours.

** The mechanical parts of our experiment, i.e. computer graphics and the automated data-output, were programmed by Takatsugu Kojima. We express deep thanks to him.

References

Carlson-Radvansky, L. A. and Irwin, D. E. (1993) Frames of reference in vision and language: Where is above? *Cognition* 46: 223–244.

Clark, H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive Development and the Acquisition of Language* 27–63. New York: Academic Press.

Evans, V. (2004) *The Structure of Time*. Amsterdam: John Benjamins.

Harris, L. and Strommen, E. (1972) Understanding 'FRONT', 'BACK', and 'Beside': Experiments on the meaning of spatial concepts. In M. H. Siegel and P. Zeigler (eds) *Psychological Research*: *The Inside Story* 198–212. New York: Harper and Row.

Hill, C. (1975) Variation in the use of 'front' and 'back' by bilingual speakers. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society* 196–206.

Hill, C. (1978) Linguistic representation of spatial and temporal orientation. *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society* 524–539.

Hill, C. (1982) Up/down, front/back, left/right: A contrastive study of Hausa and English. In J. Weissenborn and W. Klein (eds) *Here and There*: *Cross-linguistic Studies on Deixis and Demonstration* 11–42. Amsterdam: Benjamins.

Imai, M. and Ishizaki, S. (1999) Mae, ushiro, hidari, migi-no imi (The meanings of front/back, right and left). *Keio SFC Review* 4: 81–88.

Imai, M., Nakanishi, T., Miyashita, H., Kidachi, Y. and Ishizaki, S. (1999) The meanings of FRONT/BACK/LEFT/RIGHT. *Cognitive Studies* 6(2): 207–225.

Inoue, K. (1998) *Moshi migi-ya hidari-ga nakattara* (If there were no rights or lefts). Tokyo: Taishukan Shoten.

Inoue, K. (2002) Zettai to Soutai no Hazama de – Kuukan Shijiwaku niyoru Communication (Between absoluteness and relativity – communication through spatial reference frame). In T. Ohori (ed.) *Ninchi Gengogaku II – Categorization* (Cognitive linguistics II – categorization) 11–35. Tokyo: University of Tokyo Press.

Inoue, K. (2005) Kuukan Ninchi to communication (Spatial cognition and communication). In S. Ide and M. Hiraga (eds) *Ibunka to Communication* (Different culture and communication) 118–128. Tokyo: Hitsuji Shobo.

Kataoka, K. (2003) Kanban kookoku-ni miru kuukan shijiwaku-no hen'i-nitsuite (Variation of spatial frames of reference in commercial signboards). Paper presented at The Eleventh Conference of Japanese Association of Sociolinguistic Sciences.

Lakoff, G. (1987) *Woman, Fire and Dangerous Things*: *What Categories Reveal about the Mind.* Chicago: Chicago University Press.

Langacker, R. (1987) *Foundations of Cognitive Grammar Vol. 1*. Stanford, CA: Stanford University Press.

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

Levinson, S. C. (1996) Frames of reference and Molyneaux's questions: Crosslinguistic evidence. In P. Bloom, M. Peterson, L. Nadel and M. Garrett (eds) *Language and Space* 109–169. Cambridge, MA: MIT Press.

Levinson, S. C. (2003) *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Matsunaka, Y. and Shinohara, K. (2004) Spatial cognition and language of space: A perspective from Japanese. In A. Soares da Silva, A. Torres and M. Gonçalves (eds) *Linguagem, Cultura e Cognição*: *Estudos de Linguística Cognitiva* 59–74. Coimbra: Almedina.

Matsunaka, Y. and Shinohara, K. (2005) Cognition and language of space: Japanese words of FRONT. Paper presented at New Directions in Cognitive Linguistics: First UK Cognitive Linguistics Conference.

Moore, K. E. (2000) Spatial experience and temporal metaphors in Wolof: Point of view, conceptual mapping, and linguistic practice. PhD dissertation, Department of Linguistics, University of California, Berkeley.

Odate, J., Shinohara, K. and Matsunaka, Y. (2003) Nihongo-ni-okeru mae-to ushiro-no ninchi-to hyougen (Cognition and expression of front and back in Japanese). Paper presented at The Eleventh Conference of Japanese Association of Sociolinguistic Sciences.

Shinohara, K., Matsunaka, Y. and Odate, J. (2003) Cognition and expression of FRONT and BACK in Japanese. Paper presented at The Eighth International Cognitive Linguistics Conference.

Shinohara, K., Kojima, T. and Matsunaka, Y. (2004a) Lexical variation in frames of reference: An empirical study of Japanese space terms. Paper presented at the International Conference on Language, Culture and Mind.

Shinohara, K., Kojima, T. and Matsunaka, Y. (2004b) Nihongo-no kuukan goi '*mae, ushiro, saki and temae*' to sono sanshou-waku-ni-kansuru jikken-teki kenkyu (Japanese spatial terms '*mae* (front), *ushiro* (back), *saki* (front/ahead), and *temae* (front-of-hand)' and experimental study of their frames of reference). Paper presented at The Twenty-first Conference of Japanese Cognitive Science Society.

Shinohara, K. and Matsunaka, Y. (2004) Spatial cognition and linguistic expression: Empirical research on frames of reference in Japanese. *Annual Review of Cognitive Linguistics* 3: 261–283.

Shinohara, K., Odate, J. and Matsunaka, Y. (2003) 'Mae' to 'ushiro' no imi-to sanshouwaku: taiji-teki houryaku-to seiretsu-teki houryku-no arawarekata (Semantics and frame of reference of front and back: How Japanese spatial terms adopt ego-opposed strategy and ego-aligned strategy). Paper presented at The Twentieth Conference of Japanese Cognitive Science Society.

Svorou, S. (1994) *The Grammar of Space*. Amsterdam: John Benjamins.

Talmy, L. (1978) Figure and ground in complex sentences. In J. H. Greenberg (ed.) *Universals of Human Language* 4: *Syntax* 625–649. Stanford: Stanford University Press.

Talmy, L. (1983) How language structures space. In H. L. Pick and L. Acredolo (eds) *Spatial Orientation* 225–82. New York: Plenum Press.

Talmy, L. (2000) *Toward a Cognitive Semantics Vol. 1*. Cambridge, MA: MIT Press.

Tyler, A. and Evans, V. (2003) *The Semantics of English Prepositions*: *Spatial Scenes, Embodied Meaning, and Cognition*. Cambridge: Cambridge University Press.
Vandeloise, C. (1991) *Spatial Prepositions*: *A Case Study from French*. Chicago and London: The University of Chicago Press.
Yoshida, A. (2003) Nihonjin-no kuukan-ninchi-to mae-no imi (Spatial cognition of Japanese and the meaning of '*mae*'). Senior thesis, Department of Computer, Information and Communication Sciences, Tokyo University of Agriculture and Technology.

Part VI

Space in sign-language and gesture

13 How spoken language and signed language structure space differently

Leonard Talmy

1 Introduction[1]

This paper combines and relates new findings on spatial structuring in two areas of investigation, spoken language and signed language. Linguistic research to date has determined many of the factors that structure the spatial schemas found across spoken languages (e.g. Gruber 1965, Fillmore 1968, Leech 1969, Clark 1973, Bennett 1975, Herskovits 1982, Jackendoff 1983, Zubin and Svorou 1984, as well as myself, Talmy 1983, 2000a, 2000b). It is now feasible to integrate these factors and to determine the comprehensive system they constitute for spatial structuring in spoken language. This system is characterized by several features. With respect to constituency, there is a relatively closed universally available inventory of fundamental spatial elements that in combination form whole schemas. There is a relatively closed set of categories that these elements appear in. And there is a relatively closed small number of particular elements in each category, hence, of spatial distinctions that each category can ever mark. With respect to synthesis, selected elements of the inventory are combined in specific arrangements to make up the whole schemas represented by closed-class spatial forms. Each such whole schema that a closed-class form represents is thus a 'prepackaged' bundling together of certain elements in a particular arrangement. Each language has in its lexicon a relatively closed set of such pre-packaged schemas (larger than that of spatial closed-class forms, due to polysemy) that a speaker must select among in depicting a spatial scene. Finally, with respect to the whole schemas themselves, these schemas can undergo a certain set of processes that extend or deform them. Such processes are perhaps part of the overall system so that a language's relatively closed set of spatial schemas can fit more spatial scenes.

An examination of signed language[2] shows that its structural representation of space systematically differs from that in spoken language in the direction of what appear to be the structural characteristics of scene parsing in visual perception. Such differences include the following: Signed language can mark finer spatial distinctions with its inventory of more structural elements, more categories, and more elements per category. It represents many more of these distinctions in any particular expression. It also represents these distinctions independently in the expression, not bundled together into pre-packaged schemas. And its spatial representations are largely iconic with visible spatial characteristics.

When formal linguistic investigation of signed language began several decades ago, it was important to establish in the context of that time that signed language was in fact a full genuine language, and the way to do this, it seemed, was to show that it fit the prevailing model of language, the Chomskyan-Fodorian language module. Since then, however, evidence has been steadily accruing that signed language does diverge in various respects from spoken language. The modern response to such observations – far from once again calling into question whether signed language is a genuine language – should be to rethink what the general nature of language is. Our findings suggest that instead of some discrete whole-language module, spoken language and signed language are both based on some more limited core linguistic system that then connects with different further subsystems for the full functioning of the two different language modalities.

2 Fundamental space-structuring elements and categories in spoken language

An initial main finding emerges from analysis of the spatial schemas expressed by closed-class (grammatical) forms across spoken languages. There is a relatively closed and universally available inventory of fundamental conceptual elements that recombine in various patterns to constitute those spatial schemas. These elements fall within a relatively closed set of categories, with a relatively closed small number of elements per category.

2.1 The target of analysis

As background to this finding, spoken languages universally exhibit two different subsystems of meaning-bearing forms. One is the 'open-class' or 'lexical' subsystem, comprised of elements that are great in number and readily augmented – typically, the roots of nouns, verbs, and adjectives. The other is the 'closed-class' or 'grammatical' subsystem, consisting of forms that are relatively few in number and difficult to augment – including such bound forms as inflections and such free forms as prepositions and conjunctions. As argued in Talmy (2000a, ch. 1), these subsystems basically perform two different functions: open-class forms largely contribute conceptual content, while closed-class forms determine conceptual structure. Accordingly, our discussion focuses on the spatial schemas represented by closed-class forms so as to examine the concepts used by language for structuring purposes.

Across spoken languages, only a portion of the closed-class subsystem regularly represents spatial schemas. We can identify the types of closed-class forms in this portion and group them according to their kind of schema. The types of closed-class forms with schemas for paths or sites include the following: (1) forms in construction with a nominal, such as prepositions like English *across* (as in *across the field*) or noun affixes like the Finnish illative suffix *-:n* 'into', as well as prepositional complexes such as English *in front of* or Japanese constructions with a 'locative noun' like *ue* 'top surface' (as in *teeburu no ue ni* 'table GEN top at' = 'on the table'); (2) forms in construction with a verb, such as verb satellites like English *out, back* and *apart* (as in *They ran out / back / apart*); (3) deictic determiners and adverbs such as English *this* and *here*; (4) indefinites, interrogatives, relatives, etc., such as English *everywhere/whither / wherever*); (5) qualifiers such as English *way* and *right* (as in *It's way / right up there*); and (6) adverbials like English *home* (as in *She isn't home*).

Types of closed-class forms with schemas for the spatial structure of objects include the following: (1) forms modifying nominals such as markers for plexity or state of boundedness, like English *-s* for multiplexing (as in *birds*) or *-ery* for debounding (as in *shrubbery*); (2) numeral classifiers like Korean *chang* 'planar object'; and (3) forms in construction with the verb, such as some Atsugewi Cause prefixes, like *cu-* 'as the result of a linear object moving axially into the Figure'.

Finally, sets of closed-class forms that represent a particular component of a spatial event of motion/location include the following: (1) the Atsugewi verb-prefix set that represents different Figures; (2) the Atsugewi verb-suffix set that represents different Grounds (together with Paths); (3) the Atsugewi verb-prefix set that represents different Causes; and (4) the Nez Perce verb-prefix set that represents different Manners (see Talmy 2000b, ch. 1 and 2).

2.2 Determining the elements and categories

A particular methodology is used to determine fundamental spatial elements in language. One starts with any closed-class spatial morpheme in any language, considering the full schema that it expresses and a spatial scene that it can apply to. One then determines any factor one can change in the scene so that the morpheme no longer applies to it. Each such factor must therefore correspond to an essential element in the morpheme's schema. To illustrate, consider the English preposition *across* and the scene it refers to in *The board lay across the road*. Let us here grant the first two elements in the *across* schema (demonstrated elsewhere): (1) a Figure object (here, the board) is spatially related to a Ground object (here, the road); and (2) the Ground is ribbonal – a plane with two roughly parallel line edges that are as long as or longer than the distance between them. The remaining elements can then be readily demonstrated by the methodology. Thus, a third element is that the Figure is linear, generally bounded at both ends. If the board were instead replaced by a planar object, say, some wall siding, one could no longer use the original *across* preposition but would have to switch to the schematic domain of another preposition, that of *over*, as in *The wall siding lay over the road*. A fourth element is that the axes of the Figure and of the Ground are roughly perpendicular. If the board were instead aligned with the road, one could no longer use the original *across* preposition but would again have to switch to another preposition, *along*, as in *The board lay along the road*. Additionally, a fifth element of the *across* schema is that the Figure is parallel to the plane of the Ground. In the referent scene, if the board were tilted away from parallel, one would have to switch to some other locution such as *The board stuck into / out of the road*. A sixth element is that the Figure is adjacent to the plane of the Ground. If the board were lowered or raised away from adjacency, even while retaining the remaining spatial relations, one would need to switch to locutions like *The board lay (buried) in the road. / The board was (suspended) above the road*. A seventh element is that the Figure's length is at least as great as the Ground's width. If the board were replaced by something shorter, for example, a baguette, while leaving the remaining spatial relations intact, one would have to switch from *across* to *on*, as in *The baguette lay on the road*. An eighth element is that the Figure touches both edges of the Ground. If the board in the example retained all its preceding spatial properties but were shifted axially, one would have to switch to some locution like *One end of the board lay over one edge of the road*. Finally, a ninth element is that the axis of the Figure is horizontal (the plane of the Ground is typically, but not necessarily, horizontal). Thus, if one changes the original scene to that of a spear hanging on a wall, one can use *across* if the spear is horizontal, but not if it is vertical, as in *The spear hung across the wall. / The spear hung up and down on the wall*. Thus, from this single example, the methodology shows that at least the following elements figure in closed-class spatial schemas: a Figure and a Ground, a point, a line, a plane, a boundary (a point as boundary to a line, a line as boundary to a plane), parallelness, perpendicularity, horizontality, adjacency (contact), and relative magnitude.

In the procedure of systematically testing candidate factors for their relevance, the elements just listed have proved to be essential to the selected schema and hence, to be in the inventory of fundamental spatial elements. But it is equally necessary to note candidates that do not prove out, so as to know which potential spatial elements do not serve a structuring function in language. In the case of *across*, for example, one can probe whether the Figure, like the board in the referent scene, must be planar – rather than simply linear – and coplanar with the plane of the Ground. It can be seen, though, that this is not an essential element to the *across* schema, since this factor can be altered in the scene by standing the board on edge without any need to alter the preposition, as in *The board lay flat / stood on edge across the road*. Thus, coplanarity is not shown by *across* to be a fundamental spatial element. However, it does prove to be so in other schemas, and so in the end must be included in the inventory. This is seen for one of the schemas represented by English *over*, as in *The tapestry hung over the wall*. Here, both the Figure and Ground must be planes and coplanar with each other. If the tapestry here were changed to something linear, say, a string of beads, it is no longer appropriate to use *over* but only something like *against*, as in *The string of beads hung *over / against the wall*. Now, another candidate element – that the Figure must be rigid, like the board in the scene – can be tested and again found to be inessential to the *across* schema, since a flexible linear object can be substituted for the board without any need to change the preposition, as seen in *The board/The cable lay across the road*. Here, however, checking this candidate factor across numerous spatial schemas in many languages might well never yield a case in which it does figure as an essential element and so would be kept off the inventory.

This methodology affords a kind of existence proof: it can demonstrate that some element does occur in the universally available inventory of structural spatial elements since it can be seen to occur in at least one closed-class spatial schema in at least one language. The procedure is repeated numerous times across many languages to build up a sizable inventory of elements essential to spatial schemas.

The next step is to discern whether the uncovered elements comprise particular structural categories and, if so, to determine what these categories are. It can be observed that for certain sets of elements, the elements in a set are mutually incompatible – only one of them can apply at a time at some point in a schema. Such sets are here taken to be basic spatial categories. Along with their members, such categories are also part of language's fundamental conceptual structuring system for space. A representative sample of these categories is presented next.

It will be seen that these categories generally have a relatively small membership. This finding depends in part on the following methodological principles. An element proposed for the inventory should be as coarse-grained as possible – that is, no more specific than is warranted by cross-schema analysis. Correlatively, in establishing a category, care must be taken that it include only the most generic elements that have actually been determined – that is, that its membership have no finer granularity than is warranted by the element-abstraction procedure. For example, the principle of mutual incompatibility yields a spatial category of 'relative orientation' between two lines or planes, a category with perhaps only two member elements (both already seen in the *across* schema): approximately parallel and approximately perpendicular. Some evidence additionally suggests an intermediary 'oblique' element as a third member of the category. Thus, some English speakers may distinguish a more perpendicular sense from a more oblique sense, respectively, for the two verb satellites *out* and *off*, as in *A secondary pipe branches out / off from the main sewer line*. In any case, though, the category would have no more than these two or three members. Although finer degrees of relative orientation can be distinguished by other cognitive systems, say, in visual perception and in motor control, the conceptual structuring subsystem of language does not include anything finer than the two- or three-way distinction. The procedures of schema analysis and cross-schema comparison, together with the methodological principles of maximum granularity for elements and for category membership, can lead to a determination of the number of structurally distinguished elements ever used in language for a spatial category.

2.3 Sample categories and their member elements

The fundamental categories of spatial structure in the closed-class subsystem of spoken language fall into three classes according to the aspect of a spatial scene they pertain to: the segmentation of the scene into individual components, the properties of an individual component, and the relations of one such component to another. In a fourth class are categories of nongeometric elements frequently found in association with spatial schemas. A sampling of categories and their member elements from each of these four classes is presented next. The examples provided here are primarily drawn from English but can be readily multiplied across a diverse range of languages (see Talmy 2000a, ch. 3).

2.3.1 Categories pertaining to scene segmentation

The class designated as scene segmentation may include only one category, that of 'major components of a scene', and this category may contain only three member elements: the Figure, the Ground, and a Secondary Reference Object. Figure and Ground were already seen for the *across* schema. Schema comparison shows the need to recognize a third scene component, the Secondary Reference Object – in fact, two forms of it: encompassive of or external to the Figure and Ground. The English preposition *near*, as in *The lamp is near the TV* specifies the location of the Figure (the lamp) only with respect to the Ground (the TV). But localizing the Figure with the preposition *above*, as in *The lamp is above the TV*, requires knowledge not only of where the Ground object is, but also of the encompassive earth-based spatial grid, in particular, of its vertical orientation. Thus, *above* requires recognizing three components within a spatial scene, a Figure, a Ground, and a Secondary Reference Object of the encompassive type. Comparably, the schema of *past* in *John is past the border* only relates John as Figure to the border as Ground. One could say this sentence on viewing the event through binoculars from either side of the border. But *John is beyond the border* can be said only by someone on the side of the border opposite John, hence the *beyond* schema establishes a perspective point at that location as a Secondary Reference Object – in this case, of the external type.

2.3.2 Categories pertaining to an individual scene component

A number of categories pertain to the characteristics of an individual spatial scene component. This is usually one of the three major components resulting from scene segmentation – the Figure, Ground, or Secondary Reference Object – but it could be others, such as the path line formed by a moving Figure. One such category is that of 'dimension' with four member elements: zero dimensions for a point, one for a line, two for a plane, and three for a volume. Some English prepositions require a Ground object schematizable for only one of the four dimensional possibilities. Thus, the schema of the preposition *near* as in *near the dot* requires only that the Ground object be schematizable as a point. *Along*, as in *along the trail*, requires that the Ground object be linear. *Over* as in *a tapestry over a wall* requires a planar Ground. And *throughout*, as in *cherries throughout the jello*, requires a volumetric Ground.

A second category is that of 'number' with perhaps four members: one, two, several, and many. Some English prepositions require a Ground comprising objects in one or another of these numbers. Thus, *near* requires a Ground consisting of just one object, *between* of two objects, *among* of several objects, and *amidst* of numerous objects, as in *The basketball lay near the boulder / between the boulders / among the boulders / amidst the cornstalks*. The category of number appears to lack any further members – that is, closed-class spatial schemas in languages around the world seem never to incorporate any other number specifications – such as 'three' or 'even-numbered' or 'too many'.

A third category is that of 'motive state', with two members: motion and stationariness. Several English prepositions mark this distinction for the Figure. Thus, in one of its senses, *at* requires a stationary Figure, as in *I stayed / *went at the library*, while *into* requires a moving Figure, as in *I went / *stayed into the library*. Other prepositions mark this same distinction for the Ground object (in conjunction with a moving Figure). Thus, *up to* requires a stationary Ground (here, the deer), as in *The lion ran up to the deer*, while *after* requires a moving Ground as in *The lion ran after the deer*. Apparently no spatial schemas mark such additional distinctions as motion at a fast vs. slow rate, or being located at rest vs. remaining located fixedly.

A fourth category is that of 'state of boundedness' with two members: bounded and unbounded. The English preposition *along* requires that the path of a moving Figure be unbounded, as shown by its compatibility with a temporal phrase in *for* but not *in*, as in *I walked along the pier for* 10 *minutes / *in* 20 *minutes*. But the spatial locution *the length of* requires a bounded path, as in *I walked the length of the pier in* 20 *Minutes / *for* 10 *minutes*.[3] While some spatial schemas have the bounded element at one end of a line and the unbounded element at the other end, apparently no spatial schema marks any distinctions other than the two cited states of boundedness. For example, there is no cline of gradually increasing boundedness, nor a gradient transition, although just such a 'clinal boundary' appears elsewhere in our cognition, as in geographic perception or conception, e.g., in the gradient demarcation between full forest and full meadowland (Mark and Smith, 2004).

Continuing the sampling of this class, a fifth category is that of 'directedness' with two members: basic and reversed. A schema can require one or the other of these elements for an encompassive Ground object, as seen for the English prepositions in *The axon grew along / against the chemical gradient*, or for the Atsugewi verb satellites for (moving) 'downstream' and 'upstream'. Or it can require one of the member elements for an encompassive Secondary Reference Object (here, the line), as in *Mary is ahead of / behind John in line*.

A sixth category is 'type of geometry' with two members: rectilinear and radial. This category can apply to an encompassive Secondary Reference Object to yield reference frames of the two geometric types. Thus, in a subtle effect, the English verb satellite *away*, as in *The boat drifted further and further away / out from the island*, tends to suggest a rectilinear reference frame in which one might picture the boat moving rightward along a corridor or sea lane with the island on the left (as if along the x-axis of a Cartesian grid). But *out* tends to suggest a radial reference frame in which the boat is seen moving from a center point along a radius through a continuum of concentric circles. In the type-of-geometry category, the radial-geometry member can involve motion about a center, along a radius, or along a periphery. The first of these is the basis for a further category, that of 'orientation of spin axis', with two members: vertical and horizontal. The English verb satellites *around* and *over* specify motion of the Figure about a vertical or horizontal spin axis, respectively, as in *The pole spun around / toppled over* and in *I turned the pail around / over*.

A seventh category is 'phase of matter', with three main members: solid, liquid, and empty space, and perhaps a fourth member, fire. Thus, among the dozen or so Atsugewi verb satellites that subdivide the semantic range of English *into* plus a Ground object, the suffix *-ik's* specifies motion horizontally into solid matter (as chopping an ax into a tree trunk), *-ic't* specifies motion into liquid, *-ipsnu* specifies motion into the empty space of a volumetric enclosure, and *-caw* specifies motion into a fire. The phase of matter category even figures in some English prepositions, albeit covertly. Thus, *in* can apply to a Ground object of any phase of matter, whereas *inside* can apply only to one with empty space, as seen in *The rock is in/inside the box; in / *inside the ground; in / *inside the puddle of water; in / *inside the fire*.

A final category in this sampled series is that of 'state of consolidation' with apparently two members: compact (precisional) and diffuse (approximative). The English locative prepositions *at* and *around* distinguish these two concepts, respectively, for the area surrounding a Ground object, as in *The other hiker will be waiting for you at / around the landmark*. The two deictic adverbs in *The hiker will be waiting for you there/ thereabouts* mark the same distinction (unless *there* is better considered neutral to the distinction). And in Malagasy (Imai, 2003), two locative adverbs for 'here' mark this distinction, with *eto* for 'here within this bounded region', typically indicated with a pointing finger, and *ety* for 'here spread over this unbounded region', typically indicated with a sweep of the hand. In addition to this sampling, some ten or so further categories pertaining to properties of an individual schema component, each category with a small number of fixed contrasts, can be readily identified.

2.3.3 Categories pertaining to the relation of one scene component to another

Another class of categories pertains to the relations that one scene component can bear to another. One such category was described earlier, that of 'relative orientation', with two or three members: parallel, perpendicular, and perhaps oblique. A second such category is that of 'degree of remove', of one scene component from another. This category appears to have four or five members, two with contact between the components – coincidence and adjacency – and two or three without contact – proximal, perhaps medial, and distal remove. Some pairwise contrasts in English reveal one or another of these member elements for a Figure relating to a Ground. Thus, the locution *in the front of*, as in *The carousel is in the front of the fairground*, expresses coincidence, since the carousel as Figure is represented as being located in a *part* of the fairground as Ground. But *in front of* (without a *the*) as in *The carousel is in front of the fairground*, indicates proximality, since the carousel is now located outside the fairground and near it but not touching it. The distinction between proximal and distal can be teased out by noting that *in front of* can only represent a proximal but not a distal degree of remove, as seen in the fact that one can say *The carousel is* 20 *feet in front of the fairground*, but not, **The carousel is* 20 *miles in front of the fairground*, whereas *above* allows both proximal and distal degrees of remove, as seen in *The hawk is* 1 *foot* / 1 *mile above the table*. The distinction between adjacency and proximality is shown by the prepositions *on* and *over*, as in *The fly is on / over the table*. Need for a fifth category member of 'medial degree of remove' might come from languages with a 'here / there / yonder' kind of distinction in their deictic adverbs or demonstratives.

A third category in this series is that of 'degree of dispersion' with two members: sparse and dense. To begin with, English can represent a set of multiple Figures, say, 0-dimensional peas, as adjacent to or coincident with a 1-, 2-, or 3-dimensional Ground, say, with a knife, a tabletop, or aspic, in a way neutral to the presence or absence of dispersion, as in *There are peas on the knife; on the table; in the aspic*. But in representing dispersion as present, English can (or must) indicate its degree. Thus, a sparse degree of dispersion is indicated by the addition of the locution *here and there*, optionally together with certain preposition shifts, as in *There are peas here and thereon / along the knife; on / over the table; in the aspic*. And for a dense degree of dispersion, English has the three specialized forms *all along, all over* and *throughout*, as seen in *There are peas all along the knife; all over the table; throughout the aspic*.

A fourth category is that of 'path contour' with perhaps some four members: straight, arced, circular, and meandering. Some English prepositions require one or another of these contour elements for the path of a Figure moving relative to a Ground. Thus, *across* indicates a straight path, as seen in *I drove across the plateau / *hill*, while *over* – in its usage referring to a single path line – indicates an arced contour, as in *I drove over the hill / *plateau*. In one of its senses, *around* indicates a roughly circular path, as in *I walked around the maypole*, and *about* indicates a meandering contour, as in *I walked about the town*. Some ten or so additional categories for relating one scene component to another, again each with its own small number of member contrasts, can be readily identified.

2.3.4 Nongeometric categories

All the preceding elements and their categories have broadly involved geometric characteristics of spatial scenes or the objects within them – that is, they have been genuinely spatial. But a number of nongeometric elements are recurrently found in association with otherwise geometric schemas. One category of such elements is that of 'force dynamics' (see Talmy 2000a, ch. 7) with two members: present and absent. Thus, geometrically, the English prepositions *on* and *against* both represent a Figure in adjacent contact with a Ground, but in addition, *on* indicates that the Figure is supported against the pull of gravity through that contact, while *against* indicates that it is not, as seen in *The poster is on / *against the wall* and *The floating helium balloon is against / *on the wall*. Cutting the conceptualization of force somewhat differently (Bowerman 1996), the Dutch preposition *op* indicates a Figure supported comfortably in a natural rest state through its contact with a Ground, whereas *aan* indicates that the Figure is being actively maintained against gravity through contact with the Ground, so that flesh is said to be 'op' the bones of a live person but 'aan' the bones of a dead person.

A second nongeometric category is that of 'accompanying cognitive/affective state', though its extent of membership is not clear. One recurrent member, however, is the attitude toward something that it is unknown, mysterious, or risky. Perhaps in combination with elements of inaccessibility or nonvisibility, this category member is associated with the Figure's location in the otherwise spatial indications of the English preposition *beyond*, whereas it is absent from the parallel locution *on the other side of*, as in *He is beyond / on the other side of the border* (both these locutions – unlike *past* seen above –are otherwise equivalent in establishing a viewpoint location as an external Secondary Reference Object).

A third nongeometric category – in the class that relates one scene component to another – is that of 'relative priority', with two members: coequal and main/ancillary. The English verb satellites *together* and *along* both indicate joint participation, as seen in *I jog together / along with him*. But *together* indicates that the Figure and the Ground are coequal partners in the activity, whereas *along* indicates that the Figure entity is ancillary to the Ground entity, who would be assumed to engage in the activity even if alone (see Talmy 2000b, ch. 3).

2.4 Properties of the inventory

By our methodology, the universally available inventory of structural spatial elements includes all elements that appear in at least one closed-class spatial schema in at least one language. These elements may indeed be equivalent in their sheer availability for use in schemas. But beyond that, they appear to differ in their frequency of occurrence across schemas and languages, ranging from very common to very rare. Accordingly, the inventory of elements – and perhaps also that of categories – may have the property of being hierarchical, with entries running from the most to the least frequent. Such a hierarchy suggests asking whether the elements in the inventory, the categories in the inventory, and the elements in each category form fully closed memberships. That is, does the hierarchy end at a sharp lower boundary or trail off indefinitely? With many schemas and languages already examined, our sampling method may have yielded all the commoner elements and categories, but as the process slows down in the discovery of the rarer forms, will it asymptotically approach some complete constituency and distinctional limit in the inventory, or will it be able to go on uncovering sporadic novel forms as they develop in the course of language change?

The latter seems likelier. Exotic elements with perhaps unique occurrence in one or a few schemas in just one language can be noted, including in English. Thus, in referring to location at the interior of a wholly or partly enclosed vehicle, the prepositions *in* and *on* distinguish whether the vehicle lacks or possesses a walkway. Thus, one is in a car but on a bus, in a helicopter but on a plane, in a grain car but on a train, and in a rowboat but on a ship. Further, Fillmore has observed that this *on* also requires that the vehicle be currently in use as transport: *The children were playing in / *on the abandoned bus in the junkyard*. Thus, schema analysis in English reveals the element '(partly) enclosed vehicle with a walkway currently in use as transport'. This is surely one of the rarer elements in schemas around the world, and its existence, along with that of various others that can be found, suggests that indefinitely many more of them can sporadically arise.

In addition to being only relatively closed at its hierarchically lower end, the inventory may include some categories whose membership seems not to settle down to a small fixed set. One such category may be that of 'intrinsic parts'. Frequently encountered are the five member elements 'front', 'side', 'back', 'top', and 'bottom, as found in the English prepositions in *The cat lay before / beside / behind / atop / beneath the TV*. But languages like Mixtec seem to distinguish a rather different set of intrinsic parts in their spatial schemas (Brugman and Macaulay,1986), while Makah distinguishes many more and finer parts, such as with its verb suffixes for 'at the ankle' and 'at the groin' (Matthew Davidson, personal communication).

Apart from any such fuzzy lower boundary or noncoalescing categories, though, there does appear to exist a graduated inventory of basic spatial elements and categories that is universally available and, in particular, is relatively closed. Bowerman (e.g. 1989) has raised the main challenge to this notion. She notes, for example, that at the same time that children acquiring English learn its *in/on* distinction, children acquiring Korean learn its distinction between *kkita* 'put [Figure] in a snug fit with [Ground]' and *nehta* 'put [Figure] in a loose fit with [Ground]' she argues that since the elements 'snug fit' and 'loose fit' are presumably rare among spatial schemas across languages, they do not come from any preset inventory, one that might plausibly be innate, but rather are learned from the open-ended semantics of the adult language. My reply is that the spatial schemas of genuinely closed-class forms in Korean may well still be built from the proposed inventory elements, and that the forms she cites are actually open-class verbs. Open-class semantics – whether for space or other domains – seems to involve a different cognitive subsystem, drawing from finer discriminations within a broader perceptual / conceptual sphere. The Korean verbs are perhaps learned at the same age as English space-related open-class verbs like *squeeze*. Thus, English-acquiring children probably understand that *squeeze* involves centripetal pressure from encircling or bi-/multi-laterally placed Antagonists (typically the arm(s) or hand(s)) against an Agonist that resists the pressure but yields down to some smaller compass where it blocks further pressure, and hence that one can squeeze a teddy bear, a tube of toothpaste, or a rubber ball, but not a piece of string or sheet of paper, juice or sugar or the air, a tabletop or the corner of a building. Thus, Bowerman's challenge may be directed at the wrong target, leaving the proposed roughly preset inventory of basic spatial building blocks intact.

2.5 Basic elements assembled into whole schemas

The procedure so far has been analytic, starting with the whole spatial schemas expressed by closed-class forms and abstracting from them an inventory of fundamental spatial elements. But the investigation must also include a synthetic procedure: examining the ways in which individual spatial elements are assembled to constitute whole schemas. Something of such an assembly was implicit in the initial discussion of the *across* schema. But an explicit example here can better illustrate this part of the investigation.

Consider the schema represented by the English preposition *past* as in *The ball sailed past my head at exactly* 3 *PM*. This schema is built out of the following fundamental spatial elements (from the indicated categories) in the indicated arrangements and relationships: There are two main scene components (members of the 'major scene components' category), a Figure and a Ground (here, the ball and my head, respectively). The Figure is schematizable as a 0-dimensional point (a member element of the 'dimension' category). This Figure point is moving (a member element of the 'motive state' category). Hence it forms a one-dimensional line (a member of the 'dimension' category'). This line constitutes the Figure's 'path'. The Ground is also schematizable as a 0-dimensional point (a member of the 'dimension' category). There is a point P at a proximal remove (a member of the 'degree of remove' category) from the Ground point, forming a 1-dimensional line with it (a member of the 'dimension' category). This line is parallel (a member of the 'relative orientation' category) to the horizontal plane (a member of the 'intrinsic parts' category) of the earth-based grid (a member of the 'major scene components' category). The Figure's path is perpendicular (a member of the 'relative orientation' category) to this line. The Figure's path is also parallel to the horizontal plane of the earth-based grid. If the Ground object has a front, side, and back (members of the 'intrinsic parts' category), then point P is proximal to the side part. A non-boundary point (a member of the 'state of boundedness' category) of the Figure's path becomes coincident (a member of the 'degree of remove' category) with point P at a certain point of time.

Note that here the Figure's path must be specified as passing through a point proximal to the Ground because if it instead passed through the Ground point, one would switch from the preposition *past* to *into*, as in *The ball sailed into my head*, and if it instead past through some distal point, one might rather say something like *The ball sailed along some ways away from my head*. And the Figure's path must be specified both as horizontal and as located at the side portion of the Ground because, for example here, if the ball were either falling vertically or traveling horizontally at my front, one would no longer say that it sailed 'past' my head.

The least understood aspect of the present investigation is what well-formedness conditions, if any, may govern the legality of such combinations. As yet, no obvious principles based, say, on geometric simplicity, symmetry, consistency, or the like are seen to control the patterns in which basic elements assemble into whole schemas. On the one hand, some seemingly byzantine combinations –like the schemas seen above for *across* and *past* – occur with some regularity across languages. On the other hand, much simpler combinations seem never to occur as closed-class schemas. For example, one could imagine assembling elements into the following schema: down into a surround that is radially proximal to a center point. One could even invent a preposition *apit* to represent this schema. This could then be used, say, in *I poured water apit my house* to refer to my pouring water down into a nearby hole dug in the field around my house. But such schemas are not found. Similarly, a number of schematic distinctions in, for example, the domain of rotation are regularly marked by signed languages, as seen below, and could readily be represented with the inventory elements available to spoken languages, yet they largely do not occur. It could be argued that the spoken language schemas are simply the spatial structures most often encountered in everyday activity. But that would not explain why the additional sign-language schemas – presumably also reflective of everyday experience – do not show up in spoken languages. Besides, the different sets of spatial schemas found in different spoken languages are diverse enough from each other that arguing on the basis of the determinative force of everyday experience is problematic. Something else is at work but it is not yet clear what that is.

2.6 Properties and processes applying to whole spatial schemas

It was just seen that selected elements of the inventory are combined in specific arrangements to make up the whole schemas represented by closed-class spatial forms. Each such whole schema is thus a 'pre-packaged' bundling together of certain elements in a particular arrangement. Each language has in its lexicon a relatively closed set of such pre-packaged schemas – a set larger than that of its spatial closed-class forms, because of polysemy. A speaker of the language must select among these schemas in depicting a spatial scene. We now observe that such schemas, though composite, have a certain unitary status in their own right, and that certain quite general properties and processes can apply to them. In particular, certain properties and processes allow a schema represented by a closed class form to generalize to a whole family of schemas. In the case of a generalizing *property*, all the schemas of a family are of equal priority. On the other hand, a generalizing *process* acts on a schema that is somehow basic, and either extends or deforms it to yield nonbasic schemas. (see Talmy 2000a ch. 1 and 3, 2000b ch. 5). Such properties and processes are perhaps part of the overall spoken-language system so that any language's relatively closed set of spatial closed-class forms and the schemas that they basically represent can be used to match more spatial structures in a wider range of scenes.

Looking first at generalizing properties of spatial schemas, one such property is that they exhibit a topological or topology-like neutrality to certain factors of Euclidean geometry. Thus, they are magnitude neutral, as seen in such facts as that the *across* schema can apply to a situation of any size, as in *The ant crawled across my palm / The bus drove across the country*. Further, they are largely shape-neutral, as seen by such facts as that, while the *through* schema requires that the Figure form a path with linear extent, it lets that line take any contour, as in *I zig-zagged / circled through the woods*. And they are bulk-neutral, as seen by such facts as that the *along* schema requires a linear Ground without constraint on the Ground's radial extension, as in *The caterpillar crawled up along the filament /tree trunk*. Thus, while holding to their specific constraints, schemas can vary freely in other respects and so cover a range of spatial configurations.

Among the generalizing processes that extend schemas, one is that of 'extendability from the prototype', which can actually serve as an alternative interpretation for some forms of neutrality, otherwise just treated under generalizing properties. Thus, in the case of shape, as for the *through* schema above, this schema could alternatively be conceived as prototypically involving a straight path line for the Figure, one that can then be bent to any contour. And, in the case of bulk, as for the *along* schema above, this schema could be thought prototypically to involve a purely 1-dimensional line that then can be radially inflated.

Another such process is 'extendability in ungoverned dimensions'. By this process, a scene component of dimensionality N in the basic form of a schema can generally be raised in dimensionality to form a line, plane, or volume aligned in a way not conflicting with the schema's other requirements. To illustrate, it was seen earlier under the 'type of geometry' category that the English verb satellite *out* has a schema involving a point Figure moving along a radius away from a center point through a continuum of concentric circles, as in *The boat sailed further and further out from the island*. This schema with the Figure idealizable as a point is the basic form. But the same satellite can be used when this Figure point is extended to form a 1-dimensional line along a radius, as in *The caravan of boats sailed further and further out from the island*. And the *out* can again be used if the Figure point were instead extended as a 1-dimensional line forming a concentric circle, as in *A circular ripple spread out from where the pebble fell into the water*. In turn, such a concentric circle could be extended to fill in the interior plane, as in *The oil spread out over the water from where it spilled*. Alternatively, the concentric circle could have been extended in the vertical dimension to form a cylinder, as in *A ring of fire spread out as an advancing wall of flames*. Or again, the circle could have been extended to form a spherical shell, as in *The balloon I blew into slowly puffed out*. And such a shell can be extended to fill in the interior volume, as in *The leavened dough slowly puffed out*. Thus, the same form *out* serves for this series of geometric extensions without any need to switch to some different form.

One more schema-extending process is 'extendability across motive states'. A schema basic for one motive state and Figure geometry can in general be systematically extended to another motive state and Figure geometry. For example, a closed-class form whose most basic schema pertains to a point Figure moving to form a path can generally serve as well to represent the related schema with a stationary linear Figure in the same location as the path. Thus, probably the most basic *across* schema is actually for a moving point Figure, as in *The gopher ran across the road*. By the present process, this schema can extend to the static linear Figure schema first seen in *The board lay across the road*. All the spatial properties uncovered for that static schema hold as well for the present basic dynamic schema, which in fact is the schema in which these properties originally arise.

Among the generalizing processes that deform a schema, one is that of 'stretching', which allows a slight relaxing of one of the normal constraints. Thus, in the *across* schema, where the Ground plane is either a ribbon with a long and short axis or a square with equal axes, a static linear Figure or the path of a moving point Figure must be aligned with the short Ground axis or with one of its equal axes. Accordingly, one can say *I swam across the canal* and *I swam across the square pool* when moving from one side to the other, but one cannot say **I swam across the canal* when moving from one end of the canal to the other. But, by moderately stretching one axis length relative to the other, one might just about be able to say *I swam across the pool* when moving from one end to the other of a slightly oblong pool.

Another schema deforming process is that of 'feature cancellation', in which a particular complex of elements in the basic schema is omitted. Thus, the preposition *across* can be used in *The shopping cart rolled across the boulevard and was hit by an oncoming car*, even though one feature of the schema – 'terminal point coincides with the distal edge of the Ground ribbon' – is canceled from the Figure's path. Further, both this feature and the feature 'beginning point coincides with the proximal edge of the Ground ribbon' are canceled in *The tumbleweed rolled across the prairie for an hour*. Thus, the spoken language system includes a number of generalizing properties and processes that allow the otherwise relatively closed set of abstracted or basic schemas represented in the lexicon of any single language to be applicable to a much wider range of spatial configurations.

3 Spatial structuring in signed language

All the preceding findings on the linguistic structuring of space have been based on the patterns found in spoken languages. The inquiry into the fundamental concept structuring system of language leads naturally to investigating its character in another major body of linguistic realization, signed language. The value in extending the inquiry in this way would be to discover whether the spatial structuring system is the same or is different in certain respects across the two language modalities, with either discovery having major consequences for cognitive theory.

In this research extension, a problematic issue is exactly what to compare between spoken and signed language. The two language systems appear to subdivide into somewhat different sets of subsystems. Thus, heuristically, the generalized spoken language system can be thought to consist of an open-class or lexical subsystem (generally representing conceptual content); a closed-class or grammatical subsystem (generally representing conceptual structure); a gradient subsystem of 'vocal dynamics' (including loudness, pitch, timbre, rate, distinctness, unit separation); and an accompanying somatic subsystem (including facial expression, gesture, and 'body language'). On the other hand, by one provisional proposal, the generalized sign language system might instead divide up into the following: a subsystem of lexical forms (including noun, verb, and adjective signs); an 'inflectional' subsystem (including modulations of lexical signs for person, aspect); a subsystem of size-and-shape specifiers (or SASS's; a subsystem of so-called 'classifier expressions'; a gestural subsystem (along a gradient of incorporation into the preceding subsystems); a subsystem of face, head, and torso representations; a gradient subsystem of 'bodily dynamics' (including amplitude, rate, distinctness, unit separation); and an associated or overlaid somatic subsystem (including further facial expression and 'body language'). In particular here, the subsystem of classifier expressions – which is apparently present in all signed languages – is a formally distinct subsystem dedicated solely to the schematic structural representation of objects moving or located with respect to each other in space (see Liddell 2003, Emmorey 2002). Each classifier expression, perhaps generally corresponding to a clause in spoken language, represents a so conceived event of motion or location.[4]

The research program of comparing the representation of spatial structure across the two language modalities ultimately requires considering the two whole systems and all their subsystems. But the initial comparison – the one adopted here – should be between those portions of each system most directly involved with the representation of spatial structure. In spoken language, this is that part of the closed-class subsystem that represents spatial structure and, in signed language, it is the subsystem of classifier constructions. Spelled out, the shared properties that make this initial comparison apt include the following. First, of course, both subsystems represent objects relating to each other in space. Second, in terms of the functional distinction between 'structure' and 'content' described earlier, each of the subsystems is squarely on the structural side. In fact, analogous structure-content contrasts occur. Thus, the English closed-class form *into* represents the concept of a path that begins outside and ends inside an enclosure in terms of schematic structure, in contrast with the open-class verb *enter* that represents the same concept in terms of substantive content (see Talmy 2000a, ch. 1 for this structure-content distinction). Comparably, any of the formations within a classifier expression for such an outside-to-inside path represents it in terms of its schematic structure, in contrast with the unrelated lexical verb sign that can be glossed as 'enter'. Third, in each subsystem, a schematic structural form within an expression in general can be semantically elaborated by a content form that joins or replaces it within the same expression. Thus, in the English sentence *I drove it (– the motorcycle–) in (to the shed)* the parenthesized forms optionally elaborate on the otherwise schematically represented Figure and Ground. Comparably, in the ASL sentence '(SHED) (MOTORCYCLE) vehicle-move-into-enclosure', the optionally signed forms within parentheses elaborate on the otherwise schematic Figure and Ground representations within the hyphenated classifier expression.

To illustrate the classifier system, a spatial event that English could express as *The car drove past the tree* could be expressed in ASL as follows: The signer's dominant hand, used to represent the Figure object, here has a '3 handshape' (index and middle fingers extended forward, thumb up) to represent a land vehicle. The nondominant hand, used to represent the Ground object, here involves an upright '5 handshape' (forearm held upright with the five fingers extended upward and spread apart) to represent a tree. The dominant hand is moved horizontally across the signer's torso and past the nondominant forearm. Further though, this basic form could be modified or augmented to represent additional particulars of the referent spatial event. Thus, the dominant hand can show additional characteristics of the path. For example, the hand could move along a curved path to indicate that the road being followed was curved, it could slant upward to represent an uphill course, or both could be shown together. The dominant hand can additionally show the manner of the motion. For example, as it moves along, it could oscillate up and down to indicate a bumpy ride, or move quickly to indicate a swift pace, or both could be shown together, as well as with the preceding two path properties. And the dominant hand can show additional relationships of the Figure to the Ground. For example, it could pass nearer or farther from the nondominant hand to indicate the car's distance from the tree when passing it, it could make the approach toward the nondominant hand longer (or shorter) than the trailing portion of the path to represent the comparable relationship between the car's path and the tree, or it could show both of these together or, indeed, with all the preceding additional characteristics.

The essential finding of how signed language differs from spoken language is that it more closely parallels what appear to be the structural characteristics of scene parsing in visual perception. This difference can be observed in two venues, the universally available spatial inventory and the spatial expression.

These two venues are discussed next in turn.

3.1 In the inventory

The inventory of forms for representing spatial structure available to the classifier subsystem of signed language has a greater total number of fundamental elements, a greater number of categories, and generally a greater number of elements per category than the spoken language closed-class inventory. While many of the categories and their members seem to correspond across the two inventories, the signed language inventory has an additional number of categories and member elements not present in the spoken language inventory. Comparing the membership of the corresponding categories in terms of discrete elements, the number of basic elements per category in signed language actually exhibits a range: from being the same as that for spoken language to being very much greater. Further, though, while the membership of some categories in signed language may well consist of discrete elements, that of others appears to be gradient. Here, any procedure of tallying some fixed number of discrete elements in a category must give way to determining the approximate fineness of distinctions that can be practicably made for that category. So while some corresponding categories across the two language modalities may otherwise be quite comparable, their memberships can be of different types, discrete vs. analog. Altogether, then, given its greater number of categories, generally larger membership per category, and a frequently gradient type of membership, the inventory of forms for building a schematic spatial representation available to the classifier subsystem of signed language is more extensive and finer than for the closed-class subsystem of spoken language. This greater extensiveness and finer granularity of spatial distinctions seems more comparable to that of spatial parsing in visual perception.

The following are some spatial categories in common across the two language modalities, but with increasing disparity in size of membership. First, some categories appear to be quite comparable across the two modalities. Thus, both the closed-class subsystem of spoken language and the classifier subsystem of signed language structurally segment a scene into the same three components, a Figure, a Ground, and a secondary Reference Object. Both subsystems represent the category of dimensionality with the same four members – a point, a line, a plane, and a volume. And both mark the same two degrees of boundedness: bounded and unbounded.

For certain categories, signed language has just a slightly greater membership than does spoken language. Thus, for motive state, signed language structurally represents not only moving and being located, but also remaining fixedly located – a concept that spoken languages typically represent in verbs but not in their spatial preposition-like forms.

For some other spatial categories, signed language has a moderately greater membership than spoken language. In some of these categories, the membership is probably gradient, but without the capacity to represent many fine distinctions clearly. Thus, signed language can apparently mark moderately more degrees of remove than spoken language's four or five members in this category. It can also apparently distinguish moderately more path lengths than the two – short and long – that spoken language marks structurally (as in English *The bug flew right / way up there*). And while spoken language can mark at most three distinctions of relative orientation – parallel, perpendicular, and oblique – signed language can distinguish a moderately greater number, for example, in the elevation of a path's angle above the horizontal, or in the angle of the Figure's axes to that of the Ground (e.g. in the placement of a rod against a wall).

Finally, there are some categories for which signed language has an indefinitely greater membership than spoken language. Thus, while spoken language structurally distinguishes some four path contours as seen in section 2.3.3, signed language can represent perhaps indefinitely many more, including zigzags, spirals, and ricochets. And for the category 'locus within referent space', spoken language can structurally distinguish perhaps at most three loci relative to the speaker's location – 'here', 'there', and 'yonder' – whereas sign language can distinguish indefinitely many more within sign space.

Apart from membership differences across common categories, signed language represents some categories not found in spoken language. One such category is the relative lengths of a Figure's path before and after encounter with the Ground. Or again, signed language can represent not only the category of 'degree of dispersion' (which spoken language was seen to represent in section 2.3.3), but also the category 'pattern of distribution'. Thus, in representing multiple Figure objects dispersed over a planar surface, it could in addition structurally indicate that these Figure objects are linear (as with dry spaghetti over a table) and are arrayed in parallel alignment, crisscrossing, or in a jumble.

This difference in the number of structurally marked spatial category and element distinctions between spoken and signed language can be highlighted with a closer analysis of a single spatial domain, that of rotational motion. As seen earlier, the closed-class subsystem in spoken language basically represents only one category within this domain, that of 'orientation of spin axis', and within this category distinguishes only two member elements, vertical and horizontal. These two member elements are expressed, for example, by the English verb satellites *around* and *over* as in *The pole spun around / toppled over*. ASL, by contrast, distinguishes more degrees of spin axis orientation and, in addition, marks several further categories within the domain of rotation. Thus, it represents the category of 'amount of rotation' and within this category can readily distinguish, say, whether the arc of a Figure's path is less than, exactly, more than, or many times one full circuit. These are differences that English might offer for inference only from the time signature, as in *I ran around the house for* 20 *seconds / in* 1 *minute / for* 2 *minutes / for hours*, while using the same single spatial form *around* for all these cases. Further, while English would continue using just *around* and *over*, ASL further represents the category of 'relation of the spin axis to an object's geometry' and marks many distinctions within this category. Thus, it can structurally mark the spin axis as being located at the center of the turning object – as well as whether this object is planar like a CD disk, linear like a propeller, or an aligned cylinder like a pencil spinning on its point. It distinguishes this from the spin axis located at the boundary of the object – as well as whether the object is linear like the 'hammer' swung around in a hammer toss, a transverse plane like a swinging gate, or a parallel plane like a swung cape. And it further distinguishes these from the spin axis located at a point external to the object – as well as whether the object is point-like like the earth around the sun, or linear like a spinning hoop. Finally, ASL can structurally represent the category of 'uniformity of rotation' with its two member elements, uniform and nonuniform, where English could mark this distinction only with an open-class form, like the verbs in *The hanging rope spun / twisted around*, while once again continuing with the same single structural closed-class form *around*. Thus, while spoken language structurally marks only a minimal distinction of spin axis orientation throughout all these geometrically distinct forms of rotation, signed language marks more categories as well as finer distinctions within them, and a number of these appear to be distinguished as well by visual parsing of rotational movement.

To expand on the issue of gradience, numerous spatial categories in the classifier subsystem of signed language – for example, many of the 30 spatial categories listed in section 3.2.3.1 are gradient in character. Spoken language has a bit of this, as where the vowel length of a *waaay* in English can be varied continuously. But the preponderant norm is the use of discrete spatial elements, typically incorporated into distinct morphemes. For example, insofar as they represent degree of remove, the separate forms in the series *on / next to / near / away from* represent increasing distance in what can be considered quantal jumps. That is, the closed-class subsystem of spoken language is a type of cognitive system whose basic organizing principle is that of the recombination of discrete elements (i.e., the basic conceptual elements whose combinations, in turn, comprise the meanings of discrete morphemic forms). By contrast, the classifier subsystem of signed language is the kind of cognitive system whose basic organizing principle largely involves gradience, much as would seem to be the case as well for the visual and motor systems. In fact, within a classifier expression, the gradience of motor control and of visual perception are placed in sync with each other (for the signer and the addressee, respectively), and conjointly put in the service of the linguistic system.

While this section provides evidence that the classifier subsystem in signed language diverges from the schematizing of spoken language in the direction of visual parsing, one must further observe that the classifier subsystem is also not 'simply' a gestural system wholly iconic with visual perception. Rather, it incorporates much of the discrete, categorial, symbolic, and metaphoric character that is otherwise familiar from the organization of spoken language. Thus, as already seen above, spatial representation in the classifier subsystem does fall into categories, and some of these categories contain only a few discrete members – in fact, several of these are much the same as in spoken language. Second, the hand-shapes functioning as classifiers for the Figure, manipulator, or instrument within classifier expressions are themselves discrete (nongradient) members of a relatively closed set. Third, many of the hand movements in classifier expressions represent particular concepts or meta-concepts and do not mimic actual visible movements of the represented objects. Here is a small sample of this property. After one lowers one's two extended fingers to represent a knife dipping into peanut butter – or all one's extended fingers in a curve to represent a scoop dipping into coffee beans – one curls back the fingertips while moving back up to represent the instrument's 'holding' the Figure, even though the instrument in question physically does nothing of the sort. Or again, the free fall of a Figure is represented not only by a downward motion of the dominant hand in its classifier handshape, but also by an accompanying rotation of the hand – whether or not the Figure in fact rotated in just that way during its fall. As another example, a Figure is shown as simply located at a spot in space by the dominant hand in its classifier handshape being placed relaxedly at a spot in signing space, and as remaining fixedly at its spot by the hand's being placed tensely and with a slight final jiggle, even though these two conceptualizations of the temporal character of a Figure's location are visually indistinguishable. Or, further, a (so-conceivedly) random spatial distribution of a mass or multiplex Figure along a line, over a plane, or through a volume is represented by the Figure hand being placed with a loose nonconcerted motion, typically three times, at uneven spacings within the relevant n-dimensional area, even though that particular spacing of three exemplars may not correspond to the actual visible distribution. And finally, a classifier hand's type of movement can indicate whether this movement represents the actual path of the Figure, or is to be discounted. Thus, the two flat hands held with palms toward the signer, fingertips joined, can be moved steadily away to represent a wall's being slid progressively outward (as to expand a room), or instead can be moved in a quick up-and-down arc to a point further away to represent a wall relocated to a further spot, whatever its path from the starting location. That is, the latter quick arc movement represents a meta-concept: that the path followed by the hands does not represent the Figure's actual path and is to be disregarded from calculations of iconicity. All in all, then, the classifier subsystem presents itself as a genuine linguistic system, but one having more extensive homology with the visual structuring system than spoken language has.

3.2 In the expression

The second venue, that of any single spatial expression, exhibits further respects in which signed language differs from spoken language in the apparent direction of visual scene parsing. Several of these are outlined next.

3.2.1 Iconic representation in the expression

Spatial representation in signed classifier expressions is iconic with scene parsing in visual perception in at least the following four respects.

3.2.1.1 Iconic clustering of elements and categories

The structural elements of a scene of motion are clustered together in the classifier subsystem's representation of them in signed language more as they seem to be clustered in perception. When one views a motion event, such as a car driving bumpily along a curve past a tree, it is perceptually the same single object, the car, that exhibits all of the following characteristics: it has certain object properties as a Figure, it moves, it has a manner of motion, it describes a path of a particular contour, and it relates to other surrounding objects (the Ground) in its path of motion. The Ground object or objects are perceived as separate. Correspondingly, the classifier subsystem maintains exactly this pattern of clustering. It is the same single hand, the dominant hand, that exhibits the Figure characteristics, motion, manner, path contour, and relations to a Ground object. The other hand, the nondominant, separately represents the Ground object.

All spoken languages diverge to a greater or lesser extent from this visual fidelity. Thus, consider one English counterpart of the event, the sentence *The car bumped along past the tree*. Here, the subject nominal, *the car*, separately represents the Figure object by itself. The verb complex clusters together the representations of the verb and the satellite: The verb *bumped* represents both the fact of motion and the manner of motion together, while its sister constituent, the satellite *along* represents the presence of a path of translational motion. The prepositional phrase clusters together the preposition *past*, representing the path conformation, and its sister constituent, the nominal *the tree*, representing the Ground object. It in fact remains a mystery at this point in the investigation why all spoken languages using a preposition-like constituent to indicate path always conjoin it with the Ground nominal and basically never with the Figure nominal[5], even though the Figure is what executes the path, and is so represented in the classifier construction of signed language.

3.2.1.2 Iconic representation of object vs. action

The classifier subsystem of signed language appears to be iconic with visual parsing not only in its clustering of spatial elements and categories, as just seen, but largely also in its representation of them. For example, it marks one basic category opposition, that between an entity and its activity, by using an object like the hand to represent an object, and motion of the hand to represent motion of the object. More specifically, the hand or other body part represents a structural entity (such as the Figure) – with the body part's configuration representing the identity or other properties of the entity – while movements or positionings of the body part represent properties of the entity's motion, location, or orientation. For example, the hand could be shaped flat to represent a planar object (e.g. a sheet of paper), or rounded to represent a cup-shaped object. And, as seen, any such hand-shape as Figure could be moved along a variety of trajectories that represent particular path contours.

But an alternative to this arrangement could be imagined. The handshape could represent the path of a Figure – e.g., a fist to represent a stationary location, the outstretched fingers held flat together to represent a straight line path, the fingers in a curved plane for a curved path, and the fingers alternately forward and backward for a zigzag path. Meanwhile, the hand movement could represent the Figure's shape – e.g., the hand moving in a circle to represent a round Figure and in a straight line for a linear Figure. However, no such mapping of referents to their representations is found.[6] Rather, the mapping in signed language is visually iconic: it assigns the representation of a material object in a scene to a material object in a classifier complex, for example, the hand, and the representation of the movements of that object in the scene to the movements of the hand.

No such iconic correspondence is found in spoken language. Thus, while material objects are prototypically expressed by nouns in English, they are instead prototypically represented by verb roots in Atsugewi (see Talmy 2000b, ch. 1). And while path configurations are prototypically represented in Spanish by verbs, this is done by prepositions and satellites in English.

3.2.1.3 Iconic representation of further particular categories

Finer forms of iconicity are also found within each branch of the broad entity-activity opposition. In fact, most of the spatial categories listed in section 3.2.3.1 that a classifier expression can represent are largely iconic with visual parsing. Thus, an entity's form is often represented by the form of the hand(s), its size by the compass of the hand(s), and its number by the number of digits or hands extended. And, among many other categories in the list, an entity's motive state, path contour, path length, manner of motion, and rate of motion are separately represented by corresponding behaviors of the hand(s).

Spoken language, again, has only a bit of comparable iconicity. As examples, path length can be iconically represented in English by the vowel length of *way*, as in *The bird flew waay / waaaay / waaaaaay up there*. Path length can also be semi-iconically represented by the number of iterations, as in *The bird flew up/ up up / up up up and away*. Perhaps the number of an entity can be represented in some spoken language by a closed-class reduplication. But the great majority of spoken closed-class representations show no such iconicity.

3.2.1.4 Iconic representation of the temporal progression of a trajectory

The classifier subsystem is also iconic with visual parsing in its representation of temporal progression, specifically, that of a Figure's path trajectory. For example, when an ASL classifier expression represents 'The car drove past the tree', the 'past' path is shown by the Figure hand progressing from the nearer side of the Ground arm to a point beside it and then on to its further side, much like the path progression one would see on viewing an actual car passing a tree. By contrast, nothing in any single closed-class path morpheme in a spoken language corresponds to such a progression. Thus, the *past* in *The car drove past the tree* is structurally a single indivisible linguistic unit, a morpheme, whose form represents no motion ahead in space. Iconicity of this sort can appear in spoken language only where a complex path is treated as a sequence of subparts, each with its own morphemic representation, as in *I reached my hand down around behind the clothes hamper to get the vacuum cleaner*.

3.2.2 A narrow time-space aperture in the expression

Another way that the classifier expression in signed language may be more like visual perception is that it appears to be largely limited to representing a narrow time-space aperture. The tentative principle is that a classifier complex readily represents what would appear within a narrow scope of space and time if one were to zoom in with one's scope of perception around a Figure object, but little outside that narrowed scope. Hence, a classifier expression readily represents the Figure object as to its shape or type, any manipulator or instrument immediately adjacent to the Figure, the Figure's current state of Motion (motion or located-ness), the contour or direction of a moving Figure's path, and any Manner exhibited by the Figure as it moves. However, a classifier expression can little represent related factors occurring outside the current time, such as a prior cause or a follow-up consequence. And it can little represent even concurrent factors if they lie outside the immediate spatial ambit of the Figure, factors like the ongoing causal activity of an intentional Agent or other external instrumentality.

By contrast, spoken languages can largely represent such nonlocal spatiotemporal factors within a single clause. In particular, such representation occurs readily in satellite-framed languages such as English (see Talmy 2000b, ch. 1 and 3). In representing a Motion event, this type of language regularly employs the satellite constituent (e.g. the verb particle in English) to represent the Path, and the main verb to represent a 'co-event'. The co-event is ancillary to the main Motion event and relates to it as its precursor, enabler, cause, manner, concomitant, consequence, or the like.

Satellite-framed languages can certainly use this format to represent within-aperture situations that can also be represented by a classifier complex. Thus, English can say within a single clause – and ASL can sign within a single classifier expression – a motion event in which the Figure is moved by an adjacent manipulator, as in *I pinched some moss up off the rock* and *I pulled the pitcher along the counter*, or in which the Figure is moved by an adjacent instrument, as in *I scooped jelly beans up into the bag*. The same holds for a situation in which a moving Figure exhibits a concurrent Manner, as in *The cork bobbed past the seaweed*.

But English can go on to use this same one-clause format to include the representation of co-events outside the aperture, either temporally or spatially. Thus, temporally, English can include the representation of a prior causal event, as in *I kicked the football over the goalpost* (first I kicked the ball, then it moved over the goalpost). And it can represent a subsequent event, as in *They locked the prisoner into his cell* (first they put him in, then they locked it). But ASL cannot represent such temporally extended event complexes within a single classifier expression. Thus, it can represent the former sentence with a succession of two classifier expressions: first, flicking the middle finger of the dominant hand across the other hand's upturned palm to represent the component event of kicking an object, and next moving the extended index finger of the dominant hand axially along a line through the space formed by the up-pointing index and little fingers of the nondominant hand, representing the component event of the ball's passing over the goalpost. But it cannot represent the whole event complex within a single expression – say, by flicking one's middle finger against the other hand whose extended index finger then moves off axially along a line.

Further, English can use the same single-clause format to represent events with spatial scope beyond a narrow aperture, for example, an Agent's concurrent causal activity outside any direct manipulation of the Figure, as in *I walked / ran / drove/flew the memo to the home office*. Again, ASL cannot represent the whole event complex of, say, *I ran the memo to the home office* within a single classifier expression. Thus, it could not, say, adopt the classifier for holding a thin flat object (thumb pressed against flat fingers) with the dominant hand and placing this atop the nondominant hand while moving forward with it as it shows alternating strokes of two downward pointed fingers to indicate running (or concurrently with any other indication of running). Instead a sequence of two expressions would likely be used, for example, first one for taking a memo, then one for a person speeding along.[7]

Although the unacceptable examples above have been devised, they nevertheless show that it is physically feasible for a signed language to represent factors related to the Figure's Motion outside its immediate space-time ambit. Accordingly, the fact that signed languages, unlike spoken languages, do avoid such representations may follow from deeper structural causes, such as a greater fidelity to the characteristics of visual perception.

However apt, though, such an account leaves some facts still needing explanation. Thus, on the one hand, it makes sense that the aperture of a classifier expression is limited temporally to the present moment – this accords with our usual understanding of visual perception. But it is not clear why the aperture is also limited spatially. Visual perception is limited spatially to a narrow scope only when attention is being focused, but is otherwise able to process a wide scoped array. Why then should classifier expressions avoid such wide spatial scope as well? Further, sign languages *can* include representation of the Ground object within a single classifier expression (typically with the nondominant hand), even where that object is not adjacent to the Figure.

3.2.3 More independent distinctions representable in the expression

This third property of classifier expressions has two related aspects – the large number of different elements and categories that can be represented together, and their independent variability – and these are treated in succession next.

3.2.3.1 Many more elements / categories representable within a single expression

Although the spatiotemporal aperture that can be represented within a single classifier expression may be small compared to that in a spoken-language clause, the number of distinct factors within that aperture that can be represented is enormously greater. In fact, perhaps the most striking difference between the signed and the spoken representation of space in the expression is that the classifier system in signed language permits the representation of a vastly greater number of distinct spatial categories simultaneously and independently. A spoken language like English can separately represent only up to four or five different spatial categories with closed-class forms in a single clause. As illustrated in the sentence *The bat flew way back up into its niche in the cavern*, the verb is followed in turn by: a slot for indication of path length (with three members: 'zero' for 'neutral', *way* for 'relatively long', *right* for 'relatively short'); a slot for state of return (with two members: 'zero' for 'neutral', *back* for 'return'); a slot for displacement within the earth-frame (with four members: 'zero' for 'neutral', *up* for 'positive vertical displacement', *down* for 'negative vertical displacement', *over* for 'horizontal displacement'); a slot for geometric conformation (with many members, including *in, across, past*); and perhaps a slot for motive state and vector (with two members: 'zero' for 'neutral between location AT and motion TO' as seen in *in / on*, and *-to* for 'motion TO' as seen in *into / onto*). Even a polysynthetic language like Atsugewi has closed-class slots within a single clause for only up to six spatial categories: path conformation combined with Ground type, path length, vector, deixis, state of return, and cause or manner. In contrast, by one tentative count, ASL has provision for the separate indication of thirty different spatial categories. These categories do exhibit certain cooccurrence restrictions, they differ in obligatoriness or optionality, and it is unlikely – perhaps impossible – for all thirty of them to be represented at once. Nevertheless, a sizable number of them can be represented in a single classifier expression and varied independently there. The table below lists the spatial categories that I have provisionally identified as available for concurrent independent representation. The guiding principle for positing a category has been that its elements are mutually exclusive: different elements in the same category cannot be represented together in the same classifier expression. If certain elements can be concurrently represented, they belong to different categories. Following this principle has, on the one hand, involved joining together what some sign language analyses have treated as separate factors. For example, the first category below covers equally the representation of Figure, instrument, or manipulator (handling classifier), since these three kinds of elements apparently cannot be separately represented in a single expression – one or another of them must be selected. On the other hand, the principle requires making distinctions within some categories that spoken languages treat as uniform. Thus, the single 'manner' category of English must be subdivided into a category of 'divertive manner' (e.g. moving along with an up-down bump) and a category of 'dynamic manner' (e.g. moving along rapidly) because these two factors can be represented concurrently and varied independently.

A. Entity properties

1. identity (form or semantic category) of Figure / instrument / manipulator
2. identity (form or semantic category) of Ground
3. magnitude of some major entity dimension
4. magnitude of a transverse dimension
5. number of entities

B. Orientation properties

1. an entity's rotatedness about its left-right axis ('pitch')
2. an entity's rotatedness about its front-back axis ('roll')

3.a. an entity's rotatedness about its top-bottom axis ('yaw')
3.b. an entity's rotatedness relative to its path of forward motion

C. Locus properties

1. Locus within sign space

D. Motion properties

1. motive state (moving / resting / fixed)
2. internal motion (e.g. expansion/contraction, form change, wriggle, swirling)
3. confined motion (e.g. straight oscillation, rotary oscillation, rotation, local wander)
4. translational motion

E. Path properties

1. state of continuity (unbroken / saltatory)
2. contour of path
3. state of boundedness (bounded / unbounded)
4. length of path
5. vertical height
6. horizontal distance from signer
7. left-right positioning
8. up-down angle ('elevation')
9. left-right angle ('direction')
10. transitions between motion and stationariness (e.g. normal, decelerated, abrupt as from impact)

F. Manner properties

1. divertive manner
2. dynamic manner

G. Relations of Figure or Path to Ground

1. path's conformation relative to Ground
2. relative lengths of path before and after encounter with Ground
3. Figure's path relative to the Path of a moving Ground
4. Figure's proximity to Ground
5. Figure's orientation relative to Ground

It seems probable that something more on the order of this number of spatial categories are concurrently analyzed out by visual processing on viewing a scene than the much smaller number present in even the most extreme spoken language patterns.

3.2.3.2 Elements / categories independently variable in the expression – not in pre-packaged schemas

The signed-spoken language difference just presented was mainly considered for the sheer number of distinct spatial categories that can be represented together in a single classifier expression. Now, though, we stress the corollary: their independent variability. That is, apart from certain constraints involving cooccurrence and obligatoriness in a classifier expression, a signer can generally select a category for inclusion independently of other categories, and select a member element within each category independently of other selections. For example, a classifier expression can separately include and independently vary a path's contour, length, vertical angle, horizontal angle, speed, accompanying manner, and relation to Ground object.

By contrast, it was seen earlier that spoken languages largely bundle together a choice of spatial member elements within a selection of spatial categories for representation within the single complex schema that is associated with a closed-class morpheme. The lexicon of each spoken language will have available a certain number of such 'pre-packaged' spatial schemas, and the speaker must generally choose from among those to represent a spatial scene, even where the fit is not exact. The system of generalizing properties and processes seen in section 2.6 that apply to the set of basic schemas in the lexicon (including their plastic extension and deformation) may exist to compensate for the pre-packaging and closed stock of the schemas in any spoken language. Thus, what are largely semantic components within a single morpheme in spoken language correspond to what can be considered separate individually controllable morphemes in the signed classifier expression.

The apparent general lack in classifier expressions of pre-packaging, of a fixed set of discrete basic schemas, or of a system for generalizing, extending, or deforming such basic schemas may well accord with comparable characteristics of visual parsing. That is, the visual processing of a viewed scene may tend toward the independent assessment of spatial factors without much pre-packeting of associated factors or of their plastic alteration. If shown to be the case, then signed language will once again prove to be closer to perceptual spatial structuring than spoken language is.

4 Cognitive implications of spoken / signed language differences

The preceding comparison of the space-structuring subsystems of spoken and of signed language has shown a number of respects in which these are similar and in which they are different. It can be theorized that their common characteristics are the product of a single neural system, what can be assumed to be the core language system, while each set of distinct characteristics results from the activity of some further distinct neural system. These ideas are outlined next.

4.1 Where signed and spoken language are alike

We can first summarize and partly extend the properties above found to hold both in the closed-class subsystem of spoken language and in the classifier subsystem of signed language. Both subsystems can represent multifarious and subtly distinct spatial situations – that is, situations of objects moving or located with respect to each other in space. Both represent such spatial situations schematically and structurally. Both have basic elements that in combination make up the structural schematizations. Both group their basic elements within certain categories that themselves represent particular categories of spatial structure. Both have certain conditions on the combination of basic elements and categories into a full structural schematization. Both have conditions on the cooccurrence and sequencing of such schematizations within a larger spatial expression. Both permit semantic amplification of certain elements or parts of a schematization by open-class or lexical forms outside the schema. And in both subsystems, a spatial situation can often be conceptualized in more than one way, so that it is amenable to alternative schematizations.

4.2 Where spoken and signed language differ

Beside the preceding commonalities, though, the two language modalities have been seen to differ in a number of respects. First, they appear to divide up into somewhat different sets of subsystems without clear one-to-one matchups. Accordingly, the spatial portion of the spoken language closed-class subsystem and the classifier subsystem of signed language may not be exactly corresponding counterparts, but only those parts of the two language modalities closest to each other in the representation of schematic spatial structure. Second, within this initial comparison, the classifier subsystem seems closer to the structural characteristics of visual parsing than the closed-class subsystem in all of the following ways: It has more basic elements, categories, and elements per category in its schematic representation of spatial structure. Its category membership exhibits much more gradient representation, in addition to discrete representation. Its elements and categories exhibit more iconicity with the visual in the pattern in which they are clustered in an expression, in their observance of an object/action distinction, in their physical realization, and in their progression through time. It can represent only a narrow temporal aperture in an expression (and only a narrow spatial aperture as well, though this difference from spoken language might not reflect visual fidelity). It can represent many more distinct elements and categories together in a single expression. It can more readily select categories and category elements independently of each other for representation in an expression. And it avoids pre-packaged category-element combinations as well as generalizations of their range and processes for their extension or deformation.

4.3 A new neural model

In its strong reading, the Fodor-Chomsky model relevant here is of a complete inviolate language module in the brain, one that performs all and only the functions of language without influence from outside itself – a specifically linguistic 'organ'. But the evidence assembled here challenges such a model. What has here been found is that two different linguistic systems, the spoken and the signed, both of them undeniably forms of human language, share extensive similarities but – crucially – also exhibit substantial differences in structure and organization. A new neural model can be proposed that is sensitive to this finding. We can posit a 'core' language system in the brain, more limited in scope than the Fodor-Chomsky module, that is responsible for the properties and performs the functions found to be in common across both the spoken and the signed modalities. In representing at least spatial structure, this core system would then further connect with two different outside brain systems responsible, respectively, for the properties and functions specific to each of the two language modalities. It would thus be the interaction of the core linguistic system with one of the outside systems that would underlie the full functioning of each of the two language modalities.

The particular properties and functions that the core language system would provide would include all the spoken-signed language properties in section 4.1 specific to spatial representation, though presumably in a more generic form. Thus, the core language system might have provision for: using individual unit concepts as the basis for representing broader conceptual content; grouping individual concepts into categories; associating individual concepts with overt physical representations, whether vocal or manual; combining individual concepts -and their physical representations – under certain constraints to represent a conceptual complex; and establishing a subset of individual concepts as the basic schematic concepts that, in combinations, represent conceptual structure.

When in use for signed language, this core language system might then further connect with particular parts of the neural system for visual perception. I have previously called attention to the already great overlap of structural properties between spoken language and visual perception (see Talmy 2000a, ch. 2), which might speak to some neural connection already in place between the core language system and the visual system. Accordingly, the proposal here is that in the case of signed language, still further connections are brought into play, ones that might underlie the finer granularity, iconicity, gradience, and aperture limitations we have seen in signed spatial representations.

When in use for spoken language, the core language system might further connect with a putative neural system responsible for some of the characteristics present in spoken spatial representations but absent from signed ones. These could include the packaging of spatial elements into a stable closed set of patterned combinations, and a system for generalizing, extending, and deforming the packets. It is not clear why such a further system might otherwise exist but, very speculatively, one might look to see if any comparable operations hold, say, for the maintenance and modification of motor patterns.

The present proposal of a more limited core language system connecting with outlying subsystems for full language function seems more consonant with contemporary neuroscientific findings that relatively smaller neural assemblies link up in larger combinations in the subservience of any particular cognitive function. In turn, the proposed core language system might itself be found to consist of an association and interaction of still smaller units of neural organization, many of which might in turn participate in subserving more than just language functions.

Notes

1 Talmy (2003) has been reprinted as the present paper with the permission of Lawrence Erlbaum. The references have been updated. Since the initial publication, all of section 2 on spoken language has been greatly expanded and refined in Talmy (2006). And the implications of spoken-signed differences for the evolution of language are explored in Talmy (2007), which also appears on the author's website:

http://linguistics.buffalo.edu/people/faculty/talmy/talmyweb/index.html

2 I here approach signed language from the perspective of spoken language because it is not at this point an area of my expertise. For their help with my questions on signed language, my thanks to Paul Dudis, Karen Emmorey, Samuel Hawk, Nini Hoiting, Marlon Kuntze, Scott Liddell, Stephen McCullough, Dan Slobin, Ted Suppala, Alyssa Wolf, and others – who are not responsible for my errors and oversights.

3 As it happens, most motion prepositions in English have a polysemous range that covers both the unbounded and the bounded sense. Thus, *through* as in *I walked through the tunnel for 10 minutes* refers to traversing an unbounded portion of the tunnel's length, whereas in *I walked through the tunnel in 20 minutes*, it refers to traversing the entire bounded length.

4 The 'classifier' label for this subsystem – originally chosen because its constructions largely include a classifier-like handshape – can be misleading, since it names the whole expression complex for just one of its components. An apter term might be the 'Motion-event subsystem'.

5 As the only apparent exception, a 'demoted Figure' (see Talmy 2000b, ch. 1) can acquire either of two 'demotion particles' – e.g., English *with* and *of* – that mark whether the Figure's path had a 'TO' or a 'FROM' vector, as seen in *The fuel tank slowly filled with gas / drained of its gas*.

6 The size and shape specifiers (SASS's) in signed languages do permit movement of the hands to trace out an object's contours, but the hands cannot at the same time adopt a shape representing the object's path.

7 The behavior here of ASL cannot be explained away on the grounds that it is simply structured like averb-framed language, since such spoken languages typically *can* represent concurrent Manner outside a narrow aperture, in effect saying something like: 'I walking / running / driving / flying carried the memo to the home office'.

References

Bennett, David C. (1975) *Spatial and temporal uses of English prepositions: An essay in stratificational semantics.* London: Longman.

Bowerman, Melissa. (1989) Learning a semantic system: What role do cognitive predispositions play? In Mabel L. Rice and Richard L. Schiefelbusch (eds) *The Teachability of Language*. Baltimore PH: Brookes Pub. Co.

Bowerman, Melissa. (1996) The origins of children's spatial semantic categories: Cognitive vs. linguistic determinants. In J.J. Gumperz and S.C. Levinson (eds) *Rethinking Linguistic Relativity* 145–176. Cambridge, UK: Cambridge University Press.

Brugmann, Claudia and Macaulay, Monica. (1986) Interacting Semantic Systems: Mixtec expressions of location. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* 315–328. Berkeley: Berkeley Linguistics Society.

Clark, Herb. (1973) Space, time, semantics, and the child. In Timothy E. Moore (ed.) *Cognitive Development and the Acquisition of Language*. New York: Academic Press.

Emmorey, Karen. (2002) *Language, Cognition and the Brain: Insights from Sign Language Research*. Mahwah NJ: Lawrence Erlbaum.

Fillmore, Charles. (1968) The case for case. In Emmon Bach and Robert T. Harms (eds) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston.

Gruber, Jeffrey S. (1965) Studies in lexical relations. PhD dissertation, MIT. Reprinted as part of *Lexical structures in syntax and semantics*, 1976. Amsterdam: North-Holland.

Herskovits, Annette. (1982) Space and the prepositions in English: Regularities and irregularities in a complex domain. PhD dissertation, Stanford University.

Imai, Shingo. (2003) *Spatial Deixis: How Demonstratives Divide Space*. Doctoral dissertation. University at Buffalo, State University of New York.

Jackendoff, Ray. (1983) *Semantics and Cognition*. Cambridge, MA: MIT Press.

Leech, Geoffrey. (1969) *Towards a semantic description of English*. New York: Longman Press.

Liddell, Scott. (2003) Sources of meaning in ASL classifier predicates. In Karen Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Language* 199–220. Mahwah, NJ: Erlbaum.

Mark, David M. and Smith, Barry. (2004) A science of topography: From qualitative ontology to digital representations. In Michael P. Bishop and John F. Shroder (eds) *Geographic Information Science and Mountain Geomorphology* 75–100. Chichester, England: Springer-Praxis.

Talmy, Leonard. (1983) How language structures space. In Herbert L. Pick, Jr. and Linda P. Acredolo (eds) *Spatial orientation: Theory, research, and application*. New York: Plenum Press.

Talmy, Leonard. (2000a) *Toward a Cognitive Semantics*, volume I: *Concept structuring systems*. Cambridge, MA: MIT Press.

Talmy, Leonard. (2000b) *Toward a Cognitive Semantics*, volume II: *Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Talmy, Leonard. (2003) The representation of spatial structure in spoken and signed language. In Karen Emmorey (ed.) *Perspectives on Classifier Constructions in Sign Language* 169–195. Mahwah, NJ: Lawrence Erlbaum.

Talmy, Leonard. (2006) The fundamental system of spatial schemas in language. In Beate Hampe (ed.) *From perception to meaning: Image schemas in Cognitive Linguistics* 199–234. Berlin: Mouton de Gruyter.

Talmy, Leonard. (2007) Recombinance in the Evolution of Language. In Jonathon E. Cihlar, David Kaiser, Irene Kimbara and Amy Franklin (eds) *Proceedings of the 39th Annual Meeting of the Chicago Linguistic Society: The Panels*. Chicago: Chicago Linguistic Society.

Zubin, David and Soteria Svorou. (1984) Orientation and gestalt: Conceptual organizing principles in the lexicalization of space. With S. Choi. In David Testen, Veena Mishra and Joseph Drogo (eds) *Lexical semantics*. Chicago: Chicago Linguistic Society.

14 Geometric and image-schematic patterns in gesture space

Irene Mittelberg

1 Introduction

The human body exists, moves, interacts, and communicates in space and time. Inseparable from the human body, manual gestures, too, unfold and vanish in space and time. They derive their meaning in part from the coinciding speech and in part from particular combinations of hand shapes, hand motions, and their location in gesture space. Over the last few decades, research on co-speech gesture and signed languages has shown that these dynamic visuo-motor modalities do not only exploit various dimensions of physical space as their articulatory medium, but that they can also provide a window into how physical, conceptual, social, and discourse spaces interact (e.g., Emmorey and Reilly 1995; Liddell 2003; Kendon 2004; McNeill 1992, 2000, 2005; Müller 1998; Núñez and Sweetser 2006; Parrill and Sweetser 2004; Sweetser 2007; Taub 2001; Wilcox 2000; Wilcox and Morford 2007).

Within cognitive linguistics, gesture data from typologically different languages have proven to be a valuable source of multimodal evidence for conceptual metaphor and particularly for spatial metaphor. A considerable body of research done on metaphorical gestures, e.g., representations of abstract ideas and structures, has demonstrated their capacity to reveal source domain information not necessarily captured by concurrent verbal expression (Bouvet 2001; Cienki 1998a, 1998b; Cienki and Müller to appear; McNeill 1992, 2005; Mittelberg to appear; Müller 1998, 2004b; Núñez 2004; Sweetser 1998, 2007; *inter alia*). Moreover, a recent experimental study (Cienki 2005) suggests that basic image and force schemas manifest themselves in gesture. Due to their specific materiality and logic, gestures are particularly apt at depicting spatial and dynamic properties of conceptual structure and processes, thus supporting the theory of the embodied mind (Gibbs 1994, 2003, 2006; Lakoff and Johnson 1980, 1999).

Indeed, basic physical activities that involve hand motions and/or bodily movement through space – such as walking, grasping, touching, pointing, placing, and exchanging physical objects – exhibit metaphorical correspondences in the domains of thought and speech: we understand something if we can 'grasp' it, we 'walk' people through texts, 'point out' certain aspects, 'push an issue', or try to 'get ideas across' to our interlocutors (cf. Sweetser 1992). Exploring how such habitual actions play out in gesture, the aim of this paper is to offer insights into the ways in which scholars employ gestures to illustrate their discourse about abstract knowledge domains. On the basis of academic discourse videotaped in linguistics courses, I will show how gestural depictions may bring intangible subject matters into physical existence that can be shared by professors and their students. The main point of interest here is the spatialization of abstract information pertaining to grammatical concepts and theories. I will demonstrate that the prominent hand shapes and motion patterns that were found to recur across subject matters and speakers form a set of patterns which are reminiscent of simple geometric figures (e.g., squares, triangles, cubes, circles), as well as image and motor schemas proposed in the cognitive linguistics literature (e.g., object, path, balance, support, container, rotation; cf. Hampe 2005; Johnson 1987; Lakoff 1987; Mandler 1996, 2004; Talmy 1988). The term *geometric* here refers to basic shapes evoked by constellations of arms and hands and by forms resulting from imaginary lines drawn in the air. It will be suggested that a kind of 'common-sense geometry' (Deane 2005:245) may be, among other kinds of conceptual structures and motor routines, one of the factors that motivate what have turned out to be fairly systematic representations of linguistic form, grammatical categories, and syntactic relations. In view of the important role such embodied schemas have been found to assume in language acquisition (Mandler 1996, 2004), language per se (Talmy 1988), and also in the visual arts (Johnson 1987; Mittelberg 2002, 2006, in prep.), it might not be all that surprising to also see some of them reflected in gesture. The aim here is to show that discerning them in this dynamic bodily modality is useful in diagnosing less monitored aspects of cognition during communication.

While the work presented here is part of a larger study investigating how such patterns play into the iconic, metaphorical, and metonymic meaning construction in multimodal discourse (Mittelberg 2006, 2007, 2008), the discussion below will focus almost exclusively on the material side of the semiotic processes that seem to ground abstract thought in the speakers' bodies and the surrounding space.[1] This paper is thus about how abstract information is spatially represented through gesture – and not about the gestural depiction of spatial concepts or scenes per se (see Sweetser 2007 for an overview).

Before moving into the heart of the study, let us look at an example from the data in order to get a first impression of how gestures may ascribe meaning to chunks and regions of space. In the sequence from which the image below is taken (Figure 1), the speaker talks about the difference between main verbs and auxiliaries. During his explanation leading up to this particular gesture, he points to instances of both verb types contained in sentences projected onto the screen behind him. He then goes on to say that auxiliaries such as 'have', 'will', 'being' and 'been', 'must all belong to some subcategory'. Upon mentioning 'some subcategory', he produces the gesture shown below, consisting of two hands that seem to be loosely holding an imaginary object. The extended arms and almost flat hands jointly evoke two diagonally descending lines.

The meaning of the term 'subcategory' is effectively represented by a gesture that is produced in a comparatively low region of gesture space, low not only in relation to the speaker's body, but also in relation to preceding and subsequent gestures. In fact, the hand configuration appears well below the region where this speaker and also the other subjects of the this study produce the majority of gestures referring to grammatical categories and sentence structure. It is thus an unusual, marked usage of space (Waugh 1982), which receives some of its semantic properties in relation to the unmarked region of gesture space (in front of the speaker's torso) which indirectly functions as a point of reference.

Figure 1. Gesture representing 'subcategory' placed comparatively low in gesture space

If one were to accompany the same term ('subcategory') with the same object gesture but located, say, in front of one's chest, the effect of the gestural illustration would not be as insightful. And, if one were to produce the same gesture on the mention of a word referring to a concrete item, it would express that concrete entity and not, as in this case, an abstract category. Here, the abstract category is metaphorically represented in terms of an imaginary physical object (or container) that fills the space between the two hands. It can be seen as reflecting the metaphorical concept IDEAS ARE OBJECTS or CATEGORIES ARE CONTAINERS (Lakoff and Johnson 1980). At the same time, a second spatial metaphor is evoked: the 'subcategory' is literally placed underneath the superordinated category it relates to. In the course of the paper, we will explore various ways in which space becomes meaningful in gestural representations of grammar.

The structure of the chapter is as follows: section 2 describes the data and methodology of this study. Section 3 presents the results of the form analysis, providing an overview of the prominent hand configurations and motion patterns. In section 4, the findings are discussed in light of A) image and motor schemas proposed in the cognitive linguistics literature and B) issues of object representation and spatial relations more generally. The chapter concludes with a summary of the main characteristics of the gestures discussed and suggestions for further research.

2 Data and methodology: discourse genre, transcription, and coding parameters

2.1 Corpus

The corpus designed for this research comprises twenty-four hours of naturalistic academic discourse and co-speech gestures produced by four linguists (all native speakers of American English; three females and one male). The subjects were videotaped while lecturing in introductory linguistics courses at two American universities. The focus of attention is on the communicative behavior of the professor lecturing; student behavior and teacher–student interaction are not considered here. Topics covered include general aspects of morphology, syntax, and phonology as well as different linguistic theories: generative grammar, emergent grammar, and relational grammar. Correspondingly, a major part of the discourse revolves around the introduction of new concepts and technical terms. In this highly specialized type of multimodal discourse, the objects referred to are for the most part abstract entities and structures: linguistic units (morphemes, words, phrases, etc.), grammatical categories (verb classes, cases, semantic roles, etc.), syntactic structures (clauses, sentences, etc.), as well as operations (the active-passive transformation, subordination, reiteration, etc.). In search for multimodal representations of these entities, the corpus was assessed from a thematic point of view, selecting and capturing episodes in which gestures portraying grammatical phenomena occurred. Such 'referential gestures' may depict, according to Müller's functionalist typology of gestures (1998:110–113), objects, attributes of objects and people, actions, behaviors, etc. Müller further distinguishes referential gestures of concrete entities from gestures depicting abstract entities. As most of the gestures discussed here refer to abstract phenomena, they can be said to be essentially metaphorical in nature. In each semiotic act different iconic and indexical (i.e., metonymic) modes were found to interact to different degrees, but we will not be able to go into these issues of interpretation here (see McNeill 1992, 2005 on gesture categorization and Mittelberg 2008 and Mittelberg and Waugh 2009 for more details on the interaction of metaphor and metonymy in meta-linguistic gestures).

Not only the subject matter talked about, but also cultural practices and pedagogical routines influence the kinds of gestures that accompany meta-linguistic discourse. Given that in Western cultures language is represented as horizontally oriented strings of written words, habits of writing and reading from left to right and filling text spaces from top to bottom can be expected to motivate, among other factors, the graphic representation of language and grammar in gesture. Common practices in grammar and linguistics courses also need to be taken into account, such as diagramming sentence structure and dissecting sentences into functional parts (see Jakobson 1966 for an account of why grammatical patterns lend themselves so well for graphic representation). These factors as well as the use of mediational tools such as blackboards, whiteboards, overhead projectors, and laptops influence the kinds of gestural signs produced in this specific context as well as their exact execution in relation to the technical equipment and the spatial environment of the classroom.

Working with multimodal usage data involves a series of steps which will be only briefly sketched here. First, the speech of each segment was transcribed adapting the discourse transcription convention provided by Du Bois and colleagues (Du Bois et al. 1993). Then, the gestures were coded according to their kinetic features (see section below) and, in relation to the concurrent speech, the exact speech–gesture synchrony was documented in annotated transcripts.[2] To this end, the course of each gestural movement (which may include onset, preparation, peak, hold, and return to rest) gets translated into typographic representations, superimposed on the transcribed speech. Each gesture was traced from the moment the articulators (here hands and arms) begin to depart from a rest position until the moment when they return to rest or relaxation. Such a full movement excursion (Kendon 2004:111) is called a gesture-unit (G-unit): 'The G-unit is defined as the period of time between successive rests of the limbs; a G-unit begins the moment the limb begins to move and ends when it has reached a rest position again' (McNeill 1992:83). Only gestures articulated with hands and arms were taken into account, leaving aside facial expressions, gaze, self-grooming, and movements of the head and torso (for more details on methods and sample transcripts see Mittelberg 2007).

2.2 Physical gesture features: hand shape, palm orientation, and movement

In gesture research, the most widely used coding parameters are *hand presence* and *hand dominance*, *hand shape*, *palm orientation*, *movement* (trajectory and type), and the *location* in gesture space where a gesture is performed (cf. McNeill 1992, 2005; Kendon 2004; Müller 1998, 2004; Webb 1996). These kinetic features were also used to describe the referential gestures in the present corpus, thereby determining those qualities of a gesture gestalt that contribute most significantly to its meaning and function. For example, in certain cases, the movement proved to be more salient with respect to the meaning of a gesture than the particular shape of the hand performing the movement (e.g., in certain pointing gestures it did not matter whether the hand pointing was a relaxed flat hand or whether the index finger was extended); in other cases, the hand shape is more salient than the contextual movements (e.g., in the case of hands forming a closed fist); and in yet other cases, both dimensions are significant (e.g., a push with an open palm facing the addressee, thus building a barrier and evoking the idea of 'stop' or 'rejection'). As we saw in the subcategory example above, the location in which the gesture is produced may also significantly contribute to its meaning and function.

In order to categorize the *hand shapes*, a data-driven typology of manual signs was developed.[3] The data were searched for hand shapes and arm configurations that recurred across speakers and contexts, and a label was assigned to each prominent form. For example, one of the most frequently used hand shapes is a flat open hand with the palm turned upwards, thus building a sort of surface. Here it seemed worthwhile to build on conventions introduced by Müller (2004) in her study of forms and functions of the palm-up open hand gesture (hereafter referred to as 'puoh'). Each variant of the open hand gesture that occurred in the data was given an abbreviation such as 'puoh', indicating the *orientation of the palm*, plus a short name evoking the degree of openness of the hand ('tray', 'cup', 'lid', etc.) as well as an indication of which hand performed the gesture. For instance, 'puoh-tray-lh' stands for a flat palm-up open hand, produced with the left hand, evoking the shape of a tray. Or, 'pcoh-box-bh' stands for another frequent gesture consisting of two hands held apart, with both palms being held vertically and facing each other and thus pointing to the center of gesture space (i.e., 'pcoh' stands for palm-center open hand and the 'center' denotes the direction that the palm is facing). A variant of this gesture was discussed above in the subordination example (Figure 1).

Gestures typically involve some sort of *movement* through space and are as such a comparatively fluid medium: they usually vanish as quickly as they emerge, often melting into one other. Describing such manual actions entails the range and *trajectory* of the performed motion (for example, along horizontal, vertical, or diagonal axes) as well as the *manner* of the movement (straight line, wave, rotation of the wrist, etc.). When a gesture appeared unusually forceful, the energy level with which the movement was carried out was taken into account. Instances in which a movement is discontinued or a configuration is being held (e.g., the so-called gesture hold, cf. McNeill 1992) were also recorded. In keeping with the notational conventions used for hand shapes, the prominent movement patterns were given labels that inform about their trajectory and manner. For example, 'vert-trace-rh' signifies a line that is traced vertically with the right hand, and 'wrist-rota-lh' refers to a wrist rotation performed with the left hand.

2.3 Location in gesture space

Manual gestures take shape in physical space. The range, organization, and preferred use of a person's gesture space is conditioned by factors such as age (children vs. adults), cultural background, and personal style, among others (cf. Calbris 1990; Goldin-Meadow 2003; Kendon 2004; McNeill 1992; Müller 1998). Not surprisingly, the space parameter has entered gesture research in various ways, shedding light on spatial cognition, culturally-determined conceptualizations of space, etc. (cf. Haviland 2000; Levinson 1997, 2003; Núñez and Sweetser 2006; Sweetser 2007). Gesture space is relative to, and constituted by, the position and posture of the speaker-gesturer who, in each communicative instance, sets up the coordinates of gesture space around her, according to the dimensions and movements of her body, her gestural articulators (here arms and hands), her physical environment, and, if applicable, also according to the interpersonal, social space spanning between herself and her interlocutor(s). The location of a gesture can be described from various angles: relative to the gesturer's body, relative to previously or subsequently produced gestures, or relative to the addressee's gesture space. In gesture, space is exploited to indicate and describe the location of objects, people, places, events, and ideas, as well as the spatial relationships among entities and persons, a task that is generally more difficult to master with purely linguistic means (cf. Emmorey 1996; Emmorey and Reilley 1995 regarding the use of space in signed languages).

In terms of the perspective from which a scene or an object may be described in a given speech event, the speaker-gesturer can represent alternate viewpoints: observer viewpoint, character/participant viewpoint, as well as the addressee's viewpoint (cf. McNeill 1992:118–25; Sweetser 2007). It is probably a matter of teaching experience and pedagogical awareness whether a teacher assumes her or his own point of view or the audience's perspective. In any event, these considerations determine how the use of gesture space is organized. When freely gesturing (and not pointing at information on the blackboard or screen), the professors videotaped for the present study were most of the time facing their student audience, and both observer viewpoint and addressee's viewpoint could be made out in their gestural descriptions of grammatical categories and structures. For example, the subjects alternatively illustrated the word order in a sentence by drawing an imaginary line starting either on the left side and ending on the right side of their body, or in the opposite direction, from the students' left to the students' right side. Some cognitive and perceptive flexibility thus needs to be assumed at both ends of the speech and gesture event (for a discussion of frames of reference in ASL see Emmorey 1996; Liddell 2003; Wilcox and Morford 2007).

To document the locations where gestures occur and the trajectory they trace, gesture researchers have developed systems to compartmentalize gesture space into sectors. For example, McNeill established a shallow disk consisting of concentric squares superimposed on a drawing of a seated person, thus reflecting the semi-experimental set-up in which speakers were asked to retell animated cartoons (McNeill 1992:86–89, 2005:274). Since the conditions under which the present data were collected were not controlled in any way, and since teachers tend to walk around in the classroom and constantly change their position and the angle with which they turn towards the audience, blackboards, overhead projectors, laptops, etc., there were no stable space coordinates. Instead of investigating the relative density of occurrence of certain gesture types in particular sectors of gesture space (e.g., in relation to different body parts), or correlating gesture location and discourse function, which are possible ways to exploit the space factor in gestural communication (cf. McNeill 1992: 88ff.), one of the main interests here was to determine the ways in which the speakers' use of gesture space could reveal aspects of their spatial representations of abstract phenomena. This is, as will be shown below, where different geometric and image-schematic representations of linguistic form and structure come into play.

3 Study: prominent hand configurations and motion patterns in meta-grammatical gestures

The aim of this section is to provide an overview of the prominent gestural forms that were found to illustrate verbal explanations of linguistic form, grammatical relations and syntactic functions. The point of departure here was the physical forms of gestures exhibited in the data (i.e., hand figurations, manual actions, or imaginary lines drawn in the air). Only then did the analysis turn to the abstract ideas and structures which gestural signs stand for in a given moment, taking into account the concurrent speech.

Since the scope of the paper does not allow for a detailed account of the cross-modal distribution of semantic features and pragmatic functions, the discussion below will be mainly restricted to the material properties of the gestures (see Mittelberg 2006, 2008 and Mittelberg and Waugh 2009 for detailed content analyses).[4]

3.1 Prominent hand and arm configurations

The gestalt of a given gesture relies on the semiotic collaboration of several parameters, of which the hand shape is only one. Yet, the hand shape and/or arm configuration can be said to be salient in a gesture if it is the most notable feature in the process of its articulation. While most of the gestures to be discussed below involve some kind of movement, it is the hand shapes and arm configurations that, especially when being held for a moment, tend to stand out perceptually. As in the example discussed above (Figure 1), the movement leading up to the object-holding gesture is not as perceptually and semantically salient as the bimanual configuration produced on the mention of the term 'subcategory'. Factoring in the speech content it becomes evident that both the specific hand and arm configuration plus its location contribute key qualities to the bi-modally achieved message.

Across the four subjects, the data show recurrent representations of linguistic units as readily manoeuvrable objects. There are several different ways of holding and manipulating such imaginary items, some of which allude to the geometry and/or size of the object, while in other cases no or very little information about the size or form of the object can be inferred. One way to refer to an abstract item is to seemingly hold something placed on a palm-up open hand (puoh). The degree to which the hand is flat, relaxed, or cupped varies from case to case. The potential functions of this basic hand shape have been matched with the actions of holding, presenting, or offering an imaginary object for inspection, and these functions have been observed in diverse contexts (Müller 2004). Variants of the palm-up open hand gesture, also called 'palm presentation' gestures (Kendon 2004) or 'conduit gesture' (McNeill 1992, 2005), were frequently observed in the teaching contexts under investigation here, especially when professors talk about abstract categories or linguistic examples not visibly present in the immediate environment (an alternative would be to point to words written on the blackboard).

The following list comprises the different open-hand variants found in the data, some of which will be illustrated and discussed in more detail below. As indicated in the methods section above, each type was assigned an abbreviation referring to the openness and orientation of the palm (such as 'puoh') plus a 'name' and an indication of which hand was used (some of the palm-up open hand abbreviations follow Müller 2004). Finally, an abbreviation signals which hand was used. While, theoretically, the hand shapes listed below could be produced simultaneously by each hand, they were for the most part observed to be executed with only one hand at a time.

Single open and closed hands

rh: right hand; lh: left hand

A. puoh-tray-lh/rh	hand as flat surface, supporting imaginary objects
B. puoh-cup-lh/rh	hand with curled fingers, forming a receptacle
C. pfoh-stop-lh/rh	'f' stands for 'front', palm facing audience
D. pdoh-lid-lh/rh	'd' stands for 'down', flat hand
E. pdoh-claw-lh/rh	open hand facing down, fingers curled
F. pcoh-blade-lh/rh	'c' stands for palm facing center of gesture space
G. fist-lh/rh	closed fist

The last gesture type listed above is in fact the opposite of an open hand: it is a closed hand forming a fist. Other hand shapes involving specific finger configurations include 'measure' (thumb and index finger are stretched apart, tips pointing upwards, similar to the way one might take measure in inches), 'pinch' (the tips of index finger and thumb are pressed against one another), and 'scrunch' (fingers are held closely together, facing audience, tips pointing towards the floor).

Specific finger configurations

H. t-i-measure-lh/rh	't' for thumb, 'i' for index
I. pinch-lh/rh	fingertips of index and thumb pressed together
J. scrunch-lh/rh	similar to pinch, but different orientation and finger configuration, back of hand facing audience, tips pointing towards floor

Another category of gestural shapes engages not only hands but also parts of a speaker's arm(s). Most of the observed pointing gestures fall into this category, as they are usually produced with both an extended arm and hand, exhibiting either an extended index finger ['ind-index'] or the entire, mostly relaxed hand ['hand-index']. Together, hand and arm build a vector, or a path, leading to the targeted referent (e.g., an object, a person, information written on the blackboard, or to certain locations in gesture space right in front of the speaker). In addition, there were arm configurations depicting chunks of a syntactic tree diagram by mirroring the triangle-like shape of such diagonally downward branching structures ['diag-arm'].

Pointing gestures and other kinds of arm configurations

K. ind-index-lh/rh	[pointing with generic extended index finger]
L. hand-index-lh/rh	[pointing with full, relaxed hand]
M. diag-arm-lh/rh	[arm held diagonally, forming a triangle-like shape if both arms are involved]

Other gestures observed in the data are always performed with two hands, evoking an internal structure, or what has been called 'syntax' (cf. Kendon (2004:275ff.) on Open Hand Supine gestures with lateral movement). Examples are the gesture mentioned above in which the imaginary object is held between two hands, or a gesture conveying the idea of a balance by seemingly weighing two things, with two palm-up open hands moving alternately up and down, one on each side of the body.

Open hand variants performed with both hands (bh)

N. puoh-tray-lateral	[balance]
O. puoh-cup-lateral	[balance]
P. puoh-sym-offshoot	[hands thrown laterally up into the air, from center outward]
Q. pcoh-box-bh	[refers to an elongated object held between both hands]

The data were searched for instantiations of each of these identified shapes (and movement patterns, to be discussed below) across topics and speakers. For most of these forms, several instances were identified and assessed with regard to the concurrent speech content and the overall meaning of the multimodally achieved representation. Below, a selected set of these hand shapes will be illustrated and discussed in more detail.

3.1.1 Single open and closed hands: surfaces and containers for abstract entities

Comparatively small linguistic units, such as morphemes, words, and categories were represented as objects seemingly resting on a variant of the palm-up open hand gesture or inside a closed fist. The gestures shown in Figures 2 and 3 are instances of palm-up open-hand gestures with a flat palm or cupped hand evoking a kind of surface or a receptacle where items can be placed (i.e., imagined) and presented to the audience. From just looking at the hand shape it might not be clear whether the action the hand is performing represents an act of offering, receiving, showing, or requesting an item. In conjunction with the speech content, however, it turns out that the gesture in Figure 2, for instance, represents the action of receiving. It denotes a technical term, namely the semantic role 'recipient', by showing an open hand ready to receive an object. A similarly shaped gesture fulfills a different function in Figure 3, where the speaker is explaining the fact that an idea can materialize in discourse in the form of a noun or a verb. On the mention of 'a noun' she creates a sort of tray on which the emerged form is being presented to the audience.

Figure 2. puoh-cup stands for recipient

Figure 3. puoh-tray stands for a noun

By seemingly handling small imaginary objects, linguistic units are thus reified and made graspable for the mind. Flat open hands provide surfaces, planes, or, put more generally, support structures, exposed to the eye of the addressee, on which the item referred to in the speech modality can be imagined. Alternatively, the absence of an entity or the expectation to receive something can be signaled. Similar functions can be performed by cupped hands (with clearly curled fingers), building a sort of open container (see Figure 12 below). While the focus here is on the formal properties of open hand gestures, it needs to be kept in mind that the meaning of a gesture results from both its form and the function it plays in a given speech event (see Müller 2004 for a detailed account of forms and uses of the palm-up open hand gesture and also Kendon 2004). Abstracting from these pragmatic considerations, the central point here is that these open hand gestures seem to embody the image schemas SUPPORT (Mandler 1996) and CONTAINMENT (Johnson 1987; Lakoff and Johnson 1980) respectively.

As the next examples suggest, imaginary small objects can also be held in tightly closed hands. In Figure 4, the speaker refers to grammatical 'knowledge' while forming a fist (left hand) and to the idea that 'knowledge becomes automatized' with usage when forming a second fist (right hand). While talking about the fact that the word 'teacher' consists of two parts (the morphemes 'teach-' and '-er'), the speaker in Figure 5 encloses each component in a fist: the right hand holds the lexical morpheme 'teach-' and the left hand the grammatical morpheme '-er'. The spatial difference between the two hands evokes the conceptual difference between the two functionally distinct elements forming one word, thus instantiating the metaphorical concept PHYSICAL DISTANCE IS CONCEPTUAL DISTANCE (Sweeter 1998). At the same time, the two hands jointly allude to the internal structure of the word 'teach/er'. In both cases, the fists are first formed successively and then held simultaneously, as shown in the figures below (see Mittelberg 2008 on diagrammatic iconicity holding between the two hands).

Figure 4. Fist(s) for knowledge (grasp/mastery)

Figure 5. Fists containing morphemes (teach-er)

While here, too, the image schema CONTAINMENT manifests itself in these gestural representations, the fist seems to have, compared to the open hand variants, a different semantic import. It evokes the idea of having, literally and metaphorically, captured a concept, of having a firm grasp of it: one knows how to handle a certain phenomenon. Inside the closed hand, there is no space for maneuvering. At the same time, the object enclosed in the hand container is invisible and not much information about it is accessible, which stands in contrast to exposing an idea on an open hand for inspection and commentary, or alluding to the fact that one does not have an answer and is thus 'empty-handed' (cf. Müller 2004).

3.1.2 Different amounts of space between the articulators

We will now look at some hand shapes where the configuration of individual fingers and the existence or nonexistence of space between the articulators play a significant role. The two examples below represent cases of what was called a pinch in the list provided above. A pinch involves the index finger and thumb pressed together. For example, the gesture shown in Figure 6 expresses the idea of a precise list of categories in the theory of relational grammar, by drawing, with the index finger and thumb pressed together (indicating the idea of 'precise') a vertically descending line (depicting the idea of a 'list'). In a similar fashion, the gesture in Figure 7 features no space between the fingertips. However, unlike the gesture in Figure 6, it bears a stronger resemblance to what is generally known as the ring gesture due to the slightly more rounded fingers; this gesture occurs across cultures and contexts with different coded meanings, ranging from tangibility, to precision and perfection (Kendon 2004; McNeill 2005; Müller 1998). Here, in the context of a syntax lecture, it has a different designation: it co-occurs with the mention of the technical term 'node' which is a juncture point at the top of a branching structure in tree diagrams used in the framework of generative grammar.

Figure 6. Pinch indicating precise list of categories

Figure 7. Pinch/ring indicating 'node' (tree)

An alternative way to refer to small items is to seemingly hold them between the tips of thumb and index finger, as if one were taking measure. In other words, there is some space between the two fingers, which might suggest a virtual object filling the space. For example, in Figure 8, the small space between index finger and thumb indicates the compact nature of the pronoun 'it', alluding to the placement and function of such minimal forms in phrasal verb constructions. The gesture in Figure 9 stands for a verb form ('fell') at the end of a sentence ('Diana fell.').

Figure 8. Measure representing pronoun 'it'

Figure 9. Measure representing verb 'fell'

A gesture heavily used to represent parts of speech, words, phrases, and sentences depicts a comparatively bigger imaginary object as being held by two, relatively relaxed, open hands with palms facing each other. The examples below show two of the more expansive versions in which speakers hold the hands relatively far apart to represent a sentence (Figure 10) or a constituent (Figure 11). These gestures can also be said to reflect the image schema CONTAINMENT or, if one focuses on the fact that phrases and sentences have a beginning and an end, by the SOURCE-PATH-GOAL schema (Johnson 1987; Lakoff and Johnson 1980, 1999).

Figure 10. pcoh-box representing a sentence

Figure 11. pcoh-box representing a constituent

In view of the representations discussed so far, space seems to carry meaning in specific ways. Although there is no direct correspondence between the amount of space extending between the articulators and the physical characteristics of the elements referred to, there is a tendency for smaller individual linguistic units to be represented as being held in one hand (Figures 2, 3, 5, 8, 9) and for comparably more complex constructs such as entire phrases or sentences to be represented by objects held (or space extending) between both hands of the speaker (Figures 10 and 11). In the latter cases, the geometry of the objects held between two hands is specified to a higher degree than the shape of objects seemingly sitting on open hands and remains rather undefined. In both scenarios, however, the mind needs to fill in information according to the cues provided by the hand constellations as well as the concurrent speech content (see Mittelberg and Waugh 2009). It should be noted that it is difficult at times to decide whether one can assume objects or whether it is rather about delineating the space extending between fingers or hands.

3.1.3 Pointing gestures and specific arm configurations

While parts of the speakers' arms were involved in different fashions in many of the gestures discussed above, we now turn to configurations in which arms are instrumental in the gestural sign formation. As we will see, arms may be recruited to build signposts in pointing gestures or to directly stand for elements of the object they depict (cf. Müller's (1998) modes of gestural representation).

The spatial orientation and angle of pointing gestures depend each time on the location of the object towards which they are directed. Through the act of pointing at something in the proximity of the speaker (e.g., on the mention of a demonstrative pronoun) the object is established via a vector consisting of a path evoked by the extended arm and hand and its virtual extension leading to the targeted object. Such deictic gestures highlight spatial relationships between the speaker and objects, locations, or people, whether they are present in the environment, imagined, or previously introduced in the unfolding discourse (cf. Fricke 2002, 2007; Furuyama 2001; Kita 2003; McNeill 1992, 2005; McNeill et al. 1993; Sweetser 2007; Williams 2004).

To illustrate and anchor their explanations, the speakers frequently point to information presented on blackboards, whiteboards, or overhead screens. An example of this is given below (Figure 12). Talking about the difference between main verbs and auxiliary verbs, the speaker points with his right hand to words projected onto the screen behind him (on the mention of 'there is'), thus creating a vector between the position of his body (i.e., the deictic center or *origo* of the speech act, according to Bühler 1934) and the referent of the concurrent deictic expression. It can also be taken as an instantiation of the SOURCE-PATH-GOAL schema with the path leading the interpreting mind to the object referred to. Completing his sentence (started with 'there is'), the speaker forms with the left hand a cupped palm-up open hand gesture (on the mention of 'the main verb',). A concrete example of a 'main verb' is being pointed at on the screen ('taught'), while the abstract category as such is to be imagined as being inside the cupped hand directed towards the student audience.

Figure 12. Index ('there is') plus cup 'the main verb'

Figure 13. Semantic roles 'bounce around'

Another way of assigning meaning to space is to virtually place things in gesture space or to simply point to locations in space, for instance when enumerating a list of things. In the example above (Figure 13), the speaker talks about the different 'semantic roles that bounce around in linguistics', and represents each type of semantic role with a different gesture produced in a different place. The gesture shown here is made on the mention of the term 'agent' (we already looked at the gesture for 'recipient', cf. Figure 2). Metaphorically speaking, this gesture can be interpreted to reflect the metaphor IDEAS ARE LOCATIONS; it can also be seen as an instance of metonymy of place (PLACE FOR OBJECT). By dispersing categories in space, the physical distance between the assigned locations represents the conceptual distance between the different semantic roles and their respective functions (agent, patient, goal, recipient, experiencer) thus evoking the metaphorical mapping CONCEPTUAL DISTANCE IS PHYSICAL DISTANCE (Sweetser 1998).

As for gestural constellations involving both arms, let us look at Figure 14. Here the speaker illustrates a part of a syntactic tree diagram by forming a triangle-like shape, achieved with the fingertips touching at the center top and both forearms held diagonally with elbows pointing outwards. Mirroring a part of the diagram on the blackboard behind the speaker, the evoked pyramid directly imitates a tree chunk. Put differently, the arms of the speaker embody conceptual structure. Such depictions provide more substance than lines quickly traced into the air and lend, as such, otherwise relatively fleeting representations a higher degree of stability in space and time.

Figure 14. Phrase structure as a lateral diagonally branching tree chunk

Figure 15. index pointing to ground: subordination

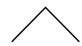
Illustrating the idea of subordination, the speaker in Figure 15 combines a pointing gesture with a representational gesture that can also be interpreted as standing for a tree branch descending to the right lower side of her body. She indicates that there are certain cases in which embedded sentences go 'all the way down', at which point she directs her fully extended right arm towards the floor and points with her index finger straight to the ground. As it was the case in the subcategory example discussed above (Figure 1), the descending arm evokes a spatialization of the idea of subordination by reaching into comparably low regions of gesture space. Alternatively, the speaker drew the same kind of geometric configurations in the air, tracing either only a single diagonal line or two diagonal lines downward, one to each side of her body (see also Figure 14). This kind of dynamic representation serves as a bridge into the section below where motion patterns will be discussed.

3.2 Motion patterns

In gesture, form can be created not only by hand and arm constellations, but also by fleeting hand movements that draw simple lines or contours of objects in the air, thus leaving imaginary traces in gesture space. Identifying significant motion patterns recurring in the data entailed determining for each dynamic gestural gestalt those qualities that contribute most significantly to its meaning. Again, this can ultimately only be done in correlation with the concurrent speech content and particularly with those speech segments that coincide with the peak, or 'stroke' phase, of a gesture (McNeill 1992). In most cases, hand shape and movement do interact in one way or another; yet, the discussion below concentrates on the different trajectories and/or manners of those hand motions that appear constitutive of the gestural signs (which in turn stand for the abstract ideas and structures they convey).

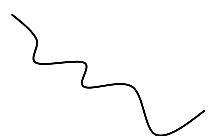
One can generally distinguish between several types of gestural movements. For example, the movement of a hand can result in the evocation of a form (such as the size and shape of a guitar). It may also be influenced by the object that is involved in the action imitated by the hand movement (such as the unlocking of a door with an imaginary key), or it can simply imitate a manual action (such as waving at somebody) or the manner and/or speed of a movement executed by a person or an object (for research on motion events and the description of movement and manner in gesture see McNeill 1992, 2000; Müller 1998; Slobin 2003). The hand movements observed in the present data were also found to exhibit several intrinsic logics: first, movements carried out by hands tracing straight lines or curved lines imitating the shape of a wave, circle, or arch (these movement types bring to bear the different planes in the gesture space such as horizontal, vertical, and front-back); second, there are pointing gestures whose direction and range depend on the location of the object or person pointed at (cf. section 3.1.3); third, object-oriented actions such as placing something; and fourth, basic motor actions with no object involved, such as two hands rotating around each other. These distinctions concur with previously made observations that a large number of gestural shapes and movements originate in concrete object manipulation and are abstracted from and structured by routinized interactions between the human body and the physical and social world. Accounting for the noted variety, Müller's (1998) system of modes of gestural representation include manual actions such as such as drawing, molding, enacting or embodying (see also Calbris 2003; LeBaron and Streeck 2000; Streeck 2002). Although the schematic representations and drawings provided below only render frozen visualizations of dynamic gestural gestalts, and while this sort of qualitative approach needs to be complemented by quantitative investigations across subject matters and speakers, the identified patterns offer a window into some of the ways in which hand movements unfolding in a teacher's gesture space may reveal aspects of the underlying conceptualizations of abstract concepts and structures. The following typology of gestural motion patterns was established:

Linear movements (horizontal/vertical/diagonal)

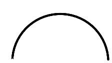
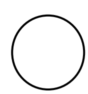
A. hori-trace-lh/rh	horizontal line	
B. vert-trace-lh/rh	vertical line	
C. diag-trace-rh	diagonal line	
D. diag-trace-ll	diagonal line	
E. diag-trace-lat	lateral diagonal line	
F. scale-lh/rh	hand trace vertically organized steps/levels	
G. hori-join-lat	horizontal line drawn with both hands going inward or lateral inward movement or a more forceful push	
H. hori-part-lat	horizontal line drawn with both hands, (lateral outward movement)	
I. push-lh/rh/bh	push away from body along a straight line, not curved (exploiting depth along sagittal axis)	
J. pull-lh/rh/bh	pull toward body along a straight line, not curved (exploiting depth along sagittal axis)	

While the movements listed above exhibit linear trajectories along the major axes, non-linear representations along the horizontal and the vertical axes also occurred; additional non-linear configurations include both half and full circles:

non-linear traces

K. hori-wave-lh/rh	wavy line traced in the air, along a horizontal axis	
L. diag-wave-lh/rh/	wavy line traced in the air, along a diagonal axis	

curves and circles

M. curve-up-lf/rh	hand(s) move(s) along upper half of circle	
N. curve- dn-lf/rh	hand(s) move(s) along lower half of circle	
O. circle-lh/rh/bh	hand(s) complete(s) one full cycle, rotation	

Other motor actions of hands, not involving simple traces of the manipulation of imaginary objects, include the following two types of rotations:

P. rotation-lateral	both hands (and arms) draw circles repeatedly rotating around one another
Q. wrist-rota-lh/rh/bh	wrist rotation, occurs with different orientations

Below, I will discuss several examples of hand movements that evoke dynamic images of abstract entities and processes – even if it is just via an imaginary trace left in the air (the dynamic nature of the movements can unfortunately not be fully appreciated without viewing the video clips).

3.2.1 Linear movements (horizontal and vertical traces)

Sentences and other sequences of linguistic units were found to be represented by movements tracing the horizontal alignment of words from the left to the right of the speaker ('hori-trace'), or, if the viewpoint of the audience was assumed, from right to left. A slight variation of such schematic representations of sentence structure is shown in Figure 16 below, where the gesture starts out with both hands joined at the center of gesture space, right in front of the upper torso of the speaker. Subsequently, the hands move laterally outward until both arms are fully extended, as if they were tracing, as mentioned in the concurrent speech, 'a string of words' ('hori-part').

Figure 16. A sentence as a string of words

Figure 17. The infix goes into the middle of another morpheme

The gesture in Figure 16 depicts 'a sentence' as a 'string of words' drawn horizontally in the air, with both hands starting in front of the speaker's chest and being pulled outward to each side of the body. A vertically descending line was already shown in the gesture representing a list of categories (Figure 6); it also underlies the gesture whose beginning point is illustrated in Figure 17 above. After setting the stage by mentioning that words in English may have prefixes and suffixes, the speaker explains the position infixes take in the structure of a complex word. In the moment captured above, the speaker is just about to insert an infix into a word stem. The idea of insertion is depicted by a well-defined vertical trajectory traced by the hand (executed on the mention of 'morphemes that go right into the middle of another morpheme'), until the hand seems to hit the base form which he quickly sketches as a container by drawing its horizontal base line and then alluding to its two outer sides with a bimanual palm-center open hand gesture.

In addition to horizontal lateral outward movements such as the string depicted in Figure 16, the data also exhibit lateral inward movements that are executed with a higher energy level. For example, as shown in Figure 18, the idea that, according to the theory of emergent grammar, boundaries between grammar and language use are 'blurred' is illustrated by a gesture that starts out with two hands apart, palms facing each other, but the palms then get suddenly pushed towards each other to convey the idea of fusion. Similarly, the speaker in Figure 19 talks about the behavior of words that like to 'go together' and 'travel together to the front of the sentence', which is portrayed by two fists being quickly and repeatedly brought together. In both cases, physical closeness signals conceptual closeness and is achieved through physical forceful action. We can thus observe an interaction between image and force schemas.

Figure 18. 'hori-join', blurring boundaries

Figure 19. 'hori-join', words go concepts together (travel)

3.2.2 Non-linear traces

Let us now look at some non-linear motion patterns. The first two images below show instances of wave-like motions along a horizontal axis. In Figure 20, the speaker draws, on the mention of 'non-linearity', a wave-like graph consisting of a first curve going down and a second one going up. The speaker in Figure 21 makes an almost identical motion to represent the concept of 'intonation contour', except that the motion goes in the opposite direction.

Figure 20. Horizontal wave for 'non-linearity'

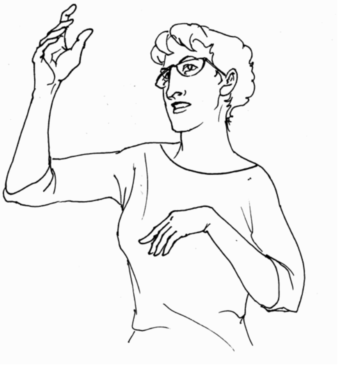


Figure 21. Horizontal wave for 'intonation contour'

There are also instances of larger arch-like structures that are executed with both hands. The following gestural demonstration, taken from a morphology lecture, provides an example of the understanding that the two elements that jointly build a circumfix (by surrounding the word stem) seem to be attached at a level above the word level. In her attempt to illustrate the hidden organization of such complex morphological structure (or, the strings attached), the speaker makes an arch-like gesture whose initial phase is captured in the image below (Figure 22). After holding both hands above head level, the speaker simultaneously draws them down to waist level, one hand to the left and one to the right of her body. The idea that the 'circumfix encompasses the front and back of the word' is subsequently represented by a bimanual palm-center open hand gesture (not shown here; it resembles the gesture in Figure 11). Her two hands seem to be holding the entire morphological structure by its front and back, where the indications 'front' and 'back' do not refer to spaces closer to or farther away from the speaker's body (which would refer to the sagittal axis that runs through her body from the space behind her back to the space in front of her). Rather, the front of the word is located to the left of the speaker and the back of the word to her right, in accordance with the conceptualization of written words and sentences as extending from left to right in front of the speaker/reader/writer (in Western cultures).

Figure 22. Circumfix as arch gesture

Another kind of arch-like gesture was found in the context of teaching the framework of relational grammar. Whereas the gesture in Figure 22 is a spontaneous depiction of morphological structure, other arch-like gestures have been observed that were motivated by a standardized diagram used in the framework of relational grammar (the diagrams are often compared to igloos or umbrellas). For example, a speaker explains the concept of 'multi-attachment' (i.e., the idea that subject and reflexive pronoun refer to the same person) as follows: first, the right (dominant) hand rises to head level and comes down making a slight arch-like swing to the right. Then, the left hand rises and makes a similar arch-like movement downward (this gesture is not reproduced here). In the videotape one sees the corresponding diagram on the blackboard in the background of the speaker; it shows exactly the kind of lines that the speaker draws in the air. Correspondingly, this gesture visualizes syntactic relations in terms of spatial structure: schematic arch-like lines cutting through several zones layered on top of each other.

It is important to keep in mind that some of the gestures discussed above are informed by a particular theoretical view of grammatical concepts and relations (i.e., generative grammar or relational grammar). Without the relevant theoretical background it would probably be difficult to make sense of such gestural diagrams. They are dynamic renditions of hypothesized conceptual relations translated into spatial configurations; without any kind of visual support, their adequate description in solely linguistic terms would probably be less economic and also less effective (see Mittelberg 2008 for a Peircean approach on image and diagrammatic iconicity in such metaphoric gestures). What all these gestures have in common is that they are based on theories that rely on a specific set of metaphors representing different understandings of language and grammar.

To conclude this section, we can say that some of the geometric gestural representations (diagonals, triangles, and arches) are in fact not the spontaneous creations of the speakers, but they instead are rooted in scientific conventions. The manual routine of literally drawing diagrams on paper or blackboards is likely to influence how speakers represent connections between words or grammatical constituents via hand movements through space. Those shapes and motion patterns that are created ad-hoc seem to be motivated, at least in part, by object-oriented actions (such as drawing, writing, and manipulating objects), and specific motor actions (e.g., wrist rotation). In all the cases, however, the anchor points for these representations are the human body and its articulators' range of possible movements as well as the dimensions constituted by the physical classroom setting and teaching tools. Embodied practices are exploited to fleetingly visualize conceptual images of abstract entities and structures in terms of physical objects, bodily actions, and locations in space. These forms of mediation between the conceptual and the embodied may offer insights, as will be detailed below, into the central role played by image and motor schemas (and their metaphorical projection) which seem to motivate and structure, at least partly, gestural representations of abstract knowledge domains and other types of intangible things such as values and beliefs in a systematic way.

4 Discussion: dynamic manifestations of geometric and image-schematic patterns

The gestures examined here are ephemeral and partial representations of objects and actions that metaphorically refer to abstract entities and operations. As the spectrum of emergent patterns discussed above suggests, some of the gestural forms and movements indeed reflect geometric and image-schematic representations of grammatical concepts and structures. Due to the fluid character of the gestural medium, the schematic images are never fully visible at once; they may find expression in a virtual trace left by a hand movement or by invoking the manual action of holding an object. It is left to the mental eye of the addressees, or to their own bodily experience with such actions, to fill in the missing pieces. In what follows, I will take these observations a step further and address some of their implications in terms of image and motor schemas (section 4.1) and regarding geometric representations of objects and spatial relations more generally (section 4.2).

4.1 Gestural instantiations of image and motor schemas

As we have seen above, hand shapes and movements collaborate in building holistic gestural gestalts. The study presented here has revealed some of the ways in which the salient properties of such multidimensional figurations give minimal information that may evoke full schemas of objects and actions. In what follows, I would like to elaborate the idea that the prominent patterns identified in the data can be recruited as tangible, non-verbal evidence for image schemas which are assumed to be part of the 'cognitive unconscious' (Lakoff and Johnson 1999:9–15). Rereading Johnson's (1987:XIV) original definition of image schemas as 'recurring, dynamic patterns of our perceptual interactions and motor programs that give coherence and structure to our experience' with gesture in mind, reinforces the assumption that gesture is a crucial source of manifestations of such embodied patterns and that in order to account for their dynamic nature one needs to consider not only visual but also kinesthetic aspects of image schemas (see also Cienki 1998 a/b, 2005 and Sweetser 1998, 2007).

Based on gestural representations of grammar, the present work offers support for the 'semiotic reality of image schemas' (Danaher 1998:190). In particular, the following correspondences between gestural patterns (cf. section 3) and basic image schemas are suggested (cf. Johnson 1987; Lakoff and Johnson 1980, 1999; Mandler 1996, 2004):

SUPPORT	('puoh-tray', 'puoh-cup')
CONTAINMENT	('puoh-cup', 'fist')
OBJECT	('puoh-tray', 'puoh-cup', 'pcoh-box',, 'fist')
SOURCE-PATH-GOAL	('hori-trace', 'vert-trace', 'diag-trace;' deictics such as 'hand-index', 'ind-index', 'diag-arm')
EXTENSION	('hori-trace', 'vert-trace', 'diag-trace;' deictics such as 'hand-index', 'ind-index', 'diag-arm')
BALANCE	('puoh-tray-bh', 'puoh-cup-bh', 'fist-bh', 'sym-offshoot')
SCALE	('scale')
CENTER-PERIPHERY	('sym-offshoot' 'hori-join', 'hori-part')
CYCLE	('circle-bh', 'wrist-rotation', 'rotation lateral')
ITERATION	('wrist-rotation', 'rotation lateral')
FRONT-BACK	('push', 'pull')
FORCE	('push', 'pull', 'hori-join', 'sym-offshoot')

The schemas PART-WHOLE, LINK, CONTACT, and ADJACENCY are not discussed here, yet contiguity relations (i.e., metonymy) proved to be particularly relevant for representing relationships between individual elements jointly constituting an entire phrase or sentence (cf. Mittelberg 2008 and Mittelberg and Waugh 2009. Furthermore, basic geometric shapes (e.g., circles, semi-circles, triangles, rectangles, squares) were identified as well as straight and curved lines traced along horizontal, vertical, and diagonal axes, as well as the sagittal axis (front-back). Perhaps not too surprisingly, the list above contains for the most part spatial and spatial relations image schemas which are assumed to structure systems of spatial relations cross-linguistically (Lakoff and Johnson 1999:35). In his experimental study on image schema manifestations in co-speech gesture, Cienki (2005) tested the potential of image schemas (i.e., PATH, CONTAINER, CYCLE, OBJECT, and FORCE) as descriptors for several types of gestures accompanying discourse on matters of honesty. Results suggest that 'images schemas are readily available, indeed "on hand" for recruitment as gestural forms' (Cienki 2005:435); they may be represented in the gesture modality as either static entities or dynamic processes. Cienki also found that gestures can invoke different schemas than the accompanying linguistic track, thus providing additional information to discourse participants.

The array of image schemas Cienki (2005) employed in his experimental study as well as the above list of image-schematic patterns found in the present discourse data contains some of the schemas that belong, according to Mandler (1996:373–8), to the preverbal, spatially structured meaning system: SUPPORT, CONTAINMENT, PATH, and CONTACT. Mandler maintains that these image schemas are spatial representations, or spatial abstractions, that result from perceptual analysis, which in the development of infants comes before object manipulation.[5] Moreover, such spatial analyses performed by the infant are supposed to be important in learning the relational aspects of language, e.g., the meaning of verbs and locative prepositions such as 'on' and 'in' (cf. Bowerman 1996; E. Clark 1973; H. Clark 1973). Although no conclusive statements can be made on the basis of the observations presented here, Mandler's assertions are relevant regarding the spatialization of grammatical relations in a modality that is utilized for communication by infants prior to language (cf. Goldin-Meadow 2003 on gesture and language development). Mandler also concedes that dynamics and internal feelings would be more difficult to analyze. Here, too, gesture research promises to further augment our understanding of the bodily logic of image and particularly of force schemas (Talmy 1988).

Recent work on image schemas comprises a variety of understandings and definitions (see contributions in Hampe 2005), but overall the notion of embodiment seems to be taken more and more literally: there is a tendency towards the realization that the human body's intuitive expressions and culturally-shaped practices represent a rich source of insight into how higher cognitive activities may be grounded in dynamic patterns not only of bodily perception and movement, but also of social behavior. Johnson (2005) strongly advocates the importance of putting flesh on image-schematic skeletons and of trying to account for the felt qualities of meanings and situations (see also Cienki 2005; Deane 2005; Gibbs 2005; Zlatev 2005). One of the central questions still seems to be how multi-faceted meanings, especially in abstract reasoning, emerge from embodied experience:

> But let us not forget that the truly significant work done by image schemas is tied to the fact that they are not merely skeletons or abstractions. They are recurring patters of organism-environment interactions that exist in the felt qualities of our experience, understanding, and thought. Image schemas are the sort of structures that demarcate the basic contours of our experience as embodied creatures. […] Their philosophical significance, in other words, lies in the way they bind together body and mind, inner and outer, and thought and feeling. They are an essential part of the embodied meaning and provide the basis for much of our abstract inference. (Johnson 2005:31)

In view of Johnson's (2005:31) exhortation to 'analyze various additional strata of meaning, such as the social and affective dimensions, to flesh out the full story of meaning and thought', it seems safe to say that bodily semiotics generally bear the potential to inform us about qualities that are difficult to access via purely linguistic inquiry. Gesture data remain a promising source to explore both structured and intuitive aspects of how we make meaning and also of how we make sense of what others try to convey. In light of these considerations, we can perhaps better appreciate the extent to which the present gesture data bring out the dynamic and embodied aspects of image-schematic and geometric representations of abstract objects and structures: gestures are not simply visual, but visuo-motoric and a bodily medium; hence, they have the capacity to shed additional light on the assumed multimodal character of concepts and image schemas (cf. Evans and Green 2006). Contrary to static visual representations of words, sentences, and diagrams captured on paper or blackboards, these gestures afford a 'representation of abstract processes as dynamic patterns' (Kendon 1997:112) through a 'dynamic visuo-spatial imagery' (McNeill et al. 2001:11). Linguistic form and structure seem to come to life: branches branch out, words move or travel together to the front of a sentence, and boundaries between concepts get blurred. Instrumental hand actions seemingly manipulating items highlight the process character of operations such as prefixation, suffixation, infixation, or the construction of a sentence. In addition, grammatical operations such as 'reiteration' and 'recursion' were found to be represented by the rotation of a single hand or by two hands revolving around each other, and a similar motor schema was observed to signify the function of a morphological case or the idea of active language use as opposed to the knowledge of grammar. In gesture research**,** the 'bodily basis of meaning, imagination and reason', the title of Johnson's (1987) ground-laying book, may be taken literally, thus trying to illuminate not only the relationship between the gesturer's body and the imaginary objects and forces it interacts with, but also to explore how meanings are conveyed through minimal movements or forceful hand actions.

4.2 Dynamic representations of objects in places: some preliminary considerations on the 'what' and 'where' in gesture space

Being aware of the preliminary character of the following reflections, I would like to draw together two central aspects that make co-speech gesture a promising source of insights into the relationship between cognition, space, and language: its spontaneous, unreflective character on the one hand and its tendency to reflect schematic imagery and basic geometric forms on the other.

Due to the attention gesture draws to what I like to think of as the '*ex*-bodiment' (Mittelberg 2006, 2008) of internalized imagery and experiences with the physical and social world, and due to its propensity to directly portray spatial and sensory-motor aspects of concepts and source domains of metaphorical mappings, gesture research has yielded insights into our understanding of abstract knowledge domains (Calbris 2003; Cienki 1998, 2005; McNeill 1992; Müller 1998, 2004; Sweetser 1998, 2007; Núñez 2004; Taub 2001). Since gestures unfold in space, they are naturally apt at illuminating spatial metaphor, not only regarding linguistic form and structure, but also regarding, for instance, the spatial representation of moral concepts (Cienki 1998 a/b), mathematical thought (McNeill 1992; Núñez 2004; Smith 2003), and concepts belonging to the domain of speech communication (Sweetser 1998).

It is because of their unreflective character that gestural representations of abstract phenomena can offer fresh insights into the metaphorical nature of the conceptual system and, more generally, into less monitored aspects of cognition during communication. Crucially, in the present data, metaphorical understandings of abstract entities are frequently expressed in the gesture modality even if the accompanying speech is non-metaphorical. The technical term 'subcategory' (shown in Figure 1) is a good example of this kind of multimodal representation of abstract concepts: the metaphorical understanding of a category in terms of a container or object is conveyed only in the gesture modality, not in speech. Other examples would be technical terms such as 'noun', 'constituent', 'node', 'sentence', and 'morpheme' or words or parts of words such as 'fell', 'teach-,' and '-er'. In contrast to carefully planned and executed pictorial metaphors deployed in advertisements, cartoons, and paintings, spontaneous metaphorical gestures may provide more intuitive renditions of mental imagery, created locally and online (see Mittelberg and Waugh 2009 and Müller and Cienki 2009 on multimodal metaphor).

Arguing in favor of a multimodal approach to spatial representations, Deane (2005:245) discusses instances in which spatial prepositions evoke a 'common-sense geometry'; he asserts that 'the same spatial relation may receive distinct representations in multiple representational modalities' (p. 247). In view of the configurations observed in the present gesture data, it seems that the speakers do apply a sort of common-sense geometry when ascribing basic shapes to linguistic entities (e.g., in the form of bounded objects) and structures (e.g., in the form of lines and diagrams, the latter exploiting both horizontal and vertical axes to spatially portray hierarchical relations).

A question that poses itself here concerns the degree to which the imaginary metaphorically construed objects are geometrically specified. Talmy (1983) suggested universal constraints as to how figure object and ground object are geometrically schematized in locative expressions; he noted an asymmetry to the effect that the figure object tends to be relatively shapeless and the ground object tends to be more precisely defined (cf. Landau 1996:321ff.; Landau and Jackendoff 1993). Investigating how the visual-spatial modality might condition descriptions of the relation between two objects, Emmorey (1996:175–9) found the tendencies identified by Talmy to hold in ASL, where, in fact, 'the use of space to directly represent spatial relations stands in marked contrast to spoken languages' (p. 175). She also found that signers tend to express the ground first and then the figure object, conceiving of the figure as a point with respect to a more complex ground (p. 179). In the present data, this process was found in gestural descriptions in which, for instance, a string of words (as in Figure 16) was first drawn in the air and subsequently functioned as a sort of virtual reference structure in which the word order of particular linguistic units was pointed out. The same is true in regard to tree diagrams which, once they are sketched out in air, provide slots where elements such as embedded clauses may be placed (cf. Mittelberg 2006). However, much more research is needed to develop a better understanding of the mechanisms of what one could call, with recourse to Landau and Jackendoff (1993), the 'what' and 'where' in gesture space.

Now, if we wanted to describe the relationship between objects and gestural articulators in light of figure/ground relationships as well as the relative specification of objects in terms of their geometry, we could, in a first approach, say the following: in cases where an imaginary object (i.e., the figure) is sitting on a palm-up open hand (i.e., the ground), it exhibits a less specific geometry than the hand itself (see Figures 2, 3, 12). In most of these scenarios, details of size or shape are not provided for the figure object, except for the fact that a single hand cannot hold a very large object. In gesture, space may carry meaning in various ways, and, as we saw above, the different amounts of space between hands or fingers may signify linguistic units of different degrees of complexity (e.g., a morpheme in Figure 9 vs. a sentence in Figure 10). The object/box gestures (Figures 1, 10, 11) seem to be more strongly profiled in terms of their size and volume and might thus qualify as geometrically idealized representations of objects, i.e. manifestations of what Talmy (1983) referred to as the 'flexible schematizing of objects' (Landau 1996:319). By contrast, different kinds of pointing gestures were found to simply assign a location, but no shape, to grammatical categories (such as semantic roles; see Figure 13). Here, we could conceive of the space in front of the speaker as the ground, this time rather vaguely defined. One could argue that these objects do not receive much specification because they signify imaginary abstract entities and that, since that which they stand for is revealed in the concurrent speech, it might be sufficient to just point to their existence and, if applicable, to their specific spatial arrangement. In fact, the gestures here take care of the 'where' of the entities, which also entails their position with respect to one another (e.g. the placement pronouns in phrasal word constructions, Figure 8, or the insertion of an infix, Figure 17). This bimodal strategy is highly economic and makes verbal paraphrases (i.e., prepositional phrases) unnecessary. While some of these observations indicate the kind of asymmetry suggested by Talmy (1983), more research is needed to correlate the geometry of objects and their relations in the gesture modality with cognitive and discourse-pragmatic factors such as, for instance, attention, perceptual saliency, information flow, pragmatic inferencing, and the exact cross-modal encoding of spatial information.

5 Concluding remarks

Gesture assigns meaning to space. It employs hand shapes, movement, and space to describe not only physical objects and their spatial relationships, but also spatial models underlying abstract knowledge domains and other concepts that are difficult to represent such as time, values or emotions. The gestures discussed in the present paper have, as I hope to have shown, the capacity to unite phenomena that at first might appear contrasting in one way or another, including the interrelation between form and motion, spontaneity and systematicity, and the abstract and the concrete.

First, in the gesture modality form may become motion and motion may become form (FORM IS MOTION, cf. Lakoff and Turner 1989). Hands may dynamically represent the form of an object by drawing its contours in the air (such as the wave-like movements representing the notion 'intonation contour', see Figure 21); or the virtual trace left by a manual motion may evoke a form (such as a virtual container in which items can be subsequently placed, see Figure 10). A gestural sign may depict the formal essence of an entity and/or its characteristic movement, both of which can be used independently of the perception or presence of the object. In addition, gestures can portray the process character of mental operations of which we often only see the final product, for example an assembled word or sentence (e.g., infixation, see Figure 17).

Second, despite their spontaneous and unreflective dimensions, gestural representations have been shown to exhibit a considerable degree of systematicity regarding both the form they take and the space they exploit. There is more and more converging evidence that the factors motivating the structure of gestures of the abstract include embodied image and motor schemas, conceptual metaphor and metonymy (Bouvet 2001; Cienki 1998, 2005; Cienki and Müller 2008; McNeill 1992, 2005; Mittelberg 2006; Müller 1998, 2004b; Núñez 2004; Núñez and Sweetser 2006; Sweetser 1998, 2007; Taub 2001), as well as routine object-oriented actions and practices of social interaction (Calbris 2003; Clark 2003; LeBaron and Streeck 2000; Kendon 2004; Müller 1998, 2004; Streeck 2002; *inter alia*).

Third, metaphoric gestures mediate between the abstract and the concrete: while being abstracted from physical objects and actions, they make abstract phenomena tangible. By isolating the essential properties of the objects and actions they represent, they provide insights into the abstractive capacities and embodied structures of the human mind, and incarnate the principles of conceptual metaphor and abstract inferencing (Johnson 2005). In the meta-grammatical discourse analyzed here, linguistic form and structure seem to propel manifestations of a set of image-schematic and geometric patterns in the gesture modality. Embodied 'common-sense geometry' (Deane 2005:245) thus manifests itself in these gestures to a certain degree, and it would be interesting to see whether such tendencies appear in gestures accompanying discourses about other abstract subject matters (cf. Cienki 2005; Núñez 2004; Smith 2003; Sweetser 2007). Such work could further attest to the embodied nature of basic image and motor schemas in general and spatial-relations concepts in particular (Lakoff and Johnson 1999:34ff.; Hampe 2005; Talmy 1988). Another promising avenue for further research would be to explore the pragmatics of the 'flexible schematizing of objects' and the relative geometry of figure and ground objects in co-speech gesture (Talmy 1983; Emmorey 1995; Landau 1996; Landau and Jackendoff 1993).

Theoretical, academic discourse might have the reputation of being dry, technical, and objective; however, the multimodal classroom discourse examined here is strikingly dynamic, immediate, and engaging. The professors' gestures convey not only visuo-spatial illustrations of grammatical concepts and theories, but also intuitive, felt qualities of thought and meaning-making processes which no doubt deserve further (cross-disciplinary) attention.

Acknowledgements

I am grateful to the editors and an anonymous reviewer as well as to Jana Bressem, Alan Cienki, Jacques Coursil, Sotaro Kita, Silva Ladewig, Cornelia Müller, Michael Spivey, Eve Sweetser, and Linda Waugh for stimulating discussions and insightful comments on earlier versions of this chapter. I also thank Allegra Giovine, Joel Ossher, and Daniel Sternberg for their valuable help with database design and data coding and Yoriko Dixon for providing the artwork.

Notes

1 The approach to multimodal discourse developed in Mittelberg (2006) combines Peircean semiotics (Peirce 1955), Jakobson's theory of metaphor and metonymy (Jakobson 1956), and contemporary cogntivist approaches to metaphor and metonymy (see also Mittelberg and Waugh 2009).

2 Gesture researchers have suggested various schemes for how to graphically capture not only the close temporal relationship between speech and co-speech gesture, but also the kinetic features of gestures (cf. Calbris 1990; Duranti 1997:144–154; Kendon 2004; McNeill 1992, 2005; Müller 1998:175–199, 284ff.; Parrill and Sweetser 2004; *inter alia*). This study has particularly been inspired by the methods of transcription, coding, and analysis developed by members of the McNeill Lab (McNeill 1992), Müller (1998, 2004a) and Webb (1996).

3 Another possibility would have been to adopt the form inventory of a signed language such as American Sign Language (c.f. McNeill 1992:86–88; Webb 1996).

4 I thank Allegra Giovine and Daniel Sternberg for their invaluable collaboration on this part of the analysis.

5 Here a link can be made to abstraction in the visual arts. Georges Braque and Pablo Picasso developed their Cubist transformations of people and everyday objects through extracting their most essential characteristics (see Mittelberg 2006 and in prep.).

References

Boroditsky, L. (2001) Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43: 1–22.

Bouvet, D. (2001) *La dimension corporelle de la parole. Les marques posturo-mimo-gestuelles de la parole, leurs aspects métonymiques et métaphoriques, et leur rôle au cours d'un récit*. Paris: Peeters.

Bowermann, M. (1996) Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and Space* 385–436. Cambridge, MA: MIT Press.

Bühler, K. (1934/1965) *Sprachtheorie. Zur Darstellungsfunktion der Sprache*. 2nd ed. Stuttgart: Fischer.

Calbris, G. (1990) *The semiotics of French gestures*. Bloomington: Indiana University Press.

Calbris, G. (2003) From cutting an object to a clear cut analysis: Gesture as the representation of a preconceptual schema linking concrete actions to abstract notions. *Gesture* 3(1): 19–46.

Cienki, A. (1998a) Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J.-P. Koenig (ed.) *Discourse and cognition: Bridging the gap* 189–204. Stanford: CSLI Publications.

Cienki, A. (1998b) STRAIGHT: An image schema and its metaphorical extensions. *Cognitive Linguistics* 9: 107–149.

Cienki, A. (2005) Image schemas and gesture. In B. Hampe (ed.) in cooperation with J. Grady, *From perception to meaning: Image schemas in cognitive linguistics* 421–442. Berlin/New York: Mouton de Gruyter.

Cienki, A. and Müller, C. (eds) (2008) *Metaphor and gesture.* Amsterdam/ Philadelphia: John Benjamins.

Clark, E. V. (1973) Non-linguistic strategies and the acquisition of word meanings. *Cognition* 2: 161–182.

Clark, H. H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive development and the acquisition of language* 27–63. New York: Academic Press.

Clark, H. H. (1996) *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2003) Pointing and placing. In S. Kita (ed.) *Pointing: Where language, culture, and cognition meet* 243–268. Mahwah, NJ: Lawrence Erlbaum Associates.

Danaher, D. (1998) Peirce's semiotic and cognitive metaphor theory. *Semiotica* 119(1/2): 171–207.

Deane, P. D. (2005) Multimodal spatial representations: On the semantic unity of over. In B. Hampe (ed.) in cooperation with J. Grady, *From perception to meaning: Image schemas in cognitive linguistics* 235–282. Berlin: Mouton de Gruyter.

DuBois, J., Schuetze-Coburn, S., Cumming, S. and Paolino, D. (1993) Outline of discourse transcription. In J. A. Edwards and M. D. Lampert (eds) *Talking data: Transcription and coding in discourse research* 45–87. Mahwah, NJ: Lawrence Erlbaum Associates.

Duranti, A. (1997) *Linguistic anthropology*. Cambridge: Cambridge University Press.

Emmorey, K. and Reilly, J. (eds) (1995) *Language, gesture, and space*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Emmorey, K. (1996) The confluence of space and language in signed languages. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and space* 171–209. Cambridge, MA: MIT Press.

Evans, V. and Green, M. (2006) *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Gibbs, R. W., Jr. (1994) *The poetics of mind: Figurative thought, language, and understanding*. Cambridge, UK: Cambridge University Press.

Gibbs, R. W., Jr. (2003) Embodied experience and linguistic meaning. *Brain and Language* 84: 1–15.

Gibbs, R. W., Jr. (2005) The psychological status of image schemas. In B. Hampe (ed.) in cooperation with J. Grady, *From perception to meaning: Image schemas in cognitive linguistics* 113–135. Berlin/New York: Mouton de Gruyter.

Gibbs, R. W., Jr. (2006) *Embodiment and cognitive science.* New York: Cambridge University Press.

Goldin-Meadow, S. (2003) *Hearing gesture: How hands help us think*. Cambridge, MA/London: The Belknap Press of Harvard University Press.

Hampe, B. (ed.) (2005) *From perception to meaning: Image schemas in cognitive linguistics.* Edited in cooperation with J. Grady. Berlin/New York: Mouton de Gruyter.

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

Haviland, J. B. (2000) Pointing, gesture spaces, and mental maps. In D. McNeill (ed.) *Language and gesture* 13–46. Cambridge: Cambridge University Press.

Jakobson, R. (1956/1990) Two aspects of language and two types of aphasic disturbances. In L. R. Waugh and M. Monville-Burston (eds) *Roman Jakobson, on language* 115–133. Cambridge/London: Harvard University Press.

Jakobson, R. (1961/1987) Poetry of grammar and grammar of poetry. In K. Pomorska and S. Rudy (eds) *Roman Jakobson, language in literature* 121–144. Cambridge/London: Belknap Press of Harvard University Press.

Jakobson, R. (1966/1990) Quests for the essence of language In L. R. Waugh and M. Monville-Burston (eds) *Roman Jakobson, on language* 407–421. Cambridge/London: Harvard University Press.

Johnson, M. (1987) *The body in the mind: The bodily basis of meaning, imagination, and reason.* Chicago: University of Chicago Press.

Johnson, M. (2005) The philosophical significance of image schemas. In B. Hampe (ed.) in cooperation with J. Grady, *From perception to meaning: Image schemas in cognitive linguistics* 15–33. Berlin/New York: Mouton de Gruyter.

Kendon, A. (1997) Gesture. *Annual Review of Anthropology* 26: 109–128.

Kendon, A. (2000) Language and gesture: Unity or duality? In D. McNeill (ed.) *Language and gesture* 47–63. Cambridge: Cambridge University Press.

Kendon, A. (2004) *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, S. (ed.) (2003) *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lakoff, G. (1987) *Women, fire, and dangerous things. What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G. (1993) The contemporary theory of metaphor. In A. Ortony (ed.) *Metaphor and thought* 202–251. 2nd ed. Cambridge: Cambridge University Press.

Lakoff, G. and Johnson, M. (1980) *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. (1999) *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Landau, B. (1999) Multiple geometric representations of objects in languages and language learners. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and space* 316–363. Cambridge, MA: MIT Press.

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

LeBaron, C. and Streeck, J. (2000) Gestures, knowledge, and the world. In D. McNeill (ed.) *Language and gesture* 118–138. Cambridge: Cambridge University Press.

Levinson, J. (1996) Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M.A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and space* 109–169. Cambridge, MA: MIT Press.

Levinson, J. (1997) Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. *Journal of Linguistic Anthropology* 7 (1): 93–131.

Levinson, J. (2003) *Space in language and cognition: Explorations in cognitive diversity.* Cambridge: Cambridge University Press.

Liddell, S. (2003) *Grammar, gesture, and meaning in American Sign Language*. Cambridge: Cambridge University Press.

Mandler, J. (1996) Preverbal representation and language. In P. Bloom, M. A. Peterson, L. Nadel and M. F. Garrett (eds) *Language and space* 365–384. Cambridge, MA: MIT Press.

Mandler, J. (2004) *The foundations of mind*. Oxford/New York: Oxford University Press.

McNeill, D. (1992) *Hand and mind: What gestures reveal about thought.* Chicago: Chicago University Press.

McNeill, D. (ed.) (2000) *Language and gesture*. Cambridge: Cambridge University Press.

McNeill, D. (2005) *Gesture and thought*. Chicago: Chicago University Press.

McNeill, D., Cassell, J. and Levy, E. T. (1993) Abstract deixis. *Semiotica* 95 (1/2): 5–19.

McNeill, D., Quek, F., McCullough, K.-E., Duncan, S.D., Furuyama, N., Bryll, R. and Ansari, R. (2001) Catchments, prosody and discourse. *Gesture* 1(1): 9–33.

Mittelberg, I. (2002) The visual memory of grammar: Iconographical and metaphorical insights. *Metaphorik.de* 2/2002: 69–89.

Mittelberg, I. (2006) Metaphor and metonymy in language and gesture: Discourse evidence for multimodal models of grammar. Unpublished Doctoral Dissertation, Cornell University.

Mittelberg, I. (2007) Methodology for multimodality: One way of working with speech and gesture data. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson and M. J. Spivey (eds) *Methods in cognitive linguistics* 225–248. Amsterdam/Philadelphia: John Benjamins.

Mittelberg, I. (2008) Peircean semiotics meets conceptual metaphor: Iconic modes in gestural representations of grammar. In A. Cienki and C. Müller (eds) *Metaphor and gesture* 115–154. Amsterdam/New York: John Benjamins.

Mittelberg, I. (in prep.) Metonymy in gesture and Cubism: A comparative study of abstraction, essence, and relativity in semiotic structure.

Mittelberg, I. and Waugh, L. R. (2009) Metonomy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In C. Forceville and E. Urios-Aparisi (eds) *Multimodal metaphor* 322–358. Berlin/New York: Mouton de Gruyter.

Müller, C. (1998) *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich.* Berlin: Berlin Verlag A. Spitz.

Müller, C. (2004) Forms and uses of the palm up open hand: A case of a gesture family? In C. Müller and R. Posner (eds) *The semantics and pragmatics of everyday gestures. The Berlin Conference* 233–256. Berlin: Weidler Verlag.

Müller, C. (2008) *Metaphors. Dead and alive, sleeping and waking. A cognitive approach to metaphors in language use*. Chicago: Chicago University Press.

Müller, C. and Cienki, A. (2009) Metaphor, gestures, and beyond: Forms of multimodal metaphor in the use of spoken language. In C. Forceville and E. Urios-Aparisi (eds) *Multimodal metaphor* 293–321. Berlin/New York: Mouton de Gruyter.

Núñez, R. E. (2004) Language, thought, and gesture: The embodied cognitive foundations of mathematics. In F. Iida, R. Pfeifer, L. Steels and Y. Kuniyoshi (eds) *Embodied artificial intelligence* 54–73. Berlin: Springer-Verlag.

Núñez, R. E. and Sweetser, E. E. (2006) Looking ahead to the past: Converging evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*.

Ortony, A. (1993) *Metaphor and thought*. 2nd edn. Cambridge: Cambridge University Press.

Özyürek, A. (2000) The influence of addressee location on spatial language and representational gestures of direction. In D. McNeill (ed.) *Language and gesture* 64–83. Cambridge, U.K.: Cambridge University Press.

Parrill, F. and Sweetser, E. E. (2004) What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture* 4(2): 197–219.

Peirce, C. S. (1955) Logic as semiotic: The theory of signs (1893–1920). In J. Bucher (ed.) *Philosophical writings of Peirce* 98–119. New York: Dover.

Richardson, D. C., Spivey, M. J., McRae, K. and Barsalou, L. W. (2003) Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27: 767–780.

Slobin, D. (1996) From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz and S. C. Levinson (eds) *Rethinking linguistic relativity* 97–114. Cambridge: Cambridge University Press.

Slobin, D. (2000) Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier and R. Dirven (eds) *Evidence for linguistic relativity*. Amsterdam/New York: John Benjamins.

Smith, N. (2003) Gesture and beyond. Unpublished Undergraduate Honors Thesis, Program in Cognitive Science, University of California at Berkeley.

Sweetser, E. E. (1990) *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure.* Cambridge: Cambridge University Press.

Sweetser, E. E. (1992) English metaphors for language: Motivations, conventions, and creativity. *Poetics Today* 13(4): 705–724.

Sweetser, E. E. (1998) Regular metaphoricity in gesture: Bodily-based models of speech interaction. *Actes du 16e Congrès International des Linguistes* (CD-ROM), Elsevier.

Sweetser, E. E. (2007) Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson and M. Spivey (eds) *Methods in Cognitive Linguistics* 202–224. Amsterdam/New York: John Benjamins.

Streeck, J. (2002) A body and its gestures. *Gesture* 2(1): 19–44.

Talmy, L. (1983) How language structures space. In H. L. Pick and L. P. Acredolo (eds) *Spatial orientation: Theory, research, and application* 225–282. New York: Plenum Press.

Talmy, L. (1988) Force dynamics in language and cognition. *Cognitive Science* 12: 49–100.

Taub, S. (2001) *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.

Waugh, L. R. (1982) Marked and unmarked – a choice between unequals in semiotic structure. *Semiotica* 39(3/4): 299–318.

Webb, R. (1996) Linguistic features of metaphoric gestures. Unpublished Doctoral Dissertation, University of Rochester.
Wilcox, P. P. (2000) *Metaphor in American Sign Language*. Washington, DC: Gallaudet University Press.
Wilcox, S. and Morford, J. (2007) Empirical methods in signed language research. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson and M. J. Spivey (eds) *Methods in cognitive linguistics* 171–200. Amsterdam/Philadelphia: John Benjamins.
Williams, W. (2004) Making meaning from a clock: Material artifacts and conceptual blending in time-telling instruction. Unpublished PhD Dissertation, University of California at San Diego.
Zlatev, J. (2005) What's in a schema? Bodily mimesis and the grounding of language. In B. Hampe (ed.) in cooperation with J. Grady, *From perception to meaning: Image schemas in cognitive linguistics* 313–342. Berlin/New York: Mouton de Gruyter.

Part VII

Motion

Final proofs with index 24 February 2010

15 Translocation, language and the categorization of experience

Jordan Zlatev, Johan Blomberg and Caroline David

1 Introduction

The phenomenon of *motion* is prevalent in experience: the rising and falling of our chests in breathing, the tapping of our feet against the floor, the flying of birds, the ripples of water in the brook. *Panta rei*. But all instances of (perceived) motion are not of the same kind. In the case of the rising chest, the tapping foot and the rippling water we do *not* experience any change of location of the moving object. On the other hand, in following by gaze the flight of birds, or perhaps a boat floating down the river, we do experience such a change of location. At the same time, there is a difference in the latter two cases: birds fly through perceived self-motion, while the boat is being moved by the flow of the river, or possibly by people rowing it.

The goal of this chapter is twofold. The first is to provide an experientially-based classification of perceived *motion situations*. We believe that the one we offer in Section 3 is more systematic than the various distinctions made in the current literature on 'motion events' (e.g. Talmy 2000, Slobin 2003, Pourcel 2005, cf. Section 2). Notice also that by emphasizing *experience*, rather than the objective *fact* of motion, we adopt a phenomenological perspective situating motion in the *lifeworld* of the human subject (Husserl 1999 [1907]), rather than in 'objective reality'. This is consistent with the assumption, often emphasized by cognitive linguists nowadays (e.g. Lakoff 1987), but with roots in antiquity (cf. Itkonen 1991), that language refers to and classifies not reality in itself – but *reality as conceived by human beings*. This brings us naturally to the second goal of the chapter: to use the proposed taxonomy of motion situations in addressing the questions of how different languages express motion, and if linguistic differences imply differences in conceptualization. Such (neo-)Whorfian questions have been explored extensively in the literature in recent years (see Pourcel 2005 and Section 4 below for a review), but unless we can define the classes of motion experiences *independently* of language, we are left without a compass in addressing the issues of linguistic relativity. Indeed, one finds an acknowledgment of the need for a language-independent characterization of experience in the writings of the father of the 'principle of linguistic relativity' himself, Benjamin Lee Whorf:

> To compare ways in which different languages differently 'segment' the same situation of experience, it is desirable to analyze or 'segment' the experience first in a way independent of any language or linguistic stock, a way which will be same for all observers. (Whorf 1956: 162)

After reviewing some of the neo-Whorfian research on motion in Section 4, we ask in Section 5 whether the different ways in which French, Swedish and Thai speakers express motion situations imply conceptual and experiential differences in tasks involving the categorization of translocation. Describing a series of experimental studies using the *Event Triads* elicitation tool (Bohnemeyer, Eisenbeiss and Naranhimsan 2001), and an extension of it (Blomberg 2006, 2007) we show that the answer to this question appears to be not unambiguous. To anticipate, our empirical findings suggest that the categorization of motion situations can be either more *direct* – and thus relatively unaffected by language – or more *mediated* (Vygotsky 1978), and that language can play a considerable role at least in the second case. As we discuss in Section 6, the change of emphasis from linguistic relativity to linguistic mediation can help interpret not only our own results, but also some of the contradictory findings reported in the recent literature.

2 Motion and 'motion-event typology'

If an essential aspect of motion is the perception of physical instability (Durst-Andersen 1992: 53) then what exactly is a 'motion event', given that this has been the dominant term in the relevant literature during the past decades? Talmy offers the following answer: 'A Motion event […] is a situation containing *motion* or the continuation of *stationary location*.' (Talmy 2000: 162, our emphasis). But whatever advantages this may have in terms of capturing commonalities across static and dynamic locative predication, it is much too general for our purposes by glossing over the major experiential division: spatial change vs. stasis.

Talmy (1985, 2000) considers the 'presence of motion', or *motion* with a small letter, along with the conceptual components *figure, ground, path* and *manner/cause* to be building blocks of a 'motion event', and depending on the way they are mapped to different constituents in the clause, formulates the basis for his well-known motion-event typology, shown schematically in Figure 1, with example sentences from English (a satellite-framed, or S-language) and French (a verb-framed, or V-language). This typology has been claimed to be exhaustive, i.e. that every one of the world's languages can be categorized as being, predominantly, an S- or a V-language.

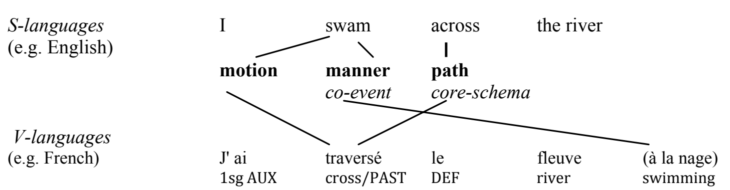


Figure 1. Different mapping patterns between the conceptual components of motion events and parts-of-speech in satellite-framed (S) languages and verb-framed (V) languages

However, it has become increasingly clear that this binary typology cannot do justice to the complexity found in the world's languages: either more 'exotic' ones such as Tzeltal (cf. Brown 2004), or more familiar ones such as Russian (cf. Smith 2003), as many of the contributions to the volume edited by Strömqvist and Verhoeven (2004), e.g. Slobin (2004) testify. In some of our own work (Zlatev and David 2003, Zlatev and Yangklang 2004), we have documented how Thai, and by extension other similar serial verb languages, constitute a distinct 'third' type (which Slobin 2004 has generalized as the 'equipollently-framed type', together with other languages which permit the easy encoding of both Manner and Path in the same clause). For example, Thai resembles V-languages in some respects (e.g. path expression by a main verb), S-languages in other respects (e.g. manner expression by a main verb), while in yet other respects it resembles neither (e.g. by having a separate 'slot' in the serial verb construction for path+manner conflating verbs), cf. Zlatev and David (2003) for discussion.

But perhaps more troublesome for the Talmian typology than the empirical problems are certain unresolved conceptual and definitional issues, such as the following:

- What exactly is 'path'? The extended trajectory traversed by the moving entity, or some sort of schematic representation of this, e.g. as in the model of Regier (1996), related to the beginning, middle and/or end of the motion trajectory?[1] And how does this relate to the concept of *direction* of motion, expressed in e.g. *up*?
- What exactly is 'manner' (of motion)? Does this include information pertaining to the vehicle of motion (e.g. fly vs. ride), the speed (e.g. stroll vs. run), the body parts (e.g. hop vs. climb), the medium (sink vs. fall) or all of these?
- Why is path regarded as the 'core schema', and is this so for all languages and for all types of motion (for this and the following point, see the discussion below)?
- What is a 'co-event'? Is it really an *event* and does it always pertain to information related to the 'manner' or 'cause' of motion?
- What exactly is a 'satellite'? Talmy (2000: 102) defines it as a constituent standing in a 'sister relation to the verb root', but it is, for example, unclear if Swedish verbal particles (e.g. *gå* <u>*in*</u>) can be thus grouped with Bulgarian verb-prefixes (e.g. <u>*v*</u>*-liza*): while both examples correspond to English 'go in', and the 'satellite' carries the meaning INTERIOR, the Bulgarian stem does not exist as an independent verb.

The basic, and yet unresolved, question however, remains 'What is motion?' and correspondingly: 'What is a motion event?' Prior to a clear answer to these questions, it is not certain that we are comparing equivalent semantic structures across languages. Talmy is clearly aware that his initial definition of a 'Motion event' needs further specification, since he repeatedly points out the difference between *translational motion*: 'an object's basic location shifts from one point to another in space' and *self-contained motion*, where 'an object keeps its basic or 'average' location' (Talmy 2000: 35) and emphasizes that the typology concerns motion only of the first kind. However, it is not altogether clear what this distinction amounts to and what is meant by 'basic location'. As examples (1) show, it is not possible to decide on the basis of the semantics of the verb alone what type of motion is involved: in (1a) John's motion is clearly 'self-contained' while in (1c) John's location has 'shifted' from outside to inside the room. But what about (1b): is the motion involved considerable enough to be 'translational'?

(1) (a) *John ran on the treadmill.*
(b) *John ran in the park.*
(c) *John ran into the room.*

In a recent monograph, Pourcel (2005) endeavors to clarify these issues through an 'alternative model', that is claimed to be based on *conceptual analysis*, rather than semantic analysis, as is the case with Talmy, or discourse analysis as done by Slobin (e.g. 1996, 1997, 2003). The core of Pourcel's proposal seems to be to distinguish between *motion events* and *motion activities*, illustrating these with examples (2) and (3) – with identical numbers in (Pourcel 2005: 153–154):

(2) *The dog ran out of the barn across the field to the house.*

(3) *The dog is running around the house.*

On this basis, it as argued that:

> [t]here is therefore a distinction between motion that is source-and-goal-oriented, as in (2), and motion that is not, as in (3). Conceptually, it is relevant to distinguish between motion event and motion activity as the conceptual emphasis of an event consists of the PATH of motion…; whereas the conceptual emphasis of an activity consists of the MANNER of motion, which specifies a motion in progress, e.g. (3). In other words, the core schema of activity is no longer PATH, but MANNER. (Pourcel 2005: 154)

In general, this proposal is quite reasonable. But if indeed the 'core schema' in activity representations is Manner rather than Path, this goes clearly against Talmy's terminology, where Path is always the core schema, irrespective of language and construction type, which brings us back to one of the conceptual/definitional problems listed earlier. Still more troublesome is that Pourcel (2005) does not provide any clear *conceptual* criterion for what distinguishes 'events' from 'activities' that would explain the corresponding focus on Path vs. Manner. The qualification 'specifies a motion in progress' for activities can hardly be correct since it is based on the progressive aspect marking of (3), while (1a) and (arguably) (1b) are representations of 'activities', even though they are not presented as being 'in progress'.

Furthermore, the concept of 'motion event' is extended by Pourcel (2005) to involve not only 'telic paths', such as those on (2), but 'atelic' or 'locative' paths, 'e.g. DOWN, ALONG, AROUND' (Pourcel 2005: 154), illustrated in the English example (4) and the French examples (5) and (6):[2]

(4) *The dog ran up the street.*

(5) *Marc monte les escaliers sur la pointe des pieds.*
Marc goes up the stairs on tiptoes.

(6) *Marc longe les bords de la rivièr.*
Marc goes along the river bank.

What are the grounds for grouping these as examples of 'events' along with (2) rather than as activities along with (3), which, note, even includes the so-called AROUND path? We believe that the reasons are twofold. First, a language-independent conceptual analysis is not provided, but rather one which is influenced by the 'grammatical features of motion event encoding in French' such as that 'PATH information is obligatory … in the main verb' (Pourcel 2005: 180), along with *a priori* classification of verbs such as *monter* and *longer* as PATH verbs (albeit 'atelic'). The second is that, as mentioned earlier, Pourcel (2005) seems to conflate lexical (i.e. *Aktionsarten*) and morphological (i.e. grammatical aspect) representations of the event/activity distinction – for example in referring to 'the variable use of the tenses, e.g. the imperfect or present tense for activities, and the past perfect or simple past … for *completed motion events*' (Pourcel 2005: 181, our emphasis). In the next section, we will propose our own conceptual analysis of *motion situations* – a term used occasionally by Pourcel (2005: 186) as well, as a hyperonym for motion events and activities – which we believe does not suffer from these problems. At the same time, we wish to express our indebtedness to Pourcel (2005) for helping bring together the 'motion' and the 'situation type' literatures, something which has been long overdue.

3 A taxonomy of motion situations

From the perspective of the analysis of (the invariants of) experience – *phenomenology* (cf. Husserl 1999 [1907]), motion as such can be defined as the experience of *continuous change in the relative position of an object (the figure) against a background*, in contrast to stasis – where there is no such change – and in contrast to a *dis-continuous* change, as when a light suddenly lights up in position A, 'disappears' and then appears in position B. As well-known, however, if the time fragment between the two discrete events is small enough then an observer will actually see the light as *moving* from A to B, in a continuous manner. Thus, motion is 'in the eyes of the beholder'. Note that 'continuous' is here meant to exclude from the definition of motion such events as disappearing at one place, and reappearing at another, as in a Star Trek case of 'teleportation', which may be in the sphere of the imaginable, but not in the ordinary human lifeworld. It does *not* exclude instances of rather abrupt types of motion, e.g. jumping, blinking, breaking or other similar 'punctual' events.

Furthermore, note that motion 'from A to B', i.e. *relocation* (Smith 2003) is not a necessary characteristic of a motion situation. First, the light could waver around A, and then there would be no change in its average position and thus there would be 'self-contained' motion in Talmy's terms. Second, the figure could be moving along a vector in an open-ended way, for all eternity perhaps – and hence there need not be any B to relocate to. Third, the figure's motion can be either spontaneous or caused by an external source. Thus, we have *three different parameters* according to which motion situations can vary, quite independent of their representation in language. These are described in the rest of this section, concluding with a summary presentation of the taxonomy, and its (perceived) advantages compared with those of Talmy or Pourcel.

3.1 Translocative vs. Non-translocative motion

We define *translocation*, which is similar to but more transparent than Talmy's term 'translational motion' (cf. Zlatev and Yangklang 2004) as *the continuous change of an object's average position according to a spatial frame of reference*. As can be seen from this definition, this is a special kind of motion, which unlike motion in general requires a spatial frame of reference (FoR):

> In the most general sense, a FoR defines one or more reference points, and possibly also a coordinate system of axes and angles. Depending on the types of the reference points and coordinates different types of FoR can be defined. (Zlatev 2007: 328)

An influential treatment of the concept FoR, especially within linguistic typology, is that of Levinson (1996, 2003), who distinguishes between *relative, absolute* and *intrinsic* FoRs. However, this distinction is only based on horizontal static relations, whereas Zlatev (2005, 2007) extends and generalizes it to involve dynamic relations, i.e. motion, as well as the vertical plane. The first type can be called *Viewpoint-centered*, which when expressed in language involves the perspective of the speaker or hearer as a reference point, as in examples (7–8).

(7) *I turned and went to the right.* **FoR**: Viewpoint-centered, Speaker

(8) *Turn and go to the/your right.* **FoR**: Viewpoint-centered, Hearer

The second type is *Geocentric*, involving the horizontal or vertical plan while relying on geo-cardinal positions as reference points, as in (9–10).

(9) *I drove West.* **FoR**: Geocentric, Horizontal

(10) *The balloon went up.* **FoR**: Geocentric, Vertical

Finally, there is the *Object-centered* FoR, which can involve the position of either the focused (and possibly moving) object, the *figure*, or that of an external object, a *landmark*, as in (11–12).[3]

(11) *I went forward.* **FoR**: Object-centered, Figure

(12) *I went to the church.* **FoR**: Object-centered, Landmark

A particular case of translocation can thus be specified according to one or more of these frames of reference, which provide the reference points allowing us (a) to judge that the object/figure has indeed changed its average position and (b) to determine its Path or Direction (see below). Similarly, in order to state that there is no change in the *average* position of a moving figure, i.e. non-translocative motion, a FoR needs to be (at least) presupposed. John's running in example (1b) is non-translocative with respect to an Object-centered FoR with the park as a whole as Landmark. But the same state-of-affairs can be construed as translocative if we, for example, adopt some more specific reference point, e.g. the viewpoint of an observer situated within the park.

On this basis, example (1c) can be classified as an expression of translocative motion, while (1a) and (1b) represent non-translocative motion. The FoRs in all three cases are object-centred, anchored in, respectively, the referents of 'the room', 'the treadmill', and 'the park'. Note how essential the choice of a particular FoR is in order to determine the type of motion. If the same external state-of-affairs described in (1b) was portrayed as (13), then the (conceptualized) situation would be translocative, involving the change of the figure's position with respect to the 'end of the park'.

(13) *John ran to the end of the park and back.*

Analogously, the same state-of-affairs can be experienced – and described – quite differently, depending on the Frame of reference, as in the examples below.

(14) *He is going to the top of the hill.* Object-centered, Landmark

(15) *He is going forward.* Object-centered, Figure

(16) *He is going uphill.* Geocentric

(17) *He is going that way.* Viewpoint-centered

While all four examples involve translocation, (15–17) do not specify the change of position in relation to a beginning (Source), middle (Via) or end (Goal) point, but rather with respect to the figure's initial position in (15), geo-centric coordinates in (16) or a deictic center in (17). Thus following the analysis presented in earlier work (Zlatev 2003, 2005), we state that of these examples only (14) involves the category Path, understood in the schematic sense (cf. footnote 1), while (15–17) express the related but different category Direction. In the case of non-translocative motion there is neither Path nor Direction, since there is no change in the figure's average position. The crucial difference is that Path implies *bounded motion*, whereas Direction implies *unbounded motion*, which brings us to the next parameter.

3.2 Bounded vs. unbounded motion

The *boundedness* of a process undergone by X implies that it will inevitably (not just possibly or probably) lead to *X undergoing a state-transition* (cf. Vendler 1967). This means that in expressions of bounded motion, X (the figure) will depart from Source, or pass through a mid-point (Via), or reach a Goal (as in 12–14) – or all three as in (2). In *unbounded motion*, nothing of the sort is implied, and in principle – though not practically – the motion can go on indefinitely, as in the situations described in examples (7–11). As pointed out above, bounded translocative motion always involves the category Path, with one or more reference points being defined through the object-centred, landmark-defined FoR. In the case of unbounded translocative motion, we have rather the category Direction, specified either as a *vector* according to one of the other FoR conditions, or as a *trajectory*, that can take particular shapes such as AROUND or ALONG, as in (3) and (4).

Note furthermore, that there is independence between the two parameters discussed so far. We have seen how translocative situations can be either unbounded, e.g. (7–11) or bounded e.g. (12) and (13). Non-translocative situations can be similarly either unbounded, as (1a) and (1b), or bounded – if the motion involved leads to a state-transition, as in (18) or the Swedish equivalent (19).

(18) *The vase broke (in pieces).*

(19)	*Vas-en*	*gick*	*sönder.*
	vase-DEF	go.PAST	broken

One might counter that (18) and (19) do not express, but rather presuppose motion, but since the 'breaking' of the vase will typically involve a perception of physical change (against a stable background) we consider these sentences representations of non-translocative bounded motion.

3.3 Self-motion vs. caused motion

The final parameter concerns whether the figure is perceived to be moving under the influence of an external cause or not. As previously stated, the relevant notion of causality concerns the (naïve) human lifeworld, and not our scientific understanding of the universe. Thus, the situation described in (20) above is one of 'self-motion' even though the motion of the raindrops is caused by gravity, objectively speaking. On the other hand, (21) clearly represents a (translocative, bounded) caused motion situation.

(20) *Raindrops are falling on my head.*

(21) *John kicked the ball over the fence.*

This parameter is likewise independent of the other two, so it is possible to have, e.g. caused translocative, unbounded motion situations (22), caused non-translocative bounded ones (23), and caused non-translocative unbounded ones (24). The self-caused correspondences to these have already been illustrated.

(22) *He pushed the car forward.*

(23) *He tore the paper up.*

(24) *She waved the flag.*

3.4 Summary

The independence of the three parameters yields the 8 types of motions situations illustrated in Table 1, with schematic representations in English.

Table 1. Illustration of the expression of 8 motion situation types in English; F = Figure, LM = Landmark, A = Agent, View-C = Viewpoint centred, Geo-C = Geocentric, Obj-C = Object centred Frame of Reference

	-CAUSED	+CAUSED
+TRANSLOCATIVE +BOUNDED	F goes to LM	A throws F into LM
+TRANSLOCATIVE -BOUNDED	F goes away (View-C) F goes up (Geo-C) F rolls forward (Obj-C)	A takes F away (View-C) A pushes F upward (Geo-C) A pushes F forward (Obj-C)
-TRANSLOCATIVE +BOUNDED	F breaks (up/down)	A breaks F (up/down)
-TRANSLOCATIVE -BOUNDED	F waves	A waves F

The tense in the examples in Table 1, the present simple, is only seldom used with any of these situation types (constructions) in English, and if so to express habitual meanings, as in (25).

(25) *Marry goes to school at 8 o'clock in the morning.*

However, it was intentionally used in the cells in Table 1 in order to highlight the fact that the different situation types (i.e. specifying the values of the three parameters) can be expressed through: (a) the lexical semantics of the verb, (b) verb-satellite (particles or affixes), (c) adpositional phrases and (d) the grammatical construction (e.g. intransitive vs. transitive). While tense and aspect markers can make the distinction between e.g. bounded and unbounded situations even clearer, i.e. by rendering the bounded ones in past simple as in (21), and the unbounded ones in present continuous as in (15–17), this is not *necessary* for making the parameter differentiations, at least for English. In fact, we broadly agree with Durst-Andersen (1992) that morphological aspect introduces *an extra dimension of meaning* over and above those expressed by (a)–(d), by allowing the profiling of situations either as *ongoing* or as *completed* – whether they are inherently bounded or not. Thus, (20) is no less a representation of a bounded situation (despite ongoing), and (22) no less a representation of an unbounded one (despite being 'in the past' and thus completed). The following three 'authentic' examples, taken from the British National Corpus (http://www.natcorp.ox.ac.uk/), show how *fall* in the past tense can be used to express unbounded translocation, despite the fact that the events are being represented as taking place in the past, and thus as 'completed'. Grammatical tense-aspect should therefore be distinguished from motion situation types, and their linguistic expression, *pace* Pourcel (2005).

(26) *She called to Hermione and Joanna and all the girls who had gone already along the paths she had rejected, called to them to wait for her and place their steady walking boots on solid earth to catch her. And still she fell and fell.*

(27) *The wind blew and the snow fell, but it didn't matter.*

(28) *… the devaluation of stock as component prices fell.*

The conceptual framework described in this section and in particular the contrast between bounded and unbounded translocative situations is highly relevant for our empirical studies involving language and translocation described in Section 5. But prior to describing these, let us first take stock.

We claim that our proposed taxonomy clarifies some of the problematic issues described earlier. First of all, we believe that we have introduced definitions of (perceived) motion in general, and specific types of motion situations that are more consistent than those used in (much of) the 'motion events' literature. Second, we consider our taxonomy to be, if not exhaustive, at least better equipped than alternatives to serve as a basis for typological investigations in the 'domain' of motion. It allows us to analyse e.g. cases such as those that were discussed in Section 2 in an unambiguous way. Thus, examples (4)–(6) can be classified as expressions of *translocative unbounded* motion situations, together with (3), while (1c) and (2) are representations of *translocative bounded* ones. On the other hand, (1a) and (1b) are neither, but rather expressions of *non-translocative unbounded* motion. This is summarized in Table 2. As pointed out in Section 2, examples such as these have been grouped and termed in various ways in the past.

Table 2. A classification of the examples discussed in Section 2, on the basis of the presented taxonomy of motion situations

Examples from Section 2	Motion situation type
(3) *The dog is running around the house.* (4) *The dog ran up the street.* (5) *Marc monte les escaliers sur la pointe des pieds.* (6) *Marc longe les bords de la rivière.*	**+ translocative** **- bounded**
(1c) *John ran into the room.* (2) *The dog ran out of the barn across the field to the house.*	**+ translocative** **+ bounded**
(1a) *John ran on the treadmill.* (1b) *John ran in the park.*	**- translocative** **- bounded**

Third, we have defined Path as always related to Source, Via or Goal (on the basis of an Object-centered, Landmark-defined FoR), while unbounded translocative situations involve Direction, and non-translocative situations involve Location. In this way, we have sharpened the conceptual apparatus used in the field. One thing this allows us is to reformulate the famous *boundary-crossing constraint* (Slobin and Hoiting 1994), stating that a motion verb expressing manner may not be used if there is a crossing of a boundary, as follows: *a Manner-verb can co-occur with an expression of Direction or Location, but not with Path in the same clause*. Assuming that French, as most V-languages, generally obeys this constraint, examples (29–32), where the first two are from Pourcel (2005: 40a-41a), and the latter two from Zlatev and David (2003: 40b-c) can be straightforwardly explained as follows.

(29) *Nous* *avons marché* *le long* *de la plage.*
We walked along the beach
MANNER DIR
'We walked along the beach.'

(30) *Nous* *avons marché* *dans* *la pièce.*
We walked in the room
MANNER LOC/*PATH
'We walked inside the room.'
* 'We walked into the room.'

(31) **Il* *a couru* *en entrant* *dans* *la maison.*
3sg+MASC run+PAST entering in DEF house
MANNER PATH LOC
'He ran entering the house.'

(32) *Il* *a couru* *pour* *entrer* *dans* *la maison.*
3sg+MASC run+PAST to enter in DEF house
MANNER PATH
'He ran in order to enter the house.'

Example (29) does not violate the constraint, since it includes a combination of Manner and Direction within the clause. In (30), only a non-translocative interpretation of walking about 'inside' the house is possible. Such an interpretation is excluded in (31) due to the participle *en entrant*, expressing Path, and the result is ungrammaticality (uncorrectness), due to semantic factors. Finally, (32) is in contrast a correct French sentence, since the Manner and Path expressions are in seperate clauses.

The reformulated boundary crossing constraint will play a role in the interpretation of the results from our experiments, described in Section 5. But prior to that, we briefly review some of the recent research on how different languages can possibly affect the experience of motion in a way that 'colours' it accordingly.

4 Neo-Whorfian research on the categorization of translocation

If Talmy made 'motion events', or as we prefer – translocative situations – into a popular subject for typology, it was Slobin (1996) who brought the subject to the attention of neo-Whorfian research on linguistic relativity. According to one of Slobin's formulations, it may even be a mistake to look for language-independent taxonomies of situations such as that presented in the previous section, since:

> The world does not present 'events' and 'situations' to be encoded in language. Rather, experiences are filtered through language into verbalized events. A 'verbalized event' is constructed online, in the process of speaking. (Slobin 1996: 75)

But at the same time, Slobin's famous 'dynamic' formulation of the Whorfian program, known as *thinking for speaking*, only concerns the 'special kind of thinking […] that is carried out, on-line, in the process of speaking' (Slobin 1996: 75) and is therefore different from Whorf's (1956) notion of 'habitual thought', according to which language should have much more pervasive effects (cf. Blomberg 2007). Methodologically, Slobin (1996, 1997, 2003) concentrated on differences in the 'rhetorical style' of speakers of V-languages such as Spanish and S-languages such as English – as something that could be explained by the languages' different ways of expressing, above all, the concepts Path and Manner. For example, due to the optional expression of Manner in V-languages (see Figure 1), their speakers were found to express Manner less often and preferred to give more static descriptions in which the Figure's motion could be inferred from the 'scene setting' and the result of the motion, while S-languages induced descriptions in which the events were presented more dynamically, with more elaborated representations of the Path. But, as pointed out by Pourcel (2005), Slobin's research gives little support for strong relativistic effects in *the categorization of experience as such*, i.e. even when 'thinking for speaking' is (apparently) not involved.

A number of other studies have attempted to demonstrate such effects using, among other methods, a classic task for studying categorization in an (apparently) non-linguistic context: *forced-choice similarity judgments*. The general method, used with various modifications, in all of these studies is to use *triads* of representations of motion situations: a target situation is presented along with two alternatives, where one differs from the target with respect to Path and the other with respect to Manner, and the subject is asked which of the two 'is most similar' to the target. The general reasoning is that if language impinges on categorization, then speakers of a V-language should be predisposed to prefer 'same-Path' rather than 'same-Manner' to a greater extent than speakers of S-languages, where both components are expressed equally easy (see Section 2). An exception to this line of reasoning was offered by Papafragou, Masely and Gleitman (2002), who suggested an alternative basis for a linguistic effect that actually runs in the opposite direction: since Manner is often expressed in a non-obligatory constituent in a V-language, when it is expressed, it would be 'foregrounded' and thus achieve more *semantic salience* (Talmy 1985) than in an S-language where it is expressed by an obligatory constituent, such as the main verb. Papafragou, Masely and Gleitman (2002) compared among other things the categorization of triads (using static pictures) by speakers of Greek (assumed to be a V-language) and English (an S-language) and despite differences in the linguistic descriptions that followed the predicted patterns (along the lines of Slobin's research), they found no bias for either Path or Manner-based judgments in either group, and thus argued against the presence of any Whorfian effect on motion event categorization.

However, other studies applying the same method, but using triads of dynamic (video-clip) representations have given different results. Finkbeiner, Greth, Nicol and Nakamura (2002) compared English (S-language) with Spanish and Japanese (V-languages) speakers' performance, and found a considerably stronger preference for Manner-based similarity in the English group, and thus support for a degree of linguistic relativity. Importantly, this effect was present only when the target clip was presented first, and the alternatives (in parallel) afterwards. When the three clips were presented simultaneously, the Manner-bias for the English group disappeared, leading the authors to conclude that 'the apparently nonlinguistic task used in Experiment 1 actually encouraged the participants to encode the scenes linguistically' (Finkbeiner et al: 454).

Gennari, Sloman, Malt and Fitch (2002) compared speakers of the two prototypical languages for Talmy's two types, English and Spanish, and established no clear difference between the groups when the represented situations were not described prior to the similarity judgments. But when they asked the subjects to provide such a description in their native tongues prior to their choice, a stronger preference for Path in the Spanish group was observed. This could be taken as offering support for a version of Slobin's thinking-for-speaking.

Pourcel (2005) reports evidence for an effect of language-type in a memory-based study, but in her categorization study with 15 triads in the form of video-clips representing people involved in various motion situations, she failed to find any difference between English and French speakers. Both without and with prior linguistic description there was a preference for same-Path categorization for both language groups. An interesting finding, however, was that two types of motion situations, corresponding to our distinction between bounded and unbounded translocation described in Section 3 gave different results: there was a strong Path bias for bounded motion ('telic Path'), but this bias was neutralized, and with linguistic description even replaced with a Manner-bias for the unbounded motion situations ('atelic Path') (cf. Pourcel 2005: 243–245). Finally, an important difference compared to the study of Finkbeiner et al. (2002) was that all three video-clips in each triad were presented *sequentially* (in different orders).

Bohnemayer, Eisenbeiss and Narasimhan (ms), conducted the most extensive study of this type, in the sense that they contrasted not just two or three languages, but 17 typologically, areally and genetically diverse languages, including Polish (S-framed with verb-prefixes), German (S-framed with verb-particles), Japanese (V-framed) and Lao (serial-verb, 'third type'). The stimuli used by Bohnemayer, Eisenbeiss and Narasimhan (ms) are identical to those used in our studies described in Section 5, where we describe them in more detail, but suffice it for now to point out that they involve an animated, smiling tomato-like figure which 'jumps', 'rolls', 'spins' or 'slides' either up/down a ramp, or left/right across a field, either with or without crossing the boundaries of the Ground objects. While Pourcel (2005) criticizes the animated 'unnatural' character of the protagonist, and the fact that it allows a limited scope of Manners of motion, we would argue that this design – similar to that of Finkbeiner et al. (2002) – has a considerable advantage: it contrasts Manner and Path (in some cases: Direction) completely systematically, so that the two choice situations are identical with the target in each triad, apart from the manipulated variable. Furthermore, given that even illiterate speakers of languages such as Jukatek living in traditional societies did not have difficulties interpreting the situations with the 'animate tomato' capable of self-motion suggest that it was not so 'unnatural'.[4]

The foremost strength of the study of Bohnemeyer et al. (ms), however, is the large number and variety of the languages involved. Accordingly, the results showed a wide variation in the produced biases in the similarity judgment task: from 85% same-Manner for the Polish group to 43% same-Manner for the Jalonke and Jukatek groups, but no general pattern for speakers of S-languages preferring Manner more than those of V-languages. This rather convincingly shows that the binary 'motion-event typology' of Talmy is not *sufficient* to predict categorization preferences (though it may be one of the factors that play a significant role) and a better conceptual and methodological basis is necessary in matching motion (i.e. translocation) typology and linguistic relativity. Interestingly, Bohnemeyer et al. (ms) also established a language-general difference between the representations of situations in which the figure moved up or down (diagonally) on a ramp from the cases when in moved either from-to, or out of-into a landmark: in the latter case the subjects were more likely to base their similarity judgment on the basis of Manner than in the first. The authors attempt to explain this in terms of the greater 'simplicity' of the ramp scenes, involving one reference object (the ramp), rather than two.

But another explanation is possible: in the case of the 'ramp' scenario, the situation was at least ambiguous between unbounded translocation (moving upward or downward) and bounded translocation (moving to 'the top' or 'the bottom' of the ramp). On the other hand, the other two types of situations involved unambiguously bounded translocative events, with or without boundary crossing. Thus, as in the study of Pourcel (2005) the similarity judgments for the bounded and unbounded translocative situations differed, implying the cognitive relevance of this distinction. However, the biases in the two studies were *converse*: stronger preferences for same-Path categorization for unbounded than bounded situations in the Bohnemeyer et al. (ms) study and stronger preferences for same-Path categorization for bounded than unbounded motion in that of Pourcel (2005). The conclusion is therefore that this factor must interact with other 'variables' such as the nature of the stimuli (animated vs. non-animate) and/or the nature of the presentation of the alternatives (sequential vs. parallel). It is possible furthermore that these factors affect the degree to which language influences the categorization process.

In sum, the studies of the categorization of motion (translocation) situations by speakers of different languages over the past few years have yielded different and somewhat contradictory results. What has become clear though is that:

a) the nature of the stimuli – static vs. motion pictures, animated vs. 'real life' video-clips, sequential vs. parallel presentation – influences the similarity judgments;

b) different types of motion situations can yield different categorization preferences;

c) the role of linguistic description, especially prior to making the similarity judgment, needs to be more carefully explored;

d) more languages than simply two representatives of the binary typology need to be taken into consideration.

Our empirical studies using the Event Triads tool of Bohnemeyer et al. (ms) (Section 5.1 and 5.2) and a modification of it (Section 5.3) with speakers of Swedish, French and Thai address the latter three points. In Section 6, we will offer an interpretation of the apparently contradictory results, suggesting a coherent explanation.

5 Three empirical studies with event triads

5.1 Study 1

In our initial study we used the original *Event Triads* elicitation tool, developed at the Max Planck Institute for Psycholinguistics, Nijmegen (Bohnemeyer, Eisenbeiss and Narasimhan 2001), which was created to investigate biases for Path or Manner in forced-choice similarity judgments. First a 5-second long animated film of the moving tomato-like figure is shown on the whole computer screen, and after one second two clips – identical to the first but differing with respect to either Path/Direction or Manner

– are shown in smaller windows in parallel (see Figure 2). The tool includes 72 such different triads, 'distributed across 6 randomized presentation lists in a Latin-square design' (Bohnemeyer, Eisenbeiss and Narasimhan, ms), where each list was presented to two participants, in reverse order.

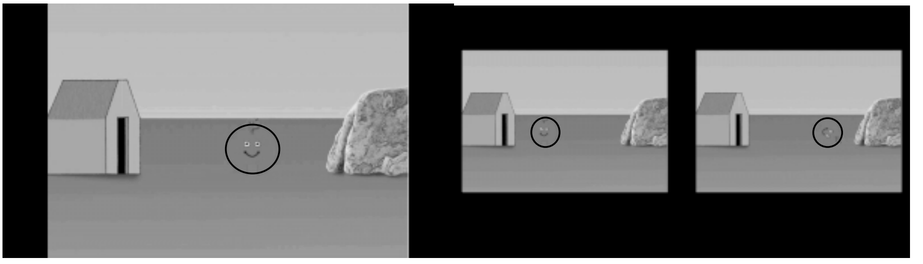


Figure 2. An example triad from the stimulus tool Event Triads. The black outline of the tomato-figure is added, so that it would be more clearly visible when viewed in a black and white printout. In the elicitation tool the red color of the tomato contrasts clearly with that (green or white) of the background and no such outlining is necessary

Thus, the Event Triads tool requires 12 participants for varying the order of presentation, for counterbalancing the left/right position of the Manner-similar and Path/Direction-similar smaller films in the second segment of the triad, and for trying all possible combinations of Path/Direction and Manner. Following three practice trials, each participant was given 50 triads. Of these, only 12 contrasted Path and Manner, while the other 38 were distracters in which the figure stops at mid-scene, or involve differences in color, or completely different situations such as one figure throwing an object to another. The 12 crucial trials can be divided in 3 groups, depending on the type of motion situation represented in the first segment (large window in Figure 2), using the terminology introduced in Section 3:

- 4 Bounded translocative situations, from landmark$_1$ to landmark$_2$ (*FROM/TO Path*)
- 4 Bounded translocative situations, out of landmark$_1$ into landmark$_2$ *(OUT/INTO Path*)
- 4 Unbounded translocative situations, up (or down) (*VERTICAL Direction*)

As pointed out, in each of these cases the second segment presents a choice between a situation in which the figure moves according to the same Path or Direction, but differs in Manner, or has the same Manner, but moves in the reverse Path or Direction. There are four different types of Manner that can be glossed in English as *jumping, rolling, spinning* or *sliding*. As mentioned in Section 4, these manners of motion are quite perceptually salient and conspicuous (especially for a 'tomato') and it was expected that there would be a relatively strong Manner bias for the similarity judgments irrespective of language. Nevertheless, one could expect this bias to be strongest (everything else being equal) for speakers of S-languages, and weaker for speakers of a V-language (i.e. relatively more

Path-based choices). As for speakers of serial-verb languages such as Thai, we expected these to show an intermediary position, given that both Manner and Path are easily codable, or alternatively equally 'backgrounded', in such a language (cf. Section 2).

Participants were 3 groups of 12 monolingual undergraduate students from Lund University (*Swedish group*), the University of Poitiers (*French group*) and Chulalongkorn University (*Thai group*). The procedure was the following: each participant was given three practice trials, followed by the 50 test triads. For the *similarity judgment task*, after every triad, the participant had to point to either the left or the right half of the second segment (cf. Figure 2) which was to serve as the answer to the question 'Which is most similar to the first film – the left or the right?' Following this and a brief pause, there was a *verbal description task*, in which the participant was asked to describe 18 video-clips of only the first fragment, representing the three kinds of translocative situations in the data: 4 Vertical, 4 FROM/TO and 10 OUT/INTO.[5] The results of the similarity judgments task were marked in a coding sheet, and the verbal description were recorded and transcribed, and both were subsequently subjected to statistical analysis.

The results for the similarity judgements are presented in Figures 3 and 4. Contrary to our expectations, it was not the Swedish, but the Thai group that had the largest proportion of same-Manner choices: the difference between the Thai group on the one side, and the French and Swedish groups on the other was statistically significant, chi2(2) = 14.415 ($p < .05$), while that between the French and the Swedish groups was not.

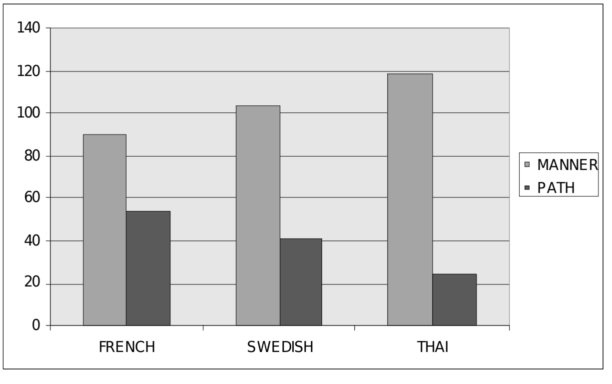


Figure 3. Distribution of Manner vs. Path/Direction biased categorization choices for the three language groups of French, Swedish and Thai. Max = 144 (12 participants * 12 choices) per language

More interesting, however, were the results when we divided the 12 test triads according to the three types listed above: FROM/TO, OUT/INTO and VERTICAL. As can be seen in Figure 4, the classification of the Vertical unbounded translocative situations for the French group differed significantly from the other two types of situation (chi2(2) =

6.933, p = 0.031), while there where no such differences for the other two languages. Given that the total number of choices of this type was 48, the French group actually displayed a weak Path bias (25 vs. 23) for this type.

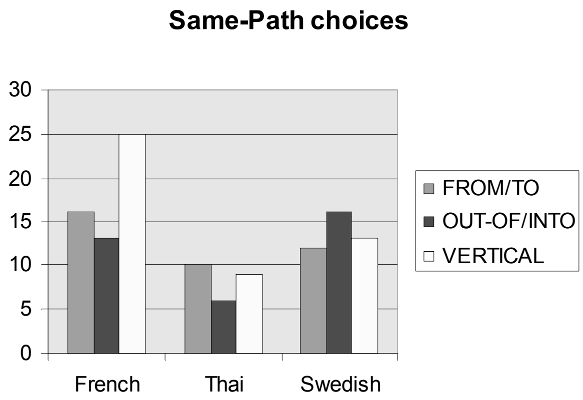


Figure 4. Same-Path based choices for the three language groups, divided by situation types: FROM/TO Path, OUT-OF/INTO Path, and VERTICAL Direction. Max = 48 (12 subjects * 4 choices) per language

To help interpret this, we analyzed the results of the linguistic description task for the French group in detail. We asked if there is a correlation between the differences in the group's similarity judgments (between the Vertical and the other two types) and *the semantic and grammatical structure of the descriptions* of the group. In analyzing the latter, we had a mini-corpus of 216 descriptions (12 participants * 18 translocative stimuli). We found indications for two such correlations. Table 3 displays all the verbs (types and tokens) in the French descriptions, divided by the categories Vertical Direction, Horizontal Path (FORM/TO + OUT/INTO), Manner and Other. The absolute number of tokens were actually mostly Manner verbs, which may appear at first hand surprising, given that French is (supposedly) a V-language, but as Pourcel (2005) and Pourcel and Kopecka (ms) show, French involves several types of constructions were Manner is expressed by the main verb (see also below). More relevant for our purposes, however, was the fact that the Direction verbs, above all *monter* and *descendre* were relatively more frequent than the Path verbs: there were only 4 stimuli (per subject) with situations that could be described with these, whereas there were 14 stimuli for the Path verbs (10 INTO and 4 TO). The ratio 8.75 vs. 5.36 in favor of Direction verbs compared to the Path verbs suggests that Direction was more readily codable than Path, and thus *possibly* also attracted relatively more attention than Path, compared to Manner in the similarity judgment task. But admittedly this is only a tentative suggestion, and it says nothing about the direction of (possible) causation involved: it is equally possible that Direction is more easily cognitively 'processable' than Path, and therefore received a higher degree of linguistic coding.

Table 3. Motion verbs produced by the French group in Study 1, in response to the linguistic task involving 4 Direction and 14 Path (4 FROM/TO and 10 OUT/INTO) stimuli

DIRECTION	PATH	MANNER	OTHER
MONTER (ascend): 18 DESCENDRE (descend): 15 GRAVIR (climb, struggle up a slope):1 DÉVALER (tumble down): 1	SORTIR (exit, go out): 23 RENTRER (enter/come back home): 23 PARTIR (leave): 14 TRAVERSER (cross): 6 PASSER (pass, go through): 2 AVANCER (move forward): 2 ARRIVER (arrive): 1 RETOURNER (go back): 2 REVENIR (come back): 2	ROULER (roll): 54 PIVOTER (pivot, revolve): 1 FAIRE DES GALIPETTES (somersault): 1 TOURNER (turn, spin): 14 FAIRE LA TOUPIE (move like a spinning top): 2 GLISSER (slide): 14 SAUTILLER (hop, skip): 11 SAUTER (jump, leap): 3 FAIRE DES BONDS (leap, spring up): 1 BONDIR (jump, bounce): 1	ALLER (go): 81 FAIRE UN DÉPLACEMENT (make a move): 1 SE DÉPLACER (move): 7 S'ARRÊTER (stop): 8
Stimuli: 4 Verb tokens: 35 Ratio: 8.75	Stimuli: 14 Verb tokens: 75 Ratio: 5.36	Stimuli: 18 Verb tokens: 102 Ratio: 5.67	Stimuli: 18 Verb tokens: 97 Ratio: 5.39

The second correlation could more easily be related to a potential linguistic effect. It turned out on analysis that in the verbalization of the bounded translocative (Path) stimuli, only 18 out of 43 Manner expressions were present in the same clause as the Path verb, while the remaining 25 (58%) occurred in an additional clause. On the other hand, in the descriptions of the unbounded translocative (Direction) stimuli, in 27 out of the 28 cases which also included an expression of Manner, the latter was expressed in the same clause, as in (33). In only one case out of 28 (3.5%) was Manner expressed in an additional clause.

(33)	*La tomate*	*monte*	*la montagne*	*en roulant*
	DEF tomato	climb	DEF mountain	rolling
		DIRECTION		MANNER

What this could be attributed to is the difficulty of encoding both Path and Manner in the same clause, as opposed to Direction and Manner, due to the boundary-crossing constraint (cf. Section 3.4) This would lead to Manner being expressed separately in the case of bounded translocation, as the main verb of a separate clause, and thus making it more *semantically salient*, somewhat along the lines suggested by Papafragou et al. (2002), mentioned in Section 4, though not in comparison to other languages, but *in comparison to other types of motion situations within the same language.*[6] The reasoning is thus somewhat paradoxical, and called for a further study in order to see if this correlation and possible explanation could be further supported.

5.2 Study 2

In this study we replicated Study 1, but using only 12 French speakers, this time of different ages (24 to 60), and professional/educational backgrounds. The linguistic descriptions were subjected to more thorough analysis. Since the Swedish group in Study 1 did not display a bounded/unbounded translocation asymmetry, this study was not designed as a comparative one.

The results from the similarity judgment task followed the same pattern as in Study 1: a general (though somewhat reduced) Manner bias but a reversal in the case of the Vertical Direction motion situation: 27 vs. 21 same-Direction choices. Furthermore, in dividing the Vertical stimuli in two groups depending on the direction of motion (24 each), it turned out that while in the case of UPWARD motion the ratio between same-Manner and same-Direction was even, in the case of DOWNWARD motion, there was a strong preference for same-Direction over same-Manner (15 vs. 9).

The verbal descriptions were this time analyzed differently. Each description was attributed to one of 5 different types: (i) Path/Direction+Manner in the same clause, (ii) Path/Direction and Manner in different clauses, (iii) Path/Direction only, (iv) Manner only and (v) Other, and each one of these was crossed with the *four* situation types (OUT/INTO, FROM/TO, Vertical-UP and Vertical-DOWN) – due to the differences in the similarity judgment task between the latter two, we decided to treat them separately. The results, displayed in Table 4, showed striking differences between the situation types. Whereas the most common type of verbal description for the bounded translocative stimuli, and especially FROM-TO, was that of Manner only, that for the unbounded translocative ones, and especially Vertical-DOWN was that of Direction+Manner in the same clause (highlighted in Table 4). Furthermore, taking together the rightmost two columns in Table 4, we can see that in the large majority of cases of FROM/TO (81,3%) Path was not expressed at all, and similarly for half of the OUT/INTO stimuli (49,2%). On the other hand, only a small minority of Vertical stimuli (16,7% and 20,8%) lacked an expression of Direction. No such conspicuous imbalance could be observed in the descriptions lacking Manner (the third and the fifth columns taken together).

Thus, we find a strong correlation between the French speakers' similarity judgments – the same-Path bias for the Vertical stimuli – and their linguistic descriptions: more frequent Path/Direction expression, particularly in the same clause. Admittedly this is again only a correlation, and given that the descriptions were produced *after* the similarity judgment task, this could not be a matter of any (direct) causation. Nevertheless, the correlation was so obvious that it seems paramount to search for an explanation.

One suggests itself once we realize that the 'Path' in the Vertical stimuli was rather Direction, and that the stimuli represented situations that were more readily interpreted as unbounded, rather than bounded. According to our redefinition of the 'boundary-crossing constraint' (Section 3.4) it is above all the boundedness of the situation that makes it difficult to express Manner and Path in the same clause in a V-language, while there is no such difficulty with respect to Manner and Direction. Thus, given that

Table 4. Classifying the data from the verbal description task in Study 2: 4 types of motion situations and 5 expression patterns, with the highest proportions highlighted

Situation \ Expression	Path/ Direction +Manner (same clause)	Path/ Direction & Manner (diff. clauses)	Path/ Direction only	Manner only	Other
FROM/TO (Tot: 48)	1 (2,1%)	6 (12,5%)	2 (4,2%)	**31 (64,6%)**	8 (16,7%)
OUT/INTO (Tot: 120)	15 (12,5%)	35 (29,2%)	11 (9,2%)	**41 (34,2%)**	18 (15%)
VERT-UP (Tot: 24)	**11 (45,8%)**	3 (12,5%)	6 (25%)	4 (16,7%)	0
VERT-DOWN (Tot: 24)	**14 (58,3%)**	3 (12,5%)	2 (8,3%)	5 (20,8%)	0

Manner is *perceptually salient* – which we know independently to be the case for the Event Triads stimulus tool – it is more likely to be expressed linguistically in a separate clause (row 2), or alone (row 4) in verbalizing bounded than unbounded translocative situations. Furthermore, this could increase the *semantic salience* of Manner, compared to the cases where it is 'conflated' in the same clause with Direction and thus lead to a stronger same-Manner bias.

Notice also in Table 4 that Manner was most often co-expressed with Direction in the case of Vertical-DOWN, and this was also the situation type that produced the most significant 'Path' (i.e. Direction) bias in the similarity judgment task. While this may be somewhat post hoc, we can interpret the difference between the two kinds of Vertical motion stimuli in terms of 'degrees of boundedness': e.g. rolling down is more open-ended than rolling-up to the top of a hill, and hence the Vertical DOWN stimuli represented the *least* bounded situation in the set. Thus we are lead to a tentative generalization (and prediction): *The more bounded a situation, the more salient Manner will be for speakers of a V-language.*

Pourcel (2005: 149) calls a similar interpretation that Zlatev and David (2004) offered of these results (though in terms of the concept of *telicity*) 'counter-intuitive', but we beg to disagree. As pointed out earlier, Bohnemeyer et al. (ms) noted a general tendency for lower same-Manner bias in the Vertical triads in the 17 languages studied, and while they did not find a general interaction with language-type, it remains unclear to what extent all the different languages in their sample abide by the 'boundary crossing constraint'. Swedish and Thai do not, and we did not find a bounded/unbounded asymmetry in their speakers' similarity judgments, which in the case of French we did. Pourcel (2005) also found an asymmetry, but in the opposite direction: greater Path salience for the bounded than for the unbounded situation. However, the design-differences between the two studies can perhaps be called on for an explanation, cf. Section 6.

Finally, note that we do not interpret the combined results of Study 1 and Study 2 in terms of a 'Whorfian' effect, since the differences in the categorization preferences between the language groups seems to be due to an *interaction* between language-independent differences in the situation types, and the constraints of a particular (type of) language. To further investigate this possible interaction, we conducted our next study, which more explicitly contrasts different contexts in which language can be thought to influence the categorization of motion situations *to different degrees*.

5.3 Study 3

For the purpose of our third study, we modified the Event Triads elicitation tool so that two groups of 12 Swedish and two groups of 12 French subjects participated: Group 1 for both languages performed the similarity judgment as in the original Event Triads tool, whereas for Group 2 there was a break after the first segment and the participant was asked to 'describe the film just seen', after which the second segment was shown and the participant was asked to make the similarity judgment. Furthermore, the number of distracters was decreased rather drastically from 38 to 8, leaving the total number of triads per participant to 20, where each first segment was described by all participants: for Group 1 after the similarity judgment task was competed, and for Group 2 prior to each judgment. In this way we could investigate possible correlations between the descriptions and the choices not only on a type-by-type basis (as in Studies 1 and 2), but also on a triad-by-triad (instance) basis. The reduction of distracter triads was necessary, since describing 50 video-clips, most of which are near-identical, would have been both tiring for the participants and could lead to a sort of 'habituation' in which they would fall into a stereotypical pattern of description that is less likely to reflect naturalistic language use.

The results were highly interesting. Whereas the similarity judgments for Group 1 (post-choice description) were similar to those in Study 1 and practically identical for the two languages ($chi2(1) = 0.14$, $p > 0.05$), i.e. a preference for same-Manner choices (albeit a weaker preference, cf. Figure 3), the situation was completely reversed for Group 2 (pre-choice description), with a surprisingly strong bias for same-Path choices, as shown in Figure 5 for both the French and Swedish groups. The difference between Group 1 and Group 2 was extremely significant ($p < 0.0001$). Furthermore, there was a stronger Path/Direction bias for SG2 than FG2, which was also significant ($chi2(1) = 4.964$, $p=0.026$).[7]

When we divided the 12 test triads according to the three types of situations as before (IN/OUT, FROM/TO and VERTICAL), we noticed, however, also a difference between Group 1 and the previous results: in the case of VERTICAL the Manner-bias was neutralized for both the Swedish and the French speakers (the slight difference between SG1 and FG1 for Vertical is not statistically significant): see Figure 6 and 7.

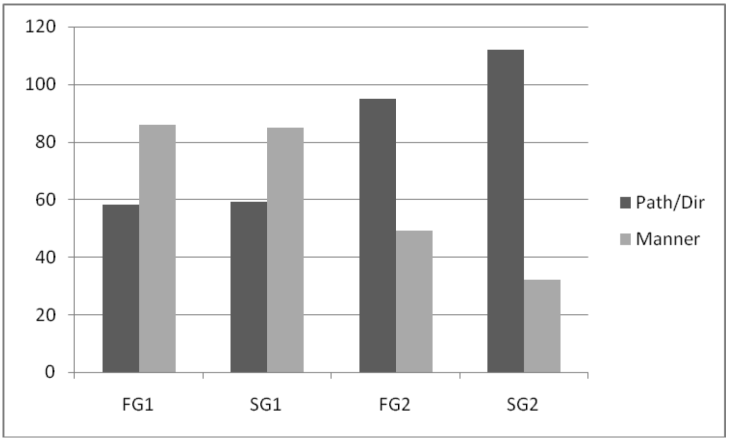


Figure 5. Total results of same-Path/Direction vs. same-Manner preference for French (FG1) and Swedish (SG1) Group 1 (post-choice description) and French (FG2) and Swedish (SG2) Group 2 (pre-choice description). Total number of choices is 144 per group

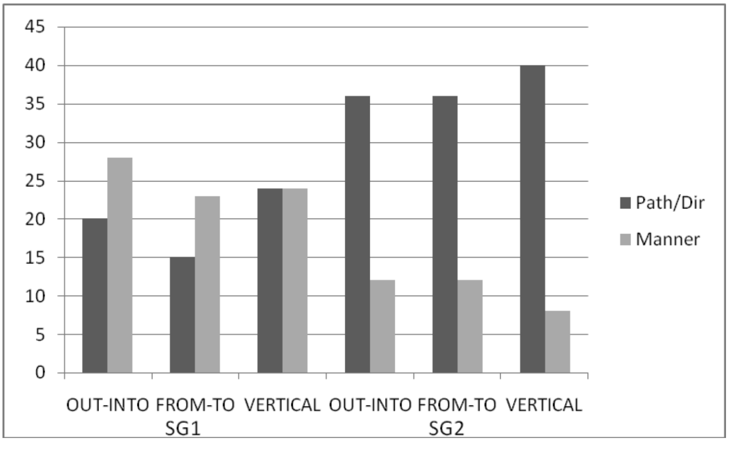


Figure 6. The results for the two Swedish groups, divided by the three different situation types: OUT-INTO, FROM-TO and VERTICAL. Total number of choices per situation type and group is 48

We coded the linguistic descriptions for the presence of Manner expressions: Manner verbs such as *hoppa* and *sautille* ('jumps') and adverbials such as *snurrande* or *en roulant* ('rolling'), Path expressions such as *från* or *de* ('from') and *till* or *a* ('to') and Direction expressions such as *upp* ('up') or *monte* ('climbs') and *ner* ('down') or *descend* ('descends') and looked for correlations between the presence of these elements and the choices of the subjects.

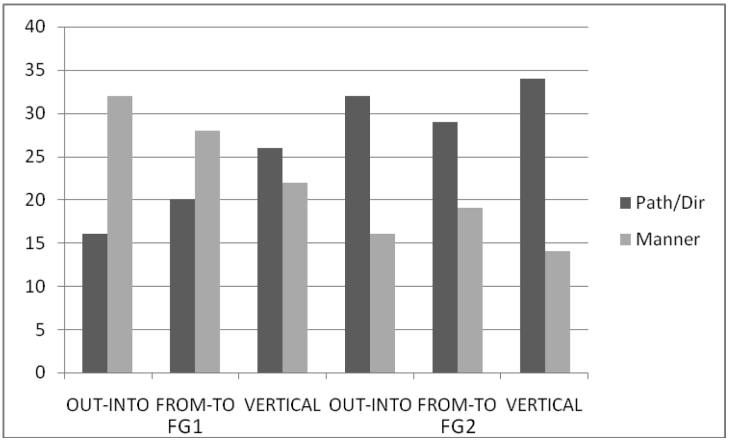


Figure 7. The results for the two French groups, divided by the three different situation types: OUT-INTO, FROM-TO and VERTICAL. Total number of choices per situation type and group is 48

Table 5. Correlations of significant value (Pearson's Correlation, significant at > ± .3 at the .05-level, two tailed) between elements in the descriptions (Direction, Path, Manner) and corresponding choice for the two groups of French speakers (FG1 and FG2) and the two groups of Swedish speakers (SG1 and SG2), divided by situation type (From/To, Out/Into, Vertical) Non-existing or non-significant correlations are marked as ×

Group	Type	Direction	Path	Manner
	From/To	×	×	×
FG1	Out/Into	×	×	×
	Vertical	×	+.308	×
	From/To	×	×	×
FG2	Out/In	×	×	- 304
	Vertical	×	×	-.329
	From/To	×	-.309	×
SG1	Out/Into	-.338	×	×
	Vertical	-.302	×	×
	From/To	×	-.307	×
SG2	Out/Into	×	-.443	×
	Vertical	+.674	×	×

Surprisingly, there were few *positive* correlations: for SG2-Vertical and for FG1-Vertical. In qualitative terms, this means that if a subject had used a Direction expression, he was more likely to make a same-Direction than same-Manner choice. We are not sure how to interpret the *negative* correlations for SG1, SG2 and FG2. On the face of it, it seems that e.g. if a French speaker had used a Manner expression (in the pre-choice description group, FG2), he was less likely to make a same-Manner choice. Also, it was surprising that the positive correlation for the French speakers was for FG1, the post-choice describing group, and it involved Path, rather than Direction expressions. In other words, the results do not lend themselves to an explanation in terms of Slobin's (1996) thinking-for-speaking hypothesis. According to the latter, and the classification of Swedish as an S-framed and thus Manner-salient language and of French as a V-framed language, one would have expected a Manner correlation for SG2 a Direction/Path correlation for FG2. In fact, the only clear correlation was for SG2, and it involved Direction rather than Manner.

A further indication that the results cannot be explained only on the basis of linguistic differences and their effects is the neutralized Direction-Manner bias in the case of the translocative unbounded (VERTICAL) situations for both the French and the Swedish groups (see Figures 6 and 7). Unlike the results from Study 2, this cannot be explained by a linguistic effect since Swedish does not obey the boundary-crossing constraint. Thus, the asymmetry in the choices between VERTICAL and the other two types of stimuli thus corroborate our claim in Section 3 that bounded and unbounded translocative situations differ (even) pre-linguistically. We may express this by saying that *Direction is conceptually simpler than Path*: all that is required is to pay attention to the *vector* (or the shape of the trajectory) of translocation, rather than perform an explicit 'parsing' of the translocative event in terms of Source, Via and/or Goal. Like Manner, Direction seems to be a category that is *more perceptually given than conceptually derived*, and thus less subject to the effects of linguistic mediation, as understood by Vygotsky (1978, 1986).

6 Discussion: from linguistic relativity to linguistic mediation

The Soviet psychologist Lev Vygotsky (1896–1934) distinguished between 'higher' and 'lower' mental functions, described by Kozulin (1986: xxv) as follows:

> Vygosky […] made a principal distinction between 'lower', natural mental functions, such as elementary perception, memory, attention, and will, and the 'higher', or cultural, functions which are specifically human and appear gradually in a course of radical transformation of the lower functions.

Thus, what is uniquely human, according to Vygotsky, is the ability to use artefacts and signs, *mediating* between perception and behaviour, and functioning as 'psychological tools' for the purpose of reflection and self-regulation: 'the central fact about our psychology is the fact of mediation' (Vygotsky 1933, quoted by Wertsch 1985:15).

The most important kind of signs, and thus psychological tools, are according to Vygotsky those of language. Like artefacts, linguistic signs are initially social and interpersonal, but with experience become *internalized* and thus intra-personal. Vygotsky argued that such internalization occurs via so-called 'egocentric speech' in early childhood, and that such speech is highly functional for the child since its presence increases with the difficulty of the task to be performed.

Applying the notion of linguistic mediation to the triad studies, both our own, and those described in Section 4, allows us to make sense of most of the results reported in the literature. First, due to the nature of the task, the similarity judgment task can be performed either more *directly* (i.e. using perceptual categorization) or more *mediatedly* (i.e. using external or internal speech). This can explain the results of both Gennari et al. (2002) and Finkbeiner et al. (2002), in which a typologically congruent bias was observed in the tasks where language was used either overtly or (apparently) covertly, but not otherwise. On the other hand, if Manner is a category which is (in general) more perceptually and conceptually *simpler* than Path, as suggested earlier, then tasks which induce categorization through less mediated processes, should bias for Manner rather than Path, and vice versa. We can thus explain the results of Study 3 for both the Swedish and the French groups through a possible 'Vygotskyan effect' of language on the categorization of (translocative) experience: linguistic mediation yields an explicit 'parsing' of the components of a motion situation, and thus attention to more *abstract* components such as Path than to more perceptually immediate components such as Manner (or Direction). Such an effect appeared to be independent of typological differences between languages. At the same time, this interpretation predicts that if Manner is expressed in French, it will be more prominent *semantically*. Whether this would lead to a cognitive effect, however, is less clear: Study 2 seems to support this, while Study 3 (e.g. the negative correlations for Manner in Table 5 for FG2), did not.

At the same time, if the Manner of motion is of a *complex type*, such as that in the stimuli used by Finkbeiner et al. (2002), while 'Path' is more a matter of 'moving left/right' and thus Direction, then the opposite tendency should be observed: a greater same-Manner bias will be observed in the more demanding task, involving sequential presentation and language-based short term memory, which again was the case established in that study.

This can furthermore even help us understand the apparently contradictory findings in the triad study of Pourcel (2005): In her first experiment with both French and English participants, the sequential presentation of stimuli possibly already induced the use of internal speech, resulting in an overall preference for 'same Path'. The second experiment used explicit written description, which 'balanced' the preferences somewhat, but still privileged Path. What remains unaccounted for, though, is why 'same-Manner' preferences *were higher* for the 'atelic Path' (unbounded) situations than for the 'telic Path' (bounded) situations, while in our studies the asymmetry was in the reverse direction: a neutralization of the Manner-bias, and thus relatively lower 'same-Manner' preferences for the 'less bounded' situations. The divergent results can perhaps be explained by the marked difference in the nature of the stimuli used: whereas the relevant kinds of Manner in Pourcel's experiment were mostly of the 'default' kind and thus less perceptually salient, those in our studies were all attention-grabbing, yielding an *overall* Manner-bias on categorization (mostly) on the basis of perceptual processes. This bias was then *reduced* for the slope scenes due to the similarly perceptually more immediate notion of Direction and possibly also due to the greater ease of verbalizing Manner in the same clause as Direction (as opposed to Path) for French. On the other hand, Path and Direction were the most relevant aspects in Pourcel's stimuli, while (simple clause) verbalization would instead have promoted Manner to higher prominence. In any case, both studies imply the importance of distinguishing between what we have analyzed as bounded vs. unbounded translocation, and thus offer support for the taxonomy presented in Section 3.

7 Summary

In this chapter we have tried to show that 'motion event' typology has suffered for quite some time from conceptual and empirical problems, and despite the indubitable contributions of scholars such as Talmy and Slobin, it is time that we move on, and establish a more coherent framework for describing our experiences of motion. Inspired by the literature on situation types (Vendler 1967), as well as Durst-Andersen (1992) and Pourcel (2005), we have attempted to provide one such framework through our taxonomy of motion situations, which, we suggest, are largely independent of the way different languages 'lexicalize' motion.

The second step, which we have only here touched, is to try to establish how as many (diverse) languages as possible express this experience. Talmy's binary typology has clearly outlived its time, but exactly how many different types of languages in terms of their expression of translocation there are is currently an open question.

In the cases where languages systematically differ in this respect, we can investigate possible linguistic effects of various sorts and strengths on seemingly 'non-verbal' cognitive tasks, and thus contribute to the neo-Whorfian program. We have described three such studies which suggest at least some effect of the differences between French on the one hand, and Swedish and Thai on the other, on the categorization of translocative situations on the basis of the components Path, Direction and Manner, arguing for the necessity of distinguishing between the first two. The effects have, however, been attributed to an interaction between language-independent factors and linguistic constraints, and cannot support a strong version of the Whorfian hypothesis ('different languages entail different worldviews').

We have also argued that we should be open to the possibility that the differences between languages may be relatively minor compared to their similarities – at least as far as the categorization of (motion) experience is concerned – and have thus suggested possible 'Vygotskyan' rather than 'Whorfian' effects, based on the differential role of linguistic mediation in the different tasks and study designs. Further studies with (typologically) different languages are likely to shed more light on these issues. Progress in linguistic typology and psycholinguistics should thus go hand-in-hand.

Acknowledgements

We wish first of all to thank all the participants in our studies, most of who did so without any economic compensation, and furthermore the host institutions who made the studies possible, including the *Centre for Research in Speech and Language Processing* (CRSLP) at Chulalongkorn University. We also thank the participants at seminars at the Copenhagen Business School and at Lund University, an anonymous reviewer and the editors of this volume for their useful comments. Also Joost van Weijer for help with statistical analysis. Finally, we express our warmest gratitude to Jürgen Bohnemeyer, Sonja Eisenbeiss and Bhuvana Naranhimsan for generously allowing us to use (and modify for our purposes) their Event Triads tool.

Notes

1 Zlatev (2005, 2007) refers to this as the distinction between 'elaborated' and a 'schematic' concept of Path, and argues for the need to separate the latter from the concept of Direction, as in the present chapter.

2 The original examples in Pourcel (2005, Chapter 5) are respectively (6), (90) and (92).

3 Note that our use of the term 'figure' corresponds to that used by Talmy (2000) and Levinson (2003), the term 'trajector' (Lakoff 1987; Regier 1996; Zlatev 1997) or 'referent' (Miller and Johnson-Laird 1976). On the other hand, our use of the term 'landmark', is more specific than that used in much of the cognitive linguistic literature (Langacker 1987), in referring to some physical object, which is typically expressed through a noun phrase in language (cf. Zlatev 2005, 2007).

4 One thing to be borne in mind, however, is that this design has been shown to give a general bias towards Manner-based categorization, probably due to the conspicuousness of the motion of the 'tomato' figure, so that the results produced using this stimulus tool cannot be directly compared with results obtained using another elicitation tool (cf. Kopecka and Pourcel 2005).

5 This unequal distribution was due to the fact that at the time of our first study we had not yet realized the importance of distinguishing between the three types.

6 Notice that this also helps explain the high proportion of Manner verbs produced by the French group, shown in Table 3.

7 However, since the Group 2 data was both compared both with Group 1, and within the two sub-groups (FG2 and SG2), Bonferroni correction (here, p-value * 2) would be required, placing the difference between FG2 and SG2 on the border of significance.

References

Bohnemeyer, J., Eisenbeiss, S. and Naranhimsan, B. (2001) Event triads. In S. C. Levinson and N. Enfield (eds) *'Manual' for the field season* 100–114. Nijmegen, Max Planck Institute for Psycholinguistics.

Bohnemeyer, J., Eisenbeiss, S. and Naranhimsan, B. (ms) Ways to go: Methodological considerations in Whorfian studies on motion events.

Blomberg, J. (2006) Rörelse och verbalisering – En studie av språkets påverkan på kategoriseringen av rörelsesituationer. BA thesis. Centre for Languages and Literature, Lund University.

Blomberg, J. (2007) Linguistic relativity, mediation and the categorization of motion. MA thesis. Centre for Languages and Literature, Lund University.

Brown, P. (2004) Position and motion in Tzeltal frog stories: The acquisition of narrative style. In S. Stromqvist and L. Verhoeven (eds) *Relating events in narrative: Cross-linguistic and cross-contextual perspectives* 37–58. Mahwath, NJ: Earlbaum.

Durst-Anderson, P. (1992) *Mental grammar: Russian aspect and related issues.* Columbus, Ohio: Slavica Publishers, Inc.

Finkbeiner, M., Nicol, J., Greth, D. and Nakamura, K. (2002) The role of language in memory for actions. *Journal of Psycholinguistic Research* 31(5): 447–457.

Gennari, S. P., Sloman, S. A., Malt, B. C. and Fitch, T. (2002) Motion events in language and cognition. *Cognition* 83: 49–79

Husserl, E. (1999) [1907] *The idea of phenomenology.* (Trans. Lee Hardy, *Collected Works* Vol. 8.) Dordrecht: Klewer.

Itkonen, E. (1991) *Universal history of linguistics.* Amsterdam: John Benjamins.

Kopecka, A. and Pourcel, S. (2005) Figuring out figures' role in motion conceptualization. Presentation, International Cognitive Linguistics Conference, Seoul, Korea. 21 July 2005.

Kozulin, A. (1986) Vygotsky in context. Preface to L.S. Vygotsky *Thought and Language.* Cambridge, MA: MIT Press.

Lakoff, G. (1987) W*omen, fire and dangerous things: What categories reveal about the mind.* Chicago: Chicago University Press.

Langacker, R. (1987) *Foundations of cognitive grammar Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.

Levinson, S. C. (2003) *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Miller, G.A. and Johnson-Laird, P.N. (1976) *Language and perception*. Cambridge, MA: Harvard University Press.

Papafragou, A., Masely, C. and Gleitman, L. (2002) Shake, rattle, 'n' roll: The representation of motion events in language and cognition. *Cognition* 84: 184–219

Pourcel, S. S. (2005) Relativism in the linguistic representation and cognitive conceptualisation of motion events across verb-framed and satellite-framed languages. PhD Thesis, Department of Linguistics and English Language, University of Durham.

Pourcel, S. S. and Kopecka, A. (ms) Motion events in French. Typological intricacies.

Regier, T. (1996) *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Slobin, D. I. (1996) From 'thought and language' to 'thinking for speaking'. In J. Gumperz and S. C. Levinson (eds) *Rethinking linguistic relativity* 70–96. Cambridge: Cambridge University Press.

Slobin, D. I. (1997) Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani and S. Thompson (eds) *Grammatical constructions: Their form and meaning* 195–220. Oxford: Oxford University Press.

Slobin, D. I. (2003) Language and thought online: Cognitive consequences of linguistic relativity. D. Gentner and S. Goldin-Meadow (eds) *Language in mind* 157–192. Cambridge: MIT Press.

Slobin, D. I. (2004) The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist and L. Verhoeven (eds) *Relating events in narrative: Cross-linguistic and cross-contextual perspectives* 219–257. Mahwath, NJ: Earlbaum.

Slobin, D. I. and Hoiting, N. (1994) Reference to movement in spoken and signed languages 487–505. *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society*.

Smith, V. (2003) Motion at the sugar factory: Is Russian a genuine Manner language? In *Proceedings of the 14th Symposium on LSP*. University of Surrey, 18–22 August 2003.

Strömqvist, S. and Verhoeven, L. (eds) (2004) *Relating events in narrative: Cross-linguistic and cross-contextual perspectives.* Mahwath, NJ: Earlbaum.

Talmy, L. (1985) Lexicalisation patterns: Semantic structure in lexical forms. In T. Shopen (ed.) *Language typology and syntactic description Vol.* 3 57–149. Cambridge: Cambridge University Press.

Talmy, L. (1991) Path to realisation: A typology of event conflation 480–520. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*.

Talmy, L. (2000) *Toward a cognitive semantics: Concept structuring systems Vol. 1.* Cambridge, MA: MIT Press.

Vendler, Z. (1967) *Linguistics in philosophy.* Ithaca: Cornell University Press.

Vygotsky, L. S. (1978) *Mind in society.* Cambridge, MA: MIT Press.

Vygotsky, L. S. (1986) *Thought and language*. Cambridge, MA: MIT Press.

Wertsch, J. V. (ed.) (1985) *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

Whorf, B. L. (1956) *Language, thought and reality: Selected writings of Benjamin Lee Whorf.* (ed. J. B. Carroll.) Cambridge: MIT Press.

Zlatev, J. (2003) Holistic spatial semantics of Thai. In E. Casad and G. Palmer (eds) *Cognitive linguistics and non-Indo European languages* 305–336. Berlin: Mouton de Gruyter.

Zlatev, J. (2005) Semantics of spatial expressions. *Encyclopedia of language and linguistics.* (Second edition, Article 00298.) Oxford: Elsevier.

Zlatev, J. (2007) Spatial semantics. In D. Geeraerts and H. Cuyckens (eds) *Oxford handbook of cognitive linguistics* 318–350. Oxford: Oxford University Press.

Zlatev, J. and David, C. (2003) Motion event constructions in Swedish, French and Thai: Three different language types? *Manusya* 6: 18–42.

Zlatev, J. and David, C. (2004) Do Swedes and Frenchmen view motion differently? Presentation at the conference *Language, Culture and Mind*. Portsmouth, 20 July 2004.

Zlatev, J. and Yangklang, P. (2004) A third way to travel: The place of Thai in motion event typology. In S. Stromqvist and L. Verhoeven (eds) *Relating events in narrative: Cross-linguistic and cross-contextual perspectives* 159–190. Mahwath, NJ: Earlbaum.

16 Motion: a conceptual typology

Stéphanie Pourcel

1 Introduction

The domain of motion in space has been extensively studied in cognitive linguistics (e.g. Talmy 1985, Slobin 2004) and in investigations of the language-cognition relationship (e.g. Gennari, Sloman, Malt and Fitch 2002, Hohenstein 2005, Papafragou, Massey and Gleitman 2002, Pourcel 2004, 2005, 2009a, b, Slobin 1996, 2000). Likewise, the domain has received attention in psychological studies of perceptual and conceptual aspects of motion information processing (e.g. Mandler 2004). Yet, comparatively few efforts – if any – have sought to adduce a comprehensive framework for the *conceptual* definition of motion variables and motion types. In this chapter, I examine the domain of motion independently of language in order to design a conceptual typology of motion that will serve linguistic analyses and applications of these analyses. Importantly, the typology attempts to categorise types of motion in a way that is concordant with how human minds categorise motion types in conceptualisation – using cognitive data obtained from categorisation tasks. In other words, the conceptual typology is not based on language patterns, as are linguistic typologies, but on conceptual categories. The conceptual typology therefore offers a classification of events and sub-events pertaining to one domain, in which the types are isolated on conceptual grounds, rather than on linguistic, cultural, folk, or otherwise arbitrary grounds. By virtue of being based on conceptual, rather than symbolic, representations, the typology characterises a fundamentally human understanding of given domains. The typology should thus inform potentially universal patterns of classification, rather than language- or culture-specific ones. The present research tests this possibility only to an extent by using native populations of typologically diverging languages – French, Polish and English.

Sketching a conceptual typology for the domain of motion is important for several reasons. First, motion is a complex domain with dynamic and variable schematic components. These components do not simply comprise moving figures following spatial paths. They also comprise physical manners of displacement, force dynamics, landscapes, locations, causal motivations, goals, resultative endstates, objects (e.g. vehicles, instruments, buildings), and more, which are situated within ideational contexts, comprising cultural cognitive models, ideologies, emotions, symbolisms, as well as expectations concerning motion properties and contextual embedding. Each of these motion component is rich and variable in nature, which entails that motion types are in turn numerous, rich and complex. Crucially, this suggests that the domain of motion is unlikely to be conceptualised in a unitary fashion. The specific claim made in this paper is that this complex and variable range of event realisations reflects conceptually distinct types of motion. Motion types have consequences for the cognitive representation of the event in question, e.g. in memory, categorisation, analogy, etc. On this basis, it should be possible to isolate different types of motion, which may correspond to distinct conceptual realities.

Besides its rich complexity, motion is a pervasive domain of experience, which is conceptualised and also expressed in language with high frequency in human daily life. As a result, the domain of motion has received extensive attention in cognitive linguistics, for instance, in typological work (e.g. Talmy 1985, 1991, 2000), in lexico-discursive studies (e.g. Slobin 2004), and in explorations of the language-cognition relationship (e.g. Gennari et al. 2002, Papafragou et al. 2002, Pourcel 2009a, b,, 2005, Zlatev and David 2004, 2005). Most studies have faithfully adhered to the usage-based tradition embraced by cognitive linguistics, and have therefore been comparative in their treatment of linguistic data. However, the generalisations drawn – though useful – are typically emergent from the language data and take little account of the conceptual reality of motion independently of language. This does not mean that existing means for analysing motion are incorrect, but that they largely remain products of the reality depicted by the languages observed. These means are therefore anchored in hermeneutics and correspond to language-embedded categories, rather than to conceptual categories pertaining to an 'objective' reality. To the extent that cognitive linguistics seeks to commit to an understanding of language that is concordant with the workings of the human mind, generalisations that are mainly or solely language-based prove potentially problematic.

In addition, the categories identified in linguistic work are often applied beyond the realm of linguistic description to questions of conceptual relativism (see Whorf 1956, Lucy 1992). For instance, a number of studies have investigated whether richer lexical resources and more systematic means of encoding manners of motion in the grammar of a particular language render those manners more cognitively salient to their native speakers (e.g. Gennari et al. 2002, Papafragou et al. 2002, Zlatev and David 2004, 2005, Pourcel 2009a, b). To guide their assumptions, hypotheses and experimental designs, these studies have departed from cross-linguistic differences. In other words, they have adopted the categories found in language to investigate matters of conceptualisation, which are richer and more complex than the linguistic means used to encode them. Few of these studies have reached successful conclusions or even consensus across their respective findings. The outcome of these studies might have proved altogether different had they considered, from the outset, the conceptual reality of motion – independently of language – and had they examined their data relative to conceptually real categories, rather than to solely linguistically-defined parameters such as path and manner. In addition, though their methodologies were often closely similar (e.g. triad-based similarity judgement tasks), research teams have often used different figures in their motion stimuli. For analytical purposes, therefore, the design of a conceptual typology of motion with greater attention to figure types and other related schematic properties is paramount to avoiding linguacentric tendencies in research, which go against the very stance of cognitive linguistics as a language-analytic enterprise. The issue of linguacentrism in linguistic, cognitive and behavioural research has long been the subject of methodological discussions (e.g. Whorf 1956: 162), yet this key pitfall is still not entirely addressed in contemporary studies investigating the language-cognition interface (Lucy 2003: 25).

This paper attempts to address the notion of linguacentrism in linguistic and relativistic research, with special attention to the domain of motion. To do so, it offers a preliminary sketch of a conceptual typology for the domain of motion, based on experiential facts and cognitive data, with the types isolated corresponding to important conceptual differences and, potentially, to actual categories of events. In doing so, it aims (i) to represent comprehensively the complexity of the domain of motion, (ii) to offer a 'language-neutral' analytic template (or metalanguage) for the comparative study of motion linguistics, and (iii) to allow for greater discernment in applications of motion linguistics to questions of cross-linguistic conceptual representations (e.g. linguistic relativity).

To achieve this, the paper proceeds with an initial illustration of the diversity existing across types of motion, followed by a proposal to place motion figures at the centre, or basis, of the conceptual typology. The choice for basing the typology on figure type is motivated by the fact that the physical properties of figures constrain and determine to a large extent the types of motion figures undergo, with impacts on types of manners, paths, causal relations and so on. A detailed outline of figure types and their properties makes it possible to establish a conceptual typology of motion types. Within these motion types, special consideration is paid to schematic variables such as paths, manners, causality, agency and intentionality. This proposal, together with the conceptual distinctions between motion types are substantiated with experimental data drawn from categorisation tasks implemented with speakers of various languages, including English, French and Polish. The tasks at hand systematically represented instances of motion in a visual medium, e.g. film clips and film extracts. The stimuli were then non-linguistic. This paper presents examples of these motion scenes in the written medium required by the static publication format. Yet, it is key to understand that the discussion is not focused on the linguistic examples (here, in English), but instead, on the conceptual reality encoded by these examples.

The paper ends with a suggestive sketch for a conceptual typology of motion, based on the figure types identified.

2 Some important distinctions in motion types

A few examples of motion situations suffice to illustrate the variability of possible motion types and of the internal properties of individual motion schemas. Consider the following scenarios:[1]

(1) Helen is jogging.

(2) Helen walked to the store.

(3) Tom pushed the pram along the street.

(4) Mum rocked the baby to sleep.

These four examples illustrate a variety of types of motion. For instance, (1) depicts a *motion activity* where the ongoing nature of the motion, together with its manner of completion are central to conceptualising the event. In (2), the motion type is no longer an activity, but corresponds more adequately to a *motion event*, as commonly discussed in the literature. An important difference between (1) and (2) concerns the presence of directionality and the reaching of a goal or destination in the second example. In (2), the directionality, or path endpoint, represents the goal of the figure's motion, whereas in (1), the motion itself constitutes the figure's goal. (1) and (2) thus contrast instances of activity versus event, respectively (Pourcel and Kopecka 2005, to appear).

In (3) and (4), the goal of the figure's motion no longer applies to the motion figure itself, but to an external entity. For instance, in (3), the figure seeks to alter the spatial location of an object – here, a pram; and in (4), the figure seeks to alter the state or condition of another being. In both cases, motion is instrumental in completing these goals, in that motion is used as a means to an end. The pushing and rocking events depicted in (3) and (4) are instances of *causal motion*.

The two situations in (3) and (4) are distinct, however. In (3), the figure's motion causes the change of location of another figure (i.e. effectively the pram's motion along a path), whereas in (4), the figure's motion causes the change of state of another figure, where the rocking causes the baby to fall asleep. A situation like (4) may be characterised as *caused state*, that is, the motion of Figure 1, i.e. the mother, causes the change of state of Figure 2, i.e. the baby. The state is dependent on the motion. In (3), we also have two figures. The two figures are moving figures which undergo a change of spatial configuration. The motion of Figure 2, i.e. the pram, is dependent upon the motion of Figure 1, i.e. Tom. (3) thus represents an instance of *caused motion*.

In addition, (3) illustrates that moving figures, such as people and objects, may be agentive or passive. *Agentive figures* are capable of *self-motion*, whereas *passive figures*, such as objects, are subject to caused motion. This distinction reflects a fundamental difference in terms of animacy and, often, intentionality. In caused motion, the primary, or motion-causing, figure is animate. This points to another potential distinction between *animate* and *inanimate motion*.

Additional distinctions are present when contrasting motion instances (2) and (3). Both are directional motion events. Yet, (2) indicates a clear path endpoint, whereas (3) does not. In other words, the motion goal is overt in (2), but not in (3). This difference may be characterised in terms of telicity (Aske 1989, Talmy 1991), and yields another contrast between *telic motion* where the motion goal is explicit, e.g. (2), and *atelic motion* where the motion goal is not apparent, e.g. (3).

Finally, when contrasting the motion situations in (1)-(4), we also note differences in terms of the physical properties involved in motion performance. These properties differ relative to aspects such as degree of control, muscular effort, speed, and general force dynamics. Types of motion may therefore be distinguished relative to physical characteristics. For instance, walking – to an average adult human being – involves relatively little effort or control compared to jogging. In fact, walking is a typical manner of adult human motion. (2) represents an instance of *default motion* (Pourcel 2004).

Other manners such as jogging, or limping, waltzing, staggering, and so forth, represent less typical, or non-default instances of motion. Distinctions linked to the physical manner of displacement of a figure thus lead to additional types of motion.

In sum, motion types may be classified relative to directional, aspectual, causal, agentive, or physical properties, and possibly more. The few types of motion reviewed so far present important conceptual characteristics, despite the fact that they are mostly encoded in language (here, English) using the same constructional pattern (Talmy 1985). Indeed, all the above examples illustrate the satellite-framed pattern of linguistic encoding for motion typical of the English language. This pattern maps the figure onto the subject position of the sentence, the manner of motion onto the main verb, the path onto a verb particle, or satellite, and the ground onto a nominal complement to the satellite. Additionally, in causal cases, the manner verb is followed by the second motion figure in a direct object. In other words, one main grammatical formula, or construction type, serves to encode the distinct types of motion outlined above. Formulaic tendencies in language do not preclude, however, the existence of significant conceptual distinctions across the events expressed. It is in conceptual, rather than linguistic, terms that this article aims to analyse the domain of motion. To do so, it is necessary to unpack and deploy apparent schematic differences between the types of events that motion may entail, regardless of the linguistic patterns available for the expression of this domain.

2.1 Directionality

A broad distinction may be drawn between motion types where motion is incidental to a change of location, i.e. motion events, and motion types where motion is the essence of the figure's action, i.e. motion activities. Consider, for instance:

(5) a. Helen is jogging.
  b. Helen jogged to the store.

Examples in (5) both involve motion, but (5a) depicts an activity and (5b) an event. The main difference between activities and events relates to directionality, as briefly mentioned above. Activities lack overt directionality, whereas events necessitate directionality by virtue of entailing a change of locational grounds.

Motion activities then describe motion with an emphasis on the type of motion itself – typically its manner. Activities, like all types of motion, include spatial reference, i.e. a ground, but the notion of directionality, or path, is not salient. In fact, there may not be any overt directionality at all in some instances, as in the act of running on a treadmill. The essential schemas in activities consist of the figure and its actual motion as characterised by a specific manner. The high salience of the manner schema versus the low salience of the path schema in motion activities is incidentally examplified in language, where activities may be expressed without mention of path details, e.g.

(6) Stop kicking your sister!

(7) Mary swims every morning.

Motion events, on the other hand, describe directional or goal-oriented motion. In this case, paths are salient conceptual dimensions of motion, whereas manners are merely instrumental to following the course of the motion path. The core schema of a motion event is therefore its path (Talmy 1991). To further illustrate the core schematicity of path and the secondary importance of manner, linguistic mappings of motion events in English may leave manner information unexpressed, using a neutral motion verb such as *go* or a path verb instead, e.g.[2]

(8) Mary went to the store.

(9) Tom crossed the bridge.

The distinction between event and activity is important as it contrasts the salience of the manner and of the path of motion, two central elements of motion. It is possible that this distinction is conceptually real, so that activities and events are systematically conceptualised as different types of actions. This possibility remains to be tested at this stage.

2.2 Telicity

Besides directionality, motion involves an aspectual dimension, in that motion may be ongoing or completed. This distinction is typically discussed in terms of 'telicity' (from Greek *telos* meaning 'end') – as briefly mentioned earlier. In the domain of space, telicity, or event completion, is understood as the reaching of locational goals and the obvious change in the locational state of the figure. Motion activities are typically ongoing, uncompleted acts of motion and therefore are atelic. Motion events, on the other hand, can be either telic or atelic depending on the type of motion path. Indeed, although motion events involve directionality, the path goal may not necessarily be apparent and salient. Consider, for instance, UP events in which the top of the ground is not readily at hand, visible, or reached. Likewise, ALONG events entail by definition an endpoint-free path, e.g.

(10) John walked along the beach for hours.

These types of events are atelic due to the lack of a path endpoint, yet they differ from activities by virtue of having directionality.

Motion events are most frequently telic, however. They typically involve an endpoint and a change of location or state. Consider the instance of an UP event in which the top of the ground is readily visible or reached. The trajectory endpoint is highly salient in this case. Telic events may also involve change of location via the crossing of boundaries, as in crossing, entering, or exiting events, e.g.

(11) a. The plane flew across the Atlantic.
 b. Mary dived into the pool.
 c. Jo walked out of the room.

Note that telic events lead to resultative interpretations of motion. By undergoing motion along a path, the figure reaches a goal, which corresponds to a particular location in space. This resultative aspect is all the more obvious in cases of caused state events, where the motion of a primary figure causes the resultative end-state of a secondary figure, e.g.

(12) The attacker stabbed the victim to death.

Aspect is thus an inherent property of motion, and we may distinguish an additional two types of motion relative to this property, namely *telic* and *atelic motion*.

2.3 Causality

A further distinction across motion types is effected by the figure's relation to its own motion, to the motion of other figures, or to the endstate of other figures. Indeed, a figure may instantiate its own motion or that of other figures. Consider,

(13) a. Felix crossed the street.
 b. Tom pushed the pram across the street.
 c. Mum rocked the baby to sleep.

Example (13a) presents a single figure initiating its own motion via its physical motor abilities. (13a) is an instance of *self-motion*. In (13b) and (13c), on the other hand, we have two interacting figures in each motion scene, i.e. Tom and the pram in (13b), and Mum and the baby in (13c). As far as the motion is concerned, the two figures are interdependent in their interaction. A primary figure initiates motion and, in so doing, causes an alteration in a secondary figure. As a result, the secondary figure may undergo a change of spatial location, as in (13b), or it may undergo a change of state, as in (13c). Examples in (13b) and (13c) display instances of *causal motion* – from the perspective of the primary figure. (13b) may also be characterised as *caused motion*, and (13c) as *caused state* – these terms reflect the perspective of the secondary figure.

Self-motion then describes the physical motion of a figure which initiates its own motion independently of external elements, whilst causal motion describes the motion of a primary figure causing the alteration of a secondary figure. Causal motion may cause either the motion of a secondary figure, or the change of state of a secondary figure.

In the case of caused motion, we obtain the motion of two figures with a cause-effect relation between motion$_1$ and motion$_2$ (see (14) and (15) below). In effect, the motion of Fig$_1$ causes the motion of Fig$_2$, so that Fig$_1$ moves Fig$_2$. Note that, linguistically, motion$_1$ characterises the manner of motion and motion$_2$ captures the path of motion – in English, e.g.

(14)	The wind	blew	the hair	into her face.
	Fig$_1$	motion$_1$	Fig$_2$	motion$_2$
		manner		path

(15)	Mary	drove	her sister	to the airport.
	Fig$_1$	motion$_1$	Fig$_2$	motion$_2$
		manner		path

Effectively, in (14), the wind causes the hair to go into her face by blowing, and in (15), Mary causes her sister to go to the airport by driving her there.

In the case of caused state, on the other hand, we obtain the motion of one figure and the state of a secondary figure with a cause-effect relation between the motion of the primary figure and the state of the secondary figure. So, the motion of Fig$_1$ results in the state of Fig$_2$, e.g.

(16)	Sara	kicked	the door	shut.
	Fig$_1$	motion	Fig$_2$	state

In sum, causal motion entails an interdependent relationship between two figures, with the primary figure engaged in a motion act. Causal motion results either in the change of location or in the change of state of a secondary figure.

In addition, the causal nature of motion impacts on the agentive value of the motion path. In the case of self-motion, path is agentive, in that the path is instantiated and followed by the motion-initiating, or primary, figure, e.g.

(17) Jim dived INTO the pool.

Causal motion, on the other hand, involves *resultative paths*, as in (14)-(15) above and (18) below.

(18) Carol poured her drink OVER Gary's head.

In causal motion then, the path corresponds to the result of Fig$_1$'s motion and refers to what is effectively happening to Fig$_2$. Fig$_1$ does not necessarily follow Fig$_2$'s path. Note that it may, as in (15) above.

This outline points to important differences in motion types relative to causal properties. These differences mainly concern the physical nature of the figure, as well as the interactions between distinct figures. Causal properties, in addition, determine the agentive value of the path of motion, which may be either agentive or resultative. These differences demonstrate the dynamic complexity of motion as a domain of experience, which may, as a result, engender divergent conceptual representations for the types observed.

2.4 Animacy

As highlighted above, in causal motion, the secondary figure may be both an object, e.g. a pram, or a person, e.g. a baby, but the primary figure cannot be an object, such as a pram. The primary figure needs motor capacities or inherent force dynamics to initiate its own motion and that of other figures. To fulfill this requirement, the primary figure needs to be animate. If we take the capacity for self-motion as a defining criterion for judging animacy, then we may categorise as animate any live creature with motor abilities (e.g. humans, animals), natural elements (e.g. water, fire), as well as natural forces in the universe (e.g. magnetism, electricity), and force-animated entities (e.g. planets, currents), and so on.[3] On the other hand, dead organisms, objects, and even vehicles can only undergo displacement in space when set into motion by animates. It follows that inanimate figures may only be subject to causal motion. On this basis, we may distinguish *animate* (e.g. (19a)-(b)) from *inanimate motion* (e.g. (19c)).

(19) a. Jayne kicked the ball.
b. The earth rotates around its own axis.
c. The ball rolled down the hill.

Examples in (19) present conceptual differences relating to animacy properties. In terms of self-motion capacities, only (19a) and (19b) offer figures that may cause their own motion and that of other figures. In (19c), the figure is in motion presumably because it was set into motion by a human or animal being, or by a natural element such as the wind or other. We may therefore distinguish animate motion from inanimate motion.

2.5 Agency

Figure properties seem significant in conceptualising motion types. The foregoing has defined a distinction in terms of animacy. However, the motion of a live creature versus that of a natural force seem to present equally important differences. Both types of figures may be considered animate and capable of self-motion. However, only the former may be characterised as agentive. That is, only live creatures, e.g. humans and animals, may instantiate motion as an intentional act. Forces, on the other hand, do not operate on an intentional basis, but on a purely mechanical one as relating to the laws of physics in the universe. The intentional dimension of motion present with live beings relates to agency, and may correspond to a figure's explicit goals or instinctive impulses for spatial displacement. Whether to the foreground or background of consciousness, the motion of live beings is typically a cognitive response to the world (to the exception of reflex motion), rather than a purely physical and mechanical one, and is often a volitional act. Consider, for instance,

(20) The cat pounced on the butterfly.

(21) The tide is coming in.

We may therefore establish a further distinction between *agentive* and *non-agentive motion*, as relating to figure properties. Note, too, that the motion of object figures is non-agentive. This is also the case for animates undergoing causal motion, in the sense that the figure, though animate and potentially agentive, is not in this case cognitively active in engaging the act of motion.

2.6 Force dynamics

In addition, one may distinguish motion types relative to the figure's physical capacity for motion. An examination of the force dynamics of motion instances directs us to manners of displacement. The case is most obvious with human figures, who are highly flexible and creative in using their bodies and other objects to undergo motion. Consider,

(22) a. Chris walked across the car park.
b. Chris hopped across the car park.
c. Chris skated across the car park.

Examples in (22) demonstrate that one figure may employ several manners of motion with distinctions present in terms of force dynamics such as speed, level of control, impacts on physical aspects (e.g. tiredness), muscular effort, use of external objects, and more. These distinctions may correspond to differing goals on the part of the figure. Importantly too, a conceptualising observer may have given expectations relating to motion performance based on the nature of the figure. Manners may be highly variable, whilst at the same time, they are constrained by the physical properties of the moving figure, e.g. pigs cannot fly, snakes cannot walk, and so forth. Relative to figure type then, we may elaborate a broad manner-based classification of motion types, including (a) *typical*, or *default, motion* (e.g. (22a) so long as Chris is a human figure), (b) *atypical*, or *forced, motion* (e.g. (22b)), and (c) motion types requiring a form of support, instrument, or vehicle, which may be referred to as *instrumental* instances of *motion* (e.g. (22c)) (see Pourcel 2004). According to this classification, default motion depicts the typical manner of displacement for a given figure, e.g.

(23) a. The ball ROLLED across the lawn.
b. The bird FLEW out of the nest.
c. John WALKED home.

Forced motion, on the other hand, involves atypical manners that may require a special effort, impediment, or high degree of motor control for performance, e.g.

(24) a. The bird HOPPED across the motorway.
b. John LIMPED home.

Finally, instrumental motion depicts manners necessitating vehicles or extra elements besides the figure's body for motion, eg

(25) a. Tim SLEDGED down the slope.
b. We SAILED across the Mediterranean.

Interestingly, forced and instrumental motion are more commonly encountered in instances of human motion, rather than animal, force or object motion.

Given that motion always involves a manner of displacement, we may thus draw an analytical distinction between default, forced, and instrumental types of motion – based on the manner of motion.

2.7 Summary

Motion types and the properties of motion-constituting components, such as paths and manners, display great variability, overall. In fact, this variability is so great that widely divergent situations are understood as instances of motion. At the same time, however, not one instance of motion may be thought of as prototypical motion, or motion *par excellence*. The rich variability of this domain in large part constitutes its complexity. It may thus be useful to partition the domain of motion into types of motion, not so much for the sake of analytical elegance, but because they may reflect distinct conceptual representations in cognisers' minds. The point is a simple one: motion is such a vast and diverse domain that conceptualising motion cannot be a straight-forward and unitary process, with the same schematic properties, e.g. path, receiving fixed levels of salience in conceptualisation, regardless of the motion type. The nature of the motion scene must in itself impact on the conceptualisation of its schematic properties, and this may be, as suggested above, relative to directional, aspectual, causal, agentive, or force dynamic properties of the scene, for instance.

Motion properties are not variable in a vacuum however. Instead, these properties often appear to interact with and constrain other properties. For instance, motion telicity constrains path types, cause-effect relations impact on path value, causality necessitates figure animacy, figure type constrains agency as well as manner types, and so on.

3 Figures as a basis for typological modelling

One central element of motion that seems to impact on most, if not all, types of motion and motion schemas consists of the figure performing, or undergoing, the motion. I will here use figure types as the basis for the elaboration of the conceptual typology, because figure properties make possible and also constrain types of motion (and motion properties). This section seeks to demonstrate how, for instance, the 'existential status' of figures defines the scope for motion types (e.g. Superman can fly across the Atlantic in a few minutes but my neighbour cannot), and how figures' physical properties limit the range of manners of motion (e.g. sharks can swim but pigs cannot), and also the types of path followed (e.g. some insects can naturally move on vertical surfaces but humans cannot normally do so). In addition, only animate figures capable of self-motion may cause the motion of external figures, so that figure type effectively impacts on causal potential. In short, I will seek to demonstrate that conceptualisation of motion, and possibly of other events, seems to be centered around figures, rather than grounds, paths, manners or causal motivations. By identifying the figure as the motion component determining other aspects of motion, I thereby suggest that any conceptual model of the domain of motion should therefore be based on, or centered around, the figure schema. In this section, I review how the properties of given figures interact and determine the various other components of motion, including motion types, as outlined in the previous section.

3.1 On figures and reality

Figure types may be distinguished relative to their 'existential status'. Indeed, the 'world out there' presents a perceptual reality as well as a fictional reality composed of fictional, or artificial, entities pertaining to popular and personal imagination. In this world, therefore, motion figures may be either real or fictional. Real figures, on the one hand, have physical existence: they are perceptually real and may be physically interacted with. They range across:

(i) humans, including babies, children, and adults, as well as their body parts;
(ii) animals, e.g. reptiles, birds, quadrupeds;
(iii) objects, e.g. bottles, bicycles, stools;
(iv) natural elements, e.g. water, fire;
(v) forces and currents, e.g. magnetism, electricity;
(vi) force-animated entities, e.g. planets, stars, currents, tornadoes.

Fictional figures, on the other hand, are man-made in the sense that they are created by human minds. Fictional figures are not perceptually real therefore. Instead, they exist in the world of fiction and may be found in myths, story books, cartoons, films, and so on. Fictional figures therefore consist of virtual creations, or objects of popular and individual imagination. Examples include:

(vii) humans, e.g. Cinderella, Charlie Chaplin, Mary Poppins, the Incredibles;
(viii) human-like figures, e.g. Homer Simpson, Hulk, Hobbits, fairies, Michelin Man;
(ix) human-like animals, e.g. Donald Duck, Nemo, Chicken Little, Bugs Bunny
(x) animals, e.g. Silvester the Cat, Coyote;
(xi) animal-like humans, e.g. werewolves, centaurs, Spiderman, Catwoman, mermaids;
(xii) objects, e.g. Star Trek Enterprise spaceship, Cinderella's mops and buckets;
(xiii) human-like objects, e.g. Thomas the Tank Engine, Christine the car, R2D2;
(xiv) artificial creations, e.g. Thing in the Addams family, Teletubbies, aliens, monsters, Pacman;
(xv) folk superstition figures, e.g. gods, ghosts, witches, angels.

Numerous fictional figures involve the selective blending of characteristics pertaining to different types of figures, e.g. human and animal, and to different representational formats for these figures, e.g. film versus cartoon (Fauconnier and Turner 2002). For instance, if we take the category of animal-like humans, we may note that mermaids and centaurs blend the physical physiognomy of both humans and animals in equivalent measure, whereas Spiderman and Catwoman retain the human physiognomy but integrate physiological properties of animals, such as spider or feline eye sight, claws, and so forth. Two points are interesting in these blends with regard to motion. First, the blending of physical properties entails new possibilities for figure motion. These possibilities may be enabling, such as Spiderman's ability to climb up vertical structures. They may also be constraining, such as mermaids' inability to walk on solid ground. In other words, the scope for motion types is redefined relative to the blended figure. The second point of interest here is that within one category of fictional figure, we obtain different selections of features in the resulting blend, as shown by the contrast between centaurs and Catwoman. This differential selectivity means that fictional figures are highly diverse and present a scope for motion types which is absent in the case of real-life motion as performed by real figures.

In sum, figures capable of undergoing motion in space are wide-ranging and highly diverse. A preliminary distinction may therefore be drawn between figure types that are perceptually 'real' and those that are man-made, either virtually, or in popular and/or individual imagination. From this distinction ensues one between *real motion* and *fictional motion*.

3.2 On figures and manners of motion/ force dynamics

One of the most obvious aspects of motion which figure type determines concerns likely manners of displacement. Real figures are constrained by their intrinsic physical properties with respect to manners of motion, e.g. humans cannot fly, babies cannot walk, dogs cannot tiptoe, birds cannot jog, balls cannot limp, books cannot roll, feathers cannot pounce, winds cannot jump, and so on. Interestingly, human figures appear to be the most flexible of all real figures and the most creative ones in terms of the types of manners they may and do use for motion. Indeed, humans may walk, run, jump, crawl, swim, dance, limp, and use instruments to drive, cycle, ice-skate, roller-blade, ski, sail, paraglide, windsurf, fly, sledge, and so on. Animals, on the other hand, appear to perform a much more limited range of manners of motion. This range typically excludes instrumental motion and is very restricted in terms of forced types of manners. Even our primate cousins do not seem capable of performing the multiplicity of manners of motion humans can perform, and it takes years to train animals to perform controlled manners, as is apparent from circus displays. Objects, natural elements and forces appear even further restricted in their potential for manner variability.

Fictional figures, on the other hand, have seemingly infinite scope for manners of motion. Most fictional figures have arguably become cultural icons with figure-specific potential for manner types, and this potential is at once enabling and restricting. This is very clear when contrasting figures pertaining to the same type. Take, for instance, fictional figures that are human, such as Cinderella, Mary Poppins and Charlie Chaplin. Only Cinderella seems to conform to prototypical real-life human manners of motion. Mary Poppins, on the other hand, also has the capacity to fly in the air. Finally, Charlie Chaplin's manner of motion is characterised by his duck-like walk, which prevents him from running or going up and down staircases in a typical adult-human fashion.

In sum, manner types are more restricted and less diverse for non-human real figures than for human figures, and they are almost infinite in scope for fictional figures. This point is important to conceptualisation because it entails that cognisers may have expectations regarding likely manners of displacement, and these expectations are more or less varied depending on the figure type. This means that motion conceptualisation in the case of fictional figures, for instance, is likely to be different in nature to that of real-life figures, as there may be no expectation concerning the range of possible motion types. Even if there were any expectations, these may be easily violated, given the unpredictable nature of fictional figures. In fact, one of the delights and reasons for the success of fictional motion scenes as displayed in films and cartoons is the very scope of possibilities for motion and the exploration of imaginary scenarios. For instance, a number of successful characters in cinematography are characterised by their non-veridical manners of motion, e.g. Superman, Wonder Woman, Spiderman, the Incredibles, Batman, Peter Pan, Mary Poppins, Jumbo. In fact, most fiction seems to rely on conceptual blends of figure properties, either granting human figures animal-like properties, e.g. flying, or giving animals and inanimates human properties, e.g. speech, bipedal motion, material culture, and so forth (Turner 1996).

3.3 On figures and animacy

As mentioned in the previous section, types of figures may be further distinguished relative to their animacy. Given the variety of figures outlined above, we may classify as animate the following categories of figures:

(xvi) real and fictional human and human-like
(xvii) real and fictional animal and animal-like
(xviii) fictional objects
(xix) fictional creations
(xx) natural elements and forces
(xxi) force-animated entities

Inanimate figures, on the other hand, comprise:

(xxii) real and fictional objects

The crucial distinction between animate and inanimate figures is that animates are capable of initiating their own motion, and that of other figures, whereas inanimates cannot perform self-motion. Inanimates may only undergo caused motion.

3.4 On figures and agency

Figure animacy has consequences for agency in motion. Indeed, agentive motion necessitates an animate figure with the capacity for self-motion. This capacity is underpinned by implicit instincts or explicit goals to either alter one's location in space or to undergo motion for the sake of it. In other words, agency entails some level of volition or intentionality in instantiating and undergoing motion. Hence, only figure types (xvi)-(xix) above may be agentive (e.g. humans, animals, fictional creations), but natural elements, forces and force-animated entities in the universe do not display agentive motion, in that sense, even though they may be capable of self-motion.

Given this understanding, non-agentive motion is therefore typical of non-sentient entities, including inanimates and animates such as natural elements and forces. Note, however, that non-agentive motion also occurs with animates when caused to move by an external figure or force. In this case, motion is not intentional and thus non-agentive. Consider, for instance:

(26) The ball rolled down the hill.

(27) The horse threw the rider flying off the saddle and into the air.

(26) illustrates an instance of inanimate object motion which is non-agentive. Indeed, the ball rolled down the hill, not of its own agency, but because it was set into motion by an external figure, such as an agentive human figure, or an external force, such as the wind, gravity, or other. In (27), on the other hand, we have two animate figures in motion; yet, only the primary figure, the horse, is agentive, whilst the secondary figure, the rider, is unwittingly undergoing motion as a result, and is thus non-agentive in its own motion.

Figure animacy may therefore be considered to correlate with properties of agency in motion, in that only animate motion may generate agentive motion, whereas inanimate motion can only be non-agentive. It does not follow, however, that all animate motion is agentive, as exemplified in (27). Likewise, the animate motion of natural elements and forces is not agentive. In other words, animacy is necessary for agency, but it is not sufficient. The agentive figure must also consist of a self-moving sentient entity. Animacy and agency are thus distinct notions constrained by figure properties and by the nature of the interactions between given figures.

3.5 On figures and causality

Figure animacy and agency have, in turn, consequences for causal properties in motion, in that agentive figures only may cause the motion or change of state of other figures, as in example (27) above. Also capable of causal motion are natural elements and forces. However, non-agentive figures can only undergo caused motion, as in (26). In addition, causality may be either overt, as in (27), where the primary motion-causing figure is conceptually salient, or it may be covert, as in (26), where the primary agent is not conceptually in focus, and yet it must be present for the motion act to take place. Note that linguistic mappings often eclipse the causal dimension of inanimate motion, as in (26), yet causality is physically real and is conceptually more or less overt as a result.

Because the agentive nature of the figure type impacts on (and constrains) causality, the figure type also impacts on path values. In agentive self-motion, the notion of path has an agentive value and it refers to spatial information, e.g.

(28) Jenny ran DOWN the hill.

The path is agentive in example (28) as it represents the path followed by the sole and primary agentive figure in this motion act: Jenny effectively descended the hill.

In causal motion, on the other hand, the path becomes resultative and it may refer to either spatial or stative information, e.g.

(29) The wind blew the napkin OFF the table.

(30) Mary tore the napkin TO SHREDS.

Indeed, in (29) and (30), the blowing and tearing actions exercised on the secondary motion figure, the napkin, result in a change-of-location path in (29) and a change-of-state path in (30).

In sum, the animate and agentive nature of the figure has a direct relation to causal potential in motion, and to path value. Only animate figures may generate causal motion together with the resultative path of a secondary figure.

3.6 On figures and directionality

As previously mentioned, motion may involve directionality or it may not. In the case of motion events, the directionality, or path, is conceptually salient. All kinds of figures undergo directed motion, either of their own doing, or as a result of a causal interaction with an animate figure, e.g.

(31) Mary walked to church this morning = animate and agentive motion

(32) The earth rotates around its axis = animate and non-agentive motion

(33) The book fell off the shelf = inanimate and non-agentive motion

(34) The door swung open = inanimate and non-agentive motion

On the other hand, non-directed motion, or motion activities, are more typical of animate figures. Indeed, inanimate figures do not readily undergo motion activities. Consider,

(35) Mary is driving = animate and agentive motion

(36) The wind is blowing = animate and non-agentive motion

(37) The buoy is floating = inanimate and non-agentive motion

(38) The ball is rolling = inanimate and non-agentive motion

Examples in (37) and (38) display instances of non-directed inanimate motion. These types of example appear to be few, and to depend importantly on the physical properties of the moving object. Indeed, non-directed, ongoing motion may be undergone by rolling items, for instance, but would be ad hoc with objects such as books, spoons, or baskets. I suggest that activities are atypical of inanimate figures because inanimate motion is by definition non-agentive, which means that inanimate motion must therefore be caused – whether overtly or covertly. As detailed above, caused motion typically makes salient the path followed by the secondary figure. This path may be either a change-of-location path, or a change-of-state path. In either case, the presence of a path is characteristic of motion events, and not activities. In other words, the causality entailed by inanimate motion leads to expectations of directionality. Conceptually, objects do not undergo non-directed motion for the sake of it, in the fashion of animates. So, it appears that motion activities are atypical with inanimate figures, though they are common in the case of animate figures, whether agentive or not.

3.7 On figures and telicity

Telicity relates to the reaching of endpoints in space or to resultative states. Consider,

(39) Mike drove to the supermarket.

(40) Jane kicked the ball into the net.

In the case of agentive figures, telicity importantly correlates with the figure's purpose in undergoing motion. This purpose may be to arrive at a specific location, or to cause a secondary figure to change its spatial or stative configuration. Given this understanding, telic events typically entail agency. However, non-agentive animates may also generate telic motion events, e.g.

(41) The fire spread across the entire forest.

(42) The wave engulfed the dinghy (under the water).

There is a correlation therefore between telic potential and animacy. Inanimates, on the other hand, undergo telic motion events by virtue of being set into motion by an animate figure, as in (40) and (42).

3.8 Summary

This section has argued that figure types determine motion types. The types of figures identified correspond to properties of animacy and agency, as well as artificiality and physicality. These fundamental properties constrain motion potential relative to aspects of manners and force dynamics, causality, path value, directionality, and telicity. This deterministic understanding of figure characteristics lends support to domain analyses of motion based on figures, rather than on any other motion-constituting element. Figures should therefore be the centre of any conceptual modelling of this domain. In order to demonstrate this point more fully, I now propose to offer cognitive evidence in support of the centrality of figures in motion conceptualisation.

4 Experimental evidence

Given the above-suggested distinctions concerning the diversity of motion types, the research question now becomes whether this diversity is conceptually real. That is, do these motion types correspond to actual conceptual categories of events? The aim of this section and of the studies it includes is to offer preliminary cognitive evidence in support of some of these distinctions in motion types. The evidence reviewed demonstrates that types of figures, paths, manners, and other motion properties (e.g. causality) influence event conceptualisation, and that they do so in non-random patterns of behaviour across speakers of distinct languages. In other words, this section offers data in support of a conceptual typology of motion which is valid regardless of the subject's native language, and which may therefore be utilised as a cross-linguistic tool for linguistic and cognitive research, and as a metalanguage of analysis.

4.1 Method

The aim of the present research is to show that a motion of type *x* is conceptualised as distinct from a motion of type *y*. In effect, the research seeks to show that motion types *x* and *y* constitute distinct conceptual categories, or classes, of motion. Given this objective, the experimental approach tests event categorisation and does so by using a sorting task in which subjects categorise motion stimuli together on the basis of perceived similarity. The stimuli consist of event triads in digital video format. These triads present a target motion stimulus, such as a man walking up a hill, followed by two alternate events which resemble the target, yet differ in one variable, such as the path or the manner of motion, as illustrated in (43).

(43)	TARGET:	a man walking up a hill	
	ALTERNATE 1:	a man running up a hill	(manner variable altered)
	ALTERNATE 2:	a man walking down a hill	(path variable altered)

This triad example displays an animate, real-life, human figure performing instances of self-motion in which the path is agentive and the manner is a default one for this type of figure. Should we intend to test for possible effects of figure types on motion conceptualisation, this type of triad would be used alongside other triads presenting similar events performed by differing types of figures, such as inanimate or fictional figures, e.g.

(44)	TARGET:	a ball rolling up a hill	
	ALTERNATE 1:	a ball bouncing up a hill	(manner variable altered)
	ALTERNATE 2:	a ball rolling down a hill	(path variable altered)

The triad in (44) is comparable to the triad in (43) insofar as the figure is real, its motion is seemingly independent from that of other figures, and the manner is of a default type for a ball. The main difference between (43) and (44) thus concerns the nature of the figure, which is human and animate in (43) and object and inanimate in (44).

By using several triads of types (43) and (44) in one experiment, it is possible to see whether subjects perform similar association choices for both types of triads (e.g. either in terms of path or manner), or whether they perform differently – and consistently so – for each triad type. Should subjects perform similarly, it may then be concluded that figure type does not cause different conceptualisation of motion. However, should subjects perform associations differently for the two types of triads, it may then be concluded that they are conceptualising the two types of motion – human and object motion – differently. In this latter case, we may infer that figure features cause distinctive motion conceptualisation.

Likewise, should we wish to test for possible effects of manner types on motion conceptualisation, triads of type (43) which displays default manners should be contrasted with triads displaying non-default manners, e.g.

(45)	TARGET:	a man limping up a hill	
	ALTERNATE 1:	a man tiptoeing up a hill	(manner variable altered)
	ALTERNATE 2:	a man limping down a hill	(path variable altered)

Again, should default and non-default motion be conceptualised differently, we would expect subjects to perform different categorisation choices for each triad type.

This experimental set-up was implemented in a number of studies to test the effects of figures, paths, manners and agentive properties of motion on event conceptualisation. Table 1 below details the properties of motion examined, together with the types and sub-types contrasted in the triadic stimuli.

Table 1. Motion properties and types examined

Motion property	Type	Sub-type	Example of motion situation
Figure	Real	Animate	Human person
		Inanimate	Basket ball
	Fictional	Animate	Virtual tomato
Motion agency	Self-motion		A person jumping over a wall
	Causal motion		A person kicking a door open
Path	Atelic		Along, up, down
	Telic		Across, into, out
Manner	Default		Rolling (ball), walking (person)
	Non-default	Forced	Limping (person)
		Instrumental	Cycling (person)

Experiments were implemented using speakers of different native languages in order to ensure that performance is independent of linguistic patterns, and to assess the comparability and reliability of findings across different language populations. In the exercise of extrapolating a *conceptual* typology, it was deemed particularly important to obtain robust findings across language populations, rather than findings subject to local fluctuations. Should categorisation behaviour be similar across different linguistic populations, we may suggest that the categories obtained are strongly indicative of universal trends in the human conceptualisation of motion.

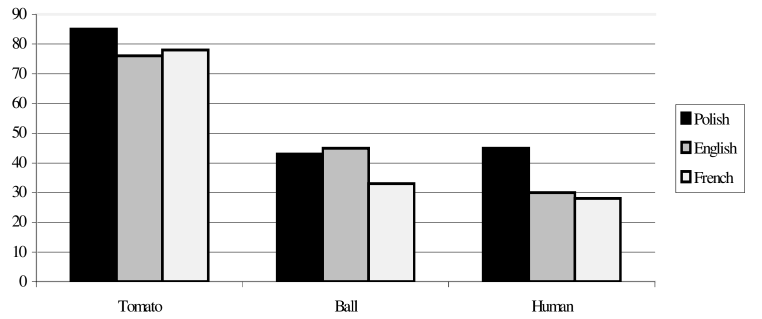
4.1.1 Kopecka and Pourcel (2005, 2006)

Kopecka and Pourcel (2005, 2006) tested the effects of figure type on event conceptualisation with 69 subjects representing the following native populations: Polish (N=24), English (N=21) and French (N=24). As detailed in Table 1, three distinct figures were used:

(i) 'Tomatoman' [- real] [+ animate] [- human]
(ii) Human person [+ real] [+ animate] [+ human]
(iii) Object ball [+real] [- animate] [- human]

The stimuli consisted of the now-famous virtual tomato designed by the Max Planck Institute for Psycholinguistics. This tool represents a two-dimensional computer animation with agentive properties, for instance, the tomato has eyes and it smiles, and importantly it is capable of self-motion. Tomatoman corresponds to a fictional figure – though one that most subjects would not be familiar with and may therefore not have a priori expectations for in terms of motion potential. In addition, the stimuli included digital triads of a man performing motion instances, and other triads presented the motion of a real-life plastic ball (yet note that no motion-causing agent was visible for the performance of the ball's motion).

Each triad contrasted two default manners of motion relative to the figure properties and two paths of displacement, as exemplified in triads (43) and (44) above. Association choices were therefore made either in terms of path or manner similarity. The results, shown in Graph 1, indicate that all three language groups – Polish, English and French – performed very similarly in the task. This similarity suggests robust trends for a universal level of conceptualisation. In addition, the results show that subjects categorise the motion stimuli differently depending on the nature of the figure. The virtual tomato, for instance, yields categorisation choices based on the manner variable to an 80 percent extent. On the other hand, the object figure yields a 40 percent only preference for manner, whilst the human figure averages a 34 percent preference for manner. In other words, the fictional tomato figure prompts high manner salience in conceptualisation, whereas the real-life figures prompt higher path salience, with path salience being higher in the case of human figures than in the case of object figures (Mann-Whitney U-test, p_E=0.0004, p_P=0.0002, p_F=0.0002 for Tomato-Human scores; p_E=0.003, p_P=0.005, p_F=0.003 for Tomato-Ball scores).[4] The three figure types thus trigger distinctive conceptual categorisation, with the most notable difference between fictional and real-life figures – over 35 percent. We may therefore suggest that fictional and real-life motion constitute distinct conceptual categories of events.



Graph 1. Proportions of manner association choices in the figure studies

In addition, these results may be explained in terms of the interaction between motion variables. Indeed, on the one hand, fictional figures unknown to subjects (such as virtual tomatoes) fail to have default manners of motion, that is, manners known to and expected by the participant. Any manner of displacement is therefore likely to be conceptually salient to the cogniser in the case of fictional motion. This salience is likely to decrease when subjects are conceptualising the motion of known entities, such as real-life objects and persons. On the other hand, the path of motion informs us of the destination the figure eventually reaches. In the case of a human figure, that destination is typically intended and corresponds to the figure's goal, or purpose of motion. In human motion, path thus conceptually correlates with the figure's intentions. This is so because human figures are animate and agentive. Their actions are typically not random, but purposeful, and much of human interaction comes down to deciphering other people's goals and intentions. This much is part of the subject's knowledge and it may explain why subjects found path the most conceptually salient variable in human motion. This categorisation bias towards the goal-loaded dimension of motion is likely to decrease in cases where the figure is not intentioned, and where paths do not correspond to actual goals. This may partly explain the low salience of path in the conceptualisation of tomato motion (in addition to the non-default nature of the manners displayed, as just discussed). This suggestion would also predict low levels of path salience in the conceptualisation of object motion. However, the data indicate a 60 percent bias towards path in object motion conceptualisation (at least in the present experiment). In this case, it may be suggested that the real-life nature of the figure caused the subjects to infer the presence of an agentive force behind the motion of the objects shown (though that agent was not visible in the stimuli). Indeed, plastic balls rarely – if ever – bounce across bridges and other types of ground of their own volition. The inference of a causal element behind the object's motion may have rendered the object path relatively salient. In addition, note that the balls used in the stimuli displayed default rolling and bouncing manners of motion, which might not be expected to trigger manner salience as a result. These various factors may thus add up to a relatively low level of manner salience across the findings.

In sum, these preliminary suggestions support the idea of a correlation between real-life animacy and goal-directed behaviour (see also Mandler 2004). They also show that there are complex interactions between motion variables in event conceptualisation. For instance, the default nature of manners seems to decrease attention to manner in favour of path, whereas unexpected and non-default manners direct attention towards that variable to the detriment of path. The following study seeks to clarify the validity of these suggestions by examining path and manner types, as well as causal relations more closely.

4.1.2 Pourcel (2004, 2005)

Pourcel (2004, 2005) tested the effects of path type, manner type and causality on event conceptualisation with 69 subjects representing the following native populations: English (N=34) and French (N=35). The stimuli consisted of fifteen video triads using a human figure to perform motion instances.

As detailed in Table 1, two types of path were displayed in the stimuli:

(i) Atelic paths showing no trajectory endpoint or crossed boundaries,
(ii) Telic paths showing explicit endpoints or crossed boundaries.

Three types of manners were displayed:

(iii) Default manners, such as walking or running casually,
(iv) Non-default manners showing heightened degrees of control for performance, or an impediment, such as tiptoeing or limping,

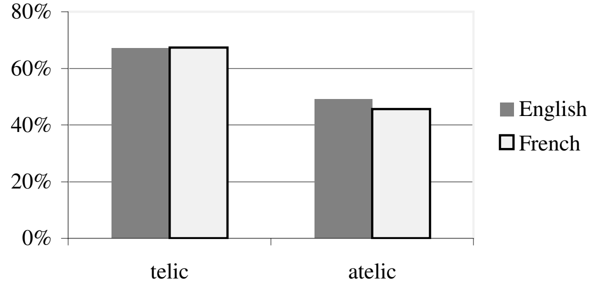
(v) Non-default manners showing the use of a vehicle for the performance of motion, such as a bicycle or a scooter.

Finally, the stimuli showed instances of self-motion and of causal motion:

(vi) Self-motion showing the intentional motion of an agentive figure,
(vii) Causal motion showing the motion of an agentive figure generating the motion of a non-agentive figure, such as an object.

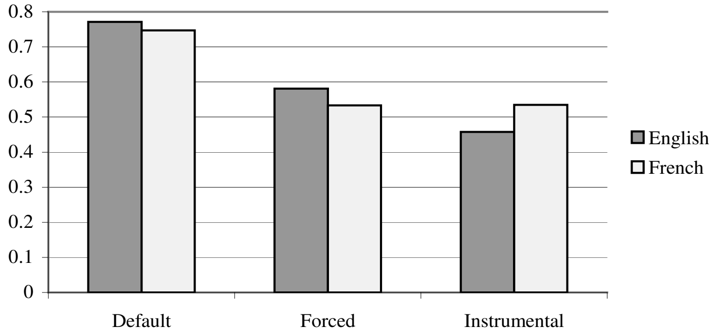
As in the previous experiments, each triad contrasted two manners of motion and two paths of displacement. Association choices were therefore made either in terms of path or manner similarity. The results, shown in Graphs 2, 3 and 4, indicate that the two language groups performed very similarly in the task. This similarity is again suggestive of universal trends in motion conceptualisation. In addition, the graphs show that subjects categorise the motion stimuli differently depending on the nature of the path, manner, and causal properties displayed in the scenes.

In the case of path properties, Graph 2 illustrates how the presence of a path end-point encourages path-salient conceptualisation of events (almost 70 percent of choices were made relative to path similarity). On the other hand, the absence of telicity does not reverse path salience in favour of manner salience. Instead, it reduces path-salient conceptualisation, and we obtain mixed performance in conceptualisation in atelic cases, with other factors likely to interfere with variable salience, such as manner type and causal relations. What is apparent, then, from Graph 2 is a clear difference in the conceptualisation of telic and atelic events – a 20 percent difference in associative performance (Wilcoxon test, p_E=0.001, p_F<0.001). We may suggest therefore that the two types of events are conceptually distinct.



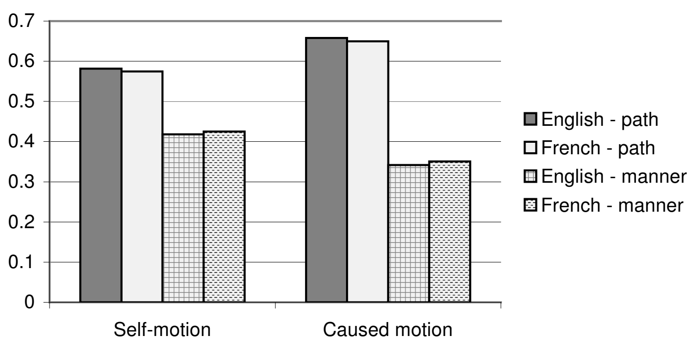
Graph 2. Proportions of path association choices relative to path telicity

In the case of manner properties, Graph 3 shows that default manners of motion prompt path-salient conceptualisation of events (76 percent of choices overall were made relative to path similarity). On the other hand, non-default manners, involving either higher control, impediment, or vehicle, reduce path salience by about 20 percent. This difference in associative performance is indicative of differential conceptualisation across default-manner motion and non-default instances (Wilcoxon test, p_E=0.001, p_F=0.0001 for a comparison of default and forced scores; p_E<0.0001, p_F=0.001 for default and instrumental scores). Note, however, that the distinction is marginal between forced and instrumental manner types (Wilcoxon test, p_E=0.069, p_F=0.781). We may thus posit two broad types of manners that are conceptually distinct, namely default and non-default types.



Graph 3. Proportions of path association choices relative to manner type

In the case of causal properties, Graph 4 demonstrates a slight difference in associative performance. On average, 65 percent of caused motion instances prompt path salience, in contrast to 58 percent of self-motion instances. We may suggest therefore a tendency towards path salience when causal relations are apparent between the motion of two or more figures. In addition, note that this analysis does not sort responses relative to path and manner types, which suggests that differences between self- and caused motion may be more pronounced in more controlled conditions, for instance, in conditions where all manners were default and all paths were telic, and where the sole contrast between the triads would be relative to causality. Although the score differences are not quite as marked as in previous tests (Wilcoxon test, p_E=0.067, p_F=0.029), the preferential tendency for path salience in caused motion is obvious when we contrast the differences between path and manner associations in each type of triad. This difference averages 15 percent in self-motion and 30 percent in caused motion.

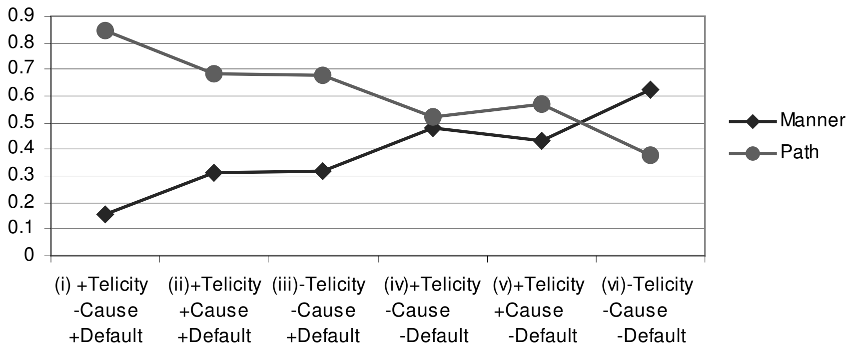


Graph 4. Proportions of path and manner association choices relative to causality

The above analyses do not isolate each motion variable neatly and are therefore likely to be skewed by the effect of other variables. These graphs are thus suggestive of generic trends. These trends make it possible to predict conceptual salience for particular types of motion, based on path telicity, manner defaultness, and motion causality. It has so far been demonstrated that telicity, default manners and causal relations prompt higher degrees of path salience, than atelicity, non-default manners and instances of self-motion. We may therefore hypothesise the following levels of conceptual salience for the following events:

(i) [+telicity] [-causality] [+default] = high path salience
(ii) [+telicity] [+causality] [+default] = high path salience
(iii) [-telicity] [-causality] [+default] = high path salience (though lower than in (i) and (ii))[5]
(iv) [+telicity] [-causality] [-default] = mixed path and manner salience
(v) [+telicity] [+causality] [-default] = mixed path and manner salience
(vi) [-telicity] [-causality] [-default] = low path salience

Subject responses were analysed according to these predictions. The analysis supports these predictions, as shown in Graph 5.



Graph 5. Proportions of path and manner association choices relative to telicity, causality, and default manner properties[6]

Graph 5 presents a more fine-grained analysis of the data, and in so doing, it usefully illustrates the scope of conceptual variability across different motion types. For instance, path is more salient than manner to an 85 percent extent in motion events of type (i), such as a walking-across-a-road event, whereas that salience drops by 50 percent down to 38 percent in motion events of type (vi), such as a tiptoeing-up-a-staircase event. We may therefore conclude that the nature of motion properties impacts in significant ways on the conceptualisation of events.

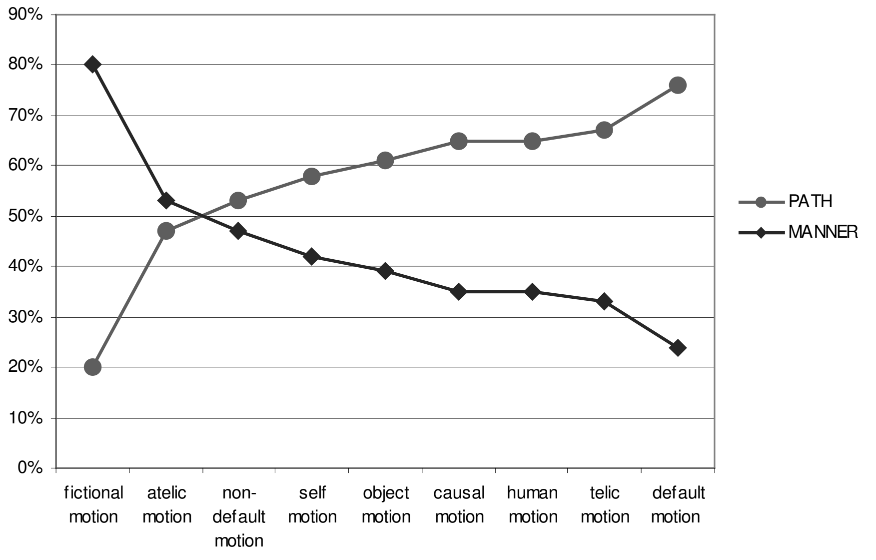
The conceptual distinctions observed may be explained in similar terms to the ones highlighted in the discussion of the figure studies by Kopecka and Pourcel (2005, 2006). That is, differential conceptualisation of events may be explained in terms of (a) the cogniser's expectations of likely manners of displacement relative to a given figure, and (b) the explicitness of the figure's goals and intentions in performing motion. Indeed, default manners of motion attract little attention to the manner variable, as illustrated in motion types (i)-(iii), whilst the overtness of the figure's locational goal in telic events renders the path dimension of motion particularly salient, as exemplified in motion types (i)-(ii) in particular. Likewise, causality entails an intention on behalf of the agentive figure to achieve a particular result, and as such it denotes a purposeful action akin to the notion of telicity. Finally, it is interesting to note that these properties interact in the overall conceptualisation of an event, so that telic events performed with a non-default manner of motion (e.g. types (iv) and (v)) yield mixed path-manner biases in conceptualisation. The conceptualisation of motion events is thus neither straightforward nor unitary, but highly dynamic and dependent upon discrete properties of motion. Motion conceptualisation is not random, however, but consistent and predictable, and it may be explained in terms of perceived intentionality and expectations based on the cogniser's knowledge of the world, as suggested above. It is equally noteworthy to recall that performance on the present cognitive tasks was strikingly similar across all 69 subjects, including subjects of different native languages. The present conceptual trends may thus be indicative of universal tendencies in human cognition.

4.1.3 Summary

This section asked whether conceptualisation behaviour supports the distinctions and categories proposed earlier in this article. The present experiments examined a number of motion properties, including figure animacy, figure reality, path telicity, manner defaultness, motion agency and causality. Results indicate that each of these fundamental property influences motion conceptualisation in significant ways. We may suggest, as a result, that the motion types discussed constitute distinct conceptual classes of events. These conceptual classes include[7]:

real-life motion
animate motion
inanimate motion
fictional motion
telic motion
atelic motion
default motion
non-default motion
self-motion
causal motion

Path- and manner-based conceptual distinctions in event classification between these types of motion are summarised in Graph 6.



Graph 6. Proportions of path and manner association choices relative to motion type

5 Sketching a conceptual typology of motion

This study has reported preliminary evidence for conceptual distinctions between

real-life and fictional figures
animate and inanimate figures
default and non-default manners of figure motion
telic and atelic events relating to figure goals
self-initiated and caused motion

These motion types are based on properties relating to the figure, including basic existential properties (e.g. real/ fictional, animate/ inanimate), physical properties (e.g. bodily capacity for performing given manners, relative strength for causing the motion of other figures), and cognitive properties (e.g. goals and intentions). It appears that motion is fundamentally about the figure that performs or undergoes it. Motion may thus be modelled as a figure-centred domain, and I propose to devise a figure-centred conceptual typology for the domain of motion, as a result. Table 2 offers a summary of the possible types of motion, based on figure properties.

Table 2. Typological modelling of motion, based on figure properties.

FIGURE TYPE				MOTION TYPE				
	Animacy	Agency	Figure type	Motion	Path	Directionality	Telicity	Force dynamics
REAL FIGURES	animate							
		Agentive						
			human	self / causal	agentive	events / activities	telic / atelic	default / non-default
			animal	self / causal	agentive	events / activities	telic / atelic	default / non-default
		non-agentive						
			natural element	self / causal/ caused	mechanical/ resultative	events / activities	telic / atelic	default / non-default
			human	caused	resultative	events	telic	default / non-default
			animal	caused	resultative	events	telic	default / non-default
	inanimate							
			object	caused	resultative	events	telic	Default
FICTIONAL FIGURES	animate							
		Agentive						
			human	self / causal	agentive	events / activities	telic / atelic	default / non-default
			human-like	self / causal	agentive	events / activities	telic / atelic	default / non-default
			animal	self / causal	agentive	events / activities	telic / atelic	default / non-default
			animal-like	self / causal	agentive	events / activities	telic / atelic	default / non-default
		non-agentive						
			natural element	self / causal	mechanical/ resultative	events / activities	telic / atelic	Default
			human	caused	resultative	events	telic	default / non-default
			human-like	caused	resultative	events	telic	default / non-default
			animal	caused	resultative	events	telic	default / non-default
			animal-like	caused	resultative	events	telic	default / non-default
	inanimate							
			object	caused	resultative	events	telic	Default

6 Conclusion

This research was initially triggered by methodological concerns in relativistic investigations of the relationship between language and cognition. Paramount among these concerns is the need to avoid linguacentric analyses of experiential domains. Motion, one of the most extensively used domain in contemporary relativistic research, remains to this day largely characterised in linguistic and typological terms. That is, little linguistic research has sought to acquire a conceptual understanding of this domain and to develop a neutral metalanguage for the analysis of motion and for the application of these analyses. This research does not pretend to develop one such metalanguage. Rather, it seeks to direct awareness towards the necessity to develop conceptual typologies and analytic metalanguages systematically in answer to calls for scientific rigour in multi-disciplinary research (e.g. Lucy 1992, 2003). In so doing, it is hoped that this research will trigger further inquiries into the conceptualisation of motion, which may in turn be of use to linguistic applications. Such inquiries would be useful in terms, for instance, of additional experimental research bringing to the fore new cognitive data to establish the reality of the categorial distinctions suggested. Indeed, the present research has offered data in support of a certain number of distinctions. More data is needed to substantiate the distinction between event and activity, for instance. According to the analysis reported in this article, experimental hypotheses would predict the following in the case of human figures: activities generate manner salience, whereas events generate path salience. In addition, more data is required to support the distinction between caused and self-motion events, as well as between further motion types based on figure characteristics, e.g. animal vs. human motion. Data is also needed to reach a better understanding of the interactions between motion variables in conceptualisation. It was suggested that motion variables do not occur in a vacuum, but instead, interact and constrain other variables in systematic ways, which are far from transparent at this juncture. Cross-linguistic data is also required to offer conclusiveness to the conceptual typology proposed, so far only supported by data from English, Polish and French subjects. Further support is also required to fully substantiate the suggestion that intentionality, for instance, is an explanatory factor for path salience in human motion conceptualisation. Finally, it would also be key to explore domain extensions and their conceptual reality, e.g. fictive motion, or the domain of TIME, and so forth.

References

Aske, J. (1989) Path predicates in English and Spanish: A closer look. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society* 1–14.

Fauconnier, G. and Turner, M. (2002) *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* New York: Basic Books.

Gennari, S. P., Sloman, S. A., Malt, B. C. and Fitch, W. T. (2002) Motion events in language and cognition. *Cognition* 83: 49–79.

Hohenstein, J. (2005) Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development* 6: 403–425.

Kopecka, A. and Pourcel, S. (2005) Figuring out figures' role in motion conceptualisation. Paper presented at the 9th International Cognitive Linguistics Conference, Seoul, Korea.

Kopecka, A. and Pourcel, S. (2006) Understanding figures in conceptualising motion. Paper presented at the 2nd Biennial Conference on Cognitive Science, St Petersburg, Russia.

Lucy, J. A. (1992) *Language Diversity and Thought*. Cambridge: Cambridge University Press.

Lucy, J. A. (2003) Semantic accent and linguistic relativity. Manuscript.

Mandler, J. (2004) *The Foundations of Mind*. Oxford: Oxford University Press.

Papafragou, A., Massey, C. and Gleitman, L. (2002) Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84: 189–219.

Pourcel, S. (2004) Motion in language and cognition. In A. Soares da Silva, A. Torres and M. Gonçalves (eds) *Linguagem, Cultura e Cognição: Estudos de Linguistica Cognitiva* Vol. 2 75–91. Coimbra: Almedina.

Pourcel, S. (2005) Relativism in the linguistic representation and cognitive conceptualisation of motion events across verb-framed and satellite-framed languages. Unpublished doctoral dissertation, University of Durham, UK.

Pourcel, S. (2009a) Relativistic application of thinking for speaking. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura and S. Özçalişkan (eds) *Crosslinguistic Approaches to the Study of Language: Research in the Tradition of Dan Isaac Slobin* 493–503. Hove: Psychology Press.

Pourcel, S. (2009b) Motion scenarios in cognitive processes. In V. Evans and S. Pourcel (eds) *New Directions in Cognitive Linguistics* 371–391. Amsterdam: John Benjamins.

Pourcel, S. and Kopecka, A. (2005) Motion expression in French: Typological diversity. *Durham and Newcastle Working Papers in Linguistics* 11: 139–153.

Pourcel, S. and Kopecka, A. (To appear) Motion events in French: Typological intricacies. *Linguistic Typology.*

Slobin, D. I. (1996) From 'thought and language' to 'thinking for speaking'. In J.J. Gumperz and S.C. Levinson (eds) *Rethinking Linguistic Relativity* 70–96. Cambridge: University Press.

Slobin, D.I. (2000) Verbalised events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier and R. Dirven (eds) *Evidence for Linguistic Relativity* 107–138. Amsterdam: John Benjamins.

Slobin, D. I. (2004) The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist and L. Verhoeven (eds) *Relating Events in Narrative* Vol. 2 219–257. Mahwah, NJ: LEA.

Talmy, L. (1985) Lexicalisation patterns: Semantic structure in lexical forms. In T. Shopen (ed.) *Language Typology and Syntactic Description* Vol. 3 57–149. Cambridge: Cambridge University Press.

Talmy, L. (1991) Path to realisation: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society* 480–520.

Talmy, L. (2000) *Toward a Cognitive Semantics: Typology and Process in Concept Structuring* Vol. 2. Cambridge, MA: MIT Press.

Turner, M. (1996) *The Literary Mind*. Oxford: Oxford University Press.

Whorf, B.L. (1956) *Language, Thought, and Reality*. Cambridge, MA: MIT Press.

Zlatev, J. and David, C. (2004) Do Swedes and Frenchmen view motion differently? Conference Presentation, Language, Culture and Mind, Portsmouth.

Zlatev, J. and David, C. (2005) Motion event typology and categorisation. Conference Presentation, International Cognitive Linguistics Conference, Seoul.

Notes

1 Note again that the examples here and below seek to represent event situations. The examples, unless otherwise stipulated, are not linguistic examples. Instead they seek to prompt a conceptual image of a particular event. It is in this non-linguistic frame of mind that examples are used and discussed in this article.

2 Note that this point merely serves as an illustration, but is not used as an argument for the conceptual distinctiveness of motion events vs. motion activities. As outlined before, a conclusion of this type would be linguacentric, not to mention that it would also ignore the plethora of satellite-framed languages which cannot do away with manner verbs, regardless of the conceptual salience of the manner and path schemas in the said events.

3 Note that the latter two entities may be posited as being animate despite lacking intentionality. The contrast between animacy and intentionality is discussed further in the next section on agency.

4 The initial following the p score corresponds to the language group being tested. E for English, P for Polish and F for French.

5 Note that the present stimuli did not include atelic caused motion. Such motion events exist nonetheless, e.g. pushing a pram along the pavement. The prediction may be one of high path salience, though lower than that expected in (ii) for instance, due to the lack of telicity.

6 Note that due to the high comparability of the French and English data, responses from the two groups were conflated in this analysis, for clarity of presentation.

7 Note that the list is by no means exhaustive, but is based on the experimental findings reported in this article. Additional classes may include artificial animate motion, artificial inanimate motion, real-life animal motion, real-life animal non-default motion (e.g. motion as performed by animals in circuses), human sports motion (as opposed to everyday motion), force motion, and so forth.

Part VIII

The relation between space, time and modality

17 Space for thinking

Daniel Casasanto

1 Introduction

How do people think about things they can never see or touch? The ability to invent and reason about domains such as time, ideas, or mathematics is uniquely human, and is arguably the hallmark of human sophistication. Yet, how people mentally represent these abstract domains has remained one of the mysteries of the mind. This chapter explores a potential solution: perhaps the mind recruits old structures for new uses. Perhaps sensory and motor representations that result from physical interactions with the world (e.g., representations of physical space) are recycled to support abstract thought. This hypothesis is motivated, in part, by patterns observed in language: in order to talk about abstract things, speakers often recruit metaphors from more concrete or perceptually rich domains. For example, English speakers often talk about time using spatial language (e.g., a *long* vacation; a *short* meeting). Cognitive linguists have argued such expressions reveal that people conceptualize abstract domains like time metaphorically, in terms of space (see Lakoff and Johnson, 1999; c.f., Evans, 2004). Although linguistic evidence for Metaphor Theory is abundant, the necessary nonlinguistic evidence has long been elusive; people may *talk* about time using spatial words, but how can we know whether people really *think* about time using mental representations of physical space?

This chapter describes a series of experiments that evaluate Metaphor Theory as an account of the evolution and structure of abstract concepts and explore relations between language and nonlinguistic thought, using the abstract domain of time and the relatively concrete domain of space as a testbed. Hypotheses about the way people mentally represent space and time were based on patterns in metaphorical language, but were tested using simple psychophysical tasks with nonlinguistic stimuli and responses. Results of the first set of experiments showed that English speakers incorporate irrelevant spatial information into their estimates of time (but not vice versa), suggesting that people not only talk about time using spatial language, but also think about time using spatial representations. The second set of experiments showed that (a) speakers of different languages rely on different spatial metaphors for duration, (b) the dominant metaphor in participants' first languages strongly predicts their performance on nonlinguistic time estimation tasks, and (c) training participants to use new spatiotemporal metaphors in language changes the way they estimate time. A final set of experiments extends the experimental techniques developed to explore mental representations of time to the domain of musical pitch. Together, these studies demonstrate that the metaphorical language people use to describe abstract ideas provides a window on their underlying mental representations, and also shapes those representations. The structure of abstract domains such as time appears to depend, in part, on both linguistic experience and on physical experience in perception and motor action.

1.1 Time as an abstract domain

> *For what is time? Who can readily and briefly explain this? Who can even in thought comprehend it, so as to utter a word about it?*
>
> *If no one asks me, I know: if I wish to explain it to one who asketh, I know not.*
>
> Saint Augustine, *Confessions*, Book 11

How long will it take you to read this chapter? The objective time, as measured by the clock, might depend on whether you're scrutinizing every detail, or just skimming to get the main ideas. The subjective time might vary according to physiological factors like your pulse and body temperature (Cohen, 1967; Ornstein, 1969), psychological factors like how much the text engages your interest and attention (Glicksohn, 2001; James, 1890; Zakay and Block, 1997), and some surprising environmental factors like the size of the room you're sitting in (DeLong, 1981).

Although subjective duration is among the earliest topics investigated by experimental psychologists (Mach, 1886), the cognitive sciences have yet to produce a comprehensive theory of how people track the passage of time, or even to agree on a set of principles that consistently govern people's duration estimates. An excerpt from a review by Zakay and Block (1997) illustrates the current state of confusion:

> People may estimate filled durations as being longer than empty durations, but sometimes the reverse is found. Duration judgments tend to be shorter if a more difficult task is performed than if an easier task is performed, but again the opposite has also been reported. People usually make longer duration estimates for complex than for simple stimuli, although some researchers have found the opposite. (pg. 12)

What makes time perception so difficult to understand? Ornstein (1969) argues that although we *experience* the passage of time, the idea that time can be *perceived* through the senses is misleading (cf. Evans, 2004):

> One major reason for the continuing scattering of [researchers'] effort has been that time is treated as if it were a sensory process. If time were a sensory process like vision…we would have an 'organ' of time experience such as the eye. (pg. 34)

Although time is not something we can see or touch, we often talk about it as if it were (Boroditsky, 2000; Clark, 1973; Gruber, 1965; Jackendoff, 1983; Lakoff and Johnson, 1980). Consider the following pair of sentences:

i) They moved the truck forward two meters.

ii) They moved the meeting forward two hours.

The truck in sentence i is a physical object which can move forward through space, and whose motion we might see, hear, or feel, from the staring point to the ending point. By contrast, there is no literal motion described in sentence ii. The meeting is not translated through space, and there is no way to experience its 'movement' through time via the senses. Events that occur in time are more abstract than objects that exist in space insomuch as we typically have richer perceptual evidence for the spatial than for the temporal.[1]

In this chapter, I will argue that (a) the language people typically use to talk about duration reveals important links between the abstract domain of time and the relatively concrete domain of space, (b) people use spatial representations to conceptualize time even when they're not using language, and (c) although the domains of space and time provide a particularly useful testbed for hypotheses about the evolution and structure of abstract concepts, *time* is only one of many abstract domains of knowledge that depend, in part, on perceptuo-motor representations built up via experience with the physical world.

1.2 Metaphor and the problem of abstract thought

The mystery of how people come to mentally represent abstract domains such as time, ideas, or mathematics has engaged scholars for centuries, sometimes leading to proposals that seem unscientific by modern standards. Plato (*Meno*, ca. 380 B.C.E.) argued that we cannot acquire abstract concepts like *virtue* through instruction, and since babies are not born knowing them, it must be that we recover such concepts from previous incarnations of our souls. Charles Darwin contended that evolution can explain the emergence of abstract thought without recourse to reincarnation, yet it is not immediately obvious how mental capacities that would have been superfluous for our Pleistocene forebears could have been selected for. What selection pressures could have resulted in our ability to compose symphonies, invent calculus, or imagine time travel? How did foragers become physicists in an eyeblink of evolutionary time? The human capacity for abstract thought seems to far exceed what could have benefited our predecessors, yet natural selection can only effect changes that are immediately useful. The apparent superfluity of human intelligence drove Alfred Wallace, Darwin's co-founder of the theory of evolution by natural selection, to abandon their scientific theory and invoke a divine creator to explain our capacity for abstract thought (Darwin, 1859/1998, 1874/1998; Gould, 1980; Pinker, 1997; Wallace, 1870/2003).[2]

Darwin's own formulation of evolutionary theory points toward an elegant potential solution to Wallace's dilemma: sometimes organisms recycle old structures for new uses. An organ built via selection for a specific role may be fortuitously suited to perform other unselected roles, as well. For example, the fossil record suggests that feathers were not originally 'designed' for flying. Rather, they evolved to regulate body temperature in small running dinosaurs, and were only later co-opted for flight (Gould, 1991). The process of adapting existing structures for new functions, which Darwin (1859/1993) gave the misleading name *preadaptation*, was later dubbed *exaptation* by evolutionary biologist Steven Jay Gould and colleagues (1982). Gould argued that this process may explain the origin of many biological and psychological structures that direct adaptation cannot.

Are abstract concepts like dinosaur feathers? Can exaptation account for mental abilities in humans that could not have been selected for directly? If so, how might this have happened: which *ad*apted capacities might abstract domains be *ex*apted from? Steven Pinker (1997) sketched the following proposal:

> Suppose ancestral circuits for reasoning about space and force were copied, the copies' connections to the eyes and muscles were severed, and references to the physical world were bleached out. The circuits could serve as a scaffolding whose slots are filled with symbols for more abstract concerns like states, possessions, ideas, and desires. (pg. 355)

As evidence that abstract domains arose from circuits designed for reasoning about the physical world, Pinker appeals to patterns observed in language. Many linguists have noted that when people talk about states, possessions, ideas, and desires, they do so by co-opting the language of intuitive physics (Clark, 1973, Gibbs, 1994; Gruber, 1965; Jackendoff, 1983; Lakoff and Johnson, 1980; Langacker, 1987; Talmy, 1988). In particular, words borrowed from physical domains of space, force, and motion, give rise to linguistic metaphors for countless abstract ideas. For each pair of expressions below, *l* illustrates a literal use and *m* a metaphorical use of the italicized words.

1*l* a *high* shelf
1*m* a *high* price

2*l* a *big* building
2*m* a *big* debate

3*l* *forcing* the door
3*m* *forcing* the issue

4*l* *pushing* the button
4*m* *pushing* the limit

5*l* *keeping* the roof *up*
5*m* *keeping* appearances *up*

The concrete objects described in the literal sentences (e.g., shelf, building, door, button, roof) belong to a different ontological category than the abstract entities in the metaphorical examples, according a test of what physical relations they can sensibly be said to enter into. For example, it is sensible to say 'the cat sat on the shelf / building / door

/ button / roof', but it may not be sensible to say that 'the cat sat on the price / debate / issue / limit / appearance'. This test is similar to a test of sensible predicates for concrete vs. abstract entities devised by Fred Sommer (1963; cf., Turner, 2005).

Based on examples like these, linguists have argued that people create abstract domains by importing structure from concepts grounded in physical experience. Although anticipated by others (e.g., Lafargue, 1898/1906), this idea appears to have been first articulated as the Thematic Relations Hypothesis (TRH) in 1965, by Jeffery Gruber. TRH was later elaborated by Jackendoff (1972; 1983) who wrote:

> The psychological claim behind [Gruber's linguistic discovery] is that the mind does not manufacture abstract concepts out of thin air…it adapts machinery that is already there, both in the development of the individual organism and in the evolutionary development of the species. (1983, pg. 188–9)

Not all theorists agree on the significance of metaphorical language for theories of mental representation. Gregory Murphy (1996; 1997) raised concerns about both the vagueness of the psychological processes suggested by linguists and about the limitations of purely linguistic evidence for metaphorical conceptual structure. Murphy (1996) proposed that linguistic metaphors may merely reveal similarities between mental domains: not causal relationships. Across languages, people may use the same words to talk about space and time because these mental domains are structurally similar, and are therefore amenable to a common linguistic coding. He argued that in the absence of corroborating nonlinguistic evidence, his *Structural Similarity* proposal should be preferred on grounds of simplicity. His view posits that all concepts are represented independently, on their own terms, whereas the metaphorical alternative posits complex concepts that are structured interdependently. It is evident that people *talk* about abstract domains in terms of relatively concrete domains, but do they really *think* about them that way?

1.3 From conceptual metaphor to mental metaphor

The idea that conventionalized metaphors in language reveal the structure of abstract concepts is often associated with Conceptual Metaphor theory, proposed by linguist George Lakoff and philosopher Mark Johnson (1980, 1999). Lakoff and Johnson described 'conceptual metaphors' as one of 'three major findings of cognitive science' (1999, pg. 3). Yet, their claim that people *think* metaphorically was supported almost entirely by evidence that we *talk* metaphorically. Despite the impressive body of linguistic theory and data that Lakoff and Johnson summarized (and the corroborating computational models of word meaning), they offered little evidence that the importance of metaphor extends beyond language. In the absence of nonlinguistic evidence for metaphorically structured mental representations, the idea that abstract thought is an exaptation from physical domains remained 'just an avowal of faith' among scientists who believe that the mind must ultimately be explicable as a product of natural selection (Pinker, 1997, pg. 301).

The term 'conceptual metaphor' is used ambiguously, sometimes to refer to patterns in language, and other times to nonlinguistic conceptual structures that are hypothesized to underlie these patterns in language. To avoid this ambiguity, I will refer to patterns in language as *linguistic metaphors* and to the hypothesized nonlinguistic metaphorical structures in the mind as *mental metaphors* (Casasanto, 2008, 2009a). This terminological shift allows several critical questions to be framed clearly. Part 1 of this chapter will address the question, 'Do people use mental metaphors that correspond to their linguistic metaphors in order to conceptualize abstract domains, even when they're not using language?' Part 2 asks, 'If so, do people who tend to use different linguistic metaphors also rely on different mental metaphors?' and further, 'Does using different linguistic metaphors *cause* speakers of different languages to rely on different mental metaphors?' Finally, distinguishing linguistic metaphors from mental metaphors allows us to pose other questions that lie beyond the scope of this chapter (see Casasanto, 2008, 2009a, 2009b), such as, 'Are there any mental metaphors for which no corresponding linguistic metaphors exist?' This question has received virtually no attention from linguists or psychologists. This could be due, in part, to the fact that it is nonsensical when phrased in the traditional terminology: 'Are there any conceptual metaphors for which no corresponding conceptual metaphors exist?' Whereas Conceptual Metaphor theorists treat patterns in language as a source of *evidence* that people think metaphorically, the research presented here takes patterns in language as a source of *hypotheses* about conceptual structure.

1.3 Experimental evidence for mental metaphors

Boroditsky (2000) conducted some of the first behavioral tests of the psychological reality of mental metaphors. Her tasks capitalized on the fact that in order to talk about spatial or temporal sequences, speakers must adopt a particular frame of reference. Sometimes we use expressions that suggest we are moving through space or time (e.g., *we're approaching Maple Street*; *we're approaching Christmas*). Alternatively, we can use expressions that suggest objects or events are moving with respect to one another (*Maple Street comes before Elm Street*; *Christmas comes before New Year's*). In one experiment, Boroditsky found that priming participants to adopt a given spatial frame of reference facilitated their interpretation of sentences that used the analogous temporal frame of reference. Importantly, the converse was not found: temporal primes did not facilitate interpreting spatial sentences. This priming asymmetry parallels a well established asymmetry in linguistic metaphors: people talk about the abstract in terms of the concrete (e.g., time in terms of space) more than the other way around (Lakoff and Johnson, 1980). Based on these results Boroditsky proposed a refinement of Conceptual Metaphor Theory, the Metaphoric Structuring View, according to which (a) the domains of space and time share conceptual structure, and (b) spatial information is useful (though not necessary) for thinking about time. A second set of experiments showed that real-world spatial situations (e.g., riding on a train, or standing in a cafeteria line) and even imaginary spatial scenarios can influence how people interpret spatiotemporal metaphors

(Boroditsky and Ramscar, 2002). These studies rule out what Boroditsky (2000) calls the Dubious View, that space-time metaphors in language are simply 'etymological relics with no psychological consequences' (pg. 6).

If people use spatial schemas to think about time, as suggested by metaphors in language, then do people who use different spatiotemporal metaphors in their native tongues think about time differently? To find out, Boroditsky (2001) compared performance on space-time priming tasks in speakers of English, a language which typically describes time as horizontal, and speakers of Mandarin Chinese, which also commonly uses vertical spatiotemporal metaphors. English speakers were faster to judge sentences about temporal succession (e.g., *March comes earlier than April*) when primed with a horizontal spatial event, but Mandarin speakers were faster to judge the same sentences when primed with a vertical spatial stimulus. This was true despite the fact that all of the sentences were presented in English. In a follow-up study, Boroditsky (2001) trained English speakers to use vertical metaphors for temporal succession (e.g., *March is <u>above</u> April*). After training, their priming results resembled those of the native Mandarin speakers.

Together, Boroditsky's studies provide some of the first evidence that (a) people not only talk about time in terms of space, they also think about it that way, (b) people who use different spatiotemporal metaphors also think about time differently, and (c) learning new spatial metaphors can change the way you mentally represent time. Yet, these conclusions are subject to a skeptical interpretation. Boroditsky's participants made judgments about sentences containing spatial or temporal language. Perhaps their judgments showed relations between spatial and temporal thinking that were consistent with linguistic metaphors only because they were required to process space or time *in language.* Would the same relationships between mental representations of space and time be found if participants were tested on nonlinguistic tasks?

The fact that people communicate via language replete with anaphora, ambiguity, metonymy, sarcasm, and deixis seems proof that what we say provides only a thumbnail sketch of what we think. Most theorists posit at least some independence between semantic representations and underlying conceptual representations (Jackendoff, 1972; Katz and Fodor, 1963; Levelt, 1989; cf., Fodor, 1975). Even those who posit a single, shared 'level' of representation for linguistic meaning and nonlinguistic concepts allow that semantic structures must constitute only a subset of conceptual structures (Chomsky, 1975; Jackendoff, 1983). Because we may think differently when we're using language and when we're not, well-founded doubts persist about how deeply patterns in language truly reflect – and perhaps shape – our nonlinguistic thought. According to linguist Dan Slobin (1996):

> Any utterance is a selective schematization of a concept – a schematization that is in some ways dependent on the grammaticized meanings of the speaker's particular language, recruited for the purposes of verbal expression. (pg. 75–76)

Slobin argues that when people are 'thinking for speaking' (and presumably for reading or listening to speech), their thoughts are structured, in part, according to their language and its peculiarities. Consequently, speakers of different languages may think differently when they are using language. But how about when people are not thinking for speaking? Eve Clark (2003) asserts that:

> [When people are] thinking for remembering, thinking for categorizing, or one of the many other tasks in which we may call on the representations we have of objects or events – then their representations may well include a lot of material not customarily encoded in their language. It seems plausible to assume that such conceptual representations are nearer to being universal than the representations we draw on for speaking. (pg. 21)

Clark predicts that results may differ dramatically between tests of language–thought relations that use language and those that do not:

> …we should find that in tasks that require reference to representations in memory that don't make use of any linguistic expression, people who speak different languages will respond in similar, or even identical, ways. That is, representations for nonlinguistic purposes may differ very little across cultures or languages. (2003, pg. 22)

Clark adds:

> Of course, finding the appropriate tasks to check on this without any appeal to language may prove difficult. (2003, pg. 22)

Clark's skepticism echoes concerns raised by Papafrougou, Massey, and Gleitman (2002) regarding the difficulty of studying the language–thought interface:

> …domains within which language might interestingly influence thought are higher–level cognitive representations and processes, for instance, the linguistic encoding of time […] A severe difficulty in investigating how language interfaces with thought at these more 'significant' and 'abstract' levels has been their intractability to assessment. As so often, the deeper and more culturally resonant the cognitive or social function, the harder it is to capture it with the measurement and categorization tools available to psychologists. (pg. 191–192)

For the studies reported here, new experimental tools were developed in order to (a) evaluate Metaphor Theory as an account of the structure and evolution of abstract concepts, and (b) investigate relationships between language and nonlinguistic mental representations. The first two sets of experiments used the concrete domain of space and the relatively abstract domain of time as a testbed for Metaphor Theory, and the final set extended these findings beyond the domain of time. These experiments used novel psychophysical tasks with nonlinguistic stimuli and responses in order to distinguish two theoretical positions, one which posits *shallow* and the other *deep* relations between language and nonlinguistic thought (table 1):

Table 1.

The Shallow View:	**The Deep View:**
i. Language reflects the structure of the mental representations that speakers form for the purpose of using language. These are likely to be importantly different, if not distinct, from the representations people use when they are thinking, perceiving, and acting without using language.	**i.** Language reflects the structure of the mental representations that speakers form for the purpose of using language. These are likely to be similar to, if not overlapping with, the representations people use when they are thinking, perceiving, and acting without using language.
ii. Language may influence the structure of mental representations, but only (or primarily) during language use.	**ii.** Patterns of thinking established during language use may influence the structure of the mental representations that people form even when they're not using language.
iii. Cross-linguistic typological differences are likely to produce 'shallow' behavioral differences on tasks that involve language or high-level cognitive abilities (e.g., naming, explicit categorization). However, such behavioral differences should disappear when subjects are tested using nonlinguistic tasks that involve low-level perceptuo-motor abilities.	**iii.** Some cross-linguistic typological differences are likely to produce 'deep' behavioral differences, observable not only during tasks that involve language or high-level cognitive abilities, but also when subjects are tested using nonlinguistic tasks that involve low-level perceptuo-motor abilities.
iv. Although the semantics of languages differ, speakers' underlying conceptual and perceptual representations are, for the most part, universal.	**iv.** Where the semantics of languages differ, speakers' underlying conceptual and perceptual representations may differ correspondingly, such that language communities develop distinctive conceptual repertoires.

2 Do people use space to think about time?

Do people use mental representations of space in order to mentally represent time, as metaphors in language suggest they do – even when they're not using language? The first six experiments reported here tested the hypothesis that temporal thinking depends, in part, on spatial thinking (Casasanto and Boroditsky, 2008). In each task, participants viewed simple nonlinguistic, non-symbolic stimuli (i.e., lines or dots) on a computer screen, and estimated either their duration or their spatial displacement. Durations and displacements were fully crossed, so there was no correlation between the spatial and temporal components of the stimuli. As such, one stimulus dimension served as a distractor for the other: an irrelevant piece of information that could potentially interfere with task performance. Patterns of cross-dimensional interference were analyzed to reveal relationships between spatial and temporal representations.[3]

Broadly speaking, there are three possible relationships between people's mental representations of space and time. First, the two domains could be *symmetrically dependent*. John Locke (1689/1995) argued that space and time are mutually inextricable in our minds, concluding that, 'expansion and duration do mutually embrace and comprehend each other; every part of space being in every part of duration, and every part of duration in every part of expansion' (p. 140). Alternatively, our ideas of space and time could be *independent*. Any apparent relatedness could be due to structural similarities between essentially unrelated domains (Murphy, 1996, 1997). A third possibility is that time and space could be *asymmetrically dependent.* Representations in one domain could be parasitic on representations in the other, as suggested by their asymmetric relationship in linguistic metaphors (Boroditsky, 2000; Gentner, 2001; Gibbs, 1994; Lakoff and Johnson, 1980, 1999).

These three possible relationships between space and time predict three distinct patterns of cross-dimensional interference. If spatial and temporal representations are symmetrically dependent on one another, then any cross-dimensional interference should be approximately symmetric: line displacement should modulate estimates of line duration, and vice versa. Alternatively, if spatial and temporal representations are independent, there should be no significant cross-dimensional interference. However, if mental representations of time are asymmetrically dependent on mental representations of space, as suggested by spatiotemporal metaphors in language, then any cross-dimensional interference should be asymmetric: line displacement should affect estimates of line duration more than line duration affects estimates of line displacement.

For Experiment 1, native English speaking participants viewed 162 lines of varying lengths (200–800 pixels, in 50 pixel increments), presented on a computer monitor for varying durations (1–5 seconds, in 500 ms increments). Lines 'grew' horizontally from left to right, one pixel at a time, along the vertical midline. Each line remained on the screen until it reached its maximum displacement, and then disappeared. Immediately after each line was shown, a prompt appeared indicating that the participant should reproduce either the line's displacement (if an 'X' icon appeared) or its duration (if an 'hourglass' icon appeared), by clicking the mouse to indicate the endpoints of each temporal or spatial interval. Space trials and time trials were randomly intermixed.

Results of Experiment 1 showed that spatial displacement affected estimates of duration, but duration did not affect estimates of spatial displacement (Figure 1a). For stimuli of the same average duration, lines that travelled a shorter distance were judged to take a shorter time, and lines that travelled a longer distance were judged to take a longer time. Subjects incorporated irrelevant spatial information into their temporal estimates, but not vice versa. Estimates of duration and displacement were highly accurate, and were equally accurate in the two domains. The asymmetric cross-dimensional interference we observe cannot be attributed to a difference in the accuracy of duration and displacement estimations, as no significant difference in was found.

Experiments were conducted to assess the generality of these results, and to evaluate potential explanations. In Experiment 1, participants did not know until after each line was presented whether they would need to estimate displacement or duration. They had to attend to both the spatial and temporal dimensions of the stimulus. Experiment 2 addressed the possibility that cross-dimensional interference would diminish if participants were given the opportunity to attend selectively to the trial-relevant stimulus dimension, and to ignore the trial-irrelevant dimension. Materials and procedures were identical to those used in Experiment 1, with one exception. A cue preceded each growing line, indicating which stimulus dimension participants would need to reproduce. Results of Experiment 2 (Figure 1b) replicated those of Experiment 1. Participants were able to disregard line duration when estimating displacement. By contrast, they were unable to ignore line displacement, even when they were encouraged to attend selectively to duration. The cross-dimensional effect of space on time estimation in Experiment 1 was not caused by a task-specific demand for subjects to encode spatial and temporal information simultaneously.

Experiments 3–5 addressed concerns that spatial information in the stimulus may have been more stable or more salient than temporal information, and that differences in stability or salience produced the asymmetrical cross-dimensional interference observed in Experiments 1 and 2. One concern was that participants may have relied on spatial information to make temporal estimates because stimuli were situated in a constant spatial frame of reference (i.e., the computer monitor). For Experiment 3, stimuli were also situated in a constant temporal frame of reference. Temporal delay periods were introduced preceding and following line presentations, which were proportional to the spatial gaps between the ends of the stimulus lines and the edges of the monitor. Results (Figure 1c) replicated those of Experiments 1 and 2.

Experiment 4 addressed the possibility that space would no longer influence participants' time estimates if stimulus duration were indexed by something non-spatial. For this experiment, a constant tone (260 Hz) accompanied each growing line. Materials and procedures were otherwise identical to those used in Experiment 2. The tone began sounding when the line started to grow across the screen, and stopped sounding when the line disappeared. Thus, stimulus duration was made available to the participant in both the visual and auditory modalities, but stimulus displacement was only available visually. Results (Figure 1d) replicated those of the previous experiments. Displacement strongly influenced participants' duration estimates, even when temporal information was provided via a different sensory modality from the spatial information.

Experiment 5 was designed to equate the mnemonic demands of the spatial and temporal dimensions of the stimulus. Materials and procedures were identical to those used in Experiment 2, with one exception. Rather than viewing a growing line, subjects viewed a dot (10x10 pixels) that moved horizontally across the midline of the screen. In the previous experiments, just before each growing line disappeared participants could see its full spatial extent, from end to end, seemingly at a glance. By contrast, the spatial extent of a moving dot's path could never be seen all at once, rather it had to be imagined: in order to compute the distance that a dot travelled, participants had to retrieve the dot's starting point from memory once its ending point was reached. The spatial and temporal dimensions of the dot stimulus had to be processed similarly in this regard: whenever we compute the extent of a temporal interval we must retrieve its starting point from memory once the end of the interval is reached. Results (Figure 1e) replicated those of previous experiments.

Experiment 6 investigated whether motion or speed affected participants' time estimates in Experiments 1–5, rather than stimulus displacement. Materials and procedures were identical to those used in Experiment 2, with the following exception.

Rather than growing lines, participants viewed stationary lines, and estimated either the amount of time they remained on the screen or their distance from end to end, using mouse clicks. Results replicate those of previous five experiments (Figure 1f), indicating that stimulus displacement can strongly modulate time estimates even in the absence of stimulus motion.

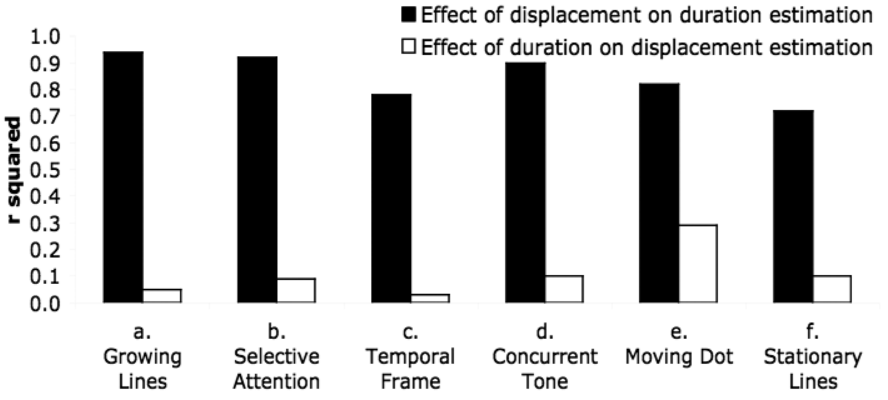


Figure 1. Summary of cross-dimensional interference effects for Experiments 1–6. The effect of distance on time estimation was significantly greater than the effect of time on distance estimation for all experiments. (1a, Growing lines: difference of correlations = 0.75; z = 3.24, p <.001. 1b, Growing lines, selective attention: difference of correlations = 0.66; z = 2.84, p < .003. 1c, Growing lines, temporal frame of reference: difference of correlations = 0.71; z =2.09, p <.02. 1d, Growing lines, concurrent tone: difference of correlations =0.63; z = 2.60, p <.005. 1e, Moving dot: difference of correlations = 1.45; z = 3.69, p <.001. 1f, Stationary lines: difference of correlations = 0.54; z = 1.62, p <.05.) Figure reproduced with permission from Casasanto, D. and Boroditsky, L. (2008). Time in the Mind: Using space to think about time. *Cognition, 106, 579–593.*

Results of all six experiments unequivocally support the hypothesis that people incorporate spatial information into their time judgments more than they incorporate temporal information into their spatial judgments. These findings converge with those of Cantor and Thomas (1977), who showed that spatial information influences temporal judgments but not vice versa for very briefly presented stimuli (30–70 msecs). Previous behavioral tests of Metaphor Theory have used linguistic stimuli (Boroditsky, 2000, 2001; Boroditsky and Ramscar, 2002; Gibbs, 1994; Meier and Robinson, 2004; Meier, Robinson and Clore, 2004; Richardson, Spivey, Barsalou and McRae, 2003; Schubert, 2005; Torralbo, Santiago and Lupiáñez, 2006). While these studies support the psychological reality of mental metaphors, they leave open the possibility that people only think about abstract domains like time metaphorically when they are using language (i.e., when they are 'thinking for speaking' (E. Clark, 2003; Slobin, 1996)). Experiments described above used nonlinguistic stimuli and responses, and demonstrated for the first time that even our low-level perceptuo-motor representations in the domains of space and time are related as predicted by linguistic metaphors.

Although English speakers describe time in terms of space almost obligatorily (Jackendoff, 1983; Pinker, 1997), we can also optionally describe space in terms of time. For example, in English we could say *my brothers live* 5 *minutes apart* to indicate that they live a short distance apart. Thus, the relationship between time and space in linguistic metaphors is asymmetrical, but not unidirectional. Accordingly, asymmetrical cross-dimensional interference between space and time was predicted in these experiments. This prediction does not entail that time can never affect spatial judgments: only that the effect of space on time estimation should be greater than the effect of time on space estimation when the effects are compared appropriately. Results of Experiments 1–6 did not show any significant effect of time on distance estimation, but such a finding would still be compatible with the asymmetry hypothesis, so long as the effect of distance on time estimation was significantly greater than the effect of time on distance estimation.

It is noteworthy that space influenced temporal judgments even for spatiotemporal stimuli that participants could experience directly. Growing lines are observable, and are arguably less abstract than entities like the 'moving meeting' described in section 0.1. Brief durations could, in principle, be mentally represented independently of space, by an interval-timer or pulse-accumulator (see Ivry and Richardson, 2002 for review), yet these data suggest that spatial representations are integral to the timing of even simple, observable events. Thinking about time metaphorically in terms of space may allow us to go beyond these basic temporal representations. Mentally representing time as a linear path may enable us to conceptualize more abstract temporal events that we cannot experience directly (e.g., moving a meeting *forward* or pushing a deadline *back*), as well as temporal events that we can never experience at all (e.g., the *remote* past or the *distant* future). Metaphorical mappings from spatial paths, which can be traveled both forward and backward, may give rise to temporal constructs such as *time–travel* that only exist in our imagination.

Together, these experiments demonstrate that the metaphors we use can provide a window on the structure of our abstract concepts. They also raise a further question about relations between linguistic metaphors and nonlinguistic mental representations: if people think about time in terms of space (the way they talk about it), then do people who use different space-time metaphors in their native languages think differently – even when they're not using language?

3 Does language shape the way we think about time?

The first set of experiments supports the *Deep View* of language-thought relations by showing that temporal representations depend, in part, on spatial representations, as predicted by metaphors in English – even when people are performing low-level, nonlinguistic psychophysical tasks (see Table 1, number i). However, it is not clear from these data whether linguistic metaphors merely reflect English speakers' underlying nonlinguistic representations of time, or whether language also shapes those representations. According to the *Shallow View,* it is possible that speakers of a language with different duration metaphors would nevertheless perform similarly to English speakers on nonlinguistic tasks. Thus, the first set of experiments leaves the following question unaddressed, posed by the influential amateur linguist, Benjamin Whorf:

> Are our own concepts of 'time', 'space', and 'matter' given in substantially the same form by experience to all men, or are they in part conditioned by the structure of particular languages?' (1939/2000, pg. 138.)

This *Whorfian* question remains the subject of renewed interest and debate. Does language shape thought? The answer *yes* would call for a reexamination of the 'universalist' assumption that has guided Cognitive Science for decades, according to which nonlinguistic concepts are formed independently of the words that name them, and are invariant across languages and cultures (Fodor, 1975; Pinker, 1994, Papafragou, Massey and Gleitman, 2002). This position is often attributed to Chomsky (1975), but has been articulated more recently by Pinker (1994) and by Lila Gleitman and colleagues (Papafragou, Massey and Gleitman, 2002; Snedeker and Gleitman, 2004). The Shallow View proposed here can be considered a variety of the universalist view that can still plausibly be maintained despite recent psycholinguistic evidence supporting the Whorfian hypothesis (e.g., Boroditsky, 2001).

Skepticism about some Whorfian claims has been well founded (see Pinker, 1994, ch. 3, for a review of evidence against the Whorfian hypothesis). A notorious fallacy, attributable in part to Whorf, illustrates the need for methodological rigor. Whorf (1939/2000) argued that Eskimos must conceive of snow differently than English speakers because the Eskimo lexicon contains multiple words that distinguish different types of snow, whereas English has only one word to describe all types. The exact number of snow words the Eskimos were purported to have is not clear. This number has now been inflated by the popular press to as many as four hundred. According to a Western Greenlandic Eskimo dictionary published in Whorf's time, however, Eskimos may have had as few as two distinct words for snow (Pullum, 1991).

Setting aside Whorf's imprecision and the media's exaggeration, there remains a critical missing link between Whorf's data and his conclusions: Whorf (like many researchers today) used purely linguistic data to support inferences about nonlinguistic mental representations. Steven Pinker illustrates the resulting circularity of Whorf's claim in this parody of his logic:

> [They] speak differently so they must think differently.
> How do we know that they think differently?
> Just listen to the way they speak! (Pinker, 1994, pg. 61).

Such circularity would be escaped if nonlinguistic evidence could be produced to show that two groups of speakers who talk differently also think differently in corresponding ways.

A series of experiments explored relationships between spatiotemporal language and nonlinguistic mental representation of time. The first experiment, a corpus search, uncovered previously unexplored cross-linguistic differences in spatial metaphors for duration. Next, we tested whether these linguistic differences correlate with differences in speakers' low-level, nonlinguistic time representations.[4] Finally, we evaluated a causal role for language in shaping time representations.[5]

3.1 1-Dimensionsal and 3-dimensional spatial metaphors for time

Literature on how time can be expressed verbally in terms of space (and by hypothesis, conceptualized in terms space) has focused principally on linear spatial metaphors. But is time necessarily conceptualized in terms of unidimensional space? Some theorists have suggested so (Clark, 1973, Gentner, 2001), and while this may be true regarding temporal succession, linguistic metaphors suggest an alternative spatialization for duration. English speakers not only describe time as a line, they also talk about *oceans of time, saving time in a bottle*, and liken the 'days of their lives' to *sands through the hourglass*. Quantities of time are described as amounts of a substance occupying three dimensional space (i.e., volume).

Experiment 7 compared the use of 'time as distance' and 'time as amount' metaphors across four languages. Every language we examined uses both distance and amount metaphors, but their relative prevalence and productivity appear to vary markedly. In English, it is natural to talk about *a long time*, borrowing the structure and vocabulary of a linear spatial expression like *a long rope.* Yet in Spanish, the direct translation of 'long time', *largo tiempo*, sounds awkward to speakers of most dialects.[6] *Mucho tiempo*, which means 'much time', is preferred.

In Greek, the words *makris* and *kontos* are the literal equivalents of the English spatial terms *long* and *short*. They can be used in spatial contexts much the way *long* and *short* are used in English (e.g., *ena <u>makry</u> skoini* means 'a <u>long</u> rope'). In temporal contexts, however, *makris* and *kontos* are dispreferred in instances where *long* and *short* would be used naturally in English. It would be unnatural to translate *a <u>long</u> meeting* literally as *mia <u>makria</u> synantisi*. Rather than using distance terms, Greek speakers typically indicate that an event lasted a long time using *megalos*, which in spatial contexts means physically 'large' (e.g., a big building), or using *poli*, which in spatial contexts means 'much' (e.g., much water). Compare how English (*e*) and Greek (*g*) typically modify the duration of the following events (literal translations in parentheses):

1*e* long night
1*g* megali nychta (*big night*)

2*e* long relationship
2*g* megali schesi (*big relationship*)

3*e* long party
3*g* parti pou kratise poli (*party that lasts much*)

4*e* long meeting
4*g* synantisi pou diekese poli (*meeting that lasts much*)

In examples 1*g* and 2*g*, the literal translations might surprise an English speaker, for whom *big night* is likely to mean 'an exciting night', and *big relationship* 'an important relationship'. For Greek speakers, however, these phrases can also communicate duration, expressing time not in terms of 1-dimensional linear space, but rather in terms of 3-dimensional size or amount.

To quantify the relative prevalence of distance and amount metaphors for duration across languages, the most natural phrases expressing the ideas 'a long time' and 'much time' were elicited from native speakers of English (*long time, much time*), French (*longtemps, beaucoup de temps*), Greek (*makry kroniko diastima*, *poli ora*), and Spanish (*largo tiempo, mucho tiempo*). The frequencies of these expressions were compared in a very large multilingual text corpus: www.google.com. Each expression was entered as a search term. Google's language tools were used to find exact matches for each expression, and to restrict the search to web pages written only in the appropriate languages. The number of google 'hits' for each expression was tabulated, and the proportion of distance hits and amount hits was calculated for each pair of expressions, as a measure of their relative frequency. English and French, distance metaphors were dramatically more frequent than amount metaphors. The opposite pattern was found in Greek and Spanish (Figure 2).

Although all languages surveyed use *both* distance and amount metaphors for duration, the relative strengths of these metaphors appears to vary across languages. This simple corpus search by no means captures all of the complexities of how time is metaphorized in terms of space within or between languages, but these findings corroborate native speakers' intuitions for each language, and provide a quantitative linguistic measure on which to base predictions about behavior in nonlinguistic tasks.

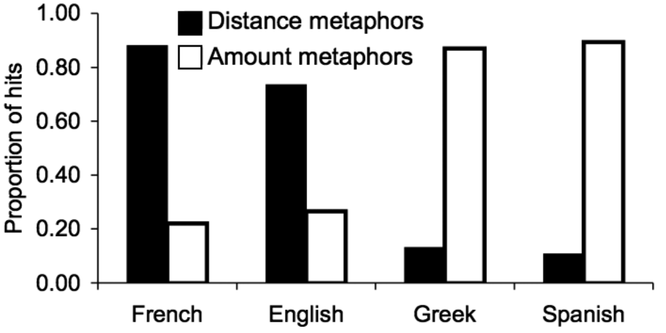


Figure 2. Results of Experiment 7. Black bars indicate the proportion of Google 'hits' for expressions meaning *long time*, and white bars for expressions meaning *much time* in each language.

3.2 Do people who talk differently think differently?

Do people who use different spatiotemporal metaphors think about time differently – even when they're not using language? Experiments 8 and 9 explored the possibility that speakers who preferentially use distance metaphors in language tend to co-opt linear spatial representations to understand duration, whereas speakers who preferentially use amount metaphors tend to co-opt 3-dimensional spatial representations. Speakers of two languages surveyed in Experiment 7 (i.e., English and Greek) performed a pair of nonlinguistic psychophysical tasks, which required them to estimate duration while overcoming different kinds of spatial interference (i.e., distance or amount interference). If people's conceptions of time are substantially the same universally irrespective of the languages they speak, as suggested by the *Shallow View*, then performance on these tasks should not differ between language groups. On the *Deep View*, however, it was predicted that participants' performance should vary in ways that parallel the metaphors in their native languages.

The 'distance interference' task was modeled on the 'growing line' task described in Experiment 2. English participants in the previous growing line studies may have suffered interference from distance during duration estimation, in part, because distance and duration are strongly conflated in the English lexicon. Would the same confusion be found in speakers of other languages? It was predicted that native English speakers would show a strong effect of distance on time estimation when performing the growing line task, whereas speakers of Greek would show a weaker effect, since distance and duration are less strongly associated in the Greek language .

A complementary 'amount interference' task was developed, in which participants watched a schematically drawn container of water filling up gradually, and estimated either how full it became or how much time it remained on the computer screen, using mouse clicks as in the growing line tasks. Spatial and temporal parameters of the stimuli were equated across tasks. Behavioral predictions for the Filling Tank task were the mirror image of predictions for the Growing Line task: speakers of Amount Languages like Greek should show a strong influence of 'fullness' on time estimation, whereas speakers of Distance Languages like English should show a weaker effect.

Results showed that effects of spatial interference on duration estimation followed predictions based on the relative prevalence of distance and amount metaphors for time in speakers' native languages. English showed a strong effect of line length but a weak effect of tank fullness on duration estimation; Greek speakers showed the opposite pattern of results (Figure 3). A 2 x 2 ANOVA compared these slopes with Language (English, Greek) and Task (distance interference, amount interference) as between-subject factors, revealing a highly significant Language by Task interaction, with no main effects ($F(1,56)=10.41$, $p=.002$).

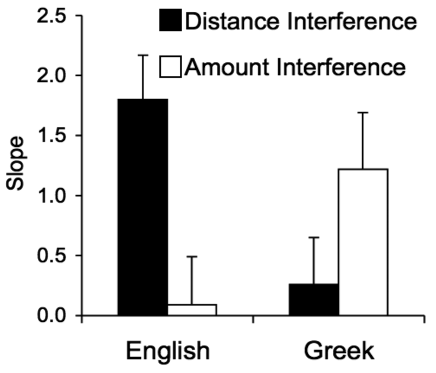


Figure 3. Results of Experiments 8 and 9. Black bars indicate the slope of the effect of line displacement on duration estimation. White bars indicate the slope of the effect of tank fullness on duration estimation. The relationship between the effects of distance and volume on time estimation was predicted by the relative prevalence of distance and amount metaphors in English and Greek (see figure 2).

The observed differences in the effects of spatial distance and amount on duration estimation cannot be attributed to overall differences in performance across tasks or across groups. Within-domain performance (i.e., the effect of target duration on estimated duration, and the effect of target distance or fullness on estimated distance or fullness) was compared across tasks and across groups: no significant differences were found between correlations or slopes, even in pairwise comparisons.

One difference between the Growing Line and Filling Tank tasks was that the lines grew horizontally, but the tanks filled vertically. To determine whether the spatial orientation of the stimuli and responses gave rise to the observed cross-linguistic differences in performance on the Growing Lines and Filling Tank tasks, an Upward Growing Lines task was administered to speakers of English and Greek. No significant difference was found in the effect of vertical displacement on time estimation across languages, suggesting that the orientation of stimuli cannot account for the between-group differences observed in Experiments 8 and 9.

Overall, Experiments 7–9 show that the way people talk about time correlates strongly with the way they think about it – even when they're performing simple nonlinguistic perceptuo-motor tasks – as predicted by the *Deep View* of language-thought relations. (See Table 1, ii.- iv.) Much of the literature on temporal language has highlighted crosslinguistic commonalities in spatiotemporal metaphors (e.g., Alverson, 1994). The studies presented here begin to explore some previously neglected crosslinguistic differences, and to discover their nonlinguistic consequences. The corpus search reported in Experiment 7 provides one measure of how frequently different languages use distance and amount metaphors for duration; the relative frequencies of *long time* and *much time* expressions across languages proved highly predictive of performance on nonlinguistic duration estimation tasks. Often, however, spatial metaphors describe *events* rather than describing *time*, per se. Preliminary data from a questionnaire study suggest that English consistently prefers distance metaphors for describing both time (e.g., *a long time*) and events (e.g., *a long party*), whereas Greek consistently prefers volume metaphors for time (e.g., *poli ora* tr.'much time') and for events (e.g., *parti pou kratise poli* tr. 'party that lasts much'), corroborating the results of the corpus search. Ongoing studies seek to characterize these crosslinguistic differences more fully, and to specify which features of language correspond to 'deep' differences in nonlinguistic mental representations of time.

3.3 How might perceptual and linguistic experience shape abstract thought?

How do people come to think about time in terms of space? How do speakers of different languages come to conceptualize time differently? Turning to the first question, some mappings from concrete to abstract domains of knowledge may be initially established pre-linguistically, based on interactions with the physical world (Clark, 1973). For example, people are likely to track the kinds of correlations in experience that are important for perceiving and acting on their environment; they may learn associations between time and space by observing that more time passes as objects travel farther, and as substances accumulate more. This proposal entails that although time depends in part on spatial representations, time can also be mentally represented *qua* time, at least initially: in order for cross-dimensional associations to form, some primitive representations must already exist in each dimension. Primitive temporal notions, however, of the sort that we share with infants and non-human animals, may be too vague or fleeting to support higher order reasoning about time. Grafting primitive temporal representations onto spatial representations may make time more amenable to verbal or imagistic coding, and may also import the inferential structure of spatial relations into the domain of time (Pinker, 1997).

If metaphorical mappings are experience-based, and are established pre-linguistically, what role might language play in shaping abstract thought? Since the laws of physics are the same in all language communities, prelinguistic children's conceptual mappings between time, distance, and amount could be the same universally. Later, as children acquire language, these mappings are adjusted: each time we use a linguistic metaphor, we activate the corresponding conceptual mapping. Speakers of Distance Languages then activate the time-distance mapping frequently, eventually strengthening it at the expense of the time-amount mapping (and vice versa for speakers of Amount Languages). Mechanistically, this could happen via a process of competitive associative learning.

Did language experience give rise to the language-related differences in performance reported for the Growing Line and Filling Tank experiments? A perennial complaint about studies claiming effects of language on thought is that researchers mistake correlation for causation. Although it is difficult to imagine what nonlinguistic cultural or environmental factors could have caused performance on Experiments 8 and 9 in English and Greek speakers to align so uncannily with the metaphors in these languages, the data are nevertheless correlational. Using crosslinguistic data to test for a causal influence of language on thought is problematic, since experimenters cannot randomly assign subjects to have one first language or another: crosslingusitic studies are necessarily quasi-experimental.

For Experiment 10, a pair of training tasks (i.e., true experimental interventions) was conducted to provide an in principle demonstration that language can influence even the kinds of low-level mental representations that people construct while performing psychophysical tasks, and to test the hypothesis that language shapes time representations in natural settings by adjusting the strengths of cross-domain mappings. Native English speakers were randomly assigned to perform either a Distance Training or Amount Training task. Participants completed 192 fill-in-the-blank sentences using the words *longer* or *shorter* for Distance Training, and *more* or *less* for the Amount Training task. Half of the sentences compared the length or capacity of physical objects (e.g., An alley is *longer / shorter* than a clothesline; A teaspoon is *more / less* than an ocean), the other half compared the duration of events (e.g., A sneeze is *longer / shorter* than a vacation; A sneeze is *more / less* than a vacation). By using distance terms to compare event durations, English speakers were reinforcing the already preferred source-target mapping between distance and time. By using amount terms, English speakers were describing event durations similarly to speakers of an Amount Language (see Greek examples in section 2.1), and by hypothesis, they were activating the dispreferred volume-time mapping. After this linguistic training, all participants performed the nonlinguistic Filling Tank task from Experiment 9. We predicted that if using a linguistic metaphor activates the corresponding conceptual mapping between source and target domains, then repeatedly using amount metaphors during training should (transiently) strengthen participants' nonlinguistic amount-time mapping.

Consistent with this prediction, the slope of the effect of amount on time estimation was significantly greater after amount training than after distance training (difference of slopes = 0.89, $t(28) = 1.73$, $p<.05$; Figure 4). Following about 30 minutes of concentrated usage of amount metaphors in language, native English speakers' performance on the Filling Tank task was statistically indistinguishable from the performance of the native Greek speakers tested in Experiment 9. By encouraging the habitual use of either distance- or amount-based mental metaphors, our experience with natural language may influence our everyday thinking about time in much the same way as this laboratory training task.

These findings help to resolve apparent tensions between the proposal that perceptuo-motor image schemas underlie our abstract concepts and the notion of linguistic relativity. Johnson (2005) defines an image schema as 'a dynamic recurring pattern of organism-environment interactions' (pg. 19). Presumably, people from all language communities inhabit the same physical world and interact with their environment using the same perceptuo-motor capacities, therefore the image schemas they develop should be universal. Yet, even if we all develop similar image schemas initially, based on our physical experiences, Experiments 8–10 suggest the way we deploy these image schemas depends on our linguistic experiences. Duration can be mentally represented *both* in terms of distance and in terms of amount. The extent to which each of these conceptual space-time mappings is activated in a given speaker or community of speakers varies with the strength of the corresponding linguistic metaphors. The structure of abstract concepts like *duration* appears to be shaped both by perceptuo-motor experience (which is plausibly universal) and by language use (which is culture-specific).

Figure 4. Results of Experiment 10. Bars indicate the slope of the effect of tank fullness on duration estimation after training with distance metaphors (left), amount metaphors (right), or with no training (middle) prior to performing the Filling Tank task. The cross-dimensional effect of amount on time estimation was significantly greater after training with amount metaphors than with distance metaphors.

4 Beyond space and time: Spatial representation of musical pitch

Time and space provide a model system for exploring connections between abstract and concrete mental representations, but time is just one among many domains that we spatialize in language; time may be just one of many abstract domains that import their structure or content, in part, from the domain of space. In Experiment 11, the psychophysical tasks that were developed to investigate space and time were adapted to explore relationships between space and musical pitch.[7]

Like time, pitch is often described in English using linear spatial terms. Unlike time, pitch tends to be described using vertical rather than horizontal metaphors. Pitches can be *high* or *low*, and can *rise*, *fall*, *soar*, or *dip below* the staff. Yet, the fact that we talk about pitch in terms of vertical space doesn't necessarily mean that we think about it that way. One possibility is that pitch is mentally represented on its own terms, and is only coded into the same words that we use to describe space as a matter of convenience: domains that share structural similarities may be amenable to common linguistic description, obviating multiple domain-specific vocabularies. Alternatively, the spatialization of pitch in language may serve as a clue that leads us to a fuller understanding of how pitch is mentally represented.

The 'growing line' task described in Experiment 2 was modified for a nonlinguistic test of the hypothesis that our mental representations of musical pitch depend, in part, on spatial representations. Nine displacements ranging from 100 to 500 pixels (in 50 pixel increments) were fully crossed with nine different pitches ranging from middle

C4 to G#4 (in semitone increments). For each trial, participants heard a constant pitch while watching a line grow up the screen from bottom to top (for half of the subjects) or across the screen from left to right (for the other half of the subjects). Before each stimulus, participants were informed whether they would need to estimate distance or pitch, to encourage them to attend to the trial-relevant stimulus dimension and, if possible, to ignore the trial-irrelevant dimension. Participants estimated line displacements using mouse clicks, as in previous experiments. To estimate pitch, participants used the mouse to adjust a probe tone until it matched the remembered target pitch.

Watching vertical lines significantly modulated subjects' pitch estimates: tones of the same average frequency were judged to be higher in pitch if they accompanied lines that grew higher on the screen (effect of actual distance on estimated pitch: slope=.37; r^2=.77, p<.003). By contrast, watching horizontal lines did not significantly modulate pitch estimates. This finding is consistent with the occurrence of vertical but not horizontal metaphors for pitch in English. Further analyses showed that whereas vertical displacement affected estimates of pitch, pitch did not significantly influence estimates of vertical displacement. Thus, the relation between nonlinguistic mental representations of space and pitch appears to be asymmetrical, as predicted by the directionality of space-pitch metaphors in language.

While these results support the claim that musical pitch is mentally represented in part metaphorically, in terms of vertical space, they are agnostic as to the direction of causation between language and thought. Further studies (such as those described in sections 2.1–2.3) are needed to investigate whether linguistic metaphors merely reflect the spatial schemas that partly constitute pitch representations, or whether the way we talk about pitch can also shape the way we think about it.

5 Conclusions

Direct evidence that spatial cognition supported the evolution of abstract concepts may forever elude us, because human history cannot be recreated in the laboratory, and the mind leaves no fossil record. However, the studies reported here demonstrate the importance of spatial representations for abstract thinking in the mind that evolution produced. For decades, inferences about the perceptual foundations of abstract thought rested principally on linguistic and psycholinguistic data. These psychophysical experiments show that even nonlinguistic representations in concrete and abstract domains are related as linguistic metaphors predict: we think in mental metaphors.

Together, the experiments described in this chapter suggest that people not only talk about abstract domains using spatial words, they also think about them using spatial representations. Results are incompatible with the *Shallow View* of language-thought relations, and provide some of the first evidence for the view that language has *Deep* influences on nonlinguistic mental representation (see table 1). Experiments 1–6 show that people use spatial representations to think about time even when they're not producing or understanding language. Experiments 7–9 show that people who talk differently about time also think about it differently, in ways that correspond to their language-particular metaphors. Experiment 10 shows that language not only reflects the structure of underlying mental representations, it can also shape those representations in ways that influence how people perform even low-level, nonlinguistic, perceptuo-motor tasks. Experiment 11 shows that these findings extend beyond the 'testbed' domains of space and time.

These findings are difficult to reconcile with a universalist position according to which language calls upon nonlinguistic concepts that are presumed to be 'universal' (Pinker, 1994, pg. 82) and 'immutable' (Papafragou, Massey and Gleitman, 2002, pg. 216). Beyond influencing *thinking for speaking* (Slobin, 1996), language can also influence the nonlinguistic representations we build for remembering, acting on, and perhaps even perceiving the world around us. It may be universal that people conceptualize time according to the spatial metaphors, but since these metaphors vary across languages, members of different language communities develop distinctive conceptual repertoires. The structure of abstract domains like time depends, in part, on both perceptuo-motor experience and on experience using language.

Acknowledgments

Thanks to Lera Boroditsky, Herb Clark, and the citizens of Cognation for helpful discussions, and to Webb Phillips and 'Smooth' Jesse Greene for their help with programming and data collection. Thanks also to Olga Fotakopoulou and Ria Pita at the Aristotle University of Thessaloniki, Greece for sharing their time and expertise. This research was supported in part an NSF Graduate Research Fellowship and an NSF dissertation grant to the author.

Notes

1 Like our mental representations of time, some of our spatial representations may also be quite abstract. For example, our conception of the Milky Way galaxy's breadth is no more grounded in direct experience than our conception of its age.

2 Cultural evolution alone cannot explain our capacity for abstract thought because, as Wallace noted, members of 'stone age' societies who were given European educations manifested abilities to similar those of modern Europeans: the latent capacity to read, to perform Western art music, etc. was present in the minds of people whose cultures had never developed these abstract forms of expression.

3 Experiments 1–6 are described in full in Casasanto, D. and Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition* 106: 579–593.

4 A preliminary report on Experiments 7–9 appeared in Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., Fotokopoulu, O., Pita, R. and Gil, D. (2004). *How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish*. Proceedings of the 26th Annual Conference of the Cognitive Science Society, Chicago, IL.

5 A preliminary report on Experiment 10 appeared in Casasanto, D. (2005) *Perceptual foundations of abstract thought*. Doctoral dissertation, MIT.

6 Native speakers of European and South American Spanish report that *largo tiempo* is only used in poetic contexts (e.g., the Peruvian national anthem) to mean 'throughout the length of history'. By contrast, some bilingual North American Spanish speakers report that *largo tiempo* can be used colloquially, much like long time, perhaps because the construction is imported from English.

7 A preliminary report on Experiment 11 appeared in Casasanto, D., W. Phillips and L. Boroditsky, *Do we think about music in terms of space: Metaphoric representation of musical pitch*. Proceedings of 25th Annual Conference of the Cognitive Science Society, 2003. Boston, MA.

References

Alverson, H. (1994) *Semantics and experience: Universal metaphors of time in English, Mandarin, Hindi, and Sesotho*. Baltimore: Johns Hopkins University Press.

Boroditsky, L. (2000) Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* 75(1): 1–28.

Boroditsky, L. (2001) Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43(1): 1–22.

Boroditsky, L. and Ramscar, M. (2002) The roles of body and mind in abstract thought. *Psychological Science* 13(2): 185–189.

Cantor, N. and Thomas, E. (1977) Control of attention in the processing of temporal and spatial information in complex visual patterns. *Journal of Experimental Psychology: Human Perception and Performance* 3(2): 243–250.

Casasanto, D. (2005) *Perceptual foundations of abstract thought.* Doctoral dissertation. MIT, Cambridge.

Casasanto, D. (2008) Similarity and proximity: When does close in space mean close in mind? *Memory & Cognition* 36(6): 1047–1056.

Casasanto, D. (2009a) When is a linguistic metaphor a conceptual metaphor? In V. Evans and S. Pourcel (eds) *New directions in cognitive linguistics* 127–145. Amsterdam: John Benjamins.

Casasanto, D. (2009b) Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General* 138(3): 351–367.

Casasanto, D. and Boroditsky, L. (2008) Time in the mind: Using space to think about time. *Cognition* 106, 579–593.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., et al. (2004) *How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish.* Paper presented at the Cognitive Science Society, Chicago.

Casasanto, D., Phillips, W. and Boroditsky, L. (2003) *Do we think about music in terms of space: Metaphoric representation of musical pitch*. Proceedings of 25th Annual Conference of the Cognitive Science Society. Boston, MA.

Chomsky, N. (1975) *Reflections on language*. New York: Norton and Company.

Clark, E. (2003) Languages and representations. In D. Gentner and S. Goldin-Meadow (eds) *Language in mind* 17–23. Cambridge: MIT Press.

Clark, H. H. (1973) Space, time, semantics and the child. In T. E. Moore (ed.) *Cognitive development and the acquisition of language* 27–63. New York: Academic Press.

Cohen, J. (1967) *Psychological time in health and disease*. Springfield: Charles C. Thomas.
Darwin, C. (1859/1998) *The origin of species by means of natural selection or the preservation of favored races in the struggle for life*. New York: The Modern Library.
Darwin, C. (1874/1998) *The descent of man*. Amherst: Promethius Books.
DeLong, A. (1981) Phenomenological space-time: Toward an experiential relativity. *Science* 213(4508): 681–683.
Evans, V. (2004) *The structure of time: Language, meaning and temporal cognition*. Amsterdam: John Benjamins.
Fodor, J. (1975) *The language of thought*. Cambridge: Harvard University Press.
Gentner, D. (2001) Spatial metaphors in temporal reasoning. In M. Gattis (ed.) *Spatial schemas and abstract thought* 203–222. Cambridge: MIT Press.
Gibbs, R. W., Jr. (1994) *The poetics of mind: Figurative thought, language, and understanding.* Cambridge: Cambridge University Press.
Glicksohn, J. (2001) Temporal cognition and the phenomenology of time: A multiplicative function for apparent duration. *Consciousness and Cognition* 10: 1–25.
Gould, S. and Vrba, E. (1982) Exaptation – a missing term in the science of form. *Paleobiology* 8: 4–15.
Gould, S. J. (1980) Natural selection and the brain: Darwin vs. Wallace. In *The panda's thumb* 47–58. New York: Norton.
Gould, S. J. (1991) Not necessarily a wing. In *Bully for bronntosaurus* 139–151. New York: Norton and Co.
Gruber, J. (1965) *Studies in lexical relations.* Cambridge: MIT Press.
Ivry, R. and Richardson, T. (2002) Temporal control and coordination: The multiple timer model. *Brain and Cognition* 48: 117–132.
Jackendoff, R. (1972) *Semantic interpretation in generative grammar*. Cambridge: MIT Press.
Jackendoff, R. (1983) *Semantics and cognition*. Cambridge: MIT PRess.
James, W. (1890) *Principles of psychology*. New York: Rinehart and Winston.
Johnson, M. (2005) The philosophical significance of image schemas. In B. Hampe (ed.) *From perception to meaning: Image schemas in cognitive linguistics* 15–33. Berlin: Mouton de Gruyter.
Katz, J. and Fodor, J. (1963) The structure of a semantic theory. *Language* 39(2): 170–210.
Lafargue, P. (1898) Ursprung der abstrakten ideen. *Die Neue Zeit, XVIII*(2).
Lakoff, G. and Johnson, M. (1980) *Metaphors we live by*. Chicago: University of Chicago Press.
Lakoff, G. and Johnson, M. (1999) *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Chicago: University of Chicago Press.
Langacker, R. (1987) An introduction to cognitive grammar. *Cognitive Science* 10: 1–40.
Levelt, W. (1989) *Speaking: From intention to articulation*. Cambridge: MIT Press.
Locke, J. (1689/1995) *An essay concerning human understanding*. Amherst: Promethius Books.

Mach, E. (1896/1897) *Contributions to the analysis of sensations*. Chicago: Open Court Publishing Company.

Meier, B. and Robinson, M. (2004) Why the sunny side is up: Associations between affect and vertical position. *Psychological Science* 15(4): 243–247.

Meier, B., Robinson, M. and Clore, G. (2004) Why good guys wear white: Automatic inferences about stimulus valence based on color. *Psychological Science* 15(1): 84–87.

Murphy, G. (1996) On metaphoric representation. *Cognition* 60: 173–204.

Murphy, G. (1997) Reasons to doubt the present evidence for metaphoric representation. *Cognition* 62: 99–108.

Ornstein, R. (1969) *On the experience of time*. Hammondsworth: Penguin.

Papafragou, A., Massey, C. and Gleitman, L. (2002) Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84: 189–219.

Pinker, S. (1989) *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pinker, S. (1994) Mentalese. In *The language instinct* 55–82. New York: Harper.

Pinker, S. (1997) *How the mind works*. New York: Norton.

Price-Williams, D. R. (1954) The kappa effect. *Nature* 173(4399): 363–364.

Pullum, G. (1991) *The great Eskimo vocabulary hoax*. Chicago: University of Chicago Press.

Richardson, D., Spivey, M., Barsalou, L. and McRae, K. (2003) Spatial representations activated during real-time comprehension of verbs. *Cognitive Science* 27: 767–780.

Schubert, T. (2005) Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology* 89(1): 1–21.

Slobin, D. (1996) From 'thought and language' to 'thinking for speaking'. In J. Gumperz and S. C. Levinson (eds) *Rethinking linguistic relativity* 70–96. Cambridge: Cambridge University Press.

Snedeker, J. and Gleitman, L. (2004) Why is it hard to label our concepts? In Hall and Waxman (eds) *Weaving a lexicon*. Cambridge: MIT Press.

Sommer, F. (1963) Types of ontology. *Philosophical Review* 72: 327–363.

Talmy, L. (1988) Force dynamics in language and cognition. *Cognitive Science* 12: 49–100.

Torralbo, A., Santiago, J. and Lupiáñez, J. (2006) Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science* 30(4): 745–757.

Turner, M. (2005) The literal versus figurative dichotomy. In S. Coulson and B. Lewandowska-Tomaszczyk (eds) *The literal and nonliteral in language and thought* 25–52. Frankfurt: Peter Lang.

Wallace, A. (1870/2003) Contributions to the theory of natural selection. In *Contributions to the theory of natural selection* 400. London: Routledge.

Whorf, B. L. (1939/2000) The relation of habitual thought and behavior to language. In J. B. Carroll (ed.) *Language, thought and reality: Selected writings of Benjamin Lee Whorf* 134–159. Cambridge, Massachusetts: MIT Press.

Zakay, D. and Block, R. (1997) Temporal cognition. *Current Directions in Psychological Science* 6(1): 12–16.

18 Temporal frames of reference

Jörg Zinken

1 Introduction

Do people understand time in the same way across languages and cultures, or is our understanding of time culturally specific? On the one hand, anthropologists have often emphasised differences between the ways cultures interpret time (see Gell, 1992, and Munn, 1992, for reviews of the literature). On the other hand, some of the problems that conceptions of time address must be addressed by humans in all environments: human life is finite all around the globe, and humans live in groups which need to coordinate their activities. Maybe then there is some cognitive bedrock of thinking about time that is the same across languages and cultures?

One way of finding out is to look for universals in the way people across languages *talk* about time (Bloch, 1989). But although the anthropology of time is a vast research field with a long history, a systematic *linguistic* anthropology of time is less developed than one might expect (Levinson, 2004). This chapter discusses possibilities of making one aspect of such a linguistic anthropology of time more systematic. In particular, I will discuss possible heuristic contributions that typologies of *spatial* frames of reference might make to typologies of *temporal* frames of reference.

The observation that in English and many other languages the vocabulary used to talk about the location of objects in space is also used to talk about the location of events in time has attracted considerable interest (Clark, 1973; Fillmore, 1997 [1971]; Jackendoff, 1983; Lakoff, 1993). More recently, the *universality* of such vocabulary sharing has been hypothesised within the framework of Conceptual Metaphor Theory (Lakoff and Johnson, 1999). Within this framework, cross-linguistic studies assess the presence in the studied language(s) of metaphorical models such as TIME IS SPACE (Radden, 2003), TIME AS SPACE (Yu, 1998), or TIME PASSING IS MOTION (Ahrens and Huang, 2002).

These studies have begun to provide semantic evidence for universals in the cognition of everyday time to supplement the abundant anthropological evidence for diversity in time cognition in ritual contexts (Bloch, 1989; Senft, 1996). However, these studies have also highlighted methodological problems. Global models such as TIME IS SPACE or TIME PASSING IS MOTION need to be qualified and specified before they can be appropriate frameworks for typological research, but such qualifications and specifications have not been systematically made. These models need to be qualified because, as they stand, they might suggest that abstract English concepts such as *time*, *space*, or *motion* are universally relevant, which they clearly are not. The contexts in which the word 'time' is used by speakers of English are diverse; although some of these contexts might be universally relevant, others are unlikely to be (Evans, 2004).

Furthermore, these models need to be specified, because they are so general that constructions with diverse functions can be used as evidence for, for example, a TIME IS SPACE model. Taken together, the cross-linguistic irrelevance of terms such as 'time', 'space' and 'motion', and the generality of the proposed models, can lead to research that mirrors the unfortunate model of research into the universals of 'colour' terms: constructions with diverse functions are forced into a framework that has no validity for the languages studied (Saunders, 1995; Wierzbicka, 1996). As far as research on 'spatial time' is concerned, this global approach has indeed led to the formulation of universals on a rudimentary empirical basis. After all, nearly all of the languages in which the polysemy of spatiotemporal lexemes has been studied are spoken by urban speakers in industrialised societies (such as Chinese, English, Japanese, Turkish). Rare exceptions are Malotki's (1983) study of Hopi time, Moore's (2000) study of spatial metaphors for time in Wolof, and Nuñez and Sweetser's (Núñez and Sweetser, 2006) study of Aymara. The anthropological literature, in turn, contains many studies of time in non-industrialised cultures (Munn, 1992). However, the linguistic descriptions provided in these studies are usually not very detailed.

Models such as TIME PASSING IS MOTION might be inappropriate as a framework for a linguistic anthropology of 'spatial time'. Some framework, of course, is necessary if we want to draw any (cross-linguistic) generalisations. The more detailed our framework, the better our chances that we describe genuine cases of conceptualisation rather than researcher-induced artefacts (Lucy, 1997). With this in mind, I want to suggest here that existing typologies of spatial frames of reference can help in making useful conceptual distinctions for a semantic typology of 'spatial time'.

In the remainder of this introduction, I will briefly describe a distinction between two *kinds of time* commonly used in the philosophy of time. Further, a typology of *spatial frames of reference* will be briefly introduced. It will then be the aim of the main body of this chapter to bring the two together in developing a typology of temporal frames of reference that is detailed enough to serve as a framework for cross-linguistic investigation and generalisation.

The A-series and the B-series of time

What is time, anyway? In order to make sense of diversity across languages and cultures, we first need to have a good grasp of what we assume to be the universally experienced aspect of the world that English speakers refer to when they employ the word 'time'. Philosophical answers to this question can be categorised into two broad groups: The 'A-series' view of time and the 'B-series' view of time. This classification can be traced back to the philosopher McTaggart (1908), and it has been taken up more recently by Gell (1992) in his anthropology of time. The brief discussion in this section is based on Gell's work.

Time can be thought of as a series of events. But what exactly is it about events that gives them a temporal quality? Some philosophers argue that events constantly change their status, from belonging to the future to belonging to the present to belonging to the past. Time is this constant change in the status of events. The series of events constituting time conceived of in this way is referred to as *A-series time*, and theorists arguing that time is the flux of events from futurity through presentness into the past are referred to as *A-series theorists*. Other philosophers argue that events never change their status; they do not 'become' and 'fade', but simply 'are', like beads strung together on a necklace. Time, on this view, is the set of relations of anteriority and posteriority holding between events. The series of events constituting time conceived of in this way is referred to as *B-series time*, and theorists arguing that time is a never-changing network of anteriority/posteriority-relations are referred to as *B-series theorists*.

We hence end up with two kinds of time: the time of our subjective experience, for which future events have a different meaning than past events (the A-series), and the network of events as they objectively occur, quite independently of our interest or lack of interest in them (the B-series). Philosophers debate which of these characterisations reveals the ontological reality of time. A-series theorists argue that the A-series captures the ontological reality of time: futurity is an intrinsic property of future events, and pastness is an intrinsic property of past events. B-series theorists argue that the B-series is ontologically real: events occur when they occur; futurity and pastness are assessments which we bring to events due to our active orientation towards the world.

The distinction between an *A-series* and a *B-series* of time originates from a metaphysical debate, i.e. the question which 'kind' of time is 'basic' and ontologically real. Of course, the aim in this chapter is not to enter into metaphysical debates, but to provide a framework for comparing the semantics of everyday time reference across contexts. The distinction between A-series and B-series is only useful in the current context if it can be translated into different types of everyday time reference.

Intuitively, it does seem that we make a distinction between the two kinds of time – the A-series and the B-series – in our everyday life. The A-series is what we experience as we coordinate our everyday activities and grapple with the finiteness of our existence – or of our time until the next deadline. The B-series is the real-world foundation for a culture's inventory of event-types embodied in calendars. Furthermore, the future-present-past stream of the A-series and the before-after chain of the B-series are expressed using different vocabularies, both of which are also employed for talking about spatial relations. The A-series is the kind of time grammaticalised in many languages in the category of tense, which in many languages is marked by morphemes derived from motion verbs corresponding to English 'come' and 'go' (Bybee, 1994; Traugott, 1978). Also, consider the following expressions of A-series time:

(1) a. I have a fun afternoon in front of me.
You have a hard week behind you.

b. I am looking forward to tomorrow.
I look back at my childhood.

In (1a), events are marked as being in the experiencer's future or past by placing them in front of the experiencer or behind him, respectively.[1] In (1b), the experiencer's active orientation towards events in the future or in the past is expressed using perception verbs: In the perceptual field in front of the experiencer, future events can be anticipated, in the field behind the experiencer, past events can be scrutinised.

Notions of *front* and *back* are also used to talk about anteriority/posteriority relations (the B-series of time), however, using different expressions in English:

(2) a. The 21st April is before the 22nd April.
Thursday comes after Wednesday.

b. We'll meet in the week following Easter.
Tuesday is ahead of Wednesday.

It is an unchanging quality of the 21st of April that it occurs before the 22nd of April (within the year), and it is an unchanging quality of Thursday that it occurs after Wednesday (within the week). The time at which I make the statements in (2a) does not matter for the interpretation of the temporal reference – i.e. it is a reference to B-series time. In these examples, a form historically expressing the spatial relation *front* is used to express the temporal relation *anteriority*, and a form historically expressing the spatial relation *back* is used to express the temporal relation *posteriority*. While *before* and *after* express static relations, the same conceptualisation of 'spatial time' can be conventionally expressed in English with terms expressing relations in motion events: In (2b), the posteriority relation of the meeting to Easter is expressed by locating it 'behind' Easter using the form *following*, and the anteriority relation of Tuesday to Wednesday is expressed by locating it 'in front of' Wednesday using the form *ahead of*.

Intuitively, and on the basis of some suggestive data as discussed above, it seems that the distinction between two kinds of time is cognitively real for speakers of English. For the purpose of this chapter, it will be assumed that both the experiencer-centred understanding of time as a series of future, present, and past events, and the experiencer-independent understanding of time as a series of before/after relations between events are universal temporal experiences. Furthermore, we have seen that in both contexts concepts of *front* and *back* are involved (at least historically) in temporal conceptualisation in English. The distinction between A-series and B-series might therefore be of value for a typology of temporal frames of reference. However, the characterisations provided in the philosophical and anthropological literature to explain how people make temporal sense of these event-series are not precise enough for our purposes. A-series time is characterised as a *stream* of events going past the experiencer. B-series time is characterised as a static *chain* of events (Gell, 1992). While these metaphors are suggestive, they are hardly a good basis for cross-linguistic comparison. We need a more precise language to address our question: Where does the association between the ideas of *front* and *future* in the case of A-series reference, and *front* and *(temporal) anteriority* in the case of B-series reference come from? More specifically: what exactly is the analogy between locating objects in space and locating events in time? To answer these questions, we first need to find out what the logic of the reference systems is which are used to locate objects in space.

Spatial frames of reference

Three frames for locating objects and places in space are commonly used across languages: The *intrinsic* or *ground-based* frame of reference, the *absolute* or *field-based* frame of reference, and the *relative* or *projector-based* frame of reference (e.g., Levinson, 2003; Talmy, 2000).[2] The brief description in this section is based mainly on the work of Levinson (1996a; 2003).

Spatial frames of reference are constituted by three logical entities: the object to be located (the *figure*), an object with a known location which is used to locate the figure (the *ground*), and an object which determines the search space to be projected from the ground (the *origin of the coordinate system*).

In the *intrinsic frame of reference*, ground and origin are conflated: the ground object is also the origin of the coordinate system. For example, *The computer is in front of me* locates the computer using an intrinsic frame of reference. 'I' am the referential ground, and the asymmetry of my body also determines what is to be understood by the relator *front*, i.e. how the search space is to be projected from me. Asymmetric inanimate objects are also often thought of as having intrinsic *fronts* and *backs*. An utterance such as *The bike is in front of the house* can be understood in this way. The *front* of an inanimate object is often the side that people canonically interact with. In the case of *houses*, the *front* side would typically be the side facing the street, where the door to the house is located. Intrinsic frames of reference take diverse forms across languages, but the logic of an intrinsic frame of reference seems to be universally used to locate objects in space. The reason for this might be that intrinsic frames of reference are relatively simple: because ground and origin are conflated, reference within an intrinsic system requires the understanding of only a *binary* relation (Levinson, 1996a,b).

In the *absolute frame of reference*, the environment in which the ground object is located provides a field which is organised in such a way that it can be used to determine a search space; the environment here constitutes the origin of the coordinate system. Familiar examples are the cardinal points *north*, *west*, *south*, and *east*. The utterance *Hamburg is north of Bielefeld* is comprehensible because the cardinal directions provide a grid running across the globe (and through Bielefeld, the referential ground). But the environment used for absolute reference can also be more concrete and localised. For example, a bowling lane can provide an absolute origin. Suppose a team of weak bowlers have only managed to toss the bowls about half-way towards the pins. Bowls lying still do not have intrinsic fronts, so I cannot use an intrinsic frame of reference to locate a particular bowl in relation to another one. Still, I can refer to the blue bowl lying *behind* the red bowl, meaning that it is further away from the pins. Due to its directedness, the lane can serve as the field (or origin) of reference. Finally, absolute origins can also be temporary: in the utterance *John is behind Mary in the queue*, the directedness of the queue determines how the relator *behind* is to be understood (Talmy, 2000).

In the *relative frame of reference*, an observer constitutes the origin of the coordinate system. The speaker's coordinates *front*, *back*, *left* and *right* are projected onto the referential ground. The details of this projection differ across and within languages. For example, the speaker's coordinates can be 'reflected' from the ground, as if the ground object was another observer 'facing' the actual observer. The utterance *The ball is in front of the tree* is understood in this way: the ball is *between* the tree and the observer, the tree's *front* is the side 'facing' the observer. However, in other contexts, the projection involves not *reflection* but *translation*, where the orientation of the observer is 'carried over' onto the ground object: A ball *to the left of* the tree is to the left from the observer's point of view, not from the point of view of an observer 'reflected' in the tree.

2 Temporal frames of reference

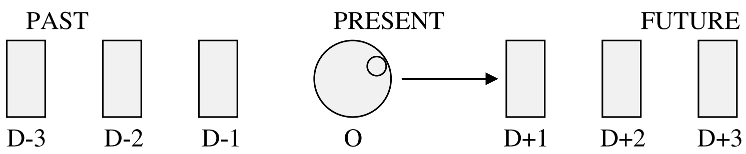
Do analogous *temporal frames of reference* exist? Can the technical terms as elaborated in work on spatial conceptualisation be of heuristic value in the description of space-time analogies used for temporal reference across languages? In this section, I aim to develop a typology of everyday (spatio-)temporal frames of reference on the background of the philosophical distinction between the A-series and the B-series of time.

Locating events in A-series time

'A-series' time is the subjective experience of a constant change in the status of events, from their futurity to presentness, to pastness. The futurity or pastness of an event can conventionally be expressed in English by placing the event *in front of* or *behind* the experiencer respectively, as in (1) above, repeated here:

(1) I have a fun afternoon in front of me.
You have a hard week behind you.

As we can now see, these utterances locate events within an intrinsic frame of reference. The defining feature of intrinsic frames of reference is their binary structure (Levinson, 2003): a figure entity is located in relation to a ground entity, and the ground is also the origin of the coordinate system. From *meetings*, through *afternoons* and *years*, to a whole *life*, events of varying regularity and temporal scope can conventionally be referred to as lying *in front of* us or *behind* us in English. It seems that for speakers of English, large-scale time intervals (such as *days*, *seasons*, *the duration of the world*) are abstracted from the actual environment as an additional, imaginary 'landscape' on which events can be (quasi-)visualised. [3] The conceptualisation of large-scale temporal intervals as a landscape affords the use of a viewer-centric, *relative frame of reference* for the localisation of events. The relative frame of *spatial* reference locates an object with respect to the ground from the point of view of an observer. With respect to the location of events in time, some authors have proposed that expressions such as *the day after tomorrow* are understood in a relative frame of reference (Radden, 2003: 12). Radden illustrates the spatial logic of this expression with the following figure:



D-3 – D+3: Days
O: Location of the observer at day 0

Figure 1. A vision-based understanding of temporal relations (adapted from Radden, 2003)

The blocks in Figure 1 symbolise days, and the *day after tomorrow* is the one 'behind' the one the observer is 'looking' at, *tomorrow*. In other words, *tomorrow* is the (primary) referential ground, the speaker's *now* is the origin of the coordinate system. The deictic nature of *tomorrow* surely supports a relative reading in Radden's example, but the present argument should apply also to *the day after Tuesday*. This relation can be understood in a relative, quasi-visual manner, if it is understood that I am talking about a particular *future* Tuesday.[4]

The function of using a relative frame of reference to locate events in A-series time might be that it allows more precision when talking about large-scale time intervals beyond the *now* than the intrinsic frame of reference and deictic expressions do. When talking about plans for the immediate future, we are more likely to use a deictic expression without a frame of reference: We would say *I'll send this e-mail in a moment* rather than, e.g., *I'll send this e-mail* 20 *seconds before a minute has passed*. However, in time-scales which go beyond the *now* (a border that is itself likely to vary across cultures, within and across language communities), simple deictics are not very useful: *I'll get some crisps before the match* is a more relevant information than the deictic *I'll get some crisps in* 3 *hours' time*.

Locating events in B-series time

The B-series of events is the time of anteriority/posteriority relations. While events arise in and fade from the field of our experience, their temporal relations to all other events never changes. In English, the words 'before' and 'after' express relations of anteriority and posteriority respectively. Let's take a closer look at the example just used:

(3) I'll get some crisps before the match.

Example (3) can be understood in a *relative*, quasi-visual manner, as discussed in the preceding section. When a future ground event ('match') and the observer 'face' each other as in a canonical encounter, a figure event ('getting crisps') between the ground event and the observer is 'in front of' the ground event. The temporal relation between crisps-getting and match in (3) might therefore be understood in a way that is analogous to the spatial reference in *The ball is in front of the tree*.

However, in a relative frame of reference the origin of the coordinate system is the observer. A spatial scenario anchored in an observer is incoherent with the observer-independent nature of relations in B-series time. If it was the case that we could understand expressions such as *I'll get some crisps before the match* only in an observer-centred manner, it would mean that we could not express the observer-independent nature of B-series time. But this is implausible: we can easily be aware of the unchanging anteriority of the crisps-getting event relative to the match. The question is therefore whether the temporal relation between getting crisps and watching a match in (3) is *necessarily* understood in a relative manner.

In their work on the use of spatial frames of reference, Levinson and colleagues have employed rotation tasks to distinguish between different frames of reference (Levinson, 2003). Would it be possible to use an analogue of such rotation tasks to find out what kind of coordinate system is used in temporal reference? In order to do this, we would need to refer to the same figure-ground relation from the opposite temporal perspective.

Suppose, then, that I want to refer to the figure event in (3) in relation to the same ground event a day later. 'Looking back' onto the same events, I now say: 'Yesterday, I got crisps _____ the match'. If we understood the temporal relations between these events in an observer-anchored manner, the correct word to fill the gap would now have to be 'after' or another expression of *behind*-ness, because the figure event (getting crisps) is now no longer between the match-watching event and me. However, the correct word to fill the gap remains 'before'.[5]

This suggests that temporal relations between *past* events are understood in English employing a coordinate system that is *independent* of the observer, and while temporal relations between *future* events can be understood in a relative manner, they, too, should be understandable in an experiencer-independent way. In other words, the spatial logic of the temporal reference in *I'll get some crisps before the match* might be ambiguous. While deictic cues such as future tense and adverbials such as 'tomorrow' might prompt a relative understanding, non-finite expressions (*I always get crisps before a match*) might prompt an experiencer-independent understanding.

But how do speakers locate events 'in front of' ('before') other events in an observer-independent frame of reference? One possibility is that speakers make use of an intrinsic frame of reference in these contexts (Bender, Bennardo and Beller, 2005; Yu, 1998). When locating objects in space in English, expressions of *front* ('in front of') and *back* ('behind') can be used in this way. For example, the spatial reference in *He is sitting in front of the TV* in most situations is intended in its intrinsic, rather than relative interpretation. Spatial reference in an intrinsic frame of reference is independent of the position of the observer: the referential ground object constitutes the origin of the coordinate system.

In order to locate events in time in an intrinsic frame of reference where the observer's *now* is not part of the referential scene, we would need to be able to identify the search interval in which the figure event (*getting crisps* in (3)) takes place on the basis of intrinsic features of the ground event (*watching a match*). Can an event have an intrinsic front? The fact that this sounds like a funny idea should not deter us from entertaining the possibility. After all, the notion of *intrinsic front* is problematic also when applied to objects: The front side of a TV is not really intrinsic to the physical object, but determined by the way people canonically interact with TVs (Levinson, 2003).

It could be that the conceptualisation of events as 'moving' suffices to assign a *front* side to them. Consider symmetrical objects. Like events, balls do not have intrinsic fronts. Nevertheless, when a ball is rolling, we easily assign a *front* and *back* based on the direction of motion (Fillmore, 1997 [1971]; Svorou, 1994); thus, football players run *behind* a ball. Similarly, events in B-series time might be conceptualised as a train, with each carriage representing an event (Yu, 1998). Irrespective of the position of an observer, the first cabin, defined by the direction of motion, will always be in front of the second one; similarly, anterior events always remain 'in front of' posterior ones. This account suggests that by invoking the idea of 'moving events' we can understand the space-time analogy in an expression like *I always get crisps before a match* within a binary figure-ground frame, i.e. in an intrinsic frame of reference. Ultimately, intrinsic temporal reference might be based in quite literally spatial front/back relations: The sun moves across the sky *ahead of* the moon. *Day comes before night*, *night comes after day*, and *one day comes after the other* might be the clearest cases of such motion-based intrinsic temporal reference.

But is it plausible to assume that speakers of English conceive of the events in *I always get crisps before a match* as moving? This seems counterintuitive, and, to be sure, it is not evident from linguistic data. While it is conventional to speak of calendaric event types (*Christmas*, *spring*) and other event types that are part of a natural cycle (*the evening*, *the morning*) as *coming* and *going by*, the same is much less felicitous when applied to singular events (?*The match is coming*). A more prudent account might be to suggest that *I always get crisps before a match* is understood in an absolute frame of reference, with the *day* as the origin of the coordinate system. The *before* relation means that the *crisps-getting event* is closer to the beginning of the *day*, the implicit secondary reference interval, than *the match*, the primary reference event. Conventionalised intervals, such as the *day*, provide a directed field for such absolute reference, in analogy to *people in a queue* (Talmy, 2000). Although the directedness of a queue ultimately derives from the canonical movement towards the goal of this queue, it maintains its directedness even when there is no motion. Similarly, events throughout a day can be thought of as 'adding up' one after the other at their specified dates[6] along the temporal field much like people forming a queue do, rather than as moving like bowls rolling one behind the other on a bowling lane. Such a 'motionless' account is advantageous also because some languages do not seem to use the idea of objects moving through space to think about temporal relations between events (Bohnemeyer, 1997). However, we would not want to deny speakers of such languages the ability to speak or think about unchanging relations in B-series time.

Absolute reference to sequentiality relations does not require the speaker/hearer to specify a particular directedness of the field along which events are located. However, such directedness is necessary when communicating visually, rather than vocally, about (B-series) time, e.g., in co-speech gesture. In cultures with a writing system, the direction seems to be imported from the relevant conventions of using visual media, such as written language or comics. For example, speakers of Spanish assume by default that events displayed on the left side of a computer screen happened earlier than events displayed on the right side of a computer screen (Santiago, 2005). Arabic speakers asked to arrange objects representing a day's activities on a plane arrange these from right to left (Tversky, Kugelmass and Winter, 1991). Speakers of Mandarin produce downwards gestures when talking about a time in the afternoon (irrespective of whether it is the afternoon of the same, a future, or a past day), and upwards gestures when talking about the morning.[7] With respect to the question of the metaphoricity of temporal understanding, it is important to bear in mind that such figurative specifications of temporal 'directions' are not part of the conceptual structure employed in thinking for speaking, but of that employed in thinking for gesturing, i.e. in a visual medium of communication, which is by necessity one big 'spatial metaphor'.[8]

3 Temporal frames of reference and other generalisations

Probably the most widely used generalisations in the study of space-time analogies across languages are the Moving Time and Moving Ego metaphors, first introduced by Clark (1973) and Fillmore (1997 [1971]). These two models have been reformulated in various ways, as TIME PASSING IS MOTION OVER A LANDSCAPE, TIME PASSING IS MOTION OF AN OBJECT, and the further generalisation TIME PASSING IS MOTION (Lakoff, 1993). In the Moving Time model, time is viewed as a 'highway consisting of a succession of discrete events' that are 'moving past us from front to back'. In the Moving Ego model 'we are moving along [time], with future time ahead of us and the past behind us' (Clark, 1973: 50). Both of these models thus describe the A-series of time. In terms of the frames of reference introduced here, both models involve an intrinsic (or possibly relative) frame of reference, and combine this with the idea of motion – either the motion of events, or the motion of the experiencer. However, while such models might indeed be operative for speakers of English, it is usually not possible to conclude this from linguistic data. Temporal reference often employs motion constructions, as for example in the utterance *the evening is coming*. This is a deictic reference which does not employ a frame of reference. Temporal reference also often does employ frames of reference, for example the intrinsic one in the utterance *I have a great evening in front of me*. But it is not common to talk of time as *moving* past the experiencer *from front to back* (?*A great evening is coming in front of me*). In cross-linguistic research, it would be more prudent to treat these two examples as different types of temporal reference rather than as evidence for one general TIME PASSING IS MOTION model. Such generalisations are better treated as complex models, i.e. as combinations of several more fundamental conceptualisations. Evidence for such complex models must be sought in non-verbal data.

Some authors have proposed distinguishing two frames of reference used for locating events in time: an ego-based or ego-reference-point (ego-RP) frame and a time-based or time-reference-point (time-RP) frame (Moore, 2000; Núñez and Sweetser, 2006). This terminology is somewhat unclear in so far as there are two reference points (or reference intervals) in temporal reference: a primary one, the ground, and a secondary one, the origin of the coordinate system (Talmy, 2000, Levinson, 2003). The explication in Núñez and Sweetser (2006) suggests that the reference point they have in mind is the primary reference point, or ground of reference. If the RP in the suggested distinction between ego- and time-RP is to be understood as the primary reference point, the English examples discussed in this chapter should be classified as follows:

Table 1. The primary reference point as the basis for classification

ego-RP (ego=primary RP)	time-RP (event=primary RP)
I have a fun afternoon in front of me	The 21st April is before the 22nd April
-	Wednesday is after Tuesday
	One day comes after the other
-	The day after tomorrow
-	I'll get some crisps before the match
-	I always get crisps before a match

This classification seems wrong. It is the explicit aim of Núñez and Sweetser (2006) to separate reference to subjective past or future from reference to anteriority/posteriority relations. In this respect, *Wednesday is after Tuesday*, which describes sequentiality, should not be in the same category with *The day after tomorrow*, which refers to the speaker's future.

It seems to me that what Núñez and Sweetser (2006) actually have in mind is a distinction that is similar to the one between the A-series and the B-series in the philosophy and anthropology of time. If this is correct, the reference point in question would be the secondary reference point, or the origin of the coordinate system. The English examples discussed in this chapter would then fall into the two categories as follows:

Table 2. The secondary reference point as the basis for classification

ego-RP (ego=secondary RP)	time-RP (event=secondary RP)
I have a fun afternoon in front of me	The 21st April is before the 22nd April (RP=month)
The day after tomorrow	Wednesday is after Tuesday (RP=week)
	One day comes after the other (RP=day)
I'll get some crisps before the match	I always get crisps before a match (RP=day)

The distinction between the A-series and the B-series, or between types of secondary reference points as conceived in table 2, is an important one, the lack of which has led to a confusion of (A-series) *past* with (B-series) *anteriority*, and *future* with *posteriority* in earlier research (see the next section). However, as a typology of systems for locating events in time it is less precise than the distinction between the intrinsic, relative, and absolute frame of reference. In this classification, *I have a fun afternoon in front of me* and

I'll get some crisps before the match are grouped together as the same type of reference. However, they do differ in the way they locate events in A-series time. The relative frame of reference allows more specific reference to the temporal location of events in A-series time beyond the *now*. The price for this is an increase in cognitive complexity: while the intrinsic relator 'in front of' specifies a binary relation, the relative relator 'before' specifies a ternary relation. Similarly, the statement that *Wednesday is after Tuesday* is only true within the absolute frame of the week, whereas the reference *one day comes after the other* probably makes use of an intrinsic frame of reference.

Ultimately, it seems that frameworks that are based on the quality of a particular reference point run into problems. In this chapter, I have argued that it might be better to use a typology that is based on types of coordinate systems. In sum, the classification that I propose looks like this:

Table 3. Temporal frames or reference

A-series		B-series	
I have a fun afternoon in front of me	Coordinates: intrinsic Origin: speaker PRP: speaker	One day comes after the other	Coordinates: intrinsic Origin: day PRP: day
You have a tough week behind you	Coordinates: intrinsic Origin: addressee PRP: addressee	Wednesday is after Tuesday	Coordinates: absolute Origin: week PRP: Tuesday
I'll get some crisps before the match	Coordinates: relative Origin: speaker PRP: match	I always get crisps before a match	Coordinates: absolute Origin: day PRP: match
The day after tomorrow	Coordinates: relative Origin: speaker PRP: tomorrow		

The distinction between experiencer-centred (A-series, ego-RP) and experiencer-independent (B-series, time-RP) time *together* with a typology of frames of reference that are used to construct these 'kinds' of time provide a reasonably fine-grained framework for the systematic exploration of universals and diversity in space-time analogies.

4 Universals and diversity in spatial time

We are now in a position to discuss how the distinctions made in this chapter can help us to integrate existing data, ask new questions, and formulate hypotheses about universals of spatial time.

Forms expressing spatial relations of *front* and *back* regularly express anteriority and posteriority across languages. Furthermore, it seems that, as in English, expressions of *front* always express *anteriority*, and expressions of *back* always express posteriority (Haspelmath, 1997).[9] A few examples are presented in (4–6).

(4) Kwaio (Keesing, 1991)
(a) na'o-na mae i Gwee'abe
‚before the battle at Gwee'abe', literally, front-of battle'

(b) buri-na mae i Gwee'abe
‚after the battle at Gwee'abe', literally 'back-of battle'

(5) Hopi (Malotki, 1983)

pam	put	hihin	a-pyeve	tìi-ti-wa
that	that	somewhat	he-before[10]	child-CAUS-PASS.PERF

'He was born a little bit before him'

(6) Wolof (Moore, 2000)

Ci	gannaaw	la ñów.	
LOCPREP	back/behind	NONSUBJ.FOC.3	come.

'At back she came.'
'She came afterwards'

Temporal relations of sequentiality (B-series time) using the relators *front* and *back* can be understood in an absolute frame of reference, and possibly in an intrinsic frame of reference, if speakers think of events as 'moving'. As in the domain of space, the linguistic data alone do not allow us to decide whether the expressions in (4–6) are understood absolutely or intrinsically. We need additional data sources to answer this question. Unfortunately, the use of rotation tasks, which make it possible to distinguish between spatial frames of reference, has its limits in the domain of time. Alternatively, co-speech gesture might be a valuable source of data, which can answer the question, e.g., whether speakers habitually think of events as moving.

Similarly, it seems premature on the basis of our current knowledge to be too sure that all languages use the relators *front* and *back* to express sequential relations. Some languages might not at all explicitly mark anteriority and posteriority relations, relying instead on context and iconicity: the event mentioned earlier happened earlier (Bohnemeyer, 1997). Furthermore, languages which prefer an absolute frame for spatial reference based on the movement of the sun might use the same vocabulary to talk about sequentiality in time (*the morning is east of the evening*).

Absolute temporal reference requires that temporal intervals, which provide the secondary ground, i.e. the origin of the coordinate system, be understood as bounded entities, or 'fields'. The fundamental space-time analogy is that between the *beginning* of the unfolding of a temporal interval and a field's *front*. The reason for the possible universality of absolute temporal reference might be that relations within an absolute system remain the same when the 'viewpoint' of the observer changes. In a relative frame of reference, relations between figure and ground change when the observer's viewpoint changes. Of course, for human experiencers it is quite impossible to hold their temporal 'viewpoint' onto the world constant. An absolute frame of reference might therefore be the only viable system for talking about unchanging anteriority/posteriority relations.

Events in A-series time can be located in an intrinsic frame of reference. Could the particular association of an experiencer's *front* with his or her *future*, and an experiencer's *back* with his or her *past*, be universal? There is some ground for entertaining such an assumption. Anthropologists of time maintain that thinking about the *immediate future*, i.e. the work at hand, the time that is still part of the *now* rather than a *then*, is the fundamental context from which societal organisations of time arise (see Gell, 1992). The immediate future is constantly apprehended and enacted in (spatial) practice, e.g., manual work and gaze. It seems plausible enough to think that speakers universally might use their front space symbolically, e.g., for gesture-supported planning of imminent tasks, but we know too little about temporal cognition across cultures to say. In this context, it is questionable whether we are dealing with a *metaphorical* association between front space and immediate future. After all, the idea that we do actually perceive the immediate future has been discussed at least since Husserl introduced the notion of *protentional consciousness*, which anticipates what lies just at the boundaries of the *now*, the current time interval. In so far as gaze and manipulation are relevant at all for protention, this form of consciousness will be directed to the body's front space.

However, not only the immediate future which is still part of the *now* is conceptualised as being in front of the experiencer in English. Large intervals of subjective time, such as the *future*, also 'lie' in front of us. The orientation towards large time-'scales', such as the relatively abstract English concept of *future* is very different perceptually, conceptually, and linguistically from thinking about *immediate future* – it involves an imaginary 'leap', the abstraction of conventional time intervals as an additional dimension. With relation to such 'larger-scale' temporal concepts, the association of *front* with *future* is not universal. Several authors have claimed that in particular languages and cultures, subjective future is conceptualised as lying *behind* the speaker, whereas past events are in the observer's visual field on a temporal landscape (Alverson, 1994; Clifford, 2004; Dahl, 1995; Klein, 1987; Miracle and Yapita Moya, 1981). The linguistic analyses supporting such arguments are often sketchy, and seem at times to have been misguided due to conceptual confusions between A-series and B-series time (for critical reviews, see Moore, 2000; Núñez and Sweetser, 2006; Shinohara, 1999). However, the analysis by Miracle and Yapita Moya in relation to Aymara has been supported by converging evidence from co-speech gesture research (Núñez and Sweetser, 2006). Núñez and Sweetser found that Aymara speakers would produce hand gestures forward from their body when talking about past events in the community's history, but would produce gestures towards their back when explicating the meaning of the word *future*.

It seems plausible that the grade of figurativity of a temporal conceptualisation is related to its cultural specificity, such that the more figurative the analogy between spatial and temporal relations, the more restricted it is across cultures. Metaphoric 'leaps' seem to require a strong cultural scaffolding to be successful (see also Evans and Wilkins, 2000). A cultural factor that might contribute to the gestures of Aymara speakers is the emphasis that is placed on being precise about the source of one's knowledge, which is grammaticalised in Aymara in the category of evidentiality (see Aikhenvald, 2004). When Aymara speakers make a predication, their grammar requires them to mark whether they have seen the reported event themselves or not (Miracle and Yapita Moya, 1981). Since predications about future events are necessarily predictions, they cannot have been eye-witnessed, which might contribute to their being conceptualised as lying behind one's back.

Finally, events in A-series time can be located in a relative frame of reference in English and related languages. Whether this also occurs more widely across languages is impossible to say, because, as in English, the linguistic form alone might not be sufficient to tell whether an utterance is understood in a relative or an absolute frame of reference.

Relative frames of *spatial* reference differ across cultures in the way in which the observer's coordinates are mapped onto the referential ground, as described earlier. In Hausa, this mapping of coordinates involves *translation* whereas it involves *reflection* in English. *The ball is on front of the tree* under a 'translational' understanding means that it is on the other side of the tree from where the speaker is standing. Temporal reference in a relative coordinate system seems to show analogous diversity (Bender, Bennardo and Beller, 2005; Hill, 1978). In Hausa, speakers 'view a later day of the week as *gaba da* 'in front of/before' an earlier one, an earlier day as *baya da* 'in back/ of after' [sic] a later one' (Hill, 1978: 528). Hill does not specify whether Wednesdays are *always* 'in front of' Tuesdays (in which case it would be an expression of 'B-series' time, and as such independent from any experiencer, i.e. it could not be an instance of a relative frame of reference at all), or whether this applies only to days in Ego's future, in which case it is a relative (translational) expression of relations in A-series time. Hill's analysis, though, suggests that the latter is the case.

Why is the future beyond the *now* located in front of us in English and related languages? One factor might be the importance that is placed on the precise planning of one's (future) time in our culture. The quasi-visual, if only imagined, access that the relative frame of reference imposes on temporal conceptualisation of future supports such planning by providing a 'space' that can be used for planning in imagination, and for temporal reference in language and in co-speech gesture. The use of a relative frame of reference converts relations in B-series time (relations of anteriority and posteriority between events) into an imagined space that is subject to personal planning and time-'reckoning'.

Languages might differ not only in terms of the overall repertoire and the precise characteristics of temporal frames of reference, but also in terms of the preferred frame of reference for a given context. Such differences might exist even between closely related languages. For example, it seems that speakers of German prefer an absolute frame of reference where speakers of English frequently use a relative frame of reference. When asked to disambiguate the sentence *The meeting planned for next Wednesday has been moved forward two days*, some speakers of English interpret *forward* to mean *later*, as would be expected when using a relative (translational) perspective, whereas others interpret *forward* to mean *earlier*, as would be expected when using an absolute perspective, with the (beginning of the) week as the origin of the coordinate system

Final proofs with index 24 February 2010

(see McGlone and Harding, 1998). Speakers of German, in a separate experiment, consistently chose the absolute solution (Bender, Bennardo and Beller, 2005).

These differences might be indicative of a more general preference for viewer-centred time reference in the case of English, and event-centred time reference in the case of German. Stutterheim, Carroll, and Klein (2003) found that speakers of English predominantly chose a viewer-centred strategy when re-telling the events in a short film, in which the film is retold as if it was playing again before the mind's eyes, with new events introduced with a 'and now I see' phrase. Speakers of German, on the other hand, predominantly chose a strategy which meant that they seemed to arrange the events 'like a string of pearls' (Stutterheim, Carroll and Klein, 2003: 108) and mark the posteriority of a new event with an 'and then' phrase. Stutterheim, Carroll, and Klein (2003) relate these differences to the grammaticalisation of an 'ongoing' aspect in the English progressive *–ing* form, which is absent in German.

5 Conclusion

In this chapter, I have suggested some conceptual distinctions that might be useful for systematic data collection and analysis in cross-linguistic research on 'spatial time'. The suggested framework integrates previous data and opens a range of new questions: Are anteriority/posteriority relations always understood in an absolute frame of reference, or can they be understood in an intrinsic frame of reference? Are the A-series and the B-series universal kinds of time, or do contexts of temporal reasoning exist which are constructed as 'A-series' by using a relative frame of reference in one language, but are constructed as 'B-series' by using an absolute frame of reference in another language? What are the relations between types of time intervals (cyclic vs. non-cyclic; 'punctual' moments vs. longer events; events in the immediate future vs. events in the further future vs. events in the past) and the use of different frames of reference? Systematic data from a more varied sample of languages and cultures are needed before we can attempt empirically grounded conclusions about possible universals in the domain of space-time analogies.

The implicit or explicit (Bloch, 1989) assumption in the anthropology of time has been that time reference in everyday contexts, as opposed to ritual contexts studied intensively in the anthropological literature, might display many universalities across languages and cultures. However, everyday life is a complex beast, and to make sure that we compare like with like across languages, we need to distinguish not only experiencer-centred (A-series, ego-RP) time from experiencer-independent (B-series, time-RP) time, but also frames of reference and the contexts in which they operate. We might find that a relativised 'view' of a temporal landscape stretching out into the future in front of us is not so much a universal and natural feature of human mind, but rather an exotic development in cultures that have developed a strong interest in 'reckoning' and 'telling' time.

Acknowledgements

I am grateful to Alan Costall, Kevin Moore, Vera da Silva Sinha, Chris Sinha, and an anonymous reviewer for their valuable comments on an earlier draft. Work on this chapter was supported by grants from the British Academy and the European Union's 6th framework programme 'What it means to be human'.

Notes

1 This chapter discusses only one aspect of the complex ways in which the temporal structures of events are communicated, and the example utterances are kept simple to avoid some of these complexities. In terms of Klein's (1994) approach to the communication of temporal relations, we deal here only with temporal relations between time intervals in TSIT, the 'time of the situation' talked about. The complex relations between TSIT, 'topic time' (TT) and 'temporal anchor' (TA) of the speech event are beyond the scope of this chapter. But these complexities at least need to be acknowledged. Thus, in an utterance like '*Once, I had a great future in front of me*', we might still say that 'front' expresses *futurity* within TSIT, but tense and the temporal adverbial 'once' establish TT as lying in the past relative to TA, the time of the speech event.

2 Levinson and his research group distinguish between intrinsic, absolute, and relative frames of reference, while Talmy has introduced the distinctions between ground-based, field-based, and projector-based reference. These terms overlap to a large extent and can, for present purposes, be treated as synonymous. To minimize confusion I will employ the terminology of intrinsic, absolute, and relative frames of reference.

3 It may be that in cultures where vision is not conceptualised as the most central modality in the acquisition of knowledge (see Evans and Wilkins, 2000) concepts of temporal intervals as a 'landscape' are less relevant.

4 The quasi-visual conceptualisation of future time in English is further illustrated by the use of visual perception verbs in conventional expressions such as *I'm looking forward to the time after Easter*.

5 Of course, we might argue that the spatial meaning of 'before' is not relevant in this case, and that the relevant meaning is, say, 'earlier than X_{ground}'. However, the next question then becomes: why has 'before' acquired the general meaning 'earlier than X_{ground}' rather than 'temporally between *now* and X_{ground})'?

6 The term 'date' as used here derives from the philosophy of time (see Gell, 1992). It refers to the real-world spatio-temporal coordinates of an event, and does not imply the existence of a calendar, as the everyday use of the word 'date' does.

7 I am only aware of anecdotal evidence for this so far. However, the association between up-down relations and anteriority-posteriority relations in Mandarin is also evidenced in conventional expressions, such as 'shang-ban-tian', literally upper-half-day, meaning 'morning; forenoon' (Yu, 1998, p. 110).

8 The claim that temporal relations between events beyond immediate future can be understood in an absolute as well as a relative manner is currently based on intuition and linguistic data. It should be possible to obtain independent evidence by operationalising co-speech gesture. Speakers of English and related languages tend to make left-to-right gestures when talking about sequences in B-series time, but they produce forward

gestures when talking about sequences in A-series time. If event sequences which are part of a speakers time plan are conceptualized in a 'visualised', relative way (such as the tasks on a given day), but event sequences not immediately relevant to personal time planning (such as, maybe, the seasons) are located in a field-based frame of reference, speakers' co-speech gestures should differ across these two contexts.

9 Haspelmath (1997) provides examples of adverbials expressing both temporal and spatial anteriority and posteriority from a sample of 55 languages. He states that 'almost all cases' (p. 56) follow this path, but he does not provide an example of a different case.

10 Malotki (1983) points out that the morpheme -*pyeve* 'before' itself is related to the locative suffix –*ve*, meaning 'before a (moving) object'. According to Malotki (p. 92–93), the antonymic suffixal element –*ngk* 'after a (moving) object' means temporally *after* in a sequence. However, Malotki does not provide an example for this use.

References

Ahrens, K. and Huang, C.-R. (2002) Time passing is motion. *Language and Linguistics* 3(3): 491–519.

Aikhenvald, A. Y. (2004) *Evidentiality*. Oxford: Oxford University Press.

Alverson, H. (1994) *Semantics and experience: Universal metaphors of time in English, Mandarin, Hindi, and Sesotho.* Baltimore and London: John Hopkins University Press.

Bender, A., Bennardo, G. and Beller, S. (2005) Spatial frames of reference for temporal relations: A conceptual analysis in English, German, and Togan. In B. G. Bara, L. Barsalou and M. Bucciarelli (eds) *Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society* 220–225. Mahwah, NJ: Lawrence Erlbaum.

Bloch, M. (1989) The past and the present in the present. In M. Bloch (ed.) *Ritual, history and power: Selected papers in anthropology* 1–18. London and Atlantic Highlands, NJ: The Athlone Press.

Bohnemeyer, J. (1997) *Yucatec Mayan lexicalization patterns in time and space.* Paper presented at the CLS opening academic year 1997/1998.

Bybee, J., Perkins, Revere and Pagliuca, W. (1994) *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Clark, H. H. (1973) Space, time, semantics, and the child. In T. E. Moore (ed.) *Cognitive development and the acquisition of language* 27–63. New York: Academic Press.

Clifford, J. (2004) Traditional futures. In M. S. Phillips and G. Schochet (eds) *Questions of tradition* 152–168. Toronto: Univeristy of Toronto Press.

Dahl, O. (1995) When the future comes from behind: Malagasy and other time concepts and some consequences for communication. *International Journal of Intercultural Relations* 19(2): 197–209.

Evans, N. and Wilkins, D. (2000) In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3): 546–592.

Evans, V. (2004) *The structure of time: Language, meaning and temporal cognition*. Amsterdam, PA: John Benjamins.

Fillmore, C. (1997 [1971]) *Lectures on deixis*. Stanford, CA: CSLI Publications.

Gell, A. (1992) *The anthropology of time. Cultural constructions of temporal maps and images*. Oxford: Berg.

Haspelmath, M. (1997) *From Space to time: Temporal adverbials in the world's languages*. München: LINCOM.

Hill, C. A. (1978) *Linguistic representation of spatial and temporal orientation.* Paper presented at the Proceedings of the Fourth Annual Meeting of the Berkeley Linguistic Society.

Jackendoff, R. (1983) *Semantics and cognition*. Cambridge, MA: MIT Press.

Keesing, R. M. (1991) *Time, cosmology, and experience*. Unpublished manuscript.

Klein, H. E. M. (1987) The future precedes the past: Time in Toba. *Word* 38: 173–185.

Klein, W. (1994) *Time in language*. London: Routledge.

Lakoff, G. (1993) The contemporary theory of metaphor. In A. Ortony (ed.) *Metaphor and thought* 202–251. (Second edition.) Cambridge: Cambridge University Press.

Lakoff, G. and Johnson, M. (1999) *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

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

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

Levinson, S. C. (2003) *Space in language and cognition. Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Levinson, S. C. (2004) Time for a linguistic anthropology of time. *Current Anthropology* 43(Supplement): 122–123.

Lucy, J. A. (1997) The linguistics of 'color'. In C. L. Hardin and L. Maffi (eds) *Color categories in thought and language* 320–346. Cambridge: Cambridge University Press.

Malotki, E. (1983) *Hopi time: A linguistic analysis of the temporal concepts in the Hopi language*. Berlin: Mouton.

McGlone, M. S. and Harding, J. L. (1998) Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, memory, and cognition* 24: 1211–1223.

Miracle, A. M., Jr. and Yapita Moya, J. d. D. (1981) Time and space in Aymara. In M. J. Hardman (ed.) *The Aymara language in its social and cultural context* 33–56. Gainesville, FL: University Presses of Florida.

Moore, K. E. (2000) *Spatial experience and temporal metaphors in Wolof: Point of view, conceptual mapping, and linguistic practice.* University of California, Berkeley.

Munn, N. D. (1992) The cultural anthropology of time: A critical essay. *Annual Review of Anthropology* 21: 93–123.

Núñez, R. E. and Sweetser, E. (2006) With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30: 1–49.

Radden, G. (2003) The metaphor TIME AS SPACE across languages. In C. B. N. Baumgarten, M. Motz and J. Probst (eds) *Uebersetzen, Interkulturelle Kommunikation, Spracherwerb und Sprachvermittlung – das Leben mit mehreren Sprachen. Festschrift fuer Juliane House zum* 60. *Geburtstag.* (Vol. Zeitschrift fuer Interkulturellen Fremdsprachenunterricht 8(2/3): 226–239.

Santiago, J. (2005) *Time also flies from left to right*. Paper presented at the First UK Cognitive Linguistics Conference, October 2005, University of Sussex.

Saunders, B. (1995) Disinterring *basic color terms*: A study in the mystique of cognitivism. *History of the Human Sciences* 8(4): 19–38.

Senft, G. (1996) Past is present – present is past: Time and the harvest rituals on the Trobriand Islands. *Anthropos* 91: 381–389.

Shinohara, K. (1999) *Typology of space-time mappings*. Manuscript.

Stutterheim, C. von, Carroll, M. and Klein, W. (2003) Two ways of construing complex temporal structures. In F. Lenz (ed.) *Deictic conceptualisation of space, time and person* 97–133. Amsterdam, PA: John Benjamins.

Svorou, S. (1994) *The Grammar of space*. Amsterdam: John Benjamins.

Talmy, L. (2000) *Toward a cognitive semantics*. Cambridge, MA: MIT Press.

Traugott, E. C. (1978) On the expression of spatio-temporal relations in language. In J. H. Greenberg (ed.) *Universals of human language* 369–400. Stanford, CA: Stanford University Press.

Tversky, B., Kugelmass, S. and Winter, A. (1991) Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology* 23: 515–557.

Wierzbicka, A. (1996) The meaning of colour terms and the universals of seeing. In A. Wierzbicka (ed.) *Semantics: Primes and universals* 287–334. Oxford: Oxford University Press.

Yu, N. (1998) *The contemporary theory of metaphor: A perspective from Chinese*. Amsterdam, PA: John Benjamins.

19 From mind to grammar: coordinate systems, prepositions, constructions

Paul Chilton

1 Introduction

Suppose I want to pick up the pencil on my desk. How do I do it? Here are some basic elements of what is a complex process. Imagine there is a set of lines relating the object to its image on my retina, and a set of lines centred on my hand. My eyes and my hand are bits of me. Somehow my brain has to relate these two 'perspectives' on the pen, so that I can execute the reaching and grasping movement.

Let us consider prepositions. My student has lost her pen. I say to my student: 'The pen is in front of the computer'. I can say this no matter where I am standing and probably she will fix on the same location for the pen, one that arises because we give a front and a back to objects like computers and because pens are relatively small. Suppose I say: 'The pen is in front of the waste basket'. Because my waste basket is roughly cylindrical and has no orienting features, this sentence will probably convey that the pen is located in a spatial region between me and the waste basket (or between the addressee and the waste basket). Some uses of some prepositions fix locations of objects relative to other objects; others fix them relative to the speaker's position. It's a question of viewpoint.

Take a more abstract linguistic structure, the counterfactual conditional construction. One may say: 'If John had gone to the party he would have seen Sarah'. Here we have, within the two parts of the sentence, affirmative clauses. But the sense is negative: John did not go to the party and he did not see Sarah. Conversely, one may say: if John had not gone to the party, he would not have seen Sarah. Here the clauses are lexically negative, but John did go to the party and he did see Sarah.

What I aim to do in this paper is to demonstrate the plausibility of making connections between all three of the above scenarios. That is to say, I want to explore the way in which attested neural operations may motivate linguistic structures, at two different levels of abstraction.

2 Egocentric and allocentric operations in visual perception and spatial orientation

The separation of dorsal and ventral pathways in schematic form has been the object of research for some years (for summary see Hartley and Burgess 2003). It is now well established that input form the eyes is fed to the visual cortex (V1), but subsequently is routed to diverse areas of the cortex – various regions in the parietal lobe (broadly called the dorsal stream) and various regions of the temporal cortex including the hippocampal and parahippocampal structures (the ventral stream) (Milner and Goodale 1995). This kind of separation is not of course arbitrary but has a functional basis. The different functions have been characterised in terms of the brain and body's relationship to physical space. Geometrically, the dorsal and ventral streams correspond to egocentric coordinate systems and allocentric systems, respectively (cf. Goodale and Milner 2005, Klatzky 1998, Rolls 1999, among others).

The broad function of the dorsal stream is related to action, specifically actions of reaching and grasping. This implies that the representations in the parietal areas utilise egocentric coordinate systems in order to locate objects the organism is about to act upon. This is why the dorsal stream has been called the 'where' stream. These egocentric coordinates are actually centred (have their origin at) various bodily parts including the retina, the hands, the mouth, the feet. It follows also that there must be geometric transformation between these egocentric coordinate systems. The nature and localisation of these transformation systems is at present not fully understood (though Burgess 2002 indicates posterior parietal cortex, Brodman's area 7a).

The broad function of the ventral stream is object recognition. It is the 'what' stream. It is possible for individuals to have lesions in the temporal area, and thus be unable to recognise or name objects, while still being able to perform spontaneous motor activities upon them (reaching, navigating, etc.). It is the ventral stream and its complex connections that enable the organism to understand a scene and the objects in it: that is, to know the categories and properties of objects as well as to know how they relate to one another in scene-based (i.e. allocentric) coordinates (cf. Goodale and Milner 2005: 101). In addition, the ventral system is connected to systems in the hippocampus that enable navigation through physical space by using allocentric landmarks within a reference frame. It has been shown that areas CA1 and CA3 of the hippocampus (anatomically similar to that of primates including humans) contain 'place cells', which fire differentially in response to real world locations (O'Keefe and Nadel 1978). Structures linked to the hippocampus (mamillary bodies, presubiculum, anterior thalamus) contain head direction cells, which represent the individual's heading relative to landmarks in the environment. The hippocampus also makes it possible to lay down long-term (episodic, i.e. individual) memories of locations and of locations in relation to events Now for this to be possible, there must be a transformational geometric to-and-fro between the egocentric parietal processing and the allocentric temporal-hippocampal processing. Again the exact nature and localisation of such process is still being researched.

The existence of these two interrelated brain systems has already excited the interest of some linguists. Givón sees a correspondence between the ventral stream and lexical semantics on the one hand and between the dorsal stream and propositional information about states or events (Givón 1995: 408–410). Similarly, but in more detail, Hurford (2003) argues for roughly the following: that the ventral and dorsal streams are the evolutionary basis of predicate-argument structure, the ventral stream being the basis of argument concepts, the dorsal the basis of predicate concepts. In the present paper I am arguing something different. I am arguing that it is the distinction between egocentric frames and allocentric frames that is of interest. I am further arguing that it is the geometric transformations between the two frames that is important. They are important because the neuroscience and experimental psychology research clearly indicates that geometric coordinate systems are instantiated neurologically and behaviourally (cf. Gallistel 1999). And it is also indicates that geometric transformations from ego- to allocentric coordinate systems is neurologically and behavioural instantiated. It surely behoves cognitive linguists to take account of this evidence. The claim I want to pursue here is that the transformations I have alluded to should be encountered in lexical meaning. This is perhaps not controversial, since the study of spatial propositions has long since been couched in terms of coordinate systems. It should however be noted that I am not here concerned with the cross-linguistic and Whorfian issues raised by the work of Levinson and others (Levinson 2003). Whatever the cross-linguistic differences in coding, egocentric and allocentric spatial cognition, alongside transformation between the two types of system, is a property of human brains and thus universal.

A more speculative claim – and the one that I shall outline in more detail below – is that the egocentric and allocentric spatial frames, together with their transformation operations, are found in grammatical constructions. Indeed, I want to suggest that many syntactic, semantic and pragmatic phenomena that have hitherto received unconvincing, superficial or controvertible expositions can be naturally explained in a motivated way, in a framework that uses an abstract discourse space representation that is derived from the physical spatial representations, the evidence for the existence of which is well attested, as noted above.

3 Spatial prepositions

It is widely recognised that spatial adpositions across languages exploit three-dimensional coordinate systems whose axes correspond to the sagittal, vertical and lateral axes of the human body.[1] That such coordinate systems are also transformed in various ways is widely acknowledged: the most explicit account in the linguistics domain is Levinson (2003). The present account draws attention to the neurobiological evidence that the human brain alternates between, and integrates, egocentric and allocentric coordinate systems (see above). Egocentric systems locate one or more objects by a position vector in a coordinate system with origin at the speaker/self, oriented by the heading of the speaker/self. This happens when the speaker is the reference location or landmark with any proposition (e.g. 'X is over me'), but also in 'relative frames' where the speaker's orientation is projected onto an object without intrinsic orientation (e.g. 'X is to the left of the tree', cf. Levinson 2003: 43–47). Allocentric systems relate one object to another object, either in coordinates centred on the landmark ('X is over Y', etc.) or by reference to an environmental feature (as in the Tzeltal equivalent of 'X is uphill of Y' for horizontal plane location: Levinson 2003). Levinson's tripartite approach downplays the underlying cognitive operations of egocentric and allocentric operations, while the present approach highlights them, as suggested by the table below.

Table 1

Type of reference frame	Egocentric coordinate system	Allocentric coordinate system
Verticality, determined by canonical position of objects but predominated by gravitational field	'X is over me' (3-dimensional coordinates, with origin at me)	'X is over Y' (3-dimensional coordinates, with origin at Y)
Intrinsically orientated coordinate system	'X is in front of me' (I have intrinsic front-back orientation)	'X is in front of the horse/chair, etc' (horse and chair, etc. have 'front' and 'back')
Relative orientation of coordinate system	'X is in front of the horse/chair, etc' (between me and the horse)	
	'I am in front of the tree' (i.e. my oriented axes are reflected in the non-intrinsically oriented object, so that this object now has a 'front')	'John is in front of the tree' (i.e. John's oriented axes are reflected in the non-intrinsically oriented object, so that this object now has a 'front')
Absolute orientation of coordinate system	'X is north of me' (coordinate system is fixed geophysically)	'X is north of Y'

This table makes clear the obvious fact that egocentric (or deictic) and allocentric conceptualisations can be made within oriented and non-oriented coordinate systems with origins at various locations. The orientational prepositions *in front of/behind* are of interest because they sometimes give rise to egocentric, sometimes to allocentric representations. Indeed, in a non-contextualised sentence they can give rise to conceptual oscillation, a little like an optical illusion. Thus

(1) John is in front of the tree

can make us think either Figure 1 or Figure 2.

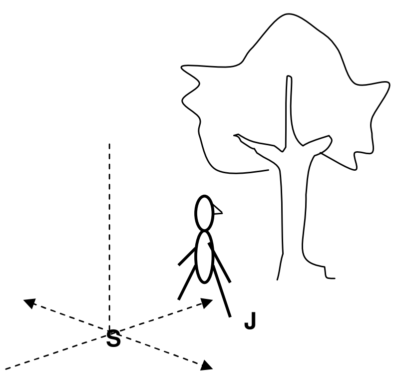


Figure 1. Speaker-centred axis system

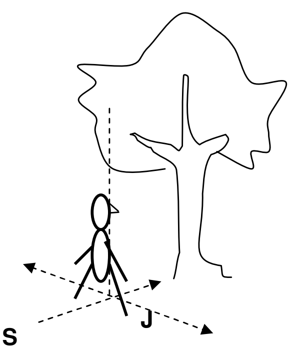


Figure 2. Allocentric axis system

Figure 1 shows the conceptualisation in which the coordinates are centred on the speaker, S, i.e. John is between S and the tree. This conceptualisation also seems to involve the 'occlusion' understanding discussed by Evans (Chapter 9 of the present volume). These are egocentric coordinates. Alternatively, sentence (1) can evoke something like Figure 2: the coordinates are now translated away from the speaker onto John. These are allocentric coordinates. However, the analysis needs to be slightly more complicated, since 'is X in front of Y' should mean that it is Y's front that is the landmark – here, that it is, so to speak, the tree's front that is the landmark with respect to which John is being located. We can describe this geometrically as a reflection transformation, as shown in Figure 3.[2]

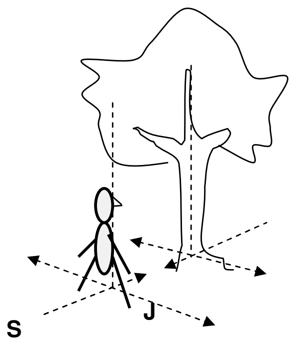


Figure 3. Reflection of axis system

The reflection transformation maintains John's left-right directedness. To say a locandum is to the left of the tree then means that it has the same coordinate on the tree's lateral axis as it does on John's. If the landmark is human or humanoid, then it is possible to conceptualise, and to linguistically encode, a 180 degree rotation that transfers the speaker's right to the landmark's left. For example, if John is facing S, then S may say, 'the tree is to John's left'. If John is not facing S, then there will be a translation of S's coordinates.

The meaning of (1) will flip between the egocentric and the allocentric conceptualisation. In real utterances the denoted location of John will vary: in the egocentric conceptualisation John will always be between S and the landmark, while in the allocentric one John may be anywhere on the circumference of a circle with centre at the tree and the frontal axis its radius. If the landmark has its own 'intrinsic' frontal orientation (e.g. it's a cat, bus, the town hall …), then there is an additional possible allocentric conceptualisation, but this does not affect the main point being made – that egocentric and allocentric representations are recruited by spatial prepositions.

Once we introduce coordinate systems it is a natural step to introduce vectors – mathematical objects drawn as arrows that have direction and magnitude. In a coordinate system the position of a point can be given by the length and direction of a vector from the origin to the point. Giving the coordinates on the axes of the system is equivalent to specifying the vector. This approach can be used for explicating the denotation of certain prepositions, including 'in front of', by specifying a vector space in which all vectors have the same origin in some coordinate system (Zwarts 1997, O'Keefe 1996). This space will be the 'search domain' within which an object can be said to be, for example, 'in front of John'.

4 Discourse space

The claim that I want to make is that the coordinate transformations described above for spatial prepositions are also found in grammatical constructions. To see this we have to move to an abstract (or metaphorical) space. In principle, this is not a controversial move within Cognitive Linguistics. But the notion of an abstract discourse space is new and needs a cautious introduction (see also Chilton 2005).

We call this new abstract space, the Discourse Space (DS), and diagrams of this space are called discourse space models (DSMs). The DS has three scalar axes: d (discourse distance), t (time) and m (modality), as in Figure 4. This is the base space of the speaker (self, subject), S. Other coordinate systems of the same type can be set up at points other than S. The t-axis points in two directions from time 0, the time of utterance. The d-axis points in one direction and allows us to represent geometrically the foreground and background distinction (figure-ground separation) in discourse, i.e. the difference between what is made grammatically salient and what is not. The m-axis also points in one direction and represents epistemic modality. The point maximally distant from S on the m-axis is *irrealis* or counterfactual. The m-axis points in one direction only (i.e. has no negative half-line) because it models modality in terms of distance from S, who is at the point of maximal certainty, coinciding with present time on t and maximum salience on d. The m-axis has an obvious mid-point corresponding to conceptualisations of 'possible' and 'if'.

The DS is thus not a direct analogue of the three-dimensional systems for physical space discussed in section 2. The way the DSM is drawn should not be taken to correspond to up/down, left/right, front/back axes. The claim is that it is the minimum space need to account for a significant number of grammatical and discourse phenomena in an insightful way that is linked to the cognitive motivation outlined in sections 1 and 2.

This does not mean that three dimensions are going to be sufficient to model all phenomena of discourse meaning. On the contrary, discourse processing certainly includes many dimensions.[3] It is possible, however, that three-dimensionality has a special part in human cognition. Be that as it may, the aim here is show how even modest dimensionality can yield insightful modelling of lexical and grammatical phenomena, precisely because the geometry enables us to model transformations of coordinate systems relative to a speaker.

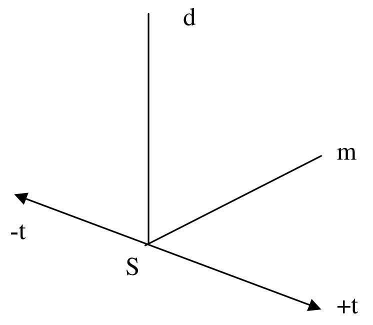


Figure 4. Basic discourse space axis system

In the discourse space referents are 'located' as points; they may be real to S (m=0), hypothetical or counterfactual, or, as we shall see, in some embedded axis system with origin at some location in the base system. Coordinates keep track of anaphoric relations. 'Locations' are of course abstract locations that express concepts in grammar, broadly as outlined in 'localist' theory (Anderson 1972). Relationships between discourse referents are vectors, postulated as unit vectors unless specified otherwise in the context. Vectors are interpreted in various ways standard physical applications, and these applications are followed here. Thus force vectors enable us to model causal relations as directed force, translation vectors allow us to model movement (physical and abstract), position vectors locate referents with respect to other referents.

5 Grammatical constructions and axis transformations

In the rest of this paper I shall treat discourse referents as points with coordinates in the DS, and relations between entities as unit vectors, whether spatial relations or not. But the focus will be on the coordinate systems in which vectors are located, and the transformational relationships between coordinate systems. Some of the constructions that were once predominantly described in generative syntax as 'transformations' can also be so described in the theoretical framework I am outlining here. But there is a big difference. The present framework has simultaneously a cognitive and mathematical motivation, and possibly, as suggested in section 1, a neurological one too. The next three subsections demonstrate three different constructions that can be elucidated by working with three-dimensional DSMs with axis transformations.

5.1 Reflection: active and passive

Passive and active constructions have the function in discourse of foregrounding an entity that has undergone some action and the resultant state. The resultant state can be a physical location, but states, like properties, are also treated in DST as conceptual locations (cf. Anderson 1971, and examples like 'in a broken condition'). The two constructions were treated in early generative grammar as transformations in an idiosyncratic sense of the term. Here we treat them as related by a transformation that has its standard sense in coordinate geometry. Within DST the effect of the transformation is, as required, to bring one discourse entity conceptually 'closer' in the discourse to S.

The example in Figure 5 also illustrates the way the discourse space described in section 3 is used to represent verb semantics. In the sentences

(2) John broke the vase

(3) John moved the vase

both verbs are analysed as having two component vectors, one a force vector, the other a translation vector. The force vector causes the translation. The verb 'move' is represents a physical change of place. The verb 'break' represents a change of physical state, with states analysed as (abstract) locations. Sentences (2) and (3) thus have parallel structure. This shows that, under localist assumptions, we can understand at least two types of transitive verb in the DST framework.[4] With these preliminaries, we can now show how an account of the active-passive relation can fall out naturally from the theory.

First, consider the active construction depicted in Figure 5. The vector v_1i s force vector with tail at the discourse referent John (realis for S) and whose impact on the referent 'the vase' causes a 'translation', v_2, to a new state, 'broken'. And let us assume a context that has this resultant state (i.e. the vector v_3 mapping onto itself) continuing to t=0. The geometric formalism allows us to show John's causing act as foregrounded, i.e. 'closer;' on the d-axis than the other parts of the event structure.

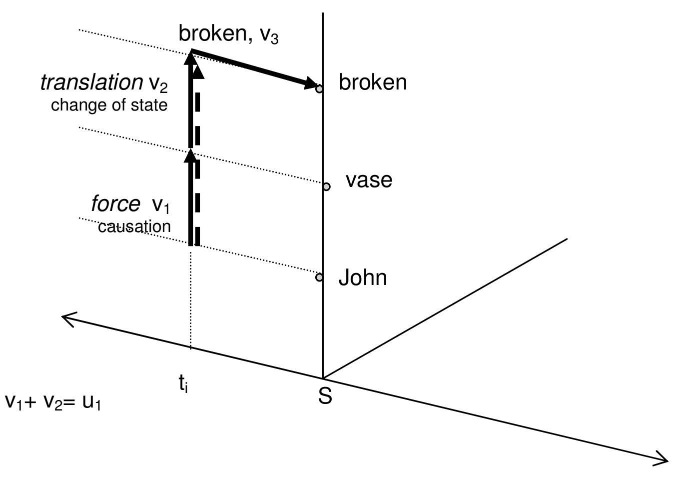


Figure 5. Event structure as vectors: *John broke the vase*

It is now a simple matter to use a transformation of axes – specifically, a reflection transformation – to represent the passive construction, as shown in Figures 6 and 7. The effect is now to background John and his causing action. In 'the vase was broken by

John' it is the referent, together with its change of state and ongoing state that are foregrounded. This is precisely what the reflection transformation produces. Furthermore, an explanation of the use of the prepositional phrase also falls into place.

In the English passive, the agent, if it is expressed, is expressed in a prepositional phrase with 'by'. Why is this spatial preposition used? All we need to do to see the answer is to look at Figure 7 and interpret v_1 as a location vector. The event $v_2 + v_3$ (the breaking of the vase) is 'located at' the backgrounded referent 'John'.[5] What this geometric analysis shows clearly is change of perspective. The speaker can bring into focus either the agent or the undergoer of in an event, just as spatial prepositions can alternate between viewpoints, e.g. *John is in front of the tree*, *the tree is in front f John*.

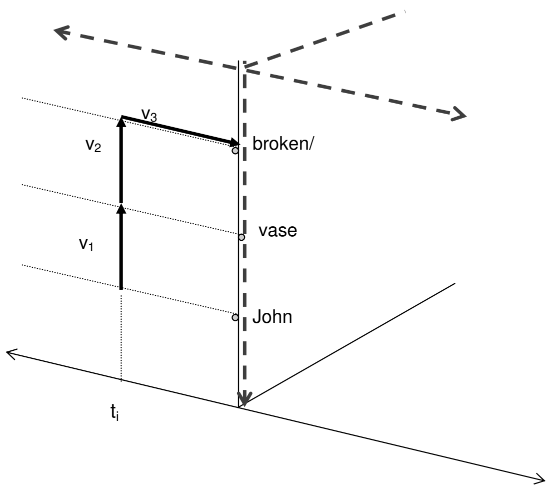


Figure 6. Passive as reflection of axes: *The vase was broken by John*

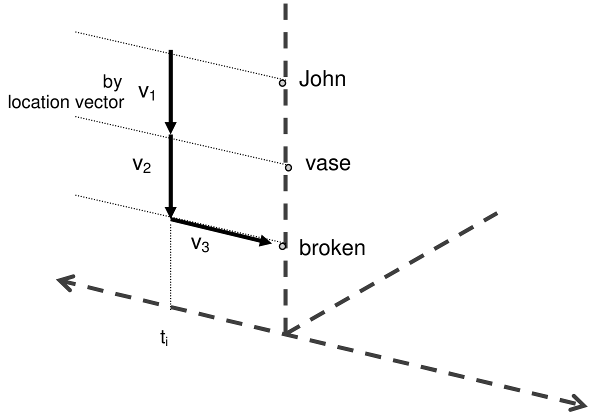


Figure 7. Reflected axes of Figure 6 in normal view

5.2 Translation: factive verbs and epistemic verbs

Just as language reflects the cognitive ability to adopt alternate viewpoints on a single physical-spatial relation (cf. section 2), and just as alternate viewpoints can be taken on events (cf. section 5.1), so language structure corresponds to a cognitive ability to adopt the 'point of view' of another human agent's mental state.[6] Certain lexical items produce semantic constructions that identify S's epistemic state with that of another human (or humanised) agent; others attribute mental states to human agents different from that of S. So-called factive verbs typify the former, as in (4), while epistemic verbs such as *believe* typify the latter, as in (5):

(4) John knows that Mary wrote the report

(5) John believes that Mary wrote the report

Representing the core proposition 'Mary wrote the report' as points and vectors, the difference between (4) and (5) can be represented in terms of transformation of axes. The DST framework naturally incorporates the presupposition triggered by the factive verb in (4), viz. that Mary did indeed write the report. This presupposition I interpret as S's belief, i.e. mental representation, for which the DSM gives the fundamental scaffolding. In Figure 8, which models (4), 'know' can be understood as a function (transformation) that creates a secondary coordinate system with origin located at the point for the discourse referent 'John' (propositions are 'located' in a mind: cf. Anderson 1972, Lyons 1977). The coordinate for the new origin is m=0 in S's base system. This means that what John, S′, knows is also epistemically identical with what S knows, which is the same as saying that (4) presupposes the truth of the complement clause for S.

In (5) there is no presupposition, which in effect simply means that S does not accept, and does not communicate that she or he accepts, that p, i.e. that Mary did indeed write the report. However, S does hold it to be true that John exists and that John believes that p. As in the Figure 8, the main verb produces secondary axes by translation, and its origin has a coordinate on d at 'John'. However, verbs like *believe* (as distinct from know) create a secondary axis system with origin at m>0. The value of m depends on contextual factors. For example, S might think that p may or may not be true, in which case the origin of the new set of axes will be located at the mid-point, as in Figure 9; alternatively, S might think that p is in fact counterfactual, in which case the origin of the new system will be at the extreme of m. Thus in this DSM, proposition p is simultaneously true for John, which is what S asserts, but not necessarily true for S, as the interpretation of (5) requires.

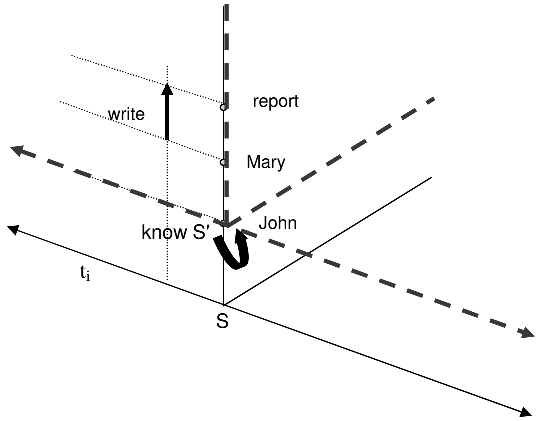


Figure 8. Translation of axes: *John knows Mary wrote the report*

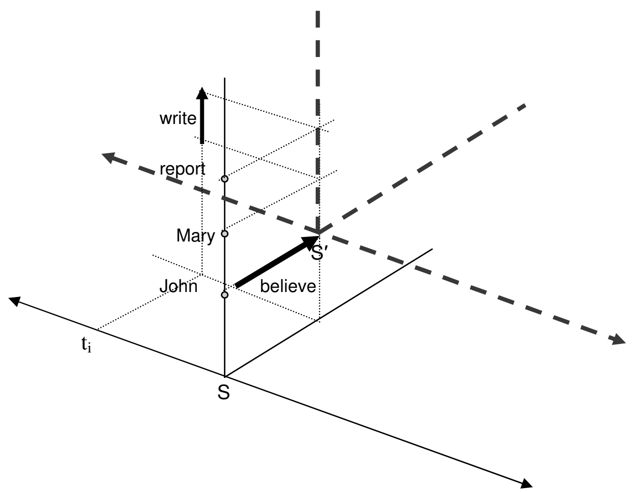


Figure 9. Translation of axes: *John believes Mary wrote the report*

5.3 Conditional sentences

A number of puzzles relating to counterfactual conditionals have been discussed (Fauconnier 1994, and Dancygier and Sweetser 2005). One puzzle that has not been much explored is the following. In a lexically 'positive' counterfactual conditional sentence, the meaning is 'negative': in (6) S is communicating that John did *not* go to the party. On the other hand, in a lexically 'negative' sentence, S is communicating that John *did* go to the party.

(6) If John had gone to the party, he would have seen Sarah

(7) If John had not gone to the party, he would not have seen Sarah

It is possible to give an account of this form-meaning relationship within the framework of DST. In both Figure 10 and Figure 11 John, Sarah and the party are real for the S and for S′ coordinate systems. Notice that the party and Sarah have the same coordinates (in both systems). This reflects the cognitive processing of the sentences: there is a basic inference that Sarah was at the party.[7] The predicates *go* and *see* are modelled as vectors relating John and the party in the protasis, likewise *John* and *Sarah* in the apodosis. The precise characterisation of the relationship between the two clauses is not our concern here, but this relationship has a temporal component, as indicated.

The word *if* is a transformation projecting a reflection of the base system axes, around the mid point on m, giving a new origin 0′ at the distal end of m. Note that this holds constant the time relationships in both the 'real' world of S and the counterfactual world of S′; the same goes for the discourse perspective given by the d coordinates.

Now, in the positive case (6), the positive verb 'had' indicates that the vector relating John and the party and the vector relating John and Sarah are realis in the reflected counterfactual coordinates, i.e. located at m′ = 0′. Simultaneously, these vectors are at the negative end of m in the base system of S. As shown in Figure 10, the reflection transformation economically represents the linguistic-cognitive facts: we have a 'positive' sentence with a 'negative' meaning.

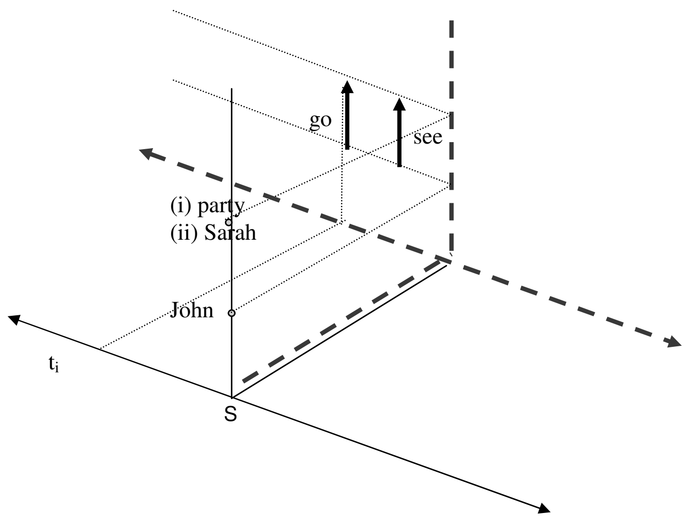


Figure 10. Counterfactual (positive): *If John had gone to the party, he would have seen Sarah*

The converse is true in the case of (7), modelled in Figure 11. The negative verb places the predicates *go* and *see* at the negative end of the reflected axis system, which means that it is simultaneously *realis* for the base system, as required. It is a simple matter to demonstrate that the model provides the right analysis of sentences with combinations of negated and non-negated protasis and apodosis.

This result emerges naturally from the framework we have adopted. The simple geometrical properties we have used are inherent in the standard geometrical framework.

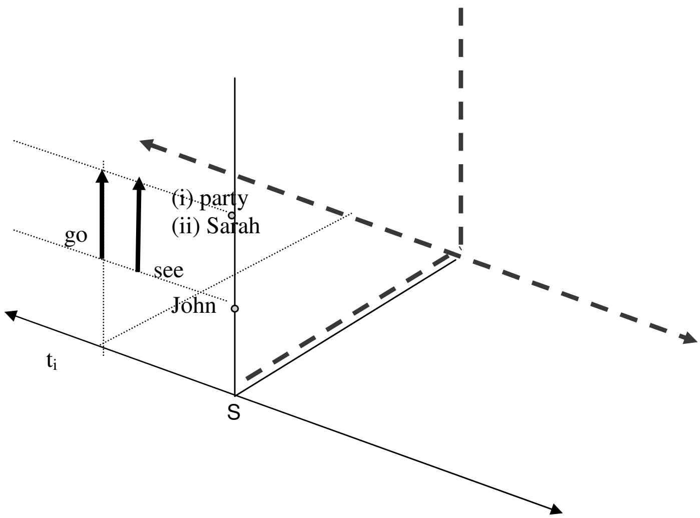


Figure 11. Counterfactual (negative): *If John had not gone to the party, he would not have seen Sarah*

6 Conclusion

Only three grammatical constructions have been demonstrated, but they are all important. What has been discovered in the DST approach is that the concept of transformation, as standardly defined in geometry, appears to be relevant to explaining grammatical constructions, once we have hit on the notion of an abstract discourse space with a 'modal' axis. Transformations involving the other two axes can be used to model constructions not considered here. What we have seen is that coordinate systems, vectors and transformation of axes appear to be descriptively powerful at the level of mind, spatial semantics and grammatical constructions. It is remarkable that the rather simple geometric concepts we have used yield models that integrate linguistic form and conceptualisation. How far this claim can be extended is, however, a matter for further investigation.

It is important to note that the geometrical axis system (frames of reference) used in DST is not the same as that used for describing the workings of spatial prepositions. Certain kinds of frame of reference serve for relations in physical space. The discourse space of DST is an abstract space based on linguistically relevant concepts. It is an abstract space but one that is deictically anchored on the speaker S. It is not meant to be a description of all aspects of discourse processing but only of the fundamental 'scaffolding' on which discourse rests. I do want, however, to suggest that geometrical principles are appropriate to describe it precisely because geometrical principles are appropriate to describe the physical space that prepositions refer to. I am also suggesting that the human mind is using spatial principles, which we can describe using three-dimensional coordinate geometry and vectors, for apprehending both physical space and the more abstract relationships such as figure-ground 'distance', temporal 'distance' and epistemic 'distance'.

This chapter began by describing some aspects of the way the brain processes the spatial environment, proceeded to describe the linguistically mediated conceptualisation of spatial relations, and then applied geometrical principles to the description of grammatical constructions. Whether the sequence in which I have presented these three domains should be taken to suggest a causal or an evolutionary sequence is another matter. But there are at least grounds for speculating that the egocentric-allocentric alternation, which appears to be neurally attested, may be of significance beyond spatial cognition. What I have suggested in this paper is the following. First, neurally embodied spatial cognition can be described geometrically and in terms of egocentric and allocentric relativisation. Second, spatial prepositions can at least in large part be described in the same terms. Thirdly, at least some key grammatical constructions can also be insightfully described in the same terms, given the abstract discourse space model. Finally, these parallelisms, one may speculate, indicate an underlying motivation for grammar in the structure of spatial cognition.

Notes

1 This is not to say that 'functional' concepts are not also involved, as argued by several contributors to the present volume.

2 Levinson (2003: 44–45) describes something similar. It is important to note that the axis system located on John is allocentric and also embedded in that of a speaker S. I have not attempted to represent S's axes explicitly in the diagrams, since S can of course occupy an infinite set of spatial positions relative to John.

3 This is no problem in algebraic format; high dimensional vector spaces are the basis of connectionist modelling and the design of textual search engines (cf. Widdows 2004).

4 Sentences such as *The message was seen by John* need a slightly different analysis.

5 How far this extends cross-linguistically requires investigation, but note that some languages need to use a translation vector: the event 'proceeds from' the agent (German *von*, Latin *ab*). In French, later Latin and other Romance languages (*per*, *par*), conceptualisation also involves spatial movement albeit motion through a medium (= the agent). It should be noted that the analysis seems to apply also to the antipassive construction in such languages as Dyirbal.

6 This linguistically attested ability corresponds to what psychologists call 'theory of mind' (Baron-Cohen 2001).

7 It is inferred that Sarah is in the same location as the party in this particular example. The DSM gives the same coordinate point for both party and *Sarah* and it is important to note that the d-axis does not represent physical space. Obviously, this is not the case in all conditional sentences: e.g. *If John had gone to the party, Sarah would have gone to Manchester*. Spatial relations, including those represented by prepositions and discussed in section 2 above are not part of DST, though there is an important analogy between them. DST does not model physical spatial relationships. DST describes certain fundamental aspects of discourse processing. I assume other systems handle representations arising from the many other aspects of discourse processing, including relations

in physical space. So far as DST is concerned, *Sarah* and *party* in (6) and (7) occupy the same 'place' in terms of discourse distance (or figure-ground separation), as explained in section 3 above. It is also important to note also that the DSM can be regarded as a composite of two successive DSMs, one for each clause, in which coordinates are assigned to referents as the sentence is sequentially processed – whence the marking of the two referents as (i) and (ii) in figures 10 and 11.

References

Anderson, J. (1971) *The Grammar of Case. Towards a Localist Theory*. Cambridge: Cambridge University Press.

Baron-Cohen, S. (2001) *Mindblindness. An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Burgess, N. (2002) The hippocampus, space, and viewpoints in episodic memory. *Journal of Experimental Psychology* 55A(4): 1057–1080.

Chilton, P. A. (2005) Vectors, viewpoint and viewpoint shift: Toward a discourse space theory. *Annual Review of Cognitive Linguistics* 3: 78–116.

Dancygier, B. and Sweetser, E. (2005) *Mental Spaces in Grammar*. Cambridge: Cambridge University Press.

Fauconnier, G. (1994) *Mental Spaces*. Cambridge: Cambridge University Press.

Gallistel, C.R. (1999) Coordinate transformations in the genesis of directed action. In B. O. M. Bly and D. E. Rummelhart (eds) *Cognitive Science* 1–42. New York: Academic Press.

Givón, T. (1995) *Functionalism and Grammar*. Amsterdam: John Benjamins.

Goodale, M. A. and Milner, A. D. (2005) *Sight Unseen*. Oxford: Oxford University Press.

Hartley, T. and Burgess, N. (2003) Spatial cognition, models of. In L. Nadel (ed.) *Encyclopedia of Cognitive Science* Vol. 4 111–119. London: Nature Publishing Group.

Hurford, J. R. (2003) The neural basis of predicate-argument structure. *Behavioural and Brain Sciences* 23(6): 261–283.

Klatzky, R. L. (1998) Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel and K. F. Wender (eds) *Spatial Cognition: An Interdisciplinary Approach to Representation and Processing of Spatial Knowledge* 1–17. Berlin: Springer-Verlag.

Levinson, S. (2003) *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Lyons, J. (1977) *Semantics.* (2 vols.) Cambridge: Cambridge University Press.

Milner, A. D. and Goodale, M. A. (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

O'Keefe, J. (1996) The spatial prepositions in English, vector grammar, and the cognitive map theory. In P. Bloom et al. (eds) *Language and Space*. Cambridge, MA: MIT Press.

O'Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.

Rolls, E. T. (1999) Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus* 9: 467–480.

Widdows, D. (2004) *Geometry and Meaning*. Stanford, CA: CSLI Publications.

Zwarts, J. (1997) Vectors as relative positions: A compositional semantics of modified PPs. *Journal of Semantics* 14: 57–86.

Index